AD 713 500

Army Science Conference Proceedings

16 - 19 June 1970

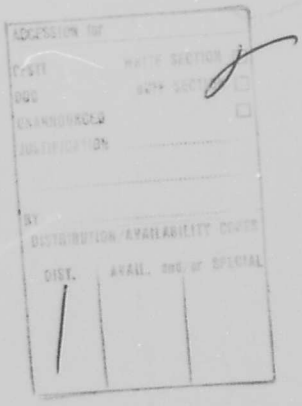
Volume I

Principal Authors A thru G

Distribution of this
document is unlimited

OFFICE, CHIEF OF RESEARCH AND DEVELOPMENT
DEPARTMENT OF THE ARMY

All experiments involving live animals that are reported in the Proceedings were performed in accordance with the principles of laboratory animal care as promulgated by the National Society of Medical Research.



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

IN REPLY REFER TO

AGDA (M) (8 Sep 70) CRDDM 23 September 1970

SUBJECT: 1970 Army Science Conference, Volume I

SEE DISTRIBUTION

1. Inclosed for your information and use is Volume I of the 1970 Army Science Conference Proceedings. This volume contains the unclassified papers by principal authors A thru G which were presented at the conference, 16-19 June 1970, U. S. Military Academy, West Point, New York.

2. Volumes II, III and IV of the Proceedings are being distributed separately.

3. Chiefs of Army Staff agencies and major commanders on the distribution list will insure that copies of the documents are placed in Technical Libraries where they will be available for reference.

BY ORDER OF THE SECRETARY OF THE ARMY:

Donald L. Geer

1 Incl
Vol I, 1970 Army Science
Conference Proceedings

DONALD L. GEER
Colonel, AGC
Acting The Adjutant General

(See page 2 for distribution)

P R O C E E D I N G S

O F T H E

1970 ARMY SCIENCE CONFERENCE

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

16 - 19 JUNE 1970

V O L U M E I

Principal Authors A through G

INTRODUCTION

1. This seventh in a series of Army Science Conferences was held at the United States Military Academy, 16-19 June 1970. "The Needs of the Modern Soldier in His Current and Future Environment", was the theme of this year's conference. In carrying out this theme, the conference presented a cross section of the many significant scientific and engineering programs carried out by the Department of the Army.

2. These Proceedings of the 1970 Army Science Conference are a compilation of all papers presented at the conference and the supplemental papers that were submitted. The Proceedings consist of four volumes, three unclassified and one classified.

3. Our purposes for soliciting these papers were:

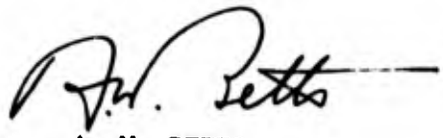
a. To stimulate the involvement of scientific and engineering talent within the Department of the Army.

b. To demonstrate Army competence in research and development.

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

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

4. It is hope that the information contained in the volumes will be of benefit to those who attended the conference and to others interested in Army research and development.



A. W. BETTS
Lieutenant General, GS
Chief of Research and Development

TABLE OF CONTENTS
PROCEEDINGS OF THE 1970 ARMY SCIENCE CONFERENCE

AD 713 501

TESTING FOR AN ORGANIC SUPERCONDUCTOR

FORREST J. AGEE, JR. AND GLENN E. SPANGLER
U.S. ARMY MOBILITY EQUIPMENT RESEARCH AND DEVELOPMENT CENTER
FORT BELVOIR, VIRGINIA

Introduction: Superconductivity and Polymers: Superconductivity is a quantum electromagnetic phenomenon exhibited by a number of metals at very low temperatures. It is characterized in part by an expulsion of magnetic flux from the superconductor and by the ability to conduct a current with zero resistive loss. The unique properties of superconductors make it possible to construct devices ranging from electromagnets with fields in excess of 10^5 Gauss to magnetometers capable of sensing the changes in magnetic field of 10^{-10} Gauss. To date superconducting materials have been found which have critical temperatures (upper temperature limit for superconductivity) ranging from about 20°K to a few millidegrees above absolute zero. There are theoretical reasons to believe that metallic superconductivity will not be found at temperatures much above 20°K.[1] The consequence of this thermal prerequisite for superconducting devices is a cooling requirement which limits the applicability of superconductivity.

The Bardeen Cooper Schrieffer (BCS) theory explains superconductivity as arising from the pairing of electrons at low temperatures due to an attractive interaction which binds the electrons into pairs which obey Bose-Einstein statistics and can occupy a common superconducting ground state. The attractive mechanism arises from a polarization distortion of the metallic lattice due to the presence of an electron, which distortion propagates a virtual (induced) phonon (quantum sound wave) making it energetically favorable for distant electron to move coherently in an electron pair. The pairing mechanism, which is dependent upon the motion of lattice sites, makes the critical temperature and critical magnetic field dependent upon the effective mass of the lattice site. Measurements of the critical temperatures for different isotopic masses of superconductors have shown this dependence to be $T_c \sim \frac{1}{\sqrt{M}}$, which was also predicted by the (BCS) theory.[2] Theoretically, the low transition temperature is due in part to the large mass of the lattice sites whose motion carries the correlation information leading to binding

between electrons.

In an article published by W.A. Little in 1964,[3] the suggestion that superconductivity was possible in certain organic polymers at higher temperatures (even above room temperature) sparked an international flurry of theoretical activity and debate which continues today and a number of attempts to synthesize a superconducting polymer. The idea behind Little's theory was that if a system could be found in which an electron could cause polarization by moving a charge with much smaller mass than a metallic lattice site the transition temperature would be much higher.

In this initial paper,[3] the macromolecule considered was a weakly conducting central spine with side chains attached to it at intervals, which could be readily polarized by an electron in the spine, involving the motion of only a delocalized electron in the chain. The result of the calculations was that an attractive binding interaction between electrons would result, and that the critical temperature for superconductivity would be in the range of several hundred to several thousand degrees Kelvin, depending upon bond localization in the spine.

Some subsequent articles concluded that Little's model would not achieve superconductivity on the bases of screening of the repulsive coulomb interaction to obtain a net attractive binding,[4,5] the effects of thermodynamic fluctuations upon superconductivity in a one dimensional system,[6,7] and the oscillation frequency required to produce pairing between electrons.[4] Other articles[8,9,10,11,12,13] concluded that the first two objections did not in fact rule out the possibility of superconductivity in macromolecules. Another argued[14] that Little's conducting polyene spine example did not satisfy a prerequisite inequality and suggested a different conducting path which would satisfy it.

Other workers were attracted to the field and investigated another polymer[15] and possible interactions[16,17,18] between organic polymers and superconductors. The foregoing did not, however, resolve the problem into inactivity. Rather, it appears that Little's organic molecular model was a specialization of a more general excitonic mechanism[19] in which an electrically polarizable entity can provide a medium for an attractive interaction between electrons.[20,21,22] Several review articles have appeared[23,1,24] which offer further discussion and references. Recently work has been reported[25,26] concerning the possibility of superconductivity in DNA molecules.

Following publication of Little's paper[3] suggesting that a polymeric system might be capable of superconductivity, a number of research groups began efforts to synthesize such a system. The groups affiliated with the U.S. Army Mobility Equipment Research and Development Center include those of Mr. S. Goldfein, Materials Research Support Division, USAMERDC, Ft. Belvoir, Va. 22060; Dr. R. Wiley, RAI Research Corporation, 36-40 37th St., Long Island City, N.Y. 11101; Dr. H.A. Pohl, Department of Physics, Oklahoma State University, Stillwater, Oklahoma 74074; Dr. R. Liepins, Research Triangle Institute, P.O. Box 12194, Research Triangle Park, N.C.

27709; and Dr. M. Goodman, Polytechnic Institute of Brooklyn, 333 Jay St., Brooklyn, N.Y. 11201.

Some of these groups submitted powdered samples for test for superconductivity at USAMERDC which formed the motivation for the work reported in this paper.

The problem of devising a test for an organic superconductor is complicated by the fact that, as far as we know, no one has ever seen one; and no one really knows what one would be like. We can, however, speculate upon probable properties based upon what we know about polymers and superconductors, and hope to devise a test sensitive to these properties.

Polymers are groups of organic molecules which link chemically together to form, in some cases, very large complicated molecules. (My apologies to organic chemists, as I am sure those of you who work in this field will wince whenever I venture into it.) Some polymers consist of only several monomers which link together and do not further polymerize, while others readily link up into vast chains and form networks in three dimensions. One can get an idea of the possible complexity of these molecular networks from proton surface scanning micrographs of textile fibrils.

Superconductors (Type I) exhibit a property called the Meissner effect in which magnetic flux is expelled from the interior of the superconductor by the mechanism of supercurrents induced at the surface which exactly cancel the magnetic field from the interior of the superconductor. In Type II (high magnetic field) superconductors, the Meissner effect is modified somewhat, in that magnetic flux is excluded from the superconductor inhomogeneously with penetration by flux bundles at normal spots within the material. Supercurrents surround the flux bundles and eliminate the magnetic field from the rest of the superconductor. In both cases the Meissner effect appears as a result of supercurrents forming closed paths within the superconducting material at the surface.

Both Type I and Type II superconductors exhibit zero resistance to passage of electric current while superconducting (i.e. at sufficiently low temperature, magnetic field and current density, which vary among different materials). For a continuous superconductor no voltage appears when a current is carried by the superconductor. For two or more superconductors in close proximity, the situation is more complicated and forms one of the most striking examples of the difference between quantum mechanics and classical physics. It is possible for superconducting electrons to tunnel from one superconductor into another across an insulating barrier which will not pass a classical current. This quantum mechanical contact between superconductors is referred to as a Josephson junction[27,28] which has an important application in the detection of small changes in magnetic fields which will interest us later.

The superconducting properties, zero resistivity and expulsion of magnetic flux, are characteristic of known superconductors, but may or may not be characteristic of unknown ones. If we consider a superconducting polymeric situation in which we have a

dilute solution of small polymers, there is little likelihood that we could detect either of these two bulk properties since there would be no continuous path for electrical superconduction, and no closed paths for supercurrents to flow in to oppose a magnetic field - hence - no Meissner effect. Clearly something else would be required for detection of superconductivity in this situation. In the case of a dilute solution of very large multiply connected superconducting polymers, there might be a Meissner effect (although it would have to be a tremendous molecule for us to detect the expulsion of flux from its volume), but a resistance measurement would require us to attach probes to the molecule to see anything. As we go to increasing concentrations of the polymers we come to a range of separation in which Josephson tunnelling between the polymers is possible, and this offers the hope that in very dense concentrations, a polymeric superconductor might assume something approaching the properties of metallic ones. If Josephson tunnelling can link closed paths through molecules between resistance probes, one can hope to see zero resistance (though perhaps only at low currents). The same process could lead also to multiple links of the polymers into quantum mechanical complexes capable of exhibiting a macroscopic Meissner effect. The ultimate degree of concentration for our purposes would arise if the polymer were continuous macroscopically in long strands resembling nylon. Atherton[29] has considered this case, and the mechanism of Little's model is such that the application of an electric field perpendicular to the conducting path would polarize the side chains freezing their ocillation and stopping the superconduction. London's[30] suggestion that in the quantum physics of superfluids might lie an understanding of complex macromolecular structure and biological mechanisms comes to mind again as Atherton points out that voltages available in the human brain would be adequate to perform this electrostatic switching if there are macromolecular superconducting applications operative there. But this is speculation which so far has no confirming basis, since no one claims to have found an organic superconductor; and from the foregoing, it is easy to see that it is not clear what one should expect it to look like.

Since the organic samples which required testing were in powder form, they fell in the relatively dense concentration range. One might therefore hope to see something of the bulk properties of a superconductor. Resistance measurements were considered and rejected, because of the low probability of achieving closed paths between probes inserted into the powders if a superconductor were present. The possibility that only a portion of a given sample might be superconducting due to possible inhomogeneity also recommended against resistance measurements, as some (all) of the probes might miss the interesting region of the sample. Magnetic susceptibility measurements were selected to test for any unusual diamagnetism indicating a Meissner effect.

Experimental Techniques: Detecting Diamagnetism: Magnetic susceptibility measurements can be performed by a.c. and d.c.

techniques, with varying sensitivity. In the conventional a.c. techniques,[31,32] a sample of interest is introduced into the center of a coaxially wound set of mutual inductance coils. The presence of the sample alters the inductive coupling of the primary and secondaries, which is measured by a Hartschorn bridge. An a.c. signal impressed upon the primary gives rise to a signal from the secondaries whose phase and amplitude are balanced by the inductive and resistive networks of the bridge until a null is achieved. The inductive reading is calibrated to known samples and then used to measure samples of interest. The secondaries consist of two identical coils of several thousand turns wound in the opposite sense and separated spacially within the primary coil. The opposite sensing of the coils reduces the signal from the secondaries to a small value, enhancing the effect due to the presence of a sample.

This opposing coil arrangement also leads to greater repeatability, because it permits a difference measurement to be made from the readings taken with the sample in each side of the coil, eliminating spurious drift due to changes between the coils and the outside world. It is possible to make measurements using this method[31,32] of susceptibilities of the order of 3×10^{-6} CGS units to an accuracy of 1%.[32]

Recent developments in applications of superconductivity to magnetometers[33,34,35,36,37,38] have extended the sensitivity of magnetic susceptibility measurements to changes as small as 10^{-10} CGS/gm[33] in a d.c. technique. A sample is introduced into a superconducting d.c. transformer located in a d.c. magnetic field $\underset{\sim}{B}_o$. (see figure 1.) The sequencing of turning on the d.c. field and cooling the superconducting loop is such that $\underset{\sim}{B}_o$ is frozen into the superconducting transformer and maintained by supercurrents generated within the superconductor which oppose any change in $\underset{\sim}{B}_o$ (this arises from the Meissner effect which excludes flux from crossing the superconducting material of the loop). The other part of the transformer passes through a superconducting shield to a superconducting Quantum Interference Detector (SQUID) operating in a low field region, which serves as a very sensitive detector of magnetic field changes. The sample is withdrawn from the superconducting loops, and a supercurrent appears in the transformer to prevent a change in $\underset{\sim}{B}_o$, which is detected by the SQUID. The relevant equations are:

With sample present: $\underset{\sim}{B}_o = \underset{\sim}{H}_o + 4\pi\underset{\sim}{M}$

Where $\underset{\sim}{M} = X\underset{\sim}{H}$, assuming a linear susceptability. When the sample is withdrawn, $\underset{\sim}{M}$ disappears, and a supercurrent appears such that $\underset{\sim}{B}_o$ remains unchanged.

With sample removed: $\underset{\sim}{B}_o = \underset{\sim}{H}_o + \underset{\sim}{H'}_{\text{induced supercurrent}}$.

The sensitivity is dependent upon the SQUID used as a current detector, and can be enhanced by moving the sample between two sets of opposing coils.

This method has been used to observe the onset of superconductivity as a superconducting sample was cooled through T_c, with

successful observation of diamagnetism amounting to 10^{-5} that of a superconductor.[38]

The method of testing selected for the initial testing of the USAMERDC samples was a modified a.c. susceptibility measurement, in which a lock-in amplifier was used to sense the voltage and phase of the output of a set of opposing secondary coils. A block diagram appears in figure 2. The mutual inductance coils consisted of two 2" long secondaries, each with 9,000 turns, wound on a 3/8" OD Lucite tube with an ID of 1/4" with a coil separation of 1/2 inch. The secondaries were located within a primary coil consisting of 60 turns/in wound onto a 7 & 1/2" long threaded Lucite tube 7/8" in dia. The magnetic field of the primary was verified to be uniform over the region of the secondaries to within 1%. The mutual inductance coils were located at the bottom of a glass liquid helium dewar which was vacuum sealed by an O ring to a vacuum manifold which permitted the region of the coils to be evacuated and filled with helium gas or cryogenic liquids for the measurements to be made at liquid nitrogen or liquid helium temperatures. The coils were mounted coaxially with the dewar axis at the lower end of a thinwall stainless steel tube. This arrangement permitted samples to be inserted without disturbing the coils through a sliding O-ring seal at the top of the tube, which extended above the top of the dewar vacuum flange. A micrometer adjustment and spacer provided an adjustable stop which provided repeatable access to the magnetic center of each of the secondary coils. The sample holder extension tube consisted of a 1/4" OD thinwall stainless steel tube with a screw attached to the bottom and an extension arm at the top which came in contact with the adjustable stop or spacer.

A 1000 hertz signal from the oscillator of the lock-in amplifier was applied through a power amplifier to the primary coil. One channel of the lock-in amplifier was used to measure the signal from the secondaries. The other channel was used to permit zeroing of the residual signal from the empty coils by means of an AC zero offset device synchronized with the lock-in signal. This device permitted a signal of adjustable phase and amplitude to be applied to the channel opposing the signal from the secondaries. The current in the primary was monitored, and kept constant at 100 ma at all temperatures. The empty coil signal from the secondary was zeroed out at each temperature.

The samples being tested were encapsulated in tubular lucite sample holders with a sample volume of 0.2 cm^3 (1/8" ID tube, 1" long). A lucite rod was affixed to each end of the sample tube, extending at least 4" in each direction, so that when the sample was located in either coil, the other was filled with lucite. In this fashion, the magnetic response of the sample was compared to that of lucite for both coil positions and for all samples. One end of each sample holder was tapped so that it could be attached to the stainless steel support tube. Care was taken to exclude oxygen from the samples during preparation and from the region of the coils during tests.

The experimental procedure was to achieve the desired temperature in region of the empty coils and use the AC zero offset to cancel the residual signal from the secondaries in amplitude and phase to the sensitivity desired. In making measurements, the sample was first located in the lower coil, then in the upper coil, and the difference in the voltage across the secondary coils formed the data of the measurement. An iron dust sample was used to find the magnetic center of the lower and upper coils and the correct phase for the inductive part of the signal. Resistive losses in iron are small if the particles are small and the frequency low (say less than 15 Khz) so that for iron dust and the frequency used (1 Khz), the resistive losses should be negligible.[39] For ferromagnetic materials, the lock-in phase was adjusted for maximum positive deflection with the sample in the lower coil. The phase angle of maximum deflection did not change with temperature. It was observed that the same phase angle gave a maximum negative reading when a superconductor was present. All samples were tested in a similar manner, with the lock-in amplifier phase adjusted to the angle (270°) of maximum ferro- and dia-magnetism.

Calibration of the apparatus was accomplished as follows. Samples of weighed quantities of magnetic materials whose susceptibilities were known were prepared and measured. The calibration samples included a ferromagnetic material (iron dust), some paramagnetic materials (Nd_2O_3, $MnSO_4$), some diamagnetic materials ($ZnSO_4 \cdot 7H_2O$, Ag I), and a superconductor (Pb at 4.2°K). The calibration measurements indicated that the apparatus was easily capable of detecting a superconducting volume comprising less than 1/2000 of the sample volume or 10^{-4} cm^3. The smallest lead sample tested (a few specks of lead embedded in lucite) verified that a superconductor comprising 1/1000 of the sample volume could be discerned with no difficulty.

The samples prepared by Research Triangle Institute were tested by Dr. H.A. Fairbanks at Duke University, who used a Hartschorn bridge in the a.c. susceptibility measurements. The sensitivity of their measurements was limited by the bridge noise to a superconducting volume of 10^{-4} cm^3.[40]

It might be mentioned in passing that a commercial mutual inductance bridge was available for these measurements, but that preliminary testing indicated that it was not as sensitive to magnetic materials and it was not possible to distinguish between dia- and para-magnetism as was the lock-in amplifier technique discussed above.

In view of the recent work of Gollup et al,[38] in which an almost superconducting material has been diamagnetically detected, it may be worthwhile to proceed to the more sensitive (and more difficult) d.c. superconducting magnetometer method. The a.c. method used in the measurements reported here would not be sensitive enough to obtain this very helpful information.

Experimental Results: The samples which have been tested to date consisted of several Na doped pyropolymers treated at

various temperatures (prepared at Oklahoma State Univ.), a variety of dyes, sugars, aromatics, and polymers (prepared at USAMERDC), and poly(diphenylacetylene) and poly(diphenyldiacetylene) prepared by R.A.I. Research Corporation. Some of the samples were diamagnetic, some paramagnetic, and three of the pyropolymers tested were ferromagnetic. The measurements shown in figure 3 were made with an earlier version of the a.c. apparatus described in this paper. The earlier measurements were sensitive to superconducting volumes larger than 7 parts per 1000. Future measurements will be made with the improved apparatus which has greater sensitivity.

Conclusions: The samples which have been tested by us to date have not shown any anomalous diamagnetism indicative of superconductivity within the limits of sensitivity cited above. The polymers tested, however, were not exactly the chemical embodiment of Little's model. It turns out that it is very difficult to proceed from the theory to the substance, since theory is not encumbered by the same problems which face the synthesizer. One of these problems is an old one of communication between physicists and chemists and vice-versa. Each group involved in the synthesis effort has attempted to create a polymer meeting the requirements of the model; however, none has claimed to have succeeded in doing so. The synthesis efforts revealed some of the complexities involved and produced some new compounds, but the results only show that the polymers tried did not exhibit superconductivity as we know it at temperatures above 4.2°K. The present status is somewhat analogous to that facing Edison (and less successful inventors) prior to a breakthrough (or eventual abandonment of their schemes). Thus far, the materials tried have apparently not been successful. The motivating theory has neither been verified nor vanquished.

Little[19] has suggested an interdisciplinary approach which would hopefully result in a polymeric structure satisfying the requirements of the model. There must first be an interaction between chemists and physicists to select candidate polymers for the spine and side groups for detailed mathematical calculations and iterative modification of the structure to achieve the desired polarization interaction. The type of calculations required[41,42,43,44] are not trivial. If some side groups can be constructed, theoretically and chemically, a criterion has been suggested [19] for testing for the desired electromagnetic properties involving infrared (IR) absorption measurements. It may also be possible to detect the interaction of the side chain polarization with electrons in the spine as well.[19] This is the kind of information which can provide insight as to the next step to be tried in synthesis. When the calculations, synthesis, and IR absorption measurements indicate that a successful polymer has been constructed, it will be interesting to see if a polymeric superconductor has been achieved; and if so, whether it in any way resembles a metallic one.

As our understanding of superconductivity and polymeric systems progresses, an ultimate evaluation of Little's model and/or higher temperature superconductors may be achieved. Recent work[38]

in the field of thermodynamic fluctuations in the vicinity of the critical temperature of superconductors has disclosed a diamagnetic effect due to fluctuation induced Cooper pairing above T_c. This work is significant, because the approach of superconductivity was detected in a system which was almost superconducting. This result recommends further study of the d.c. technique with its greater sensitivity to weak diamagnetism as a possible approach to detecting an almost superconducting polymeric system.

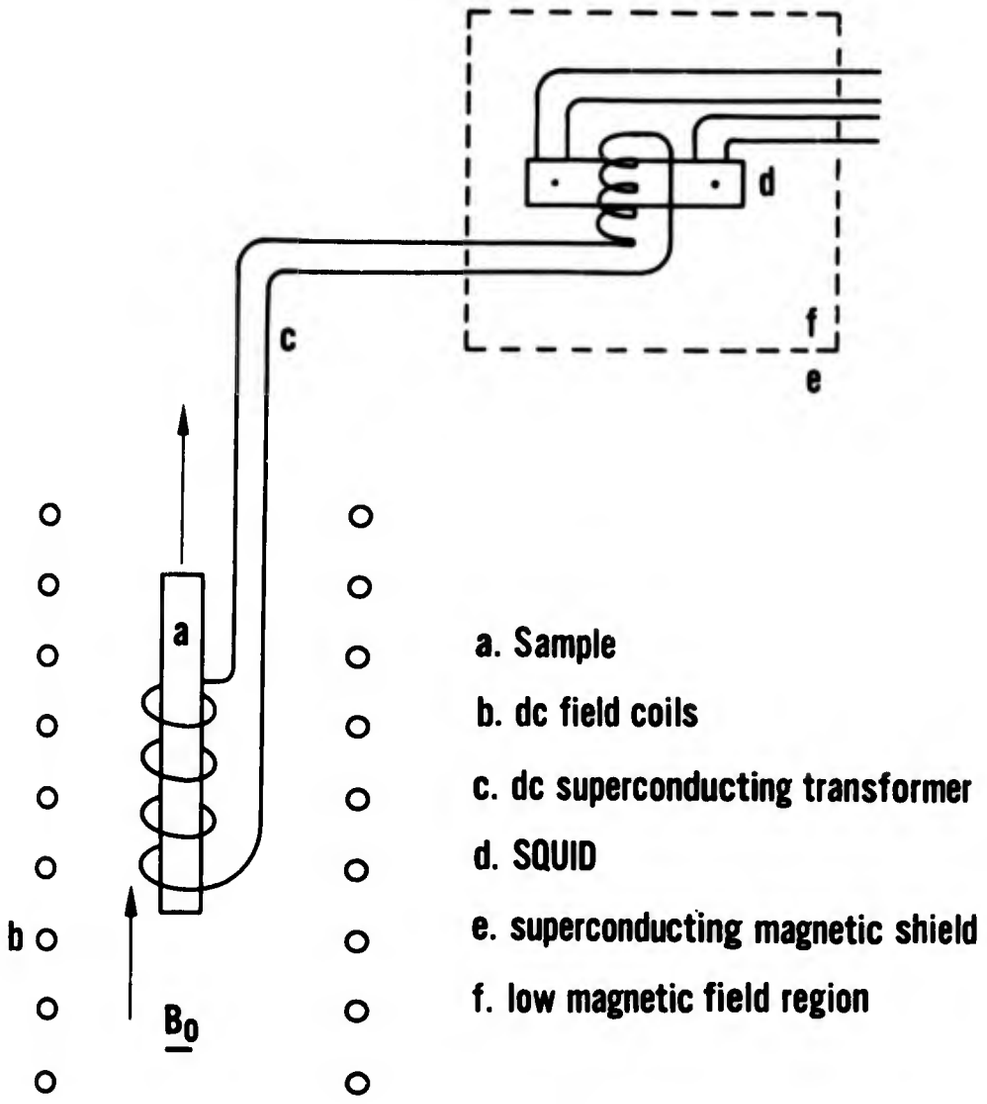
The possibility that other excitonic mechanisms might lead to a superconducting state at higher temperatures is an intriguing one which presently leads to interest in the properties of thin films and dielectrics in close proximity and to surface phenomena.[19,22] To date, however, no such superconductor has been found, and the subject remains open to further study.

Acknowledgements: The authors wish to express their appreciation to Mr. Harry W. Eisenhart and Mr. John M. Wilson for their assistance in building the apparatus and to Mr. Julio N. Calderon and SP5 J.R. Nelson for encapsulating many of the samples.

References

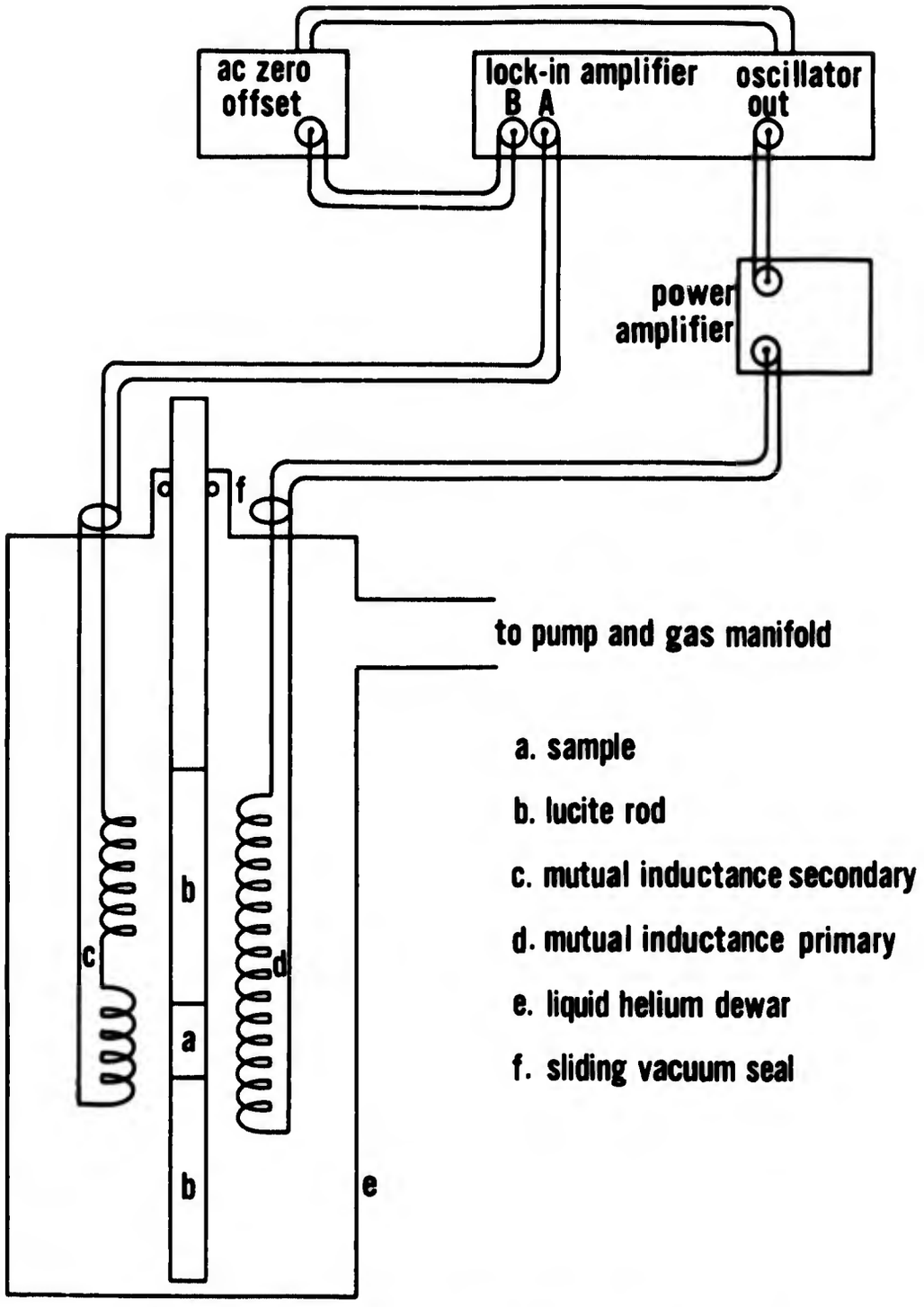
1. L.V. Keldysh, Soviet Physics USPEKHI 8, 496 (1965).
2. J. Bardeen, L.N. Cooper, and J.R. Schrieffer, Phys. Rev. 108, 1175 (1957).
3. W.A. Little, Phys. Rev. 134, A1416 (1964).
4. C.G. Kuper, Phys. Rev. 150, 189 (1966).
5. K.F.G. Paulus, Molecular Phys. 10, 381 (1966).
6. T.M. Rice, Phys. Rev. 140, A1889 (1965).
7. R.A. Ferrell, Phys. Rev. Letters 13, 330 (1964).
8. R.E. Dewames, G.W. Lehman, and T. Wolfram, Phys. Rev. Letters 13, 749 (1964).
9. Yu. A. Bychov, L.P. Gor'Kov, and I.E. Dzyaloshinskii, JETP Letters 2, 92 (1965).
10. Yu. A. Bychov, L.P. Gor'Kov and I.E. Dzyaloshinskii, Soviet Physics JETP 23, 489 (1966).
11. L. Salem, Molecular Phys. 11, 499 (1966).
12. W.A. Little, Phys. Rev. 156, 396 (1967).
13. W.A. Little, Jour. Polymer Sci. Part C 17, 3 (1967).
14. W.L. McCubbin, Phys. Letters 19, 461 (1965).
15. J. Ladkik, G. Biczo and A. Zawadowski, Phys. Letters 18, 257 (1965).
16. F. Meunier, J.P. Burger, G. Deutscher, and E. Guyon, Phys. Letters 26A, 309 (1968), USAMERDC Translation T-1937-68.
17. H.M. McConnell, B.M. Hoffman, D.D. Thomas, and F.R. Gamble, Proc. Nat. Acad. Sci. US, 54, 371 (1965).
18. H.M. McConnell, F.R. Gamble, and B.M. Hoffman, Proc. Nat. Acad. Sci., US, 57, 1131 (1967).

19. W.A. Little, "The Exciton Mechanism in Superconductivity," Paper presented at Conference on Physical and Chemical Problems of a Possible Organic Superconductor, University of Hawaii, Sept., 1969, to be published.
20. Y.M. Balkarei and D.I. Khomskii, JETP Letters 3, 181 (1966).
21. B.T. Geilikman, Soviet Physics JETP 21, 796 (1965).
22. V.L. Ginzburg and D.A. Kirzhnits, Soviet Phys. JETP 19, 269 (1964).
23. J. Bankufi, I. Kirschner, and I. Kovacs, Fizikai Szemle 17 (1967) Hungary, USAMERDC Translation T-1953-68.
24. M. Matsuzawa, Kobunski 16, N 184 (1967) Japan, USAMERDC Translation T-1938-68.
25. R.M. Pearlstein, Phys. Rev. Letters 20, 594 (1968).
26. J.L. Ladik, G. Biczo, and J. Redly, Phys. Rev. 188, 710 (1969).
27. B.D. Josephson, Phys. Letters 1, 251 (1962).
28. P.W. Anderson and J.M. Rowell, Phys. Rev. Letters 10, 230 (1963).
29. D.L. Atherton, Nature, Lond. 205, 687 (1965).
30. F. London, "Superfluids," (John Wiley & Sons, New York, 1950) Vol I.
31. R.A. Erickson, L.D. Roberts, and J.W.T. Dabbs, Rev. Sci. Instr. 25, 1178 (1954).
32. F.R. McKin and W.P. Wolf, J. Sci. Instr. 34, 64 (1957).
33. B.S. Deaver and W.S. Goree, Rev. Sci. Instr. 38, 311 (1967).
34. R.L. Forgacs and A. Warnick, Rev. Sci. Instr. 38, 214 (1967).
35. R.C. Jacklevic, J. Lamb, J.E. Mercereau and A.H. Silver, Phys. Rev. 140, A 1628 (1965).
36. M.R. Beasley and W.W. Webb, "Operation of Superconducting Interference Devices in Appreciable Magnetic Fields," Proceedings of the Symposium on the Physics of Superconducting Devices, AD 661848, University of Virginia, Charlottesville, Va. ed. B. Deaver and W. Goree, April 1967, p. V1-V8.
37. J.E. Zimmerman and A.H. Silver, Phys. Rev. 141, 367 (1966).
38. J.P. Gollub, M.R. Beasley, R.S. Newbower, and M. Tinkham, Phys. Rev. Letters 22, 1288 (1969).
39. D. DeMaw, ed., The Radio Amateurs Handbook, 1968, pp. 27-29.
40. R. Liepins, "Development of High Polymeric Superconductor," Final Report, Dec., 1969, Contract #DAAK02-69-C-0191.
41. W.A. Little, J. Chem. Phys. 49, 420 (1968).
42. H. Gutfreund and W.A. Little, J. Chem. Phys. 50, 4478 (1969).
43. H. Gutfreund and W.A. Little, Phys. Rev. 183, 68 (1969).
44. H. Gutfreund and W.A. Little, J. Chem. Phys. 50, 4468 (1969).

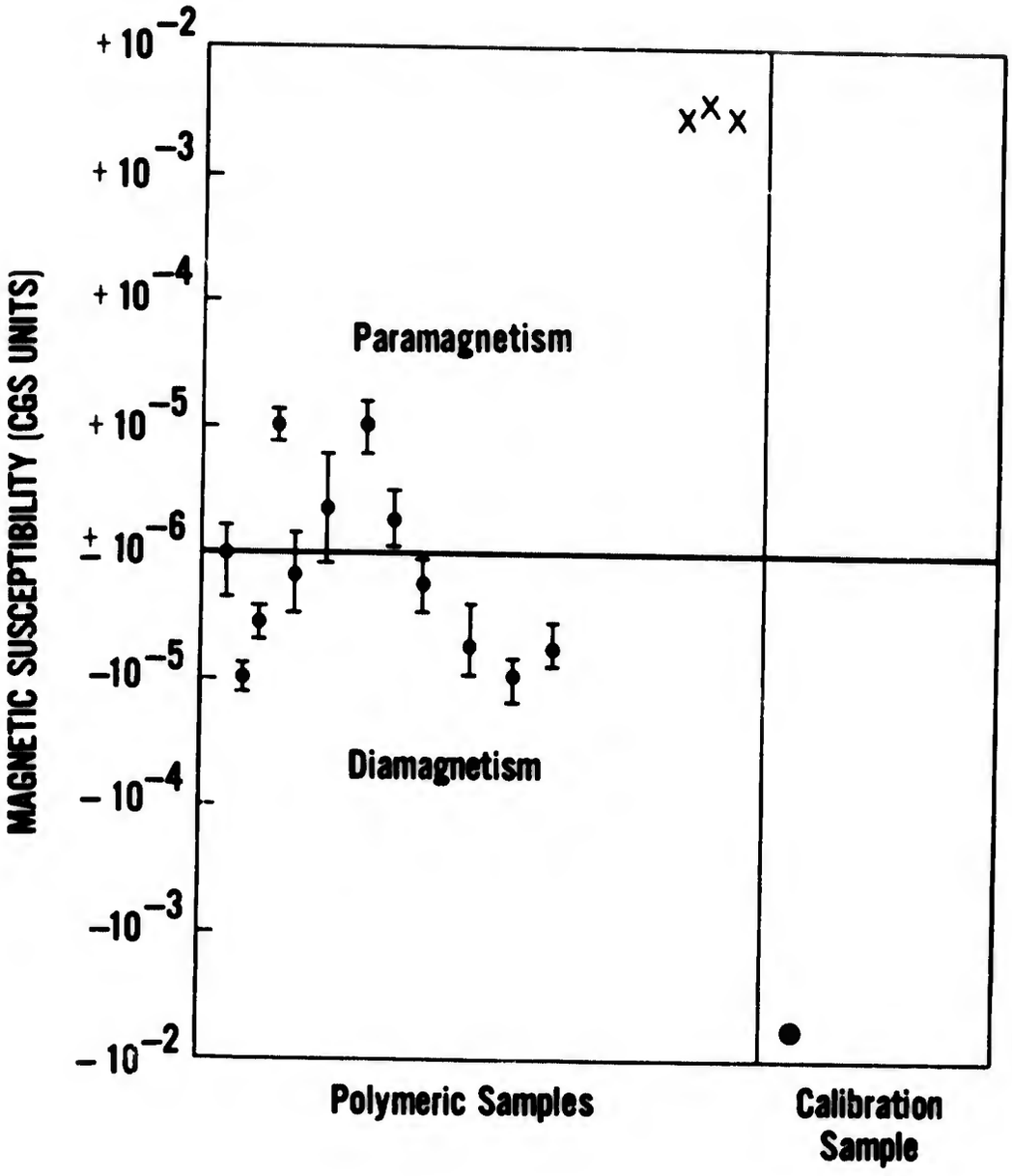


dc Superconducting Magnetometer

Figure 1



a.c. Magnetic Susceptibility Apparatus
Figure 2



Magnetic Susceptibility of Materials Tested

Figure 3

AD 713 502

RAPID ASSESSMENT OF AIRCRAFT LANDING SITES

RICHARD G. AHLVIN AND GEORGE M. HAMMITT, II
U. S. Army Engineer Waterways Experiment Station
Vicksburg, Mississippi

INTRODUCTION

If an aircraft landing site is to satisfactorily carry the user aircraft traffic, it must have sufficient supporting strength. This required strength will, of course, be greater as the aircraft is larger and is greatly dependent on tire pressure. The normal procedure (1,2,3) for determining strength of a landing site is to sample materials of which the site is composed and to conduct suitable strength tests on these materials to arrive at an aircraft-supporting evaluation. This procedure requires special test equipment and specially trained personnel to sample and test to evaluate landing site strength. It also requires substantial time to accomplish the sampling, testing, and evaluation.

Recognizing a need for more rapid and less involved means of landing site strength determination, a probe-type instrument called a "penetrometer" has been developed to give a direct measure of soil strength for use in site evaluation (4,5). This greatly expedites site evaluation but still requires special equipment and special training of personnel.

Obviously, the normal means of sampling and testing or the expedient alternate of using the penetrometer will give suitable landing site evaluations when equipment and trained personnel are available to permit their application. This paper presents a third alternative, which, while not so accurate as either of the first two, requires no special equipment and no more training than the guidance presented herein.

The concept is quite simple. It involves driving a ground vehicle along a landing site, observing its behavior, and using this to predict the supporting capability of the site for aircraft. While the concept is simple, the research on which it is based is both broad and complex. The method is possible because many years of research have gone into the relating of aircraft behavior to ground supporting strength, and during much the same period, extensive study has been made of the relation between ground-vehicle behavior and

unimproved soil strength. By relating these two complexes of criteria through the common soil-strength parameter, it becomes possible to assess aircraft landing sites from behavior of ground vehicles as they traverse the sites.

The criteria developed in this study permit very rapid site assessments by untrained personnel using commonly available ground vehicles. Not only can evaluations be made for initial selection of sites, but reevaluations can rapidly be made to determine effects of rainfall on reducing site strength. The power of the method is in its dependence on only a brief descriptive document with no special training or equipment requirements.

GENERAL CONCEPT

The procedure described in this paper permits rapid survey of entire landing sites by use of any available standard ground vehicle. Specifically, the process is one of selecting a standard vehicle, such as a 2-1/2-ton cargo truck, traversing a potential landing area, and noting the rutting. From the degree of rutting, a direct projection of allowable operations of any aircraft can be readily determined.

SPECIFIC CONCEPT

To accomplish the specific objective of this study, the dimensionless ground mobility parameters developed at the U. S. Army Engineer Waterways Experiment Station (WES) were employed. These parameters consist of clay and sand mobility numbers that reduce the variables of wheel load, soil strength, tire size, and tire deflection to a dimensionless ratio of soil and wheel characteristics as shown in following equations:

Clay mobility number is equal to:

$$\frac{CIbd}{W} \cdot \left(\frac{\delta}{h}\right)^{1/2} \tag{1}$$

The sand mobility number is equal to:

$$\frac{G(bd)^{3/2}}{W} \cdot \left(\frac{\delta}{h}\right) \tag{2}$$

Sinkage number is equal to:

$$z/d \tag{3}$$

where

CI = cone index
b = cross-sectional width of tire, in.
d = outside diameter of tire, in.
W = vertical wheel load, lb
δ = tire deflection, in.

h = tire section height, in.
G = cone index gradient (average increase in CI per in. over a depth equal to the tire width)
z = one-pass rut depth, in.

The relation of sinkage number to clay mobility number is shown in Figure 1. This relation is applicable to all sites except those consisting of clean, cohesionless, free-running sand. The sand mobility number relation is also shown in Figure 1. This paper presents application of the rapid assessment procedure using data from traffic tests on a cohesive soil to predict aircraft operation capability; however, any prediction based on rut measurement on a sandy soil would be conservative using these results. A comparison study (6), Figure 2, showed very close agreement between the relation of sinkage number and clay mobility number as established by small wheels and light-load tests and ground-flotation studies completed with full-scale aircraft wheels and loadings. The degree of agreement of this comparison allows the results to be used for full-scale predictions. The soil-strength variable may be expressed in terms of California Bearing Ratio (CBR), cone index (CI), or airfield index (AI). The conversion curve shown in Figure 3 is used to convert CBR or AI to CI for use in determining the clay or sand mobility number. Data acquired from full-scale testing and evaluation of actual landing sites and used in this study are shown in Figure 4.

APPLICATION

Standard military ground vehicles chosen for study included 1/4-ton M151, 3/4-ton M37, 2-1/2-ton M34, and 5-ton M55 trucks. Using the empty and loaded front-wheel loads and tire dimensions, data were calculated to produce a soil-strength versus rut-depth curve for each vehicle. A typical curve is shown in Figure 5. (The front-wheel loads were used because the ground mobility parameters were developed for front-wheel loading only.) The dimensionless ground mobility parameters were employed to reduce the variables of wheel load and tire dimensions and to produce the data in Table 1. Using the values of sinkage coefficient shown in Table 1, corresponding clay mobility numbers were determined from Figure 2. The clay mobility number for this value yields a soil-strength value, cone index, for the appropriate sinkage depth. The soil strength CI can be converted to AI, and the performance of any aircraft in terms of coverages may quickly be determined by use of Figure 4. Similar plots could be prepared for any pneumatic-tired ground vehicle at any loading. By measuring the front-tire rut depth of a particular military ground vehicle after one pass, the soil strength may be approximated from an appropriate plot of cone index versus rut depth. With this strength determination, one can predict the feasibility of a given aircraft operating at a site (Figure 4).

SAMPLE PROBLEM

Determine if a C-7A aircraft having a single-wheel load of 6420 lb and tire-inflation pressure of 39 psi can successfully

Table 1

Determination of Soil Strength from Rut Depth for a Clay Soil

Vehicle	Vehicle Weight lb	Tire Size	Unloaded Outside Tire Diam (d) in.	Tire Pressure psi	Unloaded Tire Width (b) in.	Rut Depth (z) in.	Sinkage Coefficient z/d	Clay Mobility No.	Cone Index
1/4-ton, 4x4 M151 truck	2,473 Empty with driver	7:00x16	30.5	20	7.17	0	0.0000	>20.00	165
						Trace (0.1)	0.0033	10.00	82
						0.25	0.0082	7.70	63
						0.50	0.0163	6.00	49
						1.0	0.0327	4.00	33
						2.0	0.0657	3.00	24
						3.0	0.0983	2.75	--
	3,000 Driver and 3 passengers					0	0.0000	>20.00	200
						Trace (0.1)	0.0033	10.00	100
						0.25	0.0082	7.70	77
						0.50	0.0163	6.00	60
						1.0	0.0327	4.00	40
						2.0	0.0657	3.00	30
						3.0	0.0983	2.75	--
3/4-ton, 4x4 M37 truck	5,950 Empty	9:00x16	35.2	50	9.63	0	0.0000	>20.00	250
						Trace (0.1)	0.0028	11.00	133
						0.25	0.0071	8.00	97
						0.50	0.0142	6.00	73
						1.0	0.0282	4.25	51
						2.0	0.0568	3.10	38
						4.0	0.1132	2.60	31
	7,820 Gross weight					0	0.0000	>20.00	265
						Trace (0.1)	0.0028	11.00	140
						0.25	0.0071	8.00	102
						0.50	0.0142	6.00	76
						1.0	0.0282	4.25	54
						2.0	0.0568	3.10	40
						4.0	0.1132	2.60	33
2-1/2-ton, 6x6 M34 truck	13,900 Empty	11:00x20	43.2	75	12.04	0	0.0000	>20.00	300
						Trace (0.1)	0.0023	11.50	182
						0.25	0.0057	9.00	142
						0.5	0.0115	6.50	103
						1.0	0.0231	4.75	75
						2.0	0.0462	3.50	55
						4.0	0.0925	2.75	43
						6.0	0.1388	2.50	39
	24,300 Gross weight					0	0.0000	>20.00	350
						Trace (0.1)	0.0023	11.50	218
						0.25	0.0057	9.00	171
						0.5	0.0115	6.50	123
						1.0	0.0231	4.75	90
						2.0	0.0462	3.50	66
						4.0	0.0925	2.75	52
						6.0	0.1388	2.50	--
5-ton, 6x6 M55 truck	24,004 Empty	11:00x20	43.2	75	12.04	0	0.0000	>20.00	400
						Trace (0.1)	0.0023	11.50	249
						0.25	0.0057	9.00	195
						0.50	0.0115	6.50	141
						1.0	0.0231	4.75	103
						2.0	0.0462	3.50	76
						4.0	0.0925	2.75	60
						6.0	0.1388	2.50	54
	34,004 Gross weight					0	0.0000	>20.00	400
						Trace (0.1)	0.0023	11.50	249
						0.25	0.0057	9.00	195
						0.50	0.0115	6.50	141
						1.0	0.0231	4.75	103
						2.0	0.0462	3.50	76
						4.0	0.0925	2.75	60
						6.0	0.1388	2.50	54

Note: All computations made using a b/h value of 0.16 and based on manufacturer's specified dimensions and weights.

operate on a clay site where a 1-in. rut depth resulted from one pass of the front wheel of a 2-1/2-ton M34 cargo truck.

SOLUTION

The empty M34 cargo truck has a front-axle weight of 6900 lb and a 11.00x20 tire size. With a one-pass rut depth of 1 in., the sinkage number z/d for this vehicle would be 0.0244. From Figure 2, the clay mobility number is 5.60. By substitution in Equation 1 as shown below, the strength of the soil in terms of CI would be 93.

$$5.6 = \frac{CI \cdot 11.96 \cdot 43.5}{3450} \cdot (0.16)^{1/2}$$

$$5.6 = \frac{520.3\ CI}{3450} (0.4)$$

$$5.6 = 0.06\ CI$$

$$CI = 93$$

From Figure 3, the CI of 93 is converted to an AI of 1.9. Next, by entering the nomograph (Figure 4) with the C-7A load of 6420 lb and tire pressure of 39 psi and the soil strength (1.9 AI), it is determined that the C-7A cannot successfully operate on this particular landing site.

VALIDATION

A validation of this procedure was conducted on a limited basis at the WES. The tests were conducted under shelter in a test bin 12 ft wide, 170 ft long, and 5 ft deep. The heavy clay soil was placed in 6-in. lifts to a depth of 5 ft and was determined to have a CBR strength of 2. The clay soil was practically identical in all respects to the soil used in providing the input data for the development of the method.

One-pass traffic was applied at approximately 5 mph with the following vehicles:

1/4-ton M151	
Empty	2,635 lb
Maximum cross-country load	3,035 lb
3/4-ton M37	
Empty	6,010 lb
Maximum cross-country load	7,800 lb
2-1/2-ton M35A1	
Empty	13,500 lb
Maximum cross-country load	18,500 lb
5-ton M55	20,500 lb

The vehicle characteristics are documented in Reference 7.

Each vehicle with each loading made one-pass applications; rut-depth and soil-strength measurements were recorded before and after traffic. Each vehicle was positioned to traverse fresh soil on every pass. Table 2 shows the results of the validation tests and comparison of measured and predicted values. These data indicate the predicted values agree well with the validation test results.

The operation of several standard military ground vehicles were related to the requirements for operation of military aircraft on unsurfaced fields. A series of charts were prepared as a general typical guide in the evaluation of soil strength of unsurfaced forward airfields by use of military ground vehicles. One of these charts is shown in Figure 6. A direct evaluation could be made by trafficking an area with one pass of a vehicle, measuring the resulting rut depth, and referring to the appropriate chart for the particular vehicle used. For example, it is noted that an unloaded U-6 aircraft has a capability of 10 operations on a field experiencing a trace of a rut (0.1 in.) from one pass of a loaded 1/4-ton truck.

CONCLUSIONS

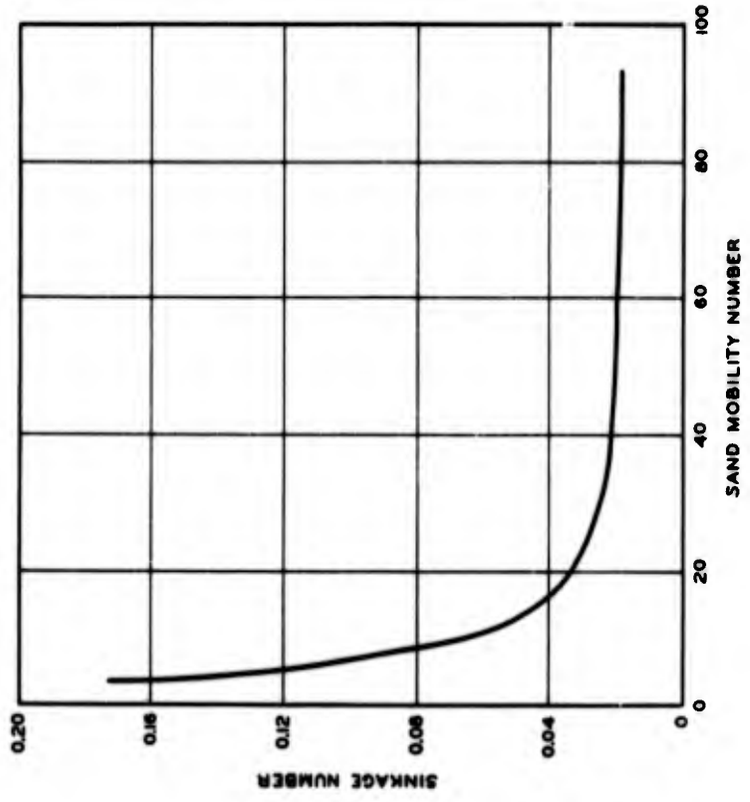
The response of a soil surface to the operation of a ground vehicle can be used to predict the ability of the soil to support aircraft operations. The method can be used to rapidly assess landing sites in forward areas when it is not possible to make conventional evaluations.

Table 2

Summary of Data

Vehicle Type	Load lb	Tire Pressure psi	Tire Size	Strength of Soil Surface CI	Strength of Soil CBR	Sta	Measured Rut Depth, in. From Original Surface	Measured Rut Depth, in. Maximum	Predicted Rut Depth in.
1/4-ton, 4x4 M151 truck	2,635	18 front 22 rear	7:00x16	63	1.8	10+00	0.05	0.05	
				71	1.2	15+00	0.12	0.12	
						20+00	0.10	0.10	
				63	0.8	25+00	0.10	0.10	
						30+00	0.10	0.15	
						35+00	0.10	0.15	
						Avg	0.10	0.11	0.10
	3,035	18 front 22 rear	7:00x16	78		10+00	0.30	0.40	
				78		15+00	0.13	0.20	
				78		20+00	0.05	0.22	
				76		25+00	0.15	0.15	
				66		30+00	0.15	0.45	
						Avg	0.16	0.28	0.22
3/4-ton, 4x4 M37 truck	6,010	40	9:00x16	63		10+00	0.37	0.50	
				72		15+00	0.47	0.50	
						20+00	0.45	0.55	
				63		25+00	0.47	0.50	
						30+00	0.40	0.55	
						35+00	0.40	0.70	
						Avg	0.43	0.55	0.40
	7,800	40	9:00x16	78		10+00	0.70	1.27	
				78		15+00	0.40	0.90	
				78		20+00	0.57	1.10	
				76		25+00	0.70	1.10	
				66		30+00	0.63	1.10	
						Avg	0.60	1.09	0.45
2-1/2-ton, 6x6 M35A1 truck W/W	13,500	35	11:00x20	63		10+00	0.60	1.00	
				72		15+00	0.54	1.00	
						20+00	0.47	0.95	
				63		25+00	0.50	0.80	
						30+00	0.45	1.05	
						Avg	0.51	0.96	1.05 0.90*
	18,500	35	11:00x20	80		12+00	1.12	1.95	
				80		17+00	0.92	1.68	
				72		22+00	0.91	1.83	
				62		27+00	1.05	2.00	
						Avg	1.00	1.86	1.60 1.30*
5-ton, 6x6 M55 truck	20,500	35	11:00x20	78		10+00	0.65	1.58	
				78		15+00	0.52	1.25	
				78		20+00	0.65	1.50	
				76		25+00	0.45	1.05	
						Avg	0.57	1.34	1.30

* Predicted rut depth based on M34, 2-1/2-ton vehicle.



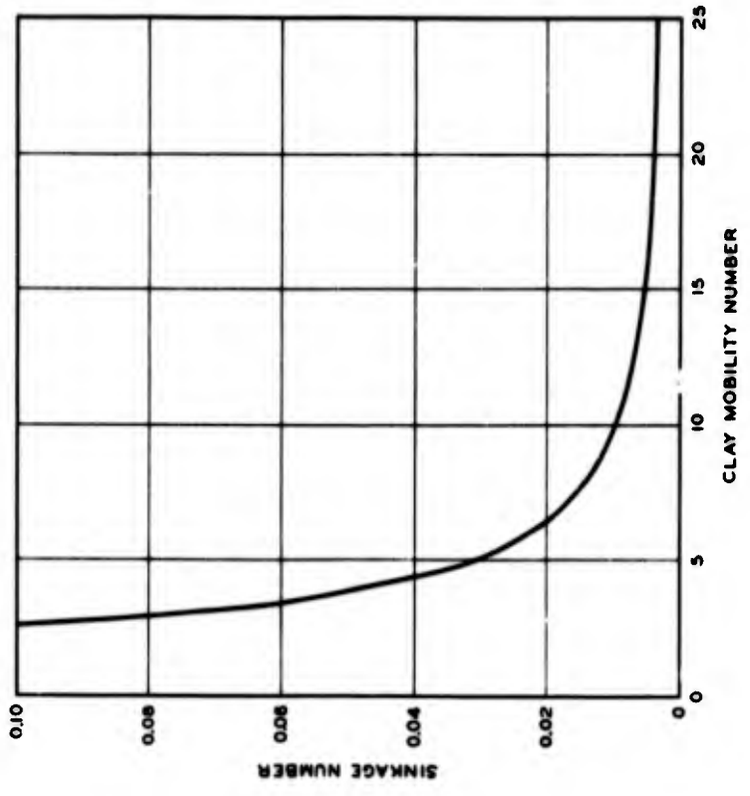
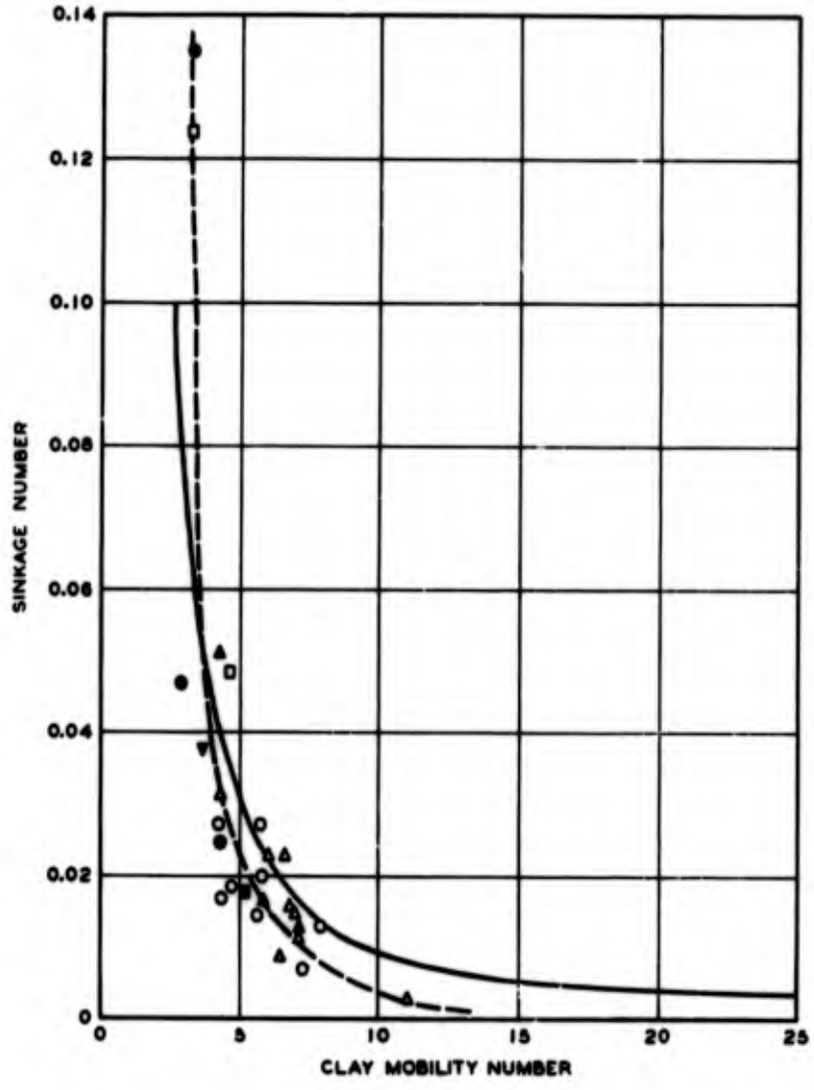


Fig. 1. Sinkage number vs mobility number (first pass)



LEGEND

GROUND-FLOTATION TEST:

	TIRE SIZE	PLY
○	56-16	24, 32, AND 38
□	30-11.5	24
△	25:00-25	30
■	17:00-16	12
▼	34-99	14
▲	20-20	22
●	30-7.7	16

SINGLE-WHEEL LOADS RANGED FROM 19,000 TO 35,000 LB.

AMRB MODEL STUDY

Fig. 2. Sinkage number vs clay mobility number (comparison study)

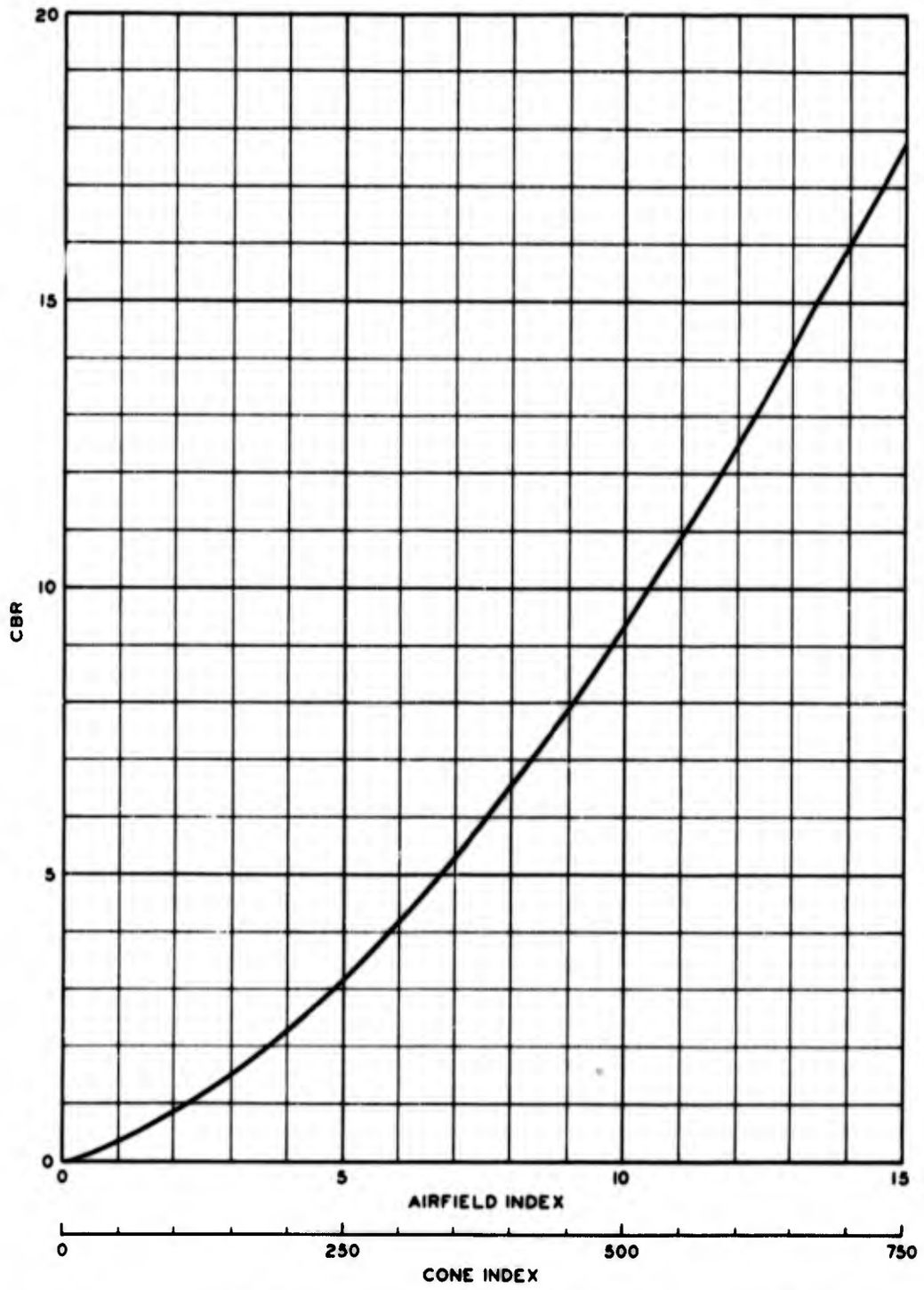


Fig. 3. Correlation of CBR and airfield index or cone index

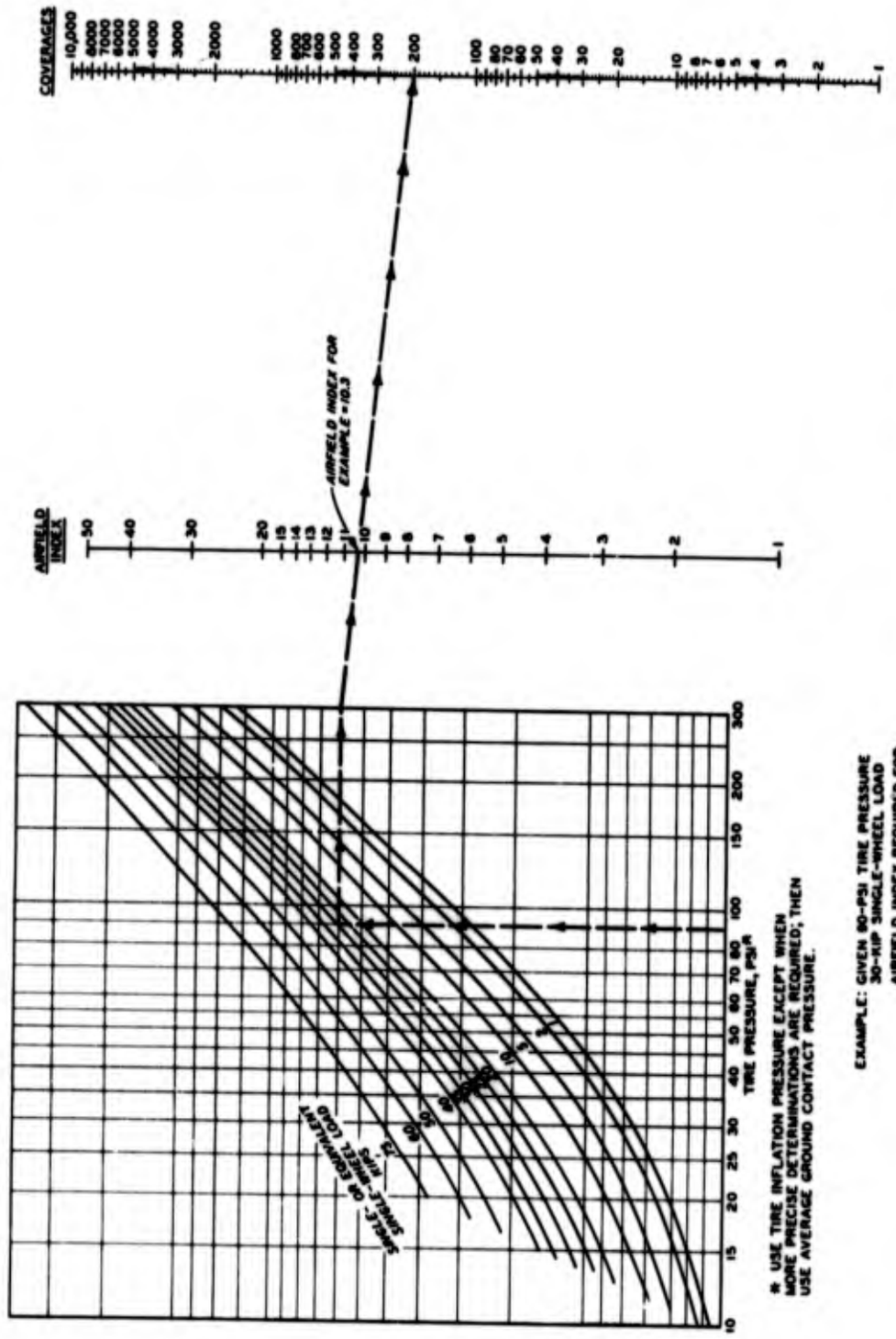


Fig. 4. Airfield index required for operation of aircraft on unsurfaced soils (Reference 6)

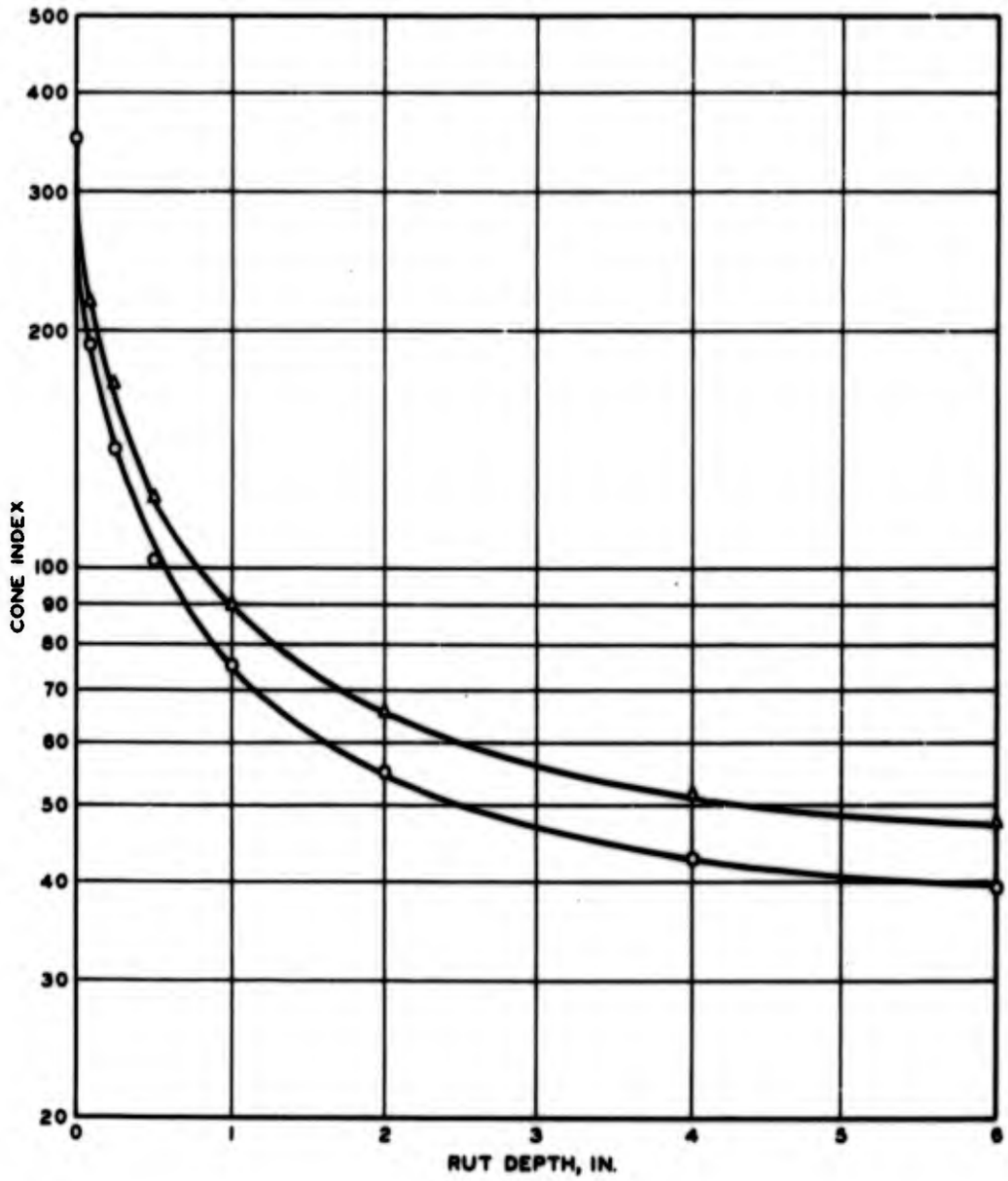
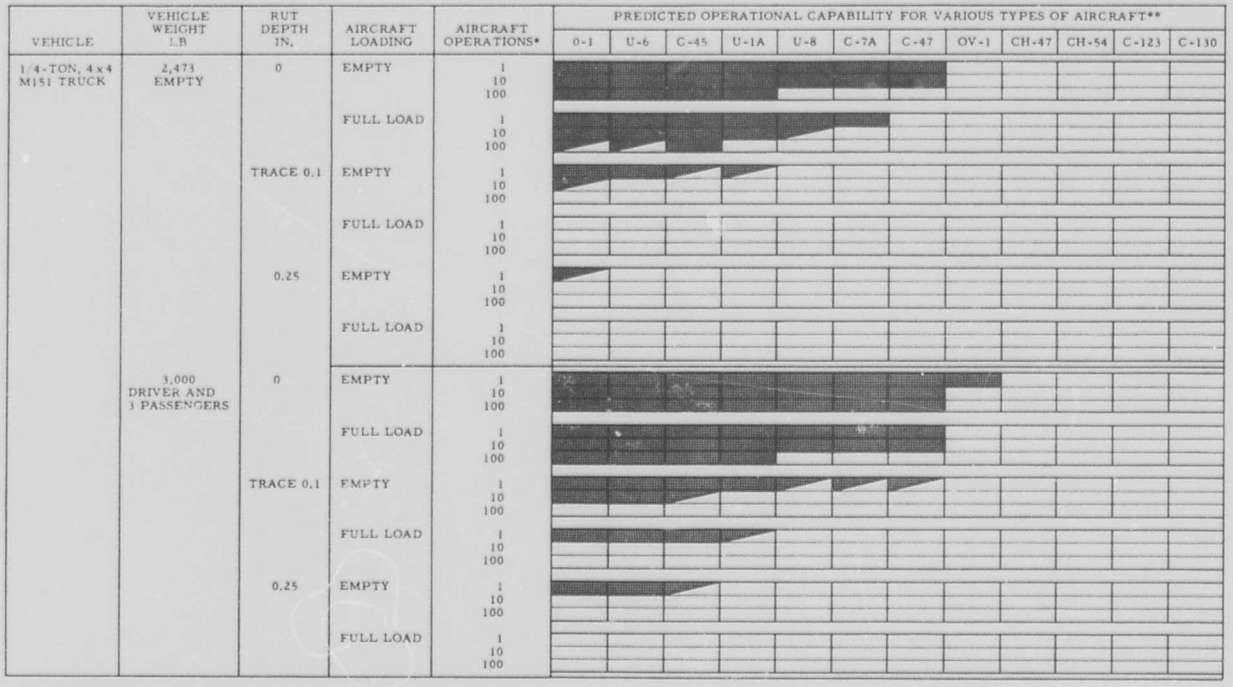


Fig. 5. Cone index vs rut depth for 2-1/2-ton M34 truck



* One operation is one takeoff and one landing.

** Aircraft can operate at indicated loading.
Aircraft cannot operate at indicated loading.
Aircraft may be able to operate at indicated loading with calculated risk.

Fig. 6. Standard military ground vehicle related to operation of military aircraft on unsurfaced airfield

LITERATURE CITED

1. Headquarters, Department of the Army, "Planning, Site Selection, and Design of Roads, Airfields, and Heliports in the Theater of Operations," Technical Manual TM No. 5-330, July 1963, Washington, D.C.

2. Headquarters, Department of the Army, "Planning and Design for Rapid Airfield Construction in the Theater of Operations," Technical Manual TM No. 5-366, November 1965, Washington, D. C.

3. Department of Defense, "Military Standard Test Method for Pavement, Subgrade, Subbase, and Base-Course Materials," MIL-STD-621A, December 1964, Washington, D. C.

4. Headquarters, Department of the Army, "Soils Trafficability," Technical Bulletin TB ENG 37, July 1959, Washington, D. C.

5. Fenwick, W. B., "Description and Application of Airfield Cone Penetrometer," Instruction Report No. 7, October 1965, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

6. Ladd, D. M., and Ulery, H. H., Jr., "Aircraft Ground-Flotation Investigation Part 1 - Basic Report," Technical Documentary Report AFFDL-TDR-66-43, Part I, August 1967, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

7. Freitag, D. R., "A Dimensional Analysis of the Performance of Pneumatic Tires on Soft Soils," Technical Report No. 3-688, August 1965, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

AD 713 503

THE FRICTION BLISTER*

WILLIAM ALEXANDER AKERS, COLONEL, MC AND
MARION BALDUR SULZBERGER, M.D.
LETTERMAN ARMY INSTITUTE OF RESEARCH
SAN FRANCISCO, CALIFORNIA

Few if any of us have not suffered from friction blisters. Athletes may lose a match or workers lose pay from this banal injury. In a soldier, blisters can substantially decrease his effectiveness and so seriously impair his mobility that they could lead to his capture or death. Thus friction blisters can be military problems, in and of themselves, and also because of their not uncommon sequelae of ulceration, fungal, staphylococcal and streptococcal infection, cellulitis and sepsis (Fig 1). President Coolidge's son died of septicemia consequent to a blister incurred while playing tennis.

Friction blisters cause high rates of disability in military populations, especially in recruits in training and in soldiers on the march. At Fort Dix, a large Army training center, 10% of all admissions to the General Surgical Service of Walson Army Hospital were for injuries to the skin of the feet and legs (1). Well over half of these admissions were either because of friction blisters or their consequences of ulcers and infections. At Fort Ord, another large training center, out of 4,000 men at risk on any given day, 12 to 15 are hospitalized and 50 more come to sick call because of friction blisters. (Communication from Colonel Joseph Molloy, the Station Hospital Commander).

The history of military campaigns shows that, at one time or another, commanders have had to center their concern on their soldiers' blistered feet. Statements to this effect can be found in the German, French, Russian and other literature. In England, in 1895, there were embarrassing questions and a furor in the House

*Footnote.
AR 70-25 regarding the Use of Human Volunteers for Research and AR 70-18 regarding Experimental Animal Care were strictly observed.

of Commons concerning excessive disability caused by friction blisters during the Army Maneuvers in the New Forest. There were 133 cases of incapacitating blisters in the 824 men who participated. CPT Cortese and Specialist Peyakovich, while members of our staff, conducted a prospective study by observing participants in the annual "Walkathon" of 16 miles between Sonoma and Petaluma, California. The participants in this march were civilians of both sexes, ranging from 7 to 43 years. Of 92 who could be examined beforehand and closely followed 48 (52%) developed foot blisters, many of which were severe enough to cause the walker to fall out (Fig. 2).

Despite the recognized magnitude of the problem, prior to the studies by our Institute (2 through 8), the only scientific publications we could find that dealt with friction blisters were the excellent articles by Naylor (Refs 9, 10, 11).

OBJECTIVES

The following were some of the factors that we wished to investigate and measure as accurately as possible:

1. The intensity, duration, rate of application and quality of the frictional forces necessary to produce blisters under a variety of conditions.
2. The pathologic, histopathologic, and other changes which characterize friction blisters and distinguish them from other blisters, such as those produced by heat, vesicant chemicals and blistering skin diseases.
3. The susceptibility of different persons and different skin areas to develop friction blisters.
4. The degrees to which various kinds of fabrics, leather, rubber, plastics, and other materials differ in their friction-producing properties.
5. The ways to select and design materials, footwear and other gear so as to reduce the frictional forces they exert upon the skin under conditions of actual use.
6. The influence of humidity, skin wetness, and temperature upon the time required to produce friction blisters.
7. The ways to modify the skin prophylactically in order to reduce the susceptibility to friction blistering.
8. The best ways to treat friction blisters to promote their healing, restore function and prevent infection.

DESIGN AND DEVELOPMENT OF INSTRUMENTS

Only by developing an experimental model were we able to study the above factors and observe typical friction blisters, from their very first inception to their final healing. To this end our laboratory has designed and constructed five generations of instruments during the past five years, which produce friction by a rubbing to-and-fro motion or by twisting through an arc on the skin (Fig 3, 4). Our earliest models were modifications of the instrument employed by Naylor (9) to study blistering; and of an instrument designed at Walter Reed Army Institute of Research by Irvin

Levin, PhD and used by Goldblum and Piper (12), for scratching the skin.

Some of the earlier instruments permitted the estimation of torque and frictional forces; of temperature rises in the skin and in the rubbing materials just below the rubbing heads; and of the effects of skin moisture or wetness on friction. Some models also allowed the simultaneous rubbing of two skin sites or two individuals, in order to compare directly their different susceptibilities; or with two different materials to compare their relative friction-producing properties.

However, all of the earlier instruments had the serious drawback that the moment of blister-beginning and the moment that the cleft was complete had to be guessed. We were obliged to interrupt the rubbing from time to time, lift the rubbing head, and look at the rubbed site to see whether the characteristic signs of the formation of the blister cleft had appeared.

With our latest model* (Fig 5), coupled to a thermal rectilinear recorder (Dynagraph Model, Beckman Instrument Company) to record the successive changes in friction, the actual point of beginning blistering and the end-point of complete production of the intradermal cleft can be seen on the graph. Due to the retaining cup (Fig 6) when a change occurs in the epidermis as a result of slippage between uppermost layers and those beneath, the rubbing head does not follow the change to a deeper plateau but is opposed by a progressively diminishing resistance to friction. Therefore, a typical graph (Fig 7) shows a fairly constant resistance of the skin until a point where the skin's resistance to displacement decreases, presumably because the epidermal cells or their connections begin to break down. Then a marked and progressive decrease of force occurs until the total area of the blister top is loosened and moveable. At this point the resistance assumes a secondary (lesser) but constant force. By watching the progressively decreasing slope-off and ensuing leveling into a lesser torque, the investigator can terminate the application of friction at the earliest moment of complete production of the blister cleft, or at any selected point prior to, or after, this.

OBSERVATIONS AND RESULTS

The results obtained with the earlier instruments plus the measurements of the precise end-points with the latest instrument just described, include the following:

*FOOTNOTE. We are grateful to Robert N. Brown, M.E. and SP4 Gerald L. Blase, B.S.M.E. for their expert contributions to the design, development and testing of this model. Details are now being submitted for publication.

The time required to produce friction blisters varies from individual to individual, from skin site to skin site. These variations are due to factors which are in the main unknown. However, certain variables have been shown to influence the time required: A) Moisture: Dry skin surfaces and wet, flooded skin surfaces produce less friction than barely moist skin surfaces when rubbed with the same material (Fig 8). This finding confirms the *in vivo* studies of Naylor (9). It also confirms the studies of Appeldoorn and Barnett (13) who used not skin *in vivo*, but a model they constructed of artificial materials. B) Temperature: The friction blister is not a thermal blister - not a second degree burn. This conclusion is based not only on differences in microscopic morphology but mainly on the fact that in man we could never show a rise of more than 1.5°C, either immediately beneath the tissue being rubbed or the material rubbing the tissue. The time required to produce friction blisters is to some degree temperature dependent. For as shown by Griffin, et al (14), blisters can be produced somewhat more quickly when the tissue is warmed, than when it is cooled during the rubbing.

We have as yet no accurate measurements of thresholds for the friction blister response in different persons, at different times or at different sites. However, we have found that, to produce friction blisters at all, there must be sufficiently tough and resistant superficial layers of skin to withstand the surface rubbing and to form the intact blister top - otherwise the skin is rubbed raw with no blister cavity capable of holding free fluid. Moreover, the tissues forming the blister base must be sufficiently firmly held from below, so that they do not move freely with the upper layers which are being displaced by the frictional forces. It is the differential movement of the upper layers to-and-fro over the lower ones which produces the shearing effect between the more superficial and the deeper epidermal structures. This eventually results in a split or cleft into which the blister fluid then flows or transudes. These conditions are satisfied best by the skin of the palms, soles, heels, sides and tops of the toes and feet, and similar "tough" areas where the tissues are quite firmly held down by the underlying structures. In these areas experimental friction blisters can best be produced; and in these areas clinical friction blisters occur almost exclusively.

In some of our experiments the metal rubbing head was covered with various materials, such as smooth or rough shoe leather, plastics or rubber. Over this were fastened patches from different types of Government issue socks. The friction produced was found to be more a function of the backing material than of the sock material which actually rubbed the skin.

Friction blistering is pathogenetically a two-step process. The first step, described above, is the differential movement which creates a tear resulting in a split or potential cleft. The second

step is the pouring of the blister fluid into this cleft, thus forming the fluid-filled blister cavity. Remarkably, and for some as yet unexplained reason, both in our experimental model and in clinical friction blisters, the tear or cleft always occurs in about the same region within the epidermis - this is at the lower part or below the granular layer and above the basal cell layer (Figs 9,10). This cleft can be regularly produced on suitable areas of cadaver skin - but, of course, there is no vis a tergo and therefore no fluid filling. Furthermore, when the arterial supply to the skin site is shut off by a tourniquet or reduced by elevation, the friction still forms the potential cleft, but this cleft does not fill with fluid until the arterial pressure to the limb is restored. (Fig 11, 12, 13).

Because the foot pads of monkeys meet the conditions of toughness of upper skin layers and relative fastening down of the lower layers, Cortese, Griffin, Layton and Hutsell (8) applied our rubbing technique to these areas in two species of Rhesus monkeys. They were rewarded with success.

The production of grossly and histologically typical friction blisters in monkeys has permitted multiple experiments on temperature rises in the tissues and on the effects of environmental and skin temperature changes upon the blister production time. The results of larger series of experiments in these laboratory animals confirmed those of the smaller series in man (Fig 14). The successful production of friction blisters in laboratory animals now affords opportunities for further study of these lesions.

As stated, the entry of fluid into the blister cleft depends upon a certain amount of hydrostatic pressure in the vascular bed of the site, inferring that the blister fluid is derived in some way from the blood. The communication between blood vessels and blister cavity is not direct, for the blister cavity takes from 1 to 2 hours to become fully filled. Moreover, the friction blister fluid characteristically contains few cells, and does not clot. It contains no platelets, fibrinogen or fibrolysin or only in the minutest amounts. (Peter Schmid, PhD, in our laboratory). Nevertheless, the blister fluid does contain many of the constituents of the blood serum, e.g. its proteins, including the relatively large molecules such as syphilitic reagins, myeloma proteins, - albeit in substantially lower concentrations than in the subjects' serum at the same time (5) (Table 1).

The fluid, therefore, appears to be either a selective transudate or due to some type of active transport. The mechanisms and dynamics of the friction blister fluid formation and its passage into and out of the blister cleft are unknown. They are now accessible to investigation, utilizing the experimental model herein described.

PROPHYLACTIC AND THERAPEUTIC CONSIDERATIONS

Our latest instrument for the first time permits the precise measurements of the period required to produce the blister, including the precise moment of beginning cleavage and precise moment of completion of the blister cleft. With these end-points one can now begin to study the problems of skin "hardening", of lubrication of the skin's surface, or of reducing friction by powders, oils and other materials in order to promote slippage and reduce blistering. One can also begin to systematically investigate the selection of fabrics and materials or the design of footwear and other gear to reduce friction over the most hazarded points.

Since friction blisters are not burns but are due to shearing forces in the epidermis, on purely theoretical grounds, foot-gear which takes up the differential movement, not in the epidermis but before it reaches the skin surface, would be likely to reduce blistering.

Although our results to date have yielded no practical preventive measures, they suggest the trial of two recommendations which may prove practical in regard to treatment of the already formed blisters. Cortese, et al; (6) showed that if one drains intact blisters not earlier than 24 hours and not later than 48 hours after their production there is the greatest likelihood that the blister top will paste down firmly and remain intact despite considerable trauma and friction. The firmly pasted-down top (Fig 15), reduces pain and infection and makes an excellent functional dressing. However, on marches, in the field and in other activities, one is not usually fortunate enough to come upon and drain and treat intact blisters. Most are broken and have no tops or only tattered fragments of tops by the time the medical aidman or physician sees them.(Fig 2). Isoamyl cyano-acrylate was found useful by Akers in the early treatment of 35 soldiers with foot blisters at Ft Ord. The tattered fragments of the blister tops were removed and the liquid was applied to the raw blister base with a cotton swab. It polymerized rapidly to form an artificial top, which covered the denuded blister base for many days. Pain was lessened, walking was comfortable, and no infection occurred.

This isoamyl cyano-acrylate was selected by Cortese and Osterhout from the series prepared by Fred Leonard and others at the Army Biomechanical Laboratory. Homologues have been successfully used by Matsumoto, Hardaway and others (15) to stop bleeding of internal organs. Upon completion of the animal toxicity studies the isoamyl cyano-acrylate treatment of blisters will be tested in field trials.

REFERENCES: (1) Delaney and Travis, Military Med. 1965. (2) Sulzberger, Cortese, Fishman and Wiley, J. Invest. Derm. 1966.

(3) Sulzberger, Cortese, Sams and Wiley, Acta Derm-Vener. 1966. (4) Cortese, Griffin and Sulzberger, Proc XIII Internat. Cong. of Dermat. 1967, Springer, 1968. (5) Cortese, Sams and Sulzberger, J. Invest. Derm. 1968. (6) Cortese, Fukuyama, Epstein and Sulzberger, Arch. Dermat. 1968. (7) Sulzberger and Cortese, Brit. J. of Clin. Pract. 1968. (8) Cortese, Griffin, Layton and Hutnell, J. Invest. Derm. 1969. (9) Naylor, Brit. J. Dermat. July 1955. (10) Idem. Brit. J. Dermat. Oct. 1955. (11) Idem. Brit. J. Dermat. 1965. (12) Goldblum and Piper, J. Invest. Derm. 1953. (13) Appeldoorn and Barnett, Proc. Scien. Sect. Toilet Goods Assoc. 1963. (14) Griffin, Cortese, Layton and Sulzberger, submitted J. Invest. Derm. July 1969. (15) Heisterkamp, Matsumoto and Hardaway, Milit. Med. 1969.

TABLE I Comparison between contents of friction blister fluid and blood serum and blood plasma.

Relative Concentration of Electrolytes in Blister Fluid	Serum=1.0
Sodium	0.99
Potassium	1.04
Chloride	1.00
Calcium	0.78

Relative Concentration of Protein in Blister Fluid	Plasma=1.0
Protein+	0.22
Albumin*	0.11
IgG*	0.11
IgM*	0.11
Fibrinogen*	0.00

\+ 4 hr. blister
* 2 hr. blister

Fig 1. Friction damage and consequent infection. Combat soldier of 9th Infantry Division, Mekong Delta.

Fig 2. Denuded blisters as first seen on foot of "drop-out" during 16 mile Walkathon, Sonoma-Petaluma, California.

Fig 3. Simultaneous application of two twisting pencil erasers to produce friction blisters on symetrically situated sites of both palms

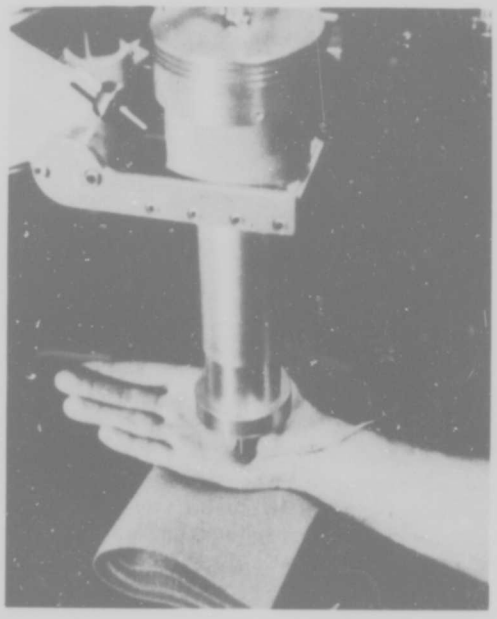


Fig 4. Apparatus for twist rubbing with measurable forces. Note thermistor in skin under site being rubbed, to measure rise in tissue temperature. Elevation was never over 1.5°C.

Fig 3. Simultaneous application of two twisting pencil erasers to produce friction blisters on symetrically situated sites of both palms

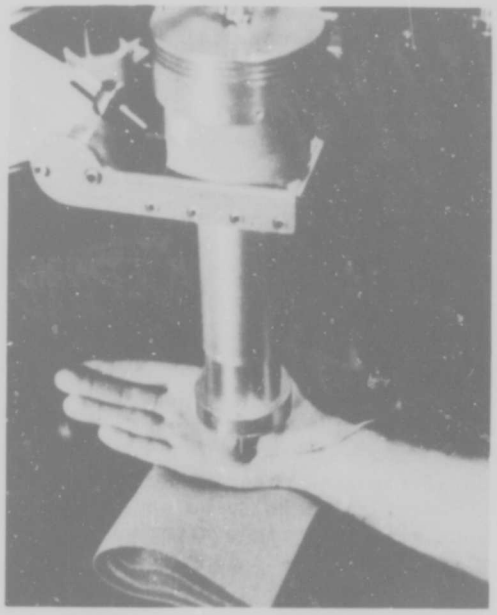


Fig 4. Apparatus for twist rubbing with measurable forces. Note thermistor in skin under site being rubbed, to measure rise in tissue temperature. Elevation was never over 1.5°C.

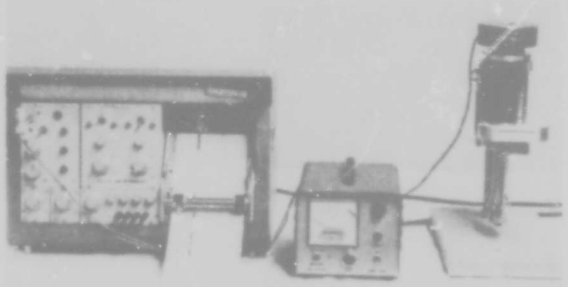


Fig 5. Latest model friction blister instrument permitting read-out of beginning and completion of formation of blister cleft.

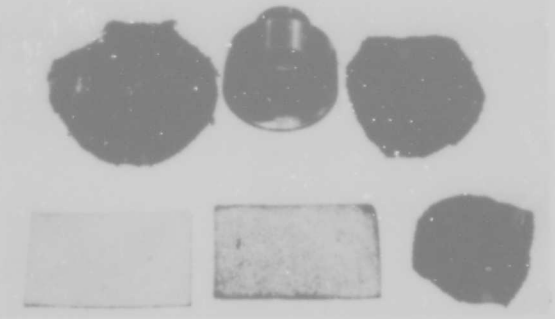


Fig 6. Rubbing head of latest model with sock material attached; Note surrounding retention ring to prevent rubbing head from going deeper. Samples of issue socks, shoe and boot leathers.

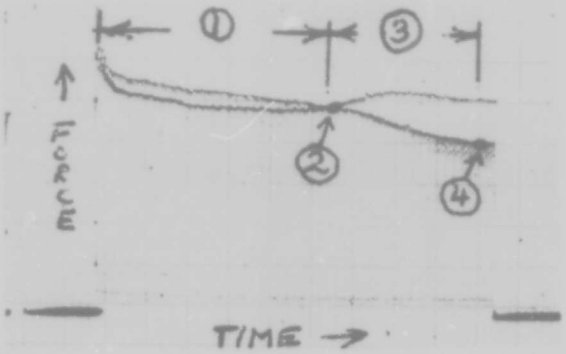


Fig 7. Typical Graph: (1) Period of repeated shearing forces. (2) Point where intraepidermal connections begin to be severed. (3) Period of increasing compliance of skin surface with movement of rubbing head and formation of cleft. (4) Cleft is complete and frictional forces assume a constant, lower, level.

Fig 8 (Above)
Graph of linear rubbing while water is being added slowly. Beginning at bottom: Dry leather and skin= low friction. Middle: Barely moist surfaces= increased friction. Top: Skin completely flooded=diminished friction.

Fig 9. Blister cleft. Typical histologic picture. Site of cleavage in epidermis always above basal cells, in and/or below granular layer.

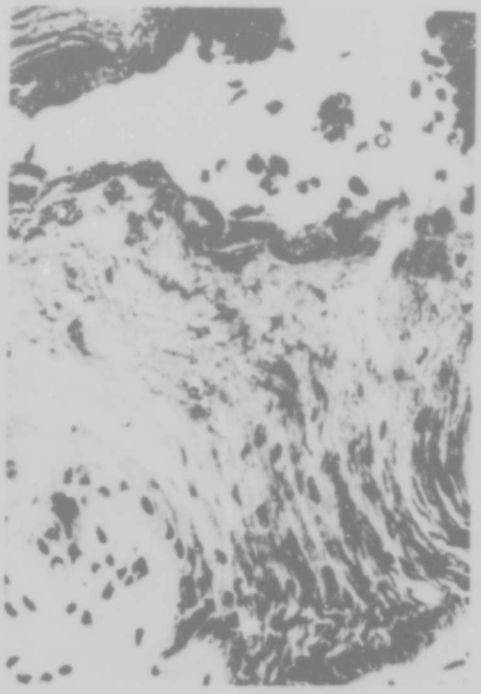


Fig 10. Histology immediately after cleft formation. Note elongation and distortion of epidermal cells which have been pulled and torn by shearing forces.

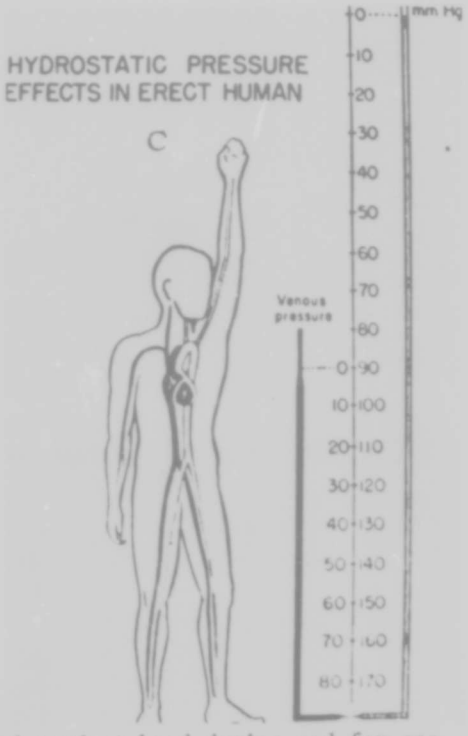


Fig. 11. Position of elevated and pendant hand during and for one hour after rubbing. On right, drawing indicating hydrostatic pressure (Rushmer).

Fig. 12. Blister on pendant palm, 2 hours old--fully filled.

Fig. 13. Palm which was elevated for 1 hour. Blister now 2 hours old and still flaccid after lowering hand.

Fig 14. Monkey palm with numerous experimental friction blisters in various stages of development and healing.

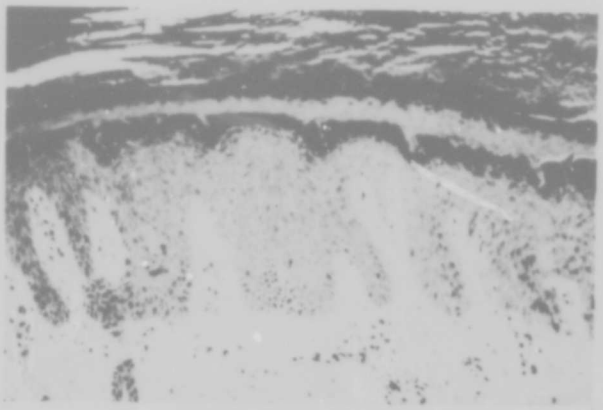


Fig 15. Histology showing pasted-down blister top adherent to new epithelium of blister base. Biopsy 48 hours after friction blistering and 24 hours after a single drainage. This firmly attached top makes an excellent dressing. When tops have been torn, application of a cyano-acrylate makes a serviceable artificial top.

AD 713 504

TEMPORARY CAVITY EFFECTS IN BLOOD VESSEL INJURY BY HIGH VELOCITY MISSILES

JOSEPH J. AMATO, MAJ, MC, NORMAN M. RICH, LTC, MC, NOEL S. LAWSON, MAJ, MC, RONALD P. GRUBER, CPT, MC, AND LAWRENCE J. BILLY, CPT, MC
BIOMEDICAL DEPARTMENT, BIOPHYSICS LABORATORY
RESEARCH LABORATORIES, EDGEWOOD ARSENAL, MARYLAND 21010

The wounding potential of high velocity missiles is greatly accentuated by the explosive effects of the temporary cavity. The missile transfers energy to the tissues within microseconds after impact. Tissue along the wound tract expands laterally forming a cavity with internal pressures as high as 100 atmospheres (1,500 pounds per square inch).[1] Description of the extent and nature of destruction to the various tissues by this phenomenon remains incomplete, and the effect upon blood vessels has been controversial.

Low velocity missiles have been known to push the blood vessels aside as the missile traverses the tissue. An early account of these injuries was described in 1862 by George Macleod in War in the Crimea.[2] Huguier in 1848 first observed the bursting effect of high velocity missiles on soft tissues.[3] He attributed this effect to the dispersion of water particles. This hydrodynamic theory was first experimentally tested by Kocher in 1876[4] and confirmed by Sir Victor Horsely in 1894.[5] Destruction of tissues appeared to correlate with the amount of fluid within the tissues. In 1898, Woodruff applied a concept of marine engineering to wound ballistics. The term cavitation was used to describe the vacuum immediately following a solid object moving rapidly through a gas or fluid.[6] He predicted the formation of a cavity following a high velocity projectile and very accurately illustrated its movements. Stevenson in 1910[7] and LaGarde in 1916[8] accepted the cavitational theory which was also described as the theory of accelerated particles. Lesions to the blood vessels were described as one of the chief characteristics of the high velocity bullet. They stated that the arteries are neatly sheared as though it were done with a cutting instrument leaving the edges without irregular lacerations. Observations of the effect of the temporary cavity upon vessels were first made by Wilson in 1921.[9] He described slit-like lesions in areas distant to the segment actually touched by the bullets. Similar lesions occurred in near misses.

Photographic documentation of the formation of the temporary cavity was first obtained by Black and associates.[10] Contrary to Wilson's observations, they stated that tissues of a highly elastic nature such as arteries were able to escape both anatomical and functional injury. Experiments by Harvey et al at Princeton in 1946 verified this opinion and demonstrated for the first time that the artery when not directly injured by the missile was pushed aside by the temporary cavity.[11] They also concluded that the blood vessel escaped gross injury.

In the recent Korean Conflict, Jahnke and Seeley[12] described microscopic changes within the blood vessel wall adjacent to the grossly injured segment. Failures of repair and increased morbidity were ascribed to these changes. They arbitrarily advocated resection of one centimeter of artery beyond the area of gross damage for more successful surgical repair of the vessels. The need for further investigation was stressed.

Experiments were designed in our laboratory to study the pathophysiology of ballistic injury to the blood vessels. The mechanism of wounding in arteries has been previously reported.[13,14] This communication is concerned with the effects of the temporary cavity including wounds produced by the M-16 bullet. Sequential events which occur in the mechanism of injury by both direct and indirect hits are documented by angiography and high speed photography. To our knowledge we have presented the first thorough documentation of the effect of the temporary cavity upon arteries.[13,14] Observations of microscopic mechanical injury are also discussed.

METHOD

In the initial experiment, the femoral arteries of thirty anesthetized dogs were exposed and injured with missiles of specific velocities. A 0.25 inch sphere weighing 16 grains was accurately calibrated to be fired at 1,000, 2,000 and 3,000 feet per second.

X rays of the extremity were taken at intervals following injection of 90 percent hypaque in some vessels. The Fexitron flash X ray system which is capable of producing pulses of one-tenth microsecond duration was used. Instrumentation was designed to inject the dye, fire at the vessels, and obtain X rays at varying intervals.

High speed photography of the mechanism of injury was taken at 3,750 to 4,500 pictures per second utilizing a Hycam Camera. This camera is a 16 mm high speed rotating prism camera used especially for photography of short-lived events.

To study the microscopic mechanical changes immediately post wounding, segments of the injured vessel approximately 20 millimeters on either side of disruption were excised. These were prepared in the normal manner with Hematoxylin and eosin stain. The Verhoeff-VanGieson method for specific elastic and connective tissue staining was used.

A second experiment was designed to visualize more adequately the vessel in an isolated environment. Segments of the femoral arteries measuring six to nine centimeters in length were removed from fourteen dogs. These vessels were filled with safranin dye or hypaque, weighted on one end and suspended within a 20 percent gelatin solution which was used to simulate normal muscle tissue density. These vessels were sighted, injured and photographed as described above.

Because of the small caliber of these arteries, a third series of experiments were performed utilizing twenty segments of calves vessels. Both width and thickness of these arterial walls approximated the size of the human femoral vessels. Injury was produced in the first ten segments with 0.25 inch spheres so that we could solely obtain the effects of the temporary cavity at specific varying velocities. The M-16 rifle bullet was used in the remaining ten segments to document other factors effecting the behavior of the military bullet. The M-16 rifle is currently used in Vietnam and utilizes a 55 grain bullet fired at a muzzle velocity of 3,250 feet per second.

RESULTS

Angiograms of the femoral artery in situ are shown at microseconds after impact (Fig. 1). The 0.25 sphere struck the artery at a velocity of 3,000 feet per second. The vessel was neatly transected by the sphere without moving the remaining artery. However, within five to ten microseconds after impact, the proximal and distal segments of the injured artery were compressed against the cavity wall (Fig. 2). High speed photography demonstrated the suspended vessel in gelatin which was struck by a sphere at 3,000 feet per second (Fig. 3) with the subsequent formation of the temporary cavity. Again, it is seen that the vessel was cut by the missile but violently hurled sideways and compressed laterally (Fig. 4,5,6).

The M-16 injury was similar to that of the sphere. The calves' vessels were stretched by the undulation of the temporary cavity in an indirect hit (Fig. 9,10). Stretching and swaying of the torn ends were documented when the vessel was struck directly by the 55 grain bullet at a striking velocity of 3,250 feet per second (Fig. 11,12).

The contrasting effects of the low velocity missile which struck the gelatin block at 1,000 feet per second were seen with minimal formation of a temporary cavity (Fig. 7) and stretching forward of the blood vessel until it was severed (Fig. 8).

Microscopic mechanical changes in these blood vessels correlate well with previous descriptions of changes along the blood vessel wall.[9,12,20] The endothelium was frequently lost. Small and large breaks were seen in the internal elastic membrane. There were microthrombi formation within the media with disruption and less commonly herniation of the media through the external elastic membrane of the adventia.

DISCUSSION

The wounding potential of high velocity missiles is a complex subject which is beyond the scope of this paper. However, several important factors, must be included in discussing ballistic behavior of the bullet in relation to wounding capacity. These are: A. The energy transmitted to the tissues by the missile. B. The direction of the transmitted energy released as determined by the shape and motion of the missile. C. The density of the tissues through which it travels.

A. The kinetic energy theory of wounding capacity indicates that the energy expended is directly proportional to the mass and to the square of the velocity ($KE = M \times \frac{V2}{2g}$. Thus, doubling of the velocity quadruples the kinetic energy of the missile. As the missile comes in contact with the artery, it bisects the wall sharply as demonstrated in Figures 1 and 3, but energy is transferred to the tissues. The arterial ends are accelerated away from the wound track radially within the formation of the temporary cavity (Fig. 2-6). The development and decay of the cavity occurs within 20 microseconds after impact causing the segments to undulate and be thrashed about within the cavity until it comes to rest (Fig. 12). Cavitation is primarily due to the energy released. However, both its size and shape are modified by the configuration of the missile and striking angle at moment of impact. To avoid the above factors a steel sphere was utilized in the majority of our experiments as was suggested by Woodruff in 1898.[6]

B. The direction of the energy released to the tissues is governed by shape of the missile and its motion in flight. A brief description of the bullet's motion would include yawing, tumbling, precession and nutation. When a bullet deviates from the straight line of flight, in its longitudinal axis, it is described as yawing. Tumbling is simply described as the action of forward rotation around the center of mass. As the bullet yaws and/or tumbles, it is also spinning laterally around its center of mass. This is termed precession. A fourth motion of the bullet in flight and perhaps the most difficult to visualize is nutation. Hopkinson appropriately describes nutation as the rotational movement of the bullet in small circles progressing forward in a rosette pattern.[15] The varying severity of wounds and amount of tissue destruction by identical missiles can be explained in most part by the angle of the bullet to the tissue at moment of impact.

The density of tissues is a third factor in the explanation of wounding effects on tissue. As mentioned earlier, Kocher in 1876[4] demonstrated that tissues which contained large quantities of water were most severely damaged. Daniel in 1944 demonstrated that the severity of high velocity wounds increased with the specific gravity of the tissue struck.[16] DeMuth stated in his review that bone injuries are most severe while lung tissues are least affected.[17] Injury to the blood filled arteries can be extensive. This perhaps, is only modified by the elastic nature of the artery and its ability to absorb the energy transmitted by the missile. When

large amounts of energy exceed the tolerable elastic limits of the cellular structures within the arterial wall, microscopic injury can occur. When the vessel is indirectly struck by the missile, the amount of injury is proportional not only to the velocity of the missile, but to the proximity of the vessel to the center of the temporary cavity, and to size of the blood vessel. This mechanism of injury is not clearly understood by a majority of surgeons faced with arterial repair.

The extent of debridement of the injured vessel has been controversial. Wilson stated that the slit-like lesions within the intima may cause thrombosis, fatal secondary hemorrhage or aneurysmal dilatation at a later date.[9] The treatment for arterial injury at that time was simple ligation. In the past the advice of resection of one centimeter of arterial wall beyond the area of gross injury by Jahnke and Seeley[12] has been followed by a majority of military surgeons. However, others have advocated the removal of one to two centimeter segments in high velocity wounds.[18] Contrary to the above, one author has shown the extent of microscopic injury to be minimal and debridement of only 3 millimeters of tissue was deemed necessary.[19]

Recently, Rich, Manion and Hughes reviewed a series of clinical cases from the Vietnam Conflict and stated that there was no correlation between the microscopic damages which were present and the success of repair.[20] Adequate debridement with good surgical judgement was advocated. In our series, microscopic examination of the segments removed post injury have demonstrated as in humans, various changes within the arterial wall. The significance of the changes mentioned earlier remain unanswered at present in our experimental model.

An important aspect of this study is to realize that the low velocity gunshot wound is a different entity. The mechanism of injury is seen as a local stretching of the vessel wall with little or no formation of a temporary cavity. Because of this difference, proper assessment of the wound and knowledge of the nature of the wounding missile are necessary requirements in proper management of missile wounds.

SUMMARY

1. Historically, the effect of the temporary cavity upon arteries has been controversial.

2. We have presented angiographic and photographic documentation of low velocity and high velocity injuries to arteries.

3. The low velocity missile pushes the blood vessel ahead stretching it slightly before penetration and can be pushed aside with little or no formation of the temporary cavity.

4. The high velocity missile neatly cuts the arterial wall at impact but then the explosive effect of the temporary cavity crushes the ends of the artery against its walls.

5. Severity of arterial damage is proportional to the velocity of the wounding missile. The energy transferred from the high velocity missile forms the temporary cavity. Other factors, such as, the configuration of the missiles vary the size and shape of the temporary cavity and extent of injury.

6. Varied tissue destruction by identical missiles at equal velocity can be explained by the shape of the missile, its motion in flight, and most important the angle to the tissue at moment of impact.

7. Knowledge of the ballistic behavior of high velocity missiles in causing arterial damage is necessary for the proper evaluation, and management of missile wounds.

REFERENCES

1. Herget, C.H. 1953. "Wound Ballistics" In Surgery of Trauma. J.B., W.B. Bowers (ed) Philadelphia, J.B. Lippincott & Co.

2. Macleod, George H.B. War in the Crimea, Philadelphia, J.B. Lippincott & Co., 1862 pp109.

3. Huguier, cited by Horsely.(5)

4. Kocher, cited by Horsely.(5)

5. Horsely, V. The Destructive Effects of Small Projectiles. Nature 50 p106 (May 31,) 1894.

6. Woodruff, C.E. The Causes of the Explosive Effect of Modern Small Calibre Bullets. The New York Medical Journal 67:p593-594, (April 30,) 1898.

7. Stevenson, W.F. Wound in War. The Mechanism of their Production and their Treatment, ed 3 London, 1910 pp54-56.

8. LaGardi, L.A. Gunshot Injuries. How they are Inflicted, their Complications and Treatment. ed 2 New York, 1916 p55.

9. Wilson, L.B. Dispersion of Bullet Energy in Relation to Wound Effects, The Military Surgeon 159:pp249-250 September 1921.

10. Black, A.N; Burns, B.D., and Zuckerman, S. An Experimental Study of the Wounding Mechanism of High Velocity Missiles. British Medical Journal 2 pp872-874 (Dec 20,) 1941.

11. Harvey, E.N. et al. Mechanism of Wounding, War Medicine 8 p102, 1945.

12. Jahnke, E.J. and S.F. Seeley, Acute Vascular Injuries in the Korean War. Annals of Surgery 138, pp158-177, 1953.

13. Amato, J.J. et al. High Velocity Arterial Injury: A Study of the Mechanism of Injury. pending publication.

14. Amato, J.J. et al. Vascular Injuries: An Experimental Study of High and Low Velocity Missile Wounds, read before the Central Surgical Association, Detroit, Michigan February 27, 1970.

15. Hopkinson, D.A., Marshall, T.K., Firearm Injuries. British J. Surg. 54 (May) 1967.

16. Daniel, R.A., Jr. Bullet Wounds of the Lungs: An Experimental Study. Surgery 15, pp744-782, 1944.

17. DeMuth, W.E.Jr. Bullet Velocity and Design at Determinants of Wounding Capacity: An Experimental Study. Journal of Trauma 6, p226, 1966.

18. Bradham, R.R., Buxton, J.T., and Stallworth, J.M. Arterial Injury of the Lower Extremity. Surg. Gyn. Obst. 118: p995, 1964.

19. Moore, H.G. et al. Gunshot Wounds of Major Arteries. An Experimental Study with Clinical Implications. Surg. Gyn. Obst. 98-145, 1954.

20. Rich, N.M., Manion, W.C. and Hughes, C.W. Surgical and Pathological Evaluation of Vascular Injuries in Vietnam. Journal of Trauma 9:279, 1969.

LEGENDS OF FIGURES

Figure 1. Angiogram of the femoral artery within microseconds after impact. Impact velocity was 3,000 feet per second.

Figure 2. Angiogram of the femoral artery compressed against the wall of the maximal temporary cavity.

Figure 3. A segment of femoral artery is neatly sheared by a 16 grain sphere at 3,000 feet per second.

Figure 4. The initial phase of the temporary cavity is seen. The arterial segments are bent laterally.

Figure 5. An intermediate phase of cavitation is demonstrated. Pressures as high as 100 atmospheres have been reached.

Figure 6. Maximum formation of the temporary cavity. Note the curled segment of vessel in the upper half of the cavity.

Figure 7. High speed photograph of a segment of femoral vessel being approached by a sphere at 1000 feet per second.

Figure 8. The artery is shown stretched forward prior to being severed by the missile.

Figure 9. A M-16 bullet was fired at 3,250 feet per second missing the femoral artery. The artery however is slightly stretched.

Figure 10. The vessel which was missed by the M-16 is thrashed about and distorted laterally by the temporary cavity.

Figure 11. A direct hit with an M-16 bullet severs the vessel and compresses the ends forcefully. The crushed segments can be seen against the wall.

Figure 12. The distal and proximal ends of the femoral vessel sway within the undulating temporary cavity until it comes to rest.

Figure 1. Microseconds after impact at 3000 feet per second.

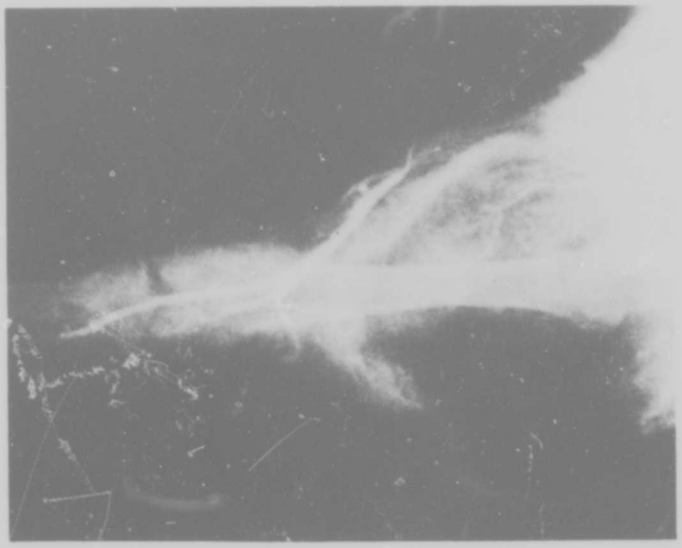


Figure 2. Maximum formation of the temporary cavity.

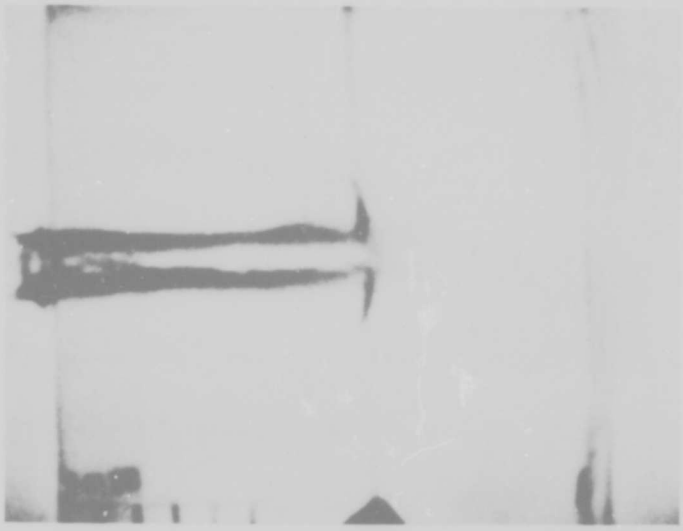


Figure 3. The vessel is transected by the missile in gelatin.

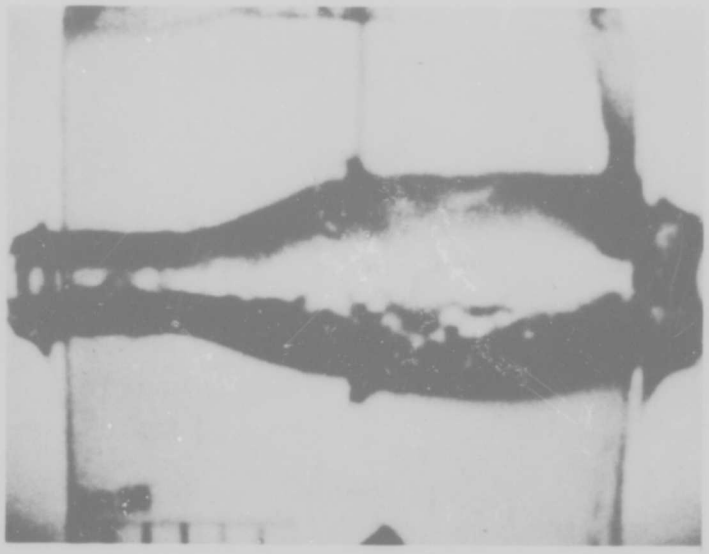


Figure 4. Early formation of temporary cavity.

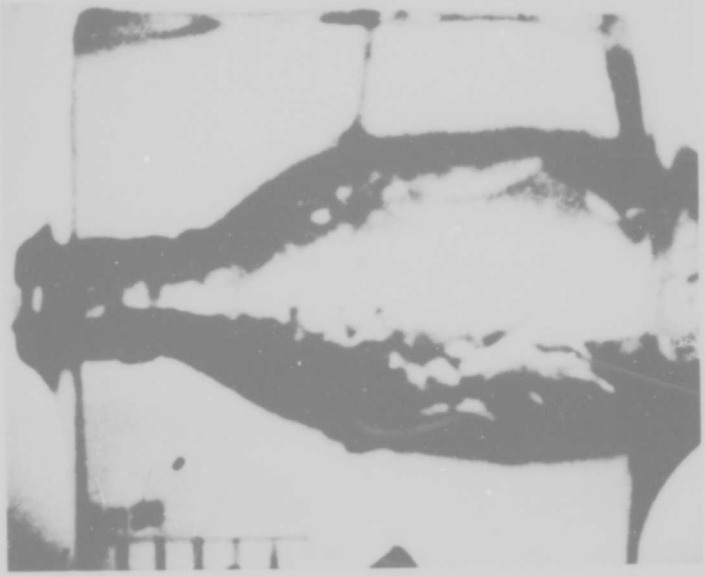


Figure 5. Intermediate phase of temporary cavity.

Figure 6. Maximum formation of the temporary cavity.

Figure 7. Vessel is approached at 1000 feet per second.

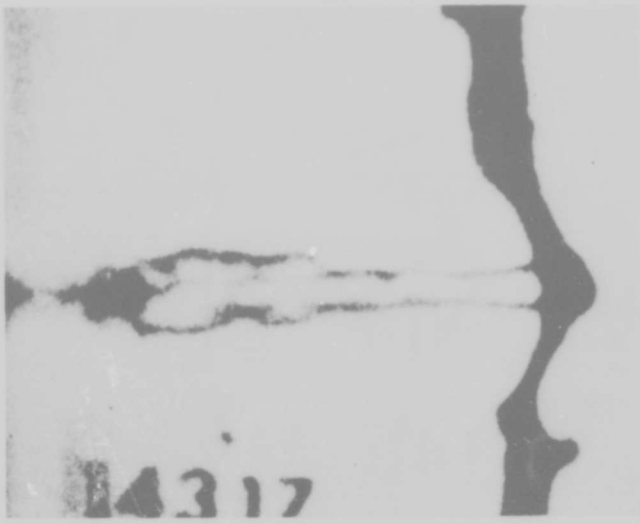


Figure 8. Vessel is stretched by the low velocity missile.

Figure 9. Near miss of vessel with M-16 bullet.

Figure 10. Stretching of the vessel - near miss M-16 bullet.

Figure 11. Direct hit with the M-16 - vessel is compressed laterally.

Figure 12. M-16 direct hit - vessel is thrashed about.

AD 713 505

AN INVESTIGATION OF THE DYNAMIC PRESSURE RESPONSE OF FLUERIC TRANSMISSION LINES

LARRY C. ATHA
ADVANCED BALLISTIC MISSILE DEFENSE AGENCY
HUNTSVILLE, ALABAMA

INTRODUCTION

In recent years, a new technology known as fluidics, or fluerics, has been developing. Flueric devices operate on the principle of using one fluid stream to control another fluid stream. By the elimination of moving mechanical parts, these devices offer promise of simplicity, high reliability, and practically unlimited life. In addition, flueric devices are operable in hostile environments such as extreme temperatures, high shock levels, high vibration levels, and nuclear radiation. Best of all, these devices promise to be inexpensive.

Even though flueric systems seem to offer so many advantages, they are in very limited use at the present time. One of the most significant reasons for this is the difficulty in connecting active flueric elements together with flueric transmission lines into systems which will operate over useful frequency ranges. A typical flueric system contains sensors and several types of flueric elements connected together by flueric transmission lines. A complete analysis of the system must include not only the flueric elements but also the flueric transmission lines. This is particularly true for dynamic operating conditions when the dynamic effects of the flueric transmission lines have, in several cases, caused otherwise well designed systems to be inoperable.

The scope of this investigation can be summarized in three parts: (1) An examination and review of pertinent literature; (2) The development of a mathematical model of a flueric transmission line suitable for use by a design engineer; and (3) The experimental determination of the dynamic response of flueric transmission lines over a wide frequency range with subsequent verification of the mathematical model.

Since the advent of flueric elements and systems, interest in flueric transmission lines has been stimulated. Most previous work in this area has been directed toward application in the process control field. The majority of this work was done for either blocked or volume terminated pneumatic transmission lines. In general, the work was directed toward determining the response of long lengths of pneumatic lines (to 2,000 feet) to slowly varying input pressures (frequencies of less than 10 hertz). Other investigators have dealt with the transmission of sound in tubes. This work was concerned with small amplitude signals in the 500 to 5,000 hertz range. Thus, a wide gap has been left in the region of current interest. A detailed review of the literature is presented(1). To summarize, the results of the previous studies were not applicable to the flueric transmission line problem for several reasons. Flueric transmission lines are short, usually not over 8 to 10 feet in length. They are not blocked or volume terminated, but are terminated by small orifices through which there is some net flow. Flueric lines operate at high frequencies, sometimes as high as several kilohertz.

In order to utilize small, high-frequency fluid elements in practical flueric systems, a better understanding of the dynamics of the flueric transmission lines which connect these elements is needed. This study was made to, at least, partially fill this need.

THE DEVELOPMENT OF A MATHEMATICAL MODEL

Practical flueric systems involve several flueric elements interconnected with fluid transmission lines. An analysis of the dynamics of this type of system must include the dynamic effects of each element as well as those of the interconnecting fluid lines. At very low operational frequencies, the dynamic effects of fluid lines are minimized, but at higher frequencies, the dynamic effects become very important. A need exists for a practical mathematical model of a flueric transmission line that can be used by the practicing engineer in the dynamic analysis of complex fluid systems. In this section, a mathematical model of a flueric transmission line, based on a system of evenly distributed parameters, is presented. The model is then used to predict the dynamic response of different lengths of flueric transmission lines to sinusoidal and square wave input pressures of varying frequencies. The digital computer programs that were used in this study are listed in Reference 1 along with sample calculations and other pertinent information.

An electrical transmission system(2) can be represented as a combination of three separate portions: the energy source, the load, and the transmission line. Four distributed parameters are required to model the transmission line which can be represented schematically as shown in Figure 1. These parameters are: R, the resistance per unit length; L, the inductance per unit length; G, the conductance per unit length; and C, the capacitance per unit length. By use of these parameters, several other quantities can be defined. They are: Z, the

series impedance per unit length; Y, the shunt admittance per unit length; γ, the propagation constant; and Z_C, the characteristic line impedance per unit length. These quantities can be defined as:

$$Z = R + j\omega L \tag{1}$$

$$Y = G + j\omega C \tag{2}$$

$$Z_C = \sqrt{\frac{Z}{Y}} = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \tag{3}$$

$$\gamma = \sqrt{ZY} = \sqrt{(R + j\omega L)\,(G + j\omega C)}, \tag{4}$$

where $(j\omega)$ is a Fourier transform. A relationship between the output voltage, E_o, and the input voltage, E_i, of the transmission line can be written as a function of the variables listed above, the length of the transmission line, L, and the load impedance, Z_L. It is given by equation (5):

$$\frac{E_o(L,\ t)}{E_{i(0,t)}} = \frac{2Z_L\, e^{-\gamma L}}{(Z_L + Z_C) + (Z_L - Z_C)\, e^{-2\gamma L}}. \tag{5}$$

By the use of a pneumatic-electrical analogy, the method described above can be used for the analysis of pneumatic transmission lines, and the considerable body of information on electric transmission lines now in the literature is made available for the analysis of pneumatic transmission lines. The pneumatic-electrical analogy consists of making pressure analogous to voltage and volumetric flow rate analogous to current. Then, distributed pneumatic parameters of resistance per unit length, R_p; capacitance per unit length, C_p; and inductance per unit length, L_p, can be defined which correspond to the analogous electrical parameters of resistance, capacitance, and inductance. These pneumatic parameters are listed below:

$$R_P = \frac{8\pi\mu}{A^2} \tag{6}$$

$$L_P = \frac{\rho}{A} \tag{7}$$

$$C = \frac{A}{k\,\overline{P}} \tag{8}$$

where: A = the effective area of a cylindrical pneumatic transmission line, one fourth of the cross-sectional area.
μ = the mean value of the viscosity of the fluid.
ρ = the mean value of the density of the fluid.
k = the ratio of the specific heats of the fluid.
$\overline{P}$ = the mean value of the pressure.

The pneumatic conductance per unit length, Gp, is zero. Then, the pneumatic parameters of series impedance per unit length, Z_p; shunt admittance per unit length, Y_p; the propagation constant, γ_p; and the characteristic line impedance per unit length, Z_{CP}, can be defined as:

$$Z_P = R_P + j\omega L_P \tag{9}$$

$$Y_P = j\omega C_P \tag{10}$$

$$\gamma_P = \sqrt{(R_P + j\omega L_P)(j\omega C_P)} \tag{11}$$

$$Z_{CP} = \sqrt{\frac{R_P + j\omega L_P}{j\omega C_P}} \quad . \tag{12}$$

The relationship between the output pressure of a flueric transmission line and the input pressure can be written as a function of the variables listed above, the length of the flueric transmission line, L, the pneumatic load impedance, Z_{LP}, and is:

$$\frac{P_o(L,t)}{P_i(0,t)} = \frac{2Z_{LP}\, e^{-\gamma_P L}}{(Z_{LP} + Z_{CP}) + (Z_{LP} - Z_{CP})\, e^{-2\gamma_P L}}, \tag{13}$$

where the pneumatic load impedance is the total load impedance seen by the fluid in the transmission line. In general, this load consists of the impedance of the load orifice in parallel with the pressure transducer impedance. The impedance of the load orifice can be found determining the slope of the pressure versus flow curve of the load orifice at the operating point. The transducer impedance can be determined experimentally or, in some cases, is available from the manufacturers' literature.

Numerical determinations of frequency response of flueric transmission lines in the form of Bode plots were made. A digital computer program was written in FORTRAN IV to determine the frequency response of flueric transmission lines by using the model that was developed. The program was then used to obtain the frequency response of flueric transmission lines of 2, 4, 6, and 8 feet in length to sinusoidal input pressure ranging in frequency from 0 to 300 hertz. The calculations were made for conditions identical to those of the experimental tests which will be described later.

Typical results of the numerical solution of the model are given in Figures 2 through 5. They are given in the form of Bode plots, i.e., curves of amplitude ratio (P_o/P_i) and phase shift (degrees) versus frequency. These curves were plotted by a curve plotter directly from the computer output.

The idea that all types of periodic waveforms can be resolved into summations of steady-state sine waves of different frequencies (Fourier series) was used to obtain an expression for the square wave pressure input signal at the inlet end of a flueric transmission line. An expression for the resulting pressure waveform at the outlet end of the flueric transmission line was obtained by using the mathematical model to determine amplitude and phase shift of each of the component sine waves, then summing up the resulting sine waves.

The Fourier series expansion for a square wave pressure signal with an amplitude of P_A, frequency of $f = \omega/2\pi$, and mean pressure of P_M is given in equation (14).

$$P_i(o,t) = P_M + \frac{2P_A}{\pi}\left[\sin(\omega t) + \frac{\sin(3\omega t)}{3} + \frac{\sin(5\omega t)}{5} + \frac{\sin(7\omega t)}{7} + \frac{\sin(9\omega t)}{9} + \frac{\sin(11\omega t)}{11} + \ldots\right] \qquad (14)$$

By using the mathematical model, the amplitude ratio and phase shift of each of the component sine waves at the outlet end of the flueric transmission line can be determined. The amplitude ratio and phase shift for the sine wave of frequency, ω, will be designated as $\left(\frac{P_o}{P_i}\right)_\omega$ and ϕ_ω, respectively. The resulting expression for the pressure waveform at the outlet end of a flueric transmission line of length, L, is given by equation (15).

$$P_o(L,t) = P_M + \frac{2P_A}{\pi}\left[\left(\frac{P_o}{P_i}\right)_\omega \sin(\omega t - \phi_\omega) + \left(\frac{P_o}{P_i}\right)_{3\omega}\frac{\sin(3\omega t - \phi_{3\omega})}{3} + \left(\frac{P_o}{P_i}\right)_{5\omega}\frac{\sin(5\omega t - \phi_{5\omega})}{5} + \left(\frac{P_o}{P_i}\right)_{7\omega}\frac{\sin(7\omega t - \phi_{7\omega})}{7} + \left(\frac{P_o}{P_i}\right)_{9\omega}\frac{\sin(9\omega t - \phi_{9\omega})}{9} + \left(\frac{P_o}{P_i}\right)_{11\omega}\frac{\sin(11\omega t - \phi_{11\omega})}{11} + \ldots\right]. \qquad (15)$$

A digital computer program was written in FORTRAN IV using equations (14), (15), and the mathematical model. The first six terms of the Fourier series expansions in equations (14) and (15) were used. The program was then used to obtain the dynamic response of a flueric transmission line 2 feet in length to square wave pressure input signals ranging in frequency from 0 to 70 hertz. The calculations were made for conditions identical to those of the experimental tests which will be described later. Typical results of the numerical solutions of the model are shown in Figures 6 and 7. They are given in the form of input pressure (P_i) and output pressure (P_o) versus time.

THE EXPERIMENTAL INVESTIGATION

In the experimental investigation, the following equipment was used: (1) a pneumatic test stand; (2) a fluidic signal generator; (3) different lengths of flueric transmission lines; and (4) instrumentation to visualize and record the necessary data. A schematic diagram of the experimental apparatus is shown in Figure 8.

Description of apparatus. A pneumatic test stand was used to furnish an accurately regulated air supply to the fluidic signal generator. The test stand consisted of three pressure regulators, three one-half percent accuracy pressure gages, and associated tubing. The test stand was supplied with compressed air from the laboratory air system.

A fluidic signal generator was used to generate pneumatic sine waves with frequencies up to approximately 300 hertz. The device consists of a rotating wobble plate driven by an air turbine which varies the distance between the wobble plate and an outlet orifice connected to a pressure chamber. When the wobble plate is rotating at a constant angular speed, the distance between the wobble plate and the outlet orifice changes sinusoidally. A corresponding pressure change occurs in the pressure chamber. Both the frequency and amplitude of the sinusoidal pressure signal are adjustable over wide ranges. A photo cell which is energized by a light beam passing through a slot in the periphery of the wobble plate supplies the signal to a frequency counter to determine the frequency of the generated pressure signal. The fluidic signal generator was a Sperry Utah Model 280-419. A Hewlett-Packard Model 712A dc power supply was used to power the light source on the fluidic signal generator. A Hewlett-Packard Model 5223L digital electronic counter was used to determine the frequency of the generated pressure wave. Piezoelectric transducers were mounted at the inlet and outlet of the fluidic transmission line. The transducers were connected in the transmission line by using small copper tees. Kistler Model 622B pressure transducers were used. The output of each transducer was amplified by a Kistler Model 566 change amplifier and then displayed on a dual beam oscilloscope. A Polaroid scope camera was used to record the oscilloscope traces. The flueric transmission lines were made of Tygon tubing, a thick-walled plastic tubing of

0.125-inch I.D. and 0.250-inch O.D. The outlet end of the fluidic transmission line was loaded by a sharpedged orifice 0.010 inches in diameter to simulate the loading condition of a fluid amplifier.

Test Procedure. The mean pressure level in the pressure chamber of the fluidic signal generator was adjusted to a pressure of 1 psig. The frequency of the generated pressure sine wave was controlled by adjusting the supply pressure of the air turbine drive on the fluidic signal generator and by observing the frequency on the digital frequency counter. The frequency could be adjusted to within 1 hertz of the desired value. The amplitude of the generated pressure sine wave was adjusted to the desired value by a micrometer adjustment that varied the distance between the outlet orifice of the pressure chamber and the rotating wobble plate. The amplitude of the generated pressure sine wave was adjusted between the limits of 0.5 and 0.20 psi for all the runs.

Tests were made using 2-, 4-, 6-, and 8-foot lengths of tubing. Each of these lengths was tested at frequencies ranging from 10 to 300 hertz, in 10-hertz increments. The tests were conducted in an ambient temperature of 78°F using air at a temperature of 78°F. For each run, the input frequency was adjusted to the desired value, and the input and output pressure waveforms were displayed on the oscilloscope and photographed.

Typical waveforms of the input and output pressures are illustrated in Figure 9. The upper trace is the input pressure; the lower trace the output pressure. The pressure amplitude ratio (P_o/P_i) and the phase shift were determined from the experimental data. Examples of the reduced experimental data are given in Figures 2 through 5.

The experimental apparatus and procedures used for the square wave inputs were similar to those used in the frequency response experiments. The primary exception was that a square wave pressure signal was introduced at the input of the flueric transmission line in these experiments in lieu of a sinusoidal pressure signal. The square wave pressure signal was generated by modifying the fluidic signal generator to generate square wave pressure signals. This was done by replacing the rotating wobble plate with a rotating flat plate which had a groove machined in it. The generator produces a square wave pressure signal by varying the distance between the rotating plate and an outlet nozzle between two distinct levels. A corresponding pressure change occurs in the pressure chamber of the signal generator. Both the frequency and amplitude of the square wave pressure signal are adjustable.

The experimental data was taken in the form of pressure versus time curves which were recorded with a Polaroid scope camera. The experimental data were taken for 2-, 4-, 6-, and 8-foot lengths of flueric transmission line. The data were taken at 10-hertz

intervals for frequencies ranging from 10 to 110 hertz. At the higher frequencies, above approximately 50 hertz, the square wave pressure signal was distorted, while at lower frequencies, the quantity of the generated signal was very good. The upper frequency limit of the square wave signal was determined by the fluidic signal generator. Typical waveforms of the input and output pressures are illustrated in Figure 10. The upper trace is the input pressure, and the lower trace is the output pressure.

CONCLUSIONS

To determine the validity of the mathematical model, predictions of the model were compared with experimental results. Comparisons of the results of the mathematical model and the experimental findings are given in the form of pressure amplitude ratio $\left(\frac{P_o}{P_i}\right)$ versus frequency and phase shift versus frequency curves. Examples of these curves are given in Figures 2 through 5.

An examination of these curves indicates that the correlation between the results predicted by the mathematical model and the experimental data is quite good. In some cases, the peaks of the calculated curves of pressure amplitude ratio versus frequency are shifted slightly from the peaks indicated by the experimental data. In other cases, the calculated curves of pressure amplitude ratio versus frequency dip to values slightly lower than the experimental data. The curves of phase shift versus frequency show that the calculated values predict the trends of the experimental data.

Graphical comparisons between the calculated and experimental curves of input and output pressure versus time for square wave pressure input signals are given in Figures 6 and 7. These comparisons were made in this form because of the lack of a general method such as Bode plots that are used to compare sinusoidal signals over wide frequency ranges.

These comparisons show that at the lower frequencies both the calculated and experimental curves of output pressure versus time exhibit the same trends with small peaks occurring at the ends of the square waves. Figure 7 shows that at higher frequencies the curves of output pressure versus time curves for both the experimental and calculated cases become triangular in shape. These comparisons indicate that the mathematical model predicts the general trends of the experimental curves. These curves also show the difficulty in producing high frequency square wave input pressures experimentally.

This study shows clearly some of the pronounced dynamic effects that occur in the higher frequency operation of flueric transmission lines. The two most pronounced effects are the resonant peaks which occur in the curves of pressure amplitude ratio (output pressure/

input pressure) versus frequency and the increasing phase shift (the angle between the output pressure signal and a corresponding point on the input pressure signal) as seen in the curves of phase shift versus frequency.

It is obvious, then, that the dynamic characteristics of flueric transmission lines must be included in the design of flueric systems along with the dynamic characteristics of the flueric elements. On the basis of the good correlation between the experimental results and the results predicted by the mathematical model, it is felt that the mathematical model constitutes a practical and useful tool for analyzing flueric transmission lines. It was demonstrated that the model could be used for both sinusoidal input pressure signals as well as for square wave input pressure signals. The same approach and basic relationships could be used to determine the response of flueric transmission lines to any periodic pressure input signal, i.e., resolve the periodic pressure input signal into a summation of steady-state sine waves of different frequencies (Fourier series); then use the mathematical model to determine the amplitude and phase shift of each of the component waves at the outlet end of the transmission line. The output pressure signal would then be determined by summing the resulting sine waves.

The immediate need in fluidics is for the development of systems synthesis techniques to allow sophisticated fluidic systems to be developed. The Advanced Ballistic Missile Defense Agency is continuing this work in a study entitled "Fluidics Systems Synthesis" being done by Stanford University. The overall objective of this work is the development of techniques for the analytical design of signal shaping networks, with particular emphasis on distributed parameter effects and their use in filter synthesis.

REFERENCES:

1. Atha, L. C., "An Investigation of the Dynamic Pressure Response of Flueric Transmission Lines," Ph.D. Dissertation, University of Missouri, Columbia, 1968

2. Skilling, H. H., Electric Transmission Lines, New York: McGraw-Hill Book Company, Inc., 1951

3. Fay, R. D., "Attenuation of Sound in Tubes," Journal of the Acoustic Society of America, 12 (1940), pp. 62-67

4. Ibervall, A. S., "Attenuation of Oscillatory Pressures in Instrument Lines," Journal of Research, National Bureau of Standards, Paper No. 2115, Vol. 45 (July 1951)

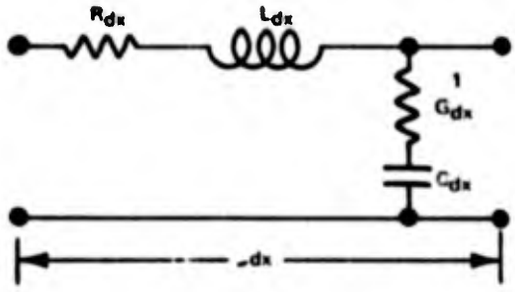


FIGURE 1. DIFFERENTIAL SEGMENT OF A TRANSMISSION LINE

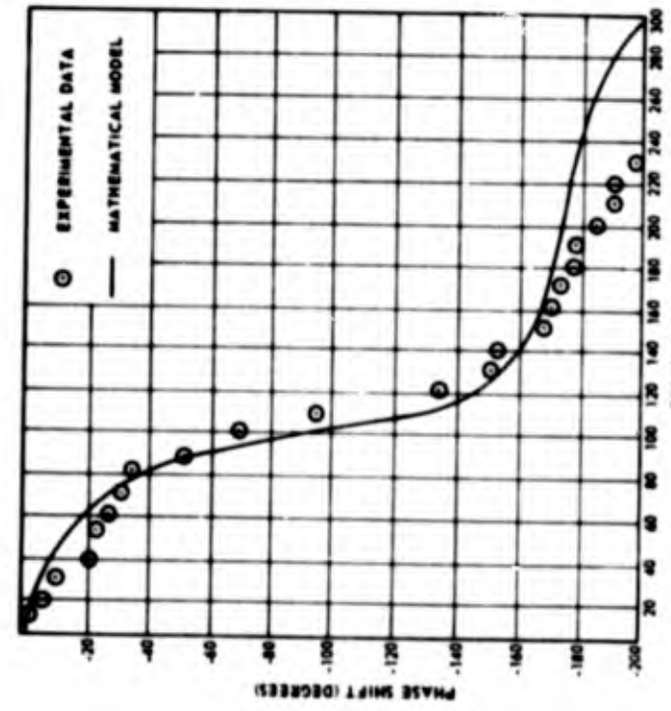


FIGURE 3. CURVE OF PHASE SHIFT VERSUS FREQUENCY FOR A 2.0-FOOT LENGTH OF FLUERIC TRANSMISSION LINE

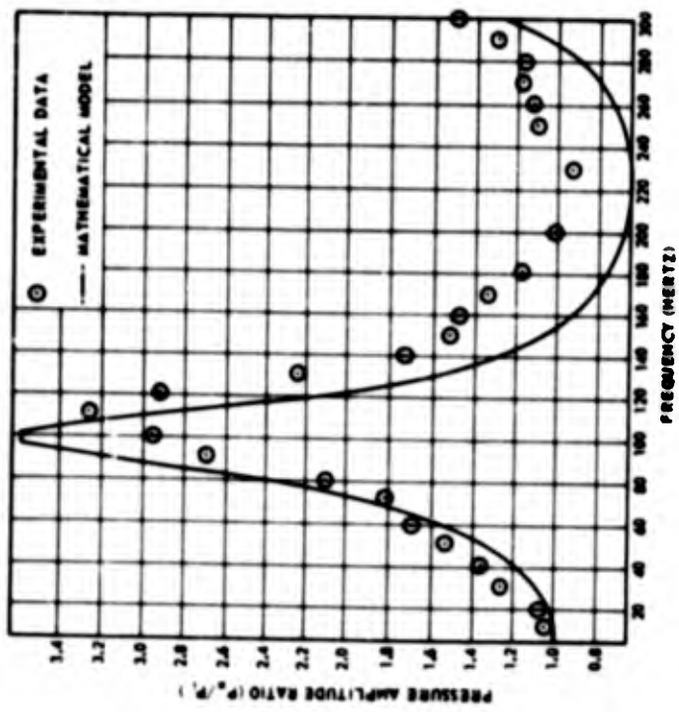


FIGURE 2. CURVE OF AMPLITUDE RATIO (P_o/P_i) VERSUS FREQUENCY FOR A 2.0-FOOT LENGTH OF FLUERIC TRANSMISSION LINE

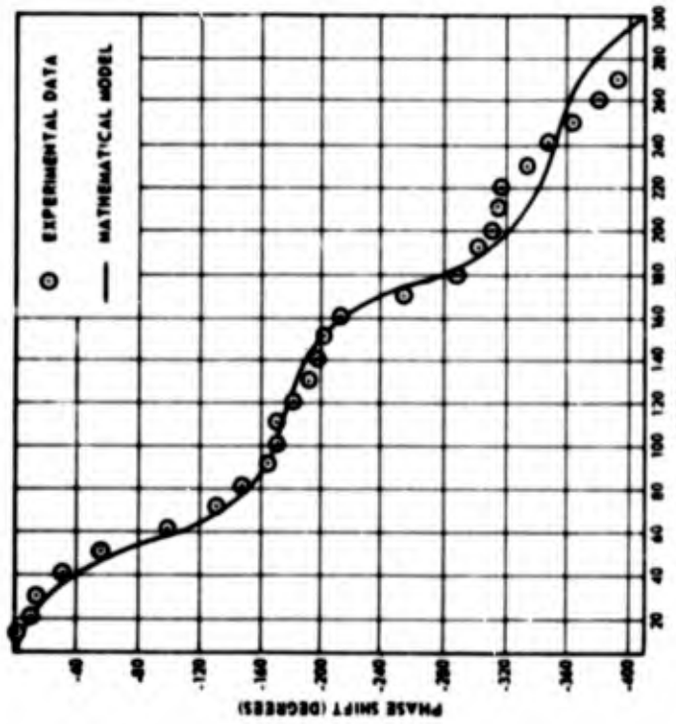


FIGURE 5. CURVE OF PHASE SHIFT VERSUS FREQUENCY FOR A 4.0-FOOT LENGTH OF FLUERIC TRANSMISSION LINE

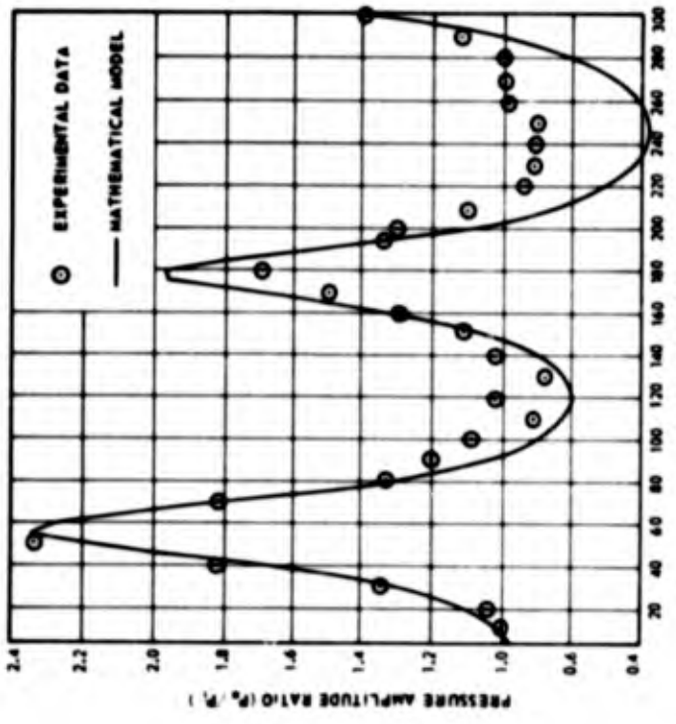


FIGURE 4. CURVE OF AMPLITUDE RATIO (P_o/P_i) VERSUS FREQUENCY FOR A 4.0-FOOT LENGTH OF FLUERIC TRANSMISSION LINE

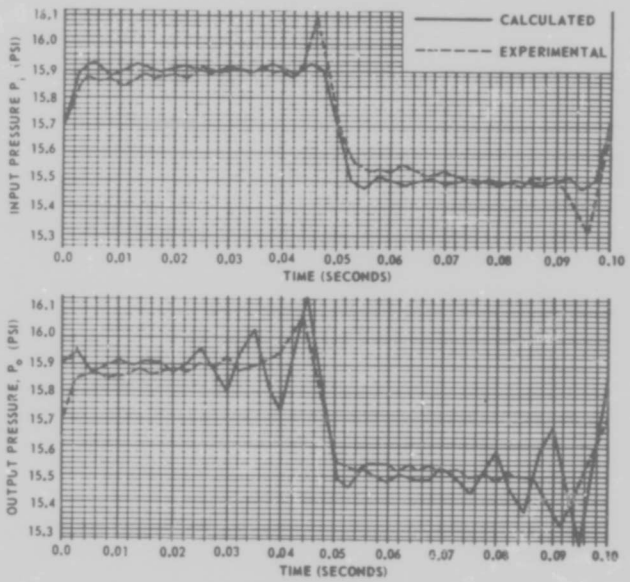


FIGURE 6. INPUT-OUTPUT PERSSURE VERSUS TIME CURVES FOR ℓ = 2.0 FEET, f = 10 HERTZ

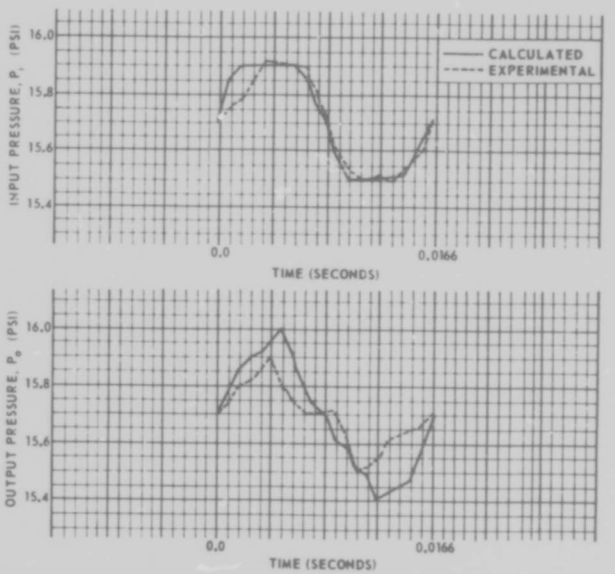


FIGURE 7. INPUT-OUTPUT PRESSURE VERSUS TIME CURVES FOR ℓ = 2.0 FEET, f = 60 HERTZ

ATHA

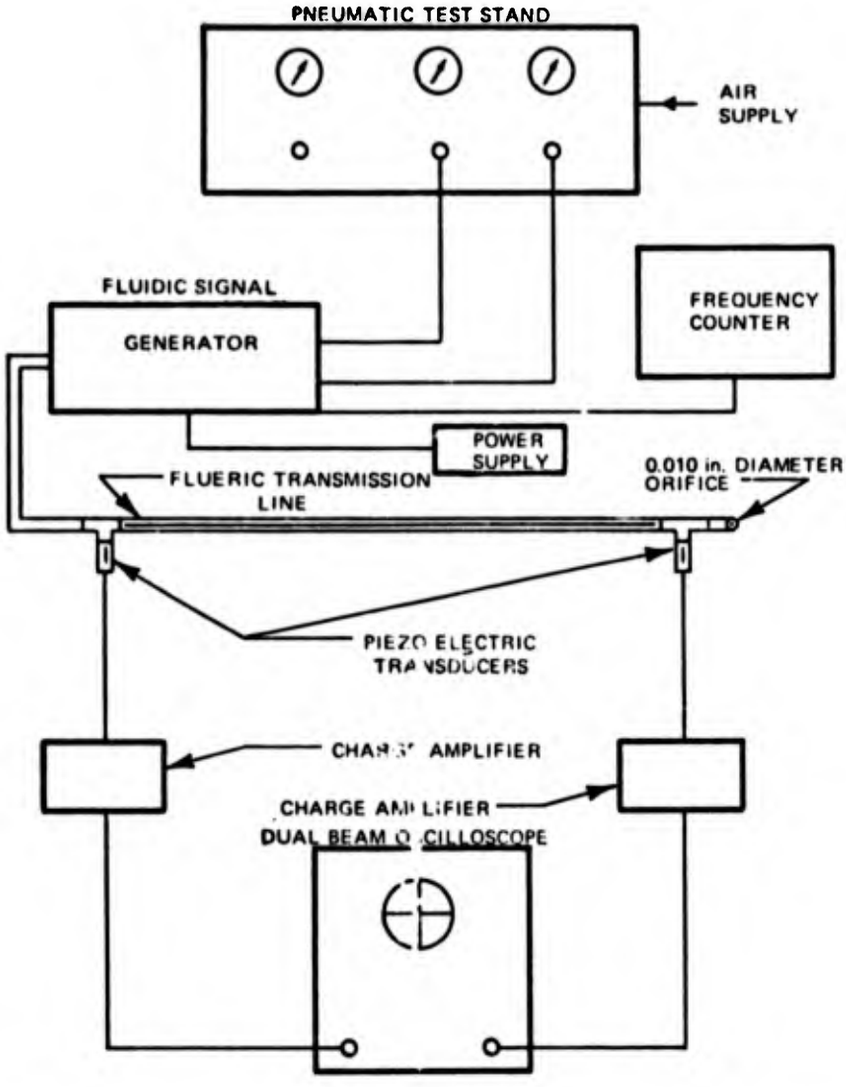


FIGURE 8. SCHEMATIC DIAGRAM OF THE EXPERIMENTAL APPARATUS

AD 713 506

PREDICTION OF HIGH VELOCITY SOLID PROPELLANT GUN PERFORMANCE BY GAS DYNAMIC COMPUTER PROGRAM

Paul G. Baer
Ballistic Research Laboratories
Aberdeen Proving Grounds, Maryland

INTRODUCTION: One of the basic objectives in the Interior Ballistic Laboratory of BRL is the development of an overall ballistic model capable of simulating the performance of entire weapon systems on the computer. Such computer simulations would be capable of exploring new weapon concepts and determining the pay off of the new concept prior to any hardware exploratory development. In addition, such a model would serve to guide the subsequent experimental programs needed to evaluate the feasibility of new weapon concepts.

A key link in the development of the overall ballistic model is the solid propellant gun gas dynamics computer program. This program is a one-dimensional Lagrangian code based on the Richtmyer Von Neuman "q" method; but unlike the basic method which considered the flow of gas only, the present program considers the flow of a fluid consisting of hot gas with burning propellant grains entrained in the gas.

This program can find use in the following potential areas of gun design:

1. Predicting the pressure and heat transfer distribution in a gun where the propellant chamber diameter is greater than bore diameter. This is a problem to the gun tube designers since conventional interior ballistic theory is based on bore diameter chambers. With the emphasis going to lighter gun tubes, errors in pressure and heat transfer distribution results cannot be compensated for by generous safety factors in gun tube thickness.

2. Predicting the acceleration loadings on saboted projectiles launched from high velocity guns. High velocity launched saboted kinetic energy penetrators have considerable military application in antitank warfare, but one of the problems in their use, is making the sabot light enough so as to attain the desired launch velocity

and yet strong enough so that the sabot does not break up under the high acceleration loadings during launch. Prediction of the acceleration loading on these projectiles will enable the sabot designer to design light weight sabots which resist the stresses encountered during a high velocity launch.

BACKGROUND: The initial study on the gas dynamics problem of gun interior ballistics was made by Lagrange in 1793. Lagrange was responsible for formulating the Lagrange problem of interior ballistics which referred to a special type of gun in which the projectile is prevented from moving until all of the solid propellant in the chamber is converted to a gas. The problem was to predict the pressure distribution in the gas after the projectile began to move and the motion of the projectile.

Throughout the 19th century and the early part of the 20th century, a series of mathematicans contributed to the solution of the Lagrange problem. This culminate in the classic work of Love and Pidduck (1). In this work, Love carried out a mathematical analysis of the wave motion developed in the propellant gas as a result of the motion of the projectile. Pidduck, using the equations developed by Love, computed a numerical wave solution for a large caliber gun. Pidduck and later Kent (2) found a limiting solution to the Lagrange problem that resulted in a theory that the ratio of pressures between the breech face and projectile base was held to be a constant during the firing cycle. This ratio is a function of the projectile weight to propellant weight ratio and the specific heat ratio of the propellant gas. This theory was applied by Kent to the case of a gun in which the propellant continued to burn while the projectile moved. This assumption, however, implied that the gas density in the propellant chamber at start of projectile motion was non-uniform. Since the experimental results of interest were maximum pressure and muzzle velocity; this assumption was sufficient for most gun interior ballistic theories.

With the advent of electronic digital and analog computers, it became possible to incorporate these interior ballistic theories into computer programs. The computer programs or models were then capable of generating gun interior ballistic trajectories which then could be compared with measured interior ballistic trajectories resulting from improved gun instrumentation. Reference (3) describes such a computer program and compares predicted interior ballistic trajectories with corresponding measured values.

While the computer program of reference (3) is able to simulate the performance of any type of solid propellant gun; the program has certain deficiencies. One of these deficiencies is that, because of application of the Pidduck-Kent constant pressure ratio, the pressure

on the projectile base during the projectile engraving is too low and thus predictions of the effect of engraving resistance on initial projectile motion in the tube will be in error. Also in this theory the propellant chamber is a bore diameter extension of the bore; thus the influence of chamber dimensions and shape on gun performance cannot be evaluated. The work of Seigel (4), while confined to guns using gas propellant, did indicate that increasing the chambrage (propellant chamber diameter-bore diameter ratio) would increase the muzzle velocity of the gun. Another deficiency was the inability of the program to predict the effect of change in gun bore cross-section on the heat transfer distribution. Experimental work at Cornell (5) had indicated that the heat transfer coefficient varied from breech face to muzzle; reaching high values in the convergent cone between the propellant chamber and the bore.

For the above reasons, it was decided to develop a one-dimensional gas dynamic gun computer model which would, given the proper data, be able to predict the interior ballistic phenomena discussed above.

THEORETICAL DEVELOPMENT: The theory of the gas dynamic method of gun interior ballistics is based on the Lagrangian form of the equations of time-dependent fluid flow. In general, finite difference methods for the calculation of time-dependent fluid flows have been based mostly on either the Eulerian or the Lagrangian form of the equations. Although these forms are essentially equivalent, the Lagrangian form gives more information (it tells where each parcel of fluid came from originally), and has the virtue that conservation of mass in the fluid parcels is automatic and exact, even in the finite difference approximations. This results in considerably greater accuracy in one-dimensional problems. In addition, moving boundaries (such as the base of a projectile) are easy to define in a Lagrangian system and are more difficult to define in an Eulerian system.

The original form of the finite difference equations describing one-dimensional time-dependent fluid flow were derived by Richtmyer and Von Neuman (6). Since in a solution of a gas dynamic problem, shocks are likely to occur, Richtmyer and Von Neuman devised an artifical dissipative mechanism of such form and strength, that the shock transition is a smooth one, extending over a small number of space intervals. This artifical dissipative mechanism (artifical viscosity), when incorporated in the finite difference equation, permits all dependent variables (pressure, velocity, etc.) to vary smoothly through the shock region and thus suppress any numerical instabilities.

The assumptions used in the model are as follows:

1. Non-ideal gas obeying the Nobel-Abel equation of state.

2. Frictional losses between projectile and gun bore to be either a constant or to vary with projectile travel.

3. Negligible frictional losses between propellant gas and bore surface.

4. Heat loss to be a function of gas velocity and gas temperature.

5. Unburned solid propellant moves as fast as local gas velocity.

6. Burning rate of solid propellant is a function of the local pressure and the local gas velocity.

The following Lagrangian hydrodynamic equations are used in the code:

Momentum:

$$\frac{\partial u}{\partial t} = -\frac{\partial (P + q)}{\partial M} A(x) \tag{1}$$

Energy:

$$\frac{\partial (Q - E)}{\partial t} = (P + q) \frac{\partial v}{\partial t} + \frac{\partial H_L}{\partial t} \tag{2}$$

where:

$$q = \begin{cases} \frac{2\rho_o}{v}^2 \left(\frac{\partial u}{\partial x}\right)^2 , & \text{if } \frac{\partial u}{\partial x} < 0 \\ 0 , & \text{if } \frac{\partial u}{\partial x} \geq 0 \end{cases} \tag{3}$$

These equations together with the following are used to describe the gas dynamic code:

Nobel-Abel Equation of State:

$$P(V - b) = nRT \tag{4}$$

Propellant Burning Rate:

$$\frac{du_p}{dt} = r = \beta P^{\alpha} + k_v v \tag{5}$$

Propellant Form Function:

$$z = 1 - \left\{ \frac{(1 - 2u_p)\left[(D - 2u_p)^2 - N(d + 2u_p)^2\right]}{L(D^2 - Nd^2)} \right\} \tag{6}$$

To develop the finite difference equations used in the computer model, the solid propellant-gas region between the breech end of propellant chamber and the projectile base is divided into J segments of equal length. This is illustrated in Figure 1. Since the cross-sectional area A(x) of the chamber may vary, the amount of solid propellant in each segment is proportional to the initial volume of the segment. Because of assumption 5, the total mass (solid propellant, propellant gas, and igniter gas) in each segment will remain fixed throughout the remainder of the problem.

The momentum equation in finite difference form is:

$$u^{n+1}_{j-\frac{1}{2}} = \left\{ \frac{\left[(P^n_{j-1} + q^n_{j-1}) - (P^n_j + q^n_j)\right] A(x^n_{j-\frac{1}{2}})g}{.5(m_j + m_{j-1})} \right\} \Delta t + u^n_{j-\frac{1}{2}} \tag{7}$$

and the gas displacement is:

$$x^{n+1}_{j-\frac{1}{2}} = u^n_{j-\frac{1}{2}} \Delta t + x^n_{j-\frac{1}{2}} \tag{8}$$

The energy equation requires a special procedure to place it in finite difference form, because of the presence of the propellant chemical energy term Q and the desire to explicitly solve for propellant gas temperature.

In finite difference form the energy equation is:

$$\Delta (Q - E) = (P + q)\, \Delta v + \Delta H_L \tag{9}$$

Using the procedure in reference 3, we define:

$$Q = J\, m\, z \int_o^{T_o} C_v\, dT + J m_I \int_o^{T_{o_I}} C_{v_I}\, dT \tag{10}$$

$$E = J\, m\, z \int_o^{T} C_v\, dT + J m_I \int_o^{T} C_{v_I}\, dT \tag{11}$$

Combining:

$$Q - E = J\, m\, z \int_{T}^{T_o} C_v\, dt + Jm_I \int_{T}^{T_{o_I}} C_{v_I}\, dT \tag{12}$$

$$Q - E = J\, m\, z\, \bar{C}_v\, (T_o - T) + Jm_I\, \bar{C}_{v_I}\, (T_{o_I} - T) \tag{13}$$

Since:

$$J\, \bar{C}_v = \frac{F}{(\gamma-1)T_o}\,, \text{ and } J\bar{C}_{v_I} = \frac{F_I}{(\gamma_I-1)T_{o_I}} \tag{14}$$

then:

$$Q - E = \frac{m_i F z}{(\gamma-1)T_o}\, (T_o - T) + \frac{m_I F_I}{(\gamma_I-1)T_{o_I}}\, (T_{o_I} - T) \tag{15}$$

$$\Delta(Q - E) = \frac{m_i F}{(\gamma-1)T_o}\left\{z_j^{n+1}\, (T_o - T_j^{n+1}) - z_j^n\, (T_o - T_j^n) + \frac{m_I F_I}{(\gamma_I-1)T_{o_I}}\, (T_j^n - T_j^{n+1})\right\} \tag{16}$$

$$(P + q)\, \Delta v = \left[\frac{(P_j^{n+1} + P_j^n)}{2} + q_j^n\right] (v_j^{n+1} - v_j^n) + SW_j^n \tag{17}$$

$$\Delta H_L = (H_{L_j}^{n+1} - H_{L_j}^n) + SH_{L_j}^n \tag{18}$$

Combining 16, 17, and 18 and substituting in the equation of state, we arrive at the following equation for the propellant gas temperature:

$$T_j^{n+1} = \frac{C - W - H}{I} \tag{19}$$

where:

$$C = \frac{m_i F}{(\gamma-1)T_o}\, (z_j^{n+1}\, T_o - z_j^n\, (T_o - T_j^n)) + \frac{m_I F_I T_j^n}{(\gamma_I-1)T_{o_I}} \tag{20}$$

$$W = (.5\, P_j^n + q_j^{n+1})(v_j^{n+1} - v_j^n) + SW_j^n \tag{21}$$

$$H = (H_{L_j}^{n+1} - H_{L_j}^{n}) + SH_{L_j}^{n} \tag{22}$$

$$I = \frac{m_j F z_j^{n+1}}{(\gamma-1)T_o} + \frac{m_I F_I}{(\gamma_I-1)T_{o_I}} + \frac{.5(v_j^{n+1} - v_j^{n})}{v_j^{n}} \left(\frac{F m_i z_j^{n+1}}{T_o} + \frac{F_I m_I}{T_{o_I}} \right)$$

$$SW_j^{n} = \left(\frac{p_j^{n} + p_j^{n-1}}{2} + q_j^{n-1} \right)(v_j^{n} - v_j^{n-1}) + SW_j^{n-1} \tag{24}$$

$$SH_{L_j}^{n} = (H_{L_j}^{n} - H_{L_j}^{n-1}) + SH_{L_j}^{n-1} \tag{25}$$

The other equations are converted into the finite difference form in a similar manner. The above equations are discussed in greater detail in reference 8.

A computer program using the equations of the theory presented above, was written for the two ARDC digital computers BRLESC1 and BRLESC2. A detailed description of this program will be given in a subsequent BRL report (9).

INSTRUMENTED GUN SYSTEM: In order to verify the predictions of the gun gas dynamic computer program described above, firings were undertaken in a high velocity 90-37mm smooth bore gun. The dimensions of the gun, the projectile weight, and characteristics of the propellant are given in Table I. Instrumentation for this gun is illustrated in Figure 2. Pressure instrumentation on the gun consists of pairs of pressure transducers at each station on the gun; each pair consisting of a Minihat strain type pressure transducer (10) and a Kistler Model 607A pressure transducer. Velocity instrumentation for the gun consists of a 10 GHz microwave interferometer to measure in-bore projectile motion (11) and three velocity screens in front of the gun to obtain muzzle velocity. Data from this instrumentation is recorded on analog magnetic tape. This data is subsequently processed on analog to digital conversion equipment and digital computers, using special data processing routines, to produce printed digital data.

RESULTS AND DISCUSSION: A total of 68 rounds have been fired in the 90-37mm gun to date. In the rounds fired, projectile weight, chamber volume, and propellant weight were varied. All other parameters were held constant. Muzzle velocities for the series varied form 4500 to 7600 f/s and maximum propellant chamber pressures ranged from 10,000 psi to 65,000 psi.

At the present time, only a few of the firings in this program have been completely reduced, so in this report, only the results which have been obtained to date will be compared with predictions from the gas dynamic computer program.

To obtain predictions from the gas dynamic computer program for the rounds which will be discussed here, it is necessary to adjust certain input parameters in the program, values for which are not available, so that predicted maximum breech pressure and projectile muzzle velocity will agree with the corresponding results from one of the experimental rounds. These parameters would then be held constant for predictions made on other rounds fired in the gun. For this study the round chosen was round 11. The constants which were adjusted to provide the agreement between theory and experiment were engraving resistance profile, heat loss coefficient k_1, and the propellant erosion constant k_v. The magnitude of the resistance profile, and the other two constants were systematically varied until the error between theory and experiment for maximum breech pressure and muzzle velocity was 1% or less. These three parameters were then held constant for the other rounds simulated. The values of the parameters obtained in this manner are given in Table I. No attempt was made in this matching process to obtain agreement between the detailed theoretical and experimental interior ballistic trajectories other than maximum breech pressure and muzzle velocity. Further details on the matching process are given in reference 11.

The detailed interior ballistic trajectory of the round used in the matching study is illustrated in Figure 3. In the figure, breech face pressure-time, projectile velocity-time, acceleration pressure-time, and projectile displacement-time are plotted both for the experimental results and the predicted results. The experimental breech pressure curve is from the Kistler pressure transducer. The velocity and projectile travel curves are from the interferometer data, the signal of which faded out after a projectile travel of 80 inches. For projectile travel beyond the 80 inch position, the times at which the projectile base passed [illegible]er the pressure transducers down bore were used.

To correlate the theoretical and experimental interior ballistic trajectories in time, the projectile travel-time curve was used as a base. This is because the agreement between experimental and theoretical projectile travel-time curves is very close except at start of motion, which is experimentally poorly defined. It will be noted that agreement between theory and experiment for velocity-time is good, the maximum error being not more than 25% between the two curves. The agreement between measured and theoretical acceleration pressure-time curves is also good, the theoretical curve falling within the spread of the experimental data. The time correlation between the measured and theoretical breech pressure-time curves is not as good. At maximum pressure, the experimental curve lags the theoretical curve by about 0.5 ms.

The reason for the lag between the two curves is not understood. A probable reason may be that the gas dynamic computer program does not correctly simulate gas flow and propellant burning rate in the nozzle connecting the propellant chamber and the gun bore.

Of interest to gun tube designers is the maximum pressure each portion of the gun tube will experience during the interior ballistic cycle. Experimentally this is determined in the 90-37mm gun by determining the maximum pressure recorded from each of the pressure transducers in the chamber and along the bore. These experimental values are plotted against pressure transducer position referred to breech face in Figures 4,and 5. These plots represent rounds 9, 11, and 12 in which chamber volume projectile weight, and propellant weight were varied. It will be noted that of these three rounds, one used a 44 cubic inch chamber, and two used a 64 cubic inch chamber. Propellant to projectile weight ratios varied from 1.4 to 3.7 and muzzle velocities ranged from 5000 to 5900 f/s. Also illustrated in these plots are the predicted maximum pressure profile from the gas dynamic code. In general, the theoretical maximum pressures predicted for positions in the chamber or along the bore either fall within the observed experimental variation in maximum pressures or fall outside of the experimental variation by an error of not more than 25%.

In table II are given projectile weight, propellant weight, chamber volume, propellant to projectile weight ratio, muzzle velocity, and maximum acceleration. Both predicted and measured muzzle velocity and maximum acceleration are given. The percentage error between experimental and predicted muzzle velocities is not greater than 5%, and the error between experimental and predicted maximum accelerations is not greater than 25%. It will be noted that predicted accelerations are higher than those experimentally observed.

In figure 6 theoretical profiles of pressure, gas velocity, heat flow rate to bore surface, and propellant fraction burnt as a function of position in the gun are given for a time at which the projectile has moved to a distance of 33 inches from start of motion. The data is taken from the simulation for round 11. Comparing these profiles to the gun geometry illustrated at the bottom of the figure, it can be seen that the 30° convergent cone between the propellant chamber and the bore creates sharp changes in the slope of the pressure and velocity profiles. It will also be noted that the heat transfer rate has a very sharp peak at a position just down stream of the convergent cone. This is the area in the gun which is subjected to the greatest amount of metal erosion and according to reference 5 experiences the greatest amount of heat addition. The fraction of propellant burnt has a high value at the breech end and a lower value at the projectile base end. The reason for this is that propellant in the breech end is subjected to higher pressures for a longer time.

CONCLUSIONS:

1. The gas dynamic computer program is able to correctly predict the interior ballistic trajectory (displacement, velocity, and acceleration versus time) of the projectile while in the gun, provided that data for resistance profile, erosion coefficient, and heat transfer coefficient is obtained by the matching of an experimental round. The velocity agreement is with an error of 10% and the error between measured and predicted maximum acceleration is less than 25%.

2. Measured breech pressure-time curves show a time displacement from the predicted breech pressure-time curves. This displacement is believed to be due to the possibility that the one dimensional model does not correctly simulate the two dimensional gas dynamic effects in the nozzle between the propellant chamber and the gun bore.

3. The gas dynamic computer program is able to predict the maximum pressures to which the gun tube is subjected along its length within an error of about 25% of maximum pressure at a position.

4. Heat transfer computations using the gas dynamic computer program indicate that the point of maximum heat transfer rate is very close to the downstream exit of the nozzle between the propellant chamber and the bore.

FUTURE PLANS: Plans are being made to reprogram the gun gas dynamic program from the FORAST language to the standard FORTRAN language so that the program can be used by other agencies. As part of this reprogramming, the theory will be revised so that the assumption that unburned propellant moves as fast as the gas will be eliminated; motion of the unburned propellant will be governed by gas drag forces. Additional changes will be the addition of provisions for the burning of more than one propellant in the gun and a provision for simulating ignition delays of portions of the propellant bed.

The detailed experimental data from the 90-37mm gun and other heavily instrumented guns in the Interior Ballistic Laboratory will be used to verify predictions made by the new program and also be used to modify the program in the interest of obtaining better predictions of new and unique guns.

REFERENCES

1. Love, A. E. H., Pidduck, F. B., "Lagrange Ballistic Problem" Phil. Trans. Royal Soc. (London) 222 (1922).
2. Kent, R. H., "The Motion of Powder Gas," BRL Report 36(Jan. 1936)
3. Baer, P. G., Frankle, J. M., "The Simulation of Interior Ballistic Performance of Guns by Digital Computer Program," BRL Report 1183, (Dec. 1962).
4. Seigel, A. E., "The Effect of Optimum Chambrage on the Muzzle Velocity of Guns with a Qualitative Description of the

Fundamental Phenomena During Gun Firing," NAVORD Report 2691 (Oct. 1952).
5. Vassallo, F. A. et al "Caseless Ammunition Heat Transfer" Cornell Aeronautical Laboratory Report G1-2433-Z-1 (Oct. 1967).
6. Richtmyer, R. D., "Difference Methods for Initial Value Problems, Chapter X, Pg. 189-229, Interscience Publishers Inc., N. Y. (1957)
7. Adams, L. H., editor, "Hypervelocity Guns and the Control of Gun Erosion," Summary Tech. Rpt. of Div. 1, NDRC, Wash., D. C. pg. 100-101.
8. Baer, P. G. "A Mass Point Computer Program for the Gas Dynamic Problems of High Velocity Gun Interior Ballistics," Proc. of 2nd ICRPG/AIAA Solid Propellant Conf., Anaheim, Calif. June 1967.
9. Baer, P. G. "A Mass Point computer Program for the Gas Dynamic Problems of Solid Propellant Gun Interior Ballistics," To be published BRL Report.
10. Gay, H. P. "The Evolution of Gages for Measuring Pressures in Guns and Rockets at the Ballistic Research Laboratories," BRL Memo Report 1402, May 1962.
11. Vest, D. C., et al, "Ballistic Studies with a Microwave Interferometer, Part II," BRL Report 1006, Feb. 1957.
12. Grollman, B. B. and Baer, P. G. "Theoretical Studies of the Use of Multipropellants in High Velocity Guns," BRL Report 1411, (Aug. 1968).

LIST OF SYMBOLS

a_o = Viscous resistance constant
$A(x_j)$ = Cross sectional area of j gas segment interface at position x_j.
b = Propellant gas covolume
C = Propellant weight
C_I = Igniter Weight
c_v = Specific heat at constant volume of propellant gas
d_v = Diameter of perforation of cylindrical propellant grain
D = Outside diameter of cylindrical propellant grain
E = Internal energy of propellant gas
F = Force of solid propellant
F_I = Force of igniter
g = Gravitational constant 386.09
H_L = Heat loss to bore surface from j gas segment
J = Mechanical equivalent of heat
k_v = Propellant velocity erosion constant
L_v = Length of propellant grain
m_j = Weight of propellant gas in j segment
m_I = Weight of igniter gas in j segment
m_j = Total weight of j segment
M_j = Mass propellant gas
m = Moles of gas

N = Number of perforations in cylindrical propellant grain
P = Gas pressure
P_j = Pressure in j gas segment
Q^j = Chemical energy of propellant
q_j = Viscous pressure in j gas segment
r^j = Linear burning rate of solid propellant
R = Gas constant 33372
t = Time
T_i = Initial temperature of gas ahead of projectile
T^i = Gas temperature
T_j = Temperature of gas in j segment
T_o^j = Isochloric flame temperature of propellant
u^o = Gas velocity
u_j = Velocity of j gas segment
u^j = Regression distance of solid propellant
v^p = specific volume of propellant gas
v_j = Volume occupied by propellant gas in j segment
V^j = Propellant gas volume
W_j = Accumulated work done on or by j segment
x_j^j = Displacement of j gas segment
z_j^j = Weight fraction of solid propellant burnt in j segment
α^j = Propellant burning rate exponent
β = Propellant burning rate coefficient
γ = Ratio of specific heats of propellant gas
Δt = Integration time step size
ΔV = Volume of segment

TABLE I

Characteristics of 90-37mm Gun, Projectiles, and Propellant

Gun:

Bore Diameter:	1.456 in.
Maximum Projectile Travel:	239.4 in.
Rated Maximum Pressure of Gun:	100,000 psi
Overall Length of Gun:	22 ft. 8 in.
Chamber Volumes with 3 inch I.D. insert:	44, 64, and 84 cu. in.
Cone Angle between Propellant Chamber and Bore:	30°
Projectile Weights:	1/4 lb., 1/2 lb., 3/4 lb.

Propellant:

Type:	M2
Force:	4,471,700 in-lb/lb
Specific Heat Ratio:	1.222
Flame Temperature:	3372.°K
Covolume:	29.67 cu. in./lb
Web:	.036 in.
Number of Perforations:	7
Igniter:	Black Powder
Igniter Weight:	019 lb

Matching Parameters:

Erosion Constant k_v: .00002
Heat Transfer Coefficient k_1: .2
Resistance Profile:

Projectile Travel, in.	Resistance Pressure, psi
0	800
.4	4415
.8	6989
1.4	2483
1.5 to end of bore	700

TABLE II

Firing Data for 90-37mm Rounds

Round No.	9	11	12
Projectile Weight - lb.	.5719	.2522	.2555
Propellant Weight - lb.	.8808	.8457	.9409
Propellant - Projectile Weight Ratio	1.540	3.353	3.683
Chamber Volume in ^{3}	64.40	44.13	64.19
Muzzle Velocity Experimental f/s	5097	5891	5707
Predicted f/s	5086	5952	5975
Maximum Acceleration Experimental kilo g's	38.8	64.6	53.7
Predicted kilo g's	47.7	76.17	66.57

Fig. 1. Gas Dynamic Model for Solid Propellant Gun

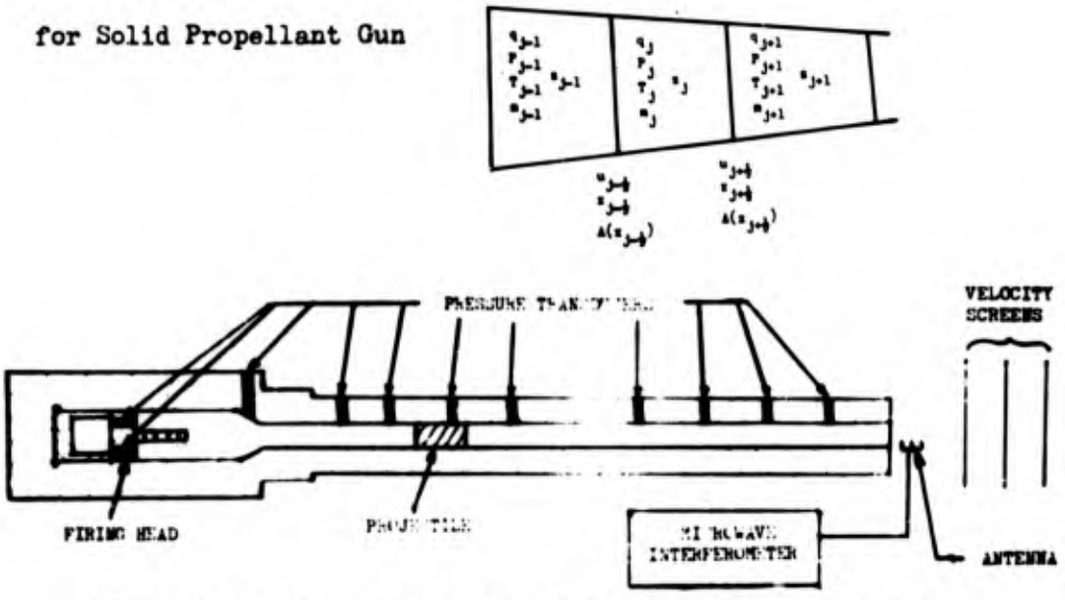


Fig. 2. Instrumentation for 90-37mm High Velocity Gun

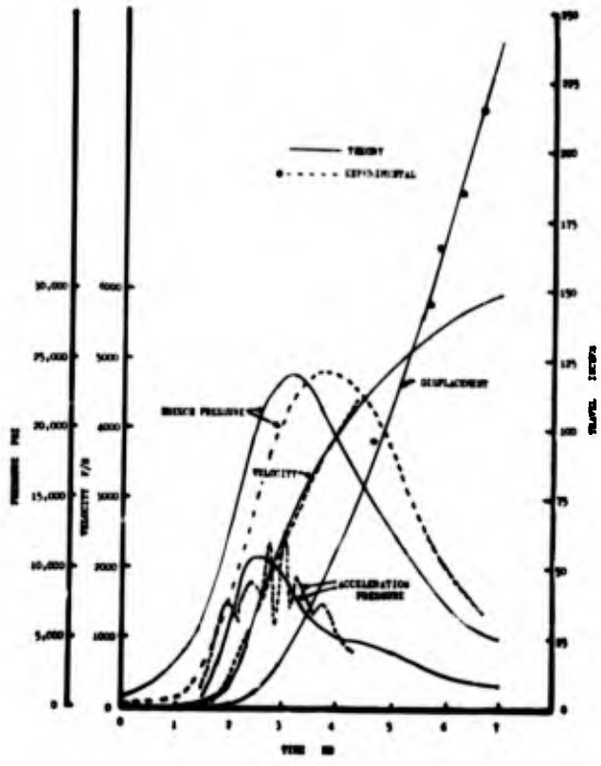


Fig. 3. Theoretical and Experimental Interior Ballistic Trajectory for Round 11, 90-37mm Gun

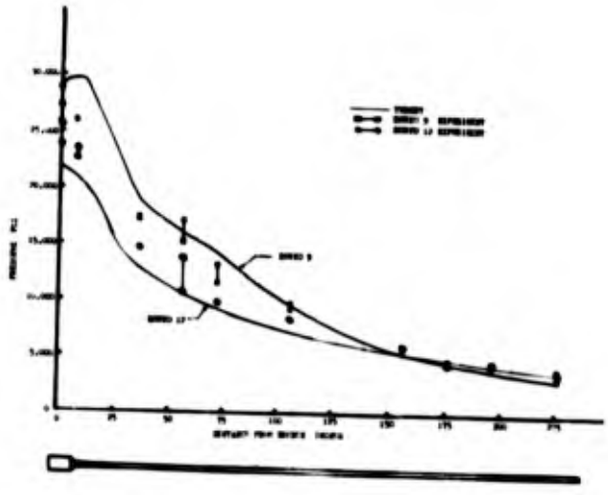


Fig. 4. Maximum Pressure Along Gun Bore, Rounds 9 and 12

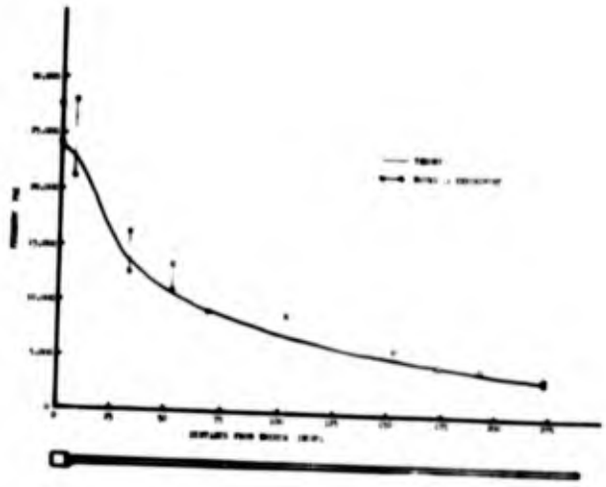


Fig. 5. Maximum Pressure Along Gun Bore, Round 11

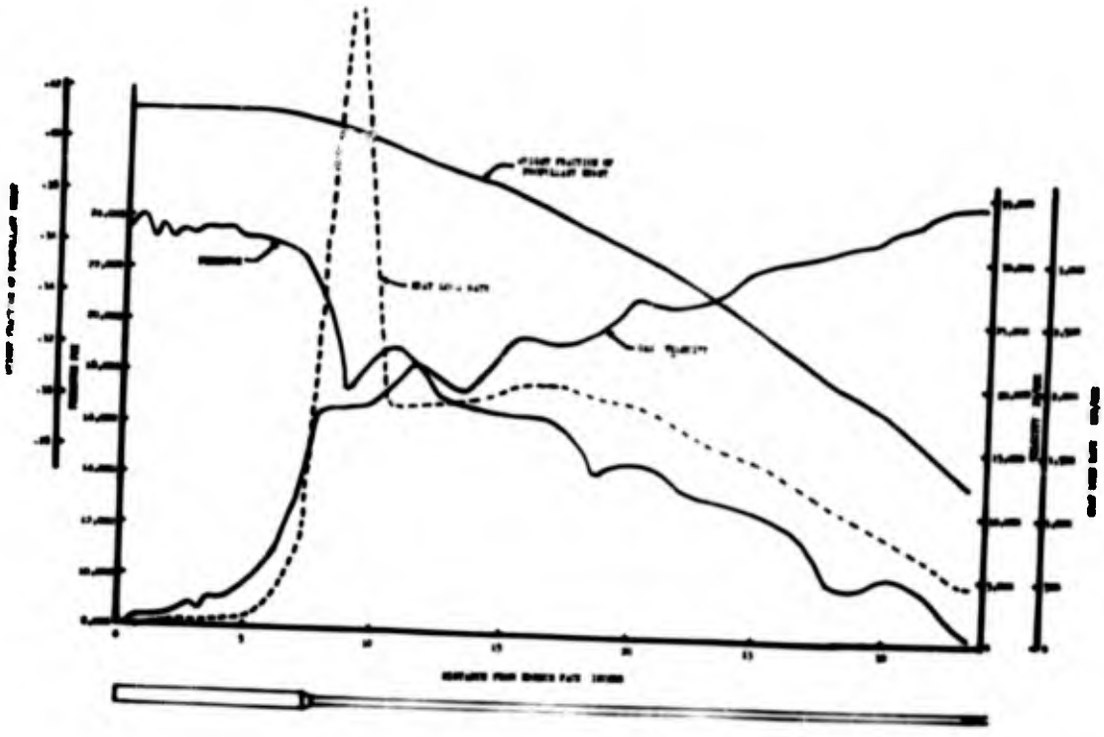


Fig. 6. Gas Dynamic Profiles Along 90-37mm Gun Bore. Round 11 Theory after Projectile Travel of 33 inches.

AD 713 507

SOCIAL STATUS VARIABLES IN THE MILITARY AND THEIR EFFECT ON EXPRESSING AGGRESSION

JOHN W. BAKER, II, MAJOR, MSC
UNITED STATES MILITARY ACADEMY
WEST POINT, NEW YORK

Baker and Schaie (1) described and demonstrated in some detail the effectiveness of a multidimensional approach to the study of frustration-aggression paradigms. This same research attempted with some success to study in combination the social, psychological and physiological variables involved in any response to "frustrating" circumstances. The general phenomenon of "arousal" indices were used as the dependent variables and suggested differences between self reported arousal and that recorded physiologically on a multi-channel polygraph. The results of that study confirmed many previous studies in which only one aspect of the paradigm was analyzed in any particular design. For example, the results of the Baker and Schaie (1) study suggested as did Thibault and Coules (9) that regardless of who carries out the aggressive activity, as long as the subject perceives injury to the instigator, he will feel some relief or tension reduction. Similarly, Baker and Schaie (1) noted as did Hokanson and Shetler (7) that there were indications of physiological tension reduction when the subject sees the instigator injured either physically or verbally. In the Baker and Schaie (1) study, the physiological tension reduction was noted regardless of whether the injury occurred directly or vicariously. Baker and Schaie (1) also attempted to assess the psychological arousal level of subjects placed in the frustration-aggression paradigm as described and commented upon by Berkowitz (2, 3, 4). The Baker and Schaie (1) results were such that little objective evidence of psychological arousal could be noted under the multidimensional design. In an effort to elicit maximal responding in the complex multidimensional design, Baker and Schaie (1) maintained the social stimulus value of those individuals in the design so that responding could and would occur. This optimal situation was enhanced by the use of highly verbal, relatively pathology-free college students as subjects. The instigator established himself as one of equal or lesser status relative to the students thus producing optimal environmental cues for the expression of aggression. This equal status criterion has been found to elicit optimal

responding by lessening the inhibitory cues introduced by high-status or authoritative individuals (Berkowitz, 2,3,4). Likewise, the social status of the vicar (ally of the subject) was that of a "friend", being equal in status and he was, at least, subjectively "likeable". This facilitated and elicited uninhibited responding as Berkowitz (4) had noted previously. The present study intends to use some of the same variables as the Baker and Schaie (1) study but at the same time will alter the social stimulus value of those people in the experiment. It is clear to many, including this author, that conditions for optimal responding seldom exist in the "real world". For that reason, social stimulus values will be altered to more nearly approximate the social environment in real life situations. One dimension of the "real world" is that provided by the social system imposed by the military. Here one sees a social structure where the instigator of aggressive feelings is often one of high status and authoritative position. Because the system prohibits overt retaliation to such individuals, one suspects the frustrated subject would take other courses of action. As the command subordinate social structure is one of the cornerstones of military life, it seemed valuable to note the types of aggression which were expressed and the concurrent physiological and psychological changes noted in the subject. Should the military system be able to discover some way for subordinates to acceptably and meaningfully express themselves toward command when they have aggressive feelings, we may see a lessening of inappropriate behavior which typically leads to incarceration, hospitalization or discharge. Beyond the scope of the military, it is of extreme practical social value to be able to define the types of tension reduction for which one searches in the face of the changing social stimulus value of those around him.

The hypothesis that aggressive experiences reduce frustration as expressed by physiological (systolic blood pressure, heart rate, respiration and galvanic skin response) and psychological (anxiety) arousal was investigated by assessing differences in arousal achieved by the subject counteraggressing alone or through aggressive responses expressed with another person (vicar). Counteraggression occurred through overt (apparent electric shock) and covert (abated) means and in the presence of a high status (Captain) or low status (E-3) instigator.

METHOD

Sixty-four enlisted men (E-3 and below) were assigned to the eight different treatment conditions.

The independent variables were a) counteraggressing alone versus counteraggression by another person in the presence of the subject, b) overt (physical) versus covert (abated) means of counteraggression, and c) high status (Captain) versus low status (E-3). It should be noted that in spite of subjective observation of varying degrees of intensity in counteraggression, no objective assessment of the quality of counteraggression was reported.

The dependent variables were a) residual psychological arousal as measured by Scheier and Cattell's IPAT 8 Parallel Form Anxiety Battery (8) and b) residual physiological arousal as measured by systolic blood pressure, pulse rate, respiration rate, and galvanic skin response recorded on a four-channel polygraph. Systolic blood pressure was monitored on the multichannel polygraph three times during the last two minutes of each phase and through the anxiety measure and then averaged for a single measure per phase. Pulse rate, respiration and GSR were continuously monitored on the polygraph. Reported measures were the average readings per minute on both dimensions beginning two minutes prior to the end of a phase and through the anxiety assessment technique.

The experiment was introduced to the subject as one involving physiological and psychological responses to working on routine intellectual tasks. Half of the subjects were exposed to a vicar who was presented as an associate who would be available to answer questions and aid the subject if the subject requested assistance. The other half of the subjects proceeded through the experiment without such an ally. Each group was further subdivided into those aggressing overtly and covertly. These subgroups were exposed to frustration-manipulation by either a high status or low status instigator.

Phase I: After a several minute adaptation period for all subjects, they were administered the Picture Completion subtest of the Wechsler Adult Intelligence Scale (WAIS - Hokanson & Burgess (6); Hokanson & Shetler (7)), following which both physiological and psychological measures were taken. Physiological measures were taken during and after the Picture Completion task. The anxiety scale was done immediately after the Picture Completion task. The elapsed time from the end of the Picture Completion task to the end of the anxiety measure and beginning of Phase II averaged 3 minutes and 5 seconds.

Phase II: The frustration-manipulation phase was then established. Here the 64 subjects were asked to count backwards from 99 to 1 as quickly as possible by two's. To make conditions highly frustrating, they were repeatedly interrupted and harassed concerning their slow performance, asked to begin again four times, and finally stopped with the statement that their data could not be used. This technique has been used by Hokanson (5) to produce subjective feelings of intense anger.

Phase III: The situation to permit both counteraggressing alone or with another was set up as follows: The subjects were told that the next task involved an interpersonal guessing game in that the subject was to think of a number between 1 and 10 and write it down, following which the experimenter was to guess the number. This would be done for 10 trials. If the subject were to counteraggress "alone", he would respond to the instigator (either high or low status) by overtly communicating his (instigator's) accuracy through the administration of apparent electric shock for an incorrect guess or the nodding of approval for a correct guess. Under covert conditions there was no feedback of results to the instigator regarding accuracy. If the subjects were to counteraggress "with another", his responses would be communicated vicariously through the associate (vicar).

RESULTS

All 64 subjects experienced the same experimental treatment up to and including the termination of measurement after Phase II. Any differences up to this point are accounted for by the exposure of the subject to either a high or low status instigator, and whether or not the subject was "alone" or "with another". A complete 2x2x2x3 analysis of variance of the arousal indices is summarized in Table 1.

Arousal indices such as systolic blood pressure ($p < .01$), pulse rate ($p < .05$) and GSR ($p < .01$) were shown to be significantly sensitive to the different phases of the experiment. Arithmetic means of all three physiological measures revealed an increased arousal state after frustration manipulation and a subsequent lowering of arousal after Phase III but not to baseline levels. It should be noted that increased arousal as measured by GSR is, in actuality, a lowering of skin resistance. The anxiety measure shows no significant variation throughout the experimental procedure. Table 1 further shows that blood pressure ($p < .05$) and GSR ($p < .01$) were significantly sensitive to the interaction between the social stimulus effects of aggressing "alone" or "with another", and whether the instigator was of high or low status. These same physiological measures reveal significant ($p < .01$) readings when the mode of aggression (overt-covert) interacts with high or low status instigators. GSR readings further reveal the significant interaction ($p < .01$) of the mode of aggression with the phases and a significant interaction ($p < .01$) among alone or vicarious experience, mode of aggression, and status of the instigator.

DISCUSSION

The expressing of aggression, whether socially approved or condemned, carries with it the possibility of retaliation which should be subjectively arousing. The fact that an associate is present to alleviate some of this arousal has been postulated wherein the ally

could respond for the subject. Baker and Schaie (1) found that the vicar's presence did not produce any significant difference in subject arousal. However, in the present study, it was noted that persons subjected to frustration-manipulation were more aroused when the vicar was present regardless of the status position of the instigator. The general pattern appears to be that the presence of the other person is such to make the subject more aware of his "intellectual difficulties" thus producing increased states of arousal. In similar fashion, when the vicar counteraggresses, whether it be to a high or low status instigator, there appears to be a general lessening of arousal thus lending further support to the notion that perceived or suspected injury to the instigator is at least subjectively comforting to an already aroused subject. It is suggested that the real life emphasis on status differences in the military is very important when that status is exercised on a person in the presence of others, notably his colleagues or contemporaries.

The results further suggest that the interaction of the numerous variables in this multidimensional analysis have further ramifications. Wherein it was just mentioned that status does not differentiate arousal in the presence of others, there are suggestions that status may be effective when people are alone. As to blood pressure in particular, there is the notion that a person is more aroused in the face of a high authority stimulus than at low authority stimulus so long as no one else witnesses the frustration-manipulation interaction. To elaborate, the point is made that high status in a military environment does produce arousal of lower ranking enlisted men. One cannot lose sight of the fact that the subject readily recognizes the high status position as being one where the authority could actively carry out retribution if the subject were to strike back as a result of his self perceived frustration.

This above situation is further elaborated upon when it is noted that those subjects who express counteraggression in an overt manner tend not to show tension reduction to the degree of those expressing aggression covertly. This finding is in direct contrast with previous research in that the tension reduction is typically seen to accompany the overt perception of injury to the instigator. It may well be that the training of responses to authority is to discourage acting out enforced by the need for respect for authority and rather concrete means of retribution should that authority not be honored. It is further suggested that perhaps one way of socially adapting to the military is the internalization of aggression and its possible sublimation or displacement upon or on more appropriate targets.

In conclusion, this study lends some doubt to many prior studies regarding the attenuation and expression of aggression in the classical frustration-aggression paradigm. This writer suspects that in an environment such as the military where authority and

status are reinforced and supported that typical modes of response and tension release may be somewhat altered. Enlisted men do tend to be aroused by high status individuals in a one-to-one interaction but are not prone to release tension by overt retaliation or counteraggression. However, when persons are frustrated in the face of others, regardless who instigates the frustration, there is a high degree of arousal. One might seriously think about the gains of punishing individuals before a group as a means of instilling further anger and hostility as opposed to leading to adaptive reactions.

Certainly the entire notion and range of status in the military has stimulus value far beyond initial observations. This study but opens the door for one to question even further and seek more evidence of the effect of social stimulus variables on the multiple means of expressing aggression.

BIBLIOGRAPHY

1. Baker, J. W., and Schaie, K. W. Effects of aggressing "alone" or "with another" on physiological and psychological arousal. Journal of Personality and Social Psychology, 1969, 12(1), 80-86.

2. Berkowitz, L. Some factors affecting the reduction of overt hostility. Journal of Abnormal and Social Psychology, 1960, 60, 14-21.

3. Berkowitz, L. Aggression: A Social Psychological Analysis, New York: McGraw-Hill, 1962.

4. Berkowitz, L. (ed) Advances in Experimental Social Psychology, New York: Academic Press, 1965.

5. Hokanson, J. E. The effects of frustration and anxiety on overt aggression. Journal of Abnormal and Social Psychology, 1961, 62, 346-351.

6. Hokanson, J. E., and Burgess, M. The effects of three types of aggression on vascular processes. Journal of Abnormal and Social Psychology, 1962, 64, 446-449.

7. Hokanson, J. E., and Shetler, S. The effects of overt aggression on physiological arousal level. Journal of Abnormal and Social Psychology, 1961, 63, 446-448.

8. Scheier, I. H., and Cattell, R. B. Handbook for the IPAT 8 Parallel Form Anxiety Battery. Champaign, Ill.: Institute for Personality and Ability Testing, 1960.

9. Thibault, J. W., and Coules, J. The role of communication in the reduction of interpersonal hostility. Journal of Abnormal and Social Psychology, 1952, 47, 770-777.

SOURCE	df	AROUSAL INDEX					
		BLOOD PRESSURE		PULSE RATE		RESPIRATION	
		MS	F	MS	F	MS	F
Alone-w/Another (A)	1	642.53	3.07	35528.	1.95	15410.	<1
Overt-Covert (B)	1	628.06	2.99	37164.	2.04	15432.	<1
High-Low (C)	1	34829.	166.35**	35984.	1.97	15397.	<1
Phase (D)	2	35757.	170.78**	67699.	3.72*	24639.	<1
AxB	1	19.156	<1	120.42	<1	1222.4	<1
AxC	1	1378.2	6.58*	397.70	<1	745.89	<1
AxD	2	204.03	<1	18480.	1.01	10148.	<1
BxC	1	1752.5	8.37**	599.88	<1	766.62	<1
BxD	2	41.375	<1	18431.	1.01	10163.	<1
CxD	2	353.88	1.69	19361.	1.06	10182.	<1
AxBxC	1	172.69	<1	34977.	1.92	17613.	<1
AxBxD	2	2.500	<1	52895.	2.90	29806.	1.06
BxCxD	2	89.625	<1	53711.	2.95	29915.	1.06
AxCxD	2	74.562	<1	54051.	2.97	29968.	1.06
AxBxCxD	2	12.625	<1	19320.	1.06	28271.	1.00
Residual	168	209.37		18217.		28159.	
Total	191						

* $p \leq .05$
** $p \leq .01$

TABLE 1: ANALYSIS OF VARIANCE FOR AROUSAL INDICES

SOURCE	df	AROUSAL INDEX			
		GSR		ANXIETY	
		MS	F	MS	F
Alone-w/Another (A)	1	984100.	7.59**	10383000.	1.47
Overt-Covert(B)	1	1955000.	1.51	10302000.	1.46
High-Low (C)	1	3665500.	2.83	10319000.	1.46
Phase (D)	2	32819000.	25.33**	16091000.	2.28
AxB	1	1295200.	<1	148350.	<1
AxC	1	18531000.	14.30**	521480.	<1
AxD	2	2997200.	2.31	6109100.	<1
BxC	1	18670000.	14.41**	443030.	<1
BxD	2	7236700.	5.59**	6076500.	<1
CxD	2	1347000.	1.04	5928300.	<1
AxBxC	1	10382000.	8.01**	11663000.	1.65
AxBxD	2	267110.	<1	16729000.	2.37
BxCxD	2	566630.	<1	17294000.	2.45
AxCxD	2	322690.	<1	17306000.	2.45
AxBxCxD	2	439030.	<1	9103700.	1.29
Residual	168	1295600.		7063100.	
Total	191				

* $p < .05$
** $p < .01$

TABLE 1 (Cont): ANALYSIS OF VARIANCE FOR AROUSAL INDICES

AD 713 508

DETERMINATION OF LATITUDE AND LONGITUDE OF UNKNOWN STATIONS FROM PHOTOGRAPHS OF A SATELLITE AGAINST STELLAR BACKGROUND INDEPENDENT OF ANY DISTANCE

ANGEL A. BALDINI
U. S. ARMY ENGINEER TOPOGRAPHIC LABORATORIES
FORT BELVOIR, VIRGINIA

The method is based on observations of an active or passive satellite, photographed against a stellar background, from a camera situated at an unknown station.

We can then fix the station position's coordinates, geocentric latitude and longitude, independently of any distance.

To make a solution, a minimum of two photographs is required, but certainly with more photographs better results can be expected.

The elapsed period of time between the two photographs can be as small as a second or as large as many hours or days. In our example of computation the elapsed period of time is approximately 5 seconds.

To make a solution the satellite's geocentric equatorial coordinates, right ascension and declination must be known for a given sidereal time or universal time UT1.

The satellite geocentric right ascension and declination can be obtained using known orbital elements, or simultaneous or quasi-simultaneous observations taken against a stellar background from two known stations, or the sequential collation of ranges, obtained from radio signals flashed from three known ground stations to the satellite transponder and returned to the stations during the period at which the satellite is within tracking range, as the SECOR Satellite does.

Since the observed topocentric equatorial coordinates, right ascension and declination of the satellite are more accurate than the geocentric coordinates as computed from orbital elements, better results can be expected using the method of simultaneity, or quasi-simultaneity of observations.

BALDINI

MATHEMATICAL RELATIONSHIP BETWEEN THE TOPO AND GEOCENTRIC EQUATORIAL COORDINATES FOR DETERMINING THE INCLINATION AND LONGITUDE OF THE ASCENDING NODE OF A FICTITIOUS ORBIT

Let α_0 and δ_0 be the geocentric right ascension and declination of a satellite, and α and δ the topocentric right ascension and declination.

With these data fictitious instantaneous orbital parameters can be defined. Due to the fact that the topocentric satellite direction changes very fast, the inclination and longitude of the ascending node does also. Therefore, a small interval of time between two photographs is sufficient for obtaining the solution.

Naming the fictitious orbit's inclination, i, and the longitude of the ascending node, Ω, we have:

$$\begin{aligned} \tan \delta_1 &= \tan i_1 \sin (\alpha_1 - \Omega_1) \\ \tan \delta_0 &= \tan i_1 \sin (\alpha_0 - \Omega_1) \end{aligned} \tag{1}$$

By developing $\sin (\alpha - \Omega_1)$ we obtain:

$$\begin{aligned} \tan \delta_1 &= \sin \alpha_1 \tan i_1 \cos \Omega_1 - \cos \alpha_1 \tan i_1 \sin \Omega_1 \\ \tan \delta_0 &= \sin \alpha_0 \tan i_1 \cos \Omega_1 - \cos \alpha_0 \tan i_1 \sin \Omega_1 \end{aligned} \tag{2}$$

Introducing as unknowns the quantities X_1, Y_1, related to the inclination i, and longitude Ω_1 through the equations:

$$\begin{aligned} X_1 &= \cos \Omega_1 \tan i_1 \\ Y_1 &= \sin \Omega_1 \tan i_1 \end{aligned} \tag{3}$$

we obtain:

$$\tan \Omega_1 = \frac{Y_1}{X_1} \tag{4}$$

$$\tan i_1 = X_1 \cos \Omega_1 + Y_1 \sin \Omega_1 \tag{5}$$

So equation (2) can be rewritten as follows:

$$\begin{aligned} \tan \delta_1 &= X_1 \sin \alpha_1 - Y_1 \cos \alpha_1 \\ \tan \delta_{01} &= X_1 \sin \alpha_{01} - Y_1 \cos \alpha_{01} \end{aligned} \tag{6}$$

With these two equations (6), we can define the values for X_1 and Y_1 and we can compute the longitude, Ω_1, of the ascending node from equation (4), and the inclination i, from equation (5).

From the second photograph we have:

$$\tan \delta_2 = X_2 \sin \alpha_2 - Y_2 \cos \alpha_2$$

$$\tan \delta_{02} = X_2 \sin \alpha_{02} - Y_2 \cos \alpha_{02} \qquad (6a)$$

$$\tan \Omega_2 = \frac{Y_2}{X_2} \qquad (4a)$$

$$\tan i_2 = X_2 \cos \Omega + Y_2 \sin \Omega \qquad (5a)$$

DETERMINATION OF THE GEOCENTRIC LATITUDE AND LONGITUDE OF THE UNKNOWN STATION

The geocentric latitude is related to X and Y and the Greenwich sidereal time in correspondence to the observed universal time UT1, through the equation:

$$\tan \beta = \cos \lambda \, (X_i \sin \theta_{gi} - Y_i \cos \theta_{gi}) - \sin \lambda \, (Y_i \sin \theta_{gi} - Y_i \cos \theta_{gi}) \qquad (7)$$

θ_{gi} = Grennwich sidereal time

$i = 1, 2, 3$

In order to determine β and λ let

$$A_1 = X_1 \sin \theta_{g_1} - Y_1 \cos \theta_{g_1}$$

$$B_1 = X_1 \cos \theta_{g_1} + Y_1 \sin \theta_{g_1} \qquad (8)$$

from the first observations, and let

$$A_2 = X_2 \sin \theta_{g_2} - Y_2 \cos \theta_{g_2}$$

$$B_2 = X_2 \cos \theta_{g_2} + Y_2 \sin \theta_{g_2} \qquad (9)$$

from the second observations.

Then from the two photographs we find the two equations:

$$\tan \beta = A_1 \cos \lambda - B_1 \sin \lambda$$

$$\tan \beta = A_2 \cos \lambda - B_2 \sin \lambda \qquad (10)$$

From which, by subtraction we obtain:

$$\cos \lambda \, (A_2 - A_1) - (B_2 - B_1) \sin \lambda = 0 \qquad (11)$$

which allows us to determine the longitude λ, from:

$$\cot \lambda = \frac{B_2 - B_1}{A_2 - A_1} \tag{12}$$

when

$$|A_2 - A_1| > |B_2 - B_1|$$

and

$$\tan \lambda = \frac{A_2 - A_1}{B_2 - B_1} \tag{13}$$

when

$$|A_2 - A_1| < |B_2 - B_1|$$

After the longitude is known, the geocentric latitude is computed through equation (10).

LEAST SQUARES SOLUTION

If one has observed more than two satellite positions on the same day or on different days, one has more than sufficient data to compute the coefficients A and B. For this case, by combining all the observations, one obtains a more accurate result than from only two observations.

The equations are independent, therefore the method of least squares can be used for computing λ and β. The processes of the least squares adjustment can be carried out as follows:

Introduce two unknown parameters ξ and η, defined by:

$$\begin{aligned} \sin \lambda &= \xi \tan \beta \\ \cos \lambda &= \eta \tan \beta \end{aligned} \tag{14}$$

so that the longitude is computed from:

$$\tan \lambda = \frac{\xi}{\eta} \qquad \text{when} \quad |\xi| < \eta \tag{15}$$

or

$$\cot \eta = \frac{\eta}{\xi} \qquad \text{when} \quad |\eta| < \xi$$

and the geocentric latitude from equation:

$$\tan \beta = \frac{\sin \lambda + \cos \lambda}{\xi + \eta} \tag{16}$$

Introducing in equation (1), the values given by equation (14), one obtains:

$$-\xi A_i + \eta A_i = 1 \tag{17}$$

$$i = 1, 2, \ldots, m$$

The A_i and B_i are computed through equation (9).

The normal equations, according to the method of least squares, will be:

$$\begin{aligned} \xi\Sigma B - \eta\Sigma(AB) + \Sigma B &= 0 \\ \xi\Sigma(AB) - \eta\Sigma A + \Sigma A &= 0 \end{aligned} \qquad (17a)$$

from which ξ and η are computed. The longitude is computed from equation (15). For determining the quadrant for λ, equation (14) is used. If the station is in the Northern Hemisphere, then the sign of sin λ is the same as the sign of ξ, and the sign of cos λ, is the same as that of η. For a station in the Southern Hemisphere, the sign of sin λ and cos λ are opposite to the signs of ξ and η.

At any rate an approximate value of λ and β or φ can be known.

METHOD OF REDUCING THE STATION TO SATELLITE DIRECTIONS GEO-CENTERED PLUMB-LINE ORIENTATIONS

It is not the purpose of this writer to describe all details concerning this method of reduction, but the fundamental concepts and equations to be used are given. Details will be published later.

As a rule, the direction of a plumb-line from a random station does not meet the earth's axis of rotation but passes close to it.

The astronomical meridian is a plane defined by the direction of the plumb-line and a parallel to the earth's axis of rotation. Therefore this plane runs parallel to the earth's axis of rotation and its separation from the axis is always small.

Consider a set of ground observing stations widely separated one from the other. Two astronomic meridians intercept and the line so formed runs parallel to the earth's axis of rotation. A set of lines so formed are all parallel, but with random position with respect to the earth's true axis of rotation. The mean of these lines will run almost coincident with the earth's true axis of rotation. On this mean line the nearest point to the plane of the true equator must be defined. This point will be called a Local Geograv Point. It is very close to the earth's true mass center.

A set of station to satellite directions resulting from the photographs of an object in space shall be referred to the Local Geograv Point. There is one such point for each set of observing stations. These points, which are considered to be the origins of Local Geograv Coordinate Systems, are taken to be first approximations of the earth's center of mass.

Photographs of a target at different times makes it possible to determine the ground station orientations in reference to a stellar position, independent of any terrestrial data other than that from the photographs themselves. But even if the satellite position is only known by ranges this position can be related to a star background through an angular adjustment.

To define the Geograv Point each local vertical must rotate through an angle ϵ, which is unknown and varies from point to point.

It is well verified that the earth can be represented by an ellipsoid of revolution. Therefore, if one chooses the shape of any ellipsoid, then the Local Geograv Point can be defined. Consequently, the target's apparent position in space, against a stellar background at a known Universal Time, can be referred to that point.

If the observing stations are distributed North and South of the equator, the shape of the ellipsoid used is not of importance. For any flattening chosen, the position of the Geograv Point will lie on the true equator. Therefore, a satellite photographed simultaneously from camera sites located North and South of the equator and reduced to its position relative to the Geograv Point would be almost consistent with a true geocentric position highly accurate in right ascension and declination. But if the stations are located all North or all South of the equator, then the satellite coordinate in declination will have a systematic error positive or negative. This value will depend on the value of f.

To define the Geograv Point each plumb-line must be rotated through an angle ϵ. This rotation is in the astronomical meridian. When both astronomic and geodetic coordinates are known, the average value of latitude and longitude must be considered. The angle of rotation ϵ is unknown, but a good approximation of it can be derived by choosing the flattening of a theoretical earth. No matter what method is used, an error in its values will affect only the declination, not the right ascension, at least for practical purposes. Therefore the best known one must be used. The best one is derived from satellite data. The flattening is f = 1/298.25.

SATELLITE GEOGRAV COORDINATES

To pass from topocentric to Geograv coordinates two parameters, ξ and η, are computed at each observing station:

$$\xi = (B_2 - B_1)\sin(\theta_1 - \alpha_1) \qquad \eta = (A_1 - A_2)/\sin(\theta_1 - \alpha_1)$$

where

$$A_1 = \tan\beta \sin\alpha_1 \qquad B_1 = \tan\delta_1 \cos\theta_1$$
$$A_2 = \tan\delta_1 \sin\theta \qquad B_2 = \tan\beta \cos\alpha_1 \qquad (18)$$

α_1 , δ_1 : satellite topocentric right ascension and declination.
θ_1 : local sidereal time.
$\tan \vartheta = [(1-f)^2 + 2f\ H/a] \tan \varphi$
a = 6378 km
f = 1/298.25
φ = astronomic latitude
H = height of the station in km.

Because the term 2 f/a is very small, it is enough to know H within an error ±200 m. With the ξ and η values so obtained for each station, the satellite position with respect to the Geograv Point can be determined. Each station provides an equation in three unknowns.

Let α_0 and δ_0 be the right ascension and declination of the satellite with respect to the Geograv Point. Then each station furnishes an equation of the form:

$$\xi_i \sin \alpha_0 + \eta_i \cos \alpha_0 + \tan \delta_0 = V_i \qquad (19)$$

$i = 1, 2, \ldots, n$

If now one lets:

m = the number of observations
A = the sum of the squares of ξ
B = the algebraic sum of the products $\xi\eta$
C = the algebraic sum of ξ
D = the sum of the squares of η
E = the algebraic sum of η

the normal equations, by the method of least squares, will be:

$$\begin{aligned} A \sin \alpha_0 + B \cos \alpha_0 + C \tan \delta_0 &= 0 \\ B \sin \alpha_0 + D \cos \alpha_0 + E \tan \delta_0 &= 0 \end{aligned} \qquad (20)$$

from which it follows:

$$\begin{aligned} \tan \alpha_0 &= (EB - CD)/(BC - AE) \\ \tan \alpha_0 &= -(C \sin \alpha_0 + E \cos \alpha_0)/m \end{aligned} \qquad (21)$$

α_0 cannot differ too much from any of the observed topocentric values α_i. Always $|\alpha_0 - \alpha_i| < 90^0$.

EXAMPLE OF COMPUTATION

As an example of the method previously described, a practical application of the computational procedure is given.

On 3 January 1964 a flash sequence of five flashes from the ANNA 1-B geodetic satellite were successfully photographed from six camera sites. Satellite photographs were taken with a PC-1000 which has a 1000 mm focal length 200 mm aperture, and 10° square field of view. The topocentric right ascensions and declinations resulting from the photographs were adjusted and the results are shown on page 10. These data and the astronomic coordinates of the camera sites indicated below were used for computing the satellite position to the Geograv System, according to formulas (18), (20) and (21).

The station astronomic coordinates used in this investigation are:

ASTRONOMIC STATION COORDINATES

STATION	LONGITUDE	LATITUDE	HEIGHT SEA LEVEL
640	90°40'45.''36	29°33'48.''44	2.0
641	95 09 15.57	29 35 41.15	8.2
643	92 31 33.02	31 19 14.88	26.8
647	88 04 44.98	30 14 49.94	1.2
648	81 09 09.13	32 00 09.50	12.2
649	80 04 51.04	26 57 17.40	6.8

The satellite position on the Geograv System computed by a pair of stations is shown on pages 11 and 12. The computation by the method of least squares is shown on page 13. The results from the least squares adjustment are:

SATELLITE GEOGRAV COORDINATES

FLASH	RIGHT ASCENSION	DECLINATION
1	100°00'53.''49	28°09'55.''53
2	100 16 04.55	27 57 14.04
3	100 33 02.30	27 42 57.99
4	100 48 15.73	27 30 04.82
5	101 03 25.35	27 17 10.59

which were used, in addition to those values resulting from the photographs for each camera site, as shown on page 10, for computing the station coordinates. On page 14 the results obtained for station 640 are shown. On the right hand side is the computation from a pair of observations according to formulas (10) and (12), and on the left side the least squares solution following equations (18)

and (20) are shown. Note our changes of sign in the second normal equation. Analogous computations were used for the remaining stations. Their results are:

GEOGRAV STATION COORDINATES

STATION	LONGITUDE	LATITUDE	HEIGHT SEA LEVEL
640	90°40'45."64	29°33'47."96	2.0m
641	95 09 22.78	29 35 45.60	8.2
643	92 31 22.21	31 19 15.79	26.8
647	88 04 42.92	30 14 51.51	1.2
648	81 09 16.33	32 00 07.46	12.2
649	80 04 48.99	26 57 14.84	6.8

Using these new values for the stations, a new computation for reducing the satellite position to the Geograv System will alter the previous results slightly. The new computations are:

FLASH	RIGHT ASCENSION	DECLINATION
1	100°00'53.45	28°04'55.48
2	100 16 04.56	27 57 13.98
3	100 33 02.34	27 42 57.96
4	100 48 15.75	27 30 04.81
5	101 03 25.30	27 17 10.59

Comparing these results with those on page 13 the discrepancies between them are:

$\Delta\alpha$	$\Delta\delta$
-0.04	-0.05
- .05	- .06
+ .04	- .03
+ .02	-0.01
- .05	0

CONCLUSIONS

The tests carried out in this investigation are very encouraging. In the writer's opinion, better results may be achieved by using large telescopes, available at many observatories, to photograph certain objects against the star background. Several classes of celestial objects can be photographed: (1) synchronous earth satellites, (2) the Io satellite of Jupiter, and (3) many of the asteroids.

BALDINI

ADJUSTED APPARENT RIGHT ASCENSIONS AND DECLINATIONS

FLASH SEQUENCE # 34

Station	Obs	Greenwich Sidereal Time	Right Ascension	Declination
640	1	192 07 06.70	92 42 12.66	21 08 44.28
	2	08 30.10	94 42 12.66	19 50 52.23
	3	10 03.60	95 50 27.83	18 22 47.98
	4	11 27.90	97 17 23.09	17 02 59.65
	5	12 52.10	98 42 26.01	15 43 4.49
641	1	192 07 06.70	115 15 53.50	20 20 29.33
	2	08 30.10	116 24 5.86	18 55 44.81
	3	10 03.60	117 38 0.50	17 21 57.34
	4	11 27.90	118 42 20.99	15 58 43.17
	5	12 52.10	119 44 37.48	14 36 50.46
643	1	192 07 06.70	101 58 40.39	12 05 23.52
	2	08 30.10	103 15 49.20	10 49 08.18
	3	10 03.60	104 40 10.18	9 24 28.06
	4	11 27.90	105 54 12.48	8 09 07.76
	5	12 52.10	107 06 26.33	6 54 50.38
647	1	192 07 06.70	81 02 02.20	16 33 09.21
	2	08 30.10	82 28 25.34	15 25 03.28
	3	10 03.60	84 04 18.75	14 07 34.71
	4	11 27.90	85 29 46.30	12 56 56.05
	5	12 52.10	86 54 14.96	11 45 43.38
648	1	192 07 06.70	61 54 01.10	6 24 32.61
	2	08 30.10	62 55 47.52	5 36 01.17
	3	10 03.60	64 05 17.53	4 40 42.97
	4	11 27.90	65 08 05.95	3 50 06.43
	5	12 52.10	66 10 58.05	2 58 51.15
649	1	192 07 06.70	51 52 44.56	22 57 38.99
	2	08 30.10	52 54 15.33	22 21 00.49
	3	10 03.60	54 03 58.81	21 38 20.38
	4	11 27.90	55 07 20.13	20 58 26.68
	5	12 52.10	56 11 09.29	20 17 12.40

SATELLITE POSITION ON THE GEOGRAV SYSTEM

FROM STATION	FLASH 1		FLASH 2		FLASH 3	
	α	δ	α	δ	α	δ
1-2	100°00'53.42	28°09'55.59	100°16'4.49	27°57'13.97	100°33'2.26	27°42'57.95
1-3	53.45	55.59	4.56	13.99	2.34	58.00
1-4	53.50	55.61	4.36	13.92	2.18	57.90
1-5	53.28	55.56	4.44	13.95	2.31	57.99
1-6	53.11	55.52	4.68	14.04	2.36	58.01
2-3	100 00 53.46	28 09 55.55	100 16 4.59	27 57 13.90	100 33 2.40	27 42 57.87
2-4	53.39	55.81	4.56	14.00	2.34	48.00
2-5	53.51	55.36	4.53	14.10	2.19	58.34
2-6	54.25	55.52	3.88	15.88	1.89	59.00
3-4	100 00 53.45	28 09 55.64	100 16 4.59	27 57 13.95	100 33 2.38	27 42 51.94
3-5	53.46	55.59	4.58	13.96	2.35	57.99
3-6	52.50	55.48	4.53	14.07	2.34	58.01
4-5	100 00 52.52	28 09 55.18	100 16 4.75	27 57 14.05	100 33 2.94	27 42 58.33
4-6	52.85	55.32	5.00	14.17	2.63	58.15
5-6	100 00 52.93	28 09 55.33	100 16 5.09	27 57 14.19	100 33 2.48	27 42 58.12

SATELLITE POSITION ON THE GEOGRAV SYSTEM

FROM STATION	FLASH 4						FLASH 5					
	α			δ			α			δ		
1-2	100°	48'	15".91	27°	30'	4".81	101°	03'	25".30	27°	17'	10".57
1-3			15.76			4.65			25.30			10.57
1-4			15.96			4.88			25.32			10.60
1-5			15.91			4.81			25.31			10.58
1-6			15.92			4.83			25.27			10.51
2-3	100	48	15.60	27	30	4.97	101	03	25.31	27	17	10.56
2-4			15.82			4.56			25.26			10.64
2-5			15.90			4.39			25.27			10.63
2-6			15.82			4.54			25.54			10.16
3-4	100	48	15.67	27	30	4.74	101	03	25.29	27	17	10.61
3-5			15.68			4.73			25.30			10.61
3-6			15.63			4.78			25.35			10.55
4-5	100	48	15.58	27	30	4.66	101	03	25.25	27	17	10.57
4-6			15.82			4.82			25.10			10.47
5-6	100	48	15.98	27	30	4.85	101	03	24.99	27	17	10.44

BALDINI

COMPUTATION OF THE GEOGRAV RIGHT ASCENSIONS AND DECLINATIONS

FLASH 1

ξ	η		NORMAL EQUATIONS
-.33009647	1.20960150		
-.62366323	-.45283611	5.80919690	+19.01433200 -4.50865650=0
-2.15645160	-9.13249490		85.94143500 -7.05817960=0
-.41546210	.72617534		
-.44022143	.58598652		α = 100°00'53.''49
-.54276166	.00539004		δ = 28 09 55.53

FLASH 2

-.24107505	1.64635050	3.88841750	+ 9.85943570 -3.89819210=0
-.62918522	-.49614555		43.54162800 -3.65603480=0
-1.67791500	-6.28530420		
-.38830543	.83358620		
-.42699457	.62002345		α = 100°16'04.''55
-.53471686	.02545485		δ = 27 57 14.04

FLASH 3

-.08980173	2.38709120	3.02600030	+ 6.36234410 -3.44493220=0
-.63457609	-.53785338		30.48556900 -1.28049490=0
-1.42996410	-4.80827960		
-.35374236	.96997467		
-.41151962	.65976706		α = 100°33'02.''30
-.52532829	.04880520		δ = 27 42 57.99

FLASH 4

.13974359	3.50939940	2.65072150	+ 5.46511050 -3.03356690=0
-.63885086	-.57044520		30.76743100 + .76703657=0
-1.30288460	-4.04998600		
-.31809940	1.11031800		
-.39695296	.69710535		α = 100°48'15.''73
-.51652271	.07064513		δ = 27 30 04.82

FLASH 5

.58819005	5.69959650	2.72497460	+ 7.38226480 -2.43987660=0
-.64267366	-.59917080		47.60048000 +3.65211590=0
-1.21898430	-3.54834710		
-.27725707	1.27082490		
-.38177416	.73594303		α = 101°03'25.''35
-.50737736	.09326937		δ = 27 17 10.59

COMPUTATION OF THE COORDINATES OF THE STATION 640

Obs	Greenwich Sidereal Time	Topocentric Right Ascension	Topocentric Declination	Geocentric Right Ascension	Geocentric Declination
1	192°07'06".70	92°42'12".66	21°08'44".28	100°00'53".49	28°09'55".53
2	08 30.10	92 42 12.66	19 50 52 23	100 16 04 55	27 57 14.04
3	10 03.60	95 50 27.83	18 22 47.98	100 33 02.30	27 42 57.99
4	11 27.90	97 17 23.09	17 02 59.65	100 48 15.73	27 30 04.82
5	12 52.10	98 42 26.01	15 43 4.49	101 03 25.35	27 17 10.59

From Equation 6 X	From Equation 6 Y	From Equation 8 A	From Equation 8 B
0.33009718	1.20958550	1.11333150	-0.57667511
.24107550	1.64634340	1.55881000	- .58195821
.08980166	2.38709100	2.31453720	- .59091899
-.13975304	3.50947220	3.45984140	- .60450316
-.58820147	5.69966920	5.69508960	- .63100240

Least Squares Solution

$$1.78400290 + 3.60199710 - 2.98505790 = 0$$
$$-8.60199710 - 53.43102600 + 14.14161000 = 0$$

$\xi = 1.77470160$
$\eta = 0.02104337$
$\cot n\lambda = \eta/\xi$

$\lambda = 90°40'45".64$
$\beta = 29\ 23\ 54.41$
$\varphi = 29\ 33\ 47.96$

$\Sigma_{BB} = 1.78400290$
$\Sigma_{AB} = -8.60199710$
$\Sigma_{B} = -2.98505790$
$\Sigma_{AA} = 53.43102600$
$\Sigma_{A} = 14.14161000$

Computation by Pair of Observations

Computation of Longitude

$a = A_2 - A_1 = 0.44547850$
$b = B_2 - B_1 = 0.00528310$
$\cot \lambda = b/a$
$\lambda = 90°40'46".05$

Computation of Latitutde

$\sin \lambda$	0.99992969	
$\cos \lambda$	0.01185852	
$A \cos \lambda$	0.01320247	-0.01848518
$-B \sin \lambda$	0.57663456	0.58191729
$\tan \lambda$	0.56343209	0.56343209

$\beta = 29°23'53".90$
$\varphi = 29\ 33\ 47.36$

AD 713 509

ANALYSIS AND APPLICATION OF GALLIUM ARSENIDE AVALANCHE TRANSIT TIME DEVICES

JOSEPH J. BARANOWSKI and VINCENT J. HIGGINS
Electronic Components Laboratory
US Army Electronics Command
Fort Monmouth, New Jersey 07703

INTRODUCTION

The generation of coherent microwave energy from a reverse biased diode was proposed by Read[1] in 1958 and first realized experimentally by Johnston[2] et. al., using a silicon p-n junction biased into avalanche, in early 1965. In that same year microwave oscillations were first reported in gallium arsenide p-n junctions by the authors[3] and later from germanium n-p junctions by Lowe.[4] Since then advances in the development of the avalanche transit time diodes have been so rapid that today this device is considered the most important new solid state source of microwave power.

Until recently most research on avalanche transit time effects centered on diodes of silicon and germanium. The primary reason for this concentration was the comparative (i.e., with gallium arsenide) advanced state of the silicon and germanium materials technology. However, the superiority of gallium arsenide as a basic material, despite the lack of material refinements and technology, had been observed experimentally and confirmed theoretically.[5-6] These views were mainly ignored until recent developments placed gallium arsenide avalanche transit time devices ahead of the state-of-the-art of their silicon and germanium counterparts. A single gallium arsenide oscillator has generated 2.1 watts of CW power in X-band with 10% DC to rf efficiency.[7] This represents the highest CW output power from a single junction device regardless of material. Before this attainment gallium arsenide avalanche devices always outperformed silicon and germanium in terms of noise and efficiency.[8]

We in this laboratory have been working continually with gallium arsenide avalanche transit time diodes since 1965. This paper will report the theory, analysis, and design with application to actual device fabrication and application of devices and circuits to USAECOM systems requirements.

ANALYSIS:

The equations which govern the dynamics of an avalanching electron-hole plasma in the space-charge region of a reverse biased gallium arsenide P-N junction are Poissons' equation

$$dE/dx = q/\epsilon \cdot (n - p - N_d + N_a) \qquad (1)$$

the continuity equations for electrons and holes

$$dn/dt = V_n\, dn/dx + \alpha(E)(V_n n + V_p p) + J_{DC}\, \alpha(E)/q + 1/q\, d\alpha(E)/dE\, E\, J_{DC} \qquad (2)$$

$$dp/dt = V_p\, dp/dx + \alpha(E)(V_n n + V_p p) + J_{DC}\, \alpha(E)/q + 1/q\, d\alpha(E)/dE\, E\, J_{DC} \qquad (3)$$

and the equation for total electric current density

$$J_{total} = q\,(V_p p + V_n n) + \epsilon \cdot dE/dt \qquad (4)$$

The following parameters are defined: q is the magnitude of the electronic charge; ϵ is the dielectric constant; N_d and N_a are the ionized donor and acceptor impurity densities; E is the electric field; n and p are the electron and hole densities; d/dt and d/dx are the derivatives with respect to time and distance; $\alpha(E)$ is the ionization rate for electrons and holes; J_{DC} is the DC electric current density; V_n and V_p represent the electron and hole saturated drift velocities.

The quantities E, n, p, J_{total} are composed of a constant term plus a small time dependent term which varies as $e^{j\omega t}$. Carrier generation consists of an average, noiseless rate, $\alpha(E)\cdot(V_n n + V_p p)$, and a stochastic rate, $J_{DC}\alpha(E)/q$. Generation and recombination through defects in the space-charge region is generally a second order effect and is neglected in this analysis.

The ionization rate reported by Sze and Gibbons[9] is expressed as: $\alpha(E) = A \cdot EXP\,(b/E)^m$. Since the ionization coefficient is dependent on E, the continuity equations are nonlinear in E. The continuity equations are linearized by adding the term $1/q\, d\alpha(E)/dE\, E\, J_{DC}$ which contains the derivative of the ionization rate with respect to electric field.

The total current density, equation (4), consists of both particle and displacement currents. Since the electric field is always very high in the space-charge region the diffusion current is negligible and not included as a term in the total current equation.

The physical model to be used in this analysis is an abrupt P-N junction with $N_a \gg N_d$. For gallium arsenide this is a very realistic choice because P-N junctions formed by either zinc-diffusion or epitaxial growth exhibit very abrupt characteristics. Since N_a is typically several orders of magnitude greater than N_d, the

depletion layer will essentially spread only into the n-region. To a good approximation then, a one-sided junction may be treated. The depletion region can be separated into an avalanche region and a drift region. Electrons and holes are generated by impact ionization in the avalanche region and drift toward the substrate and the junction respectively. The avalanche region, which is approximately 40% of the total space-charge width in gallium arsenide can be defined as that portion of the total depletion width where 99.9% of all carrier generation takes place. In this region the electric field exceeds 200 kV/cm and the carrier drift velocities are considered to be field independent. In the drift region there are no holes present and the electrons drift from the high field side to the substrate. The electron drift velocities are known to saturate at electric fields above a few KV/cm and are therefore considered to be field independent in this region also.

The solution of equations (1) through (4) is obtained by separating them into DC and AC parts. The DC equations are solved with the following boundary conditions; neglecting saturation currents (J_{ns}, J_{ps}).

P^+ - N JUNCTION: $x = 0$, $J_p = J_{DC}$, $n = 0$

N - SUBSTRATE: $x = W$, $J_n = J_{DC}$, $p = 0$

Rewriting equations (1) and (3) by eliminating terms in n and solving for electric field, voltage and hole current, one obtains

$$dE/dx = J_p/\mathcal{E} \cdot (1/V_n + 1/V_p) - J_{DC}/\mathcal{E}\, V_n + q\, N_d/\mathcal{E} \qquad (5)$$

$$E = \int_o^W dE/dx\, dx \qquad (6)$$

$$V = \int_o^W E\, dx \qquad (7)$$

$$J_p = J_{DC} \cdot \int_o^W \alpha(E)\, dx \qquad (8)$$

The results of a computer aided numerical evaluation of equations (5) through (8) are illustrated in Figure 1. Given the donor impurity density and DC current density as parameters the electric field and current distributions can be calculated as a function of distance through the depletion region. The diodes reverse operating voltage is also calculated. Details of the computer aided techniques employed will be discussed in a later section.

For the AC case $d/dt = J\omega$, $N_a = N_d = 0$, $J = \sqrt{-1}$, $\omega = 2\pi \cdot$ frequency

Terms E, n, p, J_{total} are AC quantities and are characterized by a magnitude and a phase angle. Differentiation of Poisson's equation substitution into the continuity equations yields a second order differential equation in E.

$$(D^2 - FD + G)\, E = H\, J_{total} + K \tag{9}$$

where $D = d/dx$, $D^2 = d^2 / dx^2$

$$F = J\omega(V_n - V_p)/V_n V_p \tag{10}$$

$$G = ((J\omega)^2 \varepsilon + J\omega\varepsilon\alpha(E)(V_n + V_p) - dx/dE\, J_{DC}\, (V_n + V_p))/\varepsilon V_n V_p \tag{11}$$

$$H = (\alpha(E)(V_n + V_p) - J\omega)/\varepsilon V_n V_p \tag{12}$$

$$K = J_{DC} \cdot \alpha(E)\, (V_n + V_p)/\varepsilon V_n V_p \tag{13}$$

The solution of equation (9) cannot be obtained by simple analytical means since it is nonlinear in E. The linearization technique and evaluation criteria used by the authors in the solution of this equation will be discussed in a later section. The boundary conditions employed for obtaining the diodes small signal AC impedance, transfer impedance, and Q will now be presented.

The AC impedance is evaluated with $K = 0$, and $z(x) = E/J_{total}$. Equation (9) becomes

$$(D^2 - FD + G)z(x) = H \tag{14}$$

and is solved with the following boundary conditions

P^+ - N JUNCTION: $n = 0$, $J_{total} = q\, V_p p + J\omega\varepsilon E$, $dE/dx = q/\varepsilon(-P)$

N-SUBSTRATE: $p = 0$, $J_{total} = qV_n n + J\omega\varepsilon E$, $dE/dx = q/\varepsilon \cdot (n)$

The results of the integration of the solution of equation (14) for various frequencies are shown in Figure 2. At low frequencies the real part of the impedance is positive and then becomes negative with increasing frequency. Likewise, the reactive portion of the impedance is inductive at low frequencies, then resonates and becomes capacitive with increasing frequency. The magnitude of the negative resistance is a maximum at the resonant frequency. If either the net donor density, N_d, or the DC current density, J_{DC}, are increased, the resonant frequency will increase, and vice versa.

The small signal Q of an avalanche transit time diode is also of interest. It is defined as the angular frequency times the ratio of the average stored energy to the average energy dissipated in the diode per cycle and can be written

$$Q = \omega \int_0^1 \langle W \rangle dx / - \int_0^1 \langle dW/dt \rangle\, dx \text{ where} \tag{15}$$

$$\langle W \rangle = 1/2\ \text{REAL}\ (|E|^2/2) \tag{16}$$

$$-\langle dW/dt \rangle = 1/2\ \text{REAL}\ ((J_n + J_p)\, E^*) \tag{17}$$

The Q can be treated as a figure of merit in determining the quality of an oscillation at a given frequency and current density. A smaller value of the magnitude of the Q refers to a higher quality of negative resistance. Figure 3 is a plot of the magnitude of the Q as a function of frequency. The equation for the Q exhibits a pole at the point where the real part of the impedance switches from positive to negative. At this point energy dissipation in the diode is zero. For the diode to sustain oscillations, the circuit Q must be greater in magnitude than the diode Q.

The transfer impedance, Z_t, is defined as the ratio of the open circuit voltage to the internally distributed shot current associated with the avalanche process. Equation (9) with the open circuit boundary conditions on the current, namely $J_{total} = 0$, can be written as:

$$(D^2 - FD + G)\, E = K \tag{18}$$

with boundary conditions

P^+-N JUNCTION: $n=0$, $J_{total} = 0$, $dE/dx = q/\mathcal{E} \cdot (-p)$

N-SUBSTRATE: $p = 0$, $J_{total} = 0$, $dE/dx = q/\mathcal{E} \cdot (n)$

The solution of equation (18) yields the open-circuit voltage, $V_{oc} = \int_o^w E \cdot dx$, which when divided by the internal shot current yields the transfer impedance, Z_t.

$$Z_t = \int_o^w E \cdot dx / I_{shot} \tag{19}$$

The transfer impedance calculation involves the solution of a nonlinear differential equation and will be discussed later with the other impedance calculations.

In the noise analysis the transfer impedance is used to aid in the determination of the open-circuit mean-squared noise voltage, $\langle V^2 \rangle$, and the optimum noise measure, M. An evaluation of $\langle V^2 \rangle$ and M can then lead to an understanding of the observed AM and FM performance of these devices. An equation for $\langle V^2 \rangle$ has been derived for gallium arsenide by the authors and is modeled after a similar calculation made by Gummel and Blue[10] for silicon. The equation for gallium arsenide is

$$\langle V^2 \rangle = 2q\, I_{DC}\, df \int \alpha (E)\, \left| Z_t^2 \right| dx \tag{20}$$

The equation for the optimum noise measure has been given by DeLoach[11]

$$M = \langle V^2 \rangle / df \;/\; 4\, KT\, (\text{REAL } Z_t) \tag{21}$$

An evaluation of equations (20) and (21) is shown in Figure 4 and 5. Both the open-circuit mean-squared noise voltage and the noise measure become maximum near the resonant frequency, at which the diode negative resistance is also a maximum. This indicates that an avalanche transit time osillator which is designed to give maximum power and efficiency at a given frequency would also exhibit maximum noise properties. This unwanted condition can be made less severe through proper diode design and optimization and will be discussed following the presentation of the computer aided evaluation.

COMPUTER AIDED EVALUATION

The model for the computer aided analysis is shown in Figure 6. The depletion region is divided into several hundred sections of width Δx. In each section the electric field and the hole and electron concentrations are constant. The DC equations can then be evaluated in digital form.

Initially in those equations $J_p = E = V = X = 0$, and , V_p, V_n, J_{DC}, q, and N_d are given. Equations (5) through (8) are evaluated and values for dE/dx, E, V, and J_p, are now determined for the first interval. This process is repeated for each succeeding interval until the boundary condition $J_p = J_{DC}$ is reached, thus terminating the calculation. The final calculation gives the maximum electric field, the operating breakdown voltage, and the total depletion layer width. The electric field in each interval is stored in the computer memory and is used in the AC analysis.

Since the depletion region is now represented by many intervals having constant electric, equation (9) will have a simple analytical solution in each interval. The total solution being the sum of each of the individual solutions. As an example consider the evaluation of equation (14) for the small signal AC impedance. In each interval it will have a solution of the form

$$z(x) = A(x)\,EXP\,(ALAM(x) \cdot x) + B(x)EXP(BLAM(x) \cdot x) + C(x) \qquad (22)$$

$$ALAM = F + \sqrt{F^2 - 4G}/2 \qquad (23)$$

$$BLAM = F - \sqrt{F^2 - 4G}/2 \qquad (24)$$

$$C = H/G \qquad (25)$$

where F, G, H are evaluated in each interval, and thus specifies a unique value of ALAM, BLAM, and C in each interval. It can be shown that ALAM refers to a plane wave of holes propagating toward the $P^+ - N$ junction and BLAM a plane wave of electrons moving in the opposite direction. The total diode impedance is expressed as

$$Z_{DIODE} = \int_o^1 z_1(x_1)dx + - - - \int_{m-1}^{m} z_m(x_m)dx = \sum_{x=1}^{m} z(x) \qquad (26)$$

The following terms are defined

$$a(m) = A(m)EXP(ALAM(m)\ x_m) \tag{27}$$

$$aL(m) = A(m)EXP(ALAM(m)x_{m+1}) \tag{28}$$

$$b(m) = B(m)EXP(BLAM(m)\ x_m) \tag{29}$$

$$bL(m) = B(m)EXP(BLAM(m)\ x_{m+1}) \tag{30}$$

where the m's associated with the coefficients are interval coordinates and the m's associated with the x's are position coordinates. The functions a (m), b(m), and c(m) are well behaved and linear in each interval and therefore it can be written

$$z_m(x_m) = z_{m+1}(x_m) \tag{31}$$

$$dz_m(x_m)/dx = dz_{m+1}(x_m)/dx \tag{32}$$

let N = m, and M = m + 1, and solve equations (22), (31) and (32) for aL(m) and bL(m)

$$aL(M) = \frac{a(N)(BLAM(M)-ALAM(N))+b(N)(BLAM(M)-BLAM(N))+BLAM(M)(C(N)-C(M))}{BLAM(M)-ALAM(M)} \tag{33}$$

$$bL(M)= \frac{a(N)(ALAM(N)-ALAM(M))+b(N)(BLAM(N)-ALAM(M))+ALAM(M)(C(N)-C(M))}{BLAM(M)-ALAM(M)} \tag{34}$$

Equations (27) and (28) are rewritten

$$a(M)=\frac{aL(M)EXP(ALAM(M)xm)}{EXP(ALAM(M)x_N)} = AL(M)EXP(ALAM(M)(x_M - x_N)) \tag{35}$$

let $x_M - x_N = \Delta x$, the width of each interval then

$$a(M) = AL(M)EXP(ALAM(M) \cdot \Delta x) \tag{36}$$

A similar equation is obtained for b(M) and following directly

$$z(M) = a(M) + b(M) + C(M) \tag{37}$$

The integral of equation (37) over the interval Δx then gives the diode impedance of the M_{TH} interval.

The solution of equation (14) is essentially a boundary value problem requiring two boundary conditions to determine the properties of the AC field and electron and hole densities. However, by examining equations (33) and (34) one can see that the value of the equation in the M_{TH} interval depends on the value of ALAM, BLAM, and C in the N_{TH} as well as the M_{TH} interval. Because of this dependance the boundary value problem is more conveniently solvable

as an initial value problem requiring only one boundary condition. The boundary is chosen at the N- substrate interface where the hole concentration is zero. Since it is impossible to make p identically equal to zero at this boundary, the next best thing is to make the magnitude of the hole concentration a minimum at this point.

Now consider the actual boundary conditions imposed on the diode at the P^+ -N junction at x = 0, n = 0, the boundary conditions of equations (14),(31) and (32) can be combined and written in this form

$$1/V_p\mathcal{E} = (J\omega/V_p - d/dx)z_o(x) \tag{38}$$

$$z_o(x) = a(o) + b(o) + C(o) \tag{39}$$

$$dz_o(x)/dx = ALAM(0)\ a(o) + BLAM\ (0)\ b(o) \tag{40}$$

substitute equations (39) and (40) into (38) and solve for b(o)

$$b(o) = \frac{(1/V_p\mathcal{E} - J\omega/V_p\ C(o)) - a(o)(J\omega/V_p - ALAM(o))}{(J\omega/V_p - BLAM\ (o))} \tag{41}$$

In equation (41) every term is given except a(o). If the correct value of a(o), which is a complex quantity, were known then every term a(M), b(M), and C(M) and hence Z(M) could be evaluated in turn. The solution proceeds as follows: an initial value for a(o) is selected of the form

$$a(o) = R\ (\cos\Theta + J\sin\Theta) \tag{42}$$

Then all of the a's, b's, and c's are evaluated in turn from x = 0 to x = W. At x = W, the N- substrate interface, the magnitude of p is calculated. A new value of a(o) is chosen by varying Θ and R independently and p is again evaluated at x = W. The computer program searches like a radar by sequentially varying Θ and R until an absolute minimum of p is obtained. At this point the correct value of a(o) has been chosen and the AC impedance is now calculated.

The evaluation of the transfer impedance, Z_t, is obtained in a similar manner. However, one major difference exists. The transfer impedance must be evaluated separately, over the entire depletion region, for the shot current generated in each interval. Thus if the total AC impedance is composed of the sum over M intervals, the total transfer impedance is composed of the sum over M times M intervals.

DIODE DESIGN

The optimum design of an avalanche transit time device requires a detailed knowledge of the character of the AC impedance

and the basic noise performance. The utility of the presented analysis will be indicated in an example of the design of an X-band diode to be used as a local oscillator in a portable system employing a balanced mixer. The basic requirements are 2 mW CW at 11 GHz with minimum noise characteristics and DC power consumption.

In our internal material inventory there is a wafer of epitaxial gallium arsenide with $N_d = 1 \cdot 10^{16}$ cm^{-3} and thickness of 5-microns. With this information the computer selects $J_{DC} = 10^3$ A/cm^2 for the optimum negative resistance and gives the following parameters

Operating Voltage (Volts)	60.5	47.5	30.8
Diffusion Depth (Microns)	2.18	3.5	4.16
V^2/df (V^2 sec)	$2.2 \cdot 10^{-15}$	$2.5 \cdot 10^{-16}$	$9.4 \cdot 10^{-17}$
Negative Resistance (ohms)	-180	-70	-18

The third set of data is chosen as the best design for the local oscillator since it offers the lowest noise and operating voltage. The much reduced negative resistance is also advantageous, since the low microwave power required would normally require heavy padding of the oscillator.

DEVICE PERFORMANCE AND SYSTEM APPLICATION

The analysis described in the preceding paragraphs has directly established the design criteria for fabrication of gallium arsenide avalanche diodes. Further the analysis has provided necessary cross-correlationships between microwave performance characteristics and fabrication controlled diode parameters such as doping level, junction area, breakdown voltage, depletion layer width and thermal resistance.

For an X-band medium-power diode oscillator a particular diode was fabricated as follows: Elemental zinc was diffused into 5-micron thick n on n^+ epitaxial layer of gallium arsenide doped to a concentration of 1×10^{16}/cm^3. The diffusion depth was 1.5 microns for optimum thermal resistance and ohmic contacting purposes. The elemental zinc diffusion produced a surface concentration of 10^{20}/cm^3 and very near an abrupt junction. An epitaxial n-type base of 3.5 microns remained to support a breakdown voltage of 60 volts. The zero bias junction capacitance was 2.5 pf corresponding to a junction area of 10^{-4}cm^2. After deposition of ohmic contacts the diode was bonded junction side down, for heat sinking purposes, to a copper pedestal in a microwave package.

The packaged diode was then mounted in a low Q microwave cavity biased into avalanche, tuned and characterized. This diode generated more than 500 mW at 9.0 GHz with a DC to RF conversion efficiency of 11.0%. A diode oscillator with this RF capability

could find application in moderate power microwave systems as a transmitter or TWT driver.

On the other hand as pointed out earlier, a diode for a local oscillator application would necessarily have different device parameters. Particular attention would be paid to fabrication control for low noise performance with diode parameters derived from the computer aided noise analysis.

There are many Army microwave systems which could effectively use simple reliable direct DC to RF sources. Gallium arsenide avalanche transit time diodes* could find application in these systems as transmitters, TWT drivers, reference oscillators, paramp pump sources, local oscillators and as microwave amplifiers replacing low power ($<$20 w) travelling wave tube (TWT) amplifiers.

In a number of mission oriented experiments GaAsADO's have been designed, built and characterized with respect to their practical and immediate utility in systems. Consider as a first example of this utility the application and actual performance of a GaAsADO in an avionics X-band coherent pulse doppler system. The RF assembly is shown in block form in Figure 7. System operation is straightforward and can be gleaned from the block diagram.

The detection capability of this system was severely limited by high rates incidental FM noise on the carrier signal. The noise originated in the particular type klystron used in the system as CW source, TWT driver and local oscillator. A spectrum analyzer display of the pulsed system output in Figure 8 and clearly shows the FM noise on the carrier and pulse sidebands. In sharp contrast to Figure 8 is Figure 9, the same spectrum analyzer display of the system output but with a GaAsADO as the CW source and TWT driver. The incidental FM noise is greatly reduced and all but eliminated.

This dramatic improvement in system performance is obtained with other advantages to the system, also. The ADO requires about 1 watt or 1/4 of the power required by the klystron with its other obvious disadvantages.

The high level of AM noise observed in most ADO's and perhaps their single major disadvantage had no detrimental effect on system performance. This is because of the noise cancellation properties of the balanced mixer, a conclusion consistent with previous investigations.[12] Preliminary field tests of the system equipped with a GaAsADO indicated a marked improvement in the system capability.

For these systems equipped with balanced mixers, GaAsADO's offer distinct performance advantages over thermionic devices in

*hereafter referred to as GaAsADO

local oscillator applications. Just such a system is the AN/PPS-5, a portable battlefield radar. Functioning as local oscillator is a 2.0 GHz pencil triode with an 8 times multiplier providing 200-500 μW of power. This system suffers from loss of receiver sensitivity due to local-oscillator aging and is difficult to align, very inefficient and has proven unreliable in the field.

The use of a GaAsADO as local oscillator in this system has been investigated and proven feasible. Figure 10 is a plot of the receiver sensitivity as a function of local oscillator power for the gallium arsenide device as compared with a cavity stabilized, low noise klystron. The measurement system was a simulation of the AN/PPS-5 receiver, having a sensitivity of -94 dBm. Besides having very little deviation from the stabilized klystron, the GaAsADO was easily tunable electronically with a varactor for AFC purposes, and the DC power required was less than 500 mW as compared with 3 to 6 watts for the triode multiplier arrangement.

GaAsADO's have CW transmitter application in a homodyne doppler, as velocity sensors for air cushioned vehicles. Oscillators generating 50 mW CW power operating on 60 volts or less with low rates of incidental FM were required. Three oscillators were designed and built internally and are being field tested by Avionics Laboratory, USAECOM.

GaAs avalanche diodes have also provided negative resistance amplification. Gain is obtained when the diode is operated in a circulator coupled network forming a one-port reflection type amplifier. Figure 11 is the gain versus frequency characteristics for a large signal X-band amplifier. Gain in GaAs is generally linear within 10 dB of saturated output. Amplifiers have been built and operated in Ku-band, 14.2-15.3 GHz, for source application in the MALLARD Project local distribution digital communication link. An amplifier in its waveguide circuit is shown in Figure 12, the insert is an enlarged view of the diode.

A measure of reliability is required for systems application. In an operating life test, current since April 1967, reliability and stability data on a GaAsADO has been obtained in Ku-band. Over 25,000 hours has been logged with no perceptible change in output power or frequency with a measurement accuracy of $\pm.5$ mW and ± 5 MHz.

CONCLUSIONS

The theoretical analysis of gallium arsenide avalanche transit time devices has been presented. This analysis has established the criteria for optimum design of oscillators and amplifiers and has led to the application of these devices in Army communications and radar systems.

REFERENCES

1. Read, W. T., "A Proposed High-Frequency, Negative Resistance Diode," Bell System Technical Journal, Vol. 37, March 1968, pp. 401-446.

2. Johnston, R. L., B. C. DeLoach and B. G. Cohen, "A Silicon Diode Microwave Oscillator," Bell System Technical Journal, Vol. 44, February 1965, pp. 369-372.

3. Brand, F. A., V. J. Higgins, J. J. Baranowski and M. A. Druesne, "Microwave Generation from Avalanching Varactor Diodes," Proceedings IEEE (Correspondence), Vol. 53, September 1965, pp. 1276-1277.

4. Lowe, M. E., "Microwave Diode Research," Technical Report, ECOM-01445-2, 2nd Interim, June 1966.

5. Kim, C. K. and L. D. Armstrong, "High Power High Efficiency GaAs Avalanche Diodes," Applied Physics Letters, 1 May 1969.

6. Baranowski, J. J., V. J. Higgins, C. K. Kim, and L. D. Armstrong, "Gallium Arsenide IMPATT Diodes," Microwave Journal, Vol. 12, July 1969, pp. 71-76.

7. Lee, Y. S. and C. K. Kim, "A 2 Watt CW GaAs Schottky Barrier IMPATT," to be published.

8. Brand, F. A., V. J. Higgins and J. J. Baranowski, "Characteristics of CW Silicon and GaAs Avalanche Diode Oscillators," presented before the International Symposium on Microwave Theory and Techniques, Palo Alto, California, May 1966.

9. Logan, R. A. and S. M. Sze, referenced in S. M. Sze and G. Gibbons, "Avalanche Breakdown Voltage of Abrupt and Linearly Graded p-n Junctions in Ge, Si, GaAs, and GaP," Applied Physics Letters, Vol. 8, No. 5, 1 March 1966, pp. 111-113.

10. Gummel, H. K. and J. L. Blue, "A Small Signal Theory of Avalanche Noise in IMPATT Diodes, " IEEE Transactions on Electron Devices, Vol. ED-14, No. 9, September 1967, pp. 569-580.

11. DeLoach, B. C. Jr., "The Noise Performance of Negative Conductance Amplifiers," IRE Transactions on Electron Devices, Vol. ED-9, July 1962, pp. 366-381.

12. Baranowski, J. J., V. J. Higgins, and F. A. Brand, "Investigations as to the Noise Characteristics of GaAs Avalanche Transit Time Diode Oscillators," presented before the 4th Annual Informal Conference on Active Microwave Effects in Bulk Semiconductors, New York, Jan. 1968.

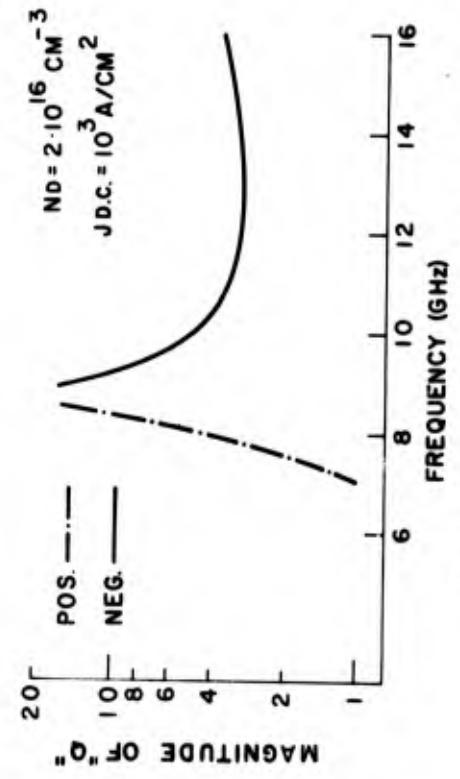


FIG. 3 IQI VERSUS FREQUENCY

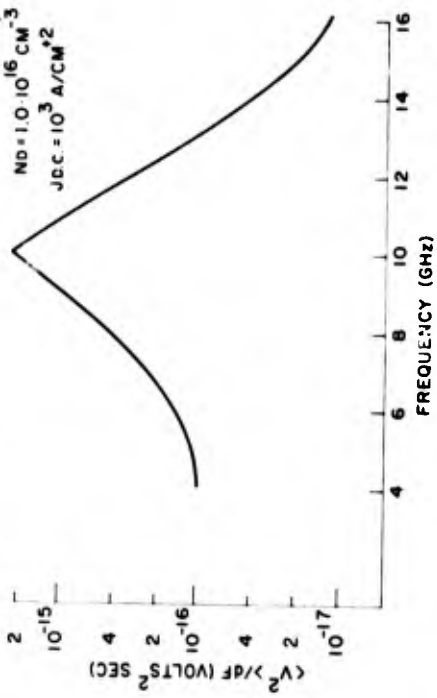


FIG. 4 OPEN-CIRCUIT MEAN-SQUARED NOISE VOLTAGE

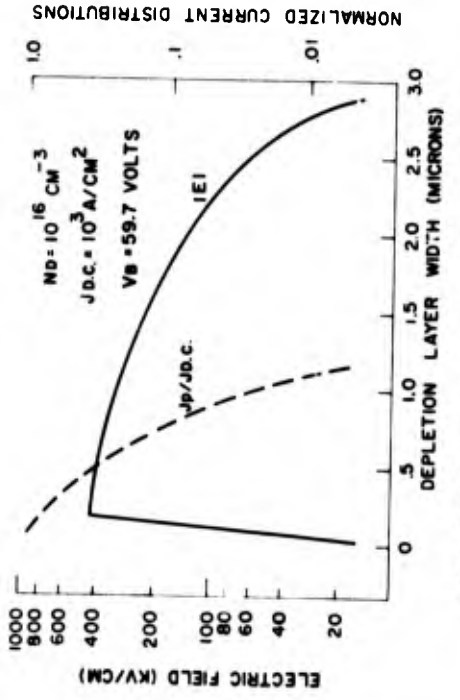


FIG. 1 ELECTRIC FIELD AND CURRENT DISTRIBUTION

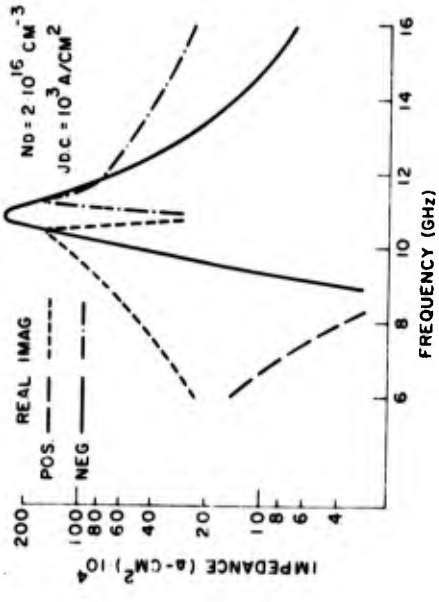


FIG. 2 SMALL SIGNAL A.C. IMPEDANCE

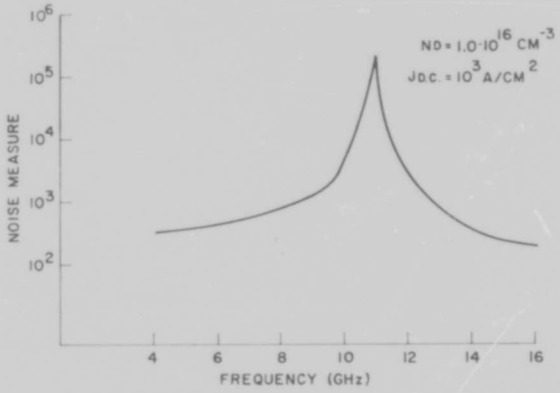


FIG. 5 NOISE MEASURE VERSUS FREQUENCY

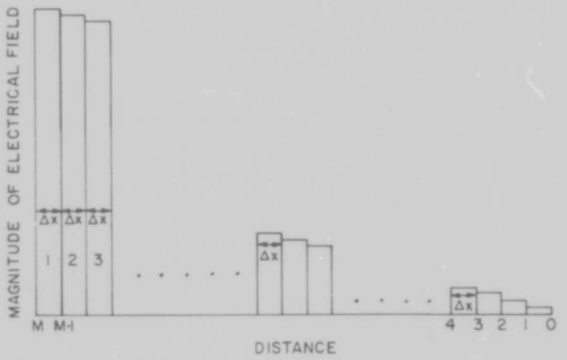


FIGURE 6. MODEL FOR NUMERICAL CALCULATIONS

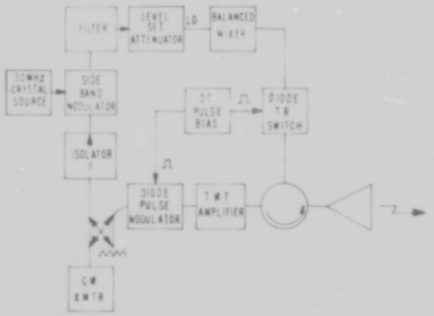
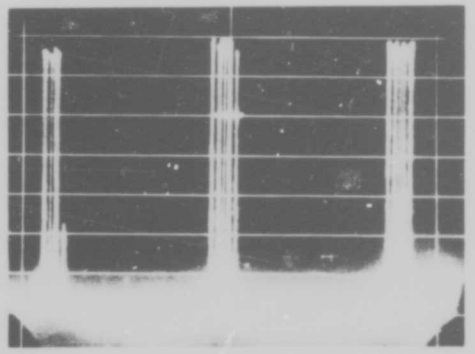


FIG. 7 BLOCK DIAGRAM OF R.F. ASSEMBLY OF AVIONICS COHERENT PULSE DOPPLER SYSTEM.



HIGH RESOLUTION SPECTRUM ANALYZER DISPLAY OF SYSTEM OUTPUT WITH KLYSTRON AS T.W.T. DRIVER

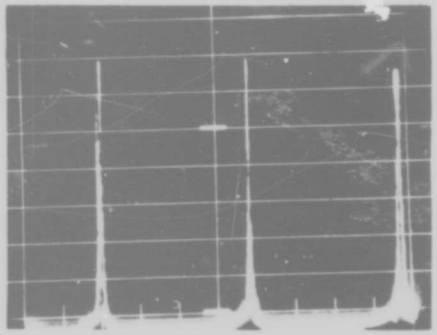


FIG. 9 SPECTRUM ANALYZER DISPLAY OF SYSTEM OUTPUT WITH A GALLIUM ARSENIDE ADO AS A TWT DRIVER.

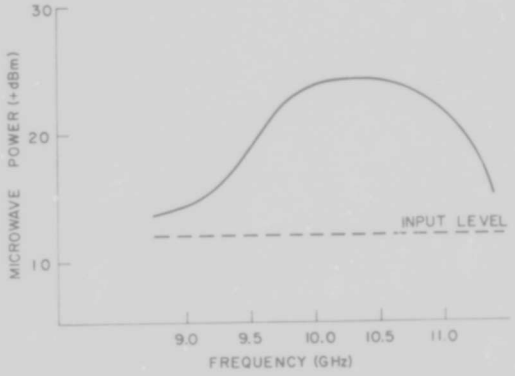


FIG. 11 SINGLE DIODE HIGH LEVEL AMPLIFICATION

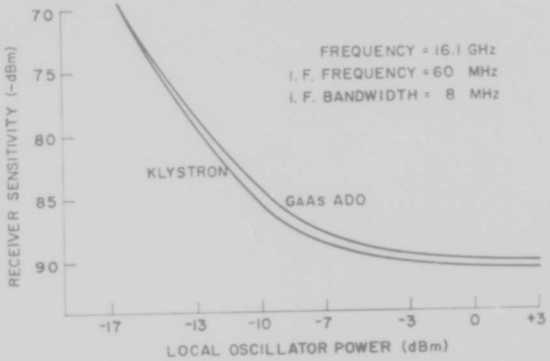


FIG. 10 RECEIVER SENSITIVITY WITH BALANCED MIXER

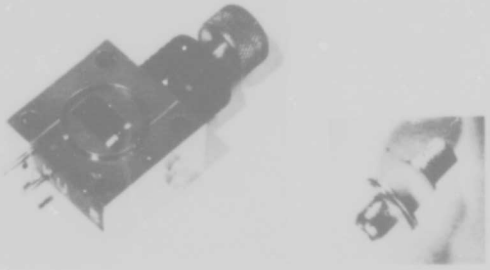


FIG. 12 AVALANCHE DIODE AND SPLIT-WAVEGUIDE OSCILLATOR CIRCUIT.

AD 713 510

ALTERATION OF TASTE QUALITIES THROUGH NATURAL PRODUCTS

LINDA M. BARTOSHUK
U.S. ARMY NATICK LABORATORIES
NATICK, MASSACHUSETTS

Feeding our military forces costs some 2 billion dollars a year; the economic importance of research devoted to maintaining the sensory qualities, the highest nutritional standards, and wholesomeness of this immense quantity of food should be clear to all. Research on the psychological, physiological, and chemical aspects of food is justified for many other reasons. One fundamental reason stands out; unpalatable food is rejected and any wholesale rejection by our military forces is costly.

The investigation to be reported concerns the sensory basis of palatability; specifically, the effects on the taste of foods of some rather remarkable substances derived from two plants, Gymnema sylvestre and Synsepalum dulcificum (more commonly called miracle fruit). These substances are of interest not only because of their potential use in enhancing flavor but also because they can help to elucidate the complex relationship between chemical structure and taste quality.

The purpose of this paper is to present historical, chemical, and sensory information on these two plants and to suggest applications.

History

Gymnema sylvestre

Gymnema sylvestre is a woody, climbing plant which runs over the tops of high trees in a large part of southern India as well as in Ceylon and tropical Africa. Many medicinal properties have been attributed to the leaves of the plant. Some of these ascribed powers are apparently without foundation (for example, ingestion of a decoction of the leaves for snake bite); however, others appear to have originated with observation of the effects of the material on taste.

The first published account of the taste properties originated over a century ago with Edgeworth (1). He reported after chewing the leaves that he could not taste the sweetness of the sugar in his tea. The effect lasted, according to his report, for 24 hours. Hooper (20) confirmed the suppression of the sweetness of sugar and

also observed a suppression of the bitterness of quinine. Shore (33) and Kiesow (24) confirmed both of these suppression effects and also reported a slight suppression of the saltiness of NaCl. Warren and Pfaffmann (37) and Kurihara (27) demonstrated, in addition, that the suppressing effects extended to the sweetness of saccharine, and the sweet taste of the salts, beryllium chloride and lead acetate.

Miracle fruit

Synsepalum dulcificum berries, known as miracle fruit, thanks to the European travelers who sampled it in its native tropical west Africa, grow on a densely foliated bush which can reach a height of 8 feet (fig. 1). The berries are oval shaped and turn a deep red when ripe. The length of the berry reaches about 3/4 inch but the flesh is rather thin, most of the size of the berry resulting from one large seed (fig. 2).

Miracle fruit was probably first described in the notes of a Frenchman in 1725. Apparently, most travelers who encountered it were amazed by the taste of fruits eaten after eating miracle fruit. Although the flesh of the miracle fruit itself is only mildly sweet, normally sour foods eaten after the miracle fruit, taste extremely sweet. Even lemons can be eaten like oranges. The natives in Africa use miracle fruit to sweeten acidulated maize bread (kankies), palm wine, and pito, a sour beer made from fermented grain. The sweetening of palm wine by miracle fruit described by Daniell in 1852 is particularly interesting. It was necessitated by the scarcity of palm trees in some areas (13). The wine was made near the trees and then transported to areas which lacked them. During the transportation the wine began to sour and the miracle fruit was used to make the wine palatable again.

An American explorer working for the department of Agriculture deserves great credit for interesting Americans in this plant. David Fairchild and several companions sampled miracle fruit on a trip to the Cameroon in the late 1920's. He did not realize at first that the "miraculous" powers resided in the effects on other tastes rather than on the taste of the berries themselves. He dismissed them as "not good enough to become excited over though not at all bad." However, sometime later he was offered beer by a chemist he was visiting and he found that it was sweet. Realizing the interest that such an effect would arouse he gathered seeds for introduction into the United States (18).

Miracle fruit is not easy to grow, however, and the initial attempts in the U.S. failed. Fortunately in about 1957 Dr. and Mrs. Otto Churney and Mr. R. G. Newcomb of Florida obtained two seeds from the Summit Garden in Panama. The plants successfully grown from these seeds are the parent plants for most of the miracle fruit now grown in the U.S.

The earliest description of the effects of miracle fruit suggested that it acts on sour and bitter tastes. Dalziel (12) and Irvine (22) reported that it mitigates the bitterness of quinine but that the effects on acidity are greater. In the opinion of Fairchild (18) "the effect is to paralyze some of the papillae of the

tongue so that many things, even acid ones, taste sweet for some time."

Miracle fruit has aroused considerable interest in the popular press, for example, stories in Popular Mechanics, Science News Letter and Food Engineering (2,3,4). Inglett et al.(21) of the research and development division of International Minerals and Chemical Corp. increased interest in miracle fruit by publishing the first effort to identify the active principle. Commercial interest in miracle fruit continues through the Unilever Co. in the Netherlands and through Meditron, Inc. in the U.S. Meditron in particular has extensively investigated growing conditions for miracle fruit plants and has been able to accelerate growth and to increase berry yields.

The great interest in miracle fruit encouraged the U.S. Army Natick Laboratories to undertake a careful examination of the exact effects of the material on taste, but obtaining samples of the berries originally proved very difficult. The first sample of berries tested was generously provided by Dr. and Mrs. Churney. Later, 55 plants were transported from Mr. Newcomb's nursery to greenhouses in Massachusetts. Although a few berries were obtained from the plants, they did not grow well and eventually all but 16 died. Dr. O.B. Dokosi of Ghana attempted to ship berries to the Natick Laboratories but the difficult plane schedules caused all of these shipments to spoil before reaching Massachusetts. Meditron's success in growing miracle fruit finally provided a stable source of material which is now utilized.

Chemical Composition of Active Principles

Work on the chemistry of Gymnema sylvestre leaf samples began with Hooper (20) who in 1887 reported that the active component, which he called gymnemic acid, is a glycoside. In 1959 Warren and Pfaffmann (37) produced a microcrystalline gymnemic acid which also appeared to be a glycoside. Yackzan (38) suggested in addition that gymnemic acid could be a saponin. Stöcklin (36) separated gymnemic acid into four components, A_1, A_2, A_3, and A_4. He concluded that they were probably B-D-glucuronides of different acylated gymnemagenins where gymnemagenin is a hexahydroxy pentacyclic triterpene. The major active constituent found in most leaf samples, A_1, has been shown by Dateo (14) to consist of at least two components.

The first work on the chemistry of miracle fruit by Inglett et al.(21) suggested that the active principle could be a glycoprotein. They were unable, however, to extract it from the fruit. Kurihara and Beidler (26) and Brouwer et al.(11) confirmed the identification and successfully extracted the protein. Kurihara and Beidler estimated the molecular weight of their "taste-modifying protein" at 44,000 and reported that it was 6.7% arabinose and xylose. Brouwer et al. estimated the weight of their "miraculin" at 48,000 and suggested that glucose, ribose, arabinose, galactose, and rhamnose were present.

Experimental Work: Sensory Analyses

Gymnema sylvestre

The first experiment was designed to measure the effects

of Gymnema sylvestre on the four "basic" taste qualities. The Gymnema sylvestre leaves were purified by a procedure developed by Dateo (14). The major component of the purified material was the fraction A_1. The study was carried out using the psychophysical method of direct magnitude estimation (35). According to this method, a subject is given a standard sulution and asked to assign an arbitrary number to its intensity, for example, "100." Then he is given other solutions and asked to rate their intensities relative to the standard. In the present experiment subjects were asked to describe the quality as well as estimate the intensity of taste solutions. The taste solutions used were sucrose for sweet, sodium chloride (NaCl) for salty, hydrochloric acid (HCl) and citric acid for sour, and quinine hydrochloride (QHCl) and quinine sulfate for bitter.

In order to increase the precision of the results, all solutions were kept at body temperature and were delivered to the tongue through a gravity flow system. Each subject sat in front of the flow system with his tongue extended slightly and his lips resting on his tongue. This prevented taste solutions from entering his mouth and also kept his tongue free of saliva. In addition, his tongue was rinsed with distilled water for 40 sec. before each taste solution was presented.

The Gymnema sylvestre leaves were obtained from the Himalaya Drug Co. in Bombay India. The active components of the leaves were purified by a procedure developed by Dr. George Dateo of the Organic Chemistry Labs at Natick. His procedure produced the active material in a water soluble salt form essentially free of carbohydrate, fats, inorganic salts, proteinaceous material and a large portion of the inactive water soluble constituents and avoided thermal degradation. The major fraction of the resulting material was determined by chromatographic comparison of the Natick sample with one provided by Dr. T. Reichstein and Dr. W. Stocklin. This fraction is designated A_1 after Stöcklin (36).

Each of the taste solutions used was tested in a separate session, with the first half of the session used to obtain responses without applying the Gymnema fraction. In the second half the subject held 8 ml. of the Gymnema fraction in his mouth for 30 sec. before each taste stimulus was presented.

The subjects were volunteers from the Behavioral Sciences Division at Natick.

The results are shown in fig. 3. The sweetness of sucrose was substantially suppressed by the Gymnema fraction but none of the other taste substances were significantly influenced. This was particularly astonishing since the bitter suppression reported by Hooper, Shore and Kiesow was completely absent. The failure of the fraction to suppress bitterness in this experiment is of especial importance because no other materials are known which suppress only one taste quality without affecting the others. The existence of such a suppressor plays a major role in taste theories and is a powerful tool in the analysis of complex tastes. It offers as well many possible applications in the area of food intake. The import-

ance of this unexpected experimental result demanded that the discrepancy with the earlier results be explained. With hindsight this is not difficult. The early work with Gymnema sylvestre was carried out by applying crude decoctions of the dried leaves to the tongue or by chewing the leaves directly. The leaves have a very intense bitter taste in themselves. The early investigators apparently did not realize that the taste of the leaves would interfere with their tests and they did not carefully rinse the tongue before testing. The exposure to the bitter taste of the leaves adapted the taste receptors to bitter. When the bitter quinine was tested it quite naturally produced very little bitterness from the adapted receptors. In the present experiment the bitterness of the leaves was decreased by the purification and any remaining bitterness was removed with the distilled water rinse so the bitterness of quinine was not suppressed. The suppression of sweetness remains after the rinse showing that it is not a consequence of any cross-adaptation. The failure of the investigators of the late 19th century to interpret properly the apparent bitter suppression of Gymnema sylvestre is readily understandable. During the years when Gymnema sylvestre first came to the attention of taste investigators (1887-1894) the effects of adaptation were only beginning to receive mention in the taste literature. A convincing demonstration of adaptation was not available until Kiesow's work on single taste papillae was published in 1898 (23). This only established the fact that continued stimulation of a single papilla resulted in loss of sensation to that stimulus. That many taste substances of similar quality could cross-adapt with one another was not established until the early 20th century (5). The next experiment was designed to test the effects on several unusual sources of sweetness chosen to be chemically different from sugar. These included Ca cyclamate as well as several naturally occurring sweeteners that have not all, as yet, been completely characterized chemically. All of the plants investigated have been used as sources of sweeteners in the areas where they grow. The leaves of Stevia Rebaudiana are used by natives in Paraguay and the berries of Dioscoreophyllum Cumminsii and Synsepalum dulcificum as well as the sweet mucilage surrounding the seeds of Thaumatococcus danielli are used by natives in tropical west Africa. Gymnema sylvestre suppressed all of these sweeteners. Sweetness can also be produced under special conditions by substances that normally have other tastes or no taste at all. Salts taste sweet when they are relatively weak and distilled water tastes sweet if the tongue has previously been adapted to bitter or sour substances (7,10,17). These sweet tastes are also suppressed by Gymnema sylvestre.

Miracle Fruit

Two theories have been formulated to explain the taste effects of miracle fruit. Dzendolet (16) suggested that the anions of some acids, for example, the citrate ion of citric acid, are sweet but are normally inhibited by the sour taste. Miracle fruit, by blocking the sour receptor sites, would allow the sweet taste of the anion to be perceived. Kurihara and Beidler (25) suggest, on the other hand, that the glycoprotein, miraculin, binds to the receptor

membrane near the sweet receptor site. Acids then change the conformation of the sweet receptor site so that it will "fit" the sugar groups attached to the glycoprotein, producing a sweet taste. These two theories were tested with a series of psychophysical investigations.

The first experiment was designed to examine the effects of miracle fruit on the four "basic" taste qualities. The procedure was very similar to that used with Gymnema sylvestre. Subjects first judged stimuli under "normal" conditions, i.e., a distilled water rinse preceding the stimulus but no miracle fruit. Then the subject's tongue was exposed to either a miracle fruit berry or a quantity of freeze-dried miracle fruit and the tests run as in the first part of the experiment. The stimuli tested were sucrose, QHCl, NaCl, HCl and citric acid. The judgments of sucrose, QHCl, and NaCl were not significantly affected by miracle fruit but the effects on HCl and citric acid were dramatic. In fig. 4 the filled circles show the functions obtained before miracle fruit and the x's show those obtained after miracle fruit. The total intensity of the taste of HCl and citric acid is not significantly changed after miracle fruit but the quality of the tastes change dramatically. The hatched area shows the part of the total taste that is sweet. The smaller unhatched area shows the part of the total taste that remains sour. The sourness of a lemon would be located near the highest concentration of cirtic acid in fig. 4. Since the citric acid in the lemon would taste much more sweet than sour as shown by the magnitude estimates, it is not surprising that observers find the effect quite startling.

Dzendolet's theory accounts for the suppression of sourness observed after miracle fruit but it cannot explain the HCl data in fig. 4. The chloride ion should not taste sweet according to his theory (15,16) but HCl is clearly sweetened by miracle fruit.

The next experiment was designed to test the effects of removing the miracle fruit induced sweetness with Gymnema sylvestre. The procedure was similar to the previous experiment. Subjects judged stimuli under normal conditions, after miracle fruit, and then again after Gymnema sylvestre. The results are shown in fig. 5. When the sweetness was removed with Gymnema sylvestre the sourness returned to the normal value. Since Gymnema sylvestre has no direct effect on the sourness of citric acid, the return of sourness to the normal level would appear to be due to the removal of sweetness.

The next experiment was designed to investigate the mechanism of the sourness suppression (9). If miracle fruit suppresses sourness only because it adds sweetness then adding sweetness another way should suppress sourness too. Subjects were given mixtures of citric acid and either xylose or arabinose (two sugars found in miraculin). The amount of citric acid in the mixture was always constant but the amount of sugar varied. After judging the sweetness and sourness of the mixtures, the subjects were given miracle fruit and asked to judge another solution. This final solution was the acid alone. The filled circles show that as the concentration of either xylose or arabinose is increased the sweetness naturally

increases; however, the sourness decreases even though the amount of acid is always constant. The x's show the judgments after miracle fruit of the acid alone. The sourness of the acid was suppressed the same amount after miracle fruit that it was by sugar of equivalent sweetness.

The data in figs. 5 and 6 are consistent with the theory of Kurihara and Beidler but this evidence is also consistent with other theories. For example, the miracle fruit protein itself rather than the taste receptor site might be altered by acids to produce a sweet taste. This is a plausible position but Kurihara and Beidler do not favor it. They scaled the sourness of several acids and the sweetness of those same acids after miracle fruit. The resulting curves were similar so they concluded that "...the mechanism of sweetness induction by acid is closely associated to the mechanism of sourness." However, their scaling procedure required the subjects to make relatively difficult judgments. The data in fig. 4 were collected with a procedure slightly easier for the subjects and contradict Kurihara and Beidler. Miracle fruit appears to have a more pronounced sweetening effect on citric acid than it does on HCl.

The last experiment was designed to test directly the statement of Kurihara and Beidler that equally sour acids are equally sweet after miracle fruit. Subjects were asked to choose one solution from a series of concentrations of an acid that was equal in sourness to .01 M HCl. This was done for seven different acids. Later the subjects were given miracle fruit and asked to choose the sweeter of several pairs of stimuli and state how many times as sweet it seemed. For each subject the pairs contained .01 M HCl and the concentration that he chose to be equally sour. The results clearly show that equally sour acids are not equally sweet. However, Dzendolet's argument could provide an explanation of the differential sweetening that would still be consistent with sourness of an acid directly predicting its sweetness after miracle fruit. Perhaps the acids that sweeten more than HCl are producing sweetness from an additional source. Perhaps the anion provides some of this sweetness. If this were true then removing the sourness of such acids should make them taste sweet. The sourness was removed by adapting the tongue with HCl and then testing with the other acids. Since adapting to HCl makes distilled water alone taste sweet (7) the other acids must taste sweeter than water in order to prove that the anion is producing the sweet taste. The acids were not sweetened more than water. Sourness alone can not be the property of acids that cause sweetness after miracle fruit. The results of these experiments suggest that neither theory available adequately explains the effects of miracle fruit.

For purposes of species comparison an experiment was designed to test the effects of miracle fruit on the chorda tympani taste nerve responses of the hamster. Robert Harvey collected the data as part of his dissertation research in the Natick Laboratories under the direction of the author. The results suggested that miracle fruit does not have sweetening effects in hamsters. This is particularly interesting because the hamster taste system is very

similar to man's in many respects. In particular, Gymnema sylvestre suppresses responses to sugars in the hamster much like it does in man (6) even though it fails to do so in some other species (30,34). The discovery of a species other than man in which miracle fruit is an effective sweetener would be extremely interesting for taste physiology.

Applications

Both Gymnema sylvestre and miracle fruit are of definite usefulness in taste research. First they are useful because they are powerful tools with which to study taste physiology. For example, the failure of Gymnema sylvestre to suppress the taste of sugar in all mammals suggests that the sweet receptor sites are not common across all species contradicting theories of sweet reception based on only one mechanism. In addition, Gymnema sylvestre suppresses diverse sources of sweetness in man suggesting that even if more than one kind of sweet receptor is present in man all these receptors are similar enough to be inactivated temporarily by the same material. An additional research benefit results from the simplification of complex tastes by the removal of sweet through Gymnema sylvestre. For example, a food product like canned fruit provides complex flavor sensations. By removing sweet the other tastes can more easily be evaluated.

Both Gymnema sylvestre and miracle fruit also have many direct applications in the area of palatability enhancement. There are occasions when sugar has desirable properties in the processing of a food but then leaves the final product too sweet. Gymnema sylvestre can be diluted to suppress sweetness by any desired amount. In addition, it can curb the intake of sweets. A snack containing the proper amount of Gymnema sylvestre would make sweets very unpalatable.

The most obvious application of miracle fruit is as a non-caloric sweetener. Since miracle fruit works on the tongue and not the food it avoids the dangers of additives like cyclamates. The recent reports on potential harm from ingesting cyclamates have focused public attention on food additives as a very dangerous source of environmental pollution. Loss of the use of cyclamates poses great problems for the diets of diabetics and weight watchers and also eliminates other benefits that are of special interest where food transportation costs are important as with military uses of food. Artificial sweeteners are usually very light and eliminate the necessity of transporting and storing large amounts of sugar. A new, safe non-caloric sweetener with excellent taste properties obviously has a very large appeal. Miracle fruit appears to be entirely safe even when eaten in quantities appropriate for a fruit (there are no reports of side effects from its consumption as a fruit in Africa and preliminary tests on mice and hamsters were unable to show any toxicity at all). It sweetens without a bitter aftertaste and makes normally sour fruits like lemons, rhubarb, and grapefruit, very palatable. Fruit based products like jams, pies, and ice creams are also sweetened very well with miracle fruit. Foods not normally sour are not affected by miracle fruit and so,

meats, soups, and similar foods are not sweetened. Observers report that the flavor of vegetables is improved with miracle fruit even though vegetables are not normally sour, but this may simply be the result of the addition of a very weak sweet.

The final potential application concerns the flavor problems encountered with some kinds of processed foods and with unconventional foods. The success of processing techniques like dehydration depends to a great extent on the ability to retain flavor when the food is reconstituted. The success of new food sources depends on the ability to give these new foods palatable flavors. Some of the sources for new foods under development - fish flour, algae, vegetable analogues of meat products, food produced by microorganisms like yeast, and protein extractions from green leaves (19,28, 29,31,32) offer enormous nutritional and economic advantages; however, they also pose serious palatability problems. Taste altering substances from plants like Gymnema sylvestre and miracle fruit suggest a new approach to the problems of increasing the acceptance of these nutritionally and economically desirable foods.

References

1. Anon. Gymnema sylvestre, an Indian plant which, when chewed, destroys the power of tasting sugar. Pharm. J., 1847, 7, 351-352.
2. Anon. "Miracle Fruit" presents new sweetness concept. Food Engineering, 1965, 114.
3. Anon. "Miraculous Fruit" makes sour foods taste sweet. Sci. News Letter, 1965, 88, 329.
4. Anon. Strange berry has a flavor secret. Pop. Mech., 1966, 125, 107.
5. Bartoshuk, L. M. Taste. In J. Kling and L. Riggs (Eds.) Woodworth and Schlosberg's Experimental Psychology, New York: Holt, Rhinehart and Winston, 1970, in press.
6. Bartoshuk, L. M. Unpublished research on the responses of the hamster after Gymnema sylvestre.
7. Bartoshuk, L. M. Water taste in man. Perception and Psychophysics, 1968, 3, 69-72.
8. Bartoshuk, L. M., Dateo, G. P., Vandenbelt, D. J., Buttrick, R. D. and Long, L. Effects of Gymnema sylvestre and Synsepalum dulcificum on taste in man. Olfaction and Taste, Vol. III, New York: Rockefeller U. Press, 1969.
9. Bartoshuk, L. M., Gentile, R., Harvey, R. J. and Moskowitz, H. R. Unpublished research on miracle fruit, 1970.
10. Bartoshuk, L. M., McBurney, D. H. and Pfaffmann, C. Taste of sodium chloride solutions after adaptation to sodium chloride: implications for the "water taste". Science, 1964, 143, 967-968.
11. Brouwer, J. N., Vander Wel, H., Francke, A., and Henning, G. J. Miraculin, the sweetness-inducing protein from miracle fruit. Nature, 1968, 220, 373-374.
12. Dalziel, J. M. The Useful Plants of West Tropical Africa. London: Crown Agents for the Colonies, 1948.

13. Daniell, W. F. On the Synsepalum dulcificum, de Cand.; or miraculous berry of western Africa. Pharm. J., 1852, 11, 445-448.
14. Dateo, G. P. Unpublished research, 1969.
15. Dzendolet, E. A structure common to sweet-evoking compounds. Perception and Psychophysics, 1968, 3, 65-68.
16. Dzendolet, E. Theory for the mechanism of action of "miracle fruit." Perception and Psychophysics, 1969, 6, 187-188.
17. Dzendolet, E. and Meiselman, H. L. Gustatory quality changes as a function of solution concentration. Perception and Psychophysics, 1967, 2, 29-33.
18. Fairchild, D. Exploring for Plants. New York: Macmillan, 1931.
19. Halstead, B. W. What we foresee from the sea by '83. Activities Report of The Research and Development Associates, Inc., 1966, 18, 78-84.
20. Hooper, D. An examination of the leaves of Gymnema sylvestre. Nature, 1887, 35, 565-567.
21. Inglett, G. E., Dowling, B., Albrecht, J. J. and Hoglan, F. A. Taste-modifying properties of miracle fruit (Synsepalum dulcificum). J. Agr. Food Chem., 1965, 13, 284-287.
22. Irvine, F. R. Woody Plants of Ghana. London: Oxford U. Press, 1961.
23. Kiesow, F. Schmeckversuche an einzelnen Papillen. Phil. Stud., 1898, 14, 591-615.
24. Kiesow, F. Über die Wirlung des Cocain und der Gymnemasäure auf die Schleimhaut der Zunge und des Mundraums. Phil. Stud., 1894, 9, 510-527.
25. Kurihara, K. and Beidler, L. M. Mechanism of the action of taste-modifying protein. Nature, 1969, 222, 1176-1179.
26. Kurihara, K. and Beidler, L. M. Taste-modifying protein from miracle fruit. Science, 1968, 161, 1241-1243.
27. Kurihara, Yoshie. Antisweet activity of Gymnemic acid A_1 and its derivatives. Life Sci., 1969, 8, 537-543.
28. Matthern, R. O. The potential of algae as a food. Activities Report of The Research and Development Associates, Inc., 1966, 18, 101-109.
29. McPherson, A. T. Microbial forming 1966-1986. Activities Report of the Research and Development Associates, Inc., 1966, 18, 90-100.
30. Oakley, B. Microelectrode analysis of second order gustatory neurons in the albino rat. Ph.D. Thesis, 1962, Brown University.
31. Odell, A. D. Meat analogues from modified vegetable tissues. Activities Report of the Research and Development Associates, Inc., 1966, 18, 85-89.
32. Oke, O. L. Protein supplement from green leaves. Food Tech., 1968, 22, 47.
33. Shore, L. E. A contribution to our knowledge of taste sensations. J. Physiol., 1892, 13, 191-217.

34. Snell, T. C. The response of the squirrel monkey chorda tympani to a variety of taste stimuli. M. Sc. Thesis, 1965, Brown University.
35. Stevens, S. S. On the new psychophysics. Scand. J. Psychol., 1960, 1, 27-35.
36. Stöcklin, W. Chemistry and physiological properties of Gymnemic acid, the antisaccharine principle of the leaves of Gymnema sylvestre. J. Agr. Food Chem., 1969, 17, 704-708.
37. Warren, R. M. and Pfaffmann, C. Suppression of sweet sensitivity by potassium gymnemate. J. App. Physiol., 1959, 14, 40-42.
38. Yackzan, Kamal S. Biological effects of Gymnema sylvestre fractions. Alabama J. Med. Sci., 1966, 3, 1-9.

Fig. 1 Miracle fruit tree approximately 8 years old.

Fig. 2 Miracle fruit. Berries are approximately 3/4 inch long.

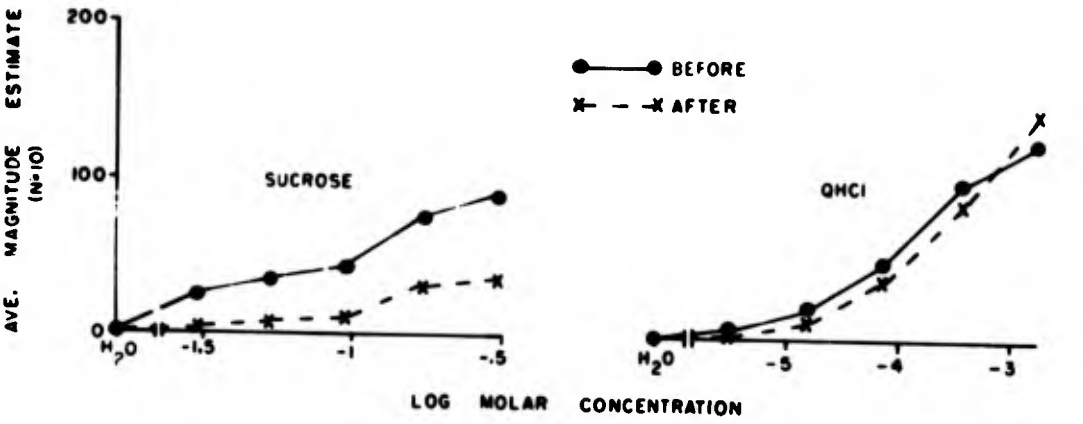


FIG. 3 MAGNITUDE ESTIMATES BEFORE AND AFTER GYMNEMA SYLVESTRE

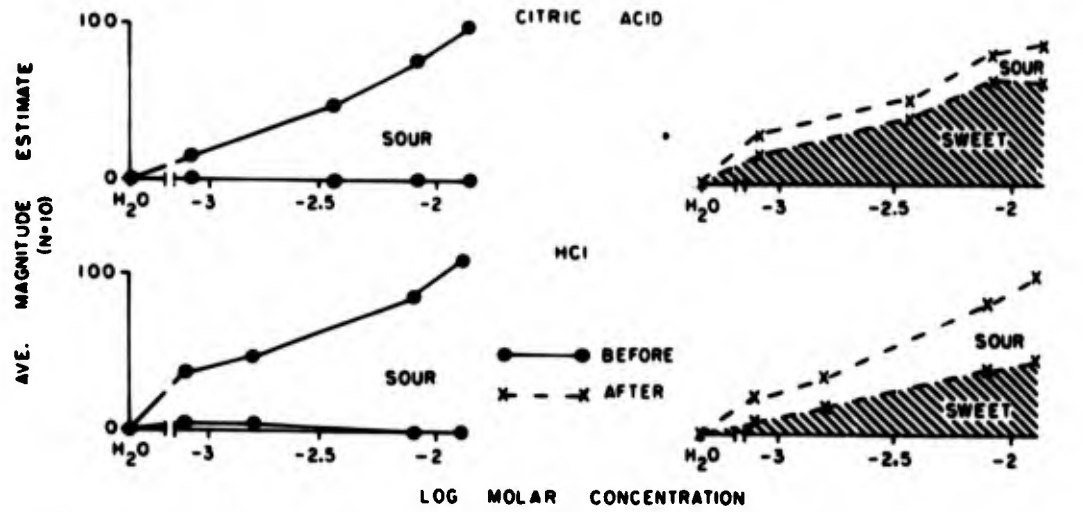


FIG. 4 MAGNITUDE ESTIMATES BEFORE AND AFTER MIRACLE FRUIT

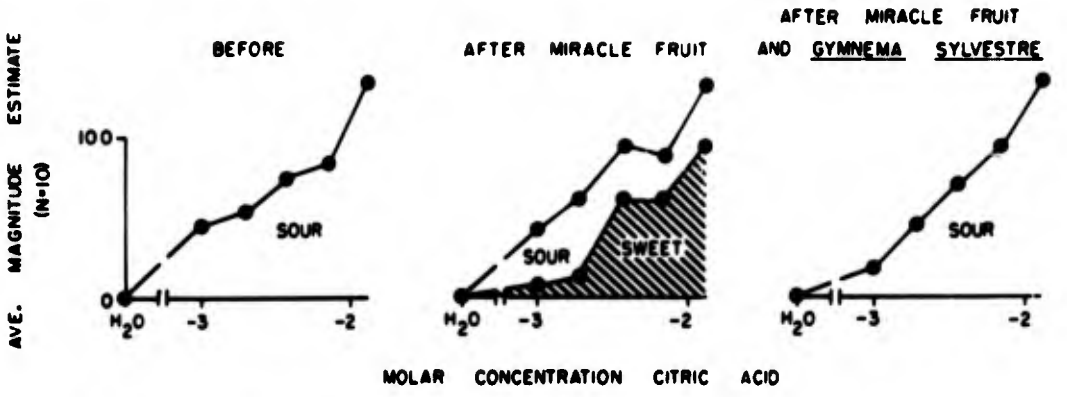


FIG. 5 MAGNITUDE ESTIMATES OF CITRIC ACID BEFORE, AFTER APPLYING MIRACLE FRUIT AND AFTER REMOVING THE SWEETNESS PRODUCED BY THE MIRACLE FRUIT WITH GYMNEMA SYLVESTRE.

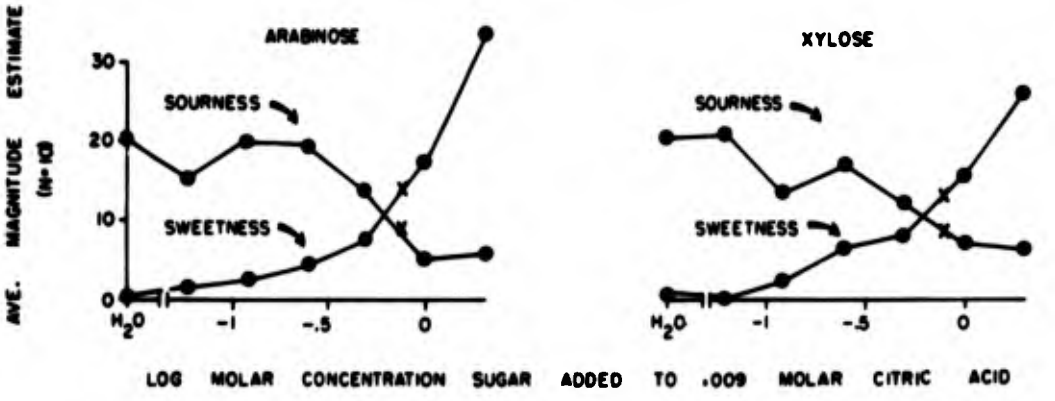


FIG. 6 MAGNITUDE ESTIMATES OF THE SWEETNESS AND SOURNESS OF CITRIC ACID SWEETENED WITH SUGAR COMPARED TO CITRIC ACID SWEETENED WITH MIRACLE FRUIT (X'S SHOW MIRACLE FRUIT VALUES)

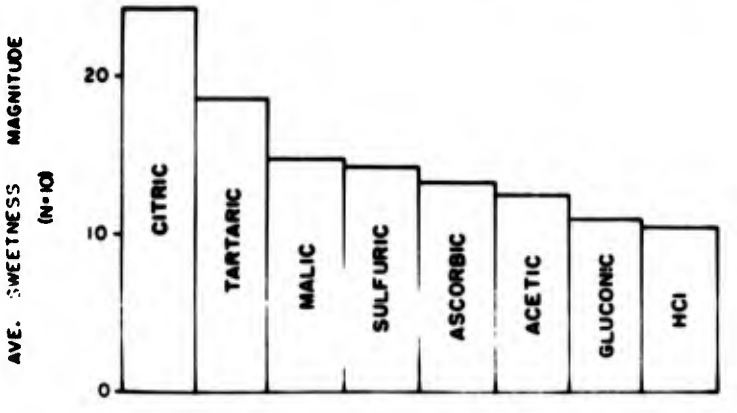


FIG. 7 SWEETNESS OF EQUALLY SOUR ACIDS AFTER MIRACLE FRUIT

AD 713 511

RETINAL DAMAGE BY Q-SWITCHED RUBY LASER

EDWIN S. BEATRICE, MAJ MC, J.O. POWELL, CPT MC,
M. B. LANDERS, MD, and G. H. BRESNICK, MAJ MC
Joint AMRDC-AMC Laser Safety Team
Frankford Arsenal, Philadelphia, Pa.

I. INTRODUCTION

The growing interest of the Army in the tactical employment of laser devices is evidenced by the incorporation of lasers in close to three dozen weapon systems currently in Army research, development and engineering phases. Coincident with the proliferation of laser applications there has been a rapid intensification of efforts to define the health hazards to personnel from laser radiation and an urgency to establish safety guidelines[1]. The Q-switched ruby laser has received considerable attention in this regard due to its effectiveness for tactical applications as well as to its potential for biological damage.

The eye is particularly vulnerable to injury from laser radiation. Lasers, such as the ruby laser (6943Å), which operate in the visible and near-visible portions of the electromagnetic spectrum are especially hazardous to the retina. This is due both to the high transmittance of the ocular media in this wavelength region and to the inherent focusing properties of the eye, which in combination may result in tremendous amplification of the energy density entering the eye. Under the so-called "worst case" conditions where a collimated laser beam is directed through a fully dilated pupil in an eye refracted to infinity, a gain at the retina of approximately 100,000 times the energy density incident at the cornea can, theoretically, be anticipated.

The primary absorption sites for ruby laser radiation entering the eye are considered to be the melanin granules located within the pigment epithelium of the retina and in the melanocytes of the uveal tract. Tissue injury appears to be mainly a thermal effect secondary to the absorption and transformation of light

energy into heat, although other factors such as sonic transients and shock waves may play a pathogenic role especially in Q-switched systems.

Because of the recognition of the ocular hazards of the Q-switched ruby laser relatively early in the history of laser research, a number of studies have been undertaken to determine the threshold energy required to produce retinal damage[2-6]. The design of threshold experiments has varied significantly from investigator to investigator, making the comparison of data from different laboratories a difficult problem. For example, various studies have employed different species and strains of animals which show important differences in the quality of the eye optics and in the anatomical structure of the retina. In addition, some studies have utilized a collimated beam in an effort to achieve a minimal spot size on the retina, whereas other studies have used a divergent beam resulting in a much larger retinal spot size. In the minimal spot experiments the dimensions of the retinal lesion produced at near threshold energy levels are so small as to make detection of the lesions in the living animal very difficult.

In our study of retinal damage by the Q-switched ruby laser, threshold values were determined for both "minimal" and large retinal spot sizes. In addition to using the conventional ophthalmoscopic methods for detecting injury in the living animal, we performed histopathological examinations on a number of irradiated eyes following enucleation. The microscopic evaluation of irradiated tissue was considered to be particularly important in the minimal retinal spot size experiments in order to support the ophthalmoscopic observations of very small retinal lesions.

Rhesus monkeys were exclusively employed in the threshold studies because of the close similarity of Rhesus and human eyes. The eye of the Rhesus monkey has an optical system comparable to that of the human eye and possesses a retina which closely resembles the human retina in terms of its architecture, vascularity and macular specialization.

In analyzing our data we have considered separately the macular and extramacular exposure sites, because of conflicting reports concerning the relative sensitivity of these two retinal areas to laser damage[7,8].

II. MATERIALS AND METHODS

The entire ruby laser delivery system was mounted on an Ealing Tech-Ops 3 meter optical bench (fig. 1). A TRG model 104A ruby laser containing both Pockel's Cell and Eastman 10220 (cryptocyanine in acetonitrile) Q-switches was used. The output beam

diameter of the TEM_{oo} mode laser was 1.7mm with a beam divergence of 1.0 milliradians.

The beam was attenuated by use of 2" x 2" x 1/8" neutral density filters calibrated at 6943Å. A 1" square beam splitter, placed 2.5cm from the center of the animal's cornea, reflected 25% of the energy to the eye. Mounted at the end of the optical bar was a MgO diffusing block with a reflectivity of 91% at 6943Å. An Optics Technology Model 620 ultrafast detector was used to record the energy from the block via a Tektronix Type 556 oscilloscope with Type 1A5 preamplifier coupled through an impedance matching network.

The detectors had been calibrated using an NBS spectral irradiance lamp with an interference filter (multi-layered) with a 140Å half-power bandpass centered at 6950Å. The sensitivity of the detector at 6943Å was 0.096 v/w.

A helium-neon laser was used to align the delivery system and served as an aiming source in exposing the eye of the experimental animals.

Three sets of conditions were employed for ocular exposures. In the first two cases an effort was made to achieve a "minimal" retinal spot size (approximately 40μ diameter) by utilizing beams 3mm and 8mm in diameter with low divergence. In the third case a beam with 30 milliradians divergence was used to produce a large retinal spot (approximately 495μ diameter). 1. Three mm beam. The diameter of the beam was measured as follows: Repeated laser exposures of photographic film (Polaroid type 55 P/N) were made with the beam attenuated by progressively increasing neutral density filters (1.0 ND to 2.0 ND in 0.2 ND increments). The film plane was set at 2.5cm from the beam splitter, coincident with the plane tangent to the center of the animal's cornea. A Leeds-Northrup densitometer was used to measure the opacification produced on the film by the beam. The beam width at 90% of the total energy was measured to be 2.93mm. The retinal spot produced by this beam was estimated to be 40μ by a comparative photographic technique utilizing a helium-neon laser and intra-ocular wires of known dimensions as previously described[9]. 2. Eight mm beam. A Galilean telescope was designed to meet the following specifications: A negative lens 13mm in diameter with a 13mm focal length diverged the incoming beam to a 5.0cm diameter, 62mm focal length plano-convex lens. The divergence of the laser beam at the output of the telescope was 0.7 milliradians, producing a beam at the cornea of 7.9mm diameter (90% energy point measured as described above). Although the retinal spot size produced by this beam has not yet been determined, one would anticipate a spot-size slightly larger than 40μ based on studies by other investigators on the optical

performance of the eye. 3. Divergent beam. In order to create a large retinal spot, a lens of +10.00 diopters was placed 20cm from the cornea of the experimental animal, producing a 3mm diameter beam of 30 milliradians divergence in the plane tangent to the center of the cornea. Theoretical calculations based on diffraction limited theory, assuming a focal length of 1.5cm for the Rhesus monkey eye, predict an estimated retinal spot size of 495μ diameter. Photographic techniques are presently being employed to verify this estimate.

The Rhesus monkeys (Macaca mulatta) used in these experiments weighed between 2 and 5 kilograms. Preanesthetic medication consisted of a sedative dose of phencyclidine hydrochloride (0.25 mg per kg) intramuscular and atropine sulfate (0.2 mg) subcutaneously. Anesthesia was induced with sodium pentobarbital (5 mg per kg) via the saphenous vein. The pupils were dilated with phenylephrine hydrochloride 10% and cyclopentolate hydrochloride 1%. Corneal transparency was maintained by frequent irrigation with physiological saline. The eyes were refracted with a Copeland Streak Retinoscope, and the appropriate correcting lenses were placed in the path of the beam, proximal to the final beam splitter.

The animals were placed in a goniometer type animal holder which gave precise two dimensional rotation around the geometric center of the animal's cornea, allowing 60° displacement in the vertical and horizontal planes. By means of a Zeiss Fundus Camera, the fundus could be viewed immediately before and after laser exposures.

Exposures were placed in a regular grid pattern to facilitate the ophthalmoscopic and histological evaluation for damage. Exposure sites included the macula and extramacular areas within 10° from the optic disc on the nasal and temporal sides.

Each retinal site was examined following exposure at five, ten and fifteen minutes with the fundus camera and at thirty and sixty minutes with a direct ophthalmoscope. The minimal criterion for retinal damage was defined as the presence of an ophthalmoscopically visible burn seen sixty minutes after exposure.

Histopathological examination was performed on a random sample of eyes in the 3mm beam experiment. Both flat preparations of the retina and pigment epithelium as well as serial sections of paraffin embedded tissues were studied by light microscopy as previously described[10]. Typical threshold lesions are depicted in figures 2 and 3.

III. RESULTS

The data from each set of experiments are presented in the form of histograms and retinal damage probability curves. The histograms were constructed by dividing the range of total energy delivered to the eye (microjoules) into equal logarithmic intervals and computing the proportion of damage for each interval (= number of exposures producing damage/total number of exposures within the interval). Smooth curves were fitted to the histogram in order to approximate the probability curves for damage over the range of energies tested. The 50% damage probabilities were then obtained from the curve.

The damage probability curves for the 3mm beam, "minimal" retinal spot are shown in figures 4 through 7. Using ophthalmoscopic criteria for damage, the 50% damage probability points were 18 microjoules for extramacular and 9 microjoules for macular exposure sites. In contrast, when histological criteria were used the 50% points were virtually the same for macular and extramacular sites (9 microjoules and 10 microjoules respectively).

The data for the 8mm beam, "minimal" retinal spot (figures 8 and 9) showed 50% damage points of 26 microjoules for extramacular and 13 microjoules for macular sites using ophthalmoscopic criteria. In both cases the total energy requirements are greater for the 8mm beam than for the corresponding sites in the 3mm beam series.

Data for the divergent beam, 495μ retinal spot (figure 10) showed a 50% damage point of 110 microjoules.

Tables 1 and 2 show the summary of data from the entire series of experiments. All levels represent the 50% damage points obtained from the probability curves expressed in several ways: 1. Measured total energy delivered to the eye. 2. Calculated corneal energy density (total energy/area of corneal exposure). 3. Calculated retinal energy density (total energy/estimated retinal spot area).

IV. CONCLUSIONS

On the basis of the data collected the following conclusions may be drawn:

1. Using ophthalmoscopic criteria the macula would appear to be approximately twice as sensitive as extramacular sites to damage by the Q-switched ruby laser. However, using histological criteria, no difference in sensitivity is noted. This might be explained by the fact that the more darkly pigmented macula affords better contrast for ophthalmoscopic detection of retinal injury

whereas the histological method is less dependent upon such contrast factors. It should be stressed that when one utilizes ophthalmoscopically derived data for the establishment of safety criteria, the portion of the retina exposed in the experiments must be known. If the data pertains to extramacular sites, one would expect at least one-half the energy to be capable of producing ophthalmoscopically visible damage in the macular region.

2. More total intraocular energy is required to produce threshold retinal damage in the "minimal" spot experiments when the energy is delivered in an 8mm, as opposed to a 3mm, diameter beam. Most likely this is due to greater reflectance from the surface of the cornea and poorer performance of the optical system of the eye at the larger beam diameter.

3. The retinal energy density required to produce damage appears to vary as a function of retinal spot size. For larger retinal images (495μ diameter), one requires a lower energy density than for smaller retinal images (40μ). This finding is consistent with results reported separately by other investigators[2,3]. It is beyond the scope of this paper to discuss the possible biological mechanisms underlying this phenomenon. Suffice it to say that previously proposed mathematically derived thermal models of Q-switched laser damage do not explain the data.

Currently available techniques for determining retinal spot size leave much to be desired. Until better methodology is evolved, caution must be employed in any calculations of damage threshold in terms of retinal energy density. At the same time caution must similarly be exercised in attempting to derive extrapolated values for damage thresholds for small retinal spots from experimental data obtained with large retinal spots, in view of the present work and the studies of previously cited investigators.

REFERENCES:

1. Dept. of the Army Technical Bulletin - TB MED 279 - Control of Hazards to Health from Laser Radiation.

2. Ham, W.T., Jr., Williams, R.C., Mueller, H.A. et al: Effects of Laser Radiation on the Mammalian Eye, Trans. N.Y. Acad. of Sci., 28: 517-526, 1966.

3. Vassiliadis, A., Rosan, R.C., Peabody, R.R., Zweng, H.C., and Honey, R.C.: Investigation of Retinal Damage Using a Q-Switched Ruby Laser, Stanford Research Institute Special Technical Report, August, 1966.

4. Zaret, M.M.: Ocular Exposure to Q-Switched Laser Irradiation, Technical Report AFAL-TR-65-279, April, 1966.

5. Kohtiao, A.I., Resnick, I., Newton, J., and Schwell, H.: Threshold Lesions in Rabbit Retinas Exposed to Pulsed Ruby Laser Radiation, Amer. J. Ophth. 62: 664, 1966.

6. Bergquist, T., Kleman, B. and Tengroth, B.: Retinal Lesions Produced by Q-Switched Lasers, Acta Ophth. 44, 1966.

7. Ham, W.T., Jr., Clarke, A.M., Cleary, S.F., Mueller, H.A., Williams, R.C. and Geeraets, W.J.: Biological Applications and Effects of Optical Masers, Report Number 1968-6, Contract No. DA 49-193-MD-2241, April, 1969.

8. Vassiliadis, A., Rosan, R.C. and Zweng, H.C.: Research on Ocular Laser Thresholds, Stanford Research Institute, Final Report, August, 1969.

9. Landers, M.B., Beatrice, E.S., Byer, H.H., Powell, J.O., Chester, J.E. and Frisch, G.D.: Determination of Visible Threshold of Damage in Retina of Rhesus Monkey by Q-Switched Ruby Laser, Frankford Arsenal Memorandum Report M69-26-1, October, 1969.

10. Yanoff, M., Landers, M.B. and Bresnick, G.H.: Flat-preparations of the Retina and Pigment-Epithelium; Annals of Ophth., in press.

Table I. Ophthalmoscopic fifty percent probability values for varying input beams.

Beam Diameter (Cornea) (mm)	Total Intraocular Energy (μj)	Corneal Energy Density (j/cm^2)	Retinal Spot Diameter (μ)	Retinal Energy Density (j/cm^2)	Total No. Exposures
		Extramacular Exposures			
3.0	18	257×10^{-6}	40	1.44	200
8.0	26	54×10^{-6}	Not Estimated		180
3.0 (30mr divergence)	110	1598×10^{-6}	495	0.055	150
		Macular Exposures			
3.0	9	126×10^{-6}	40	0.72	120
8.0	13	27×10^{-6}	Not Estimated		120

Table II. Histologic fifty percent probability values for 3mm (minimal spot) beam.

Total Intraocular Energy (μj)	Corneal Energy Density (j/cm^2)	Retinal Energy Density (j/cm^2)	Total No. Exposures
	Extramacular Exposures		
9.5	135×10^{-6}	0.79	70
	Macular Exposures		
10.0	142×10^{-6}	0.82	67

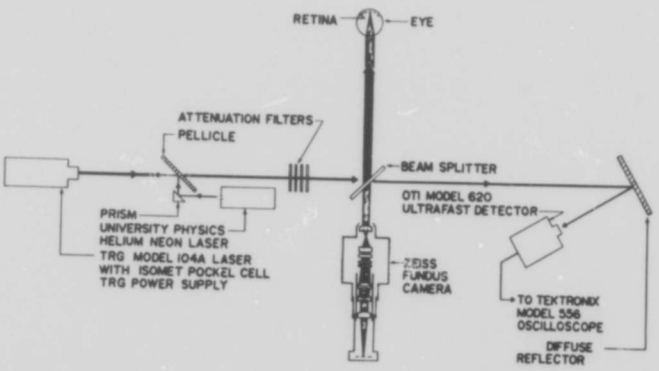


Figure 1. Schematic diagram of ruby laser exposure delivery system.

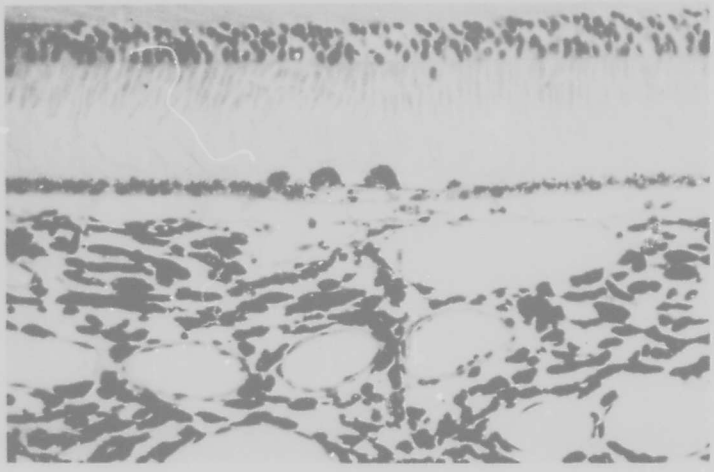


Figure 2. H & E section of three week old Q-switched ruby lesion (10μj). Note disruption of pigment epithelium.

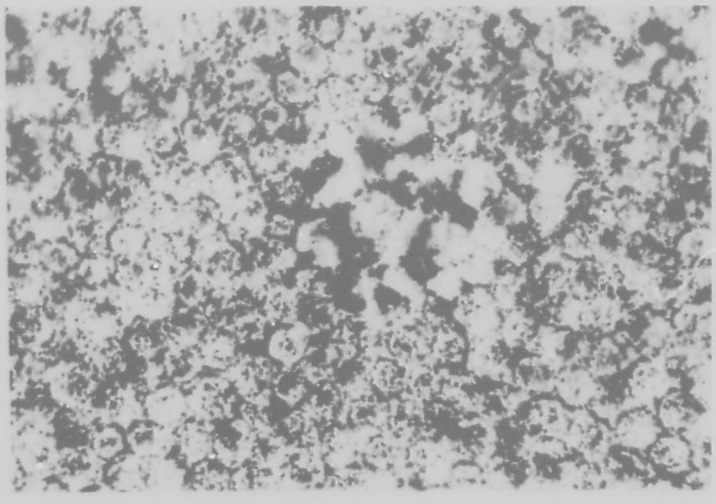


Figure 3. Flat mount of Rhesus monkey retina showing a three week old Q-switched ruby lesion (12μj). Note disruption and dispersion of pigment granules.

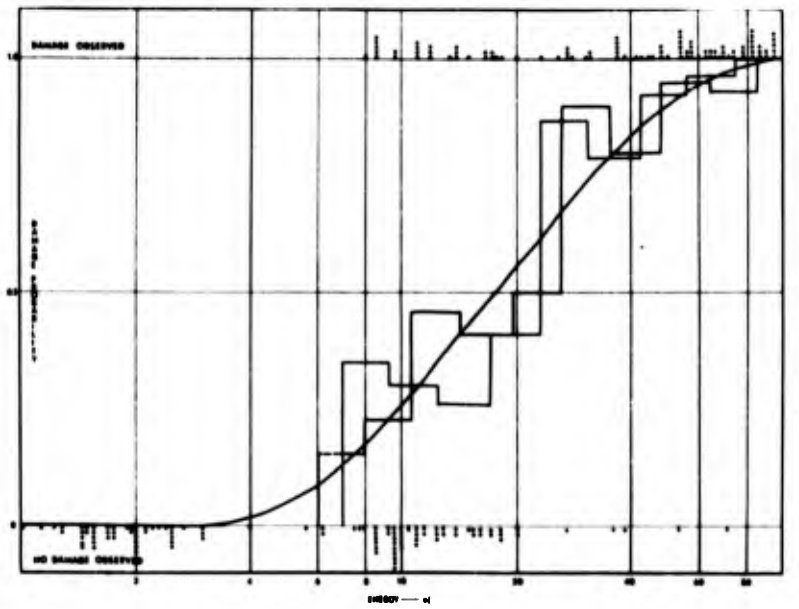


Figure 4. Damage probability for extramacular exposures (temporal to the optic nerve). Three millimeter beam diameter.

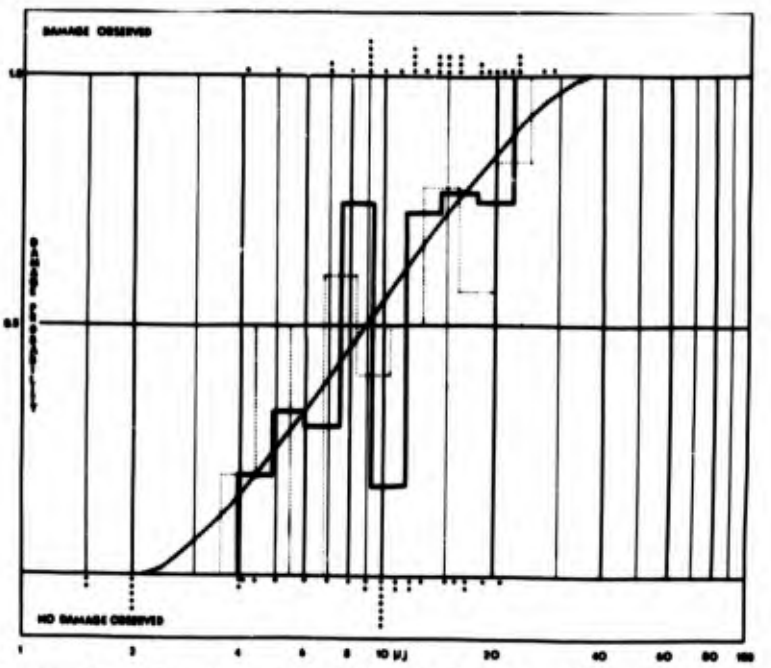


Figure 5. Damage probability for macular exposure. Three millimeter beam diameter.

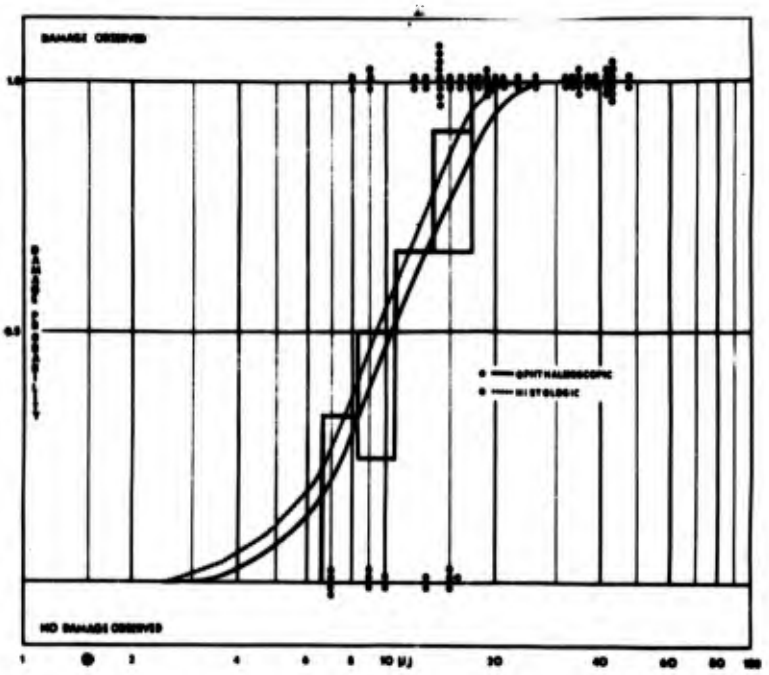


Figure 6. Comparative histograms of damage probability (3mm beam) between ophthalmoscopic and histologic evaluation for identical extramacular exposures.

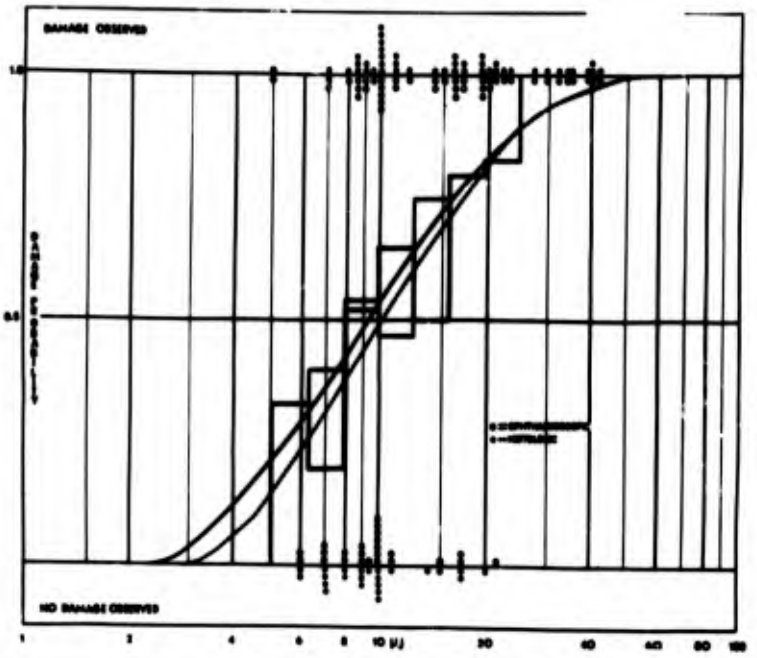


Figure 7. Comparative histograms of damage probability (3mm beam) between ophthalmoscopic and histologic evaluation for identical macular exposures.

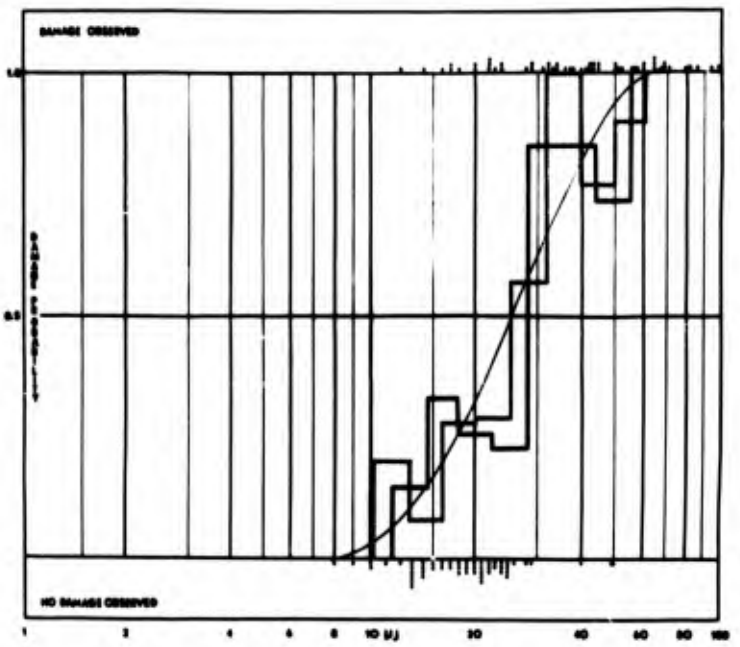


Figure 8. Damage probability for extramacular exposures (temporal to optic disc). Eight millimeter beam diameter.

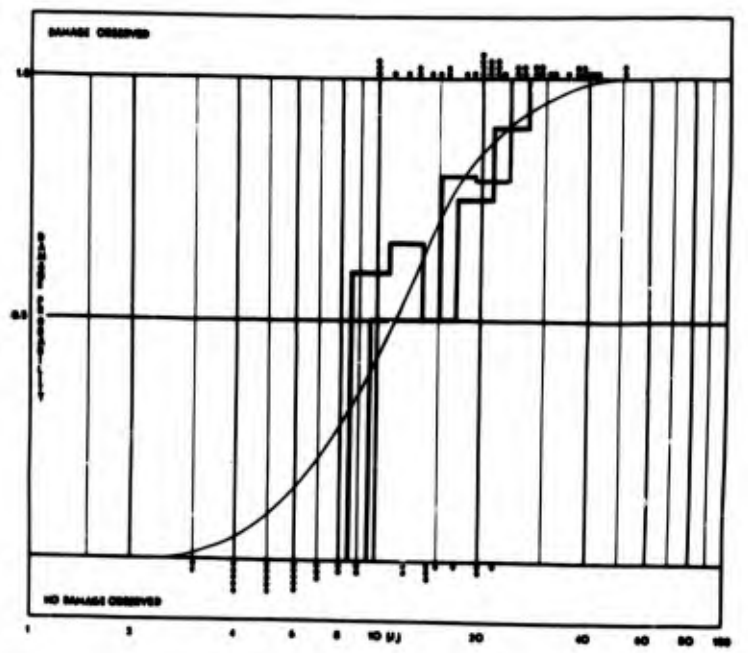


Figure 9. Damage probability for macula exposures. Eight millimeter beam diameter.

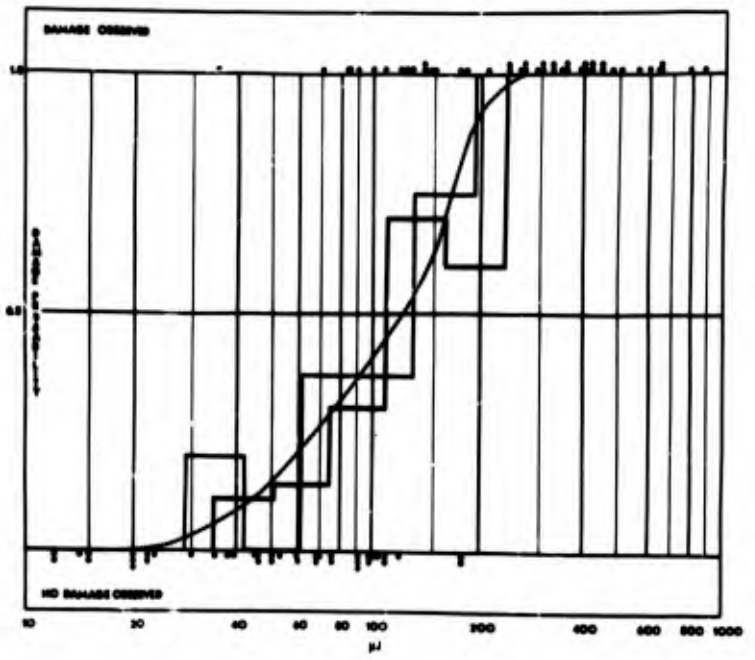


Figure 10. Divergent beam histogram of damage probability for extramacular exposures (temporal to the optic disc).

BLUHM, GORDON and
MORRISSEY*

AD 713 512

EXPLOITATION OF CONTOURED DOUBLE CANTILEVER BEAM SPECIMENS IN CRACK GROWTH AND ARREST STUDIES

JOSEPH I. BLUHM, BENNETT E. GORDON, JR. AND
ROBERT J. MORRISSEY*, ARMY MATERIALS AND MECHANICS
RESEARCH CENTER, WATERTOWN, MASS.

I. INTRODUCTION

General

Fracture toughness characteristics have been studied via a number of specimens developed specifically for that purpose. Those specimens include single edge, center cracked or edge notched plate specimens under tension; they include the "compact tension specimen." Many others have been suggested; in each case a simultaneous observation of the load (at crack instability) and the crack length, permits one to calculate the fracture toughness. Mostovoy et al (1)** introduced a double cantilevered beam specimen contoured such that the compliance varies linearly with crack length, thus yielding an energy release rate G independent of crack length. By this device, one is able to bypass the need of making the difficult crack length measurement to calculate toughness. One need only, then, measure the easily determined load at stability to calculate the fracture toughness. Experience with the constant G specimen shows that a number of instabilities or pop-ins are observed.

Bluhm (2) has suggested that by appropriately modifying the contour of the double cantilever type specimen such that the energy release rate decreases linearly with crack length, it is possible to inhibit the multiple successive pop-ins and thereby to achieve a steady state "tearing" of the specimen. On the other hand, it is anticipated that less extreme contouring will permit the control of crack propagation rate during discreet pop-ins. In this manner, a study of rate effects on toughness might be studied simultaneously with the dynamics of crack propagation.

**Numbers in ( ) refer to the list of references.

In the present study we have concentrated on attempting to develop a specimen contour such that a constant negative value of the slope dG/da, the energy release rate G with respect to crack length a, is achieved independent of crack length. The value of this slope is presumed, for a given material, to control the degree of stability of the crack. Positive values exaggerate crack instability; negative values tend to encourage stable growth. Hence suitable control of this slope, the entire range of specimen behavior from incremental dynamic crack growth to static and stable crack growth can be achieved. This span of control is expected to be effective even for those materials such as ceramics in which one would anticipate catastrophic fracture rather than stable or pop-in growth.

Although the characterization of crack stability is an important attribute of this specimen and the one which this paper treats, other desirable features are noted. Multiple toughness determinations can be obtained in a single specimen. The control of pop-in lengths and the accompanying velocity changes poses further the possibility of using this specimen for study of strain rate effects on toughness and for studying the dynamic characteristics of crack propagation.

Constant Load - Constant Deflection Considerations

In any test in which the rate of the fracturing process is to be controlled, it is obvious that consideration must be given to two extremes of test condition. For example, once the crack starts, at the critical load, the subsequent events depend to a great degree upon the "hardness" or "softness" of the total testing system. In the "hard" system, the fracture tends to propagate under conditions of constant deflection whereas in the "soft" system propagation proceeds under essentially constant load conditions. In practice, neither of these limiting conditions is achieved. A truly soft system would have to be capable of achieving unrealistically high acceleration and the hard machine would have to be an infinitely rigid device. The relative stiffness or compliance of the specimen to the testing machine and the respective mass distribution dictates which of these extreme models is most suited to a particular case.

In the present series of tests, a 10,000 lb. capacity Instron test machine with overall compliance of approximately 2×10^{-6} in/lb was utilized. Data obtained with this testing machine with various specimens having compliance greater than 2×10^{-6} in/lb suggested that crack growth behavior was relatively independent of the specimen compliance and was (at least for the stages of rapid growth) best modeled by a constant deflection constraint.

BLUHM, GORDON and
MORRISSEY*

Model for Crack Growth Behavior

In fracture testing, assuming no fracture mode change during the course of the fracturing process, one encounters generally one of three gross behavioral patterns: static tearing or stable growth, pop-in growth, and finally, catastrophic or "pop-through" growth. The static tearing mode refers to propagation which is enforced by the continued application of either increasing load and/or increasing deformation - inertial effects are insignificant. Pop-in refers to a short burst of crack extension concluded by an arrest. Pop-through is essentially an unarrested pop-in. These latter two modes of propagation involve inertial effects. Only limited studies are available, however, which have taken inertial effects into consideration.

Berry (3) and Hoagland (4) have both noted that unstable dynamic crack growth under constant deflection conditions is not possible unless the material is rate sensitive. Rate sensitivity would manifest itself in a reduced toughness with increasing crack velocity.

The representation of Fig. 1 suggests the rationale for the behavioral pattern mentioned above. The significant features are the relative shapes of the curves governing the strain energy release rate (the energy rate available for crack propagation) G and the resistance to propagation (the energy rate required for propagation) R as a function of crack length a.

Stable or static growth would be expected to occur as long as the energy rate required for fracture is greater than the energy rate available. This condition is schematically shown in Fig. 1a.

Pop-in growth, defined loosely as a dynamic crack growth with subsequent crack arrest prior to specimen failure, is shown in Fig. 1b. A typical dynamic growth which begins at crack length a_o, continues dynamically to length a_1 (due to the energy rate available being greater than the energy rate required), and remains stable beyond length a_1 when the surplus energy has been dissipated and the required energy for continuing propagation is no longer available. The exact shape of the G and R curves depends on both the beam configuration and the material properties.

Catastrophic or pop-through behavior occurs under conditions shown in Fig. 1c. Once a crack starts, a surplus of energy is always available for crack growth; hence dynamic growth continues until the specimen fractures into two parts.

The above rationale presumes a fixed (plane strain) mode of fracture. It is possible that specimens will fracture with shear lips formed, leading to an erroneous and fictitiously high value

of G_{Ic}. Flat fracture surfaces imply plane strain fracture, as opposed to the shear lip formations associated with plane stress fracture. Plain strain fracture is considered to be a material characteristic, whereas plane stress or the mixed mode of fracture wherein both flat and shear fracture are present is dependent on both specimen geometry, particularly thickness, and material properties. We have neglected considerations of rate sensitivity on the R curve. Generally, crack motion leads to a decrease in the R curve as shown in Fig. 1d. The consequences of this are discussed later.

II. SPECIMEN DESIGN

Specimen configurations aimed at providing a constant negative slope dG/da were analytically developed using a strength of materials approach suggested by Mostovoy et al (1). Presuming a configuration as sketched in Fig. 2, the specimen was considered as a pair of joined cantilever beams of length equal to the crack length a and of variable height h.

The compliance C can then be shown to be as determined from the relation

$$C \equiv \frac{\delta}{P} = 2 \left[\frac{12}{Eb} \int_o^a \frac{x^2}{[h(x)]^3}\, dx + \frac{3(1+\nu)}{Eb} \int_o^a \frac{dx}{h(x)} \right] - - - - (1)$$

where δ is the deflection between load points

P is the applied load

E is Young's modulus

b is the nominal specimen thickness

b_n is the reduced specimen thickness along the crack path

x is the coordinate measured as shown

a is the crack length from line of action of applied loads

$h(x)$ is the beam height at x

$h(a)$ is the beam height at the crack tip

ν is Poisson's ratio

Note that the compliance determination considers two contributions to the deflection, that due to bending and that due to shear. The additional contribution due to the concentration effects at the root of the crack is not considered in the present analysis.

Since the energy release rate, G, is calculable from the relation

$$G = \frac{1}{2} \frac{P^2}{b_n} \left(\frac{dC}{da}\right) \text{ - } (2)$$

Eq. (1) can be used to calculate $\frac{dC}{da}$ by merely differentiating Eq. (1) with respect to the upper limit of the integral. This leads to

$$\frac{dC}{da} = 2 \left[\frac{12a^2}{Ebh(a)} + \frac{3(1 + \nu)}{Ebh(a)}\right] \text{ - - - - - - - - - - - - } (3)$$

whence Eq. (2) becomes, after consolidation

$$G = \frac{3P^2}{Eb\ b_n} \left[\frac{4a^2}{[h(a)]^3} + \frac{(1 + \nu)}{h(a)}\right] \text{ - - - - - - - - - - - } (4)$$

Note that G is dependent only on the height of the beam at the crack tip and is otherwise independent of the contour of the specimen.

Since the objective of the specimen design is to achieve a configuration such that dG/da is a constant, it is convenient to differentiate Eq. (4) with respect to a, set it to the desired constant and then, at least in principle, to solve the resulting equation for h(a). However, since differentiation of Eq. (4) leads to a nonlinear ordinary differential equation, an indirect iterative procedure was used as follows: Beginning with a known beam thickness, beam height at a starting point and the corresponding initial crack length, the load is conceptually raised till G, say $G_1 = G_{Ic}$ is reached. At this load level, the deflection is calculated and subsequent incremental crack extension is imagined to occur at this deflection. During this extension of the crack, G falls to the level G_2 (see Fig. 3). The G vs a path during this incremental extension is governed by several criteria, (a) the deflection δ_1 is constant, (b) the slope $\Delta G \Delta a = \frac{G_2 - G_1}{a_2 - a_1}$ is constant and prescribed at the desired value, and (c) the increment of ΔG is selected arbitrarily to be such that G_2 is 80 to 90% of G_1. During this imaginary crack extension under constant deflection, the load drops off. The beam height at point (2) is then calculated and the load is again increased till a $G_3 = G_{Ic}$ is reached; and crack extension is again presumed to take place at the new deflection corresponding to point (3). The cycle is repeated till the height contours are determined for the entire beam. Computations for this iterative procedure were programmed for and carried out by use of a digital computer. The nominal thickness b (see Fig. 2) of all specimens regardless of material was $.375^{+.007}_{-.002}$ inches and a bn/b ratio of .75 was adopted.

Side grooving was incorporated to help insure that the crack would remain in the desired path. The side grooves were machined to a semicircular configuration to minimize stress concentration and to increase visibility of crack tip. This gentle groove is in contrast to the V-groove configuration used by Mostovoy (1). The actual height contours obtained by the iterative procedure mentioned in the last section, varied from material to material for a prescribed value of dG/da.

III. TESTING PROCEDURES

Testing was performed utilizing the equipment shown in Fig. 4. The basic test machine was a 10,000 pound capacity Instron machine equipped with an accessory speed reduction unit that enabled testing at a crosshead velocity of .0002 in/min. Deflection measurements were obtained at the load points by use of a clip gage consisting of two strain gaged beams. Load vs. deflection curves were displayed on an X-Y recorder.

Crack lengths were observed optically at the surface of the specimen using two measured microscopes accurate to $\pm$ .0001 inch. Additionally, crack length measurements were made directly by measurement of X-ray photographs and by compliance measurements.

Some details of the testing procedure varied from material to material. To the degree that these are significant they are described under the discussion specific to the material.

IV. RESULTS AND DISCUSSION

A. Summary of Tests:

Table I lists the various tests conducted by material and shows the total number of specimens tested for each dG/da value. Details are provided in the following sections.

Table I. Summary of Specimens Tested

MATERIAL	dG/da lb/in^2	No. of Specimens	Remarks
Aluminum 7075-T651	-50	2	Pop-in
	-100	6	
	-200	1	
	-100/-50	1	
Graphite, POCO Grade AXF-5Q	-1.0	2	Steady State Tearing ρ = 0.002 inch
	-1.0	1	ρ = 0.020 inch
Plate Glass	-0.065	2	Broke spontaneously
	-0.126	2	Broke spontaneously

B. Aluminum - 7075-T651

As indicated in Table I, ten specimens were tested. All were removed from a single plate of 3/8 inch thick 7075-T651 aluminum. Machining to a finished contour was accomplished in the as-received T651 condition. No further heat treatment was provided. All specimens were oriented so the crack propagation path was parallel to the rolling direction.

All specimens were initially machine notched with a chevron configuration to facilitate the subsequent fatigue crack initiation. Fatigue cracks were generated using an M.T.S. general purpose test system. The load used for generating the fatigue crack was periodically reduced to minimize the ultimate size of the plastic zone. Final fatigue cracks were normal to the side faces of the specimen to within approximately 5 degrees. However, the crack tips did develop thumbnail shapes after successive pop-in crack extensions. A typical pattern of growth is shown in Fig. 5. Based upon extensive measurements using measuring microscopes, X-ray photographs, as well as compliance tests, made both during testing as well as after total specimen fracture, a number of observations are noted.

1. Referring to Fig. 5, in the transition thumbnail growth region, between three and six cycles of pop-in crack growth were generally observed before the final thumbnail front was fully developed.

2. Once developed, the thumbnail front moved in a steady state with size and shape remaining essentially constant during subsequent crack propagation.

3. The size and shape of the thumbnail appears to be independent of the dG/da value.

4. No shear lips were observed indicating that plane strain fracture was the prevailing fracture mode.

As noted in Table I, specimens were prepared with contours corresponding to $dG/da = -50$, -100, and -200 lb/in^2. All exhibited pop-in growth behavior highlighting the fact that suppression of pop-in had not been fully effective within this range of dG/da. The lengths of the incremental growth Δa are shown in Fig. 6 as a function of dG/da. The vertical bar represents $\pm$ one standard direction. In Fig. 7, a least squares linear fit analysis confirms as anticipated that the lengths of the incremental growth are essentially independent of crack length so long as dG/da is kept constant.

The transition thumbnail growth regime of Fig. 5 has not been included in the data shown in Figs. 6 and 7. Only the steady state period data has been used. It is evident that at least for

the aluminum under test, stable growth of cracks would necessitate dG/da value more negative than the -200 lbs/in^2 extremes tried.

It is noted that the scatter in the data of Fig. 6 and Fig. 7 is appreciable. The degree of this scatter is believed to be attributable to the extreme sensitivity of the crack arrest phenomenon to minor changes in the shapes of the R and G curves discussed earlier. In Fig. 8a, it is noted, for example, that if the R curve and the G curve intersect at a small angle, then very slight shifts in either curve can lead to an appreciable shift in the intercept. A small upward shift in the R curve to R' caused, for example, by microstructural variation, voids any possibility of surplus energy and leads to completely stable crack growth (i.e., $\Delta a' = o$) as against a sizable incremental growth corresponding to the original R curve. This sensitivity to slight variation of the R or G curves can be expected to lead to erratic behavior or large scatter. On the other hand, if the angle of intersection is relatively large, i.e., as in Fig. 8b, then the same relative shift of say the R curve to position R' leads to only a minor change in the incremental crack growth Δa to $\Delta a'$. Typical load-deformation curves are shown in Fig. 9. The presence of incremental crack extension or pop-ins is quite evident from the sharp drops in load.

One final specimen was evaluated in which the contour was stepwise prescribed to have a dG/da alternating between -100 and -50 lbs/in^2. This configuration was motivated by the desire to check the feasibility of using such a specimen to start and/or arrest a crack at will and thus facilitate studies of dynamic crack behavior. The resulting specimen contour is shown in Fig. 10. Average values of incremental crack growth in each of the two dG/da regions were consistent with the results of the constant stability specimen having the appropriate dG/da contour, suggesting that control is indeed feasible and raising the hope that this specimen can be exploited for crack dynamics studies. The same degree of scatter in the incremental crack growth was evident here as in the earlier specimens.

The observed pop-in phenomenon generally observed in the 7075-T651 aluminum specimen raises a rather interesting point. This material, which is the classic example of how a well behaved material should behave, insofar as fracture mechanics is concerned, is normally considered to be rate insensitive. If, however, one returns to the energy rate, R, G representation of crack stability, the solid line of Fig. 1d suggests the relative shape of these curves for the present material and specimen configuration. It is noted that the R curve is in fact horizontal with an ordinate $R = G_{Ic}$, the plane strain fracture toughness. With these shapes of the R and G curves it should be evident that crack instability is not possible since the G curve is always below the R curve. How then does one account for the observed pop-ins?

If a slight degree of rate sensitivity does in fact exist, then when the crack starts to extend, the R curve may dip lower than the G curve (taking the position of R' for example) and thence lead to a condition where the observed instability occurs.

It is possible also that this so-called rate sensitivity in these otherwise stable specimens, is rather a manifestation of crack tip blunting. If, in fact, the fatigue crack tends to blunt upon loading, then the load necessary to initiate crack extension is higher than with a sharp crack even though the energy release rates are identified (Bowie)(5). Once started under this higher load, however, the energy release rate is then in excess of that required for extension of a sharp crack, and the crack becomes unstable and a pop-in extension is experienced. This phenomenon has been observed with the graphite specimen described later and serves to emphasize the fact that the Griffith energy balance criteria is only a necessary condition for crack propagation - but it is not a sufficient condition. At this writing, it is not certain which, if either, of these rationale account for the observed behavioral pattern.

C. Graphite

As indicated in Table I, three specimens were tested. Three identically oriented specimens were machined from a single block of POCO graphite, grade AXF-5Q. The specimen contour was designed to yield a dG/da value of -1.0 lbs/in^2 based upon a G_{Ic} of 1.0 in. lbs/in^2. The resulting contour configuration is identical to that of the aluminum specimen having a dG/da value of -100 lbs/in^2. "Crack tips" were machined to a 0.002 inch maximum radius in two of the specimens; the radius of the third had been inadvertently machined to 0.020 in. radius. This was not detected till after the tests were completed.

In both the sharper radius specimens, stable crack growth prevailed. The load after reaching a maximum value, dropped smoothly off as anticipated as the crack extended. It was evident that for this material the dg/da value of -1.0 was ample to assure stable growth. A typical load deformation curve is shown in Fig. 9.

On the other hand, the blunter 0.020 inch radius specimen sustained a higher load for initiation of crack growth but then spontaneous crack growth took place in a sudden burst of unstable incremental growth of 0.330 inches. After this initial pop-in, however, the crack growth proceeded in the same stable manner observed in the other two graphite specimens. The rationale for this seemingly inconsistent behavior has been discussed earlier under the Aluminum section.

To verify the validity of the earlier estimate of the G_{Ic} value, the averaged value of the data from the three specimens was used in conjunction with Eq. (2) to determine G_{Ic}. The results,

shown in Fig. 11, show an average value of G_{Ic} = .889 in. lbs/in^2, a difference from the assumed value of 13%. They furthermore suggest that there is a minor but systematic variation in G_{Ic} through the material.

D. Other Materials

As indicated in Table I, four specimens of plate glass were tested; these included two sets of dG/da values -0.065 and -0.126 lb/in^2. However, all four specimens exhibited catastrophic crack growth indicating that adequate dG/da values had not been prescribed. Extended ranges of dG/da are to be investigated.

V. CONCLUSION

Based upon the present study it appears that

a) design of specimens for a constant negative dG/da is possible and practical.

b) for aluminum 7075-T651, a dG/da = -200 lbs/in^2 is not quite sufficiently negative to assure crack stability or crack arrest.

c) the incremental crack growth is, however, markedly dependent upon the dG/da value tending toward the anticipated stability as dG/da becomes more negative.

d) for graphite, a dG/da = -1 lb/in^2 is more than adequate to provide stable crack growth.

e) for plate glass, a dG/da = -0.126 lb/in^2 is definitely not sufficiently negative to assure stable crack growth.

f) finally, the proposed specimen concept - though requiring considerably more investigation - has potential for use in dynamic crack studies.

REFERENCES

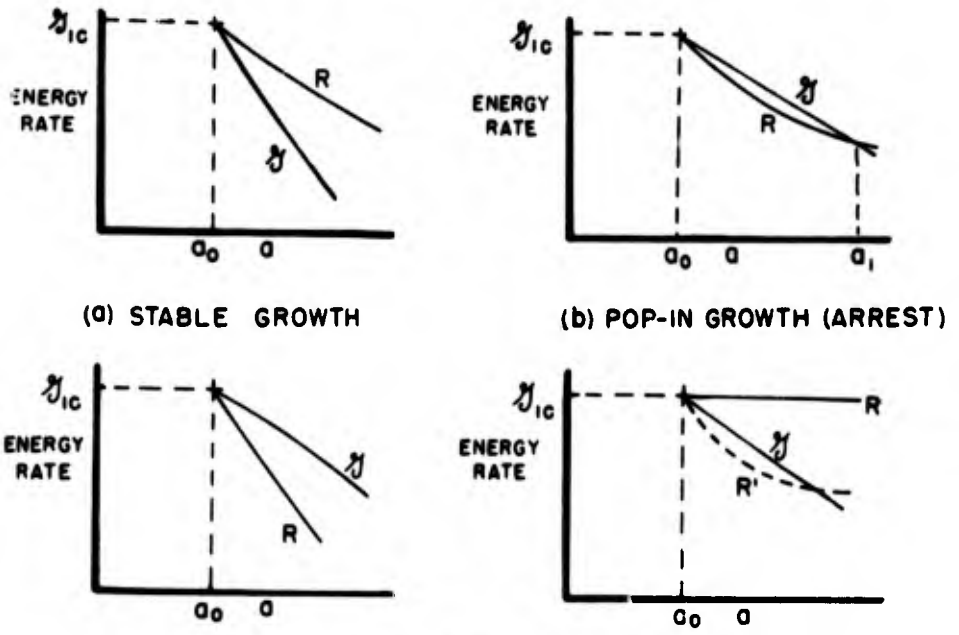
1. S. Mostovoy, P. B. Crosley, and E. J. Ripling, "Use of Crack Line Loaded Specimens for Measuring Plane Strain in Fracture Toughness," Report. Materials Research Laboratory, Inc., Richton Park, Illinois, 1966.

2. J. I. Bluhm, "Fracture Arrest," Fracture, V5, Chapter 1, Academic Press Inc., New York, 1969.

3. J. P. Berry, "Some Kinetic Considerations of the Griffith Criteria for Fracture - I Equations of Motion at Constant Force," J. Mech. Phys. Solids, V8, 1960, pp. 194-206.

4. R. G. Hoagland, "A Double Cantilever Specimen for Determining the Plane Strain Fracture Toughness of Metals," Report BNWL-168. Battelle Pacific Northwest Laboratory, Richland, Washington. 1965.

5. O. L. Bowie and D. M. Neal, "The Effective Crack Length of an Edge Slot in a Semi-Infinite Sheet Under Tension," International Journal of Fracture Mechanics, V3, No. 2, June 1967, pp. 111-119.



(a) STABLE GROWTH (b) POP-IN GROWTH (ARREST)

(c) POP-THRU GROWTH (CATASTROPHIC) (d) RATE SENSITIVITY

FIGURE 1. SCHEMATIC-CRITERIA FOR VARIOUS MODES OF CRACK GROWTH

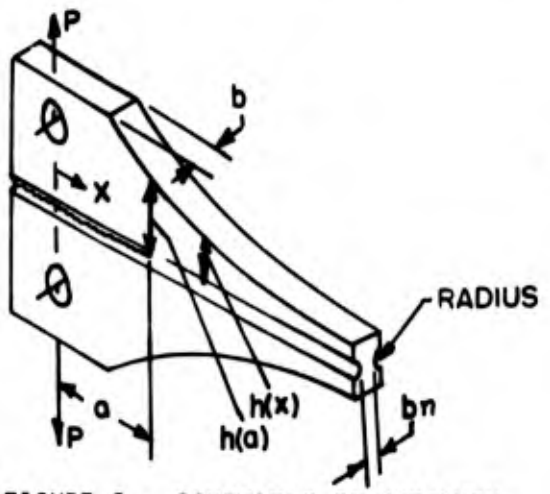


FIGURE 2. SCHEMATIC OF SPECIMEN DESIGN FOR dG/da = NEGATIVE CONST.

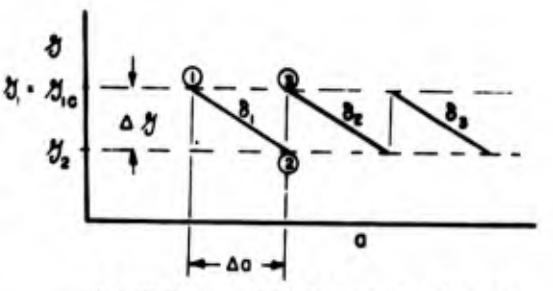


FIGURE 3. REPRESENTATION OF ITERATIVE PROCEDURE FOR BEAM DESIGN

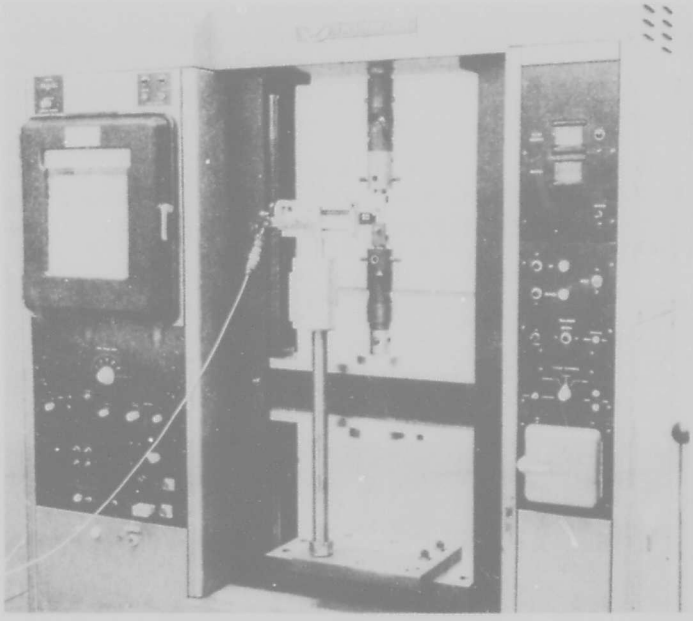


FIGURE 4. TEST SET-UP

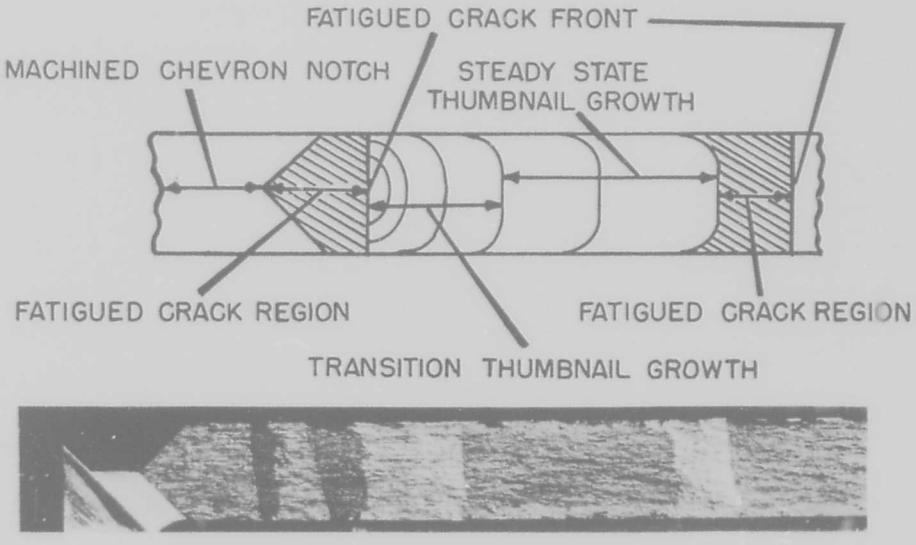


FIGURE 5. TYPICAL FRACTURE SURFACE IN ALUMINUM 7075-T651 SPECIMEN

BLUHM, GORDON and
MORRISSEY*

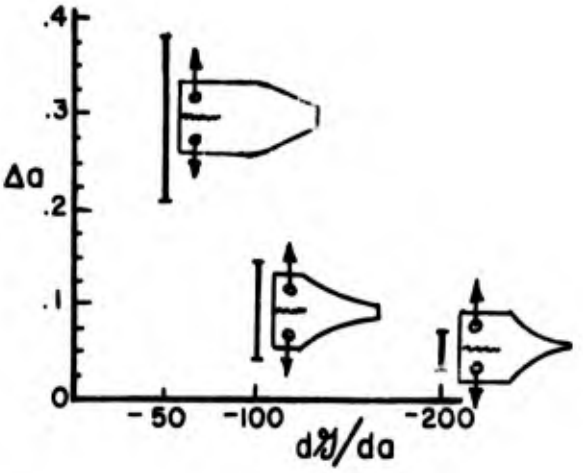


FIGURE 6. INCREMENTAL CRACK GROWTH VS dG/da (7075-T651 AL)

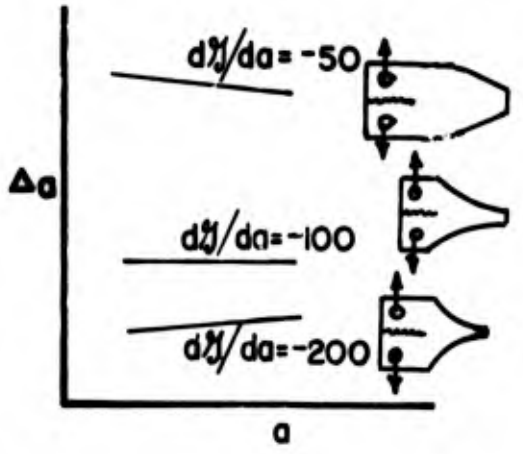
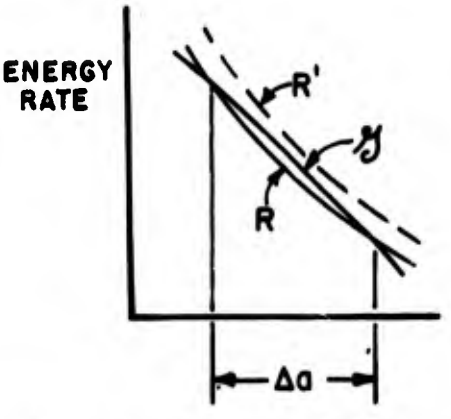
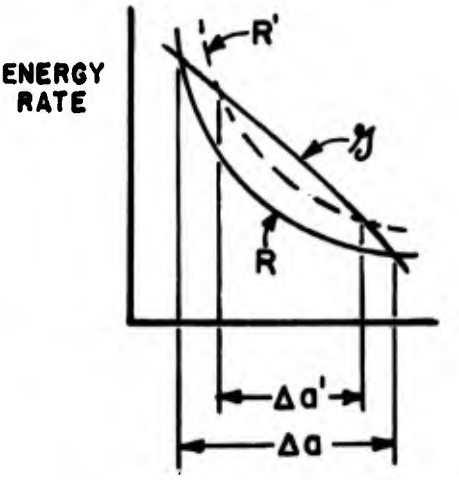


FIGURE 7. INCREMENTAL CRACK GROWTH VS CRACK LENGTH (7075-T651 AL)

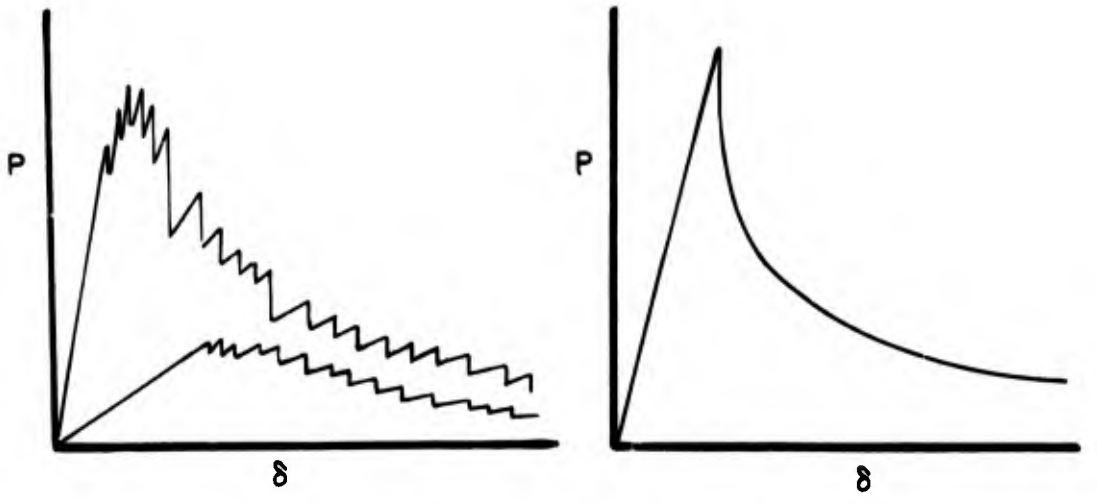


(a) CONDITION FOR INCONSISTENT CRACK EXTENSION



(b) CONDITION FOR CONSISTENT CRACK EXTENSION

FIGURE 8. CONDITION REFLECTING DEGREE OF SCATTER IN POP-IN PHENOMENON



(a) ALUMINUM (b) GRAPHITE

FIGURE 9. REPRESENTATIVE OF LOAD-DEFLECTION CURVE

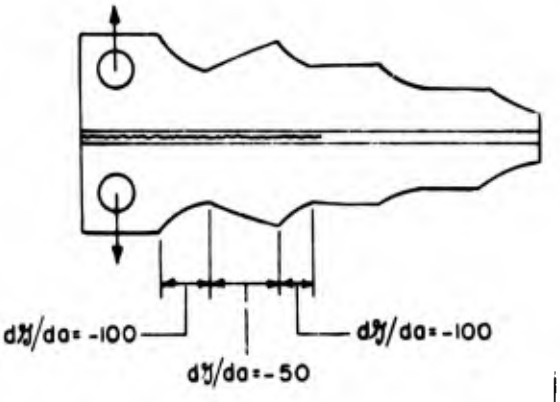


FIGURE 10. STEP-WISE STABLE SPECIMEN CONFIGJRATION (7075-T651 AL)

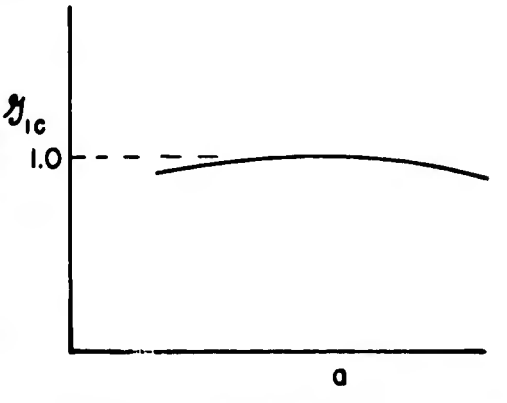


FIGURE 11. VARIATION OF FRACTURE TOUGHNESS ALONG SPECIMEN

AD 713 513

RESPONSE OF SELECTED MATERIALS TO HIGH-SPEED FRAGMENT IMPACT

JERRY W. BROWN
U. S. Army Engineer Waterways Experiment Station
Vicksburg, Mississippi

1. BACKGROUND

This paper was motivated by the Army Aircraft Protective Shelters Program and is one of four papers by Waterways Experiment Station personnel dealing with this subject. During the early phases of the Aircraft Shelters Program it became evident to researchers that sufficient information dealing with the mechanics of fragments and the defeat of fragments by various materials was not readily available. A study of fragment mechanics and the effect of fragments on various materials was conducted to provide designers with facts that could be used in solving the protection problem.

2. OBJECTIVES

The eventual objective of the study of fragment mechanics was to obtain information on the ability of various materials to stop the penetration of fragments from indirect-fire weapons and to define the optimum orientation of these materials whether used singularly or in combination with each other. Before this main objective could be realized, several intermediate goals had to be reached. A logical method of simulating a fragment by some standard projectile had to be selected, and a facility for propelling the projectile under closely controlled conditions had to be constructed. Researchers had to choose, from among a large number of possible protective materials, those few that best met Army needs regarding availability, cost, weight, ease of construction, and effectiveness. A test program had to be conducted and the accumulated data had to be analyzed in order to categorize the best of the available data. Those materials showing promise in the laboratory were selected for full-scale field testing. This paper describes the handling of each of these steps and lists the conclusions drawn from each phase of the work.

3. DESCRIPTION OF PHYSICAL FACILITIES

A fragment-simulation facility was constructed and equipped with several firing devices. A capability exists for firing several different sizes and shapes of fragment-simulating projectiles ranging

in weight from 17 to 305 grains. Most of the data collected and analyzed were obtained from the firing of a 21-grain steel cube measuring 0.218 in. on a side. This cube resembles fragments from several types of mortar rounds in various ways. It has sharp corners and lines and a small sectional density which makes it aerodynamically inefficient (Figure 1). The cube is not spin stabilized, and the 21-grain weight classes it with a wide range of fragments from both domestic and foreign mortar and rocket rounds (Reference 1). Unlike a true fragment, however, it is not hot. This could be of some significance in evaluating its effect on certain textiles such as ballistic nylon.

4. ANALYSIS OF MATERIAL RESPONSE TO FRAGMENTS

4.1 Behavior of Textile Filaments Under High-Speed Impact.

Considerable theoretical work has been done regarding the behavior of textile filaments under high-speed tensile impact (References 3-5). Some of the results of this work are useful in explaining the method by which ballistic nylon defeats fragments and in determining the best amount and orientation of the material.

When a high-speed fragment strikes a nylon filament, the filament responds by moving in the direction of the fragment motion if the fragment velocity is not too high. This motion creates a transverse wave in the filament, and, simultaneously, two tensile strain waves propagate down the filament in opposite directions from the point of impact. The configuration of the filament prior to breaking is shown in Figure 2.

In Figure 2, Point I is the impact point of the fragment and Point A shows the position of the head of the transverse wave. Point C indicates the front of the tensile wave while Point B shows the end of this wave. Point D indicates material that is neither strained nor moving with the transverse wave.

The velocity, U , of the transverse wave front at A is related to the tension, strain, and density of the filament by

$$U = \sqrt{\frac{T}{M(1 + \epsilon)}}$$

where

U = velocity of transverse wave
T = tension in the filament
M = linear density of the unstrained filament
ϵ = strain of the filament

Here U is expressed in Lagrangian rather than fixed coordinates. It is evident from the formula that whenever the local strain at the projectile is large enough to produce rupture of the filament the tension drops to zero, and the transverse wave no longer propagates. Thus, the amount of filament moving in the transverse wave and the amount of energy absorbed to produce the transverse wave are highly dependent on the time at which rupture strain is reached.

In addition to the loss of energy required to produce the transverse wave, the fragment also transfers energy to the filament in creating the tensile strain wave. Because of the interdependence

of the transverse and tensile waves, the rupture of the filament also causes an end to the propagation of the strain wave. Thus, the rate of strain, which is directly proportional to the velocity of the fragment, is very important in determining how much energy is transferred from the projectile to the filament before breaking.

One other important fact should be considered in describing the behavior of the filament during transverse impact. There is a velocity limit on the propagation of the transverse wave. This velocity limit has been termed the critical velocity, and when a filament is struck by a projectile with this velocity, the rate of strain is so high that the local strain becomes sufficient to produce rupture before the transverse or tensile waves are formed. The projectile shears through the filament immediately upon impact and the only energy lost by the projectile is that required for the shearing mechanism.

These ideas allow one to identify three distinct response patterns of a textile to a transverse impact. These response patterns are designated as tensile, transitional, and shear response. The characteristics of each response type are presented herein (Figure 3).

4.1.1 Tensile Response: At low impact velocities (1200-fps range) the local strain around the projectile does not reach the level required for breaking the filament until the entire filament has responded in tension and transverse motion. This is the response area that absorbs the maximum amount of fragment energy. Some textiles, notably nylon, can absorb very large amounts of energy at this rate of strain. The total amount of energy absorbed prior to rupture of the filament depends on the mass of the filament and its specific breaking energy (the area under the tension-strain curve from no strain to rupture strain). These are physical parameters that can be evaluated for various textiles and used in comparing their relative energy absorption characteristics.

4.1.2 Transitional Response: At intermediate velocity levels the transverse wave can form and begin to propagate. Some material is put into tension and part of the filament is set into motion. But the rate of strain is much higher than the rate of propagation of the transverse wave, and breaking strain is reached before the entire filament responds. This response absorbs less energy than the tensile response but, for nylon, the energy absorbed is still quite large.

4.1.3 Shear Response: Whenever the impact velocity is sufficiently high the filament will not begin to transmit the transverse wave before the local strain is sufficient to produce breaking. This velocity is called the critical velocity and, at or above the critical velocity, the filament shears immediately upon impact. No transverse or tensile waves are formed, and the energy absorbed during this penetration is minimal. The energy absorption at this velocity is so low that a textile should not be used to defeat fragments if fragment velocities higher than the critical velocity are anticipated.

4.1.4 Summary: From the above information it is seen that

for a textile material to be effective in defeating fragments it should possess the following characteristics: (a) the critical velocity should be high, (b) the material should stretch for a high percentage of its length before breaking, and (c) the level of energy required to stretch the fiber should be high. Table 1 (taken from Reference 4) shows that nylon possesses a better combination of these characteristics than do other synthetic fabrics.

TABLE 1 CRITICAL VELOCITY, ELONGATION, AND ENERGY FOR VARIOUS PROTECTIVE MATERIALS

Material	Transverse Critical Velocity	Breaking Elongation	Specific Breaking Energy
	fps	%	joules/gram
Acetate	1115	30.7	34.9
Glass fiber	1420	2.6	8.1
Nylon	2240	11.1	38.5
Polyester	1830	8.0	24.3
Rayon	1465	13.1	25.8

4.2 Empirical Data on Ballistic Nylon. A large number of tests were conducted to evaluate the fragment-defeating capability of ballistic nylon. This material is referred to in Army supply channels as "Federal Stock Number 8305-261-85 lb, cloth, ballistic, nylon, basket weave, 13.5 oz minimum, 15 oz maximum wt/sq yd." A 12-ply flak blanket with grommets and exterior weatherproof cover weighs approximately 21 oz/sq ft and the procurement cost is approximately $3.60/sq ft.

Various sample thicknesses and orientations were tested in the fragment-simulation facility. Projectiles were fired at the samples from a distance of 12 ft, and velocities of the projectiles were chronographed in front of and behind the sample. This arrangement allowed a determination of both the velocity needed to penetrate the sample and the velocity loss that the projectile sustained when the striking velocity was high enough to cause penetration.

The results of some of the tests have been used in preparation of Figure 4. All these curves are based on velocity change of the 21-grain cube when impacting the nylon at right angles. Notice in these curves that the nylon shows a decreasing velocity loss, and hence a decreasing loss of momentum, with increased striking velocity. However, the nylon absorbs almost constant energy over a wide range of striking velocities. It is assumed that this is the broad range of maximum tensile response. When the striking velocity is high enough to keep the projectile moving in the material at more than 2300-2400 fps, the effectiveness of the nylon has declined sharply.

In addition to these tests with the projectile impacting loose-hanging material at a 90-deg angle, the material was also

tested at various impact angles, tested while wet, under slight tension, and using separation of the plies to produce air space between the layers of the blanket. No curves are given for these tests as they showed no important changes in the behavior of the nylon. The results are summarized as follows:

a. The loss in velocity that the projectile sustains when penetrating a ballistic nylon blanket decreases if the projectile maintains velocities of above 2200 fps while passing through the blanket.
b. Doubling the thickness of a nylon blanket will not double its effectiveness in stopping fragments.
c. There is no change in the effectiveness of the nylon if it is angled up to 45 deg relative to the path of the projectile.
d. There is no change in the effectiveness of the nylon if it is hanging loose or under slight tension.
e. Wet nylon is as effective as dry.
f. Air gaps between individual or groups of nylon layers do not increase the effectiveness of the blanket.
g. At velocities greater than 2000 fps the projectile will lose as much velocity in 10 ft of air as in passing through four layers of standard nylon.
h. The projectile can be stopped in 32 plies if its striking velocity is near critical. Adding layers beyond 32 plies gives diminishing returns. Test results showed that the mortar fragment that could penetrate 32 plies could generally penetrate 64 plies as well. This indicates that the nylon blanket is effective in the low-velocity regions (below 2200 fps) and adding plies does not increase this effectiveness enough to offset the additional cost and weight.

4.3 Tests on Plywood. Both a 21-grain cube and a 305-grain cylinder were used in studying the response of 3/4-in. fir plywood. The curves in Figure 5 summarize the tests. Note that, unlike ballistic nylon, the response of the plywood seems to be independent of the velocity of the projectile. The velocity loss that the projectile sustains when passing through the plywood is nearly the same over a very broad range of velocities. Also, the effectiveness of the plywood is nearly linear with thickness.

The fact that plywood causes a constant velocity loss regardless of impact velocity while the ballistic nylon loses its effectiveness with increasing impact velocity is the basis for the following suggestion concerning orientation of plywood and nylon. If these are used in combination, the plywood should be placed in front of the nylon. This enables the velocity to be reduced by the wood to the velocity region where the nylon becomes effective. This fact is clearly seen in Figure 6.

4.4 Shots on Sand and Clay. Tests were run on both dry and saturated sand in order to gain some idea of its effectiveness under general outdoor conditions. These sand samples were contained in 1-cu-ft boxes made from 1/2-in. plywood. The sand, either wet or dry,

proved highly resistant to penetration by the 21-grain cube. The curves in Figure 7 illustrate the effectiveness of the sand in stopping fragments, and they also show the tendency of the projectile to reach a maximum depth of penetration at approximately 3000-3500 fps. Increased velocity from this point does not yield increased penetration.

Tests and preliminary mathematical investigations indicate that the response of sand to fragment impact may also be divided into three different areas depending on the fragment velocity. In the range of velocities below 1500 fps the sand tends to absorb the projectile energy by compression. Throughout most of this velocity range the sand can transmit a shock wave faster than the projectile is moving; hence the load is distributed over a large area, depending on the angle of internal friction of the sand. At velocities from 1500 to 3500 fps the fragment seems to truly penetrate the sand rather than compress it. The fragment is moving through the sand faster than the sand can propagate a compressional wave, so there is no major spreading loss. The only resistance encountered by the projectile may be that required to move the grains of sand far enough apart to effect penetration. At these velocities (below 3500 fps) the classical equation of Poncelet-Petry can be used to approximate the depth of penetration of the fragment:

$$D = \frac{w}{a} k \log_{10} \left(1 + \frac{v^2}{215{,}000}\right)$$

where

D = total penetration distance (ft)
w = projectile weight (lb)
a = cross-sectional area (sq in.)
k = constant depending on soil type
v = velocity (fps)

The range of velocities above 3500 fps shows different phenomena. There are indications that a significant amount of heat is created. Also the inertia of the sand at this loading rate is high enough to prevent the movement of the sand, so that it is pulverized and reduced to the fineness of powder. At this rate of loading it is felt that the problem requires consideration of the Rankine-Hugoniot equation of state before a solution is attempted.

The shots into clay showed other interesting tendencies. The impact of the projectile into a clay sample would cause a void in the clay in the shape of a cone with the projectile stopping in the vertex (see Figure 8). There is practically no change in the depth of this cone with a change in striking velocity. However, the volume of the cone increases with increased striking velocity. The energy of the projectile seems to be expended both in penetration and creation of the cavity, and the latter becomes more important as velocity increases. With a suitable choice for the constant, k , the Poncelet-Petry equation may also be used for clay. However, a general mathematical description of the penetration problem for clay, like that for sand, is still unavailable.

REFERENCES

1. Carre, G. L., and Huff, W. L. (1969), Army Aircraft Protective Structures Design, Technical Report N-69-8, Report 1, Vicksburg, Mississippi, U. S. Army Engineer Waterways Experiment Station.

2. Colp, J. L. (1968), Terradynamics: A Study of Projectile Penetration of Natural Earth Materials, Report SC-DR-68-215, Sandia Laboratories.

3. Maheux, C. R. et al. (1957), Dynamics of Body Armor Under High-Speed Impact, Report No. 2141, Army Chemical Center, Maryland, Chemical Warfare Laboratory.

4. Smith, J. C., McCrackin, F. L., and Schiefer, H. F. (1967), The Impact Absorbing Capacity of Textile Yarns, ASTM Bulletin No. 220.

5. American Society for Testing and Materials (1963), Behavior of Filamentous Materials Subjected to High-Speed Tensile Impact, Special Technical Publication No. 336.

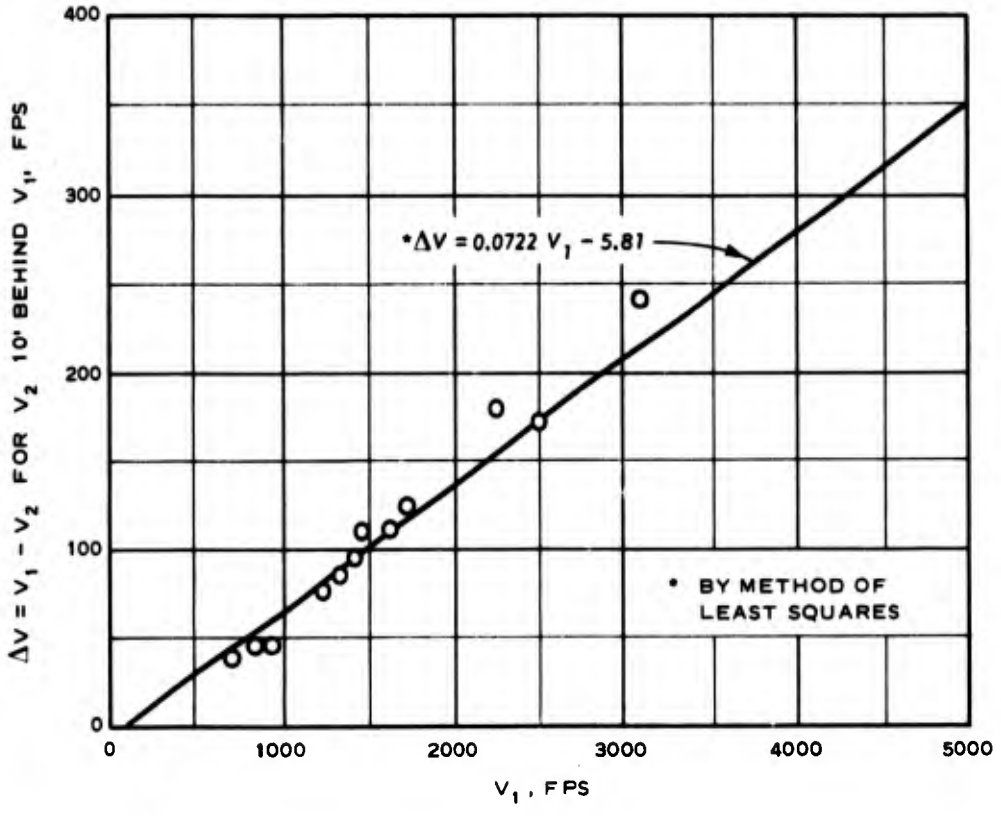


Fig. 1. Velocity loss for 21-grain cube in air

Fig. 2. Filament configuration after impact

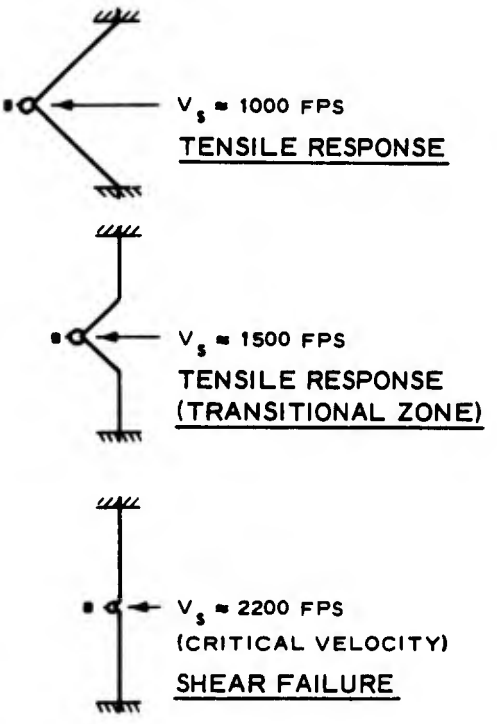


Fig. 3. Failure patterns for nylon filament

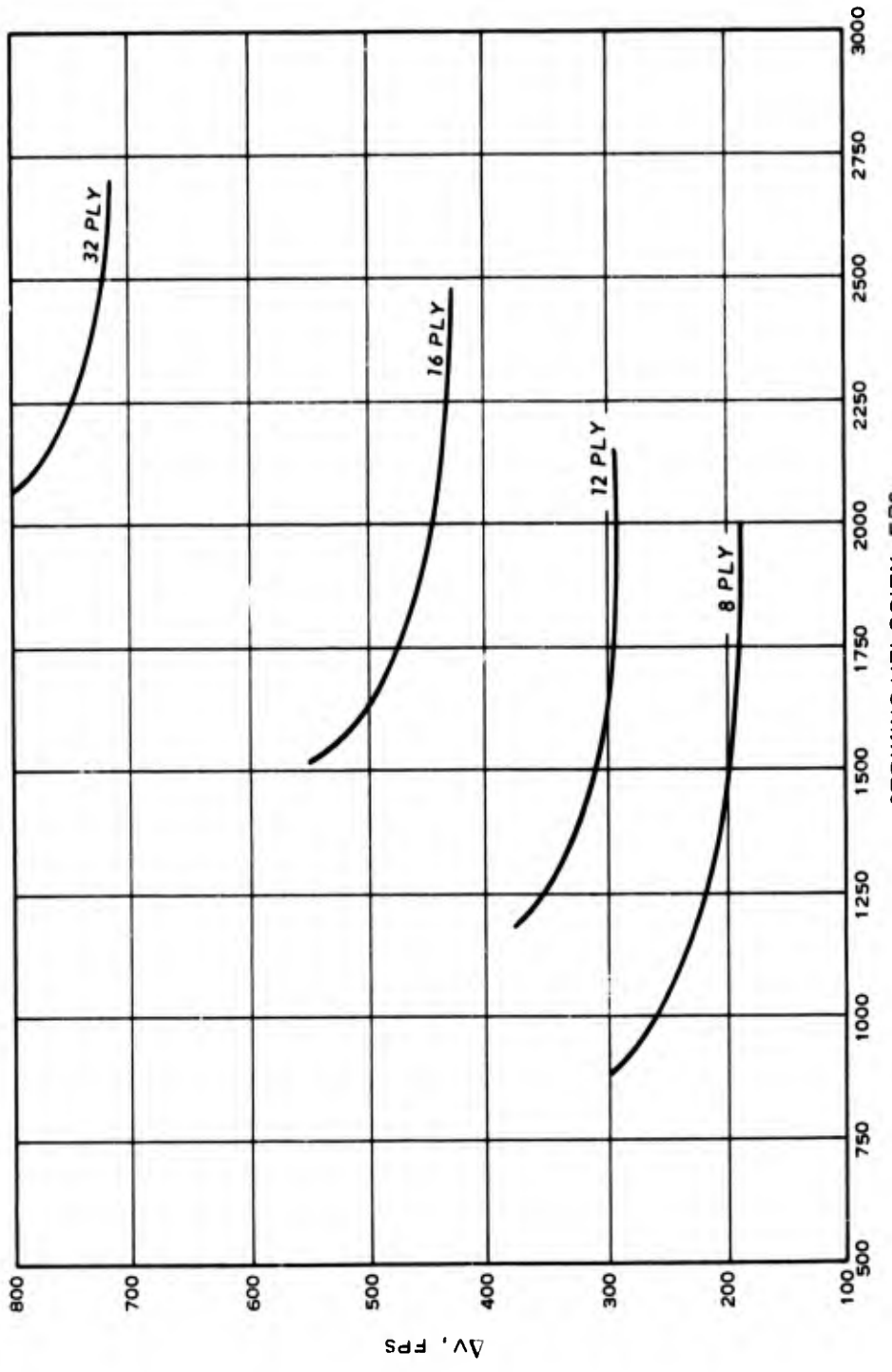


Fig. 4. Velocity loss for cube in nylon

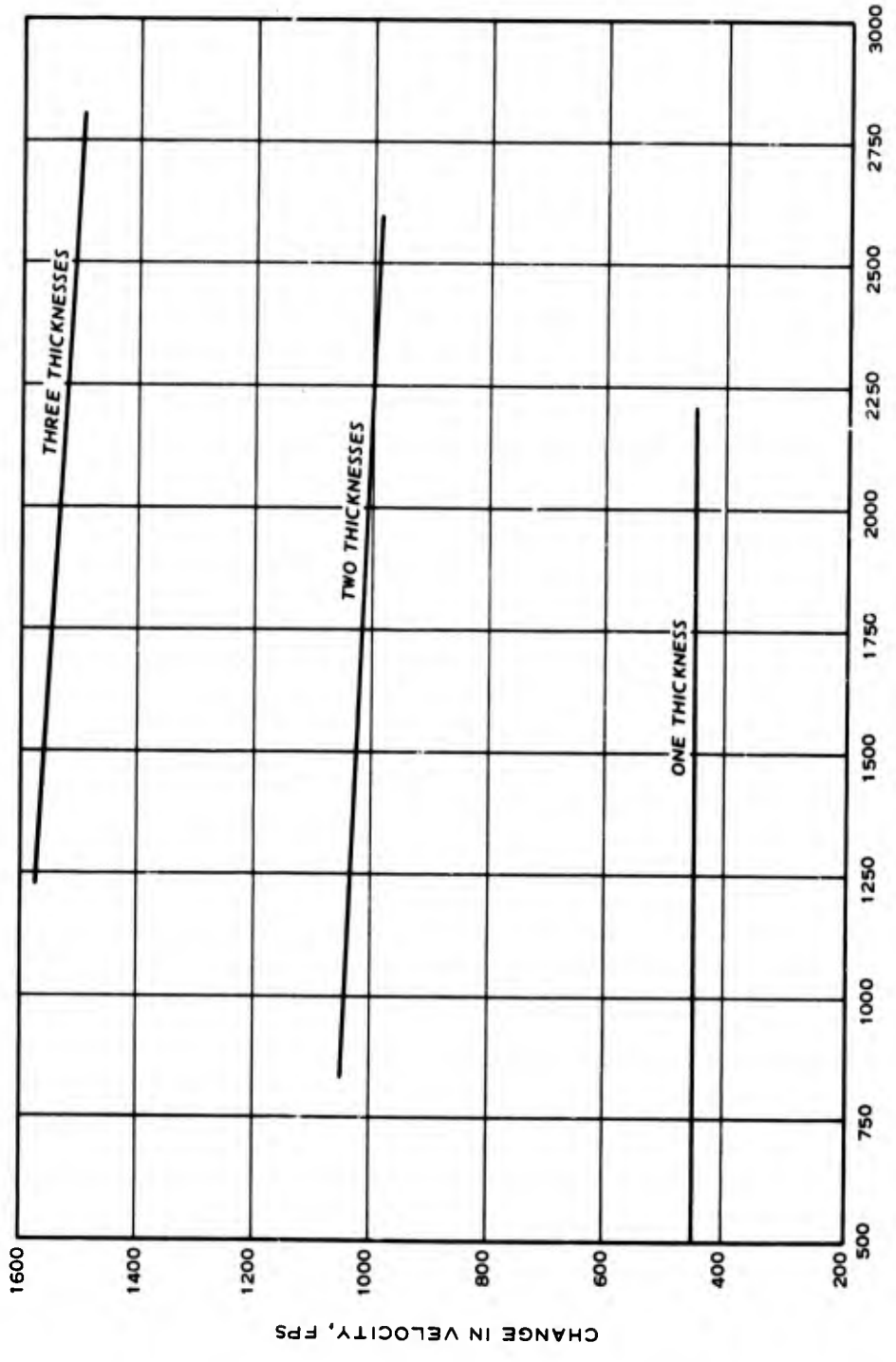


Fig. 5. Velocity loss in 3/4-in. plywood

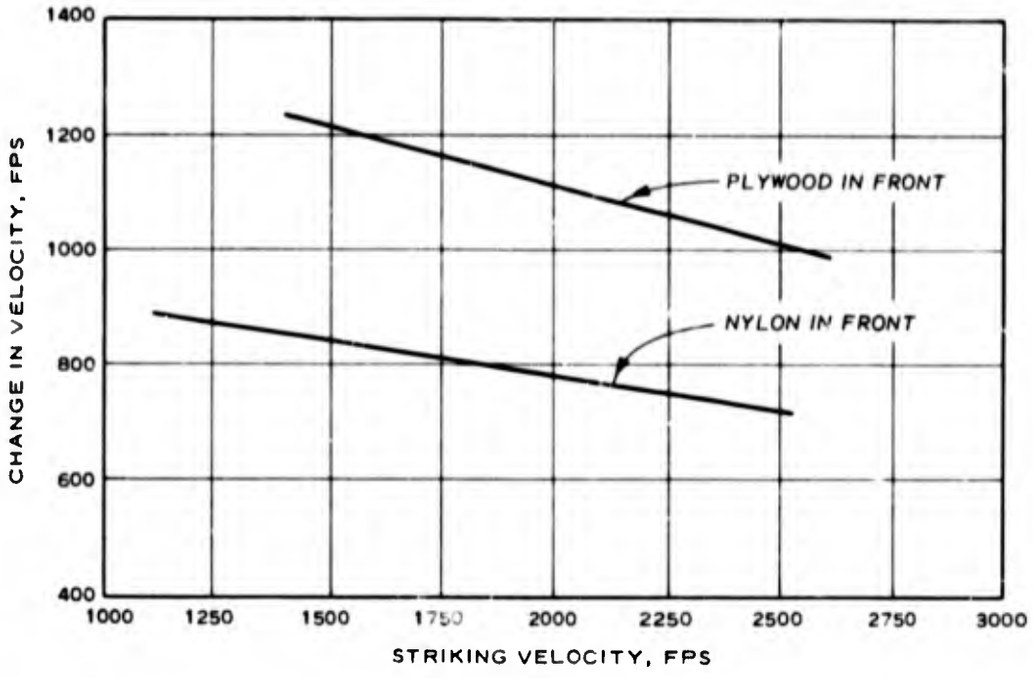


Fig. 6. Velocity loss in nylon/plywood combination

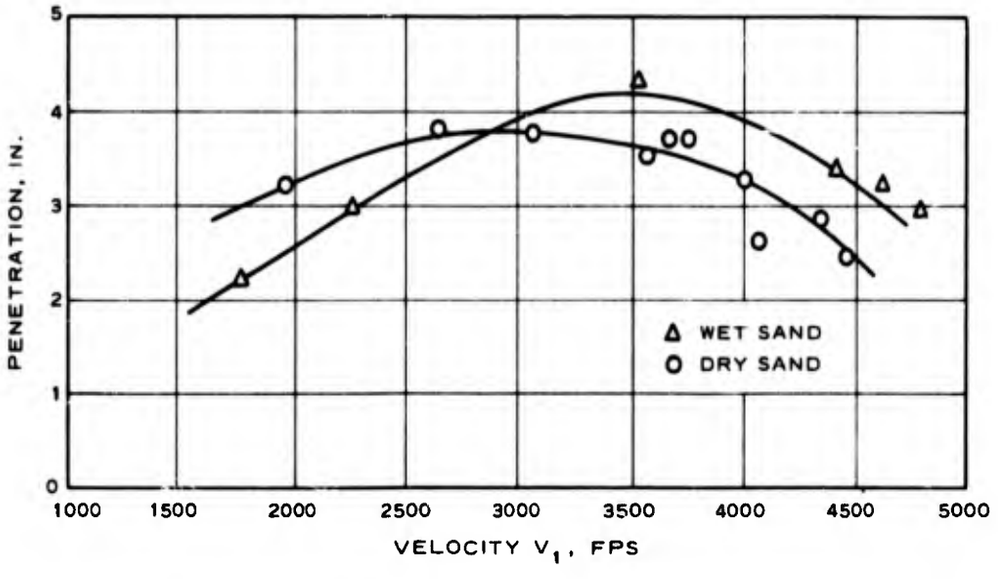


Fig. 7. Penetration of cube in sand

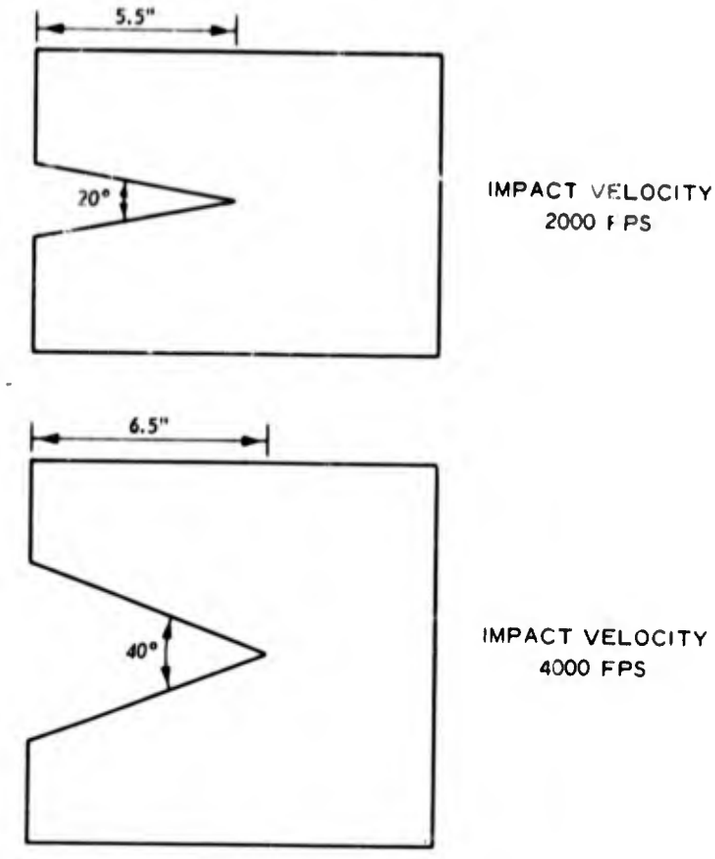


Fig. 8. Behavior of clay upon impact

AD 713 514

A FLUERIC OSCILLATOR FOR MILITARY TIMER APPLICATIONS

CARL J. CAMPAGNUOLO and STACY E. GEHMAN
Harry Diamond Laboratories
Washington, D. C. 20438

1. INTRODUCTION

Mechanical oscillators, such as the pendulum, balance wheel, and tuning fork, have been used extensively for timing applications. The invention of fluid amplification at the Harry Diamond Laboratories has generated interest in determining the feasibility of flueric oscillators as a time base, because of their ability to withstand severe environments. The output frequency of these oscillators, although insensitive to vibration and other mechanical stresses, depends on the air stagnation temperature and pressure. This dependence limits their accuracy as time bases. A flueric RC oscillator whose frequency is only minimally affected by changes in the stagnation conditions has been developed at the Harry Diamond Laboratories. This oscillator is basically a bistable amplifier in which some of the output is returned through an RC network to the controls causing periodic alternation of the power jet (fig. 1). The amount of fluid entering the capacitance is determined by the resistor R_1 and the fluid leaving it by the resistor R_2. Hence, R_1, R_2, and the capacitor volume determine the frequency of the oscillator.

2. THEORETICAL CONSIDERATIONS

An oscillator whose frequency is temperature independent but pressure dependent has been reported by Kirshner and Campagnuolo.[1] This oscillator had an R-L (resistance-inductance) feedback network, which was designed by using transmission line equations.

An oscillator whose frequency is insensitive to both pressure and temperature can be achieved with the lumped RC feedback network shown in figure 2. The criteria for pressure and temperature insensitivity can be established for this circuit from the flow conditions through each element of the feedback network.

The total flow into the network (through R_1) is $Q_1 + Q_2 + Q_b$, where Q_1 is the volume flow that charges the capacitor, and $Q_2 + Q_b$ is the flow through R_2. A bias flow Q_b is a d-c component of flow always present in the network. If the resistances are assumed linear, the pressure drops across R_1 and R_2 are $R_1(Q_1 + Q_2 + Q_b)$ and $R_2(Q_2 + Q_b)$. The pressure drop across R_2 must equal the drop across the capacitor C; hence, the total pressure P_o at the entrance of the network is written as

$$P_o = R_1(Q_1 + Q_2 + Q_b) + \frac{1}{C} \int Q_1 \, dt \tag{1}$$

where

$$\frac{1}{C} \int Q_1 \, dt = R_2(Q_2 + Q_b) \tag{2}$$

Differentiating equation (2) with respect to time and substituting the results in equation (1), the control flow $Q_2 + Q_b$ is obtained in differential form;

$$P_o = R_1 R_2 \, C \, \frac{d}{dt} (Q_2 + Q_b) + (R_1 + R_2)(Q_2 + Q_b) \tag{3}$$

Since the feedback network is driven by a bistable amplifier, the initial conditions for the charging half cycle are

$$P_o \, (0) = P_o \text{ and } Q_2 \, (0) = 0 \tag{4}$$

With this initial condition, the solution of equation (3) can be expressed as

$$Q_2^{+}(t) = \frac{P_o}{R_1 + R_2} (1 - e^{-t/RC}) + Q_b \, e^{-t/RC} - Q_b \tag{5}$$

where

$$R = \frac{R_1 \, R_2}{R_1 + R_2}$$

$(Q_2^{+} + Q_b)$ is the flow through R_2 (the amplifier's control) during the charging cycle. This flow increases until switching occurs. At switching, the capacitor discharges and the decay cycle begins.

The initial conditions for the decay cycle are

$$P_o = 0 \text{ and } Q_2(0) + Q_b = Q_s \tag{6}$$

where Q_s is the total flow through R_2 required to switch the amplifier. Applying the initial conditions (6) in equation (3), the following expression for the flow during the discharge cycle is obtained:

$$Q^{-}_2 = Q_s \, e^{-t/RC} - Q_b \tag{7}$$

Since $(Q_2 + Q_b)$ is the flow through R_2, the pressure drop across R_2 is $R_2 (Q_2 + Q_b)$. This pressure drop, which is equal to the pressure in the capacitor, can be found from equations (5) and (7); it is shown plotted versus time in figure 3a. Figure 3b is an oscilloscope trace of the pressure in the capacitor for one of the experimental oscillators, which verifies the waveform predicted by the equations. If it is assumed that the amplifier is switched by a constant value of difference flow between the two control ports, the flow difference at switching ΔQ_s may be written as

$$\Delta Q_s = (Q^+_2 (\tau_s) + Q_b) - (Q^-_2 (\tau_s) + Q_b) \qquad (8)$$

where τ_s is the time at switching, or the half period, and $Q_s = Q_s + Q_b$. By substituting the expressions for Q_2^+ and Q_2^-, equation (8) becomes

$$\Delta Q_s = \frac{P_o}{R_1 + R_2} - \left(\frac{P_o}{R_1 + R_2} + \Delta Q_s \right) e^{-\tau_s/RC} \qquad (9)$$

Equation (9) can be rewritten as

$$\tau_s = RC \ln \frac{P_o/(R_1 + R_2) + \Delta Q_s}{P_o/(R_1 + R_2) - \Delta Q_s} \qquad (10)$$

The term $P_o/(R_1 + R_2)$ in equation (10) is the maximum flow difference between the two control ports. If the switching flow is assumed small compared with the maximum flow difference, τ_s can be approximated by the following

$$\tau_s = 2R_1 R_2 C \frac{\Delta Q_s}{P_o} \qquad (11)$$

If ΔQ_s is assumed proportional to P_o, then τ_s will be independent of pressure if R_1R_2C does not depend on pressure. In addition, if ΔQ_s is not affected by temperature, τ_s will be temperature dependent only if R_1R_2C is a function of temperature.

For temperature insensitivity, the compensation of R_1R_2C takes place as follows: as the temperature of the fluid in the network rises, the resistances R_1 and R_2 increase because the viscosity of the fluid increases. Similarly, the capacitance C decreases because of the decrease in fluid density. Thus, by an appropriate choice of R_1, R_2, and C, the period of oscillation can be made temperature insensitive. A similar argument is proposed for pressure insensitivity. An exact solution of equation (10) for temperature insensitivity ($\frac{d\tau_s}{dT} = 0$) and pressure insensitivity ($\frac{d\tau_s}{dP} = 0$) can be found in reference 2.

3. DESIGN CONSIDERATIONS FOR THE RC OSCILLATOR

The preceding analysis was intended to explain qualitatively the temperature and pressure insensitivity of an RC oscillator. For the experimental analysis, the oscillator shown in figure 4 was designed so that its outputs were connected to the control ports of a bistable amplifier with bleeds. This amplifier isolates the oscillator from any downstream load. Since the oscillator and amplifier had the same supply pressure, the amplifier entrains flow from the oscillator output, thus changing the impedance at the oscillator output. This variable impedance was not considered in the theoretical analysis, but it contributes significantly to the pressure insensitivity since it determines the flow that enters the feedback loop.

4. EXPERIMENTAL RESULTS

For the experimental analysis, three geometrically similar oscillators were constructed, each having a different resistance R_2 (fig. 1) in its feedback loop. The capacitor volume (0.366 cu in.) and the size of the resistor R_1 (0.046 in. deep, 0.034 in. wide) were the same for all three.

Resistors R_2 had the same depth (0.046 in.) and length but their widths were 0.010, 0.015, and 0.018 in. The frequency was measured by a piezoelectric crystal installed at one of the two outputs of the amplifier, fully blocking this channel. This generated a distortion in the output signal, so a load was provided for the opposite port and adjusted to eliminate the distortion. Plots of the frequency as a function of input pressure are shown in figure 5. It appears that the element with a resistor of 0.018-in. width (top curve) was least sensitive to pressure in the range from 6 to 30 psig. The change in frequency in this range was, at most, 0.6 percent. The remaining two elements had a greater frequency variation over the same pressure range (1.4 percent for the 0.015-in. width and 4.3 percent for the 0.010-in. width resistors).

The three oscillators were tested for temperature insensitivity by supplying them with heated air. A copper coil approximately 10 ft long was connected to a brass tank and placed in an oven. The tank was used to measure input stagnation pressure. The oscillators were placed outside the oven and were carefully wrapped in asbestos to minimize heat loss to the surroundings. A copper-constantan thermocouple was placed between the stagnation tank and the oscillator output. The frequency output was monitored by the piezoelectric transducer (as in the previous test). The oven temperature was allowed to rise to its maximum, and when equilibrium was reached between the oven temperature and the air stagnation temperature, the frequency output of the oscillator was measured. As the oven was cooling, the frequency was monitored continuously. The results obtained using an input pressure of 10 psig are shown in figure 6. The curves shown throughout the report at 10 psig are typical of results obtained at pressures of 3, 10,

20, and 30 psig. The oscillator with the 0.010-in. wide resistor showed a frequency change of 4 percent over a temperature range from 70° to 200° F. For the same temperature range the frequency changed 9 percent for the oscillator with R_2 0.015 in. wide, and 13 percent for the one with R_2 0.018 in. wide.

The curves described thus far have shown that the oscillator with the 0.018-in. width resistor was least pressure sensitive, while the one with a 0.010-in. resistor was least temperature sensitive. Thus, none of the three oscillators was simultaneously pressure and temperature insensitive. To obtain another variable parameter, two adjusting screws were placed in feedback resistor R_1 to control the amount of fluid entering the feedback loop. The screws were placed in the oscillator having an R_2 width of 0.010 in. This oscillator was chosen because it displayed the least temperature sensitivity. The screws were adjusted until the oscillator frequency was least sensitive to changes of input pressure. Figure 7 shows that the frequency output of the oscillator was flat from 8 to 30 psig. The unit was also found to display a high degree of temperature insensitivity. Figure 8 shows the effect of temperature variation on frequency at 10 psig. The maximum change in frequency was 2.2 percent when the temperature was varied from 80° to 200° F.

The restriction caused by the screws was measured and a modified oscillator was built having a similar geometrical configuration but with R_1 only 0.032 in. wide instead of 0.034 in. Figure 9 indicates that for this new unit, the frequency varied less than $\pm$ 1 percent for a pressure input from 6 to 30 psig. The test for temperature insensitivity showed that the frequency varied less than 1 percent from 77° to 175° F, for an input pressure of 10 psig (fig. 10). This oscillator was then tested at low temperatures. The results were invalid because leaks developed in the oscillator. To overcome this difficulty, a new oscillator was constructed with the same dimensions as the modified one. The brass plates on which the oscillator was machined were gold-plated and then diffusion bonded to eliminate the leakage problem. The pressure characteristics for this oscillator are shown in figure 11. For pressures from 6 to 30 psig, the frequency varied less than $\pm$ 1 percent. For the low temperature tests, air dried at -70° F was used to prevent ice formation in the oscillator channels. During the test, the oscillator was wrapped in asbestos to minimize heat transfer with the environment. The air supplied to the oscillator was cooled by allowing it to flow through copper coils located inside an environmental chamber. The air temperature was measured with a copper-constantan thermocouple at the oscillator input. The measured frequency change at a pressure input of 10 psig did not exceed 1.5 percent for temperatures from 70° to -58° F (fig. 12). The diffusion-bonded oscillator was also tested at higher temperatures. Figure 13 indicates that the frequency varied less than 1 percent for temperatures between 75° and 175° F at an input pressure of 10 psig.

5. CONCLUSIONS

A flueric RC oscillator that is capable of operating over a wide range of stagnation conditions has been described. This oscillator, developed primarily by experimental methods, demonstrated the feasibility of using flueric oscillators as time bases. To optimize pressure and temperature insensitivity, three geometrically similar oscillators with different resistors in the feedback network were constructed and tested. A modification of one of these oscillators had a frequency change of less than $\pm$ 1 percent for pressures ranging from 6 to 30 psig and temperatures from -56° to 175° F.

Further work is being done at HDL to miniaturize this oscillator and to incorporate it with a flueric counter. This combination will result in an accurate, reliable timer capable of satisfying many military needs.

REFERENCES

1. Kirshner, J. M., and C. J. Campagnuolo, "A Temperature-Insensitive Pneumatic Oscillator and a Pressure-Controlled Pneumatic Oscillator," Proceedings of the Fluid Amplification Symposium, Vol. II, Oct. 65, pp. 5-19.

2. Campagnuolo, C. J. and S. E. Gehman, "Flueric Pressure and Temperature-Insensitive Oscillator for Timer Application," Harry Diamond Laboratories, TR-1381, Feb. 68.

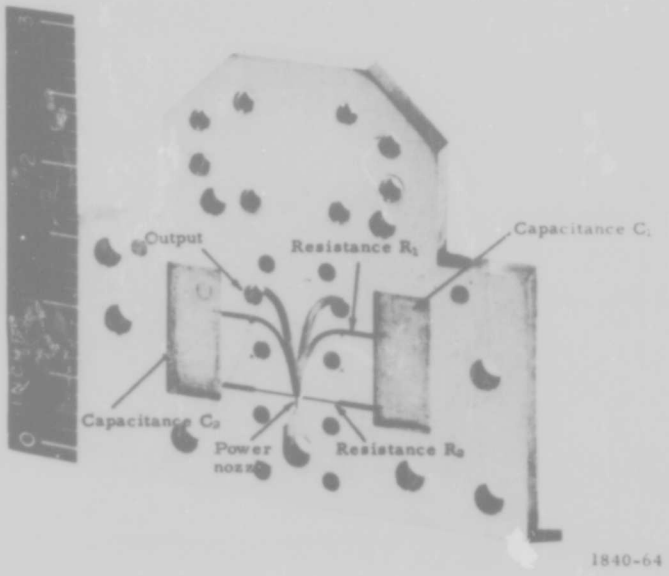


Figure 1. RC feedback oscillator

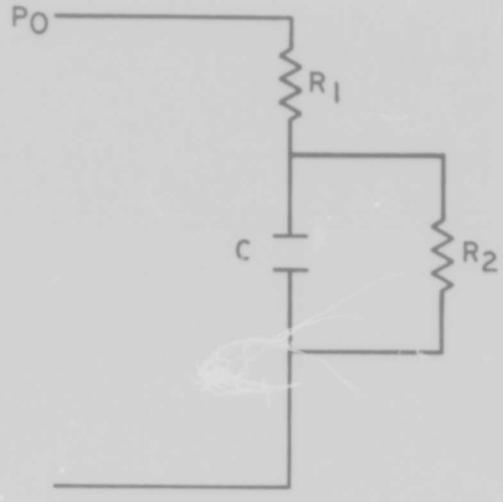
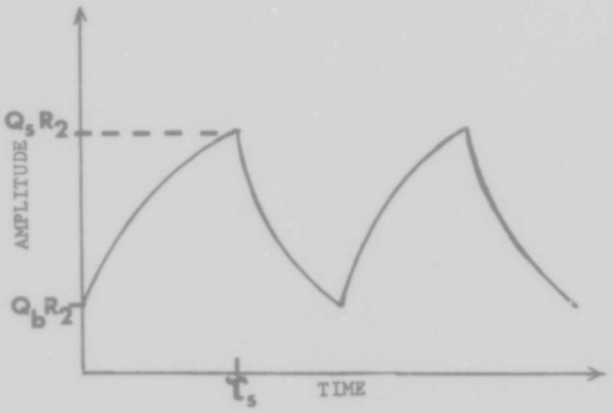
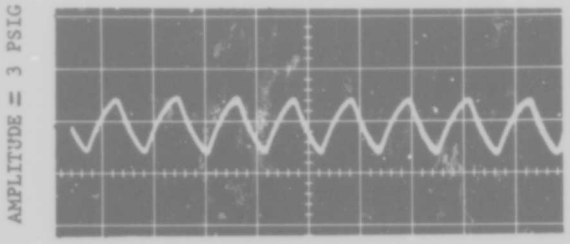


Figure 2. RC feedback network



A. Theoretical pressure response



PIN = 30 PSIG FREQ. = 97 Hz

B. Experimental pressure response

Figure 3. Pressure versus time in the capacitor of the RC oscillator.

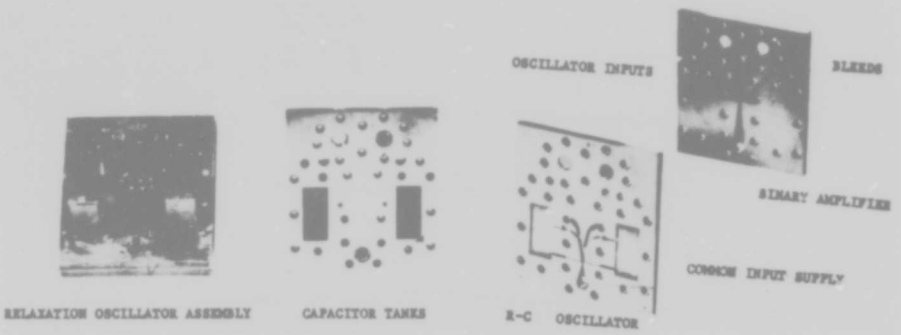


Figure 4. RC oscillator with bistable amplifier

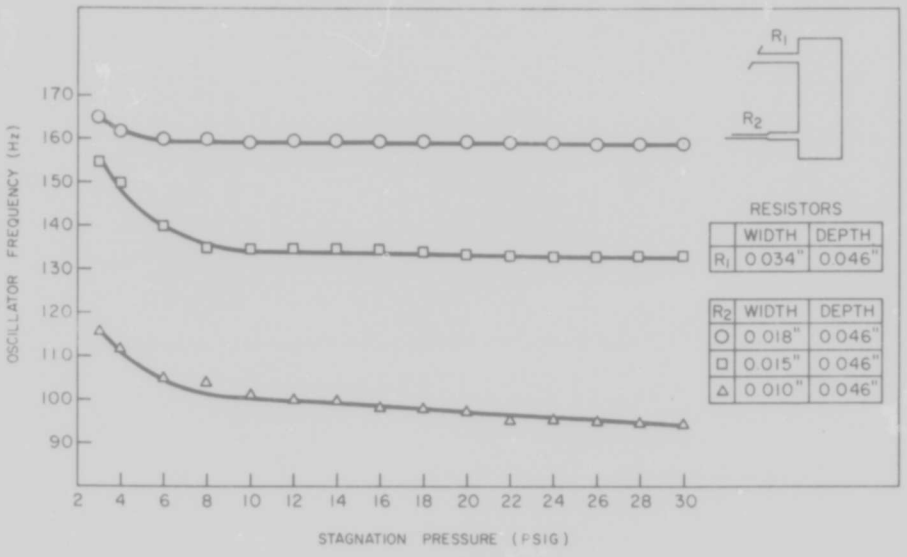


Figure 5. Oscillator frequency versus stagnation pressure

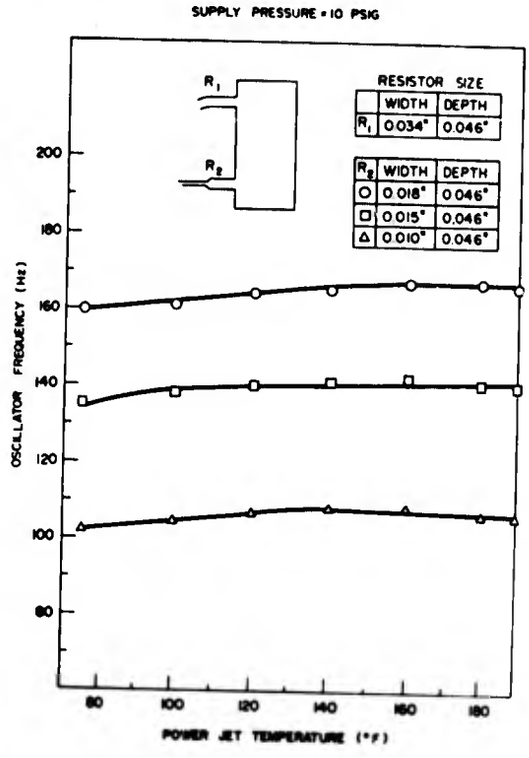


Figure 6. Oscillator frequency versus stagnation temperature

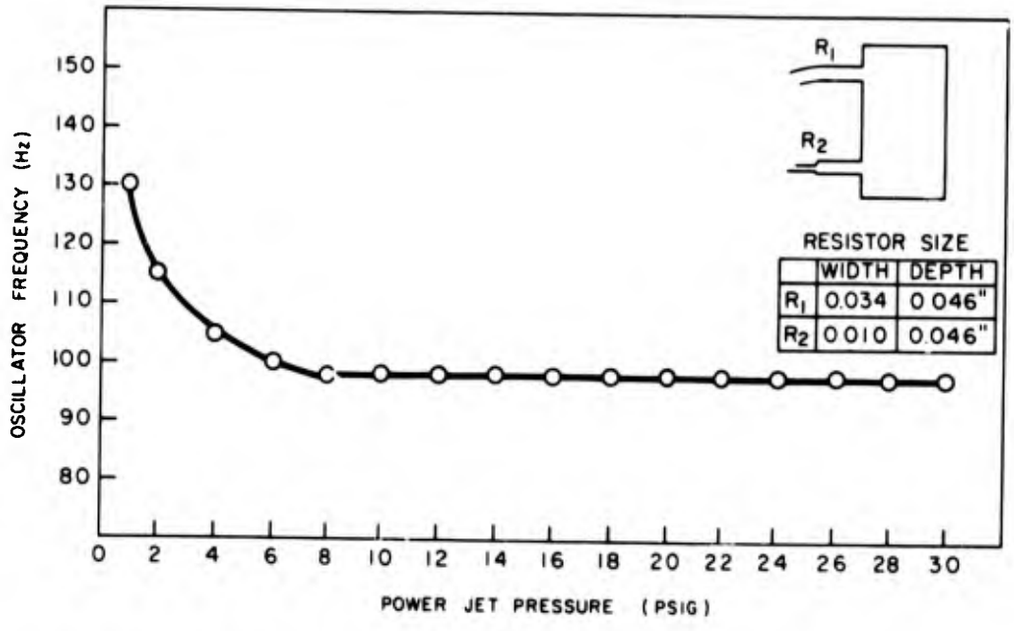


Figure 7. Frequency versus stagnation pressure for the oscillator with a variable resistor R_1

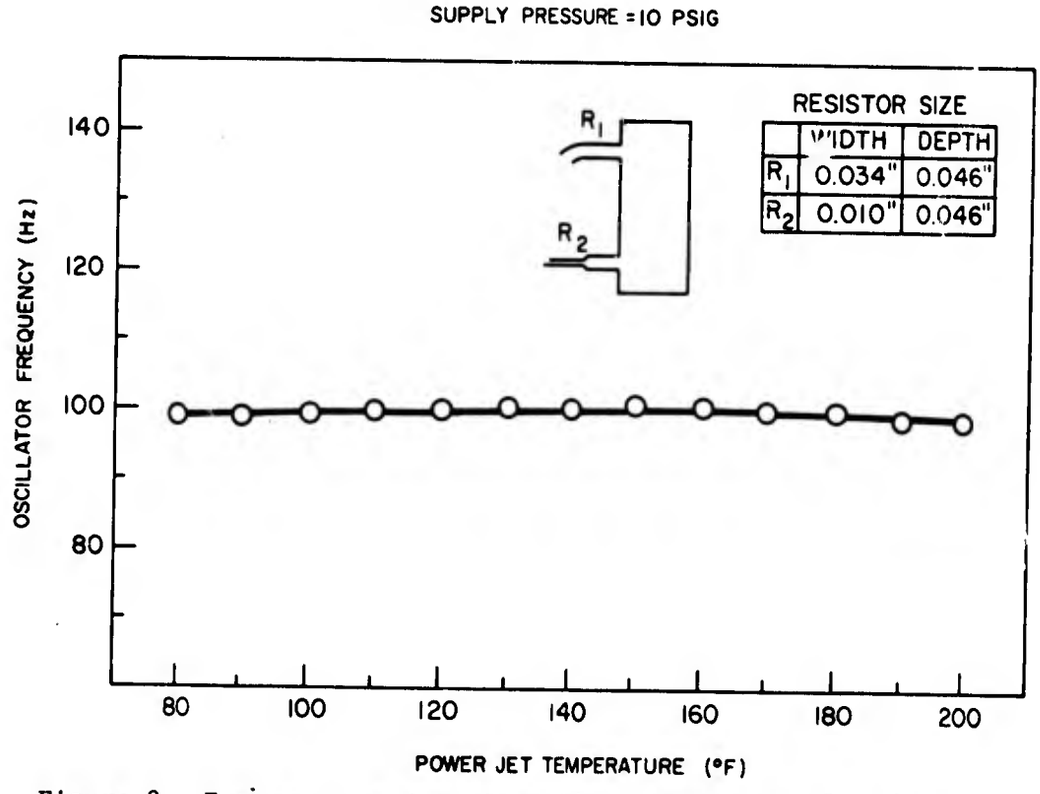


Figure 8. Frequency versus stagnation temperature for the oscillator with a variable resistor R_1.

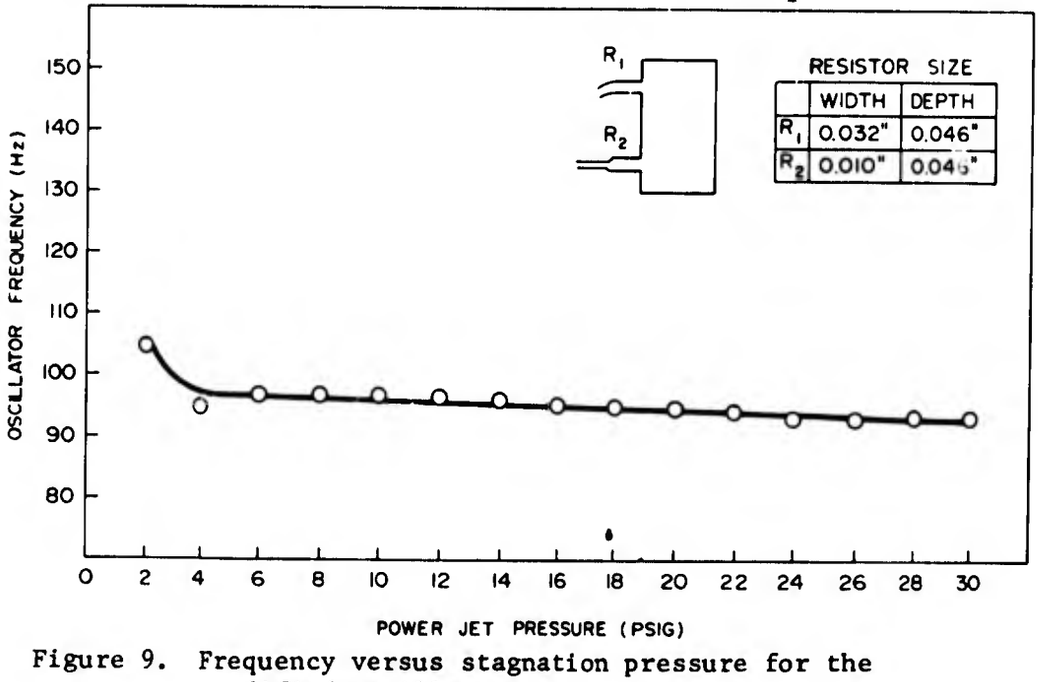


Figure 9. Frequency versus stagnation pressure for the modified oscillator.

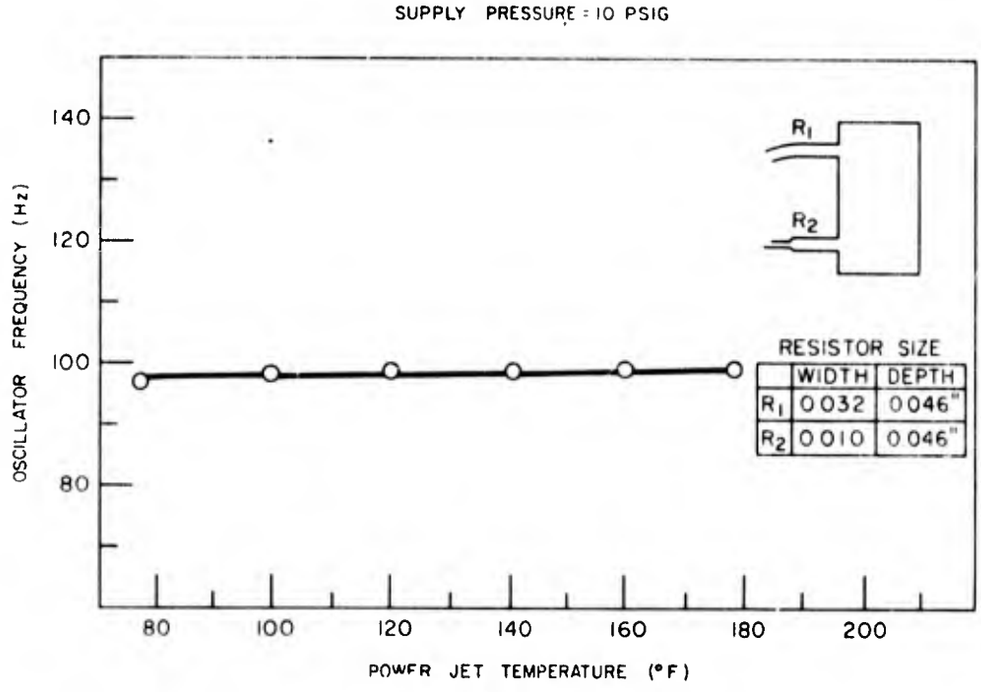


Figure 10. Frequency versus stagnation temperature for the modified oscillator.

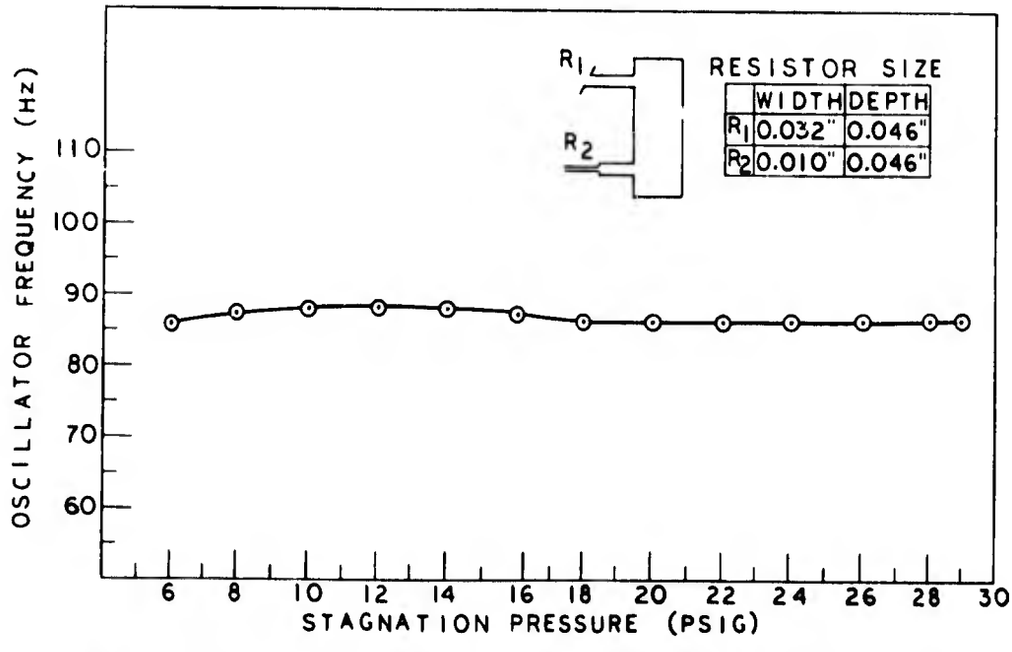


Figure 11. Frequency versus stagnation pressure for the diffusion bonded oscillator.

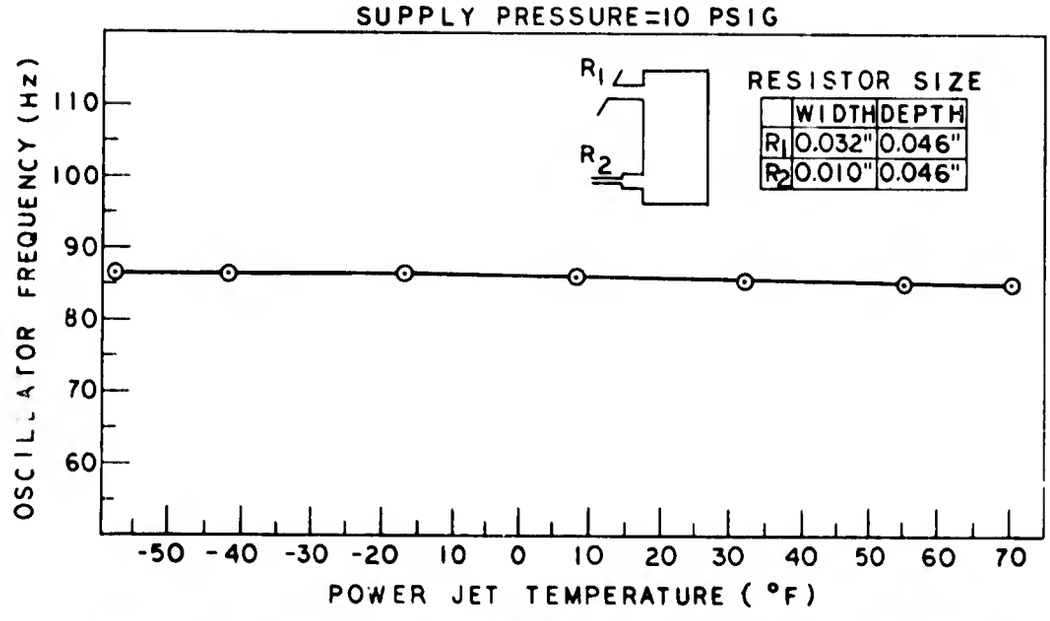


Figure 12. Frequency versus stagnation temperature for the diffusion bonded oscillator.

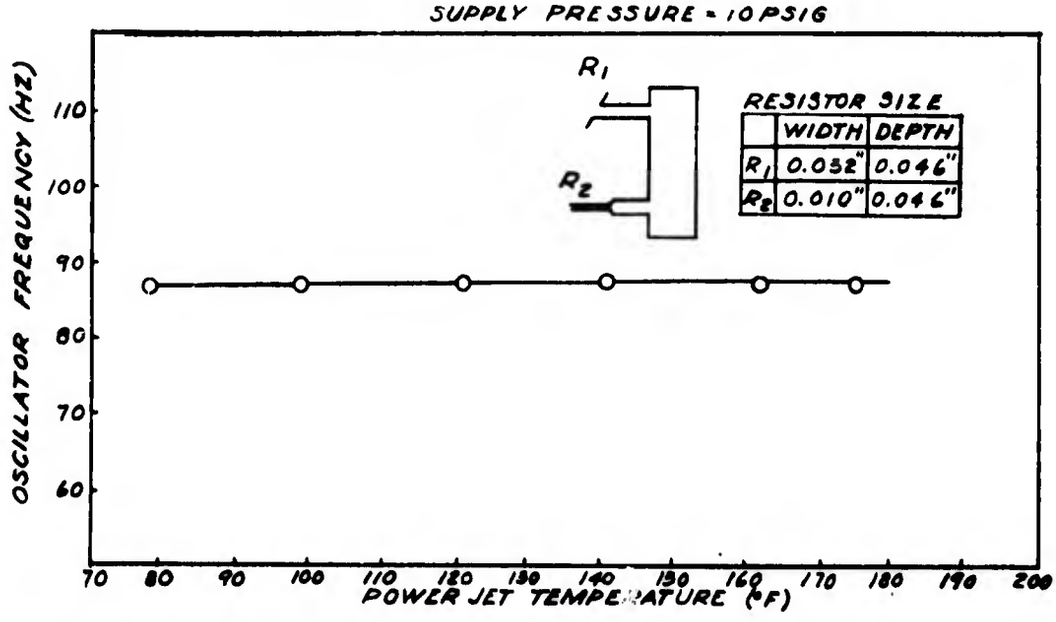


Figure 13. Frequency versus stagnation temperature for the diffusion bonded oscillator.

AD 713 515

COMPUTER OPERATED AUTOMATIC FUZE TESTING SYSTEMS

ROGER P. CHASE
HARRY DIAMOND LABORATORIES
WASHINGTON, D. C. 20438

Introduction and Theme

There has been an accelerating interest in computer-controlled test equipment in recent years. In the early 60's, it became apparent that the proliferation and cost of special purpose test equipment for military systems was getting out of hand, and steps were taken toward the development of multipurpose automatic test equipment (1). The new systems usually consist of many general purpose signal generators and monitors interfaced to a computer which operates them by remote control to perform the same operations a test technician could perform but with much greater speed and unerring adherence to prescribed procedures. This combination of hardware constitutes a very general purpose testing apparatus. It requires only a program to control it, physical connection to the test subject, and a minimal amount of interpretive intervention from the operator. Since the controlling programs for this equipment are written with the benefit of a full engineering knowledge of -- and experience with -- the equipment to be tested, operator training costs are low and skilled personnel are freed for other work. The Army's depot-based DIMATES (2) and the Navy's carrier-based VAST (3) systems are excellent examples of the fruition of these concepts. A broad range of electronics hardware may be fully tested with this equipment in as little as 1/10 the time required using ordinary manual techniques (4). While the cost of these systems is high, it seldom exceeds the collective cost of the special purpose systems they replace, and the great increase in speed and accuracy costs nothing more.

Computers and Test Equipment

It is common to speak of the operations performed by computers in anthropomorphic terms. For example, one might say, "The computer will now search its memory for all last names beginning with "M" and print them out in alphabetical order." While the analogy suggested by this mode of expression is a very apt one, it tends to obscure the fact that a computer must be instructed or programmed to do its work. A computer program, like a musical composition, leads a life of its own as an entity apart from the computer hardware it controls. To emphasize this, computer programs are often called collectively "software". While a computer by itself is a general purpose tool, with a controlling program it becomes a machine with a special purpose. A computer with a program permanently installed is a special purpose computer. All fixed program automatic test equipment belongs to this class.

As automatic test equipment has evolved, improvements have been made utilizing plug-boards, paper tape readers, and so on to enhance programming versatility. The effect of these improvements has been to make the test equipment resemble ever more closely a general purpose computer interfaced to a group of versatile signal generation and measurement devices. So, from an evolutionary point of view, it is a small step to fully --programmable --computer controlled test equipment. From a conceptual point of view, there is a decided change of emphasis from hardware to software -- the system becomes software oriented and readily adaptable to diverse tasks without hardware alterations. This single factor can, after one initial investment, allow tremendous savings in tooling costs for new development items.

Computers Large and Small

When the computer revolution began, computers were physically large and cumbersome, and interest was focused mainly on problem solving and processing large amounts of data presented on prepunched cards. Real-time applications were not a practical possibility. Today there exist a variety of small computers and even smaller "mini" computers specifically designed for system oriented real-time applications. These machines can perform a limited amount of real-time data processing. In many applications this is sufficient to render the data useable immediately without further processing by a larger computer. This is the case in most automatic testing situations. The importance of real-time (on-the-spot) processing can be seen if we consider that one day's delay in discovery of a malfunctioning high speed assembly line

could leave us with $50,000 worth of very expensive junk.

The high-volume production test requirements of the 40mm XM596 fuze have led us at HDL to investigate the application of computer supervised testing and quality control to fuze production and final acceptance. Although the present remarks and conclusions will be drawn from this experience, your attention is invited to the broader implications of these concepts.

The XM596 Fuze

The XM596, developed at HDL, is a low cost 40mm fuze that is produced on a fully automatic assembly line at the rate of one complete fuze every three seconds. The fuze is assembled automatically piece by piece on a continuous web which runs the length of the assembly line. At the end of the line, each completed circuit is cut from the web and is inserted and sealed with a safety and arming device into an outer casing to form a complete fuze. The testing requirements for this fuze include measurements of arming time, transmitter power, burst height, total power supply current and a number of voltages that indicate proper circuit operation. The circuit web is probed automatically at three points along the assembly line to make tests that assure correct operation of the circuit at these intermediate stages. A circuit that fails to operate properly is automatically marked with an identifying hole that prevents further assembly. The completed circuit with radome installed receives a final test at the fourth test position.

The XM596 Automatic Testing System

The XM596 Automatic Testing System is a general purpose data acquisition and control system with test probes and programs designed for the XM596 fuze. The major components of the system are shown in Figure 1. A central processing and control system communicates with remotely located test stations which, in turn, control the electrical operations inside the probes that contact the fuze circuit. A test station consists of a data transfer unit and the signal generators, power supplies, and so on which it controls to effect tests.

Data are transmitted between the interface and the data terminals in a bit serial pulse-space format over coaxial cables. Transformer coupling at both ends of each coaxial line provides a differential high-pass transmission link with high immunity to industrial electrical noise.

Each test position has a test program (a subordinate part of the software system) which resides in the computer memory and directs operations at the test station during the testing cycle. By transmitting sequences of 16-bit words addressed to a particular test station, the software test program can control the station and command readings of analog and digital quantities to be transmitted back to the system interface and thence to the computer for processing. An internal real-time clock permits software control and recording of the time of each test.

The test stations are interrogated in cyclic order once every eight microseconds by the interrupt scanner located in the system interface. When the production machine closes a test probe on the circuit strip, an interrupt request is transmitted to the scanner. Within a few milliseconds, the request is recognized, the currently running program is interrupted momentarily and the new test program is started. Any test programs running at the time of an interrupt operation resume concurrent operation with negligible delay or disturbance. In a sense, all test programs may operate simultaneously because they are executed in parts interleaved with programmed delays and effectively time-share the computer. Measurements require an average of 70 microseconds to complete. Most of the testing cycle is consumed waiting for the fuze to react to stimuli.

All measurements made by the system are based on measurements of time and voltage. These basic measurements are performed using start/stop counters and analog to digital converters located in the test stations and controlled from the central computer by the test programs in the software system. Other quantities of interest are transformed into the time-voltage measurement space by transducers and scaling amplifiers located in the test probes. The basic system accuracy is 0.1% (10 bits).

Data on each test parameter are accumulated statistically in histogram form. Statistical distributions of real-time raw data may be printed on the line printer on request. The computer makes pass/fail decisions under control of the software system based on all data received during a test sequence. Should a fuze fail, a command will be sent causing a hole to be punched in the defective circuit which will inhibit further assembly and testing. Consecutive failures of any particular test parameter in excess of a preset acceptable number will be recognized by the software system and the machine responsible will be

stopped automatically. An interactive control program permits the human operator to change the test acceptance limits, request data printout and make other modifications to the software via the teletype while the system is testing fuzes. Access to this program can be gained only if the correct alphanumeric password is known.

Automatic self-calibration is an experimental feature in this system. In addition to testing fuzes the high-speed testing capability of the system is used to test and calibrate the system itself. Each probe assembly contains a set of relays that permit the scaling amplifier inputs to be switched to calibrated standard voltages. A calibration program in the software system manipulates these relays and makes analog-to-digital conversions that permit a calibration of each of the analog measurement channels used by the probes. A malfunction of the interface, data link, or test station logic will nearly always cause a calibration failure, which is signaled to the operator immediately by a typed message indicating the nature of the failure and its location. The calibration program requires about 1/10 second to complete and is run at each test station before each fuze is tested.

Final Remarks

Some conclusions of general interest based on experience with the XM596 test equipment are stated here briefly.

1. Engineering Applications

Although the XM596 project schedule did not permit it, concurrent development of the test programs with fuze engineering can have a beneficial effect on both efforts. In the course of fuze design and development many hundreds of experimental fuzes are built and tested. The testing system can be used to acquire accurate data rapidly on many parameters of engineering interest and print them out in a format convenient for analysis and documentation. Also, probing and testing problems can be discovered early and fuze design changes can be made, if necessary, to assure effective, low-cost production testing.

2. Peripheral Equipment

It often happens in applications of small computers that the cost of peripheral equipment exceeds the cost of the computer itself. This is particularly true of the high-speed line printers and disk memories that are essential to efficient programming. In all but the

simplest applications, money saved on these items will be paid out in programming man-hours and frustration. Since many systems, once programmed, do not require high-speed printout or mass memory, these items can be reused on other systems and their cost spread accordingly.

3. Software Design

Since production test requirements nearly always change as experience is gained with a new product, the ability to react quickly to these changes is necessary and valuable. For this reason care should be taken that the software structure be as general as possible. The programs that control the operation of individual test positions should be modular, self-contained entities subject to the control of an overall real-time scheduler program. This makes it possible to change the order and timing of tests or whole test programs with minimum impact on the rest of the system. Software development is expensive. A few computer manufacturers offer real-time software systems in addition to the usual basic software library. A well designed, tested and documented software system that can easily be adapted to the testing application can easily save many man-months of effort.

Coda

The XM596 Automatic Testing System was a practical experiment that verified the superiority of computer software oriented fuze testing systems. In addition to immediate and continuing cost savings to the government, it provided a vehicle for training government engineering personnel in the application of digital computer technology to data acquisition, engineering, and control problems. This experience, which has great and general utility, is a valuable asset to the agency and to the personnel.

Design studies based on the XM596 testing system are now under way that will lead to the multiple production of a standard, compact, computer controlled fuze testing system for general use in HDL fuze engineering and procurement programs.

References

1. Policy Guidance Letter, General F. S. Besson, Jr., USA, Rec'd at HDL 9 March 1963.

2. SETE Report 209/42, 10 October 1964.

3. "VAST A Computerized System for Carrier Based Avionics", Capt. A. J. Stanziano, USN, 1966 IEEE Inst. Conv. Rec. Part 4, pp 81-86.

4. "Automation Speeds Tests of Radio for SE Asia", Army Research and Development Newsmagazine, April 1968.

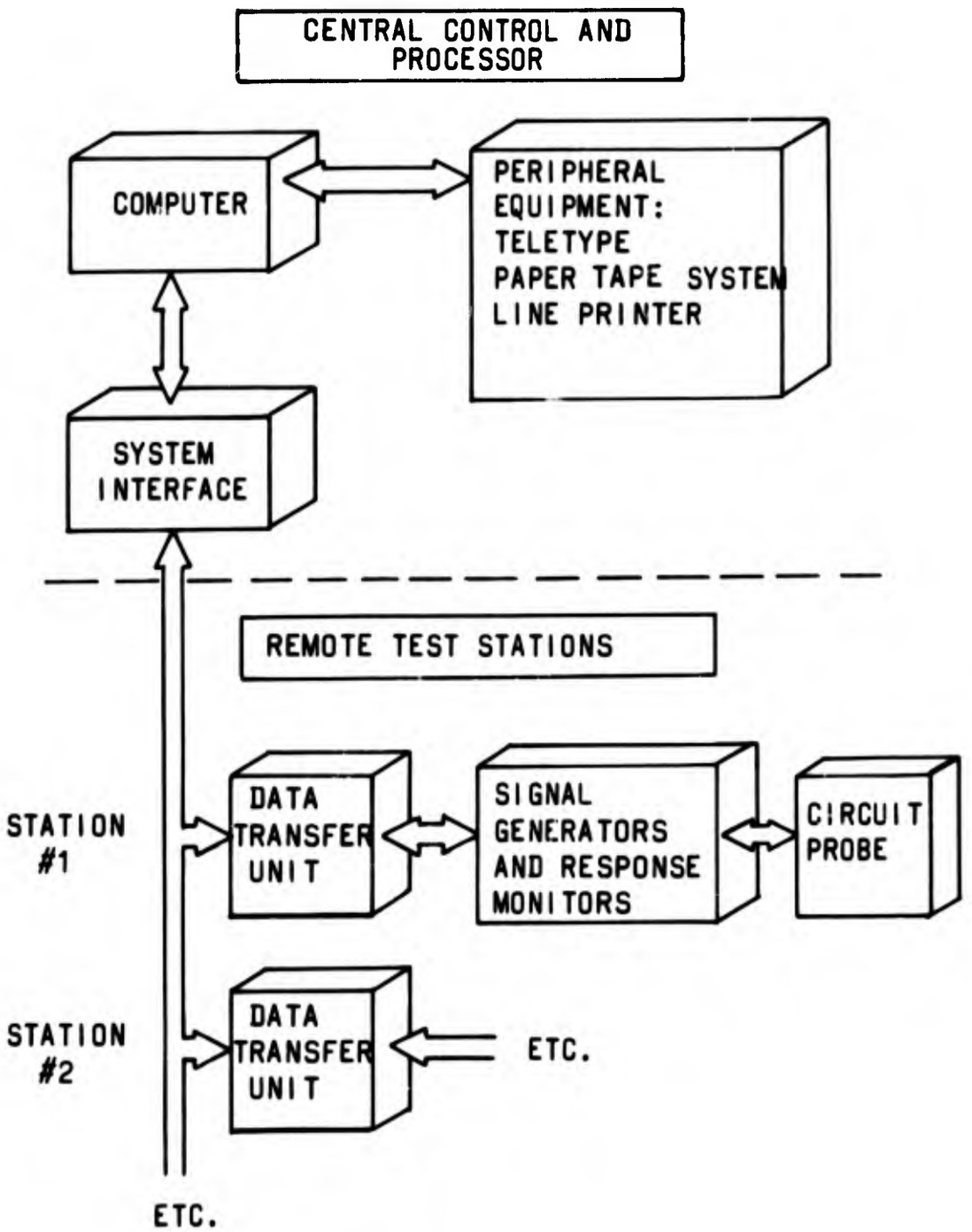

CENTRAL CONTROL AND PROCESSOR
COMPUTER
PERIPHERAL EQUIPMENT:
TELETYPE
PAPER TAPE SYSTEM
LINE PRINTER
SYSTEM INTERFACE
REMOTE TEST STATIONS
STATION #1
DATA TRANSFER UNIT
SIGNAL GENERATORS AND RESPONSE MONITORS
CIRCUIT PROBE
STATION #2
DATA TRANSFER UNIT
ETC.
ETC.

FIGURE 1

AD 713 516

PROPERTIES OF n-BUTYL-α-CYANOACRYLATE RESTORATIVE MATERIALS

SIMON CIVJAN, LIEUTENANT COLONEL, DC
U. S. ARMY INSTITUTE OF DENTAL RESEARCH
WASHINGTON, D. C.

INTRODUCTION

The alkyl-α-cyanoacrylates have recently gained increased attention in both medical (1,2) and dental professions. In dentistry, these monomers have been investigated for use as liners, pit and fissure sealants, and more recently as luting agents, hemostatic dressings and temporary tissue adhesives. (3-6) While earlier work has been limited to the rapidly degradable and more toxic methyl ester, recent studies have been expanded to include the more biostable higher homologs. (7-11) Of these, the n-butyl-α-cyanoacrylate has gained the widest attention. The purpose of this study has been to assess the usefulness of butyl cyanoacrylate in restorative dentistry. For this purpose, polymerization characteristics and physical properties of the unfilled and variously filled resin have been investigated.

PROCEDURE

The monomer was mixed with a series of metallic oxides and salts to provide working consistencies, initiate polymerization and possibly reinforce the resin. For uniformity, powders and monomer were mixed by mortar and pestle; however, mixing could also be done on a glass slab, paper pad and mechanically in gelatine, bakelite or metal capsules. Properties of the filled resin were determined for below, near and above optimum filler-monomer ratios. Setting time, compressive strength and solubility and disintegration were determined by A.D.A. Specification No. 9 for Silicate Cement. (12) Solubility and sorption of unfilled resin were measured by A.D.A. Specification No. 12 for Denture Base Polymer. Tensile strength was measured by the diametral compression method described by Civjan and de Simon. (13) Shear strength was determined by the punch method of Taylor and Margetis. (14) Modulus of elasticity was computed from relative platen displacement in compressive tests. To evaluate

effects of extreme clinical conditions, additional mixtures were allowed to set under water. To assess polymerization characteristics of the monomer, one drop of each, 0.1N as well as saturated solutions of antimony and aluminum chlorides, n,n-dimethyl-p-toluidine and ammoniated silver nitrate, were added to 0.5 ml monomer aliquots. In addition, pure monomer and monomer containing 1% benzoyl peroxide were exposed to ultraviolet light.

RESULTS

No polymerization was initiated in 24 hours by either antimony or aluminum chlorides, by the 1% benzoyl peroxide or UV light alone. (Table 1). One percent benzoyl peroxide-monomer solution, when exposed to UV light, polymerized gradually within one hour. The amine caused polymerization in one minute. Ammoniated silver nitrate polymerized the monomer in less than a second with evidence of silver reduction.

Physical properties of the unfilled resin were determined on the amine initiated bulk polymer. (Table 2). Compressive, tensile and shear strengths of 3,400, 600 and 1,200 psi were in the order of previously reported values for zinc oxide-eugenol cements. (13) The modulus of elasticity of only 120,000 psi was about one third of values for the zinc oxide-eugenol cement. Sorption and solubility were similar to those of commonly used denture resins. (12)

Of the oxides (Table 3), lead, silver, barium, calcium and mercuric oxides induced very rapid polymerization. Reactions with magnesium, titanium, nickel, zinc and zirconium oxides were somewhat slower, while mixtures with antimony and tin oxides set very slowly, if at all. The only sulfide used in this study, zinc sulfide, appeared to react somewhat faster than the oxides. (Table 3). In general, the oxides as well as the sulfide formed tough, rubbery masses that were difficult to manipulate.

Of the salts (Table 4), calcium silicate and calcium phosphate caused nearly instantaneous polymerization. Plaster of paris-monomer mixtures set in about eleven minutes. Slurries with calcium carbonate set over a range from less than one minute to 80 minutes, depending on the source of the powder. Mixtures with zinc phosphate set in seven hours, while slurries with barium sulfate, calcium oxalate and calcium metaphosphate set very slowly, it at all. Mixtures that were easy to handle could be obtained with any one of the salts that did not set too rapidly. With all fillers, water accelerated the set. At least the surface portions of the specimens tended to set on contact with water.

Calcium sulfate hemihydrate (Table 5) did not alter significantly compressive, tensile or shear strength of the resin, although it increased the modulus of elasticity from 120,000 to a maximum of

250,000 psi. The solubility and disintegration varied with filler content and ranged between 0.9 and over 3%. Early contact with water reduced mechanical properties by about 20 to 40% and nearly doubled solubility and disintegration of the plaster of paris filled resin. (Table 6).

Calcium carbonate (Table 7) raised compressive strength of the resin to a maximum of 5,600 psi and approximately doubled tensile and shear strengths as well as the modulus of elasticity. Solubility and disintegration was low, but increased nearly tenfold as the filler content rose from 0.5 to 1.0 g/0.5 ml. Allowing the material to set in direct contact with water was not as harmful to the calcium carbonate as to the calcium sulfate filled resin. (Table 8). Early exposure to water did not reduce the strength of the calcium carbonate filled materials. Solubility and disintegration remained low, but values were more than ten times higher with the wet than with the dry cured specimens.

With the exception of an increased modulus of elasticity at higher filler-monomer ratios, strength of the zinc phosphate filled resin was nearly the same as of the polymer matrix. (Table 9). The low compressive value of only 1,100 psi at 1.0 g/0.5 ml and the generally low tensile strength of 400 psi were probably due to the long setting time with this filler. Solubility and disintegration of specimens at the near optimum manipulative powder-liquid ratio was 0.21%.

Substitution of zinc oxide for part of the zinc phosphate (Table 10) gave a wide range of setting times. 0.1 to 0.3 g zinc oxide in 1.5 g of powder gave practical setting times of 4 to 10 minutes. Compressive strength and modulus of elasticity reached maximum respective values of 5,700 and 500,000 psi with 0.2 g zinc oxide in 1.5 g of powder. A combination of near optimum manipulative strength properties were thus obtained with powder containing 2 parts of zinc oxide to 13 parts of zinc phosphate at a filler-monomer ratio of 1.5 g to 0.5 ml. Other values with this formulation were as follows: tensile strength - 900 psi, shear - 2,400 psi, solubility and disintegration -0.03% (Table 11).

SUMMARY AND CONCLUSION

1. The strongly acidic Friedel Crafts reagents failed to initiate polymerization.

2. Experiments with monomer-benzoyl peroxide solutions indicate that free radical polymerization with butyl cyanoacrylate is possible but is too slow for restorative dentistry. The practical mode of polymerization with the cyanoacrylates, therefore, appears to be anionic.

3. The reactivities of the oxides with the monomer seemed to relate to the state of oxidation and position of the elements in the periodic chart.

4. Reaction rates with calcium carbonate probably depended on manufacturing processes that control particle size and shape as well as residual impurities and pH of the powders.

5. Solubility and disintegration of the filled resin increased with solubility of the filler, filler-monomer ratio and premature exposure to water.

6. Of the individual reagents used, plaster of paris, calcium carbonate and zinc phosphate offered initial promise. Mixtures of fast and slow reacting compounds widened the possibilities. The hemihydrate gave negligible reinforcement and caused excessive solubility and disintegration. Zinc phosphate filled resin had a setting time of 7 hours and did not reinforce the resin. Calcium carbonate and combinations of zinc oxide and zinc phosphate formed effectively reinforced materials with low solubility and disintegration in vitro.

7. From the values shown, one can conclude that filled poly(butyl-α-cyanoacrylate) can combine optimal compressive properties and low solubility and disintegration of zinc oxide-eugenol with the higher tensile and shear strength of zinc phosphate cements. These properties, coupled with favorable tissue response, not only render butyl cyanoacrylate materials promising as cements, bases, liners, pulp capping media and long term temporary filling materials but may well open new investigative avenues in restorative dentistry with other monomers susceptible to ionic polymerization.

BIBLIOGRAPHY

1. Matsumoto, T., Hardaway, R. M., Heisterkamp, C. A., Pani, K. C., Leonard, F. and Margetis, P. M. Cyanoacrylate Adhesive and Hemostasis, Archives of Surgery, 94:858, 1967.

2. Aaby, G. V., West, R. L., and Jahnke, E. G. Myocardial Response to Application of Tissue Adhesives. Comparison of Methyl 2-Cyanoacrylate and Butyl-Cyanoacrylate, Annals of Surgery, 165:3, 1967.

3. Cueto, E. I. and Buonocore, M. G. Sealing of Pits and Fissures with an Adhesive Resin: Its Use in Caries Prevention, Journal of the American Dental Association, 75:121, 1967.

4. Ewen, S. J. Periodontal Uses of a Tissue Adhesive, Journal of Periodontology, 38:2, 1967.

5. Buonocore, M., Azevedo, N. and Dalton, E. Evaluation of Cavity Liners for Permeability and Adhesion to Dentin, New York State D. J., 28:239, 1962.

6. Bhaskar, S. N., Frisch, J. and Margetis, P. M. Application of a New Chemical Adhesive in Periodontal and Oral Surgery, O. S., O. M. & O. P., 22:526, 1966.

7. Cameron, J. L., Woodward, S. C., Pulaski, E. J., Sleeman, H. K., Brandes, G., Kulkarni, R. K. and Leonard, F. The Degradation of Cyanoacrylate Tissue Adhesive. I. Surgery, 58:424, 1965.

8. Leonard, F., Kulkarni, R. K., Brandes, G., Nelson, J., and Cameron, J. J. Synthesis and Degradation of Poly(Alkyl-α-Cyanoacrylate), Journal of Applied Polymer Science, 10:259, 1966.

9. Lehman, R. A. W., West, R. L., and Leonard, F. Toxicity of Alkyl-2-Cyanoacrylates. II. Bacterial Growth, Arch. Surg., 93:447, 1966.

10. Bhaskar, S. N., Jacoway, J. R., Margetis, P. M., Leonard, F., and Pani, K. C. Oral Tissue Response to Chemical Adhesives (Cyanoacrylates), O. S., O. M. & O. P., 22:394, September 1966.

11. Bhaskar, S. N., Frisch, J. and Margetis, P. M. Tissue Response of Rat Tongue to Hexyl, Heptyl, and Octyl Cyanoacrylate, O. S., O. M. & O. P., 24:137, July 1967.

12. American Dental Association. Guide to Dental Materials, Chicago, Illinois: American Dental Association, 1966.

13. Civjan, S. and de Simon, L. B. Evaluation of the Diametral Compression Test, I.A.D.R., 45:126, 1967.

14. Taylor, D. F., and Margetis, P. M. Progress Report on a Setting Time Test for Dental Amalgam, National Bureau of Standards Report No. 6109, June 1958.

TABLE 1

BULK POLYMERIZATION OF n-BUTYL-α-CYANOACRYLATE

Initiator	Polymerization Time
$SbCl_3$ $AlCl_3 \cdot 6H_2O$ Benzoyl Peroxide UV Light	No Polymerization in 24 Hrs
Benzoyl Peroxide and UV Light	60 Min
N,N-Dimethyl-p-Toluidine	1 Min
Ammoniacal $AgNO_3$	<1 Sec with Reduction of $AgNO_3$

TABLE 2

24-HOUR PROPERTIES OF POLY(n-BUTYL-α-CYANOACRYLATE)

Strength, psi	
Compressive	3,400
Tensile	600
Shear	1,200
Modulus of Elasticity, psi	120,000
Solubility*, mg/cm^2	0.03
Sorption*, mg/cm^2	0.2

* By A.D.A. Specification No. 12

TABLE 3

SETTING PROPERTIES OF FILLED
n-BUTYL-α-CYANOACRYLATE

Fillers				Setting Time
PbO_2	Ag_2O	PbO	BaO	<1 Min
	CaO	HgO		
	MgO	TiO_2	ZnS	1-5 Min
	Ni_2O_3	ZnO	ZrO_2	
		Sb_2O_3	SnO_2	2-3 Hrs
			SnO	8 Hrs

TABLE 4

SETTING PROPERTIES OF FILLED
n-BUTYL-α-CYANOACRYLATE

Filler	Setting Time
$CaSiO_3$ $Ca_3(PO_4)_2$	<1 Min
$CaSO_4 \cdot 1/2H_2O$	11 Min
$CaCO_3$	<1-80 Min
$Zn_3(PO_4)_2$	7 Hrs
$BaSO_4$ $CaC_2O_4 \cdot H_2O$ $Ca(PO_3)_2$	16-24 Hrs

TABLE 5

24-HOUR PROPERTIES OF FILLED POLY(n-BUTYL-α-CYANOACRYLATE)
FILLER $CaSO_4 \cdot 1/2H_2O$

Properties	Filler-Monomer Ratio g/0.5 ml		
	0.75	1.50	1.75
Strength, psi			
Compressive	2,600	3,800	3,500
Tensile	500	700	500
Shear	1,500	1,800	1,400
Modulus of Elasticity, $X10^5$ psi	2.0	2.5	1.5
Solubility and Disintegration, %	0.94	1.83	3.28

TABLE 6

PROPERTIES OF FILLED POLY(n-Butyl-α-CYANOACRYLATE)

FILLER $CaSO_4 \cdot 1/2H_2O$

FILLER-MONOMER RATIO 1.50 g/0.5 ml

Method of Cure	Strength, psi			Modulus of Elasticity $X10^5$ psi	Solubility and Disintegration
	Compressive	Tensile	Shear		
Dry	3,800	700	1,800	2.5	1.84
Wet	2,900	400	1,400	1.5	3.42

TABLE 7

24-HOUR PROPERTIES OF FILLED POLY(n-BUTYL-α-CYANOACRYLATE)

FILLER $CaCO_3$*

Properties	Filler-Monomer Ratio g/0.5 ml		
	0.50	0.75	1.00
Strength, psi			
Compressive	3,900	4,900	5,600
Tensile	1,000	1,200	1,200
Shear	2,500	2,700	2,800
Modulus of Elasticity, $X10^5$ psi	2.5	3.0	3.0
Solubility & Disintegration, %	0.01	0.05	0.09

* Merck, Reagent, Precipitated

TABLE 8

24-HOUR PROPERTIES OF FILLED POLY(n-BUTYL-α-CYANOACRYLATE)

FILLER $CaCO_3$*

FILLER-MONOMER RATIO 0.50 g/0.5 ml

Method of Cure	Compressive Strength	Modulus of Elasticity X10^5 psi	Solubility and Disintegration %
Dry	3,900	2.5	0.01
Wet	3,800	2.5	0.12

* Merck, Reagent, Precipitated

TABLE 9

24-HOUR PROPERTIES OF FILLED POLY(n-BUTYL-α-CYANOACRYLATE)

FILLER $Zn_3(PO_4)_2$

Properties	Filler-Monomer Ratio g/0.5 ml		
	1.0	1.5	2.0
Strength, psi			
Compressive	1,100	3,200	2,400
Tensile	400	400	400
Shear	1,300	1,200	1,000
Modulus of Elasticity, X10^5 psi	1.0	3.0	3.0
Solubility and Disintegration, %	-	0.21	-

TABLE 10

24-HOUR PROPERTIES OF ZnO - $Zn_3(PO_4)_2$ FILLED
POLY(n-BUTYL-α-CYANOACRYLATE)
FILLER-MONOMER RATIO 1.50 g/0.5 ml

Filler Composition g ZnO/1.5 g Powder	Setting Time Min	Compressive Strength psi	Modulus of Elasticity X10^5 psi
0.00	420	3,200	3.0
0.05	20	-	-
0.10	10	5,200	4.0
0.20	8	5,700	4.9
0.30	4	5,600	4.7
0.50	1	5,800	5.0
0.75	<1	-	-

TABLE 11

24-HOUR PROPERTIES OF FILLED POLY(n-BUTYL-α-CYANOACRYLATE)
FILLER 0.2 g ZnO/1.3 g $Zn_3(PO_4)_2$
FILLER-MONOMER RATIO 1.5 g/0.5 ml

Strength, psi	
Compressive	5,700
Tensile	900
Shear	2,400
Modulus of Elasticity, X10^5 psi	5.0
Solubility & Disintegration, %	0.03

AD 713 517

OPTIMIZATION ANALYSIS OF A COMPACT LIGHTWEIGHT LASER RANGEFINDER

JOSEPH COSTANTINO, CARMEN PONTELANDOLFO,
DONALD REAGO, FRANKFORD ARSENAL,
PHILADELPHIA, PENNSYLVANIA

I. INTRODUCTION

Years of combat experience by field troops have emphatically proven the need for effective fire control by infantry units. This need is especially acute in the case of indirect fire weapons, such as, mortars. Present fire direction techniques used in the field rely heavily on range estimation by forward observers, a method of range determination that contains an average of at least a 20% error even for experienced personnel. This range estimation error is by far the greatest single source of error causing weapons' fire misses.

After some study of the field problem, it became obvious that a significant improvement in effectiveness could be achieved by using a polar coordinate method of fire control, if the range estimation error was eliminated by some range measurement technique. An apparently ideal device, with a proven capability for accurate range measurement, is the laser rangefinder.

To verify the suitability of a laser rangefinder to mortar type indirect fire, a test was arranged at Ft. Benning, Georgia, in conjunction with the Project Manager for Mortar Ammunition, Picatinny Arsenal, who supplied the low dispersion ammunition used. An AN/GVS-3 laser rangefinder was utilized, although this particular rangefinder is designed for use by the artillery and would not meet infantry requirements.

The conclusive results of these tests demonstrated that it is possible to achieve the ideal advantage of first round fire for effect. Even in the case where the forward observer had no idea whatsoever of his location with respect to the mortar, only one registration round was needed before any target within the forward observer's view was engaged successfully on the first round. For comparative purposes, a highly trained fire control team was assigned

to engage the same targets using procedures presently employed in combat. This team was never able to call for fire for effect until expending a minimum of three correction rounds, thus giving any potential enemy several minutes of warning time and consuming at least one man-load of ammunition per target engagement.

The purpose of this paper is to describe the optimization methods used to achieve a laser rangefinder design that would be suited for use in mortar fire control. The most significant parameters are listed below.

TABLE I

Total Weight	7 pounds
Field of View	7 degrees for monocular 2 milliradians for laser receiver
Magnification	7 power
Maximum Range	2000 meters, visibility limited
Range Accuracy	±5 meters
Laser Power Output	1 megawatt minimum
Laser Wavelength	1.06 microns
Power Source	Rechargeable Ni-Cd batteries
Minimum Number of Rangings Per Battery Recharge	100

II. SYSTEM DESCRIPTION

a. Transmitter

The laser transmitter to be used in this ranging device is a very compact, lightweight unit designed by the American Optical Corporation under contract to Frankford Arsenal. This transmitter package includes the Q-switch optical resonator and all associated electronics: the inductor of the pulse forming network, the time delayed switching circuit for the Pockel cell Q-switch, the high-voltage triggering circuit, and a start pulse detector.

This transmitter uses a 4mm square clad rod with a 3mm diameter core of neodymium doped laser glass. The Q-switch is a deuterated KDP Pockel cell operated with a bias of approximately 4000 volts. The polarizer is a 3-plate "pile-of-plates" arrangement. The flashlamp is a xenon filled 3mm I.D. Pyrex tube in which the ends with the electrodes are slightly bent to eliminate interference with the laser optical path in the compact configuration of the transmitter.

The maximum input energy to the flashlamp is 25 joules. This should result in a minimum Q-switched laser output energy of

of 50mj, 90% of which is to be contained within an 8 milliradian beamspread. At the half-power height of the laser output, the pulse width is approximately 30 nsec., yielding an output power exceeding the 1 megawatt minimum. This unit is capable of operating continuously at a 3 shots per minute repetition rate, and could be used at a much higher rate, e.g., 20 shots per minute, for short periods of time. The minimum life of this unit is expected to be 5000 shots.

The mechanical packaging of the transmitter is based on a rigid aluminium frame that is aligned to the collimating optics by two mounting pins and supported on three precision mounting pads. This arrangement was chosen to facilitate ease of replaceability and eliminate the necessity of realignment each time the transmitter is removed. The transmitter package measures 7" x 2½" x 1¼" and weighs approximately 16 ounces.

b. PFN Power Supply

The DC-DC converter used to charge the energy storage capacitor in the pulse forming network (PFN) had to be optimized drastically in size, weight, and efficiency. A single-ended converter was chosen as the basic circuit. It operates on the same principle as the standard automobile ignition system. The single-ended converter has an inherent advantage over the R-C charging circuit in that the current in the secondary is not limited resistively and, therefore, the converter efficiency is not limited to fifty percent.

Also, the single-ended converter has advantages over resonant charging curcuits where the values of L and C are critical and difficult to realize. Since the pumping requirements of the laser determine C, the value of L needed to maintain a continuous supply of charging energy to the L-C circuit during the complete charging interval is usually impractical with respect to size and weight.

The charging energy in the single-ended converter is supplied by a sequence of pulses. If the time duration of these pulses is made much smaller than the primary time constant, the Q of the circuit is high, and it is then possible to optimize the converter efficiency for a given value of L constrained by the size and weight limitations of the converter.(1)

The power supply designed for the rangefinder is capable of supplying 25J of energy at 2000 volts in less than 3 seconds at an efficiency of at least 85%. The estimated final package dimensions are 2 inches in diameter by 4.5 inches long with a weight of 7 ounces.

c. Receiver

The most critical area in the design of a practical rangefinder at 1.06u is the receiver. In the interest of size, weight,

complexity, and cost, an analysis was made to determine the feasibility of building a receiver utilizing a silicon non-avalanche pin photodiode optimized for 1.06 microns. Use of an avalanche photodiode operating at a "useful gain" of around 100 requires both critical bias stabilization and temperature control. Operation of these diodes at gains of around 10 - 30 is far less critical but the improvement in receiver sensitivity is marginal.(2)

In either case both for avalanche and non-avalanche operation, the system is amplifier noise limited for large ranges and backscatter noise limited for short ranges. More specifically, the optimized gain for the avalanche photodiode is very much dependent on the preamp noise level. This is the reason the expression "useful gain" is used since one does not arbitrarily make the avalanche gain as large as one wishes. The problem of backscatter can be more severe when using photodiodes as compared to photomultipliers. In the case of photomultipliers, one may suppress large backscatter returns via appropriate biasing of the dynode stages. Whereas with photodiodes, one has two problem areas to contend with. One basic consideration in the selection of P-N junction photodiodes is that the signal is composed of two contributions, one from photons absorbed in the depletion region of the diode and the other from the zero field region of the diode. It is the zero field region which can cause complications.

The carriers generated by photons absorbed in the zero field region diffuse to the junction (as opposed to drifting in the depletion region). This phenomenon in effect causes a delay between the incoming photon signal and the collected photoelectron carriers. Thus, a large backscatter pulse can in effect mask a smaller signal return if this effective delay time is substantial enough. Therefore, one must judiciously choose a diode optimized for the wavelength region of interest. This selection applies to both avalanche and non-avalanche diodes. The second problem arises from maintaining a constant false alarm rate throughout the ranging interval. This amounts to varying the detection threshold with range. This can take the form of time program gain for the first stage preamp for non-avalanche photodiodes or with the avalanche photodiode one can consider varying the avalanche gain with time. This leads to further complications and complexity in that we now have tight bias stabilization, a gain temperature control circuit, and time programmed gain.

With the foregoing discussion in mind, the decision was made to determine the feasibility of building a 200 - 2000 meter rangefinder using a non-avalanche diode as the receiver.

At optical frequencies one must use a statistical analysis to arrive at a required signal to noise ratio for a given probability of detection and false alarm rate. The statistical properties of radiation have been shown to Poisson distributed.(3)

When the average number of photons arriving per receiver resolution time is greater than or equal to 10 the Poisson distribution very closely approximates a Gaussian distribution.

Experience has shown that for a reliable ranging system one needs a 99% probability of detection and a .1% false alarm probability. These requirements result in signal to noise ratio of 7 to 1 and a threshold to noise ratio of 5 to 1. Having established these system parameters, one may now calculate the transmitter power commensurate with these requirements.

Taking into consideration all noise contribution in the receiver, one obtains an expression for the signal to noise power ratio as given by

$$\left(\frac{S}{N}\right)_p = \frac{\rho^2 P_r R_L G^2}{2eB(\rho P_r + \rho P_b + I_d)R_L G^2 + 2FkTB} \tag{1}$$

Where the parameters are defined in Appendix A. The background power P_b which contributed noise to the system is given by

$$P_b = \frac{H_s A_r T_a T_o \alpha_r^2 \beta}{4} \tag{2}$$

If we now solve equation 1 for the received signal power required for a given signal to noise ratio, we obtain

$$P_r = \frac{(2\rho eBR_L G^2 \frac{S}{N}) \pm \left\{(2\rho eBR_L G^2 \frac{S}{N}) + (\delta \frac{S}{N} \rho^2 R_L G^2 C)\right\}^{1/2}}{2\rho^2 R_L G^2} \tag{3}$$

Where $C = [eBR_L G^2(I_d + \rho P_b) + FkTB]$

Now making use of the laser radar range equation

$$P_r = \frac{\beta T_a T_o T_s P_t r^2}{R^2} \tag{4}$$

Where T_s represents the two-way transmission loss, due to scattering, to a target at Range R as a function of wavelength and visibility[5] and is given by

$$T_s = \exp -\left[\frac{3.91}{V}\left(\frac{.55}{\lambda}\right)^{.585\sqrt{V}} 2R\right] \qquad (5)$$

Visibility is used here in the conventional definition as the range at which a 100 percent contrast target is at the threshold of visibility. For a typical military target a contrast of 30 percent is more resonable. Middleton [6] gives visibility under such conditions as

$$V = \frac{-3.9\,R}{\ln\left(\frac{.02}{.30}\right)} = 1.44\,R \qquad (6)$$

Where R is the target range.

We now return to our discussion on the transmitter power required to meet the detection criteria.

Making use of equations 2 through 6 and R equal to a maximum range of 2000 meters, we find that

$$P_t \geq .94 \text{ megawatts} \qquad (7)$$

This coincides with the state-of-the-art of Nd laser transmitters achievable which are relatively compact and require fairly low power pumping.

If one examines the various noise contributions in the signal to noise expression of equation 1, the worst offender and by far major source of noise is that due to first stage amplifier. It is this problem area which has led to the development of the avalanche photodiode. More specifically to ranges greater than about 2000 meters and large amplifier bandwidth requirements (narrow laser pulses) the avalanche photodiode does have an advantage over non-avalanche but at the price of system complexity. Thus, the area of preamp design is a most critical one.

The signal current generated for a 2000 meter range and parameters assumed is approximately 70 nano-amperes. Remembering that the required peak signal to rms noise ratio is 7 to 1, this says that an equivalent input rms noise current of 10 nano-amperes is needed.

Until recently commercially available low noise wideband amplifiers were inadequate to satisfy these needs. Texas Instruments has recently introduced an amplifier specifically designed for use with photodiodes which shows great promise. Their model TIXL152 exhibits a wideband (18MHz) equivalent rms current input of typically 3 $pA/Hz^{\frac{1}{2}}$. For a minimum bandwidth requirement of 10 Megahertz this leads to rms noise level of approximately 10 nano-amperes.

Thus, a receiver for 1.06 microns and maximum target range of 2000 meters is realizable using state-of-art components, viz. a P-N photodiode in conjunction with a low noise wideband amplifier.

d. Optics

There are two optical systems within the rangefinder: one each for the transmitter and the combined receiver-sight.

The laser transmitter makes use of an 8-power reverse Galilean telescope to collimate the laser output, reducing the beamspread to 1 milliradian at the 90% power points. Although this telescope would be used only at the 1.06 micron wavelength, it still represented a design challenge since a path length of only 3 inches was allowed. The extreme curvature of the surfaces had to be carefully chosen to avoid spherical aberration.

The optical system used as a combined receiver-sight is basically a monocular viewing system with a beamsplitter that separates the returning laser radiation and reflects it to the photodiode detector. The optics in this section are used at both 1.06 microns and across the full visible spectrum, thus complicating their design.

The light captured by the 2 inch collecting aperture is directed through two image erecting prisms before going to the dichroic beamsplitter. At the beamsplitter, the 1.06 micron radiation is focused onto a 2.0 milliradian field stop and then impinges on the photodiode where it is converted into an electrical signal. The visible light, on the other hand, is focused onto a reticle pattern identical to the reticle presently used in forward observers' binoculars. This reticle would allow the operator to fall back to standard fire control techniques if there is a failure in the rangefinder.

The field stop for the detector and the reticle used for aiming the laser are bonded into a single optical piece at the beamsplitter. This is done to give the operator the complete assurance that the laser detector is always looking at exactly the same point as the center of the aiming reticle. In this way, all alignment problems between the laser receiver and the sighting optics are eliminated. Also, the field of view of the detector is defined exactly by the aiming reticle allowing the operator to do some target discrimination based on the small field of view of the laser rangefinder.

Incorporated into the field of view of the reticle focal plane is a micro-miniaturized numerical display. Four digit range information is presented directly to the operator, eliminating the need for removing his eye from the sight and possible losing the target. Each digit in the display is a 7 segment character measuring 0.010" square, yielding a total size of 0.010" high by 0.040" wide for all four digits. The segments of the display are actually active segments of a monolithic light emitting diode operating in the red portion of the spectrum. This display would be illuminated only when desired by the operator and can be controlled in brightness. Thus, it is useful in full daytime against bright targets as well as at twilight against dark targets.

e. Counting Electronics

An electronic time interval counter using standard medium scale integrated circuitry accurately measures the time it takes the laser pulse to travel from the rangefinder to the target (where some portion of it is reflected) and back. By using a crystal controlled clock, range accuracy in the order of 5 meters can be achieved.

The signal that initiates the counting of clock pulses is generated by the photodiode in the transmitter when it detects a laser output. A select logic circuit processes the echo signals from the receiver amplifier and uses the pre-selected signal to stop the counter. The range number stored in the counter can then be displayed whenever the operator desires.

If the target is partly obscured to a degree that precludes any way of avoiding obstacles in front of the target, the operator must evaluate the target scene and determine which echo (in the order of their return to the rangefinder) represents the target, e.g., the third echo. The target select logic would then be set to ignore the first two echoes and use the third echo to stop the time interval counter. In this manner, even targets in cluttered environments can be handled successfully.

Because of the extremely high packaging density possible with microcircuits, the entire electronics of the rangefinder weigh only 4 ounces and require less than 2 cubic inches. These figures also include a built-in test capability that would enable the operator or repair man to simulate a ranging operation so that the performance of the device can be evaluated and possible malfunctions located.

III. CONCLUSION

The laser rangefinder described represents a total package not much larger or heavier than a pair of M17 binoculars with carrying case. The advantages to be gained from using the rangefinder are tremedous: accurate surprise fire, efficient use of ammunition,

reduced time needed for each target engagement, faster and more accurate mapping and locating of battlefield points.

The entire design was based on optimization of already developed components and techniques. No high risk approaches were needed to achieve a design of a rangefinder device that would be ideally suited to the needs of the man in the field.

REFERENCES

1. D. Latshaw, "Optimal Design of Converter Circuits for Portable Laser Rangefinders", University of Pennsylvania Master's Thesis, December, 1969.

2. D. Reago, C. Pontelandolfo, "Investigation of the Feasibility of an Erbium Laser Rangefinder", U.S.A. Frankford Arsenal Report R-1875, 1968.

3. E. Wolf, C.L. Mehta, "Determination of the Statistical Properties of Light from Photoelectric Measurements", Physics Review Letters, Vol. 13, No. 24, Dec. 14, 1964, P 705.

4. RCA, Electro-Optics Handbook, 1968.

5. Ross, Laser Receivers, (Chapter 8), John Wiley & Sons, New York, 1966.

6. W.E. Middleton, Vision Through the Atmosphere, (Chapter 6), Univ. of Toronton Press, 1952.

APPENDIX A

LIST OF SYMBOLS

SYMBOL	DESCRIPTION	VALUES (UNITS)
P_t	Transmitter Peak Output Power (Exiting from Trans. Optics)	1×10^6 watts
P_b	Background Noise Power	watts
P_r	Received Signal Power	watts
R_L	Load Resistance	10^3 ohms
G	Detector Gain	1
ρ	Target Diffuse Reflectivity	.1
$\mathcal{R}$	Detector Responsivity	.35 amps/watt
T_a	Water Vapor Transmission Factor	.85
T_s	Round Trip Scattering Transmission Factor	----
T_o	Receiver Optics Transmission Factor	.7
$\left(\frac{S}{N}\right)_p$	Peak Signal to RMS Noise Power Ratio	-----
R	Range	200 - 2000meters
k	Boltzman's Constant	$1.38 \times 10^{-23} J/^oK$
F	Noise Factor of Preamplifier	----
B	Preamplifier Bandwidth(-3db)	15MHz
λ	Laser Wavelength	1.06 Microns
V	Visibility	Km
R	Radius of Receiver Optics	2.54cm
I_d	Photodiode Dark Current	.1uamps

AD 713 518

ECHO AMPLIFICATION IN MAGNETIC MATERIALS WITH APPLICATION TO PULSE COMPRESSION RADAR

A. C. DANIEL, B. D. GUENTHER, AND C. R. CHRISTENSEN
PHYSICAL SCIENCES LABORATORY, U. S. ARMY MISSILE COMMAND
REDSTONE ARSENAL, ALABAMA

INTRODUCTION

Since 1959, Kaplan (1) has been pursuing the potential applications of a microwave interaction that produces an effect known as the spin echo (2). Although his interests for several years were concerned with other applications, Mims (3) in 1964 suggested that the spin echo could be used as a pulse compression technique in chirp radar systems. However, the usefulness of the spin echo at microwave frequencies was generally restricted by the requirements of cryogenic temperatures and the large signal insertion loss. However, several of the required parameters were sufficiently satisfied that interest has been sustained and several of these studies are still being continued (4). In 1968 Kaplan (5) observed another effect which greatly increases the potential usefulness of this technique for the radar system. In certain ferromagnetic materials, such as yttrium iron garnet (YIG) and lithium ferrite (LF), it is possible to observe a spin echo, which exhibits a gain of 10^5 over that of an incident signal pulse and which operates at room temperature. This effect has since been observed in plasmas (6) and has been theoretically studied to the extent that a general criterion for the observation of the amplified echo has been established (7-9). Thus, it would seem that this technique has overcome its principal difficulties and that it should be considered for system applications. In fact, some of the underlying principles have been utilized in the design of a new device (4,10) that has the potential of offering an interesting combination of characteristics to the radar design engineer.

It has been the purpose of the efforts reported in this paper to review the experiments of Kaplan in order that a greater understanding of the characteristics, and limitations, of this device could be obtained and, additionally, to investigate the amplified echo

in other materials for the purpose of exploring the range of its validity. Experiments similar to those of Kaplan have been performed at 4.0 and 9.2 GHz on YIG at room temperature and at 0.6 GHz on MnO at liquid helium temperature. The latter experiment was performed on the nuclear spins of the Mn, as opposed to the electronic spins in the other experiments, while the MnO is in its antiferromagnetic state.

SPIN ECHO AND RELAXATION TIMES

The spin echo, as first reported by Hahn (2) in 1950, was obtained as a result of the pulsed rf excitation of nuclear spins in paramagnetic solids. In contrast to the magnetic materials to be considered later, each of the large number of spins act in a similar but essentially uncorrelated manner. Simply stated, a first rf pulse rotates the spins, as a group, from their equilibrium direction commonly defined by an external static magnetic field. Immediately following the pulse the spins precess about their equilibrium direction and generate a rf signal, referred to as the free induction decay. This signal decays with a time constant T_2^* as the spins lose their coherence. Due to the damping mechanisms in the material, the spins also lose their precessional energy and return to their equilibrium direction with a time constant T_1, where generally $T_1 \ll T_2$. Even though the spins lose their phase coherence in time T_2^*, there is a time $T_2 (T_2^* < T_2 < T_1)$ where a properly applied second pulse will cause the spins to regain their phase coherence and again radiate a rf signal, the spin echo. The time constant T_2 is determined by irreversible phase scrambling processes that are commonly termed as the inhomogeneous contributions to the resonance, whereas the reversible contributions are termed homogeneous. Thus, for two appropriate rf pulses separated by a time τ, an echo is observable at time 2τ and decays with a time constant T_2 as a function of τ. These characteristics are summarized in Fig. 1a. The echo is observed provided the pulse separation satisfies the requirement, $T_2^* < \tau < T_2, T_1$.

Since the first observation of the spin echo, this phenomenon has been found to be much more general than originally thought. The phenomenon is not restricted to the resonant excitation of nuclear and electronic spins. Echoes have been observed from optical transitions in both solids and gases that do not require the consideration of spin, from the classical cyclotron resonance in a plasma, from a superconductor, and from magnetic materials where the spins act as a collective phenomenon (for review, see refs. 7 and 9). The basic criterion for the observation of echoes has been restated in terms general enough to apply to all observed echo phenomena. The physical system can be described by a group of oscillators which interact nonlinearly.

The echo phenomena discussed in this paper are observed from both the electronic and nuclear resonances in ferro- and antiferromagnetic materials. However, the resonances observed in these magnetic materials involve nonlinear interactions between strongly

coupled spins and include the simultaneous excitation of many resonant modes. The collective electronic spin resonances, often referred to as spin wave resonances, that occur in the microwave spectrum are generally strongly coupled to the spins of any magnetic nuclei that may exist in the material. The spins of these nuclei and electrons may often be considered as a single resonant system consisting of many modes of oscillation. It is most often observed that when an external energy source excites one of these modes, several nonlinear effects are observed, such as the saturation and shift of the resonance with increasing applied rf power. This situation is contrasted with paramagnetic materials where the spins are considered as isolated oscillators and the nonlinearity required for the echo observation involves each oscillator individually.

AMPLIFIED ECHO

It should be apparent from this discussion that the radar device based upon this phenomenon will be unconventional. If the first rf pulse is the radar signal returned from some target, then the second pulse, at essentially the same frequency but displaced by time τ, is locally generated at the receiver and is termed a recall or pump pulse. Thus, the radar operator does not work with the radar return signal but instead works with an echo delayed by time 2τ and related to the radar return signal by the recall pulse. This device will not be susceptible to damage due to high rf pulse leakage from the transmitter but may exhibit more complex transmitter saturation effects.

Progress was made toward the development of this technique as a radar receiver with Kaplan's observation of signal gain between the first rf pulse and the resultant echo in YIG and LF at room temperature (5). There are, nevertheless, other characteristics of the echo in these materials that may present sufficient operational inconvenience to reduce its practical acceptability. As an example, YIG has overlapping acoustic and spin wave modes such that once the spin wave modes are excited, energy will be transferred into the acoustic modes. Thus, an incident rf pulse produces an acoustic disturbance which propagates through the material at a velocity much less than that of the spin waves. Reflected and delayed acoustic signals may cause the delayed generation of a rf signal, as in a microwave delay line (11). However, the ultrasonic echoes may occur in the same time interval as the signal echo and be of sufficient amplitude to obscure the signal echo. The strength of the coupling to the acoustic modes increases as the frequency is decreased. Thus, the usefulness of the spin echo in YIG is doubtful at frequencies much below 8 GHz. Another basic limitation is that at room temperature the echo decay time T_2 is of the order of a microsecond, which severely restricts the time during which a radar target signal may be observed to only a narrow time interval following the recall pulse (5).

Echo phenomena have been found in many physical systems. Although the generality of the echo amplification is not as well established, it has been reported for the plasma echo (6) and theoretical attempts have been made to establish the nonlinearity requirements for amplification (7-9). Gould (9) has been able to account for all the observed echo phenomena with one of the following nonlinear effects:

(a) energy-dependent driving force
(b) energy-dependent frequency (anharmonic oscillator)
(c) energy-dependent relaxation process.

He has further attempted to show that the echo amplification is a general feature of the second class of nonlinear effects (8). In the conventional spin excitation by a rf pulse, the net spin magnetization is thought of as being rotated through some angle θ by the rf pulse — the θ-pulse. With the assumption that the resonant frequency is dependent upon the energy of the rf pulse, the resonance frequency of the spin changes as θ is varied. Although this situation is not valid for paramagnetic materials, it is definitely true for both the nuclear and electronic spin wave resonances in magnetic materials. Since spin does not enter into all of the various observed echo phenomena, the concept of the rotation of the net magnetization through the angle θ may be replaced with some generalized displacement of the system oscillator and be related to the system resonance frequency. However, since the present paper is concerned only with magnetic materials, the more obvious concepts of magnetization rotation and precession will be retained.

The text book solution of the nonlinear harmonic oscillator generally assumes that the potential-energy function will be of the form $V' = V + a\theta^2 + b\theta^3 + \ldots$ and the resultant first-order solution for the resonant frequency will be given by $\omega' = \omega + \alpha\theta^2_{max}$, where a, b, α are constants defined by the nonlinear restoring forces. The nonlinear interaction is not so well known for the rotated magnetization in magnetic materials, but it is assumed that the frequency shift will be dependent upon the amount of rotation. Thus, Gould generalizes the above anharmonic equation to one of the form

$$\omega' = \omega + (\partial\omega/\partial\theta^2)\theta^2_{max} \quad .$$

If the precession of the component of the magnetization rotated into the plane normal to its equilibrium direction is taken into account, the response of the single spin system to the first rf pulse can be given by

$$m_\perp = m_\perp' \, e^{i\omega' t} = c \sin\theta_1 \, e^{i\omega' t} \approx c\theta_1 \, e^{i\omega' t} \qquad (1)$$

where the rotation angle θ is generally small in magnetic materials due to the strong coupling to other resonance modes. The resonant frequencies of the precessing moments are never given by a single frequency, but by a distribution of frequencies about this center frequency. But in order to simplify the discussion, this distribution is neglected and the frequencies are assumed single valued. Then, substitution in Eq. (1) for ω' yields

$$m_\perp = c\theta_1 \, e^{i(\omega t+\eta)} \tag{2}$$

where $\eta_1 = (\partial\omega/\partial\theta^2)\theta_1^2 t$ is the accumulated change in phase over the time t. Although the change in frequency $d\omega$ due to the rf pulse is generally small, this phase shift η_1 is generally large.

Following two rf pulses separated by time τ, the magnetization will take the form

$$m_\perp = c\left[\theta_1 \, e^{i\omega'\tau} + \theta_2\right] e^{i\omega''(t-\tau)} \tag{3}$$

where

$$\omega'' = \omega + \left(\partial\omega/\partial\theta^2\right)\left[\theta_1^2 + \theta_2^2 + 2\theta_1\theta_2 \cos \omega'\tau\right] .$$

By using the identity

$$e^{i\alpha \cos \omega'\tau} = \sum_{-\infty}^{\infty} (-i)^n J_n(\alpha) e^{-in\omega'\tau}$$

the substitution of ω'' and ω' into $m_\perp$ yields

$$m_\perp = \sum_{-\infty}^{\infty} A_{n-1}(t) e^{i\omega(t-n\tau)} \tag{4}$$

where

$$A_{n-1}(t) = (-i)^n e^{i\eta_2}\left[\theta_1 J_n(\alpha) + i\theta_2 J_{n-1}(\alpha)\right]$$

$$\alpha = (\partial\omega/\partial\theta^2) 2\theta_1\theta_2 (t - \tau)$$

$$\eta_2 = (\partial\omega/\partial\theta^2)\left[(\theta_1^2 + \theta_2^2)\, t - n\theta_1^2\tau - \theta_2^2\tau\right] .$$

Then the echo that exists at $t = 2\tau$ has amplitude

$$|A_1(t)| = |\theta_1 J_2(\alpha) + i\theta_2 J_1(\alpha)|$$

$$\approx \theta_2 J_1(\alpha) \approx \tfrac{1}{2}\theta_2\alpha = \theta_1\theta_2^2(\partial\omega/\partial\theta^2)(t - \tau) \qquad (5)$$

where it is assumed that $\theta_1\alpha$ is small. If the first pulse is small with respect to the second, $\theta \ll \theta_2$, and at the time $t = 2\tau$

$$\eta_2 = (\partial\omega/\partial\theta^2)\theta_2^2\tau$$

so that

$$|A_1(2\tau)| = \theta_1\eta_2 \quad .$$

An echo gain G may be defined by

$$G = |A_1(2\tau)|/\theta_1 = \theta_2^2(\partial\omega/\partial\theta^2)\tau = \eta_2 \quad . \qquad (6)$$

Thus, the echo gain is the accumulated phase shift following two rf pulses. It is important to note that with the assumption of small θ_1, the amplitude of the first echo given by Eq. (5) increases linearly with increasing pulse power or, over narrow limits, with pulse width. Also, in magnetic materials it is common to observe additional echoes, secondary echoes, which correspond in Eq. (4) to values of $n \geq 3$. These additional echoes are indicated in Fig. 1b. Kaplan has experimentally observed a gain defined from Fig. 1b as A_1/P_1. This factor is significantly different from that obtained from Gould in Eq. (6). P_1 refers to the actual signal applied to the material whereas θ_1 is descriptive of the resonant response. These two signals could differ in some materials by a large factor. This difference adds greater significance to the magnitude of gain observed by Kaplan.

The above description of echo amplification does not take into consideration the damping or relaxation effects. There will be an optimum value of τ at which the greatest gain can be realized before the relaxation effects begin to predominate.

EXPERIMENTAL OBSERVATIONS IN MnO

Antiferromagnetic MnO offers several opportunities to investigate the characteristics of the amplified nuclear spin echo in this different class of materials. Moreover, an independent study of the amplified echo provides a stimulus to further understand such interactions and their potential utilization. The strong coupling often

exhibited between the nuclear and electronic resonances in magnetic materials might be utilized for the direct translation of the incident frequency to a lower value.

Two Mn^{55} nuclear magnetic resonances have been reported for MnO in an external magnetic field (12). As shown in Fig. 2 the higher frequency resonance ν_H is only slightly field dependent, whereas the lower frequency resonance ν_L is strongly pulled to lower frequencies with decreasing field. The degree of such frequency pulling is normally considered as an indication of the strength of the interaction between the nuclear and electronic spin wave resonances. However, for MnO there is some doubt as to the nature of this interaction since the only observed electronic spin wave resonance is at 860 GHz, too far removed to be expected to strongly influence the pulling of the nuclear resonances. For an antiferromagnet such as $CsMnF_3$ this interaction between modes is well understood since its lowest electronic spin wave resonances are easily studied at a few GHz. Since the $CsMnF_3$ resonances are more attainable, this material has also been briefly investigated.

Although the ν_L resonance in MnO is unobservable below about 14 kOe due to the rapidly increasing linewidth (or decreasing T_2^*), above about 18 kOe this resonance is considerably stronger than the higher resonance. The ν_H has a width that remains essentially constant at 0.3 MHz and is always narrower than that for ν_L. Fig. 3 shows the field dependence of T_2 for the echo generated by two narrow rf pulses. Note that the maximum value for T_2 with ν_H is of the order of 10 μsec. After increasing the energy of the first pulse, the observation time of the echo will be seen to greatly increase.

Spin echoes from the two resonances are observed by placing an irregularly shaped single crystal of approximately 10 mm^3 in a foreshortened, reentrant cavity that is excited by a pulsed triode oscillator rated at 2.0 kW and/or by a pulsed signal generator capable of cw operation at 50 W. Fig. 4 is a recorder trace of the echo amplitude, measured with a boxcar integrater, as a function of the separation τ between the first pulse and a pulse pair, as shown in Fig. 1c. The pulse pair was used, instead of a single second pulse, so that the spin echo would be observable at all times and thus permit the relative enhancement of the echo to be easily observed (13). Thus, the pulse pair is used for observational convenience and should be considered as a single pulse. In contrast to the operational definition used by Kaplan, the gain observed here is essentially in agreement with the gain defined by Eq. (6).

As τ is increased the echo amplitude increases rapidly to some maximum value and then decreases to the amplitude obtained from an isolated pulse pair. This characteristic is essentially the same as that reported by Kaplan for the ferrimagnetic echo. Since this plot was obtained from three pulses of equal power, the initial rise

should not necessarily be expected to be linear. The initial increase could be described by an exponential with a characteristic time of no more than 13 μsec. The decay of the echo amplitude was a simple exponential with a characteristic time of 165 μsec, which is a factor of 20 longer than the T_2 shown in Fig. 3. Both of these times appear to be insensitive to large variations in the power or width of the first pulse. A recorder trace of the echo amplitude is shown in Fig. 5 as a function of the width of the first pulse t_1, while the separation τ is held constant. For a reasonably short ($t_1 < 50$ μsec) or a very long ($t_1 > 700$ μsec) first pulse the echo amplitude is equal to that obtained with a two pulse sequence. The steep rise on the leading edge of the trace in Fig. 5 may correspond to the linear response expected for small first-pulse excitations. At the extreme of long first pulses, the pulse affects only a small portion of the resonance line and additionally provides adequate time for several relaxation effects to predominate. It should be kept in mind, however, that it is difficult to critically analyze the effect of such a wide pulse. It is surprising, nevertheless, that the first pulse is effective for widths greater than 600 μsec.

Figs. 6 and 7 show the effects of the first pulse power on the echo amplitude for several pulse widths; these data were taken in a helical terminated transmission line rather than the resonant cavity. Fig. 7 exhibits the type of behavior expected from the theory developed by Gould, i.e., a linear dependence of echo amplitude on the first pulse as long as the approximation of a small first pulse is met. Fig. 6 does not exhibit the linear behavior expected from Eq. (5). The different behavior of ν_H and ν_L is due to the different amount of frequency pulling exhibited by the two resonances. The amount of frequency pulling would determine the size of the factor $(\partial\omega/\partial\theta^2)$. It should also be noted that, in contrast to the ferrimagnetic echoes, the nuclear echo amplification is not dependent on sample shape.

The observation of echo amplification in a nuclear spin system and the general agreement with the theory developed for plasma echoes and ferrimagnetic echoes suggests that echo amplification is a quite general phenomenon.

APPLICATIONS FOR CHIRP RADAR

If the radar transmitted signal is swept linearly in frequency for the duration of a long pulse, then the long pulse should permit the detection of signals reflected off targets at increased range. By utilizing the frequency variation of the transmitted pulse, the earlier portions of the reflected signal may be delayed so that the received energy is concentrated into a short pulse. The radar resolution is thereby increased by the amount of pulse compression without sacrificing the range obtained from the long transmitted pulse. The spin echo technique proposed by Mims satisfies the pulse compression requirement by sweeping the frequency of the second or recall

pulse in the manner shown in Fig. 8. The room temperature observation of echo amplification by Kaplan considerably increased the potential merit of this technique. However, the short echo decay times of $T_2 \approx 1$ μsec at room temperature, and only about 12 μsec at liquid helium temperature, still prohibits the general acceptance of this device. The echo observation time can be increased by multiple application of the recall pulse as shown in Fig. 1d. A operationally acceptable device does not appear feasible with the material characteristics presently known. However, the observation of the amplified echo in plasmas and ferrimagnetic materials holds promise that further material advances can be made.

The magnetic materials appear to hold the most promise for radar systems. It is desirable to have a wide operational bandwidth, which requires resonance absorption over a wide range of frequencies. This requirement is indicated in Fig. 8. If in addition the material has a long T_2 of the order of 100 μsec, then a large compression ratio, as defined in Fig. 8, as well as a large operational time gate for receiving the signal should be obtainable. A very inhomogeneous material, where an excited resonance in one part of the absorption spectrum is not influenced by the remaining spectrum, as indicated by $\delta\nu$ in Fig. 8, may permit longer relaxation times than otherwise expected. If T_2 is short as in YIG, then very severe restrictions are placed on the width of the transmitted pulse and on the receiver. Fig. 9 indicates the type of receiver operation that would be required with a short T_2. The raster of recall pulses would extend the receiver sensitivity over a greater radar range, but there would still be dead zones. Although this difficulty could perhaps be overcome by using two parallel complementary receivers, certain disadvantages still remain.

The paramagnetic materials at low temperatures do not encounter this difficulty. However, at present low temperatures are required and no echo amplification has been observed.

Under laboratory conditions pulse compression ratios of about 250 appear reasonable. However, this technique appears capable of attaining much higher values. As the value of the compression ratio is increased, other requirements will become more noticeable. In particular, the linearity of the frequency sweeps in Fig. 8 can greatly influence the ratio. If the range of frequency sweep is not sufficiently large, the doppler shift of a returned signal may be troublesome. However, these are engineering problems that are presently secondary to those related to the material and to the echo technique.

SUMMARY

Echo phenomena in several materials have been reviewed for the purpose of considering the recently proposed amplified spin echo as a chirp radar receiver. The technique is believed to exhibit considerable merit in that a number of desirable characteristics have

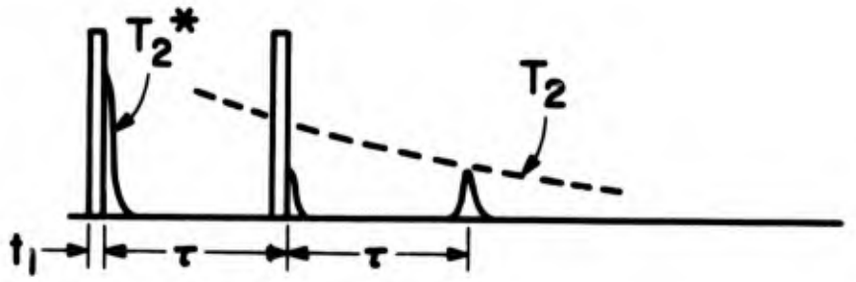
been combined in a unique manner. If properly realized, a portion of the processing of the radar signal could be accomplished at the antenna feed element. However, the available materials still limit the practical usefulness of this device as compared to more conventional techniques.

ACKNOWLEDGMENTS

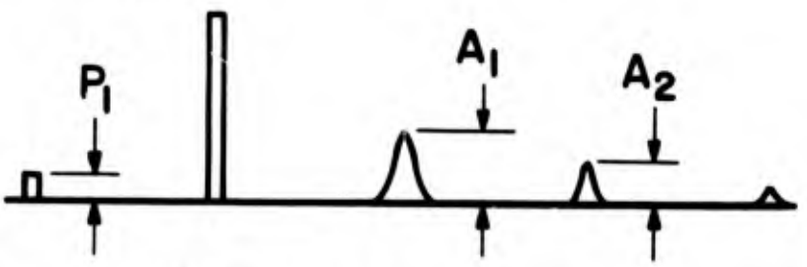
The authors gratefully acknowledge the laboratory assistance provided by Mr. Robert Jensen and the typing and manuscript assistance provided by Mrs. Hazel M. Oaks.

REFERENCES

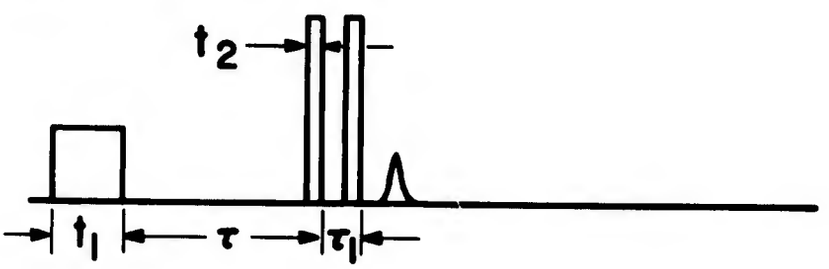
1. D. E. Kaplan, M. E. Browne and J. A. Cowen, Rev. Sci. Instr. 32, 1182 (1961).
2. E. L. Hahn, Phys. Rev. 80, 580 (1950).
3. W. B. Mims, Proc. IEEE 51, 1127 (1963).
4. P. Franklin, Microwaves 8, 13 (Oct., 1969).
5. D. E. Kaplan, R. M. Hill and G. F. Herrmann, Phys. Rev. Letters 20, 1156 (1968).
6. R. S. Harp and R. R. Smith, Phys. Letters 29A, 317 (1969).
7. G. F. Herrmann, D. E. Kaplan and R. M. Hill, Phys. Rev. 181, 829 (1969); D. E. Kaplan, R. M. Hill and G. F. Herrmann, J. Appl. Phys. 40, 1164 (1969).
8. R. W. Gould, Phys. Letters 29A, 347 (1969).
9. R. W. Gould, Am. J. Phys. 37, 585 (1969).
10. P. Franklin, Microwaves 8, 16 (Oct., 1969); __, Electro-Technology __, 26 (Oct., 1969); E. de Atley, Electronic Design 1, 34 (Jan., 1969).
11. For example, see J. R. Eshbach, J. Appl. Phys. 34, 1298 (1963).
12. B. D. Guenther, C. R. Christensen and A. C. Daniel, Phys. Letters 30A, 391 (1969).
13. C. R. Christensen, B. D. Guenther and A. C. Daniel, Bull. Am. Phys. Soc. 14, 1185 (1969).



a. Spin echo experiment



b. Pulse amplification experiment



c. Echo enhancement experiment

d. Extension of echo observation beyond T_2

Fig. 1. Pulse sequences

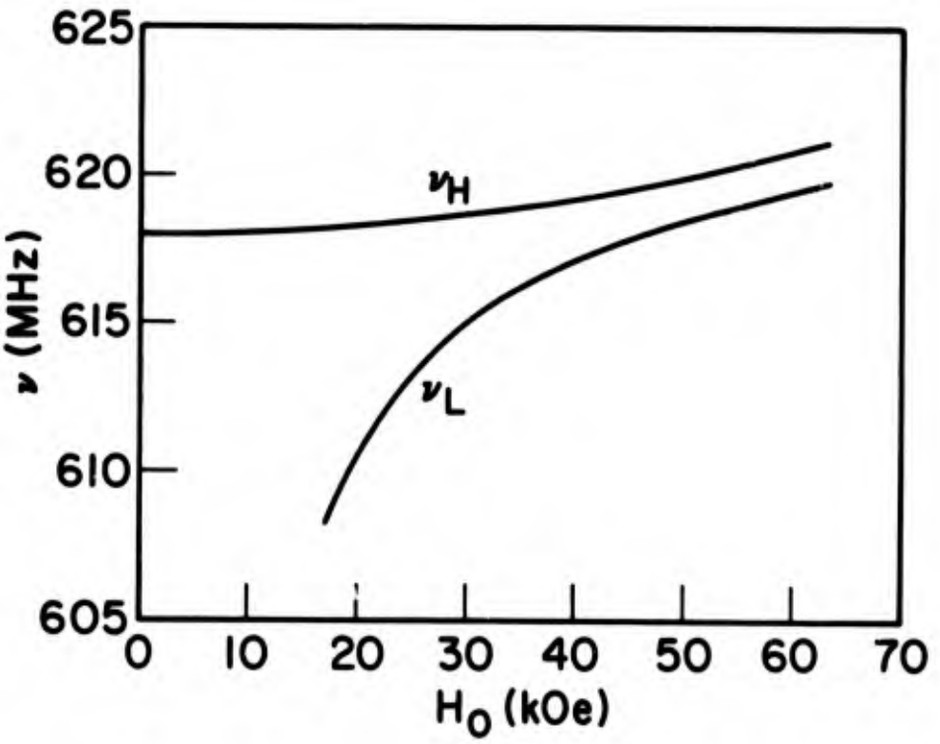


Fig. 2. Nuclear resonant modes in MnO

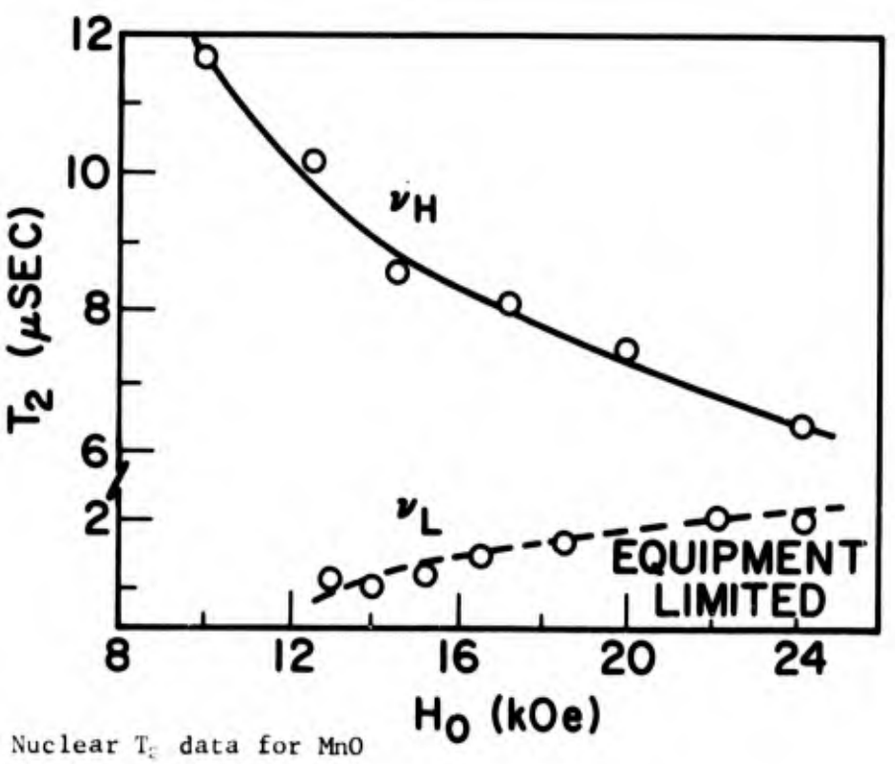


Fig. 3. Nuclear T_2 data for MnO

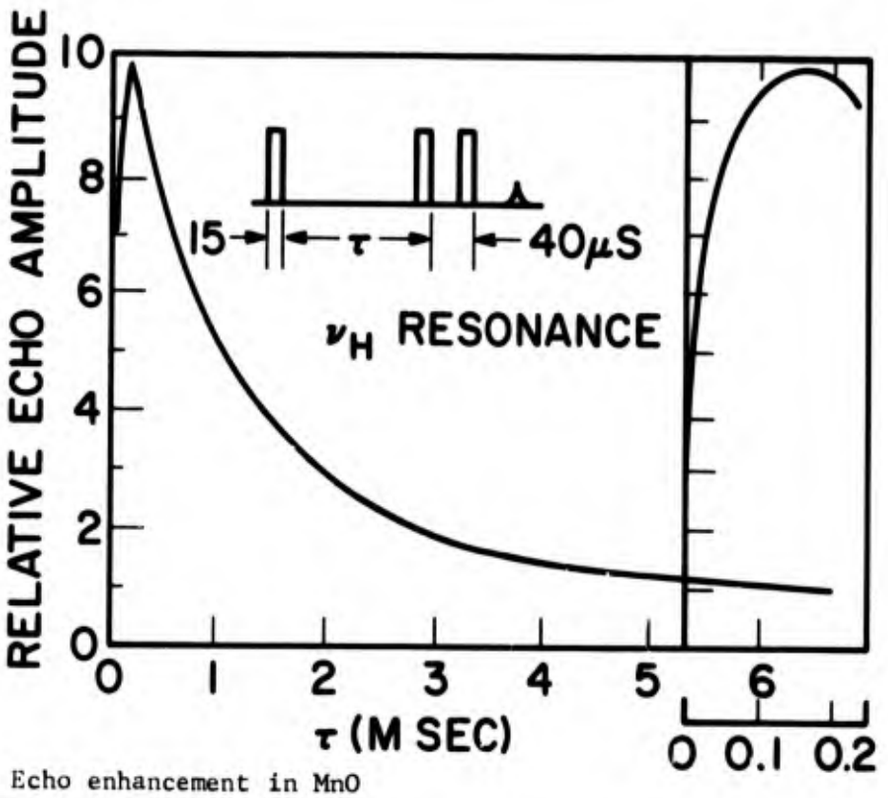


Fig. 4. Echo enhancement in MnO

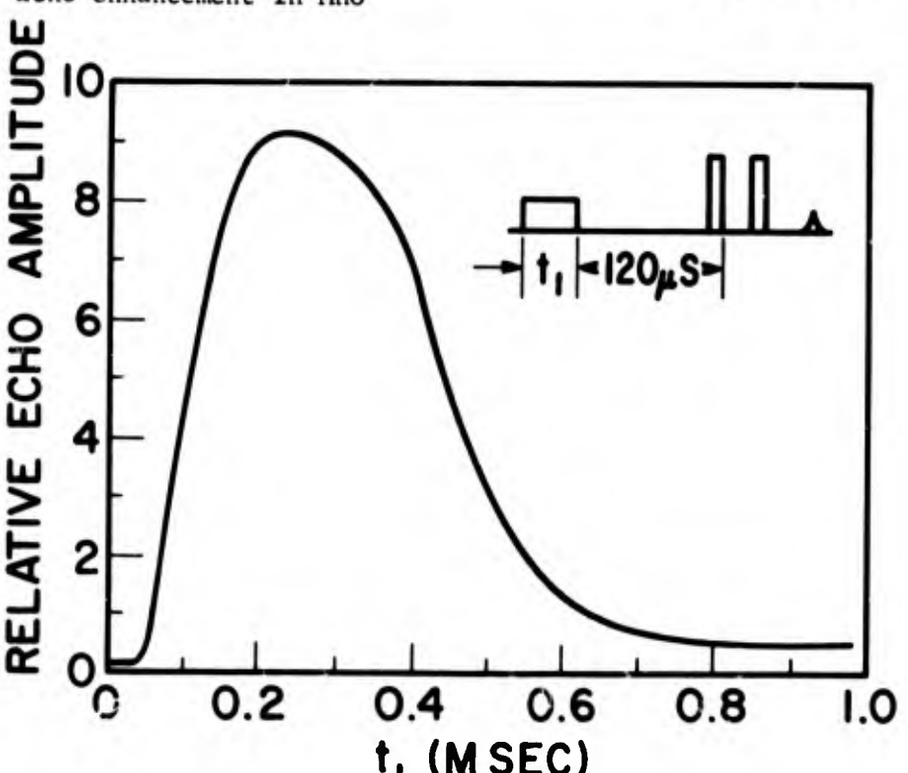


Fig. 5. Echo amplitude as a function of pulse width

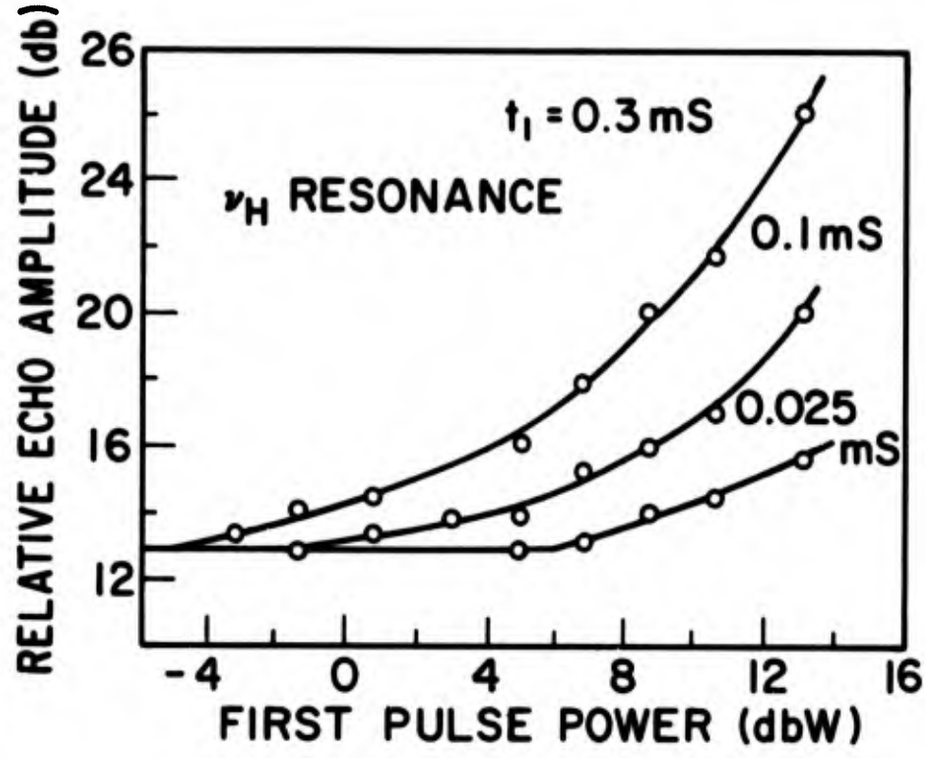


Fig. 6. Echo amplitude as a function of pulse power

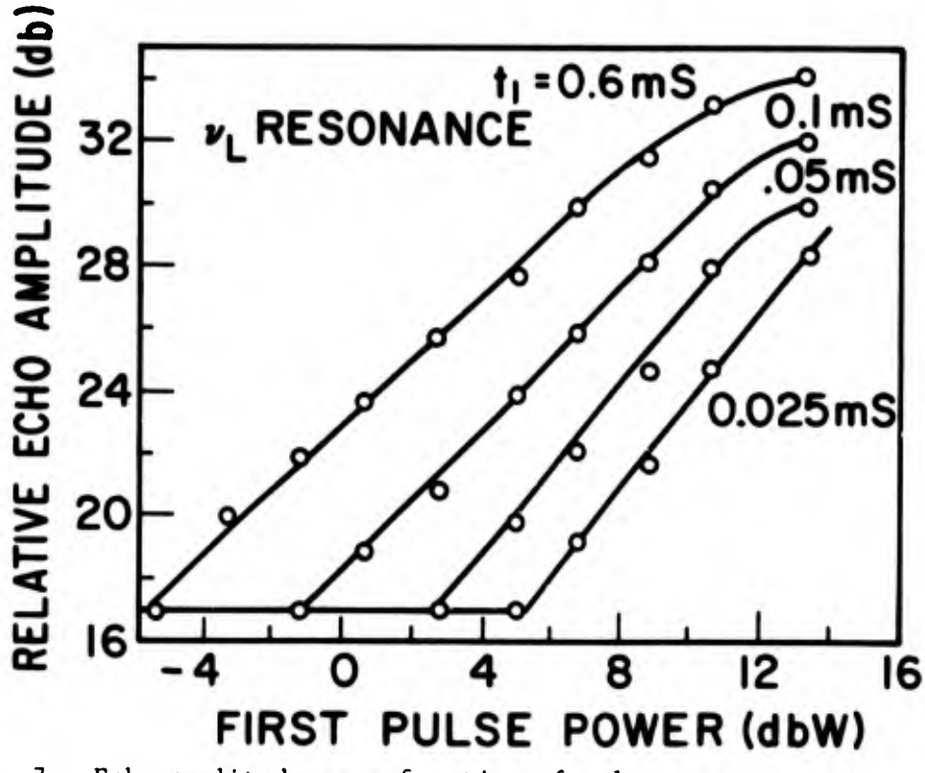


Fig. 7. Echo amplitude as a function of pulse power

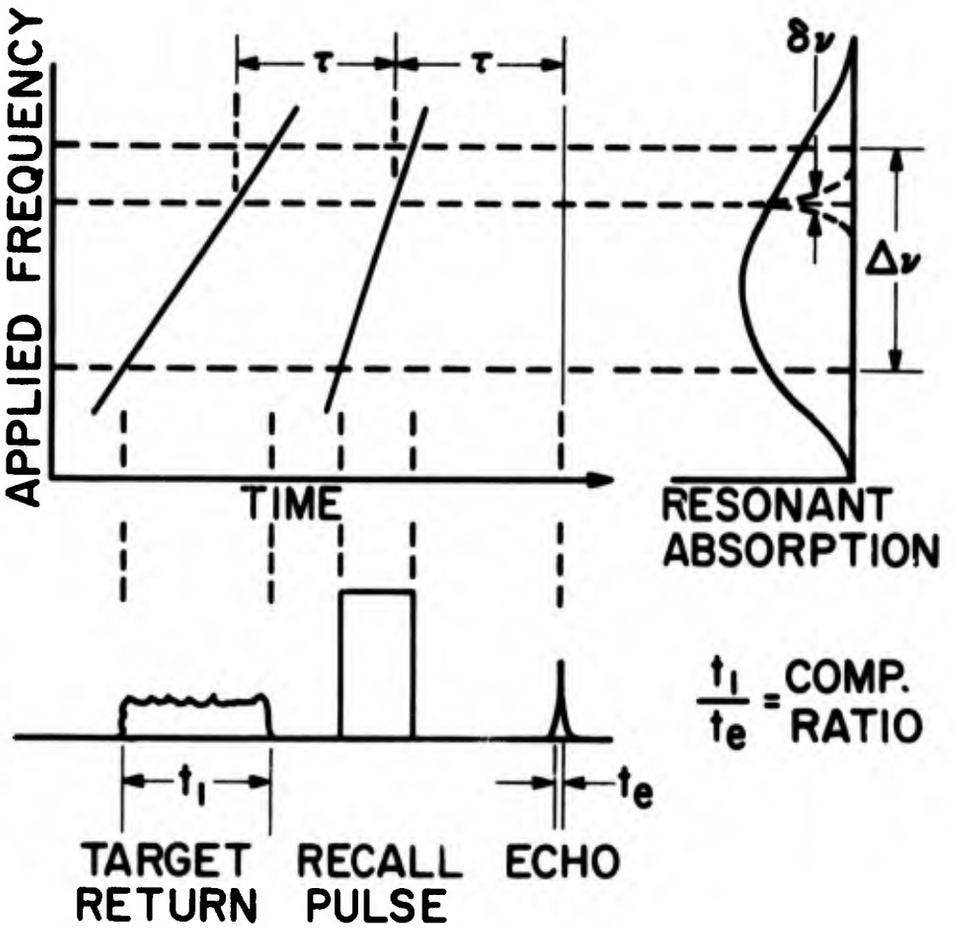


Fig. 8. Chirp pulse compression technique

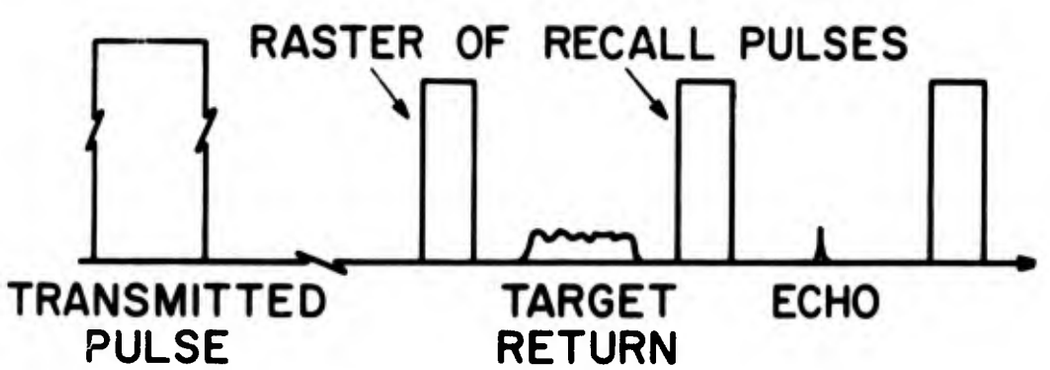


Fig. 9. Pulse sequence for long radar range sensitivity

AD 713 519

THE ROLE OF FRACTURE TOUGHNESS AND RESIDUAL STRESSES IN THE FATIGUE AND FRACTURE BEHAVIOR OF LARGE THICK-WALLED PRESSURE VESSELS

THOMAS E. DAVIDSON, JOSEPH F. THROOP AND ALBERT N. REINER
BENET RESEARCH & ENGINEERING LABORATORIES
WATERVLIET ARSENAL, WATERVLIET, N. Y.

INTRODUCTION

Both fatigue life and ultimate fracture mode become important considerations in any structure that is subjected to repetitive loading. The number of repeated loadings that can be endured is important from the standpoint of operational limitations. Equally important, from the standpoint of safety, is the assurance that any failure will be ductile in nature rather than brittle and catastrophic. Pressure vessels subjected to repeated internal pressurization represent a type of structure in which both of these considerations are especially important. Several examples of catastrophic failures might be cited in which the design and materials selection did not adequately consider these factors.

A series of recent failures in cannon tubes has motivated a critical examination of fatigue and fracture behavior in thick-walled cylinders. In general, fatigue is a process which may be divided into three stages, including crack initiation, cyclic crack propagation and, finally, fracture when the fatigue crack reaches critical size. A cannon tube may be considered as a special type of pressure vessel in which the crack initiation stage occurs quite early and rapidly, mainly as the result of thermal effects upon the bore surface caused by the high propellant temperatures. The aspects of primary concern are, therefore, how rapidly the fatigue cracks propagate and the depth to which they can grow without producing a brittle fracture.

Summarized herein are the results of an investigation into the fatigue and fracture behavior of large thick-walled cylinders identical in configuration to a 175mm cannon tube. Crack growth rates and fatigue life data are presented for materials of three strength levels and different fracture toughness levels. The effects of autofrettage were examined and found to improve the fatigue life

significantly. This improvement in life is shown to be the result of retardation of the fatigue crack growth rate at small crack depths. This observation, along with the relationship between fracture toughness, critical crack depth and fracture mode, is interpreted in terms of recent advancements in the application of fracture mechanics to the case of a cylinder under internal pressure.

The applicability of the results in explanation of the earlier failures of 175mm gun tubes and in formulating the basis for a new design is described and discussed. The results indicate that both a longer fatigue life and a leak-before-fracture type of failure can be attained in these tubes at normal operating pressures.

MATERIALS, SPECIMENS AND PROCEDURES

Materials

The materials utilized in this investigation consisted of cylinders of a modified AISI 4337 steel, all normalized at 1750°F, austenitized at 1550°F and water quenched. The tempering temperature was varied from 1040°F to 1100°F, and some of the cylinders were hydraulically autofrettaged to 50% overstrain. The designation, nominal transverse properties and condition of the four categories of specimens prepared from these materials are as follows:

TABLE I

Designation and Condition	Y.S. ksi	UTS ksi	R.A. %	Charpy Impact ft.lb.(-40°F)	Fracture Toughness ksi $\sqrt{in}$.(-40°F)
A Autofrettaged	149	163	36	23	116*
B Non-autofret.	182	198	31	12	104
C Non-autofret.	161	176	28	15	111*
D Non-autofret.	164	178	45	25	123*

*denotes K_Q

The fracture toughness values were measured using an ASTM chevron-notched three point loaded bending specimen. These were taken from the cylindrical fatigue specimens in the transverse direction. Due to the size limitations imposed by the wall thickness, the results for the lower yield strength materials, A, C and D are apparent fracture toughness, designated K_Q, rather than valid K_{Ic} values. Only the values from Charpy impact and fracture toughness tests measured at -40°F are reported here. Valid room-temperature values of K_{Ic} are expected to be correspondingly higher.

Specimens

The fatigue specimens were cylinders identical in size, diameter ratio and configuration, including internal rifling and an

external keyway, to the origin-of-rifling portion of the 175mm gun tube. The specimen and the pressure sealing system are shown schematically in Fig. 1. The pressure sealing was by a mandrel type packing system, the mandrel serving to provide an "open-end" condition[2] of the specimen.

The autofrettage cylinders for this investigation were overstrained in 120 inch lengths, each of which was subsequently divided into two fatigue specimens. The autofrettage operation was performed with a mandrel-packing system basically the same in principle as that used for the fatigue testing, except with a longer mandrel. The specimens were autofrettaged to the 50% overstrain condition using 120 ksi pressure.

Procedure

Fatigue Testing - The hydraulic pressure cycling was performed by means of an intensifier pumping system at pressure levels of 50 ksi and 60 ksi, using a synthetic instrument oil as the pressure transmitting medium. The cyclic rate was 5 cycles/minute. The inaccuracy in pressure measurement was $\pm$ 2% with a repeatability error of $\pm$ 1/2%. Pressure measurements from a manganin pressure transducer, along with outside surface tangential strain of the specimen by foil type strain gages, were monitored continuously using a 4-channel recorder.

Fatigue Crack Measurement - Fatigue crack depths were measured using an ultrasonic[8] measurement technique wherein crack depth was determined directly from the signal reflections from the crack tip by means of a longitudinal wave probe. The frequency of inspection during the pressure cycling was dictated by the crack depth. Ultrasonic measurements were made at approximately one inch intervals along the length of the specimen. An index ring affixed to the specimen was used to indicate the position of the manually held ultrasonic probe holder during circumferential scanning. The estimated inaccuracy in crack depth measurement was $\pm$ 0.050 inch, and the smallest discernible depth of cracks emanating from the bore was in the range of 0.1 to 0.2 inch depth.

RESULTS AND DISCUSSION

Fatigue Life

The fatigue life test results[6] at the two pressure levels, 50 ksi and 60 ksi, are shown in Table II. The results for the unfired cylinders are plotted in Fig. 2 with maximum hoop stress plotted versus cycles to failure. Failure is considered to be either perforation of the cylinder wall by a crack or else complete fracture of the cylinder by a running crack.

TABLE II

Mat'l Desig.	Spec. No.	P ksi	Cycles to Failure	Critical Depth (in.)	Crack Length (in.)	Fracture Toughness (ksi $\sqrt{in}$) -40°F	72°F
A	1A	50	20,547	3.5	15	116*	--
	4A		23,356	3.5	15	110*	--
	6A		25,127	3.5	15	118*	--
	7A		22,385	3.5	21	114*	--
	9A		21,677	2.7	10	107*	--
	11A		25,815	3.5	16	113*	--
	2A	60	12,220	3.5	15	116*	--
	3A		12,452	3.5	14	122*	--
	5A		13,771	3.5	13	125*	--
	8A		11,653	3.5	15	111*	--
	10A		11,537	2.0	5	101*	--
	12A		12,383	3.5	17	113*	--
B	4N	50	10,086	1.8	6	108	104
	5N		10,630	2.0	6	102	101
	6N		9,565	1.4	7	103	105
	1N	60	6,116	1.25	6	106	114
	2N		6,134	1.6	8	106	115*
	3N		6,877	1.5	8	101	115*
C	7N	50	14,291	2.0	12	114*	--
	8N		13,237	2.1	9	114*	--
	9N		13,389	1.75	10	103*	--
D	412	50	14,738	3.5	--	124*	--
	2108		12,377**	3.5	13	121*	--
	1785		10,033**	3.5	9	122*	--
	1910		7,677**	2.3	19	126*	--

* K_Q

** Includes 400 EFC fired-rounds prior to laboratory cycling.

The fatigue life of the autofrettaged cylinders of the 140-160 ksi yield strength material (A) is considerably greater than that of the non-autofrettaged cylinders of the 170-190 ksi yield strength material (B). At the hoop stress level of 80 ksi, corresponding closely to the nominal Zone 3 firing condition for the 175mm gun tube, the autofrettaged cylinders (A) exhibit a mean fatigue life approximately 2.5 times that of the latter (B) material which was originally used in the 175mm M113 gun tube. At the same stress level the non-autofrettaged cylinders of the lower strength materials (C) and (D) exhibited greater fatigue lives than the high strength (B) material. The 80 ksi stress corresponds to P = 50 ksi.

Comparison of these 175mm cylinder results with results from an earlier investigation of 105mm gun tubes(4), shown in curves (E) and (F) of Fig. 2, reveals that the (C) and (D) materials exhibited fatigue lives corresponding quite closely with those for the 105mm non-autofrettaged (E) tubes. These had similar low strength and high toughness levels and had the same diameter ratio as the (C) and (D) specimens. Furthermore, the fatigue life for the 50% autofrettaged 175mm cylinders of the 140-160 ksi yield strength (A) material is greater than that of the non-autofrettaged 105mm specimens (E) and is exceeded by that of the 100% autofrettaged 105mm specimens (F). This demonstrates the relative effects of the compressive residual bore stress resulting from partial (50%) overstrain and complete (100%) overstrain in autofrettaged cylinders as compared to the non-autofrettaged cylinders.

Macroscopic Fracture Analysis

The fracture surface of the autofrettaged specimens of the low strength material (A) show that the critical crack effectively intersects the outside surface of the cylinder, as illustrated by the final light elliptical marking on specimen 11A in Fig. 3. Thus, the critical crack depth is approximately equal to, or greater than, the wall thickness. The low strength non-autofrettaged material (D) exhibited similar behavior. In such cases the fatigue crack simply perforates the outside surface, Spec. 11A, Fig. 3, a highly desirable type of behavior since only leakage of the internal pressure will occur rather than a running crack that could result in fragmentation.

In contrast, the non-autofrettaged cylinders of high yield strength material (B) exhibited considerably different fracture behavior. As illustrated in specimen 5N in Fig. 3, the critical crack size is smaller than the wall thickness and fast fracture occurs, resulting in a crack that runs to the end of the specimen. As will be discussed, similar behavior was also observed in the non-autofrettaged material (C).

FRACTURE MECHANICS ANALYSIS

Stress Intensity Factor

In a cylinder containing internal pressure P, the stress intensity factor for a crack emanating from the bore may be expressed in the form

$$K = K_p + (\sigma_{max}) Y \sqrt{b} \qquad \ldots\ldots (1)$$

where Kp is the contribution to the stress intensity factor due to the pressure acting within the crack cavity itself, b is the crack depth, Y is a non-linear function of the relative crack depth (b/B)

and the crack shape ratio (b/2a), and σ_{max} is the tangential stress in the cylinder wall. The latter may be written

$$\sigma_{max} = \sigma_{tP} + \sigma_{tR} \quad \ldots\ldots (2)$$

in which σ_{tP} is the tangential stress associated with the maximum pressure of the loading cycle and σ_{tR} is any residual stress present. In this discussion they are the nominal local stresses at the depth b from the bore into the cylinder wall, and the function Y must be expressed accordingly.

In an autofrettaged cylinder the residual stress σ_{tR} is compression near the bore and changes to tension as the outside surface is approached, varying through the cylinder wall as shown by curve A of Fig. 4. This figure shows the residual stress gradient for the 50% autofrettaged condition used in this investigation. Two curves are shown, the solid curve being the theoretical residual stress based on the Tresca yield criterion (1), and the dashed curve A is the residual stress corrected for the Bauschinger effect(3).

In a non-autofrettaged cylinder the value of σ_{tR} is zero and $\sigma_{max} = \sigma_{tP}$, which is maximum at the bore and varies through the cylinder wall. In these 175mm cylinders at a pressure of 50 ksi, it has the gradient shown by curve B in Fig. 4. In comparison, the value of the maximum stress for the 50% autofrettaged cylinders at the same pressure is given by Equation (2) and varies as shown by curve C in Fig. 4. It is evident that there is considerable difference between curve C for σ_{max} in the partially autofrettaged condition and curve B for σ_{max} in the non-autofrettaged condition, especially at small values of crack depth b.

Critical Crack Depth

Fracture mechanics supplies the concept that a crack becomes a running crack when it reaches a critical size at which the stress intensity factor K becomes equal to the fracture toughness K_{Ic} of the material. Thus, substitution of K_{Ic} for K in Equation (1) yields an analytical relationship between critical crack depth (b_c) and fracture toughness. Four curves representing such a solution for non-autofrettaged cylinders are shown in Fig. 5. They represent the critical crack depth versus fracture toughness for four values of crack shape ratio b/2a in the 175mm cylinders at a pressure of 50 ksi.

Experimental results of measured critical crack depth versus fracture toughness are also shown by data points in Fig. 5. The solid points are data from a previous 175mm tube study(5). The open symbols are for non-autofrettaged high strength (B) cylinders tested in this investigation. The agreement between these two sets of data and the analytical curves is quite good. The other three

categories of specimens (A, C and D) are not included in Fig. 5 since, as previously mentioned, it was not possible to obtain valid K_{Ic} values for these materials.

The fracture mechanics solution shown in Fig. 5 demonstrates the role of fracture toughness in controlling critical crack depth and fracture mode. It can also be used to explain, at least qualitatively, the differences among the fracture results of the specimens contained in this current work and summarized in Table II.

In all but 2 cases, the low strength autofrettaged specimens (A) failed by the fatigue crack perforating the outside surface (critical crack depth $\geq$ 3.5 inch wall thickness). Similarly, the low strength non-autofrettaged specimens (D) also failed by perforation, the only exception being Spec. No. 1910 which failed too close to its end to permit perforation. By comparison, the high strength non-autofrettaged (B) and low strength non-autofrettaged (C) specimens had critical crack depths substantially less than the wall thickness. The reason for this behavior is explained on the basis of the fracture toughness of the materials. Materials (A) and (D) had considerably higher toughness than materials (B) and (C). Materials (A) and (D) would have indicated even higher toughness were it possible to obtain valid values. The smaller critical crack depth of Nos. 9A and 10A, as compared to the other materials of the same class (A), is a manifestation of their low fracture toughness.

Cyclic Crack Propagation

The applications of fracture mechanics to the problem of fatigue crack growth have suggested that cyclic crack propagation occurs at a rate proportional to some power of the cyclic stress intensity factor. A study of similar steels(7) has indicated that the cyclic crack propagation rate (db/dN) may be expressed in terms of the material properties and the stress intensity factor in the form

$$\frac{db}{dN} = \frac{C\,(\Delta K)^m}{E\,S_y\,K_{Ic}} \qquad \ldots\ldots\ldots (3)$$

where E is the elastic modulus, S_y is the tensile yield strength and K_{Ic} is the fracture toughness of the material. The exponent m has some value between 2 and 4 depending upon the material, and C is a material constant.

In cyclic fatigue crack growth the ΔK in Equation (3) is the change in K from minimum to maximum load. For a pressurized cylinder ΔK is the change in K corresponding to the change in tangential stress at a given crack depth, which in the case of a non-autofrettaged cylinder is simply equal to K as given in Equation (1) with $\overline{\sigma_{tR}}$ equal to zero. In an autofrettaged cylinder the value of K could be negative at zero pressure because of the compressive

residual stress $\overline{\sigma_{tR}}$ at the bore. This would be meaningless in regard to crack growth, however, because K can only apply to a crack that is "open". Therefore, in the autofrettaged cylinders ΔK is taken as the change in K corresponding to the change in tangential stress from zero to maximum during the loading cycle.

Substituting Equations (1) and (2) into (3) yields for the crack propagation rate:

$$\frac{db}{dN} = \frac{C}{E\, S_y\, K_{Ic}} \left[K_P + (\overline{\sigma_{tP}} + \overline{\sigma_{tR}})\, Y \sqrt{b} \right]^m \qquad (4)$$

For comparison of crack propagation rates in autofrettaged versus non-autofrettaged cylinders it may be assumed that for a given crack shape the same values of Kp and Y apply to both cases. Therefore, the difference in crack rates of the two cases depends directly upon $\overline{\sigma_{tR}}$ in Equation (4), all other things being the same.

As shown by curve A in Fig. 4 the residual stress $\overline{\sigma_{tR}}$ varies from compression at small values of (b) to tension at values of (b/B) greater than 0.37. Hence $\overline{\sigma_{tR}}$ is initially negative in Equation (4). Similarly, curve C in Fig. 4 shows that $\overline{\sigma_{max}}$ is less than that for the non-autofrettaged cylinder, curve B, up to (b/B) = 0.37, beyond which it is greater than curve B. The initial stress intensity factor and the early crack propagation rate for these 50% autofrettaged cylinders at 50 ksi pressure should, therefore, be less and should increase much less rapidly with increasing crack depth than in non-autofrettaged cylinders of the same mechanical properties at the same pressure level.

RELATIONSHIP OF RESULTS TO GUN TUBE BEHAVIOR

In Fig. 6 the crack rate, determined from ultrasonic crack depth measurements, is plotted versus the total cycles at 50 ksi pressure level for representative cylinders of each material. In these curves the lowest crack rate shown for each specimen is at $N_{0.1}$ for a crack depth of 0.1 inch. The upper limit was arbitrarily plotted as 2/1000 inch per cycle. Beyond this very little of the total fatigue life, N, remains. Taking $(N-N_{0.1})$ gives the following estimated firing life under Zone 3 rounds at 50 ksi:

A	Autofrettaged, 50%	140-160 ksi Y.S.	7815 Rounds
B	Non-autofrettaged	170-190 ksi Y.S.	3286 Rounds
C	Non-autofrettaged	140-160 ksi Y.S.	5091 Rounds
D	Non-autofrettaged	160-180 ksi Y.S.	7738 Rounds

As indicated by the critical crack depths in Table II, gun tubes of the (B) and (C) material may be expected to have brittle catastrophic fractures while the (A) and (D) materials may sustain perforation and leak before fracture.

Fired 175mm M113 tubes fabricated of the high strength non-autofrettaged material (B) demonstrated(5) a mean fatigue life of

4000 rounds plus cycles to failure and a brittle fracture mode. The new 175mm M113E1 tube, the design of which was based upon the low strength autofrettaged material (A), exhibited a mean fatigue life (origin of rifling) of 10,800 rounds plus cycles with a ductile fracture mode. Considering the conservative assumption ($N-N_{0.1}$), this represents quite good agreement with the estimated firing lives.

CONCLUSIONS

1. Considerable enhancement in fatigue life of thick-wall pressure vessels can be achieved through the use of autofrettage. The improvement in life is a manifestation of the reduction of the cyclic crack propagation rate resulting from the presence of compressive residual stresses.

2. Fracture toughness plays a dual role in the fatigue and fracture behavior of pressurized cylinders. First, the critical crack depth, which controls the fracture mode, increases with increased toughness. Second, increased toughness decreases the cyclic crack propagation rate. Both of these effects result in an increase in fatigue life.

REFERENCES

(1) Davidson, T. E., Barton, C. S., Kendall, D. P. and Reiner, A. N. "The Overstrain of Thick-Walled Cylinders of Intermediate Diameter Ratio". Proceedings First International Congress of Experimental Mechanics, December 1961

(2) Davidson, T. E., Reiner, A. N., and Eisenstadt, R., "The Fatigue Characteristics of Open-End Thick-Walled Cylinders Under Cyclic Internal Pressure". Journal of Basic Engineering, ASME, Vol 85, December 1963

(3) Milligan, R. V., Koo, W. M. and Davidson, T. E., "The Bauschinger Effect in High Strength Steels". Journal of Basic Engineering, ASME, Vol 88, June 1966

(4) Austin, B. A., Reiner, A. N. and Davidson, T. E., "Low Cycle Fatigue Strength of Thick-Walled Pressure Vessels". Journal of Institute of Mechanical Engineers, April 1968

(5) Davidson, T. E., Reiner, A. N., Throop, J. F. and Nolan, C. J., "Fatigue and Fracture Analysis of the 175mm M113 Gun Tube", Watervliet Arsenal Report No. WVT-6822, November 1968

(6) Davidson, T. E., Throop, J. F., Reiner, A. N., Austin, B. A., "Analysis of the Effect of Autofrettage on the Fatigue Life Characteristics of the 175mm M113 Gun Tube". Watervliet Arsenal Report No. WVT-6901, January 1969

(7) Throop, J. F. and Miller, G. A., "Optimum Fatigue Crack Resistance", Watervliet Arsenal Report No. WVT-7006, January 1970

(8) Miller, J. J., "Ultrasonic Measurement of Crack Depth in Thick-Walled Cylinders", Watervliet Arsenal Report No. WVT-7017, February 1970

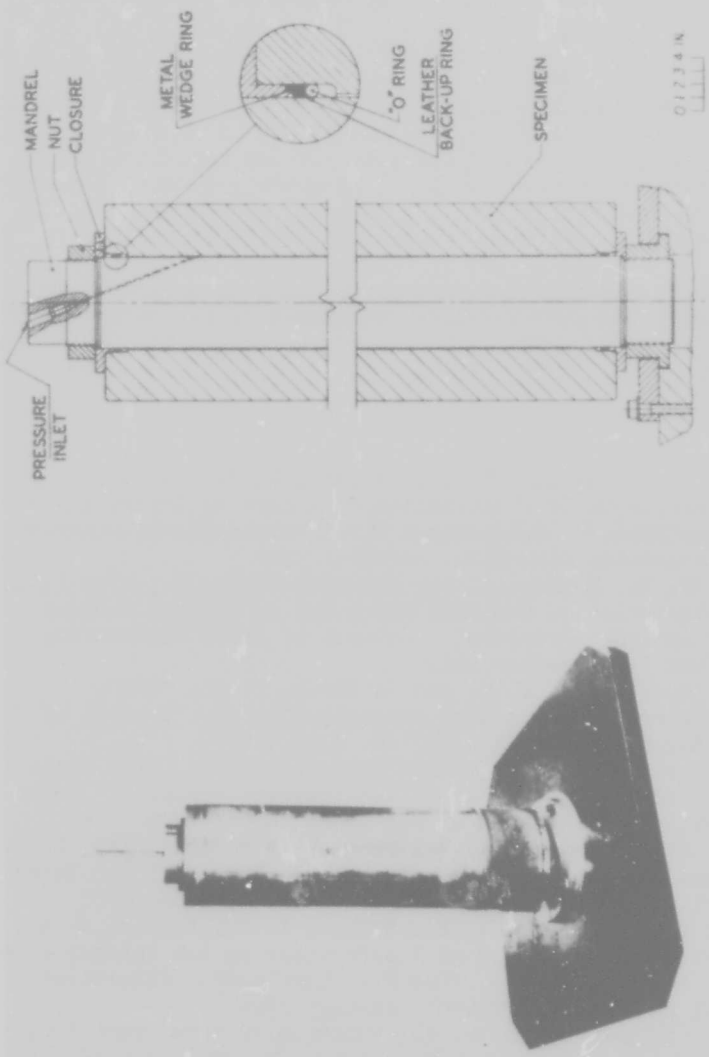


FIGURE 1 175MM FATIGUE SPECIMEN AND MANDREL-PACKING SYSTEM

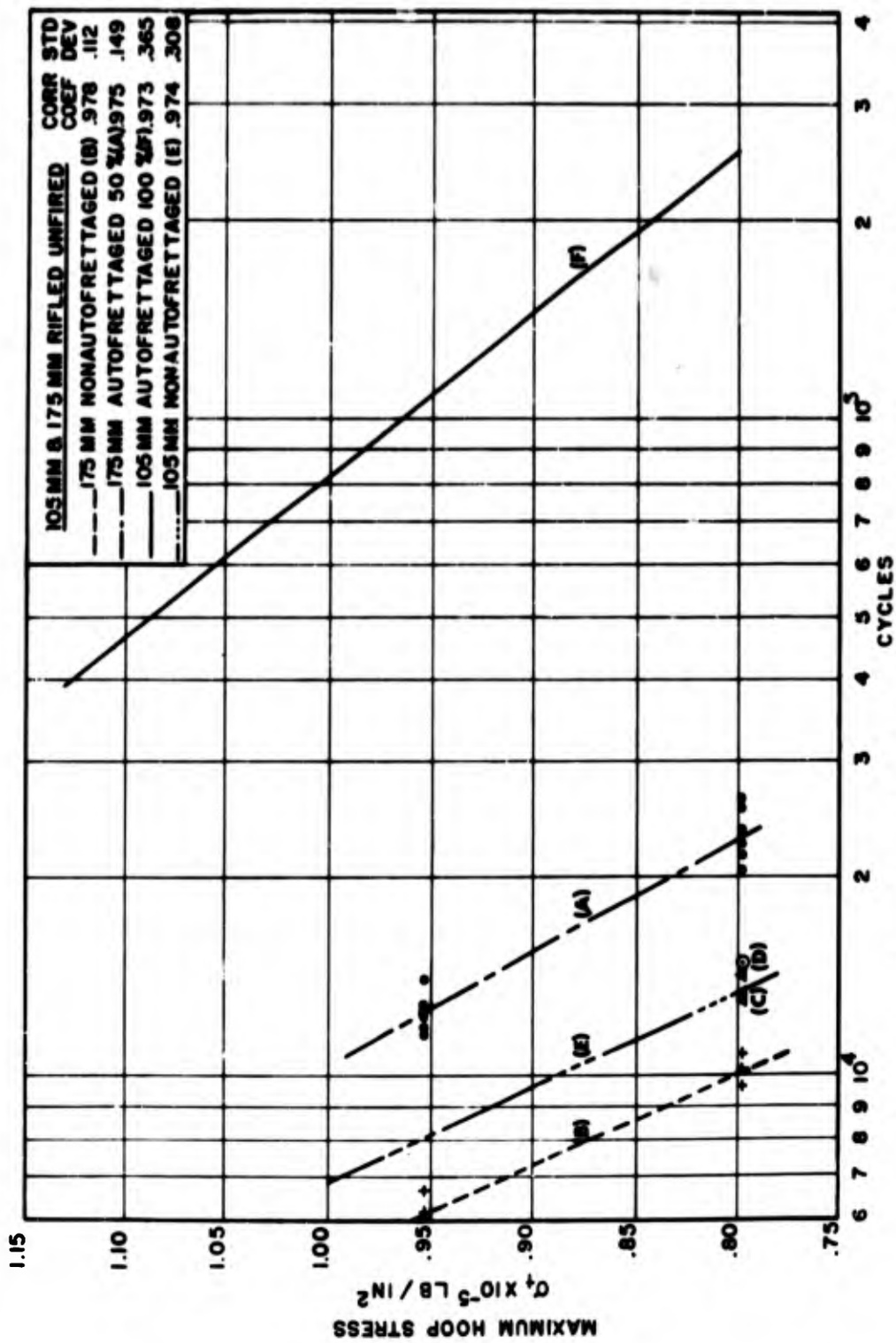


FIGURE 2 MAX HOOP STRESS VS CYCLES TO FAILURE 105 MM & 175 MM FATIGUE RESULTS

(A) - 140-160 KSI AUTOFRETTAGED

(B) - 170-190 NON-AUTOFRETTAGED

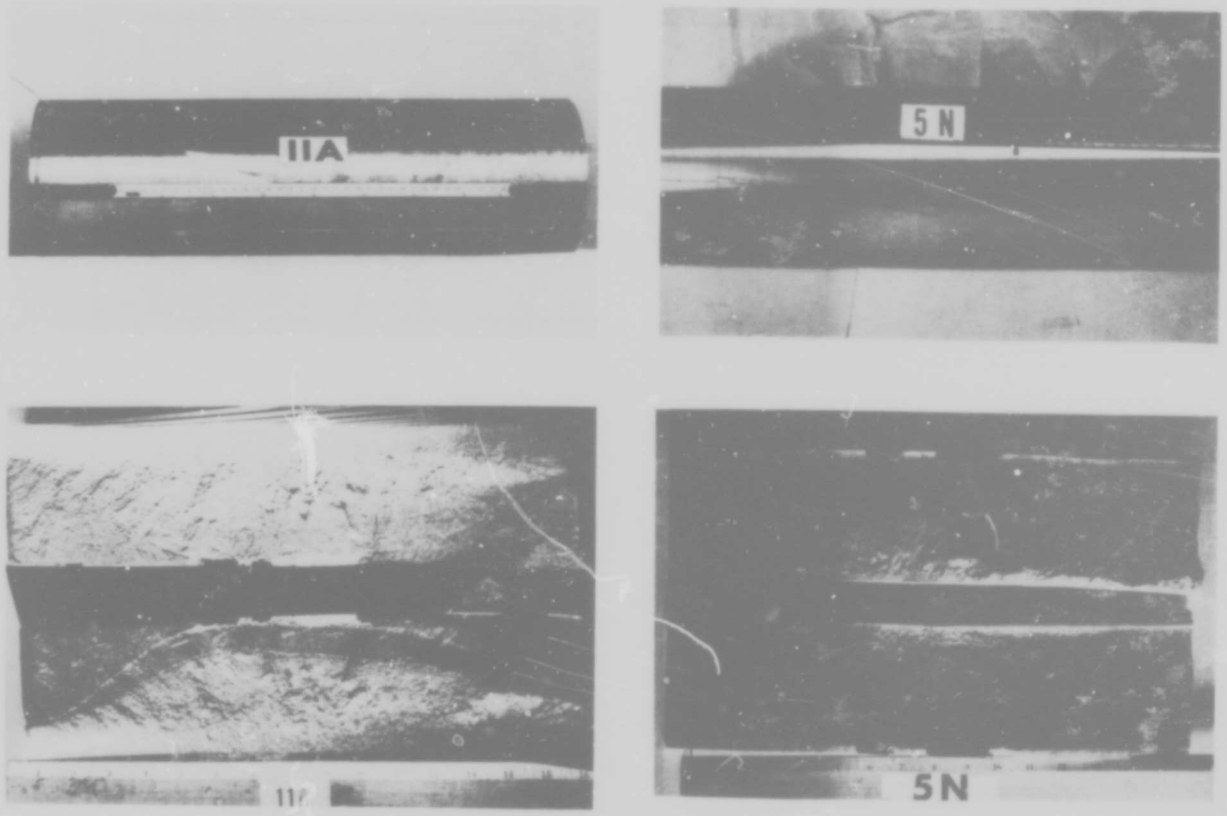


FIGURE 3 MACROSCOPIC FRACTURE APPEARANCE P=50 KSI

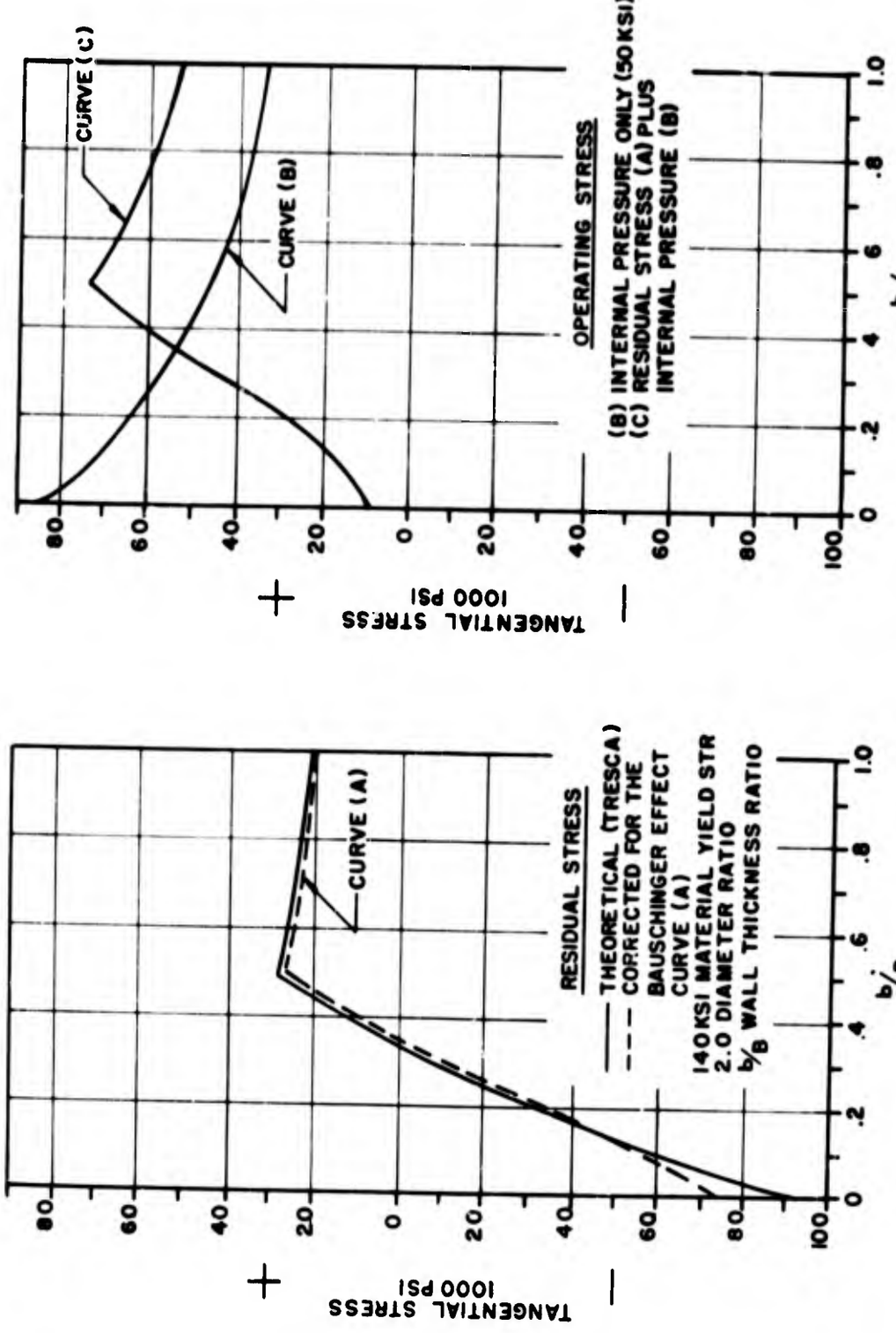


FIGURE 4 RESIDUAL AND OPERATING STRESSES 175 MM AUTOFRETTAGED TUBE

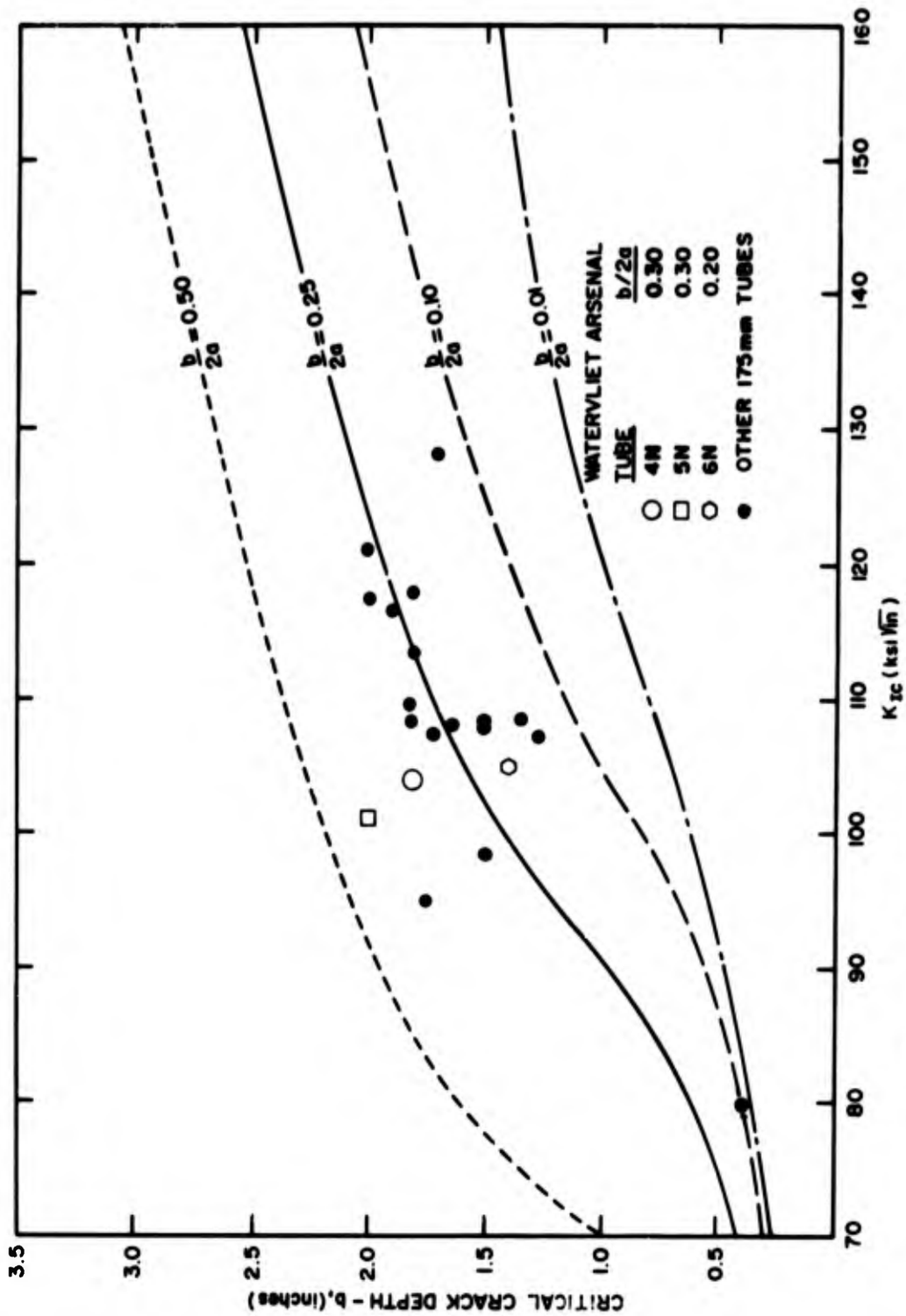


FIGURE 5 CRACK DEPTH vs FRACTURE TOUGHNESS FOR 175mm TUBE - 50 ksi

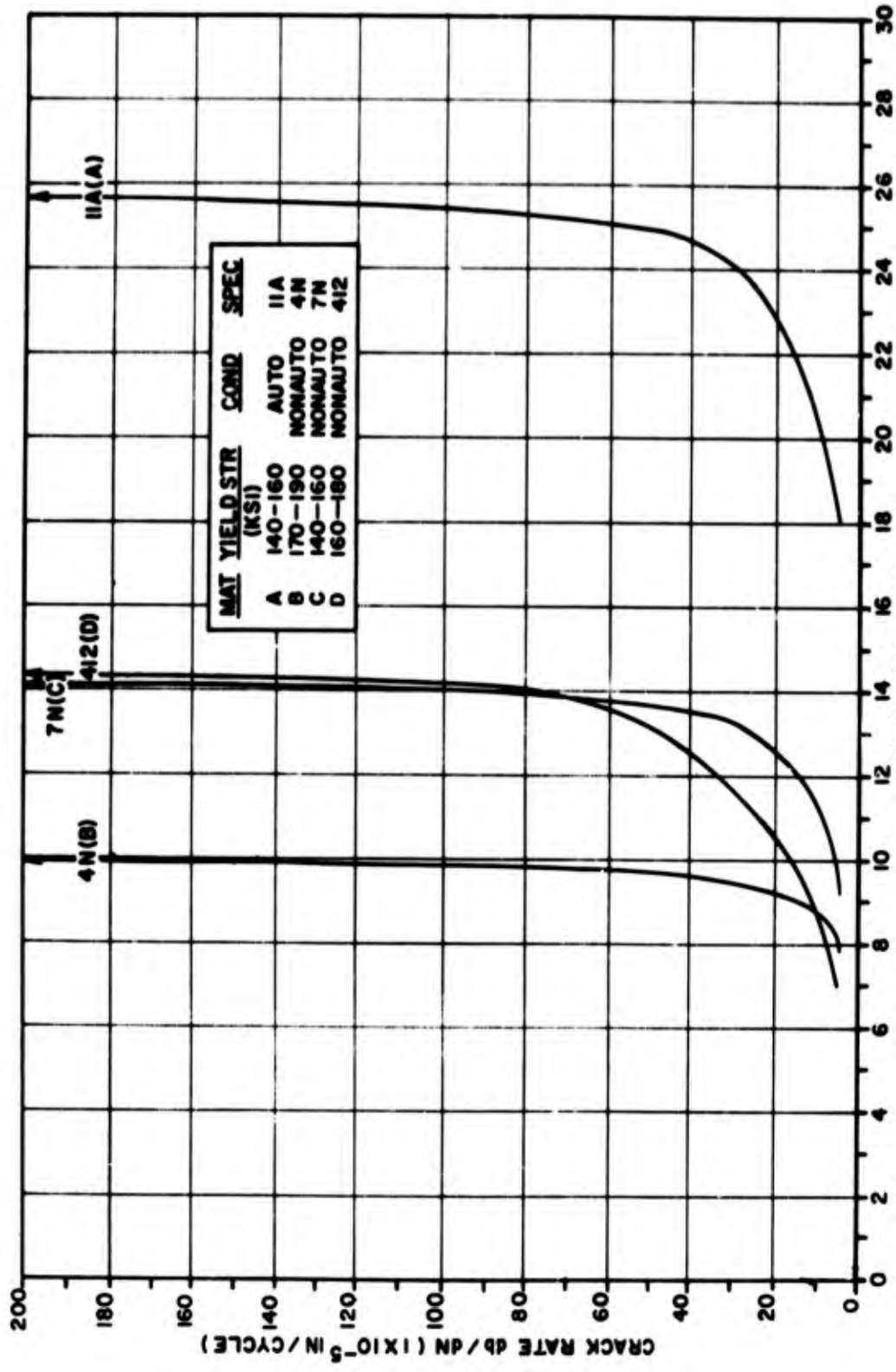


FIGURE 6 CRACK RATE VS TOTAL CYCLES - P=50KSI

AD 713 520

THE INTERNAL AND EXTERNAL FLOW FIELD ASSOCIATED WITH PARACHUTES DURING INFLATION

GREGORY C. DE SANTIS
U. S. ARMY NATICK LABORATORIES
NATICK, MASSACHUSETTS

1. INTRODUCTION

The role of the parachute in supply logistics is vital to the successful accomplishment of many military objectives. In September 1968, the leading researchers in parachute technology, gathered at El Centro, California for the American Institute of Aeronautics and Astronautics, 2nd Aerodynamic Deceleration Conference, discussed the critical problems in parachute research that remain unsolved. One result of this meeting was an article titled "Technical Voids in Aerodynamic Deceleration R&D" which appeared in the December 1968 issue of "Aeronautics and Astronautics". The article delineated a number of major problem areas in parachute research such as a definition of parachute geometry during inflation, determination of parachute filling time, improvement in model scaling and knowledge of the flow field surrounding the parachute. Relative to the latter problem area, the article states, ". . . the flow field surrounding a flexible device is poorly understood, and that (field) surrounding an inflating parachute in a transient velocity field is strictly a matter of conjecture".

Up to the time of the conference the only research into parachute flow fields involved mounting a conventional survey rake at the inlet and vent openings of a parachute model. The resulting interference caused by the rake changed the flow and affected the readings. Realizing that a major breakthrough in the flow field knowledge had not been achieved to date using conventional pressure-sensitive devices, it was decided early in the flow field study to investigate other methods of flow measurement. The result of a preliminary investigation was the decision to use the hot-wire anemometer because of its low flow interference and excellent velocity sensitivity to determine the characteristic velocities of the flow surrounding a parachute. In a rough form, the hot-wire anemometer was first used over 70 years ago to measure wind

velocity, but it has remained essentially a laboratory curiosity until just a few years ago when it began to be produced commercially. The problem of irrelevant readings caused by temperature sensitivity which had prevented wide spread use of the hot-wire anemometer in aerodynamic research was eliminated by conducting the flow field tests in the temperature controlled wind tunnel at Natick Laboratories. The last barrier to the application of the hot-wire had been removed and now the potential existed for a major impact on parachute flow field technology.

The procedures used in this study involved the use of the hot-wire anemometer to obtain data on the flow field associated with an inflating parachute. Seven models simulating various stages of inflation of the C-9 parachute were fabricated and tested in a specially constructed test section where the temperature could be held uniform. Using this method, it was possible to accurately measure the internal and external flow surrounding the canopy. Some possible applications of the data to full-scale parachute systems are presented.

II. THE HOT-WIRE ANEMOMETER

The hot-wire anemometer is capable of doing all the things conventional survey rakes can do with the additional advantages that it has the ability to measure extremely low velocities and has such a small cross-sectional area that the interference caused in the flow is negligible. Under good test conditions, it is possible to measure pressures of 0.02 lb/sq ft accurately with a pitot tube. This corresponds to a velocity of approximately 3 ft/sec. A hot-wire system is capable of measuring velocities of 0.1 ft/sec under the same conditions. The ability to measure such low velocities is especially useful in measuring the velocity immediately behind the model. A pitot tube survey rake would have had to be mounted at least 10 canopy diameters downstream from the model before the dynamic pressure would be high enough to be measured accurately. The hot-wire probes used in this study could be mounted either immediately behind the model or at some point downstream and not have their ability to measure the velocity affected.

III. WIND TUNNEL MODELS

A solid flat circular parachute model with 28 gores and a nominal diameter of 16 inches was chosen for the test because it is similar in design to a full size C-9 parachute. The C-9 canopy is also a solid flat circular parachute with 28 gores but with a nominal diameter of 28 feet. The use of models which are scaled-down versions of the full-size C-9 canopy made it possible to compare the wind tunnel data with the full size data and determine the validity of the results.

The models used for the study were designed to represent a 28 gore solid flat parachute at seven stages of inflation. Wire frames were covered with 1.1 oz/sq yd rip-stop nylon parachute fabric attached to the structure along the seams so the material would take the required inflated shape in the wind tunnel. To retain the inlet gore configuration required for the simulation, fine copper wire was inserted into the leading edge of each gore. Using these construction techniques, it was possible to fabricate the seven models used to simulate the various inflation stages of the C-9 parachute.

IV. WIND TUNNEL TEST FACILITY

The wind tunnel used to conduct the flow field study is located at the Climatic Research Laboratory of the US Army Natick Laboratories, Natick, Mass. The test facility is equipped with two wind tunnels, an Arctic Wind Tunnel and a Tropic Wind Tunnel, which are used to test military equipment and personnel under extreme climatic conditions. The facility was used for the study because it was possible to maintain uniform control over the temperature to $\pm$ 1°F throughout the test, thus eliminating the problem of temperature sensitivity inherent in hot-wire anemometers. The elimination of this major problem area resulted in highly accurate and uniform readings of the local velocities over a range of flow conditions.

To control the upstream and downstream velocities of the flow, a special 3 ft x 3 ft test section was constructed for the tests (Figure 1). A large flexible panel was fabricated which closed off the 8 1/2 x 15 ft test section with an opening into a 6 ft x 6 ft contraction section. The large reduction ratio created with this design increased the degree of control over the velocity of the flow prior to entering the 3 ft x 3 ft test section. This technique of controlling the flow resulted in repeatable flow conditions with velocities in excess of 110 ft/sec having low turbulence levels within the flow.

The free stream velocity, U_{oo}, was obtained from an upstream pitot tube, converted to an equivalent hot-wire current, I_{oo}, and used in the computation of the local velocity ratios, U/U_{oo}.

Since the models were mounted on the wind tunnel centerline the probes were adjusted to move radially outward from the test section centerline measuring the hot-wire current, I, at 1/2 inch intervals. With the problem of temperature compensation removed, the complicated equations for velocity found in Reference 5 can be greatly simplified. It was now possible to express the local velocity ratio, U/U_{oo}, as a function of the zero velocity current I_o, the free-stream velocity current, I_{oo}, and the local velocity current I.

$$\frac{v}{v_\infty} = \left[\frac{I^2 - I_o^2}{I_\infty^2 - I_o^2}\right]^{\frac{1}{2}} \qquad (1)$$

In this manner, it was possible to obtain the velocity ratios for the flow entering the inlet and gores of the canopy, as well as the flow exiting through the vent and fabric. By making holes in the fabric it was also possible to insert the hot-wire probes into the models and obtain the velocity ratios for the flow inside the canopies.

V. TEST RESULTS AND DISCUSSION

To make the results applicable to a number of solid flat circular canopies, the data was plotted in the form of dimensionless velocity ratios versus dimensionless radial distances from the centerline (Figures 2 and 3). Except for Models 1 and 2, each plot contains four curves. Two curves represent the velocity ratios at the skirt and vent planes and are continuous to the point where the velocity reaches free-stream or in terms of velocity ratios, unity. The other two curves represent the velocity ratios of the flow inside the canopies at Stations 1 and 2. Since these are inside the canopy, they are bounded by the radial distance to the canopy fabric at the given station. Below the plots are cross-sectional outlines of the specific stages of inflation represented by the parachute models. This method of presentation makes it possible to relate a specific data point with a physical feature on the model such as vent opening or edges of gores. As shown in Figures 2 and 3, the solid line on the outline of the model is the outer perimeter of the gore while the dashed line is the base of the gore or the seam where the suspension line is attached to the parachute.

A. Flow Into the Canopy

The flow into the canopy models has a marked transition from Model 1 to Model 7. Model 1 has a distinct core of higher velocity flow along the centerline which reduces to a low velocity region of flow in the skirt plane at the base of the gores. Models 2 and 3 show similar core flow of about the same magnitude but more uniformly distributed. They also have a lower velocity ratio at the base of the gores than the centerline value. It appears that during the early phase of inflation the low velocity region is in the form of an annulus and located at the base of the gores. Model 4 begins to show a transition in the flow into the canopy. Here, the velocity across the inlet remains uniform with a slight rise in velocity on the suspension line circle. Models 5, 6, and 7 continue with similar smooth transition curves. All the models have a rise in velocity across the inlet of the gore. In models 1 through 4 the rise is sharp and begins at the base of the gores. The last three models show a smooth change in velocity from the base of the gores

and out across the gores.

B. Flow Out of the Canopy

All of the models have a high velocity flow through the canopy vent. The first four models have a high velocity flow along the centerline. In the last three models, the inlet area of the parachute is much greater than the vent area, causing the internal pressure to be more uniform, and the velocity of the flow shows a more uniform region across the vent opening.

Behind the canopy, the velocity of the flow decreases to a very low level. The average velocity ratio of the flow passing through the fabric of the seven models was found to be 0.040. All of the models experience a distinct flow separation at station 1. The higher velocity ratios shown behind the canopy for Models 1 and 2 appear to be caused by the higher velocity flow coming around the outside of the canopy due to this separation of flow. From Model 3 to Model 7 the ratio remains uniform behind the canopy with a rise in the velocity at the radial distance of Station 1 from the centerline.

C. Flow Inside the Canopy

The small internal dimensions of Models 1 and 2 made it impossible to obtain data on their internal flow. The first model in which the velocity was studied was Model 3. Here the velocity curve shows a low velocity on the centerline at Station 1 with a steady increase in velocity moving outward from the centerline. A transition is more noticeable from Model 3 to Model 5 in which the low velocity core flow at Station 1 broadens until the centerline velocity increases slightly in Model 6. Model 7 does not have a Station 1 because it would coincide with the skirt of the canopy. Therefore, the flow velocities of Station 1 in Model 6 would transition to the inlet velocity curve of Model 7.

The velocity ratios for Station 2 all seem to show similar curves. That is, they have higher velocity ratios at the centerline and lower values away from the centerline. At Station 2, the flow through the vent is beginning to form and the flow steadily increases in velocity until it exits through the vent.

D. Summary of the Data

The data presented in Figures 2 and 3 can be converted into a more useful form by taking average values of the velocity ratios and relating them to the inflating canopy shape or more exactly the parachute inlet diameter ratios. By averaging the velocity ratios for the flow into the canopy at the skirt plane it is possible to get the two curves shown in Figure 4. The lower curve represents the average velocity ratio across the circle formed

by the base of the gores, while the upper curve is the average velocity ratio across a single gore. Figure 4 indicates that for a parachute with a small number of gores (12 to 28) the dominant curve affecting the early phases of inflation would be the upper one because most of the air filling the canopy would enter through the gore inlets. For parachutes with a large number of gores (64 to 160) the upper curve would only be effective in initiating the inflation process but the greatest amount of air would enter through the main inlet of the parachute.

In a similar fashion, the flow existing the canopy at the vent plane can be simplified by plotting the average velocity ratios through vent and canopy fabric (Figure 5). The dashed portion of the lower curve represents the average value for the flow as measured behind the canopy from the test data, but since this flow was affected by separated flow coming around the side of the canopy, the average velocity ratio through the fabric is assumed to be about 0.040.

VI. APPLICATION OF RESULTS

The results of the study have many implications for future research in parachute technology. For one thing, the data from the scale-model tests can be used to predict values for full-scale systems. In discussing the applications of the model data, first to be considered is the concept of dimensionless time versus projected area during inflation, secondly, values for the filling times for two full-size parachutes are predicted.

A. Dimensionless Time Versus Projected Area

The concept of dimensionless time is that it is possible to express the times during the inflation of a parachute as a function of the instantaneous time, t, and the total filling time, t_f.

$$T = \frac{t}{t_f} \qquad (2)$$

Therefore, the value of the dimensionless time, T, ranges between zero and unity. One method for determining T as a function of the projected area of the parachute during inflation is covered in Reference 2. The assumptions used in this method require that unknown factors take on linear relationships with the various parameters. The solution to the system of equations which are developed in this fashion is by a trial and error technique. This type of iterative solution occurs because the free-stream velocity, U_{oo}, varies during inflation as a function of the drag area and mass of the parachute system.

Now that the flow field data obtained in this study is available, it is possible to determine the filling time, t_f, as a

function of the change in volume and volumetric flow rate, Q, into and out of the parachute. Using the average velocity ratios from figures 4 and 5 it is possible to determine the volumetric flow rate into the parachutes, Q_{in}, through the inlet and gores of the canopy, as well as the flow out of the canopy through the fabric and vent, Q_{out}.

$$Q_{in} = U_\infty \left[\left(\frac{U}{U_\infty} \right)_{INLET} \left(\frac{\pi}{4} D_i^2 \right) + \left(\frac{U}{U_\infty} \right)_{GORE} \left(\frac{\pi^2 D_i D_o}{2n} - \frac{\pi^3 D_i^2}{8n} \right) \right] \tag{3}$$

$$Q_{out} = U_\infty \left[\left(\frac{U}{U_\infty} \right)_{VENT} \left(\frac{\pi}{4} D_v^2 \right) + \left(\frac{U}{U_\infty} \right)_{FABRIC} \left(\frac{\pi}{4n} \right) \left(\frac{6 D_i D_o}{4 D_o + \pi D_i} \right)^2 \sqrt{4\pi^2 + 9^2} \right] \tag{4}$$

Where "Do" is the nominal diameter of the parachute, "Di" is the instantaneous diameter of the parachute, "Dv" is the vent diameter and "n" is the number of gores. This results in a simplified method of determining the time between the various stages of inflation and the total inflation time.

$$\Delta t = \frac{\Delta (Volume)}{Q_{in} - Q_{out}} \tag{5}$$

$$t_f = \sum \Delta t \tag{6}$$

Using this method, a full size C-9 parachute system was analyzed for two different snatch velocities (the velocity at which parachute initially begins to open).

SYSTEM CHARACTERISTICS

Parachute: 28 ft, 28 gore C-9 canopy
Suspended Weight: 200 lb
Sea Level Standard Day Conditions
Snatch Velocity: 227 ft/sec and 103 ft/sec

The system described above was chosen for the analysis because it is similar to those studied by Berndt and De Weese (Reference 3). Figure 6 shows the comparison between the computed values of projected parachute area ratios, S/S_o, versus dimensionless time obtained from this flow field study and the average curve obtained from over thirty full scale test drops made by Berndt and De Weese. The close correlation among the three curves indicates that the results of the scale model wind tunnel study can be applied to full-scale systems. This is significant because extrapolation from model parachute data has produced results very close to those obtained from the full-scale system.

B. Predicted Parachute Filling Time

Using the same technique described above an attempt was made to predict the filling time for a C-9 parachute and a G-11A parachute over a range of snatch velocities and suspended weights. The G-11A is also a solid flat circular parachute but it is 100 ft in diameter and contains 120 gores. Both canopies were analyzed over a range of snatch velocities between 150 ft/sec and 250 ft/sec. The C-9 had suspended loads from 160 lbs to 320 lbs while the G-11A loads ranged from 2200 lbs to 4400 lbs. The results of the C-9 predictions are shown in Figure 7 along with some of the data points obtained by Berndt-De Weese (Reference 3). The predicted filling times appear to be lower than the Berndt-De Weese data, but this is probably caused by the numerical model of the inflating parachute used to compute the curves. The model assumes the ideal case in which all of the gores are equally inflated. In the real case, some of the parachute gores are collapsed or folded causing the parachute to take longer to open. The predicted filling times for the larger G-11A parachute are shown in Figure 8. Although this parachute is used extensively for airdropping heavy loads there is still a shortage of accurate information on its filling time. Estimates for this value range between 7 seconds and 10 seconds with the average being about 8 ½ seconds for snatch velocities between 200 ft/sec and 250 ft/sec.

Another interesting result of the prediction of the filling times is that doubling the weight of the suspended load results in a 10% reduction in the filling times for the C-9 parachute and a 25% reduction in the filling times for the G-11A. These differences in filling times seem to be consistent over the complete range of snatch velocities analyzed.

C. Other Applications

Other applications of the data to full-scale systems also exist. Flow field data with so high a confidence level would be of distinct value in the determination of the parachute stability angle during inflation. This would be accomplished by looking at the relationship between the diameter of the undisturbed stream-tube of flow around the parachute and the centerline of the flow. The interaction of the flow fields of parachutes in clusters would also be used to study the problem of irregular opening of clustered parachutes. Other areas include prediction of trajectories for systems, random gore opening problems and effects of reefing on opening characteristics.

VII. CONCLUSIONS

The hot-wire anemometer study has developed the first significant data on the flow field associated with an inflating

parachute. By controlling both the velocity and temperature of the flow in the test section, it was possible to identify velocities inside and immediately behind the parachute which have never been measured before. Although the objective of this study was to investigate the flow field, the results can be applied to the other problem areas in parachute technology described earlier in the study. It is now possible to relate the inflation geometry to the filling time, determine the filling time and extrapolate from scale-model data to full-scale data. Some of the other results of the study are listed below:

1. During the early stages of parachute inflation a low velocity region in the shape of an annulus forms at the base of the parachute gores.

2. As the inflation of the canopy progresses through the middle and final stages of inflation the low velocity annular region changes from the annulus shape to a region of approximately uniform velocity across the canopy inlet.

3. The velocity ratios behind the inflating canopy range between 0.03 and 0.05. The uniform distribution of the velocity ratios indicates that these values can be attributed to the flow passing through the canopy fabric.

4. The flow across the base of the hemispherical cap (Station 1) within the canopy during inflation appears to have a uniform transition from the initial to final stages of inflation.

5. The initial formation of the flow through the vent is visible in the velocity field within the canopy cap.

6. The results of the flow field study can be used to compute the filling times of various solid flat circular parachutes with a reasonable degree of agreement with full-scale data over a range of snatch velocities and suspended loads. This is significant because model simulation of full-scale systems in the field of parachutes has always been a major problem.

The state-of-art in parachute technology should be greatly advanced by the data presented in this paper, and the applications should provide better guidelines for parachute design.

VIII. REFERENCES

1. F. O'Hara: "Notes on the Opening Behavior and the Opening Forces of Parachutes," Royal Aeronautical Society Journal, November 1949.

2. "Performance of and Design Criteria for Deployable Aerodynamic Decelerators," Technical Report No. ASD-TR-61-579, December 63.

3. R. J. Berndt and J. H. De Weese: "Filling Time Prediction Approach for Solid Cloth Type Parachute Canopies," Paper presented at AIAA Aerodynamic Deceleration Systems Conference, Houston, Texas, September 1966.

4. R. J. Berndt: "Experimental Determination of Parameters for the Calculation of Parachute Filling Times," Paper presented at the Annual Meeting WGLR-DGRR, Berlin, Germany, 1964.

5. "Hot-Wire Measurement of Air Velocity, Direction and Temperature" Bulletin No. 94B, Flow Corp., Watertown, Mass.

6. A. Pope and J. J. Harper: "Low-Speed Wind Tunnel Testings" John Wiley and Sons, Inc. New York, New York, 1966.

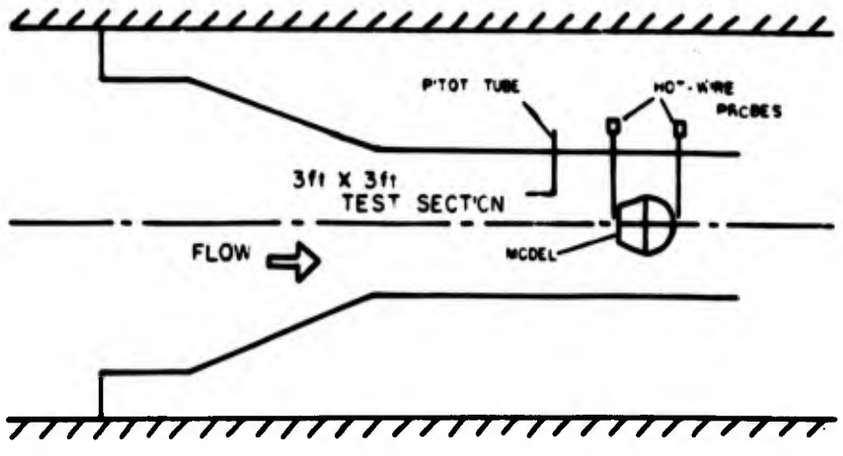


Fig. 1 Wind Tunnel Test Setup

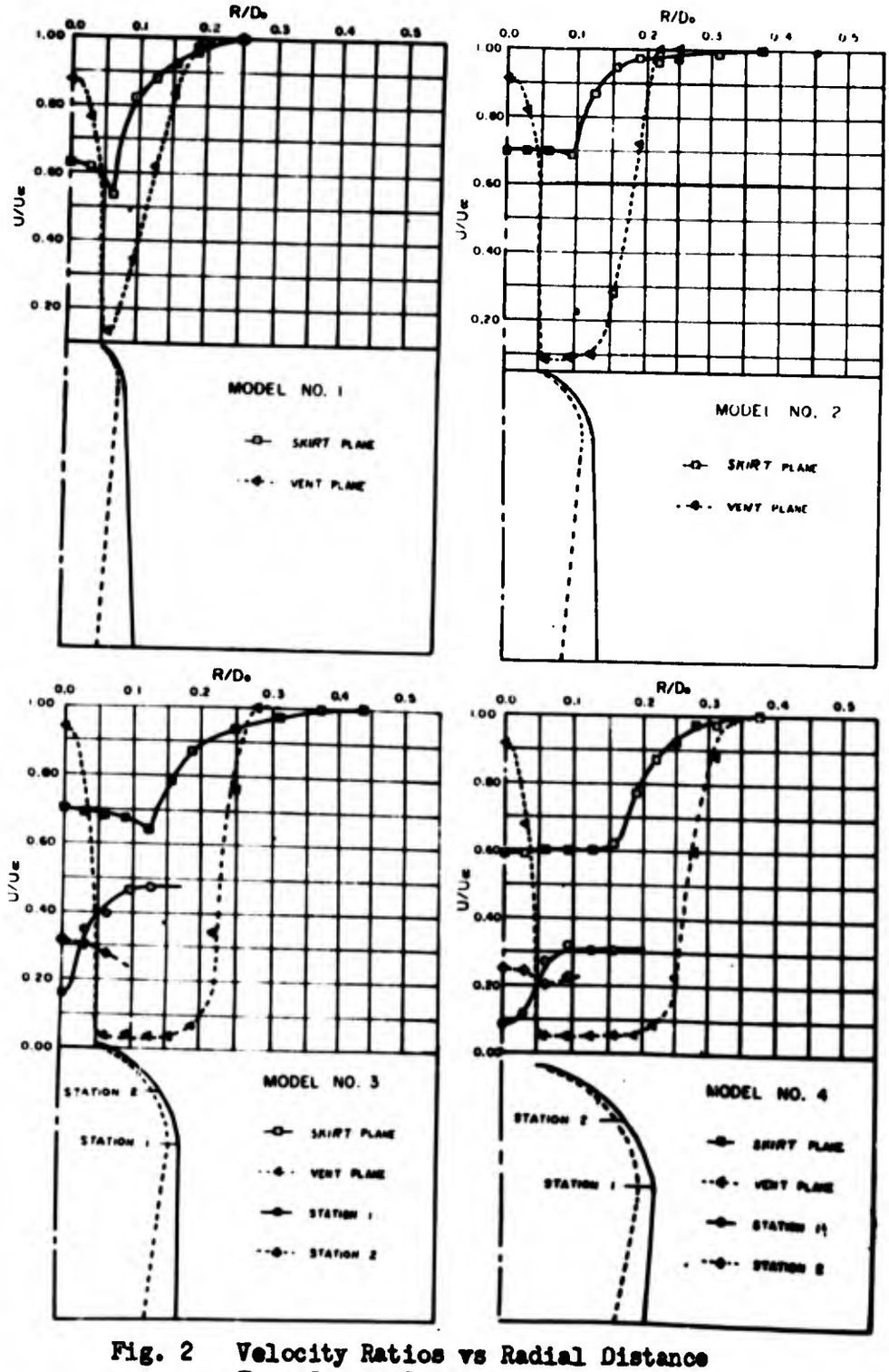


Fig. 2 Velocity Ratios vs Radial Distance From Canopy Centerline

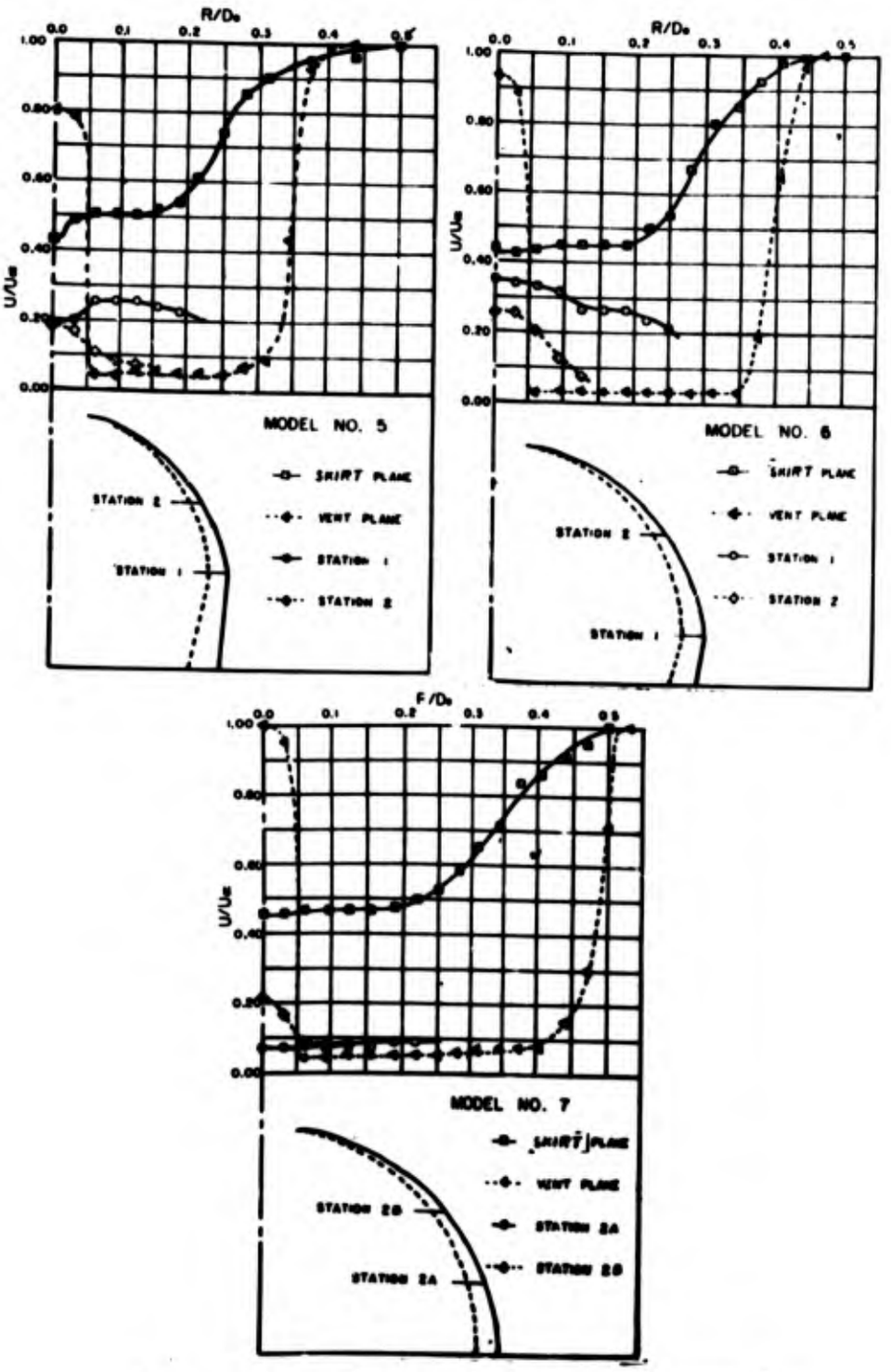


Fig. 3 Velocity Ratios vs Radial Distance From Canopy Centerline.

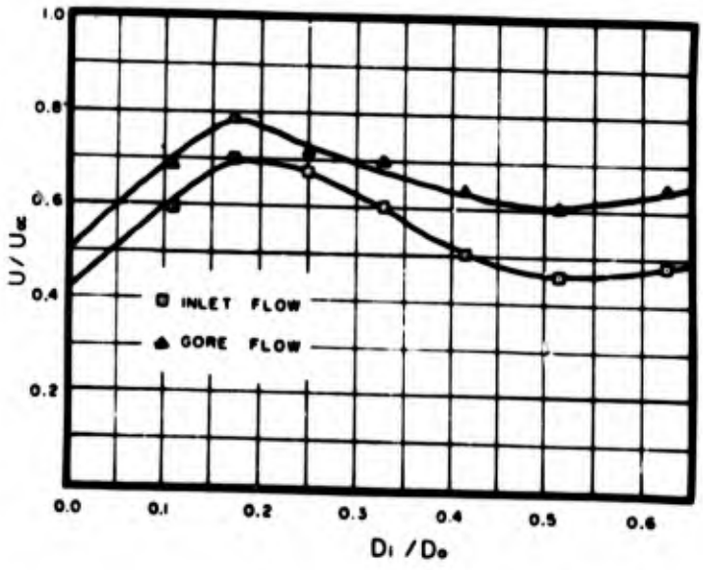


Fig. 4 Average Velocity Ratios Vs Inlet Diameter Ratio in the Skirt Plane

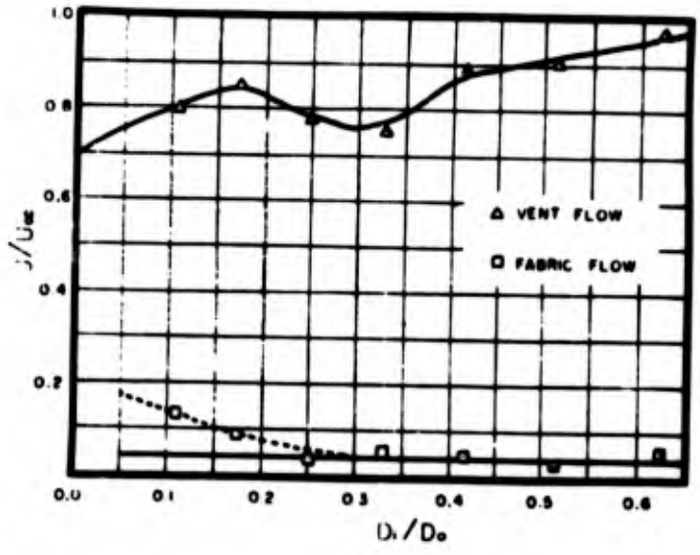


Fig. 5 Average Velocity Ratios vs Inlet Diameter Ratio in the Vent Plane

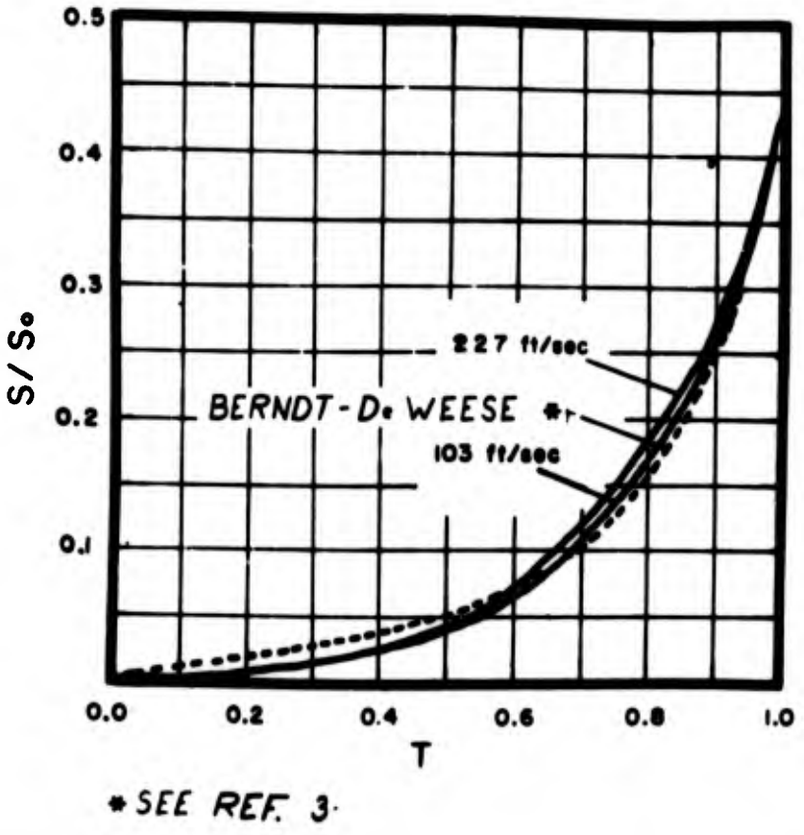


Fig. 6 Projected Area Ratio vs Dimensionless Time (Predicted Values and Full-Scale Data)

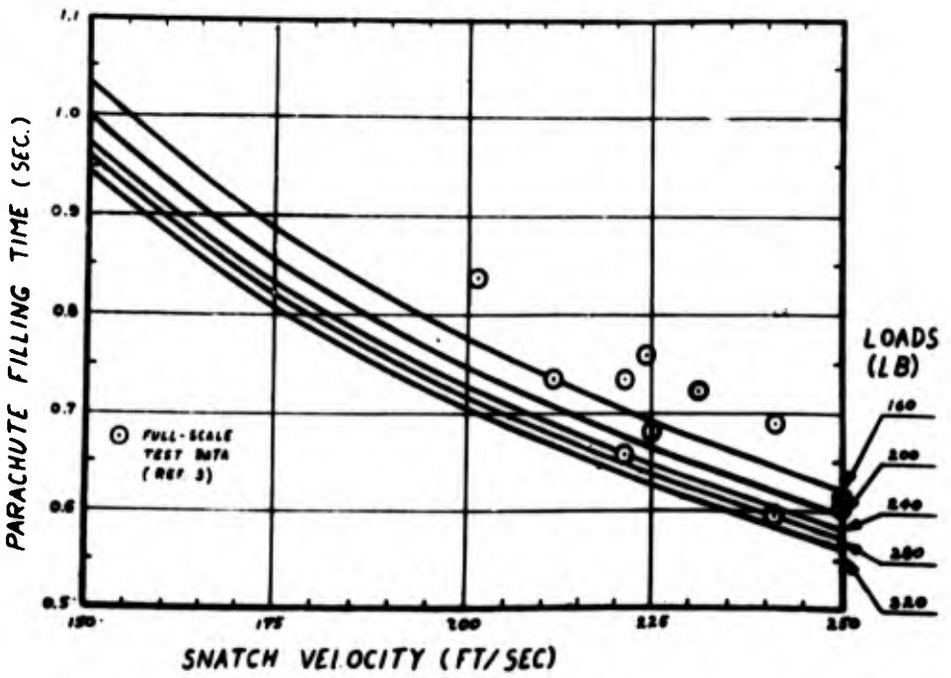


Fig. 7 C-9 Parachute - Filling Time vs Snatch Velocity (Predicted Values and Test Data)

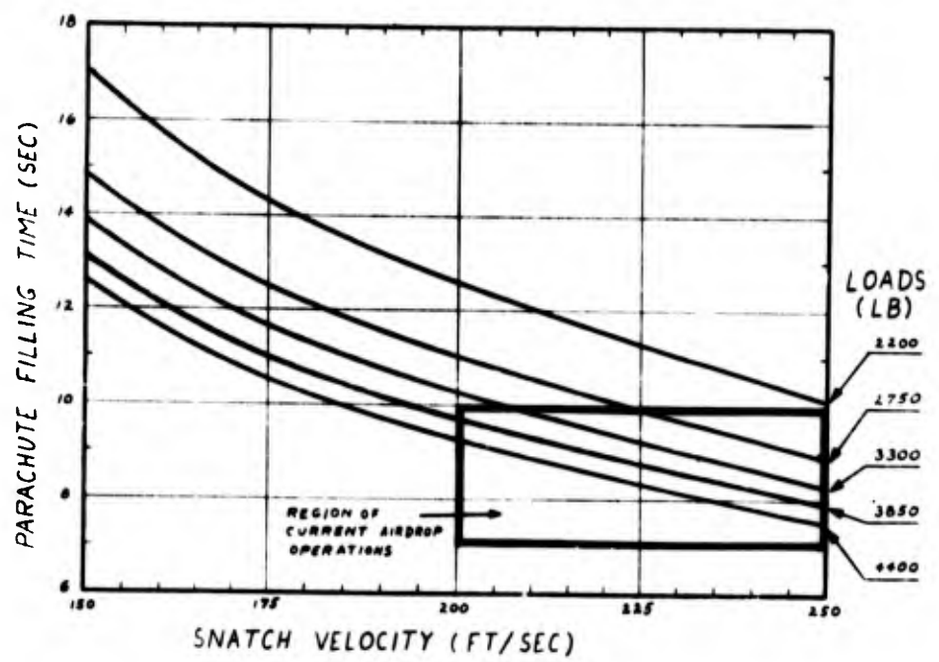


Fig. 8 G-11A Parachute-Filling Time vs Snatch Velocity (Predicted Values)

AD 713 521

FAR FIELD CHARACTERISTICS OF GROUND SHOCK INDUCED BY EXPLOSIONS

J. L. DRAKE AND A. SAKURAI
U. S. Army Engineer Waterways Experiment Station
Vicksburg, Mississippi

INTRODUCTION

Since the advent of nuclear weapons, the investigation of the ground shock effects from these devices has been the subject of much research. Reliable methods for estimating the stresses and the ground motions transmitted from a postulated nuclear detonation through earth materials are necessary for the cost effective design of hardened strategic systems.

Methods currently employed for the evaluation of ground shock effects are generally inadequate from the standpoint of the system designer. Computer code simulations are designed to predict the history of ground motion from the initial moment of the explosion, thus requiring an elaborate description of the characteristics of the material under high pressure and temperature. Because of their complexity, the extension of a numerical calculation into the range beyond the close-in field is costly and often unreliable due to the accumulating errors of the approximation. Empirically determined formulae, derived from scaled high explosive and past nuclear test events, are more generally used to provide quantitative description for specific weapons effects conditions; however, these formulae often cannot be extended to account for varying geometry of bursts and changes in the earth properties.

This situation is further complicated by the fact that the shock wave from a surface burst preserves the close-in characteristics of the individual explosion throughout the entire range of interest. Thus, slight deviations in the input values from the idealization of the real environment can result in seemingly unrealistic results. The sensitivity of the solution to variations in the source description can be mathematically related to a singularity of the solution at the origin.

This paper describes an attempt to improve the situation by utilizing a simple model which is adequate to describe the salient features of the ground motion history outside the close-in range. It is essential, to this purpose, to recognize the fact that the elastic

solution cannot account for the behavior near the explosion source and that the close-in effects are taken into account by an appropriate fit of the singularity of the solution. This approach is justified when these effects are confined to a small region in space and thus provides the key to the analysis.

SOLUTION FOR THE CONTAINED BURST

Ground shock effects research has centered around the fully contained underground weapons testing because of limitations on testing in the atmosphere. The analytical solution for the contained explosion is studied to provide a simple check case for verification of the elastic model and to determine the experimental scale factors required for fitting the close-in effects. Since the particle motion is most frequently measured in underground experiments, the particle velocity parameter will be used in this study.

The general solution for spherically diverging waves in an elastic medium expressed in terms of the particle velocity $\dot{u}(r,t)$ is (1)

$$\dot{u}(r,t) = \frac{c}{r^2}\left[F(r - ct) - rF'(r - ct)\right] \tag{1}$$

where r is the distance from the origin, t is time, and c is the dilatational wave velocity. $F(r - ct)$ is an arbitrary function of the argument $(r - ct)$. A prime denotes a derivative with respect to the argument and a dot denotes a derivative with respect to time.

Applications of this solution to waves generated by explosions were studied by a variety of authors (1,2). The explosion is commonly modeled by applying a pressure to the surface of a spherical cavity. The results are generally found to be inadequate for describing weapons effects because the far field solution predicts spatial peak attenuations of nearly r^{-1} whereas measurements show rates of about r^{-2}

A close examination of Equation 1 shows that by simply specifying the arbitrary function $F(r - ct)$ in a proper form, the first term (r^{-2} term) of the solution can be made to dominate and thus provide an expression that at least describes the peak values of an experiment correctly. The function $F(r - ct)$ was chosen as

$$F(r - ct) = Ae^{-\left[\frac{\alpha}{c}\left(t - \frac{r}{c}\right)\right]^2} \tag{2}$$

because of its simple form and fit to the measurements. The final expression for particle velocity is then

$$\dot{u}(r,t) = \frac{Ac}{r^2}\left(1 - \frac{2\alpha^2\tau r}{c}\right)e^{-\alpha^2\tau^2} \qquad \tau = t - \frac{r}{c} \geq 0 \tag{3}$$

where A and α are quantities to be determined that are characteristic of the source. The peak particle velocity occurs at $\tau = 0$

and attenuates as r^{-2} which is consistent with measurements. The wave form is an exponential which agrees with the typical pulse determined by weapons testing.

Conventional cube root scaling is introduced into Equation 3 without changing its dimensions in order to determine the variation with the weapon yield W. Equation 3 becomes

$$\dot{u}(r_o,t_o) = \frac{A_o c}{r_o^2}\left(1 - \frac{2\alpha_o^2\tau_o r_o}{c}\right)e^{-\alpha_o^2\tau_o^2} \tag{4}$$

where

$$A = A_o W^{1/3} \quad , \quad r = r_o W^{1/3} \quad , \quad \tau = \tau_o W^{1/3} \quad , \quad \alpha = \alpha_o W^{-1/3}$$

The weapon parameters A_o and α_o are easily determined from the scaled field measurements by noting

$$A_o = \frac{\dot{u}_{max}\, a_o^2}{c} \qquad \alpha_o = \frac{c}{2a_o\tau_o^+} \tag{5}$$

where $\dot{u}_{max}$ is the maximum particle velocity and τ_o^+ is the positive duration of a measured wave at the scaled range $r_o = a_o$. Evaluations of A_o and α_o show that they are constants not depending on test material or weapon yield.

A comparison between the wave form calculated by Equation 4 and a wave form from the Salmon (5 kt) (3) event in a salt medium is shown in Figure 1. The elastic calculation necessarily has an abrupt rise on the initial portion because of the need to make $F(r - ct)$ dominate its derivative in the neighborhood of the peak value. The peak particle velocity calculated using constant A_o is compared with several nuclear bursts in a variety of earth materials in Figure 2.

SOLUTION FOR THE SURFACE BURST

The object of this section is to utilize the classical linear elastic model for the prediction of the characteristics of the ground shock waves generated by explosions at the surface of the earth. The use of this model for this purpose has not been fully explored because of the lack of an adequate analytical solution to be utilized. Since the mathematical problem of the explicit solution of the surface burst explosion is very difficult, even for the simplest model of linear elasticity, the existing analytical solutions to this problem are either incomplete or unrealistic for all practical purposes; thus, a more realistic and explicit solution had to be developed. It is essential, to this purpose, to recognize the fact that the solution of the surface burst problem is very sensitive to the input conditions. Mathematically, this sensitivity is related to the singularity of the solution at the explosion source, and the

solution, to be useful, should be constructed such that the singularity properly represents the characteristics of the source.

Consider the displacement potentials φ and ψ determined from the wave equations

$$\frac{\partial^2 \varphi}{\partial t^2} = c^2 \Delta \varphi \quad , \quad \frac{\partial^2 \psi}{\partial t^2} = c_s^2 \Delta \varphi \tag{6}$$

with propagation velocities c and c_s of the compressional and shear waves and $\Delta = \frac{\partial^2}{\partial z^2} + \frac{1}{r}\frac{\partial}{\partial r}\left(r \frac{\partial}{\partial r}\right)$ using the cylindrical coordinate system r and z with axial symmetry. The formal solution of Equation 6 to the surface burst problem can be obtained in the form of the double integral of the Laplace-Hankel transforms. The direct derivation of the explicit expression of φ and ψ from this formal solution is difficult for the general region of $r > o$, but it can be obtained on the axis $r = o$ in the following asymptotic expansion form:

$$\varphi \sim \frac{a_1}{z} + \frac{a_2}{z^2} + \cdots \quad , \quad \psi \sim \frac{b_1}{z} + \frac{b_2}{z^2} + \cdots \tag{7}$$

where a_i and b_i $(i = 1,2,\ldots)$ are functions of $t - \frac{z}{c}$ and $t - \frac{z}{c_s}$ respectively and are determined from the characteristics of an individual input source condition. Equation 7 exhibits the nature of the singularities on the input source.

Approximate expressions for φ and ψ for $r > o$ are constructed from the sum of the elementary solutions of the wave equation which have singularities at the origin that are matched to the solution (Equation 7) above are

$$\varphi \sim \frac{\varphi_1\left(t - \frac{R}{c}\right)}{R} + \frac{\partial}{\partial z}\frac{\varphi_2\left(t - \frac{R}{c}\right)}{R} + \frac{1}{r}\frac{\partial}{\partial r}\left(r \frac{\partial}{\partial r}\frac{\varphi_3\left(t - \frac{R}{c}\right)}{R}\right) + \cdots$$

$$\psi \sim \frac{\psi_1\left(t - \frac{R}{c_s}\right)}{R} + \frac{\partial}{\partial z}\frac{\psi_2\left(t - \frac{R}{c_s}\right)}{R} + \frac{1}{r}\frac{\partial}{\partial r}\left(r \frac{\partial}{\partial r}\frac{\psi_3\left(t - \frac{R}{c_s}\right)}{R}\right) + \cdots \tag{8}$$

$$R = \left(r^2 + z^2\right)^{1/2}$$

where φ_i and ψ_i $(i = 1,2,\ldots)$ are determined from a_i and b_i $(i = 1,2,\ldots)$ by comparing the terms with the same powers in z at $r = o$ with those in Equation 7. The horizontal and vertical displacements u and v are then determined by

$$u = \frac{\partial \varphi}{\partial r} + \frac{\partial^2 \psi}{\partial r \partial z} \quad , \quad v = \frac{\partial \varphi}{\partial z} - \frac{1}{r} \frac{\partial}{\partial r}\left(r \frac{\partial \psi}{\partial r} \right) \tag{9}$$

The accuracy of the approximation depends upon how many terms are retained in Equation 8. The result of the application of this method to the simpler case of the liquid half-space with only one wave shows good agreements with experimental data is attained with two terms of the approximation formula (4). A cursory examination of the elastic case with two terms of the approximation is being made and compared with test data for rock environments. The results of this correlation are shown in Figure 3, where the horizontal accelerations $\ddot{u}$ near the surface are compared with the MINE ORE (5) high explosive test data.

CONCLUSIONS

General formulae for the ground motions from an explosion source were derived and their results compared with high explosive and nuclear weapons test data. The comparisons show reasonably good agreement may be obtained by using a simple elastic description for the earth as long as the explosion is properly modeled.

REFERENCES

1. Newmark, N. M., "Feasibility of Modeling Cavity Behavior in Jointed Rock Masses (U)," Contract Report No. 1-135, November 1965, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

2. Ewing, W. M., Jardetzky, W. S., and Press, F., Elastic Waved in Layered Media, McGraw-Hill, New York, N. Y., 1957.

3. Perret, W. R., "Free-Field Particle Motion from a Nuclear Explosion in Salt, Part 1 (U)," VUF-3012, June 1968, Sandia Laboratory, Albuquerque, N. M.

4. Sakurai, A., "Shock Waves Resulting from Explosions at an Air-Water Interface; Report 1, Analytical Study (U)," Technical Report TR N-69-3, May 1969, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

5. Joachim, C. E., "Mine Shaft Ground Motion Measurements: Events Mine Under and Mine Ore (U)," U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss. (unpublished).

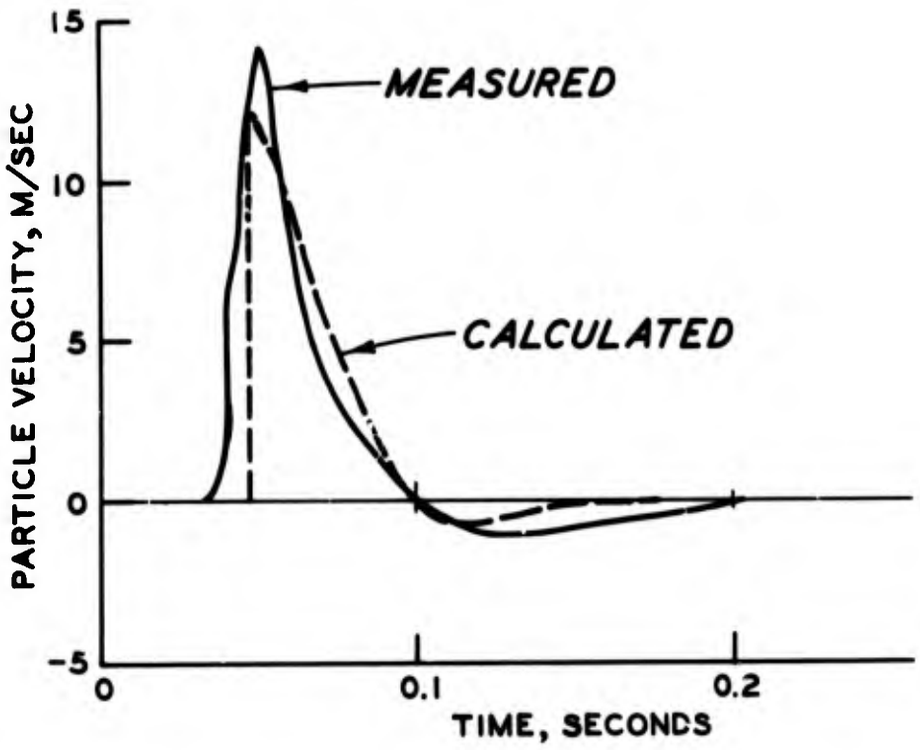


Fig. 1. Comparison of calculated particle velocity-time history with test data from Event SALMON at r = 166 meters

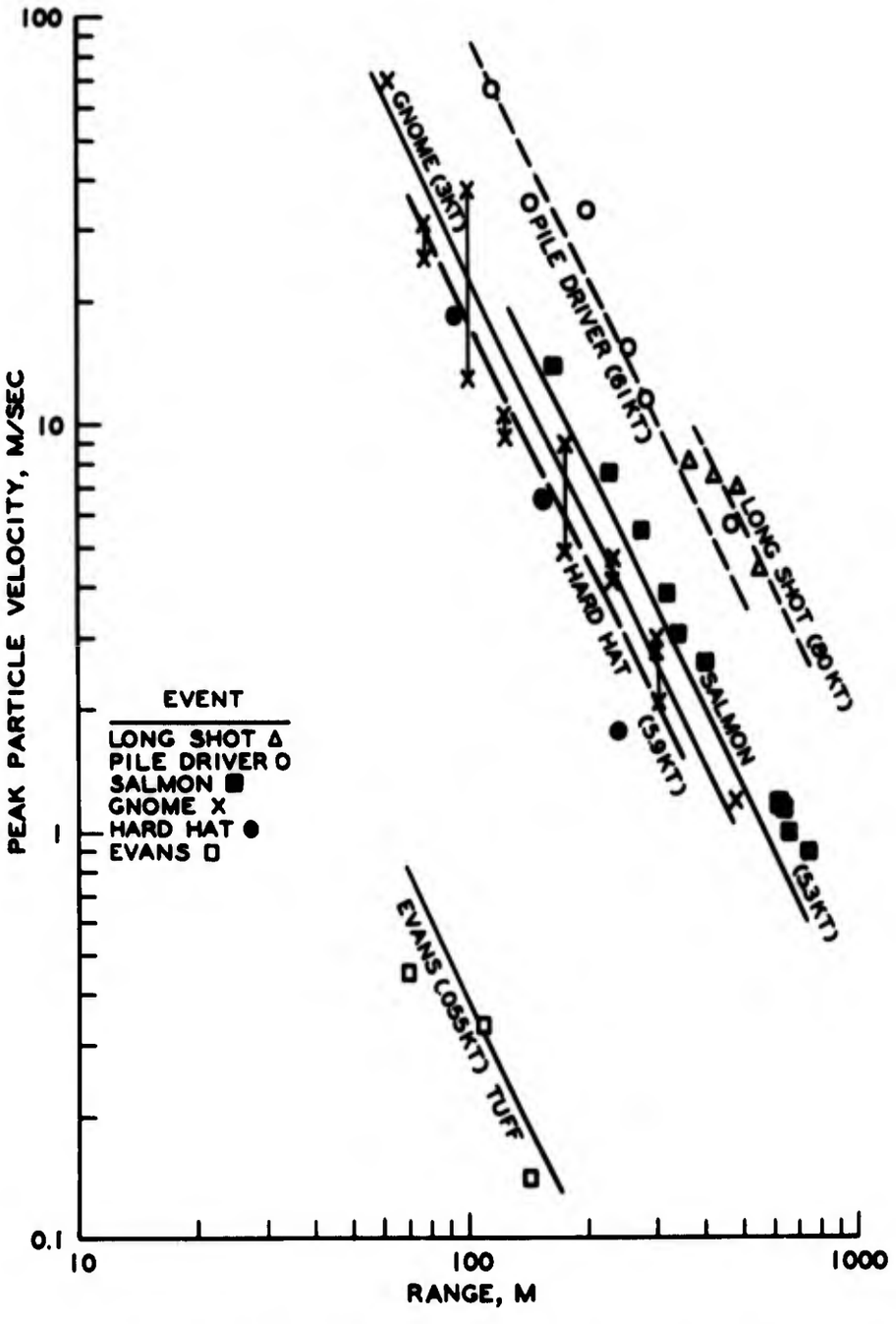


Fig. 2. Comparison of calculated peak particle velocity with test data from nuclear events in rock

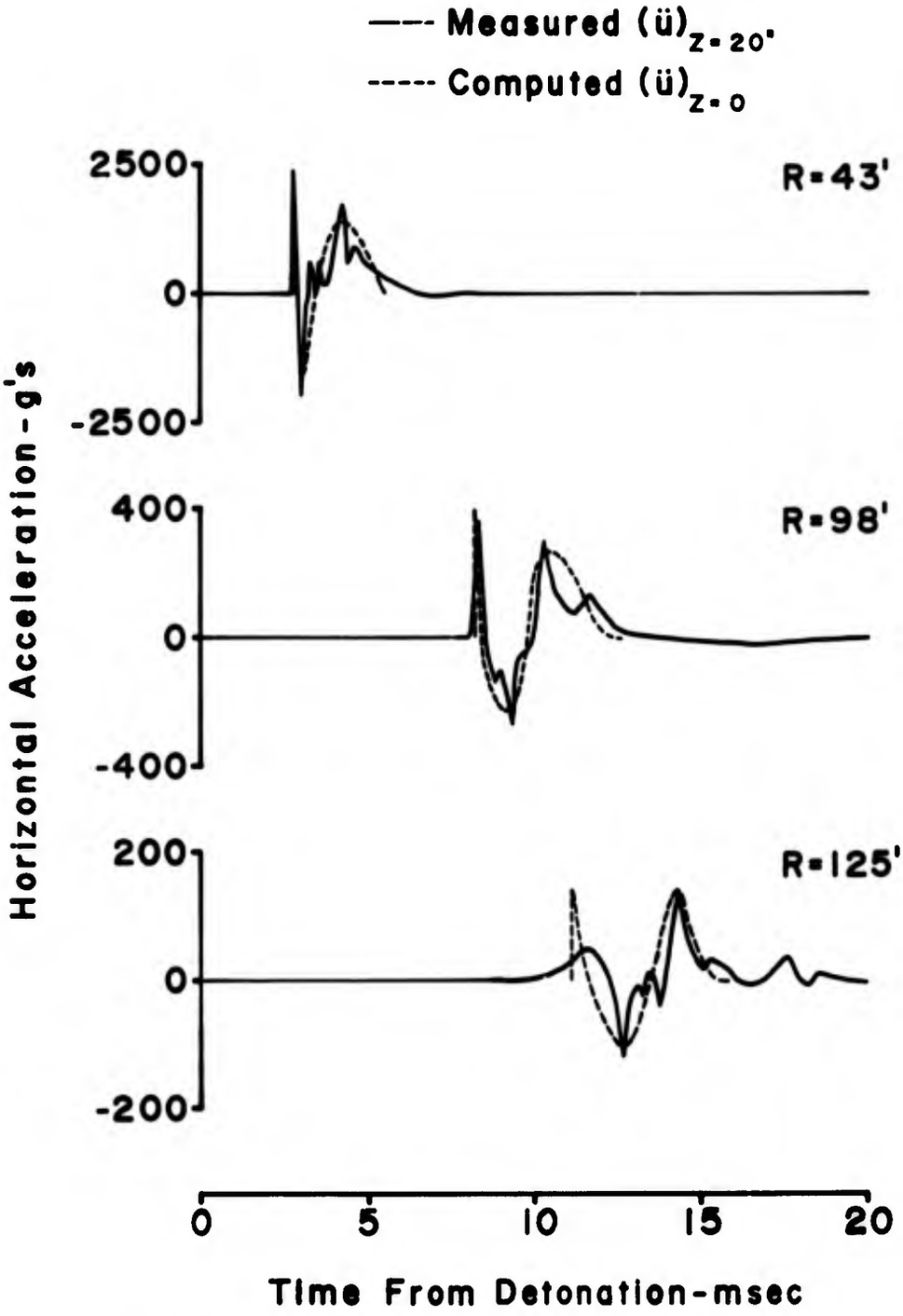


Fig. 3. Comparison of calculated horizontal acceleration-time history with test data from MINE ORE Event, MINE SHAFT Series

AD 713 522

THE CHARACTERISTIC COEFFICIENTS TECHNIQUE FOR PROBABILITY MODELS OF WIND PROFILES IN MISSILE DESIGN AND ENVIRONMENT ANALYSIS

DR. O. ESSENWANGER
PHYSICAL SCIENCES LABORATORY, RESEARCH AND ENGINEERING DIRECTORATE
U. S. ARMY MISSILE COMMAND
REDSTONE ARSENAL, ALABAMA 35809

I. INTRODUCTION

Three forms to describe global conditions of the atmosphere for computer analysis of the wind influence upon missile systems have been common in the past. The first method utilizes individual wind profiles. Although appropriate, the voluminous input into computer programs can cause storage problems, analysis may be very costly and suffer from inhomogeneity and data bias. Therefore usually only a limited number of profiles are taken into consideration, and then results lack completeness and are not conclusive.

The input problem is reduced to some degree by the establishment of intra and interlevel correlation matrices. Their establishment, however, is costly and complex, and little is gained to simplify design analysis.

The third technique employs a limited number of mostly synthetic wind profiles, which curtails the analysis costs, but practically neglects the realistic vertical wind profile relationship. Further, association with wind shear, gust, and turbulence is difficult, and a true probability level is hard to assess. Modern analysis tools such as a Monte Carlo method cannot be applied.

A new approach has been attempted by the author to overcome above-mentioned shortcomings. The individual wind profile is described by a number of characteristic coefficients, whose interrelationships were studied. It proved that the interdependence permits the reduction of the analytical model to a one-variable parameter only. Hence the complete set of global wind conditions can be expressed by three equations and a set of constants. This solves the input problem, eliminates data bias and associates wind profiles with probability; and modern analysis tools, including Monte Carlo methods, can be employed.

This new characteristic-coefficients technique solves an additional problem. It provides a relatively simple scheme to establish a set of global wind models. This task was difficult to resolve in the past and has led to synthetic models with deviations from the realistic vertical wind relationship.

The employed technique is not restricted to wind profiles but has general application to other problems of model selection, especially when modelling is based upon empirical rather than theoretical concepts. The scheme will be discussed in the subsequent sections of the manuscript.

Finally, a set of wind models will be introduced and the geographic and seasonal distribution of the wind models will be analyzed to show the rationality of the system.

II. DERIVATION OF WIND MODELS

a. Characteristic Coefficients of Individual Profiles

In previous reports (1, 2) the author has demonstrated that the wind speed profile (wind speed as function of altitude) from the surface to about 25 km can be efficiently described by terms of a Fourier series, while the wind direction profile is more easily approximated by polynomials. In a recent report O. Essenwanger and H. Boyd (3) have studied the interrelationship of the coefficients of these Fourier terms and have concluded that only one parameter is necessary to represent the wind speed profile V_h with sufficient accuracy. This leads to the form

$$(V_h - \bar{V}_h) = (A_o - \bar{A}_o)\,(1 + k_1 \sin(\alpha_h + \bar{\alpha}_1) + k_2 \sin(2\alpha_h + \bar{\alpha}_2) + k_3 \sin(3\alpha_h + \bar{\alpha}_3)) \quad (1)$$

where $$\alpha_h = \frac{2\pi h}{H} = \frac{360h}{H},\ h = 0,\ 1,\ 2 \ldots (H-1) \quad (1a)$$

and $$\bar{V}_h = \bar{A}_o\,(1 + k_1 \sin(\alpha_h + \bar{\alpha}) + \ldots) \quad (1b)$$

The $\bar{A}_o$, k_1, k_2, k_3, $\bar{\alpha}_1$, $\bar{\alpha}_2$, $\bar{\alpha}_3$ represent constants which vary with season and geographic location.

Since the α_h is merely a phase adjustment for the altitude h, the only variable parameter in eqn. (1) is A_o. This enables us to represent the frequency distribution of the wind speed profiles by a frequency distribution of the parameter A_o, and the link to

probability is found. A suitable form may be the Weibull distribution (see 2)

$$F(x) = 1 - e^{-\left(\frac{x-\gamma}{\theta}\right)^{\beta}} \qquad (2)$$

where the β, γ, θ are parameters of the distribution and the x corresponds to the A_o. Any desired threshold of the frequency distribution can be computed and the pertinent wind speed profile obtained for trajectory analysis and design studies. This method has already been applied successfully in development of new missile systems.

The third equation relates the wind direction θ to the altitude, corresponding with the set of constants of eqn. (1).

$$\theta_h = \theta_m + c_1 \phi_{1h} + c_2 \phi_{2h} + \ldots c_n \phi_{nh} \qquad (3)$$

where ϕ_1 and ϕ_2 are orthogonal (Tchebycheff) polynomials (see 2). The three equations satisfy three conditions: a realistic description of the true wind profiles, an association with probability and a limited number of input quantities. With this solution Monte Carlo studies of variations of missile components and configurations can easily be performed without introducing data bias.

b. Wind Model Concepts

The next step in the development is the derivation of a global set of wind models which can be utilized in analysis and in the tactical operation of missile systems. The establishment of models is generally a subjective task and depends on the practical application for a particular missile system. Thus the sets of wind models may differ for individual systems. Some basic concepts are common, however, and are subsequently discussed. These concepts also have widespread application beyond the narrow field of wind analysis.

As introduced with eqn. (1), the only independent parameter in eqn. (1) is the A_o. If the constants k_i and $\bar{\alpha}_i$ showed little variation with time or space (in our case month and geographic location) a global frequency of the parameter A_o could be formed by combining the individual frequencies. The models could then be based on subsequent class intervals of the A_o. Since the width of the class interval can be chosen arbitrarily, the number of models would depend upon the desired distance between the models. There would be no duplication of models, and probabilities would correspond to the probabilities of the A_o classes.

Unfortunately, this trivial scheme does not succeed for the wind speed because of seasonal and geographic variation of the constants. Thus the variation of the constants must be considered. We could construct a set of models for every individual station and combine these sets into a global scheme. The global set would almost certainly contain similar profiles, whose elimination could create an arduous task.

A close perusal of eqn. (1) and its constants reveals that the difference in constants can be expressed by a simple parameter, the percentage reduction. This can be written for the individual term of the Fourier analysis as

$$z_i^2 = \frac{A_i^2}{2\sigma^2} , \tag{4}$$

where the A_i is the amplitude of the harmonic term and the σ^2 the variance of the N observational data. In our case the following relationships exist

$$A_i = A_o \, k_i \tag{5a}$$

and

$$2\sigma^2 = \sum_1^n A_i^2 = \sum_1^n k_i^2 \tag{5b}$$

This leads to

$$z_i^2 = k_i^2 / \sum_1^n k_i^2 \tag{4a}$$

in which the contribution of the individual term to the overall explanation of the variance is expressed. This scheme is independent of the parameter A_o and contains the constants in normalized form only.

In our scheme, three coefficients have been considered a sufficient approximation of the wind profile and therefore n = 3. The scheme now resembles the system employed by the author in the derivation of sound speed profile models (see 4, 5).

An objective arrangement of the models could be made by following a similar division of dominant components as the author

has utilized for sound speed profiles. Table 1 reflects the summary of a study of the z_i^2 for four selected stations of different wind regimes and all months. It is disclosed that some combinations virtually do not occur (see last column of Table 1) and that modelling could therefore be restricted largely to conditions 1, 4 and 7 of Table 1. This leads to the four different model groups of Table 2, in which the groups have been further subdivided by different phase angles $\bar{\alpha}_1$, $\bar{\alpha}_2$, and $\bar{\alpha}_3$. If the k_i had not all been positive, more diversity in the group would have been required in the family of wind models.

Table 1

Analytical Scheme for Modelling based on the Dominance of Fourier Term Expressed by the Percentage Reduction

Model	z_1^2	z_2^2	z_3^2	n
1	1.0	0	0	28
2	0	1.0	0	1
3	0	0	1.0	0
4	.5	.5	0	9
5	.5	0	.5	0
6	0	.5	.5	0
7	.33	.33	.33	10

Figure 1 displays group 1a for varying A_o, ranging from 42 m/sec for model 1 to 7 m/sec for model 6. The variety of group conditions for a constant A_o of 28 m/sec is illustrated in Figure 2. The association of code numbers in Figure 2 is given in the last column of Table 2. This graph has been drawn for a first set of 43 models, which did not include group 2b.

Table 2 discloses the conditions of the z_i^2. They are altered from the scheme of Table 1. As a consequence, an unequal difference ϵ_Δ^2 between the models arises which is unessential, however, for our practical purpose. The ϵ_Δ^2 is defined by

$$\epsilon_\Delta^2 = \left[\frac{1}{n} \Sigma \, (y_{ij} - y_{ik})^2\right]^{\frac{1}{2}} \qquad (6)$$

where the i represents the altitude, the j and k two different models. ϵ_Δ^2 for j = k is zero.

The model groups and corresponding constants k_i and $\bar{\alpha}_i$ should be considered preliminary, as more work is necessary to derive a final set.

Table 2

Percentage Reduction z_i^2 for Chosen Model Groups and Corresponding k_i Coefficients and Phase Angles

Model Group	z_1^2 %	z_2^2 %	z_3^2 %	k_1 %	k_2 %	k_3 %	$\bar{\alpha}_1$ Degree	$\bar{\alpha}_2$ Degree	$\bar{\alpha}_3$ Degree	max A_o m/sec	C*)
1a	100	0	0	100	0	0	290	0	0	40	3
1b	100	0	0	100	0	0	310	0	0	40	9
2a	85	15	0	92	39	0	330	225	0	40	15
2b	60	30	10	77	55	32	150	180	190	32	-
3a	45	45	10	67	67	32	270	200	300	40	21
3b	45	45	10	67	67	32	200	180	250	40	27
3c	45	45	10	67	67	32	150	285	200	32	31
4a	40	35	25	63	59	50	200	210	240	32	35
4b	40	35	25	63	59	50	200	140	240	32	39

C*) Code number in Figure 2

c. Comparison of Empirical Wind Profiles with Wind Model Profiles

As previously stated, how many models we choose is an arbitrary decision depending on various conditions and assumptions, one being the desired average deviation of the empirical profile from the model. It is obvious that the larger the spacing between models the higher the expected deviation. We may measure this deviation by the squared difference

$$\epsilon_\Delta^2 = \left[\frac{1}{n}\Sigma(y_i - y_{ia})^2\right]^{\frac{1}{2}} \qquad (6a)$$

which was introduced earlier. This time the y_{ij} is replaced by the empirical profile y_i and the y_{ik} by the model profile y_{ia}. The assignment of the model to the empirical profile is then based on the minimum ϵ_Δ^2, although other methods could be employed (see 4, 5).

The first study of the ϵ_Δ^2 was mainly performed to obtain a survey of the frequency distribution of ϵ_Δ^2 and to study whether the series of models was complete. Profiles with $\epsilon_\Delta^2 > 10$ m/sec were printed out and specially examined. This study led to the inclusion of the model group 2b. Large deviations, however, are

mostly of the type illustrated by Figure 3, which exhibits the empirical and the two adjacent wind model profiles. As is shown, the empirical profile fluctuates between the two models. Other variations appear like spikes, which cannot be expected to be included in any wind model. Computation of the total percentage reduction for three Fourier terms of the empirical profile resulted in 67% with dominance of the first term. This explains first the assignment of the empirical profile to the models with code 8 and 9 (group 1b). It further reveals that the model, derived from three coefficients, cannot express more than the 67% of the variance and hence the large ϵ_{Δ}^2 must be expected. It is obvious that the construction of wind models cannot be based on a complete number of terms, in our case 13, as too many small scale fluctuations would enter.

Table 3 displays the frequency distribution of the ϵ_{Δ}^2 for three model systems. First a set of 43 models (without group 2b) has been computed with differences between A_o steps of 7 m/sec (see Figure 1). The maximum A_o in all groups is not the same (see Table 2 next to last column) and therefore only 43 models are required. Model 43 represents a condition of no wind at all levels.

Models with a low A_o in the various groups were very close, and therefore nine models were deleted to achieve an approximate spacing of 7 m/sec between models. It can be seen that the cancelling of models did not result in any appreciable increase of the ϵ_{Δ}^2. Hence the reduction to 34 models seems justified.

While studying the ϵ_{Δ}^2 at Thule it was discovered that a group of large deviations existed which were not random fluctuations or spikes as discussed before (Fig. 3) but belonged to a model group 2b. The addition of this group would have increased the number of models. An attempt was made to enlarge the spacing of the A_o within a group to 8 m/sec. Under these changed conditions a total of 42 models was necessary. The ϵ_{Δ}^2 for this system is displayed in the last four columns of Table 3. Because of the wider spacing one would expect an increase of the ϵ_{Δ}^2 compared with the 43 model system. Table 3 displays that the introduction of group 2b more than compensates for the growth and the average ϵ_{Δ}^2 remains the same, although not at every individual station.

It is intended to study the problem of group establishment further and to derive or modify the models by averaging empirical profiles for the individual models. Hence deletion of models in the 42 system was postponed. Although the set of 42 models could already be utilized for analysis of missile systems, it should still be considered of preliminary nature.

Table 3

Frequency Distribution of the Squared Deviation ϵ_{Δ}^2

Class m/sec	43 Models A %	43 Models M %	43 Models Ch %	43 Models Th %	34 Models A %	34 Models M %	34 Models Ch %	34 Models Th %	42 Models A %	42 Models M %	42 Models Ch %	42 Models Th %
1.0 - 1.9	-	.0	.2	1.2	-	.0	.0	.0	.0	.0	.2	.6
2.0 - 2.9	.5	2.0	4.9	9.8	.3	1.4	2.0	1.3	1.4	1.9	4.0	6.9
3.0 - 3.9	9.6	14.1	17.2	24.7	6.0	11.2	9.8	11.9	12.7	13.6	14.5	19.9
4.0 - 4.9	30.0	26.6	27.2	29.1	26.7	24.6	24.3	29.9	32.7	23.8	24.8	30.6
5.0 - 5.9	24.2	21.3	19.8	12.7	25.1	22.0	24.3	24.8	36.5	22.4	22.8	21.8
6.0 - 6.9	18.2	14.2	10.6	8.2	19.5	16.9	14.8	13.8	11.8	15.8	12.8	8.6
7.0 - 7.9	12.0	10.4	7.7	5.2	14.7	11.3	10.2	7.8	3.1	10.1	7.9	5.1
8.0 - 8.9	4.5	5.2	5.1	3.2	6.1	6.0	6.7	4.2	1.2	5.6	5.0	3.1
9.0 - 9.9	.9	2.7	3.3	2.2	1.4	3.0	3.5	2.4	.5	3.4	3.2	1.6
> 10.0	.1	3.5	4.0	3.7	.2	3.6	4.4	3.9	.1	3.4	4.1	1.8
Mean	5.6	5.7	5.5	4.9	5.6	5.9	5.9	5.5	5.1	5.8	6.0	5.0
N	3590	3601	3751	6859	3590	3601	3751	6859	3590	3601	3751	6859

A = Albrook (Canal Zone)
M = Montgomery (Alabama)
Ch = Chateauroux (France)
Th - Thule (Greenland)

d. Frequency Distribution of Wind Models

The merits of the technique to derive wind models from characteristic coefficients must be judged by the application to observational data. Therefore a brief survey of the geographical and seasonal variation of the groups is given below.

Table 4 contains the frequency of occurrence of the various model groups by season and by station. The seasonal and station combination was based on results exhibited in Table 5. The data were homogenized by converting the seasonal frequency into percentage and combining them to the averages shown in Table 4, eliminating the differences in the number of observations for the individual season and station.

We conclude from Table 4 that the first four groups contain the major types with a total of 68% of occurrence compared to 32% for all other groups. These four groups are associated each with one particular station, representing a climatological regime. Thus group 1a appears 50% of the time at Montgomery (subtropics), type 1b is dominant at Chateauroux (midlatitude), 2a at Thule (polar) and 2b at Albrook (tropics). While groups 3a - c and 4a apparently display no significant geographic pattern, group 4b seems to be another tropical type.

The presence of at least two major groups at each station supports a statement earlier in this article that derivation of individual station models based on the frequency distribution of the A_o alone would not have succeeded.

The second part of Table 4 displays the seasonal changes of the individual groups. As expected, seasonal shift for some groups exists, making some groups summer types (e.g. 2a or calm) and others winter types (e.g. 1a). The seasonal shift of models together with the shift in station occurrence fits into the pattern of the general circulation, which can be studied in more detail from Table 5.

The dominance of types 2b and 4b as tropical types at Montgomery in summer is in line with a seasonal shift of circulation patterns. Thus the diminuation of the subtropical type in summer at Montgomery should be expected.

The polar type 2b expands in winter and spring to midlatitudes, while in summer type 1b prevails as a midlatitude type.

Since the derived wind groups follow known features of the general circulation pattern, the established system appears rational.

Table 4

Frequency of Model Groups (in Percentage)

Model Group	by Station				by Season				Total
	Alb	Mont	Chat	Thule	12-2	3-5	6-8	9-11	
1a	10.1	50.1	13.4	2.2	24.2	21.2	9.0	21.2	18.9
1b	1.6	18.2	33.1	2.9	9.2	15.8	14.5	16.3	14.0
2a	2.7	2.5	28.0	44.9	13.8	23.8	22.8	17.8	19.5
2b	34.[illegible]	6.4	4.9	18.2	19.7	11.2	16.8	15.7	15.8
3a	7.2	6.6	9.4	4.5	8.4	5.4	6.9	7.1	7.0
3b	5.2	4.7	3.2	2.6	4.4	3.3	4.6	3.4	3.9
3c	9.6	1.6	1.0	6.2	4.1	4.2	6.6	3.5	4.6
4a	10.3	2.1	2.3	7.2	6.2	5.5	5.1	5.2	5.5
4b	16.5	7.0	1.6	.4	9.6	5.0	6.1	4.8	6.4
Calm	2.8	0.8	3.1	10.9	0.4	4.6	7.6	5.0	4.4

Table 5

Seasonal Variation of Group Frequency by Station (in °/oo)

Model Group	Albrook				Montgomery				Chateauroux				Thule			
	Wi	Sp	Su	Fa	Wi	Sp	Su	Fa	Wi	Sp	Su	Fa	Wi	Sp	Su	Fa
1a	175	45	69	114	710	677	90	525	74	97	192	171	10	28	11	39
1b	15	17	11	23	230	243	40	214	107	338	491	387	18	35	36	28
2a	16	39	26	27	12	23	32	34	303	394	195	230	219	496	659	421
2b	149	334	435	441	6	0	232	16	149	26	3	19	483	90	0	153
3a	113	98	45	34	20	18	154	72	169	64	50	95	35	34	27	84
3b	74	81	15	38	5	5	156	21	82	15	10	23	15	30	5	54
3c	29	80	212	61	2	4	43	18	24	3	4	8	110	82	4	54
4a	75	122	134	82	3	9	64	9	70	13	3	5	100	76	4	110
4b	352	169	36	104	12	14	183	71	15	11	23	13	3	6	1	4
Calm	2	15	17	76	0	7	6	20	7	39	29	49	7	123	253	53

Wi = December - February
Sp = March - May
Su = June - August
Fa = September - November

III. CONCLUSIONS

Based upon the representation of the wind profile by characteristic coefficients as described in a recent article by the author (2), a technique is developed to derive wind models on a global basis. The division into groups of models is accomplished by utilizing the percentage reduction of three coefficients (eqns. 4 and 4a), but only a limited number of the possible combinations occur. Nine major groups are introduced (Table 2), and individual models are derived by the differences in A_0. The extreme value of the A_0 in the individual groups fluctuates. Since the spacing of the A_0 depends on configurations of the particular missile system, the total number of individual wind models may vary from system to system.

A system of 42 models is presented which essentially describes the wind structure from surface to 25 km on a global basis. Such a system is an additional tool to the frequency distribution of the characteristic coefficients for systems analysis and tactical application. Global probabilities can be established similar to the ones provided for the groups (Tables 4, 5). Association with other parameters such as wind shear, gust, turbulence etc. can be studied and the prediction of models for tactical operation can be attempted. The system of wind models emerging from the application of the characteristic coefficients proves therefore an important tool in system analysis and in application of future missile systems and has been tested on a particular system.

It should further be stressed that the derived models are not synthetic profiles describing the vertical wind relationship in unrealistic ways. They are based upon the real atmospheric structure and therefore system analysis utilizing them should lead to maximum efficiency and accuracy with minimum cost to missile systems, avoiding over- or under-design. It further solves the problem of voluminous data input and eliminates data inhomogeneity and bias.

The technique of characteristic coefficients for wind profiles and how to derive models from them is not limited to the narrow field of wind analysis but has general application in statistical analysis. This has been demonstrated in some of the author's previous reports (4, 5).

LIST OF REFERENCES

1. ESSENWANGER, O. M., Mathematical Characteristics of Individual Wind Profiles, Army Missile Command Report No. RR-TR-64-12, December 1964.

2. ESSENWANGER, O. M., On Wind Profile Representation for Missile Design and Environment Analysis. In print in Proceedings of the Fourth National Conference on Aerospace Meteorology, 4-7 May 1970, Las Vegas (Nev), published by the Am. Meteor. Soc., Boston.

3. ESSENWANGER, O. M. and H. Boyd, On Methods to Establish a Set of Global Characteristic Wind Profiles, Army Missile Command Report No. RR-TR-70- , in publication.

4. ESSENWANGER, O. M., Profile Types of Sound Speed in the Lower Atmosphere and Their Relationships to Acoustic Focusing. Army Missile Command Report No. RR-TR-66-6, April 1966.

5. ESSENWANGER, O. M., Neue Konzepte fur die Vertikalstruktur in der Raketenklimatologie (New Concepts of Vertical Structure in Rocket Climatology), Raumfahrtforschung, H3, 1969, p. 101-106.

ACKNOWLEDGMENT

I am indebted to Mrs. Helen Boyd, who aided the investigation by establishing the IBM 7094 program for the assigning of model numbers to the empirical profile and the computation of the ϵ_{Δ}^2. Without her help the manuscript could not have been completed in time. I further wish to express my gratitude to Dr. Dorathy A. Stewart and Mrs. Helen Boyd for reviewing the manuscript. Mrs. Beatrice Wright deserves credit for typing the manuscript in record time.

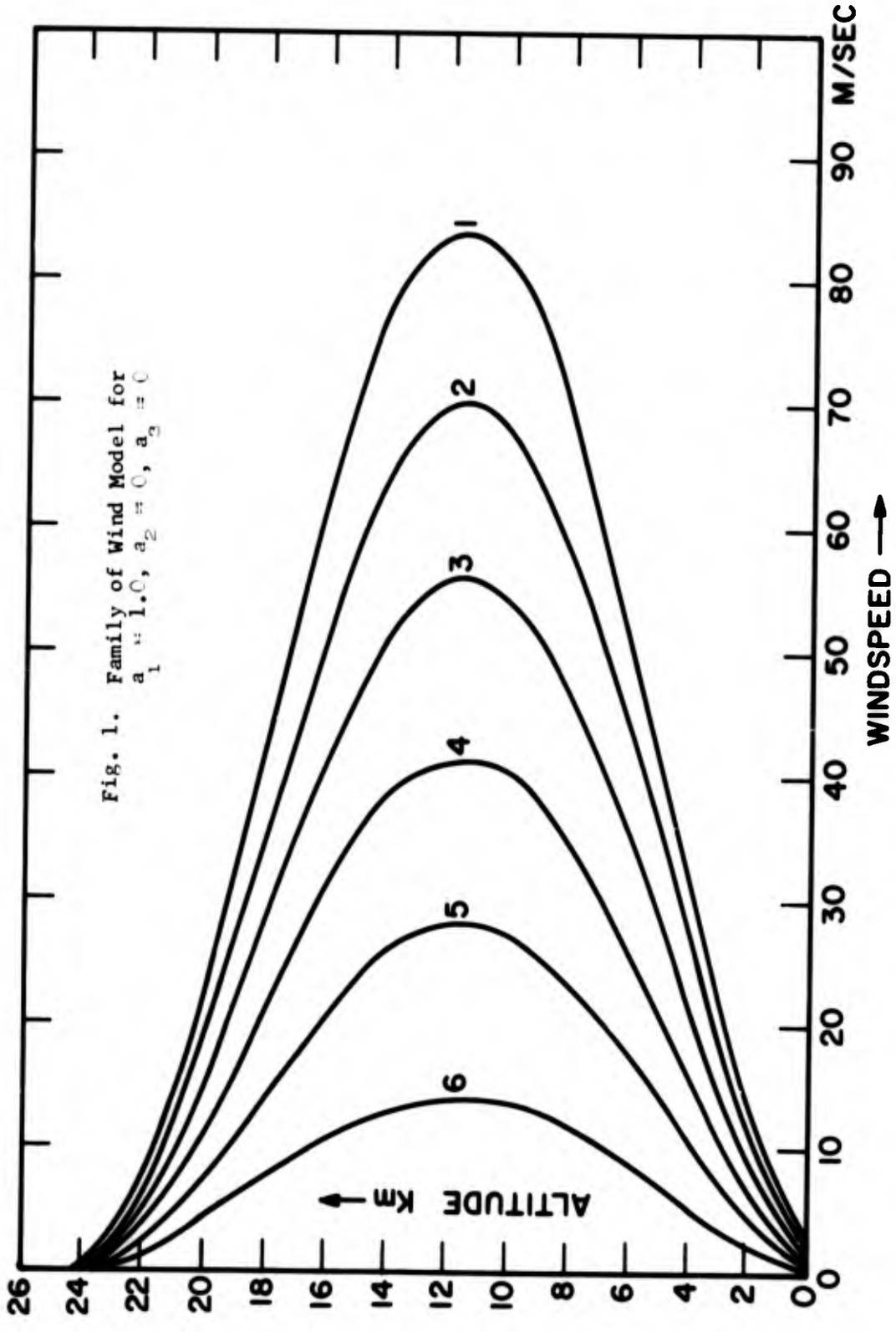


Fig. 1. Family of Wind Model for $a_1 = 1.0$, $a_2 = 0$, $a_3 = 0$

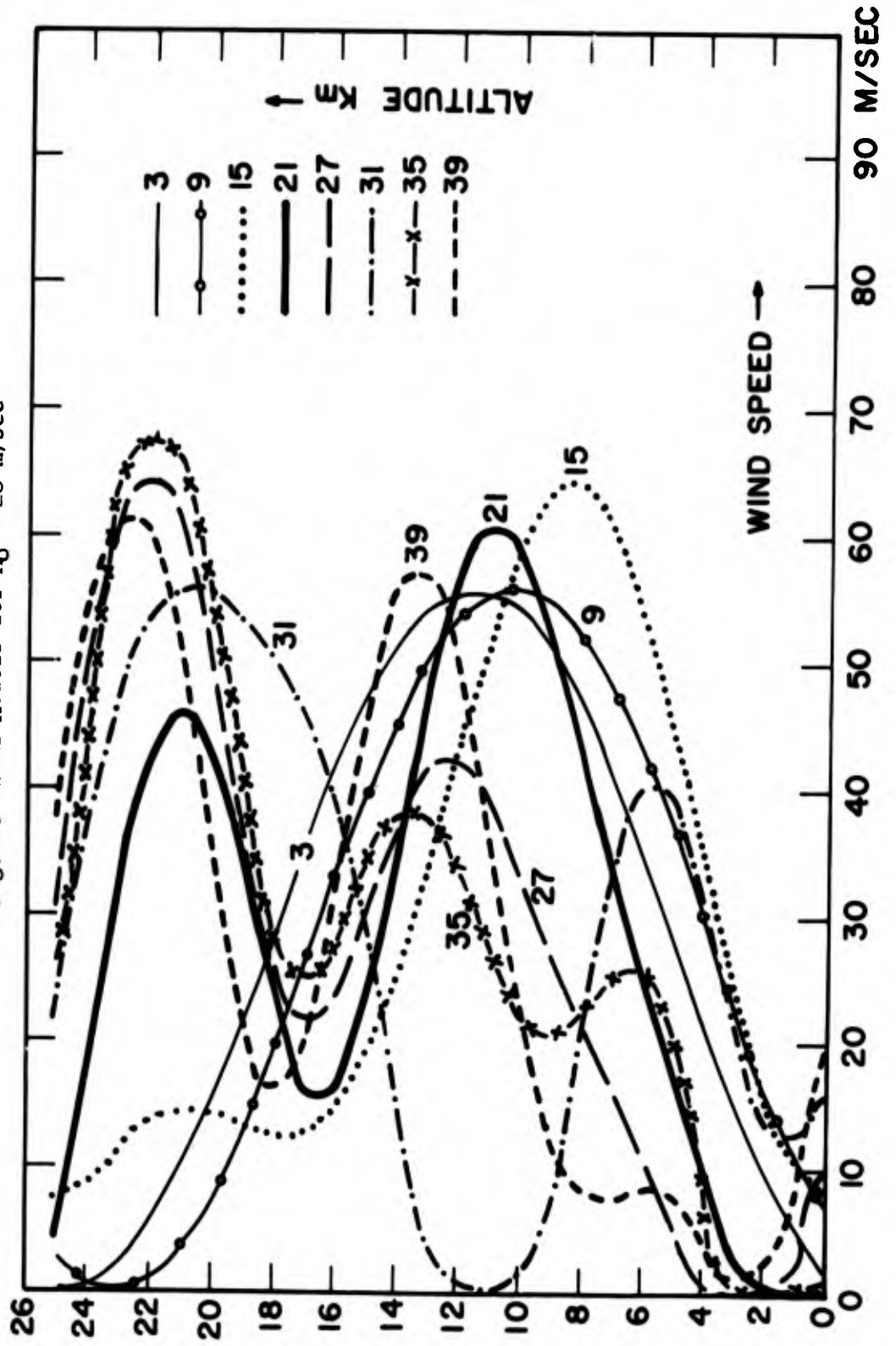


Fig. 2. Wind Models for A_0 = 28 m/sec

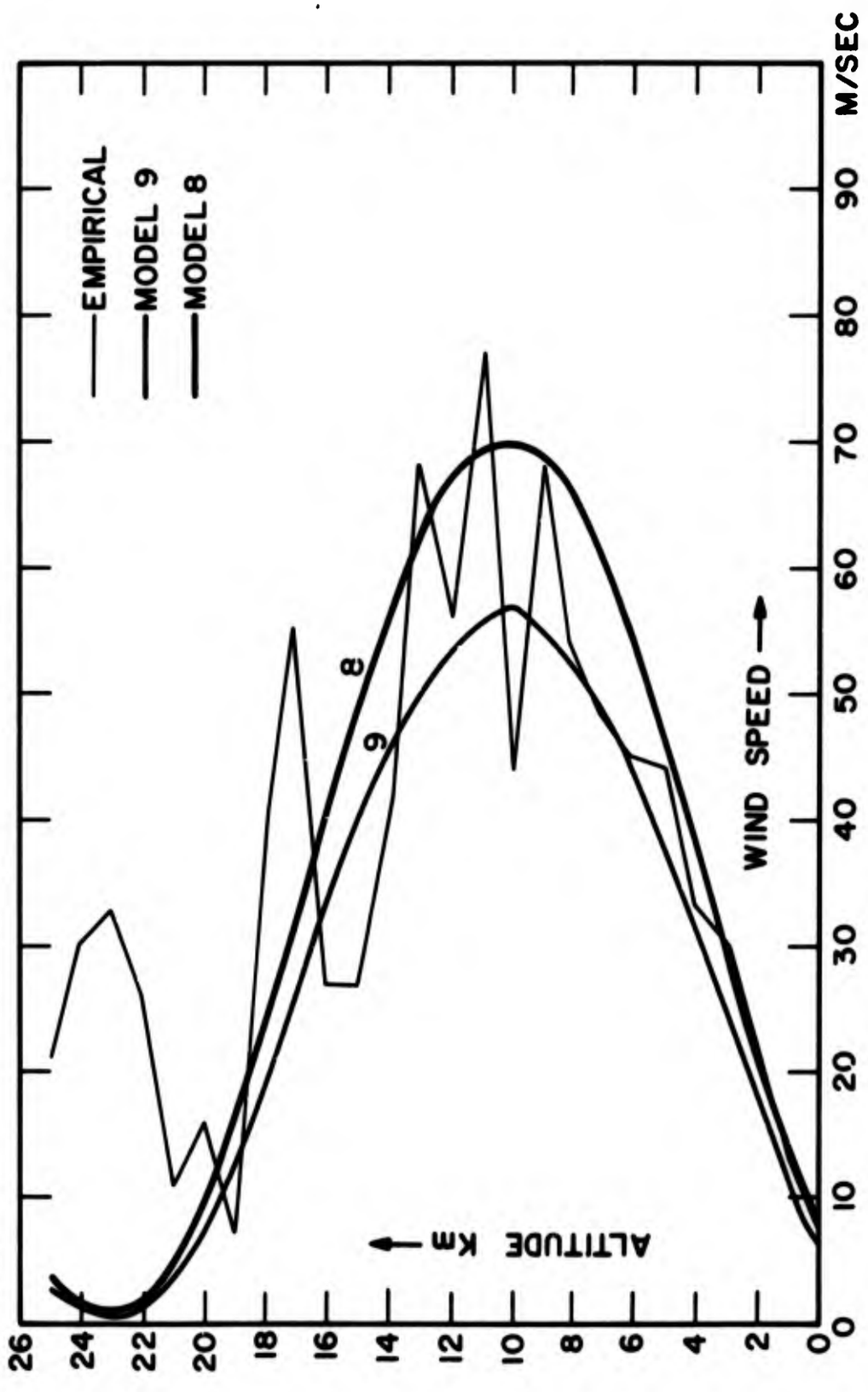


Fig. 3. Comparison of Empirical Wind Profile with Wind Models. 09 November 1956 at 09 Hour GMT.

AD 713 523

EXPERIMENTAL SUPERCONDUCTING ALTERNATORS WITH IRON-CORE AND IRON-FREE ARMATURES

JAMES H. FERRICK AND CARL J. HEISE
U.S. ARMY MOBILITY EQUIPMENT RESEARCH AND DEVELOPMENT CENTER
FORT BELVOIR, VIRGINIA

Introduction: The Army's ever increasing demand for lightweight, highly mobile electric power units has stimulated a research and development effort to adapt the emerging technology of applied superconductivity to electrical energy converters. The application to be discussed here is to rotating machinery, and in particular to alternators operating at frequencies up to 400 Hz. This paper describes a comparison test of two synchronous alternators having superconductive field windings. One alternator is made without ferromagnetic material; the other has a laminated iron core armature.

In reference 1, which describes the theoretical background for the design of the alternator without ferromagnetic material, referred to as the plastic core alternator, it was pointed out that adding iron to the armature would increase the output voltage of the machine. Design equations from which the effect of ferromagnetic material can be estimated were derived.

At the time the plastic armature alternator was made the main interest was to check the analytical investigation (reference 1) and it was decided to go ahead without the added complexity of having the ferromagnetic material. Since then a machine identical to the plastic armature alternator having a laminated iron core armature instead of plastic has been designed and constructed. The comparison test of the two machines shows that the predicted result of approximately twice the voltage generation capability was obtained with the iron core machine.

Two additional and important findings resulted from this test. When operating the plastic armature alternator over long time an appreciable temperature rise caused by eddy current loss in the armature copper conductors was measured. An identical test on the iron core armature alternator did not show this high temperature

rise, the reason being that the magnetic flux primarily flows in the iron.

The net effect of adding the iron core in this investigation resulted in a four-fold increase in power output.

Description of Experimental Alternators: The plastic armature alternator was manufactured several years ago, and a detail description is given in reference 2. Only the field of this alternator is superconductive, hence a stationary field-rotating armature type design was selected.

The alternator is vertically mounted and a cylindrical liquid helium reservoir (dewar) for supplying refrigeration to the field winding is mounted directly above the alternator as shown on figure 1. A vertical cylindrical extension in the bottom of the dewar contains the superconductive field winding, and the rest of the alternator consisting of the armature, shaft, bearings and housing, made as a unit, is mounted around the extension in such a way that the field will be located in its center. The superconductive field, which has four poles, consists of four coils of 1700 turns of 0.010 inch diameter niobium 25% zirconium wire, copper coated and formvar insulated. During operation of the alternator a torque will be transmitted to the field and it is therefore mounted on the end of a vertical support tube centrally located in the dewar and attached to the upper flange.

The rotating armature consists of a three-phase winding imbedded in plastic and forming a hollow cylinder. The winding is made up of 12 full pitch coils each having 38 turns of number 16 wire and connected in series Y giving 152 turns per phase. To give added strength against centrifugal force the cylinder containing the winding is banded with high strength wire, and to give axial stiffness it is mounted inside a sturdy fiberglas plastic cylinder. This two cylinder assembly forms the rotor of the machine and is mounted in bearings inside an aluminum housing. The generated power is transmitted to the output terminals via slip rings. The iron core armature alternator is made identical to the plastic armature alternator with the exception of the rotor. This makes it possible to use the two alternators with the same field coil dewar assembly.

The iron core rotor consists of a laminated iron core made up of .015 inch thick laminations mounted inside a steel cylinder which in turn is mounted in bearings inside an aluminum housing. The iron core has 24 slots for the three-phase windings which is made up of 24 full pitch coils each having 18 turns of two parallel number 19 wires and connected in series-parallel Y giving 72 turns per phase.

The design of the iron core armature was based on information given in reference 1.

For a four pole machine with iron core

$$B_{max} = \frac{\mu_o J_f}{\pi} \left(\left(\frac{R_2}{R_s} \right)^4 - \left(\frac{R_1}{R_s} \right)^4 \right) R_s$$

B_{max} = Flux density at the iron (Weber/m^2).

μ_o = Permeability in vacuum ($4\pi 10^{-7}$ H/m).

$$J_f = \frac{2N_f I_f}{R_2^2 - R_1^2}$$

N_f = Total number of series turns in field winding.

R_1 = Inside radius of field winding (in.).

R_2 = Outside radius of field winding (in.).

R_s = Inside radius of iron core (in.).

Calculating B_{max} for 25 amp field current gives .725 Weber/m^2 or 7.25K Gauss, which is more than twice as much as for the plastic core machine, hence requiring approximately half the number of turns per phase in order to get the same output voltage for the two machines.

Summary of Alternator Parameters:

Field: Consists of 4 coils each having 1700 turns

Overall length: 4.5 inches (11.4 cm)

Inner radius R_1 = .285 inch (.725 cm)

Outer radius R_2 = 1.09 inch (2.77 cm)

Inductance: 0.437 henry

Room temperature resistance: 1000 ohms

Critical current: 28 amperes

Armatures:

Plastic Core:		Iron Core:
12 coils, each 38 turns of number 16 wire, four coils per phase, 60° phase belts, series Y connected		24 coils, each 18 turns of 2 parallel number 19 wire, 8 coils per phase, 60° phase belts, series-parallel Y connected
Length of straight section: 3 inches		Length of straight section: 3 inches
Inner radius R_i = 1.44"		R_s = 1.44"
Resistance per phase at room temperature	.670 ohm	.185 ohm
Synchronous reactance on the base of 200 volts, 10 Amps for both machines	.23 p.u.	.18 p.u.
O.D.:	6.75"	O.D.: 6.75"
Length:	15"	Length: 15"
Weight:	40.0 lbs.	Weight: 48.5 lbs.

Test Methods and Results: Where applicable, the experimental alternators were characterized by tests conforming to MIL-STD-705A. The testing reported here was performed at 6000 RPM (200 Hz). Tests included no load saturation curves, rated load saturation curves for unity and zero power factor, voltage regulation curves for unity and zero power factor, short circuit tests, harmonic analysis, and armature temperature rise tests. The saturation curves shown in figures 2 and 3 reveal that the iron core alternator is beginning to exhibit some saturation effect, while the plastic core alternator shows only a slight non-linearity, this due to flux leakage between the superconducting field coils. The non-zero terminal voltage for zero field current shown for the plastic core alternator demonstrates a property of superconducting coils, viz., they generally exhibit some remnant field after they have been energized and de-energized while maintaining the superconductive state.

The voltage regulation curves for both machines (figures 4 and 5) show good inherent regulation characteristics. The full field curves indicate regulation to within 6% for unity power factor loading to 20 amperes and to within 13.5% for zero power factor loading to 20 amperes for the iron core alternator; for the plastic core machine the regulation percentages are 10% and 13% for 10 ampere loading.

The short circuit tests were taken mainly to determine the synchronous reactance. Beyond this, they indicate excessive current would result from short-circuiting at full field.

Table 1 gives the relative harmonic content of the output voltage waveform. Relative harmonic content is very low except for the third harmonic in line-to-neutral voltage; no effort was made to reduce the harmonic content by design.

TABLE 1: HARMONIC ANALYSIS

Iron Core Armature Alternator

Harmonic	No Load		Resistive Load		Reactive Load	
No	L-L	L-N	L-L	L-N	L-L	L-N
1	100%	100%	100%	100%	100%	100%
3	0	5.0	0	5.4	0	2.3
5	0.8	0.8	0.7	0.7	1.2	1.2
7	0.9	0.9	0.9	0.9	0.7	0.7
9	0	1.1	0	0.9	0	0.8
11	0.4	0.4	0.4	0.4	0.4	0.4

Plastic Core Armature Alternator

1	100%	100%	100%	100%	100%	100%
3	0	6.3	0	5.4	0	6.5
5	0.2	0.2	0.3	0.2	0.2	0.2
7	0	0	0.2	0.2	0	0
9	0	0	0	0.1	0	0
11	0	0	0.1	0.1	0	0

Field current set at 25 amps throughout the test.

Figure 6 shows the measured angular variation of the maximum radial component of magnetic field, and figure 7 details the radial dependence of the maximum radial component of magnetic field. The output waveforms are shown in figure 8 and the third harmonic effect is clearly visible in the line-to-neutral waveforms. Additionally, saturation effects can be noted on the iron core output waveforms.

The temperature rise data (fig. 9) for the iron core alternator follow an expected trend, increasing approximately as the square of load current. The iron core alternator could conservatively be rated at 22 amperes with an 85°C armature temperature rise.

In attempting to characterize the plastic core armature a strong dependence of average armature winding temperature on applied field was noted (fig. 10). This is attributable to eddy current losses in the copper conductors; in the plastic core armature the windings are exposed to the maximum magnetic field while in the iron core armature the windings are shielded by the teeth in the laminations. Eddy current losses in the plastic core armature are found to be 138 watts at full field. The armature heating due to the eddy current losses restricted steady loading to four amperes, this at an average armature temperature rise of 97.5°C, where the machine is rated for steady operation at a 90°C temperature rise.

The maximum steady power generated was 7.3 KW with the iron core armature and 1.72 KW with the plastic core armature, both at unity power factor.

Conclusions:

1) The design equations developed in this effort are found to be reliable. The predicted output voltage for the iron core alternator was 1.9 times that of the plastic core alternator for equal numbers of series turns; this was achieved.

2) When designing superconducting machines in which the copper windings are exposed to the full field, care must be taken to reduce the eddy current loss problem in the copper.

3) The iron core improves heat conduction from the copper windings to the cooling medium.

4) The steel structure improves mechanical integrity and reduces construction complexities.

5) The temperature rise data for the iron core armature, if extrapolated, indicate that a rating upward to approximately 33 amperes would be within the acceptable thermal limits of the Class H insulation used (operation at 180°C). This would result in a machine rating in excess of 20KVA at 12,000 RPM and unity power factor, with no additional cooling improvement considered.

Acknowledgment: We express our appreciation to Mr. Philip Pemberton for his efforts in providing valuable technical assistance in machine performance characterization.

References

1. "Research and Design of Superconductive and Cryogenic Electric Power Generator," Final Technical Report on Contract #DA-44-009-AMC-236, Avco-Everett Research Laboratory, January 1965.

2. Z.J.J. Stekly, H.H. Woodson, A.M. Hatch, L.O. Hoppie, and E. Halas, "A Study of Alternators with Superconducting Field Windings: II-Experiment." IEEE Trans. on Power Apparatus and Systems, Vol PAS-85, No. 3,March 1966, pp. 274-280.

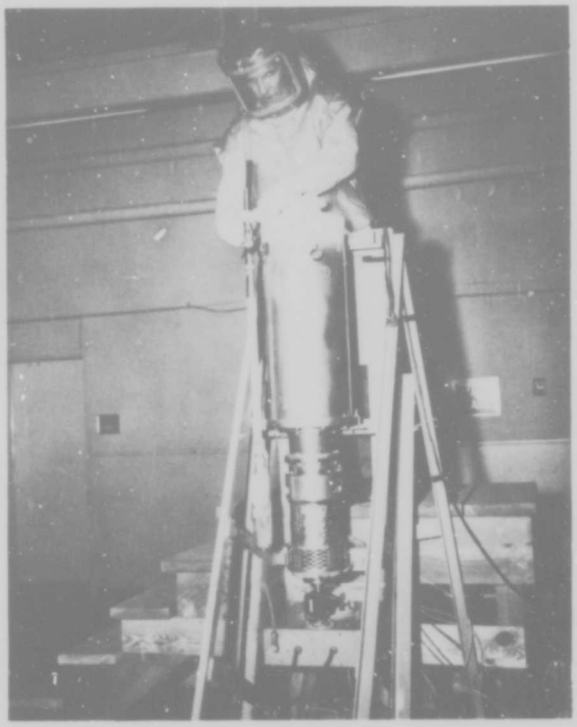
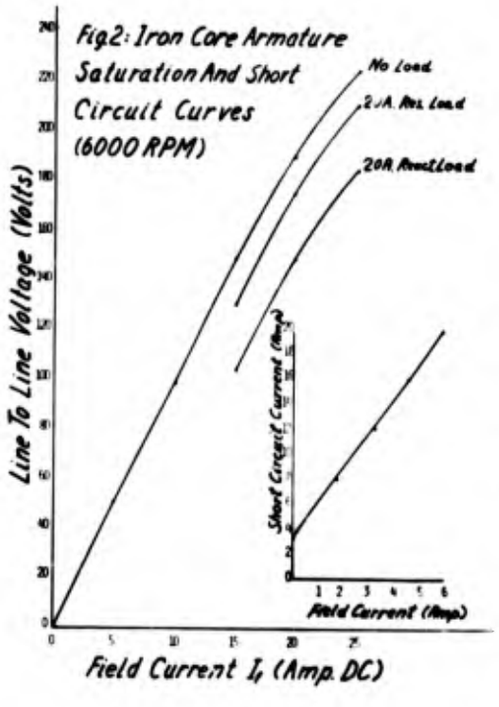
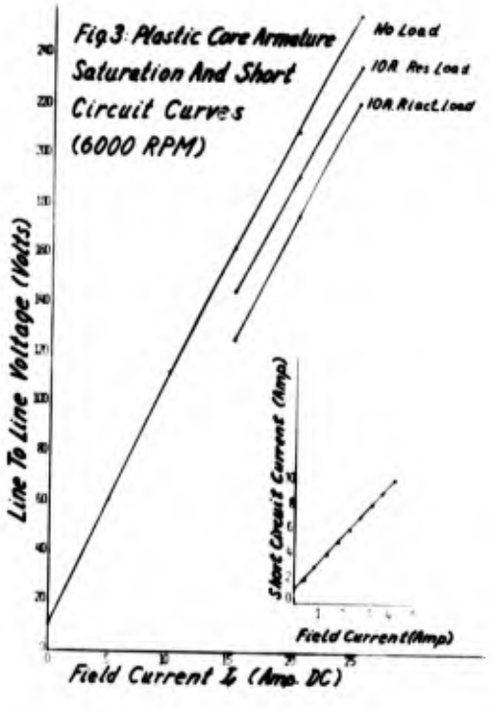


Figure 1. Superconducting Alternator Test Assembly





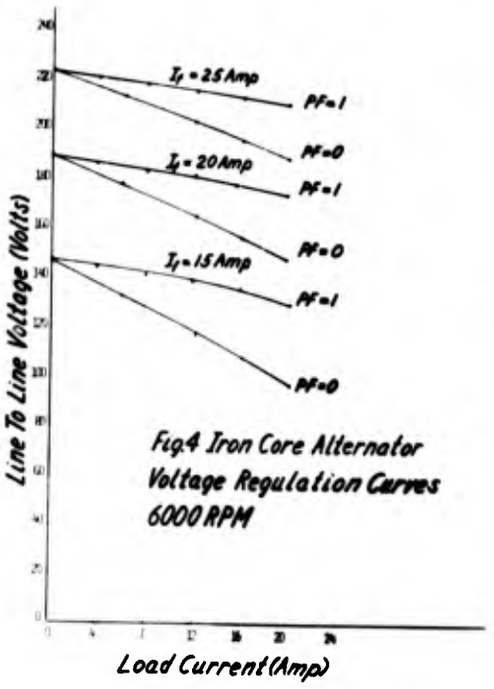


Fig.4 Iron Core Alternator Voltage Regulation Curves 6000 RPM

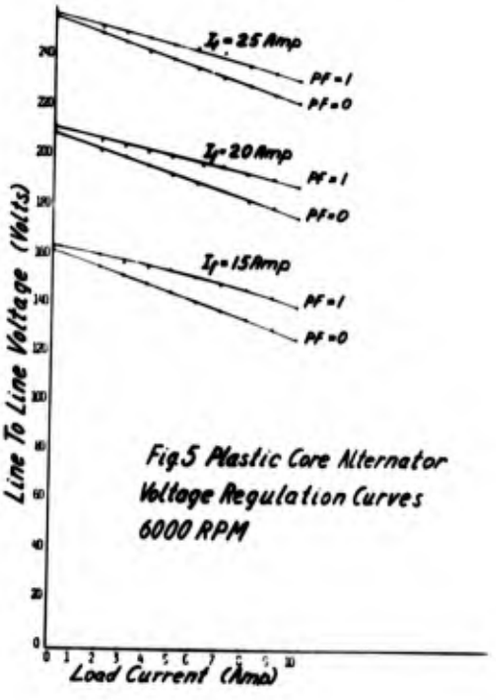


Fig.5 Plastic Core Alternator Voltage Regulation Curves 6000 RPM

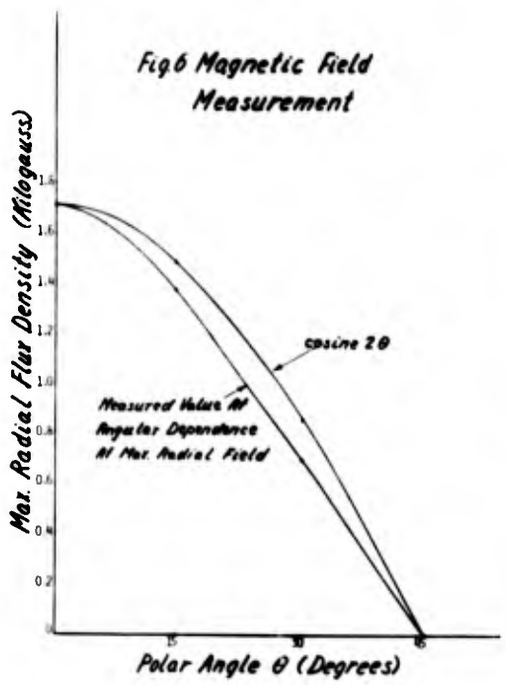


Fig 6 Magnetic Field Measurement
Max. Radial Flux Density (Milogauss)
cosine 2θ
Measured Value of Angular Dependance of Max. Radial Field
Polar Angle θ (Degrees)

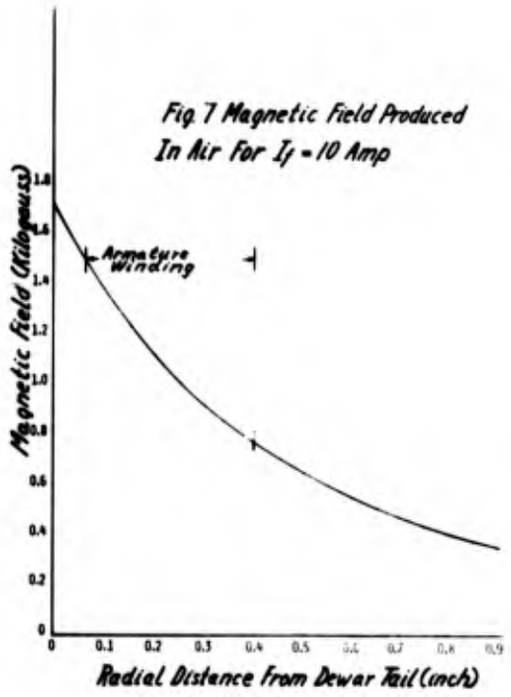
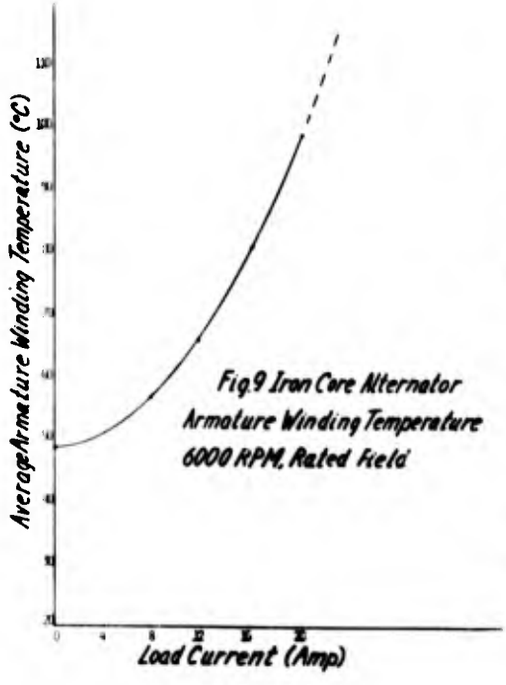
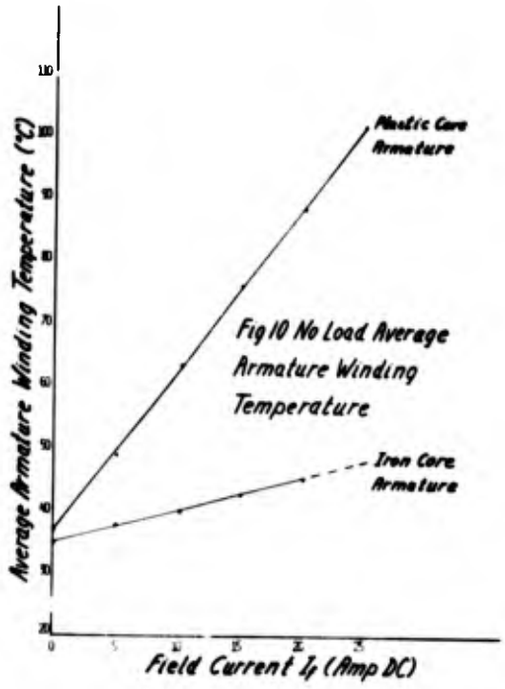


Fig 7 Magnetic Field Produced In Air For If = 10 Amp
Magnetic Field (Kilogauss)
Armature Winding
Radial Distance from Dewar Tail (inch)

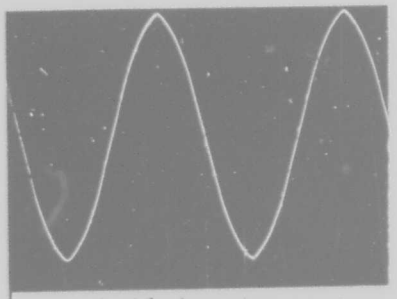

Average Armature Winding Temperature (°C)
Fig 9 Iron Core Alternator Armature Winding Temperature 6000 RPM, Rated Field
Load Current (Amp)


Average Armature Winding Temperature (°C)
Plastic Core Armature
Fig 10 No Load Average Armature Winding Temperature
Iron Core Armature
Field Current If (Amp DC)

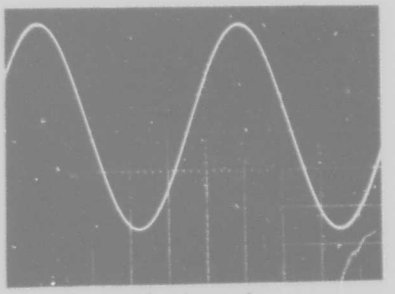
Iron Core

Fig.8

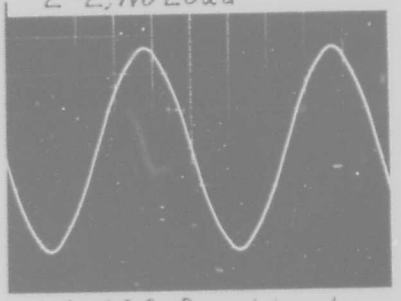
Plastic Core



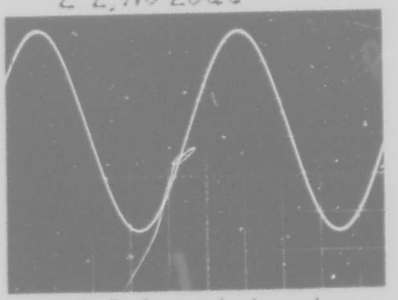
L-L, No Load



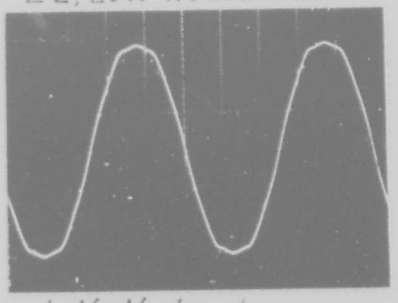
L-L, No Load



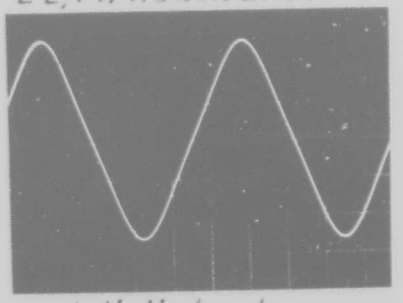
L-L, 20A React. Load



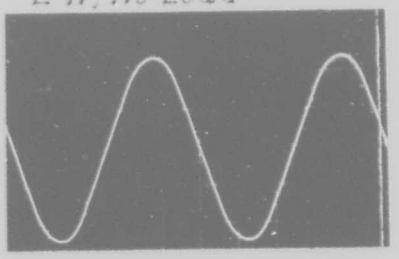
L-L, 4 A React. Load



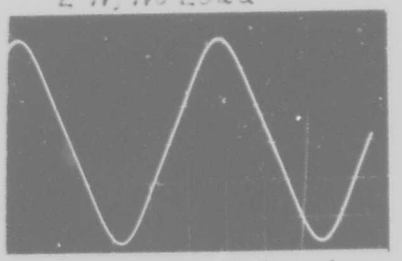
L-N, No Load



L-N, No Load



L-N, 20A React. Load



L-N, 4 A React. Load

AD 713 524

A RESEARCH CONCEPT FOR THE INTERPRETATION OF HUMAN MISSILE WOUNDS BY THE PATHOLOGIST

PIERRE A. FINCK, COL, MC, USA
ARMED FORCES INSTITUTE OF PATHOLOGY
WASHINGTON, D. C. 20305

This paper emphasizes the value of human cases in the interpretation of wounds caused by missiles such as bullets or fragments and recommends the types of information to be studied and correlated by the pathologist.

The experiments on animals or gelatin blocks have brought information on the "temporary cavity" and thus contributed immensely to the understanding of the mechanism of wounding. The temporary cavity within the tissue lasts a few milliseconds but creates pressures that are often in excess of 1,000 lbs/in.^{2} (70 kg/cm^2). Such a phenomenon, which increases with the velocity of the projectile, explains lesions or functional disturbances at a distance from the permanent missile track observed by the surgeon or the pathologist. Although human missile wounds do not allow a study of the temporary cavity, they provide a valid basis for investigating the effects of a known projectile on tissue by examining the wound track and the consequences of the temporary cavity that is away from the permanent path of the missile.

In a PENETRATING WOUND the anatomic structure shows an entry but no exit. The recovered missile should be submitted to ordnance experts for identification. Ordnance specialists can often estimate the striking velocity of a fragment when they know its weight and the distance between the point of detonation (POD) and the victim.

The opinions or assertions contained herein are the private views of the author and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

In a PERFORATING (through-and-through) WOUND the anatomic structure shows an entry and an exit. In some cases the projectile breaks up, its fragments are retained within the body, and a wound of exit is caused by a fragment from the primary missile or by a secondary missile such as a bone fragment set into motion by the primary missile. Such an injury is a "PERFORATING WOUND WITH RETENTION OF FRAGMENTS."

If there is an entry and an exit but no retention of missile, further study of the case will be of limited value unless the distance between the victim and the point of origin of the projectile is known.

GUIDE FOR THE DATA TO BE RECORDED: Regardless of the chosen dimensions--3 X 5, 5 X 8, or 8 X 10-1/2 inches--the summary should be limited to one side only. The summary, with a case number, can thus be reproduced without a loss of identification.

Below is a guide for the recording of important data from one case, which in most instances can be done on one side of a 3 X 5-inch card.

GUIDE FOR 3 X 5-INCH CARD FOR SUMMARIZING CASES OF MISSILE WOUNDS

CASE NUMBER MISSILE OR WEAPON

1. Race/Sex/Age.
2. Manner of death: died of wounds (DOW), killed in action (KIA); accident; homicide; or suicide.
3. Distance from weapon or point of detonation.
4. Interval between injury and death (survival).
5. Wounds of body areas, penetrating or perforating, with wounds of organs or tissues, penetrating or perforating; fractures.
6. Missiles recovered from organs or tissues: identification, shape, weight, and striking velocity.

In the upper right corner of the card is stated the causative agent: weapon or missile: This heading should include a category and a subcategory, for example: RIFLE, AK-47, or BULLET, 7.62-39mm. In this last example, 7.62 refers to the caliber of the bullet in millimeters and 39mm indicates the length of the cartridge case.

MISSILE CHARACTERISTICS COMPARED TO PENETRATION AND PERFORATION OF ANATOMIC STRUCTURES: A study of human cases showed that high velocity, i.e., at least 2,500 feet per second (ft/sec) or 750 meters per second (m/sec) (1), does not necessarily mean perforation. Military jacketed bullets of high velocity often break up within the human body and do not always produce a wound of exit. The fragments of the M26 Fragmentation Hand Grenade weigh less than 1 gram. Although their initial velocity is 5,000 ft/sec (2), or 1,524 m/sec, most of them are retained in the body even when the POD is within 1 meter. The steel spheres of the M18A1 Claymore Antipersonnel Mine weigh 10.4 grains (0.68 gram) and have an initial velocity of 3,800 ft/sec (1,158 m/sec) (3). At a POD within 1 meter from the fatal casualty, several spheres remain in the body.

NOMENCLATURE OF VELOCITIES: The concept of low and high velocities may vary according to the kinds of cases involved or to the type of research. For example, ordnance experts call "ordnance velocities" those up to 5,000 ft/sec for projectiles from conventional weapons. Table I shows categories of velocities used at the Ballistic Research Laboratories of Aberdeen Proving Ground:

Table I

V E L O C I T I E S

	Ft/sec	M/sec
Ordnance velocities	Up to 5,000	Up to 1,500
High velocities	5,000 to 10,000	1,500 to 3,000
Hypervelocities	Over 10,000	Over 3,000

TISSUE STUDIES: The pathologist should study the wounds as well as the tracks of the missiles radiologically, macroscopically, and microscopically and correlate his observations with the overall aspects of the injuries and the physical characteristics of the missiles.

CONCLUSIONS

The pathologist can make a contribution to the understanding of the wounding power of projectiles. In his study of human wounds caused by missiles, he should correlate the following data:

1. Composition, shape, weight, and striking velocity of missile.

2. Penetrating and perforating wounds of body and organs.

3. Radiologic, photographic, macroscopic, and microscopic aspects of wounds and their tracks.

In research projects, a pathologic examination of tissue without a complete wound track and without the knowledge of the causative missile is not recommended, since it is impossible to reach significant conclusions.

REFERENCES

1. Beyer, J. C. (Ed.): Wound ballistics, Office of the Surgeon General, Department of the Army, Washington, D. C., 1962. (For sale by the Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. 20402.)

2. Aberdeen Proving Ground Firing Record B9863, 1952, and B12328, 1954.

3. Aberdeen Proving Ground Firing Record B15215, April 1966.

AD 713 525

ADVANCED COMPUTATIONAL ALGORITHMS FOR LARGE SCALE, THREE DIMENSIONAL, ARTILLERY SURVEY APPLICATIONS

LAWRENCE A. GAMBINO
US ARMY TOPOGRAPHIC COMMAND, CORPS OF ENGINEERS
ENGINEER TOPOGRAPHIC LABORATORIES
FORT BELVOIR, VIRGINIA 22060

INTRODUCTION: It is the purpose of this paper to demonstrate how a newly derived set of computational algorithms allows complete flexibility and rigor in solving for tracking station coordinates and their associated error models in a large, simultaneous, three dimensional adjustment. The algorithms will be used to solve a hypothetical, artillery survey problem since the hardware is currently in the R&D stages at the US Army Engineer Topographic Laboratories (USAETL). The new system is called the Long Range Position Determining System (LRPDS), and it is being investigated by engineers in the Surveying and Geodesy Division, USAETL for its application in the artillery problem. Personnel in the Computer Sciences Laboratory, USAETL, are investigating data reduction schemes for the future reduction of LRPDS data, and will soon be performing computer simulations using the algorithms to be described in this paper. The new algorithms are considered to provide a major computational breakthrough for efficiently handling very large, Second Order Regression Schemes, which are to be defined later, and they allow the engineer to extend his hypotheses associated with problems of systematic errors. It is the purpose of this paper to develop a Second Order Regression Scheme for LRPDS and to show that it is computationally feasible to invert this large system of equations for its solution and error propagation.

LRPDS GEOMETRY: The geometry for a possible LRPDS artillery survey problem is shown on page 6. Here we have shown a conventional artillery area in which we find 25 forward observers, whose positions are unknown, and 5 base stations, whose positions are known. The 5 known stations are not a requirement since we could begin with only 3 known stations. Each of the 30 stations has a transceiver, and the aircraft carries a special transceiver and transmitter which continuously transmits a signal. Note that all the ground stations can be passive; that is, no transmissions. Details about the hardware can be found in reference 1 (Kiisk, 1969), where it will be noted that unless a reference distance can be established prior to

the aircraft leaving the ground, the LRPDS is an ambiguous range system. Under these circumstances, the data reduction scheme must carry at least a bias error model term in the observation equation for each one of the trackers.

Under this geometric situation, we will allow the following hypotheses:

1. Thirty pieces of ground hardware are available.
2. LRPDS is unsaturable.
3. Ambiguous range in each tracking channel.
4. Frequency bias in each tracking channel.
5. Frequency drift in each tracking channel.
6. Error model parameters in 3, 4, and 5 change for each pass of the aircraft; that is, we will reinitialize these parameters on each pass.

It is the 6th hypothesis which leads to a Second Order Regression Scheme while allowing more flexibility in system analysis. Without this hypothesis, we obtain a First Order Regression Scheme.

OBSERVATION EQUATION: Our basic observation equation is that which describes the distance between two points in three space. It is of the form:

$$R_{ijk} = \sqrt{(X^c_{jk} - X_i)^2 + (Y^c_{jk} - Y_i)^2 + (Z^c_{jk} - Z_i)^2} \qquad (1)$$

where the subscript i (i = 1,2,...,m) denotes the i th tracking station, j (j = 1,2,...,n) denotes the j th aircraft position at a given instant of time, and k (k = 1,2,...,ℓ) denotes the k th pass of the aircraft. Therefore, equation (1) is our ranging equation in the absence of systematic error and it arises from tacking by the i th ground station to the j th aircraft position at time t_j on the k th pass.

The error model is given in the following form since it is not the intent of this paper to give the functional forms associated with each unknown error model parameter:

$$\Delta R_{ik} = a_o u_{o_{ik}} + a_1 u_{1_{ik}} + \ldots + a_p u_{p_{ik}} \qquad (2)$$

in which the error coefficients (the a's) are unknowns and the u's are known functions of known arguments such as (1) adopted value of transmitted frequency, (2) adopted value of local reference frequency, (3) bias in the adopted value of the transmitted frequency at the start of the pass, (4) bias in the adopted value of the reference frequency at the start of the pass, (5) drift rates of the transmitted and reference frequencies, (6) the zero set, or ambiguity associated with each tracker, and (7) residual refraction remaining after the initial corrections. There are other corrections which can be included. In any event, we have our "p" parameter error model, (q = 1,2,...,p), representing the systematic error in the i th ranging channel for the k th pass subject to the hypothesis made earlier.

The true range on the left side of equation (1) can now be

expressed as

$$R_{ijk} = R^{\circ}_{ijk} + \Delta R_{ik} + v_{ijk} \tag{3}$$

where the v's are the observational residuals, R°_{ijk} is the observed range, and ΔR_{ik} has been previously defined. The equations are nonlinear in the coordinates of the aircraft positions and of the tracking stations. If we linearize (3) by using the Taylor series expansion, we have

$$R_{ijk} = R^{\circ\circ}_{ijk} + \left[\frac{\partial R_{ijk}}{\partial (X^c_{jk},\ Y^c_{jk},\ Z^c_{jk},\ X_i,\ Y_i,\ Z_i)}\right] \cdot \begin{bmatrix} \delta X^c_{jk} \\ \delta Y^c_{jk} \\ \vdots \\ \delta Z_i \end{bmatrix} \tag{4}$$

$$\text{and } \Delta R_{ik} = \Delta R^{\circ\circ}_{ik} + \left[\frac{\partial \Delta R_{ik}}{\partial (a_{0_{ik}},\ a_{1_{ik}},\ \ldots,\ a_{p_{ik}})}\right] \begin{bmatrix} \delta a_{0_{ik}} \\ \delta a_{1_{ik}} \\ \vdots \\ \delta a_{p_{ik}} \end{bmatrix} \tag{5}$$

where the superscript "$\circ\circ$" denotes the range and systematic error as computed from initial approximations of aircraft positions $\left[(X^c)^{\circ\circ},\ (Y^c)^{\circ\circ},\ (Z^c)^{\circ\circ}\right]_{jk}$, ground trackers $\left[X^{\circ\circ},\ Y^{\circ\circ},\ Z^{\circ\circ}\right]_i$, and error model parameters $\left[a_0^{\circ\circ},\ a_1^{\circ\circ},\ \ldots a_p^{\circ\circ}\right]_{ik}$ and the δ's denote the unknown corrections to these approximations. Again it is not a concern of this paper to actually take the partial derivatives, but we note that the partial of the systematic error function with respect to the error model parameters is linear with respect to these parameters.

The matrix notations to be used in furthering this development are those used in reference 2 (D.C. Brown). We begin our partitioning process by denoting the partial derivatives of the range in (4) with respect to the aircraft positions as follows:

$$\underset{(1,3)}{\dot{B}_{jk}} = \left[\frac{\partial R_{ijk}}{\partial (X^c_{jk},\ Y^c_{jk},\ Z^c_{jk})}\right] \tag{6}$$

and the associated correction vector as

$$\underset{(3,1)}{\dot{\delta}_{jk}} = \left[\delta X^c_{jk},\ Y^c_{jk},\ Z^c_{jk}\right]^T \quad . \tag{7}$$

Matrix equation (6) implies we have range data from the i th tracking station to the j th aircraft position on the k th pass. Similarly, the partials with respect to the ground trackers are given by:

$$\underset{(1,3)}{\hat{B}_i} = \left[\frac{\partial R_{ijk}}{\partial (X_i,\ Y_i,\ Z_i)}\right] = \underset{(1,3)}{-\ddot{B}_{jk}} \tag{8}$$

and the associated correction vector as

$$\underset{(3,1)}{\hat{\delta}_i} = \left[\delta X_i,\ \delta Y_i,\ \delta Z_i\right]^T \quad . \tag{9}$$

We do the same for the error model parameters;

$$\underset{(1,p)}{\ddot{B}_{ik}} = \left[\frac{\partial \Delta R_{ijk}}{\partial (a_{0_{ik}},\ a_{1_{ik}},\ \ldots,\ a_{p_{ik}})}\right] \tag{10}$$

and

$$\underset{(p,1)}{\ddot{\delta}_{ik}} = \left[\delta a_{0_{ik}},\ \delta a_{1_{ik}},\ \ldots,\ \delta a_{p_{ik}}\right]^T \quad . \tag{11}$$

Again we see that data from the i th tracker to the j th aircraft position on the k th pass contributes to the solution of error model parameters, which are allowed to be different for each pass of the aircraft.

If we substitute equations (4) and (5) into equation (3) and further substitute the corresponding partitioned matrix equations, we have the following linearized, matrix observation equation:

$$\underset{(1,1)}{v_{ijk}} - \underset{(1,3)}{\hat{B}_i}\ \underset{(3,1)}{\hat{\delta}_i} - \underset{(1,3)}{\dot{B}_{ik}}\ \underset{(3,1)}{\dot{\delta}_{ik}} - \underset{(1,p)}{\ddot{B}_{jk}}\ \underset{(p,1)}{\ddot{\delta}_{jk}} = \underset{(1,1)}{\varepsilon_{ijk}} \tag{12}$$

where "ε_{ijk}" is the discrepancy vector denoting the difference

between the observed range and the range computed using approximate values of the unknown parameters; that is,

$$\varepsilon_{ijk} = R^{oo}_{ijk} - R^{o}_{ijk} - \Delta R^{oo}_{ijk} \quad . \tag{13}$$

We will take the covariance matrix of the random error in the observed range, R^{o}_{ijk}, to be a simple 1 x 1 matrix and denote it as

$$\underset{(1,1)}{\Lambda_{ijk}} = \begin{bmatrix} \sigma^2_{R^{o}_{ijk}} \end{bmatrix} \quad . \tag{14}$$

Therefore, the weight matrix of the observed quantities is defined to be

$$\underset{(1,1)}{W_{ijk}} = \Lambda^{-1}_{ijk} \tag{15}$$

and the a priori weight constraints associated with the parameters are given as $\hat{W}_i$, $\dot{W}_{ik}$, and $\ddot{W}_{jk}$ for the a priori weights on station coordinates, error model parameters and aircraft positions, respectively. We will assume that these weight matrices are incorporated into the solution and will not discuss the theoretical aspects of these matrices in this paper.

At this point, we have in matrix equation (12) the smallest of the hierarchy of matrices which are required for a single observation from the i th tracker to the j th aircraft position on the k th pass. Space does not permit us to show the patterning in these lower echelon matrices, but it is with this knowledge that efficient computational algorithms begin to evolve. The store of information for two hierarchies of matrices above that for the matrix observation equation (12) is given as follows:

1. All m stations (i = 1,2,...,m) and all n aircraft coordinates (j = 1,2,...,n) for the k th pass (k = 1):

   a. v is (mn,1) column vector.

   b. $\hat{B}$ is (mn,3m) column matrix of matrices where each component is a $\hat{B}_{jk}$ matrix, which is (m,3m) diagonal matrix of (1,3) matrices.

   c. $\hat{\delta}$ is (3m,1).

   d. $\dot{B}$ is (mn,mp) column matrix of matrices where each component is a $\dot{B}_{jk}$ matrix, which is (m,mp) and the same shape as $\dot{B}_{jk}$.

   e. $\dot{\delta}$ is still (mp,1).

f. $\ddot{B}$ is (mn,3n) diagonal matrix of matrices where each component is a $\ddot{B}_{jk}$ matrix, which is (m,3) column matrix of (1,3) matrices.

g. $\ddot{\delta}$ is (3n,1) column vector.

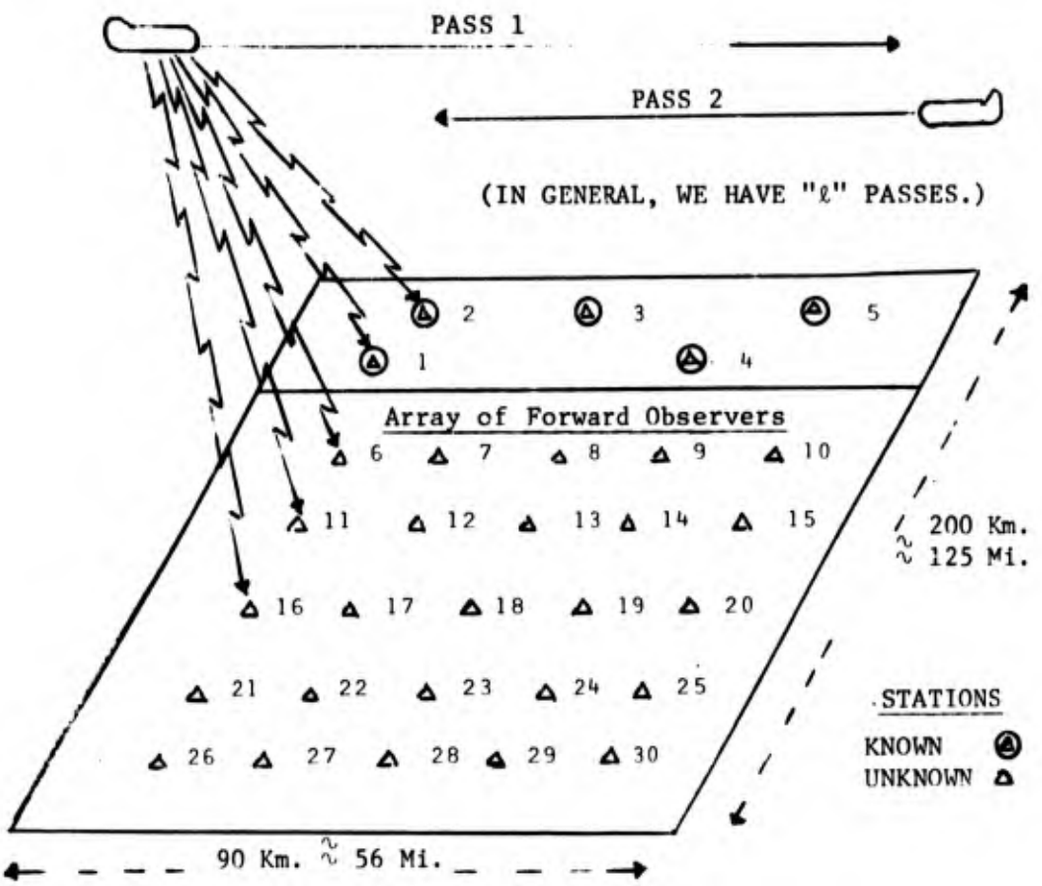
h. ε is (mn,1), same as v.

This hierarchy of matrices is given by the matrix equation

$$v_k - \hat{B}\,\hat{\delta} - \dot{B}_k\dot{\delta}_k - \ddot{B}_k\ddot{\delta}_k = \varepsilon_k \quad . \tag{16}$$

The covariance and weight matrices of the observations are (mn,mn) diagonal matrices. Thus, equation (16) contains the entire store of information for the k th pass of the aircraft where only the tracking station coordinates are common to all "ℓ" passes (k = 1,2,...,ℓ) of the aircraft.

To solve the problem depicted in the following diagram, we



assume that we will need at least 10 passes and 120 observations per pass with at least 4 trackers participating in each observation

(since LRPDS is unsaturable, we may conceivably obtain many more simultaneous observations at a given instant) and, in all cases, at least 3 known stations participating in each interrogation. We also must solve for at least the ambiguous range parameter for each tracker on each pass. For convenience, we will assume that 6 trackers participate at each of the 120 interrogations per pass and that these trackers are in such combinations so as to have all 30 stations participate during the course of a single pass. We will also allow 4 error model parameters per tracker (zero set, frequency bias, frequency drift, residual refraction). Therefore, we have in general

$$q = 3m + \ell(mp + 3n) \tag{17}$$

parameters, where

q = total number of parameters,
m = total number of trackers (30),
ℓ = total number of passes (10),
p = total number of error model parameters per tracker (4),
n = total number of aircraft positions per pass (120).

Therefore, we must solve for 4890 parameters simultaneously. However, we have "ℓmn" equations where "m" is now the number of trackers participating simultaneously at each interrogation. This leads to 7200 equations which is more than enough to solve the problem. Without knowledge of the aforementioned patterning, and the advanced solution algorithms to be introduced shortly, the problem seems horrendous indeed for we will be asked to invert a matrix of order 4890.

NORMAL EQUATIONS: We can rewrite equation (17) for the k th pass in the following matrix form:

$$\begin{bmatrix} v_k \end{bmatrix} - \begin{bmatrix} \hat{B} & \dot{B}_k & \ddot{B}_k \end{bmatrix} \begin{bmatrix} \delta \\ \dot{\delta}_k \\ \ddot{\delta}_k \end{bmatrix} = \begin{bmatrix} \varepsilon_k \end{bmatrix} \tag{18}$$

where for our problem outlined earlier, we have for the k th pass, v_k is (720,1), $\hat{B}$ is (720,90), $\hat{\delta}$ is (90,1), $\dot{B}_k$ is (720,120), $\dot{\delta}_k$ is (120,1), $\ddot{B}_k$ is (720,360), $\ddot{\delta}_k$ is (360,1) and ε_k is (720,1). Note that these dimensions refer only to a single pass, i.e., the k th pass. For all ten passes, v is (7200,1), $\hat{B}$ is (7200,90), $\hat{\delta}$ is (90,1), $\dot{B}$ is (7200,1200), $\dot{\delta}$ is (1200,1), $\ddot{B}$ is (7200,3600), $\ddot{\delta}$ is (3600,1), and ε is (7200,1). We begin to realize the importance of patterning in the system of observation equations and in the forthcoming normal equation system.

If we use equation (18) for the k th pass to form normal equations and later use the rule that independently formed sets of

normal equations involving the same vector of unknowns on additional passes can be added together for the simultaneous adjustment of merged sets of equations, we form the coefficient matrix and constant column of the normal equations by matrix multiplication as follows:

$$\begin{bmatrix} \hat{B}^T \\ \dot{B}_k^T \\ \ddot{B}_k^T \end{bmatrix} \begin{bmatrix} W_k \end{bmatrix} \begin{bmatrix} \hat{B} & \dot{B}_k & \ddot{B}_k \end{bmatrix} \begin{bmatrix} \hat{\delta} \\ \dot{\delta}_k \\ \ddot{\delta}_k \end{bmatrix} = \begin{bmatrix} \hat{B}^T \\ \dot{B}_k^T \\ \ddot{B}_k^T \end{bmatrix} \begin{bmatrix} W_k \end{bmatrix} \begin{bmatrix} \epsilon_k \end{bmatrix} \qquad (19)$$

where the dimensions of the matrices have been given previously for all m stations and all n aircraft positions on the k th pass. We represent equation (19) in the normal equations system on the k th pass for our artillery problem as follows:

$$\left[\begin{array}{c|cc} \hat{U}_k & \dot{U}_k & \bar{U}_k \\ \hline \dot{U}_k^T & \dot{N}_k & \bar{N}_k \\ \bar{U}_k^T & \bar{N}_k^T & \ddot{N}_k \end{array}\right] \begin{bmatrix} \hat{\delta} \\ \dot{\delta}_k \\ \ddot{\delta}_k \end{bmatrix} = \begin{bmatrix} \hat{c} \\ \dot{c}_k \\ \ddot{c}_k \end{bmatrix} . \qquad (20)$$

Using the rule mentioned previously for augmenting the normal equations for additional passes of the aircraft, we have for an additional (k + 1) st pass

$$\left[\begin{array}{c|cc|cc} \hat{U}_k + \hat{U}_{k+1} & \dot{U}_k & \bar{U}_k & \dot{U}_{k+1} & \bar{U}_{k+1} \\ \hline \dot{U}_k^T & \dot{N}_k & \bar{N}_k & 0 & 0 \\ \bar{U}_k^T & \bar{N}_k^T & \ddot{N}_k & 0 & 0 \\ \hline \dot{U}_{k+1}^T & 0 & 0 & \dot{N}_{k+1} & \bar{N}_{k+1} \\ \bar{U}_{k+1}^T & 0 & 0 & \bar{N}_{k+1}^T & \ddot{N}_{k+1} \end{array}\right] \begin{bmatrix} \hat{\delta} \\ \dot{\delta}_k \\ \ddot{\delta}_k \\ \dot{\delta}_k \\ \ddot{\delta}_{k+1} \end{bmatrix} = \begin{bmatrix} \hat{c}_k + \hat{c}_{k+1} \\ \dot{c}_k \\ \ddot{c}_k \\ \dot{c}_{k+1} \\ \ddot{c}_{k+1} \end{bmatrix} , \qquad (21)$$

until finally we accumulate all the components of the normal equations for all 10 passes so that we have a coefficient matrix and constant column vector of order (4890 x 4890) and (4890 x 1),

respectively. For later reference, we will write matrix equation (21) in the following condensed notation:

$$\left[\begin{array}{c|c|c} \hat{U}_k + \hat{U}_{k+1} & \hat{\ddot{U}}_k & \hat{\ddot{U}}_{k+1} \\ \hline \hat{\ddot{U}}_k^T & \ddot{U}_k & 0 \\ \hline \hat{\ddot{U}}_{k+1}^T & 0 & \ddot{U}_{k+1} \end{array}\right] \left[\begin{array}{c} \hat{\delta} \\ \hline \ddot{\delta}_k \\ \hline \ddot{\delta}_{k+1} \end{array}\right] = \left[\begin{array}{c} \hat{c}_k + \hat{c}_{k+1} \\ \hline \ddot{c}_k \\ \hline \ddot{c}_{k+1} \end{array}\right] \qquad (22)$$

where we can equate each of these new partitioned matrix components to the corresponding partitioned components in equation (21) according to the definition of matrix equality. The importance of writing equation (21) in the form of equation (22) is to highlight the fact that we have a Second Order Regression Scheme embedded in a First Order Regression Scheme. If we had restricted our initial hypotheses and stated that the error model parameters were stable, i.e., do not change from pass to pass, these parameters could be put into the outer loop, so to speak, along with the ground tracker coordinates. This results in a First Order Regression Scheme for which solution algorithms were developed by D.C. Brown about 10 years ago.

A finer structure of the coefficient matrix of the normal equations resulting from our 30 station artillery problem is shown in Figure 1.

ADVANCED SOLUTION ALGORITHMS: The examination of Figure 1 indicates that the only portion of the normal equation system which does not grow indefinitely is $\hat{U}$, which involves the ground tracker coordinates. These are common to all passes of the aircraft. We want to exploit this situation in finding the inverse of this very large coefficient matrix. However, we encounter a serious problem when we attempt to treat our problem as a First Order Regression Scheme, as implied by equation (22) where the $\ddot{U}$ matrices are very large, block diagonal matrices. In a First Order Regression Scheme, these block diagonal matrices are usually very small. Normally they are 3 x 3 matrices for aircraft position information in geometric data reduction applications. We will point out this difficulty as applied to our artillery problem. If we rewrite equation (22) for just the k th pass, we have

$$\left[\begin{array}{c|c} \hat{U} & \hat{\ddot{U}}_k \\ \hline \hat{\ddot{U}}_k^T & \ddot{U}_k \end{array}\right] \left[\begin{array}{c} \hat{\delta} \\ \hline \delta_k \end{array}\right] = \left[\begin{array}{c} \hat{c} \\ \hline c_k \end{array}\right] \qquad (23)$$

where the solution for the correction vector is given by

$$\begin{bmatrix} \hat{\delta} \\ \hline \delta \end{bmatrix} = \begin{bmatrix} \hat{U} & | & \tilde{U} \\ \hline \tilde{U}_k^T & | & \ddot{U}_k \end{bmatrix}^{-1} \begin{bmatrix} \hat{c} \\ \hline c \end{bmatrix} . \tag{24}$$

It is well known that when we multiply the coefficient matrix by its inverse we obtain the identity matrix. Therefore, we have

$$\begin{bmatrix} \hat{U} & | & \tilde{U}_k \\ \hline \tilde{U}_k & | & \ddot{U}_k \end{bmatrix} \begin{bmatrix} \hat{M} & | & \tilde{M}_k \\ \hline \tilde{M}_k & | & \ddot{M}_k \end{bmatrix} = \begin{bmatrix} \hat{I} & | & 0 \\ \hline 0 & | & \ddot{I}_k \end{bmatrix} \tag{25}$$

from which we obtain four matrix equations in the unknowns $\hat{M}$, $\tilde{M}_k$, $\tilde{M}_k^T$, and $\ddot{M}_k$. Without going through the derivation, we give the result for $\hat{\delta}$ and δ and point out the aforementioned difficulty. We have

$$\hat{\delta} = \hat{M}_k \, (\hat{c} - \tilde{U}_k \, \ddot{U}_k^{-1} \, c_k) \tag{26}$$

where

$$\hat{M}_k = (\hat{U} - \tilde{U}_k \, \ddot{U}_k^{-1} \, \tilde{U}_k^T)^{-1} , \tag{27}$$

$$\tilde{M}_k^T = -\ddot{U}_k^{-1} \, \tilde{U}^T \, \hat{M}_k , \tag{28}$$

$$\ddot{M}_k = \ddot{U}_k^{-1} + \ddot{U}_k^{-1} \, \tilde{U}_k^T \, \hat{M}_k \, \tilde{U}_k \, \ddot{U}_k^{-1} . \tag{29}$$

Therefore,

$$\hat{\delta} = (\hat{U} - \tilde{U}_k \, \ddot{U}_k^{-1} \, \tilde{U}_k^T)^{-1} \, (\hat{c} - \tilde{U}_k \, \ddot{U}_k^{-1} \, c_k), \tag{30}$$

and

$$\delta = \ddot{U}_k^{-1} \, c - \ddot{U}_k^{-1} \, \tilde{U}^T \, \hat{\delta} . \tag{31}$$

We see that the major work in solving for the correction vector involves the inverse of $\ddot{U}$ which, for First Order Schemes, is usually simply the inversion of "n" matrices of order 3 for three dimensional, geometric applications. However, when we consider our Second Order Regression Scheme, we see that $\ddot{U}$ for the k th pass is of order 480 x 480: that is,

$$\underset{(480,480)}{\ddot{U}_k} = \begin{bmatrix} \underset{(120,120)}{\dot{N}_k} & \underset{(120,360)}{\bar{N}_k} \\ \underset{(360,120)}{\bar{N}_k^T} & \underset{(360,360)}{\ddot{N}_k} \end{bmatrix}, \quad \delta = \underset{(480,1)}{\begin{bmatrix} \underset{(120,1)}{\dot{\delta}_k} \\ \underset{(360,1)}{\ddot{\delta}_k} \end{bmatrix}}, \text{ and } c = \underset{(480,1)}{\begin{bmatrix} \underset{(120,1)}{\dot{c}_k} \\ \underset{(360,1)}{\ddot{c}_k} \end{bmatrix}} . \tag{32}$$

Even though $\ddot{U}$ may be sparse to begin with, its inverse is not necessarily sparse. It becomes obvious, therefore, that the inversion and the carrying along of $\ddot{U}$ through the computations is highly impractical, especially when many passes of the aircraft have to be carried for a rigorous, simultaneous solution for thousands of parameters. It is possible to apply the first order algorithms a second time to solve for the inverse of $\ddot{U}$ by partitioning, thereby dropping to a lower order set of matrices. We begin by performing the same operation on the $\ddot{U}$ portion of equation (23) as was done to the whole matrix to find its inverse. Therefore,

$$\ddot{U}_k \, M_k = \begin{bmatrix} \dot{N}_k & \bar{N}_k \\ \bar{N}_k^T & \ddot{N}_k \end{bmatrix} \begin{bmatrix} \dot{M}_k & \bar{M}_k \\ \bar{M}_k^T & \ddot{M}_k \end{bmatrix} = \begin{bmatrix} \dot{I}_k & 0 \\ 0 & \ddot{I}_k \end{bmatrix}, \tag{33}$$

and solving for the partitioned inverse of $\ddot{U}_k$, we have

$$\dot{M}_k = (\dot{N}_k - \bar{N}_k \, \ddot{N}_k^{-1} \, \bar{N}_k^T)^{-1} , \tag{34}$$

$$\bar{M}_k^T = -\ddot{N}_k^{-1} \, \bar{N}_k^T \, \dot{M}_k , \tag{35}$$

$$\bar{M}_k = -\dot{M}_k \, \bar{N}_k \, \ddot{N}_k^{-1} , \tag{36}$$

$$\ddot{M}_k = \ddot{N}_k^{-1} + \ddot{N}_k^{-1} \, \bar{N}_k^T \, \dot{M}_k \, \bar{N}_k \, \ddot{N}_k^{-1} . \tag{37}$$

The solution vectors are given by

$$\dot{\delta}_k = \dot{M}_k \left(\dot{c}_k - \bar{N}_k \ddot{N}_k^{-1} \ddot{c}_k\right), \tag{38}$$

and

$$\ddot{\delta}_k = \ddot{N}_k^{-1} \dot{c}_k - \ddot{N}_k^{-1} \bar{N}_k^T \dot{\delta}_k \ . \tag{39}$$

If we inspect the dimension of these matrices, we find that they are of much lower order than Ü as a whole. For example, information for all 120 aircraft positions for the k th pass is given by the block diagonal matrix $\ddot{N}_k$ of order 360, but its inverse is simply the inversion of 120, 3 x 3 matrices. The matrix $(\dot{N}_k - \bar{N}_k \ddot{N}^{-1} \bar{N}_k^T)$ requiring inversion is a symmetric matrix of order 120 for our artillery problem. However, we usually carry less than 4 error model parameters per tracker so this matrix may actually be much smaller. In any event, we begin to realize that the matrix components, after applying the inversion process a second time, are much smaller than the original Ü, which is of order 480. It remains to actually substitute the inverse of Ü into equations (30 and 31); that is

$$\hat{\delta} = \left[\hat{U} - \begin{bmatrix} \dot{U}_k & \bar{U}_k \end{bmatrix} \begin{bmatrix} \dot{M}_k & \bar{M}_k \\ \bar{M}_k^T & \ddot{M}_k \end{bmatrix} \begin{bmatrix} \dot{U}^T \\ \bar{U}_k^T \end{bmatrix}\right]^{-1} \left[\hat{c} - \begin{bmatrix} \dot{U}_k & \bar{U}_k \end{bmatrix} \begin{bmatrix} \dot{M}_k & \bar{M}_k \\ \bar{M}_k^T & \ddot{M}_k \end{bmatrix} \begin{bmatrix} \dot{c}_k \\ \ddot{c}_k \end{bmatrix}\right], \tag{40}$$

and

$$\delta = \begin{bmatrix} \dot{\delta}_k \\ \ddot{\delta}_k \end{bmatrix} = \begin{bmatrix} \dot{M}_k & \bar{M}_k \\ \bar{M}_k^T & \ddot{M}_k \end{bmatrix} \begin{bmatrix} \dot{c}_k \\ \ddot{c}_k \end{bmatrix} - \begin{bmatrix} \dot{M}_k & \bar{M}_k \\ \bar{M}_k^T & \ddot{M}_k \end{bmatrix} \begin{bmatrix} \dot{U}_k^T \\ \bar{U}_k^T \end{bmatrix} \hat{\delta}. \tag{41}$$

In reference 2, (Brown, 1969), D. Brown points out efficient and inefficient ways of performing the matrix multiplications resulting from the further substitutions to be made in equations (40) and (41). Space does not permit all the matrix algebra in this paper. However,

equations (40) and (41) essentially constitute the key to the solution of Second Order Regression Systems. The matrix equation formulae for the solution vectors $\hat{\delta}$, $\dot{\delta}$ and $\ddot{\delta}$ for the k th pass after all the matrix algebra is completed are given in expanded form as follows:

$$\underset{(90,1)}{\hat{\delta}} = \left[\hat{U} - \bar{U}_k \ddot{N}_k^{-1} \bar{U}_k^T - \left[(\dot{U}_k - \bar{U}_k \ddot{N}_k^{-1} \bar{N}_k^T)(\dot{N}_k - \bar{N}_k \ddot{N}_k^{-1} \bar{N}_k^T)^{-1}\right.\right.$$

$$\left.\left.(\dot{U}_k^T - \bar{N}_k \ddot{N}_k^{-1} \bar{U}_k^T)\right]\right]^{-1}_{(90,90)} \times \left[\hat{c} - \bar{U}_k \ddot{N}_k^{-1} \ddot{c}_k - \left[(\dot{U}_k - \bar{U}_k\right.\right.$$

$$\left.\left.\ddot{N}^{-1} \bar{N}^T)(\dot{N}_k - \bar{N}_k \ddot{N}^{-1} \bar{N}^T)^{-1}(\dot{c}_k - \bar{N}_k \ddot{N}^{-1} \ddot{c}_k)\right]\right]_{(90,1}, \quad (42)$$

$$\underset{(120,1)}{\dot{\delta}_k} = (\dot{N}_k - \bar{N}_k \ddot{N}_k^{-1} \bar{N}_k^T)^{-1}(\dot{c}_k - \bar{N}_k \ddot{N}_k^{-1} \ddot{c}_k) - (\dot{N}_k - \bar{N}_k \ddot{N}_k^{-1} \bar{N}_k^T)^{-1}$$
$$(\dot{U}_k^T - \bar{N}_k \ddot{N}_k^{-1} \bar{U}_k^T)\hat{\delta}, \quad (43)$$

and

$$\underset{(360,1)}{\ddot{\delta}_k} = \ddot{N}_k^{-1} \ddot{c}_k - \ddot{N}_k^{-1} \bar{N}_k^T \dot{\delta}_k. \quad (44)$$

We notice that the solution vector for tracker coordinates, $\hat{\delta}$, never grows to anything larger than a (90,1) no matter how many passes are reduced simultaneously. This is the reason for including these parameters into the outer loop, so to speak, of the normal equation system, and the inversion of $\ddot{N}_k$ is simply the inverse of 3 x 3 matrices in the solution for $\ddot{\delta}_k$. The interested reader can correlate the various matrix dimensions in equations (42), (43), and (44) with Figure 1, which actually shows a finer matrix structure than these equations. This suggests that these equations can be broken down to even smaller matrix dimensions, and the summation of matrices can be efficiently incorporated into these equations.

CONCLUSION: We have shown how newly derived algorithms can be used to solve a large scale, Second Order Regression Scheme

whereby the simultaneous solution of 4890 parameters becomes practical. The use of First Order algorithms leaves much to be desired for we have seen that this entails carrying matrices of order 480 through the computations for each pass of the aircraft. The solution normally involves two or three iterations for convergence, thereby adding to the computational burden. However, we have shown that the application of new, advanced, algorithms reduces the largest matrix to be handled in the computations to order 120 while solving for all 4890 parameters simultaneously. This results in a significant saving in computation since the inversion of matrices of order 120 requires much less computer time and computer memory space than matrices of order 480. What may have been initially impractical to solve, now becomes tenable.

It is conceivable that LRPDS could serve as the Army's electronic surveying system in addition to the artillery survey application where aircraft flight paths are geometrically restricted. In a full blown, surveying operation, the flight path geometry can be chosen to be optimum so that all trackers are exercised in a wide excursion of geometry. This enhances the recovery of systematic errors and opens the possibility that LRPDS accuracies can approach its resolution; that is, less than two feet. The new algorithms would greatly enhance this type of operation also.

REFERENCES

[1] KIISK, A., Description and Analysis of the Range Change Method of Determining Position, Surveying and Geodesy Division, US Army Engineer Topographic Laboratories, Fort Belvoir, Virginia, 30 July 1969.

[2] BROWN, D.C., TROTTER, J.E., SAGA, A Computer Program for Short Arc Geodetic Adjustment of Satellite Observations, DBA Systems, Inc., P.O. Drawer 550, Melbourne, Florida 32901.

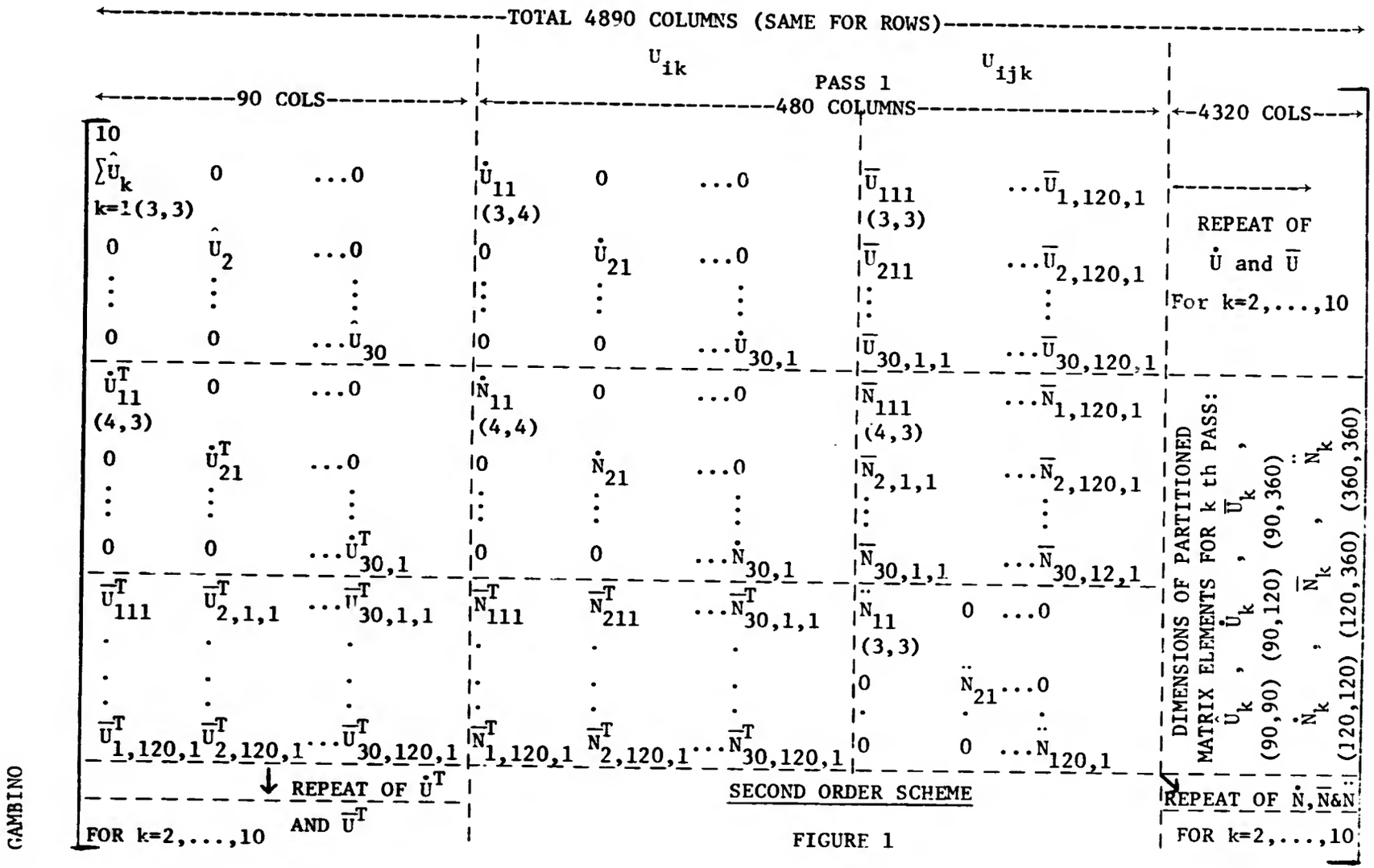


FIGURE 1

CAMBINO

AD 713 526

DENVER EARTHQUAKES

LOUIS E. GARONO
EDGEWOOD ARSENAL
EDGEWOOD ARSENAL, MARYLAND

You may wonder how the Army -- and particularly Edgewood Arsenal -- have gotten into a study of earthquakes, so I would like to give you a little background before I talk about the problem of the Denver earthquakes. In 1942 the Army established Rocky Mountain Arsenal on about 20,000 acres of dry farm land about 15 miles from downtown Denver and five miles from Stapleton Airport. The arsenal was built to manufacture a variety of chemical munitions. As with all chemical manufacture, there are waste products that must be disposed of safely.

Standard evaporation ponds were tried but these were not successful due to seepage into the soil and contamination of the eighty foot aquifer. In 1957 a blown asphalt sealed lake covering 86 acres and holding about 240,000,000 gallons was constructed but the rate of evaporation was not rapid enough to keep up with liquid waste production. Also, the fact that a large reservoir of chemical waste was stored above ground was not considered entirely safe.

A thorough study by a very distinguished industrial panel reviewed all known methods of waste disposal and recommended that the Army use deep well disposal for a permanent, safe and almost limitless reservoir for liquid wastes. The well was drilled and all the safety precautions necessary to protect the fresh water and oil producing strata were installed as shown in Figure 1. The well was triple cased and cemented to protect the fresh water to 2100 feet and double cased and cemented to protect the oil strata. It was decided to go to granite, or the bottom of the precambrian rocks of the sedimentary sequence, with the well to get the waste as far down as possible and under 6500 feet of impermeable shale. The reservoir at this level of 11,985 to 12,045 feet, or about 6800 feet below sea level, was at a low pressure 4100 lbs. per square inch instead of 5700 lbs. per square inch that would have been expected considering the overburden with a 10% water absorption capability,

mostly due to fractures (3). This made for good acceptance of waste and represented a very acceptable waste reservoir good for millions of gallons of liquid chemical waste.

During the Fall of 1961, the well went through its final testing and finishing operations including the injection of 800,000 gallons of fresh water to form a barrier to prevent precipitation as a safety precaution. By February 1, 1962, the deep well was ready to receive waste. Operations were immediately started as the sealed lake had reached capacity and there was an urgency to bring the lake down to a safe level. There was another activity in the Denver area at the same time that the disposal well was being drilled that would have a great influence on the well operations at a future date but the Army did not realize it at the time. Two new sensitive seismographs, much more sensitive than any used in the Denver area before, were being installed in the Fall of 1961 - one at Regis College in Denver and the other at the Colorado School of Mines Campus. Some Denverites became more tremor conscious after these new sensitive seismographs came into operation during December 1961 and January 1962 when the micro tremors were reported in the Denver press.

Waste was injected into the well as a rather routine operation during 1962, 1963, 1964 and 1965 at the rate of from 100-300 gallons per minute (gpm) and on one, two, or three shifts with the deep well being used as needed and when well operators were available. Over a period of four years approximately 165,000,000 gallons of chemical waste was injected. Rocky Mountain Arsenal's waste disposal problems had apparently been successfully solved. However, it was noted by the well operators that there were more micro quakes (a Richter reading of three or less) when waste was being injected than when the well was shut down. This increase in tremors was noted in the Denver papers as data from the new seismographs were reported. This did not cause concern at Rocky Mountain Arsenal as Denver had experienced earthquakes since the area had been settled (9).

However, the earthquake question was abruptly brought to everyone's attention in November 1965 by an initial release by a newspaper article and TV show giving the data shown on Figure 2 (4). This figure indicates the wastes injected into the well in millions of gallons per month versus the number of earth tremors per month recorded at the Colorado School of Mines. Mr. Evans, a Denver Geologist, pointed out the apparent correlation of millions of gallons of waste pumped into the well on a monthly basis and the number of earthquakes per month recorded at the Colorado School of Mines as the result of research being done using the new sensitive seismographs. The correlation between waste pumped per month and the number of minor earthquakes with a maximum intensity of 4.5 on the Richter scale appeared so valid that sufficient political

pressure was applied. As a result, the well was shut down in February 1966, and the matter was considered so important that it was called to President Johnson's attention by the Governor of Colorado. The well has not operated since February 1966.

The idea of being able to cause earthquakes of a minor magnitude under controlled injection appealed to a great many people and organizations, including the Army Corps of Engineers, the US Geological Survey, and most of the geologists in the Denver area. The ability to relieve the natural stresses in the rock without catastrophic earthquakes offers a great reward in preventing suffering and the destruction of property. As a result a great many studies have been completed to determine how valid the apparent correlation really is:

a. Studies by Dr. Maurice Major of the Colorado School of Mines show that Colorado is a seismic area and has been subject to earthquakes since the area was settled and is considered to be subject to tremors (5), (9).

b. Studies by the US Geological Survey, by Dr. J.H. Healy and Associates, indicate that the well was responsible for the earthquakes and that the removal of water from the deep well would relieve the rock pressures which could have been fracturing the rock structures and possibly causing the release of rock strain. However, the pumping out of 300,000-plus gallons did not change the pressure in the rock and pumping was stopped and the level of the well has returned to its original level before injection (6).

c. Earth tremors did not stop when pumping was stopped in February 1966 but have continued to date, with the largest and highest intensity tremors happening about 16 months after waste injection into the well was stopped. Figure 3 shows the number of tremors per month compared with the pumping of waste per month in millions of gallons from February 1962 to January 1, 1970. You will note that there have been several areas of high seismic activity during periods of no pumping. The waste level in the well was also recorded on a monthly basis since the well was shut down as indicated by this figure, with waste gradually injecting into the rock under the force of gravity starting at the rate of about a foot or two gallons of waste a day starting in March 1966 and decreasing to zero in 1969. The theory of the US Geological Survey was tested by removing waste from the well. After 300,000 gallons of liquid was removed, the head only dropped 100 feet or 50 lbs., thus demonstrating that the rock pressure could not be reduced by a small removal of liquid as predicted. Liquid flowed back into the well casing from the strata as fast as it was removed. You will also note there was no great falloff of water level in the well when the three 5+ earthquakes occurred during the period of no waste injection.

It should always be remembered that earth tremors are the release of built up rock strains, built up over many years, that are released or triggered and were not caused by pumping waste. It seems appropriate to point out at this point that there are more than 100 waste disposal wells in operation in this Country in many different strata and locations. In no case has injection of waste into any of these wells caused tremors that have been recorded or noticed. Further, the Army has a disposal well at its Newport, Indiana plant where since 1960 a similar waste has been injected into a fractured strata at 6300 feet below sea level with no noticed tremors resulting. Assuming that there are other forces that could possibly trigger tremors - such as sonic booms, testing of Titan missiles at the Martin-Marietta Denver plant, planes landing at Stapleton Airport, and build-up of surface weight - all of these were plotted with only surface weight or rainfall indicating correlation. The effect of weight was investigated as tremors have been recorded at a number of other locations during the filling of large reservoirs and dams. Assuming a six-inch rainfall over the active area, the result is a weight of 32.2 lbs. per square foot. With several months of high rainfall totalling to more than 12 inches of water, this may be significant. Therefore, rainfall per month in inches was plotted against tremors with the results shown in Figure 4.

You will note there is correlation between rainfall per month and tremors both during the injection period and after February 1966 when injection of waste was stopped. High rainfall occurred in 1965 and 1967 with three instances of high tremor activity in 1967. However, there was also a period of medium rainfall in 1969, not as high as previously bunched high rainfall months, without a corresponding tremor activity. To me, this indicates that tremors may be triggered by periods of high rainfall or some other natural phenomena connected with heavy rainfall.

A study made for the Corps of Engineers by a distinguished panel of earth scientists in 1966 concluded (1):

"1. Although the investigations which have been made have revealed some very suggestive correlative statistics, no scientific proof of a relationship between the operation of the deep waste disposal well at the Rocky Mountain Arsenal and the Derby earthquakes has been established."

"2. None of the earthquakes recorded during the period of injection operations have been of a magnitude that would have prompted more than passing note had they occurred in an area of previously recognized seismicity."

"3. Use of the well could continue without fear of initiating an earthquake of higher magnitude than already experienced. The injection of fluid may minimize earthquake potential, but is unlikely to accentuate it."

"4. The deep well offers great research potential for the advancement of the knowledge of seismology and for possible use as an instrument in the study of earthquake mechanisms. The risk involved in continued operation of the well is believed to be negligible. Operation of the well as a test facility in the National interest is considered desirable."

"5. The possibility exists that future testing might demonstrate that injection in the pre-cambrian or the higher fountain or lyon formations could be accomplished without rick."

"6. Based on initial evaluation of reservoir properties and study of performance of the well, it appeared that the reservoir was normal and would predictably react to fluid injection as a reservoir consisting of fractured rock."

However, these findings have not been accepted by the Corps of Engineers as a basis for action.

The US Geological Survey office of Earthquake Research and Crustal Studies in Menlo Park, California, on the other hand on March 3, 1968, warns that "An earthquake of major magnitude -- as high as 6.2 on the Richter Scale -- is possible in the Denver area." (2)

Figure 5 gives a quick summary of the data on the Denver earthquakes compared with waste injection in millions of gallons per month and against rainfall in inches per month. Denver apparently will have tremors for the foreseeable future with most tremors not being felt and could never have been recorded prior to 1962 when sensitive seismographs were installed in the Denver area.

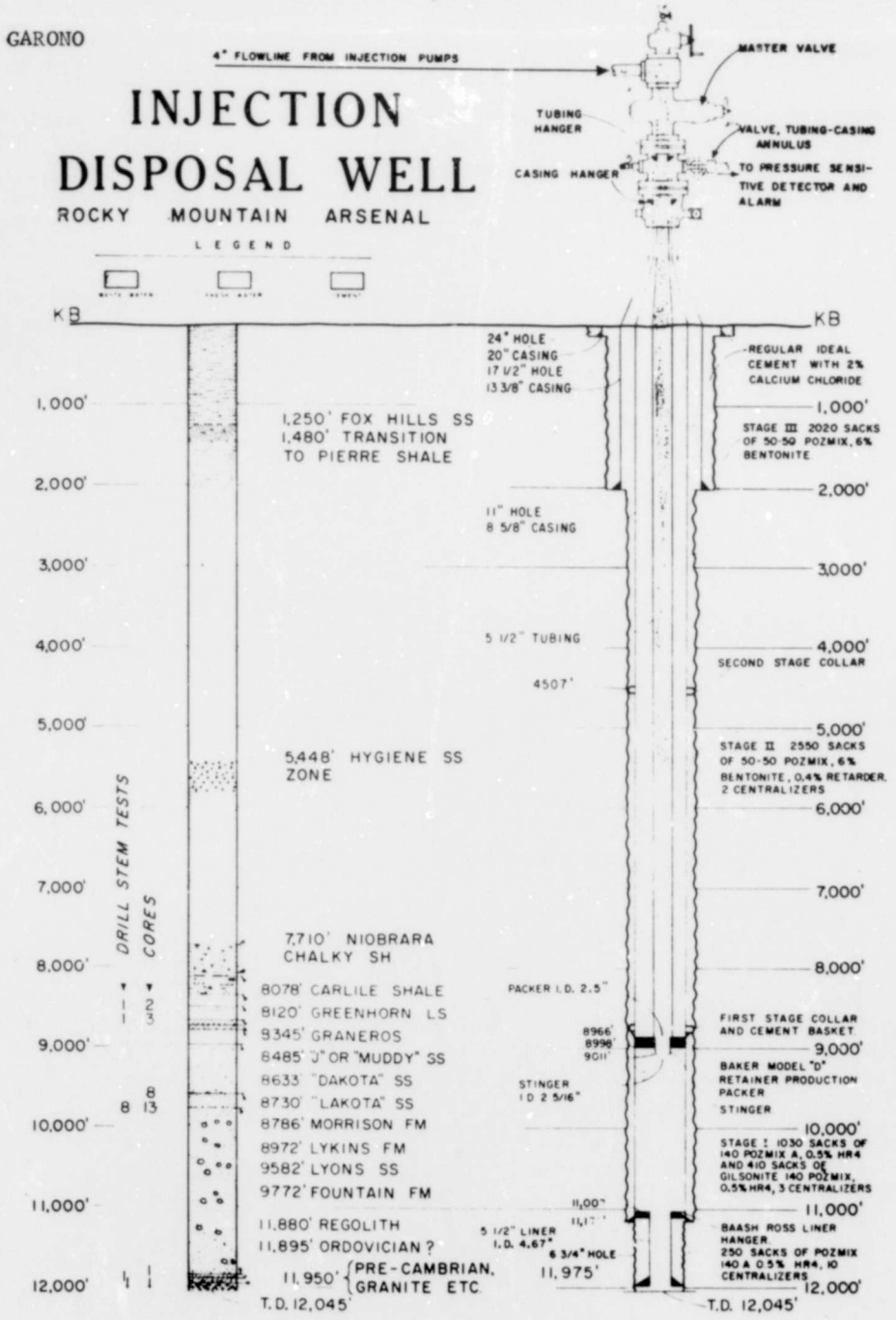
In the case of waste injection, correlation was very good during 1963 and 1965, however, there were three periods of very high earthquake activity with tremors in the range of 5+ in May and August 1967 with no injection being done and no waste injected for 14 months. Until this time the greatest tremor intensity amounted to 4.5 on the Richter scale in 1965. It should also be noted that on November 27, 1967 there were three tremors of the intensity of 5.1, 4.3 and 3.5 in a 10 minute period. In the case of inches of rainfall per month versus tremors, the correlation is very good for 1965 and 1967 where there were three months of heavy rainfall followed by two periods of high earthquake activity and high intensity - in both instances there were more than 12 inches of rainfall in a three month period. It should be remembered that the

normal rainfall in the Denver area is 11 to 13 inches per year. In 1969 there was a short period of medium rainfall without corresponding earthquake activity. In 1963 there was high tremor activity with low intensity and normal rainfall. Time is running out and it is very difficult to put together several months of high rainfall in the normally dry Denver area to test the premise of surface weight or other natural phenomena connected with rainfall as the controlling factor in causing tremors. I am not trying to prove that heavy rainfall is the cause of earth tremors but pointing out the fact that other phenomena than the injection of waste may be the responsible cause. I believe we have reached the time after 52 months of no waste injection when additional injection is desirable to physically check the concept that waste injection causes earth tremors.

A computer statistical analysis shows some correlation for both injection of waste into the deep well as well as with instances of heavy rainfall for several successive months causing or triggering earthquakes as shown on Figures 6 and 7. These analyses indicate the correlation of waste and rainfall as the cause of earthquakes are of the same order with rainfall showing slightly less correlation than waste injection on a monthly basis. Figure 7 shows the same correlation using another method. However, it must be remembered that since February 1966 there has been no waste injected and therefore for the last 52 months there has been no correlation between waste injected and earthquakes. However, in the two instances when there were three months of heavy rainfall in a row which amounts to more than 12 inches there is an excellent correlation.

Rocky Mountain Arsenal will shortly have about 12,000,000-15,000,000 gallons of waste (this amounts to less than 10% of the 165,000,000 gallons that has previously been injected) to be disposed of from the distruction of our chemical stockpile. The most economical means of disposal for this liquid waste is the deep well. The real question which must be decided is: "Can the Army use the injection well, with the possibility of saving millions of dollars in equipment, materials, and labor while facing the possibility of initiating additional tremors and possibly a large quake of the magnitude predicted by the US Geological Survey?" If the tremors increase, the fact that earth strains can be released by water injection in this particular formation will be established. Use can then be made of this system of earthquake prevention at other locations. In any case, man's knowledge is increased with a saving in dollars with little or no risk. In any case, if there is a major tremor, the Army will be blamed for the quake so why not save a large amount of money which could be as much as eight million dollars and use the well. Therefore, I along with the distinguished panel of experts assembled by the Corps of Engineers in 1966 (1) to

examine this subject, recommend that the deep well at Rocky Mountain Arsenal be put back in service for waste disposal as soon as required with the necessary recording and controlling devices to determine the true correlation of tremors with waste injection.



E A POLUMBUS, JR AND ASSOCIATES, INC.

FIGURE 1

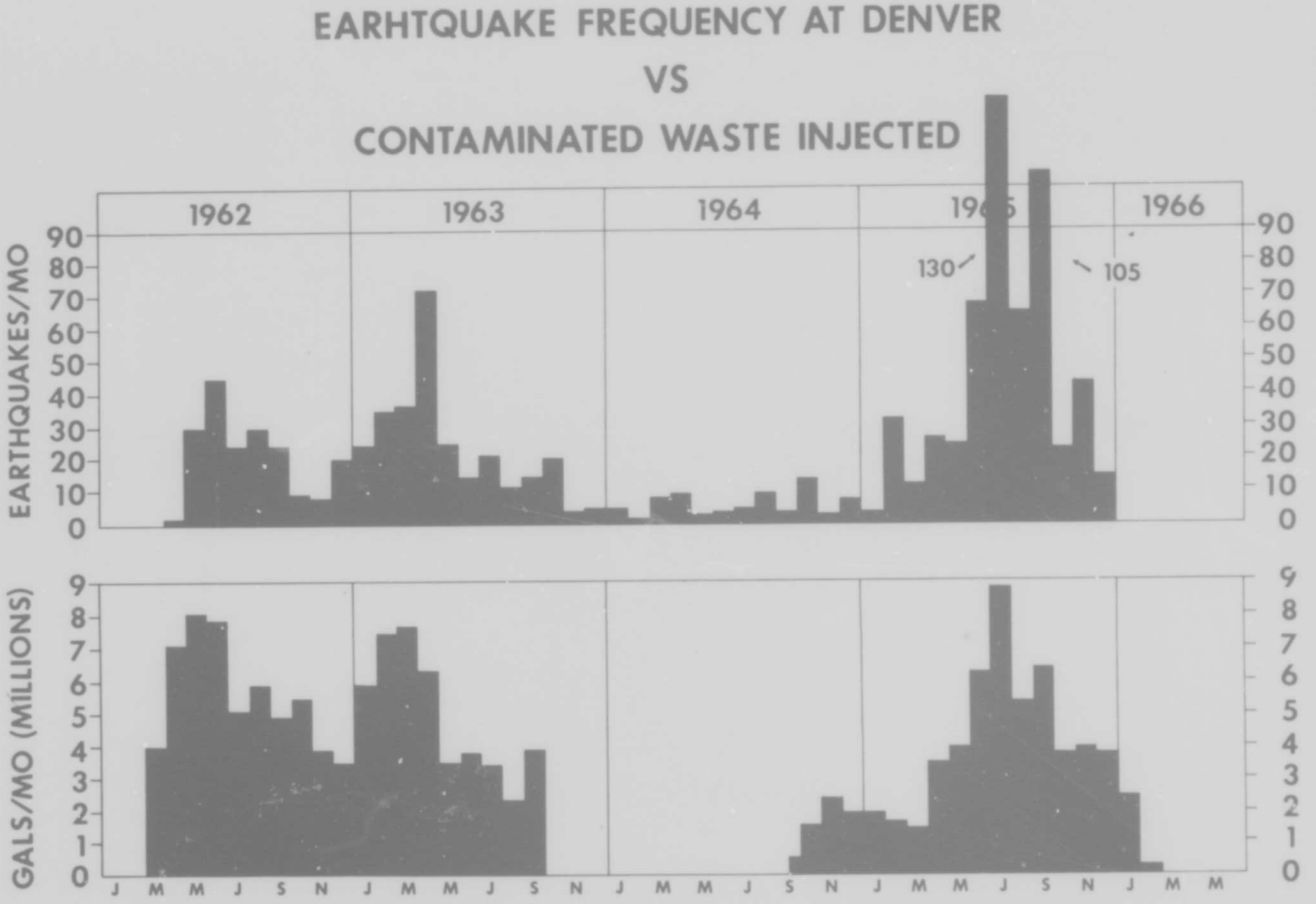

EARHTQUAKE FREQUENCY AT DENVER
VS
CONTAMINATED WASTE INJECTED
1962
1963
1964
1965
1966
EARTHQUAKES/MO
GALS/MO (MILLIONS)
130
105
J M M J S N J M M J S N J M M J S N J M M J S N J M M

FIGURE 2

EARTHQUAKE FREQUENCY AT DENVER
VS
CONTAMINATED WASTE INJECTED

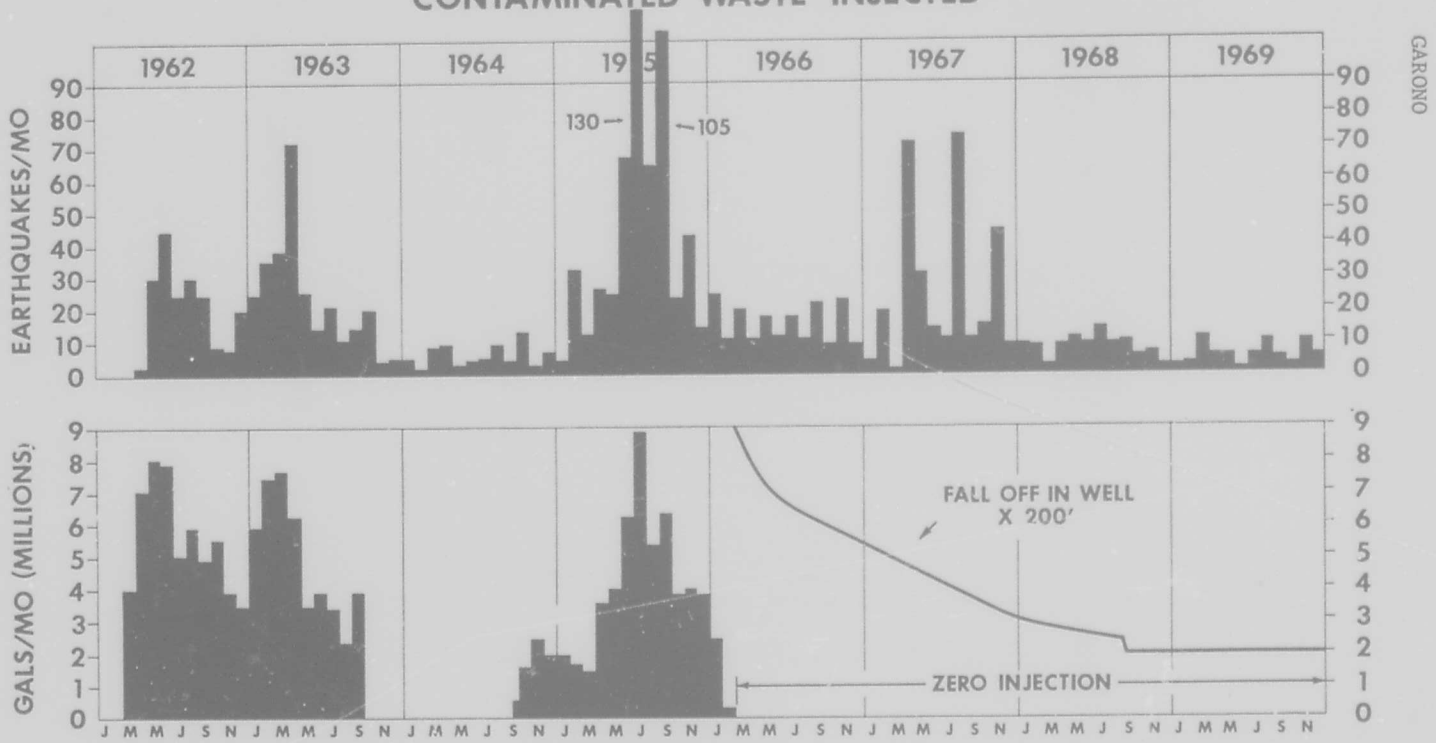


FIGURE 3

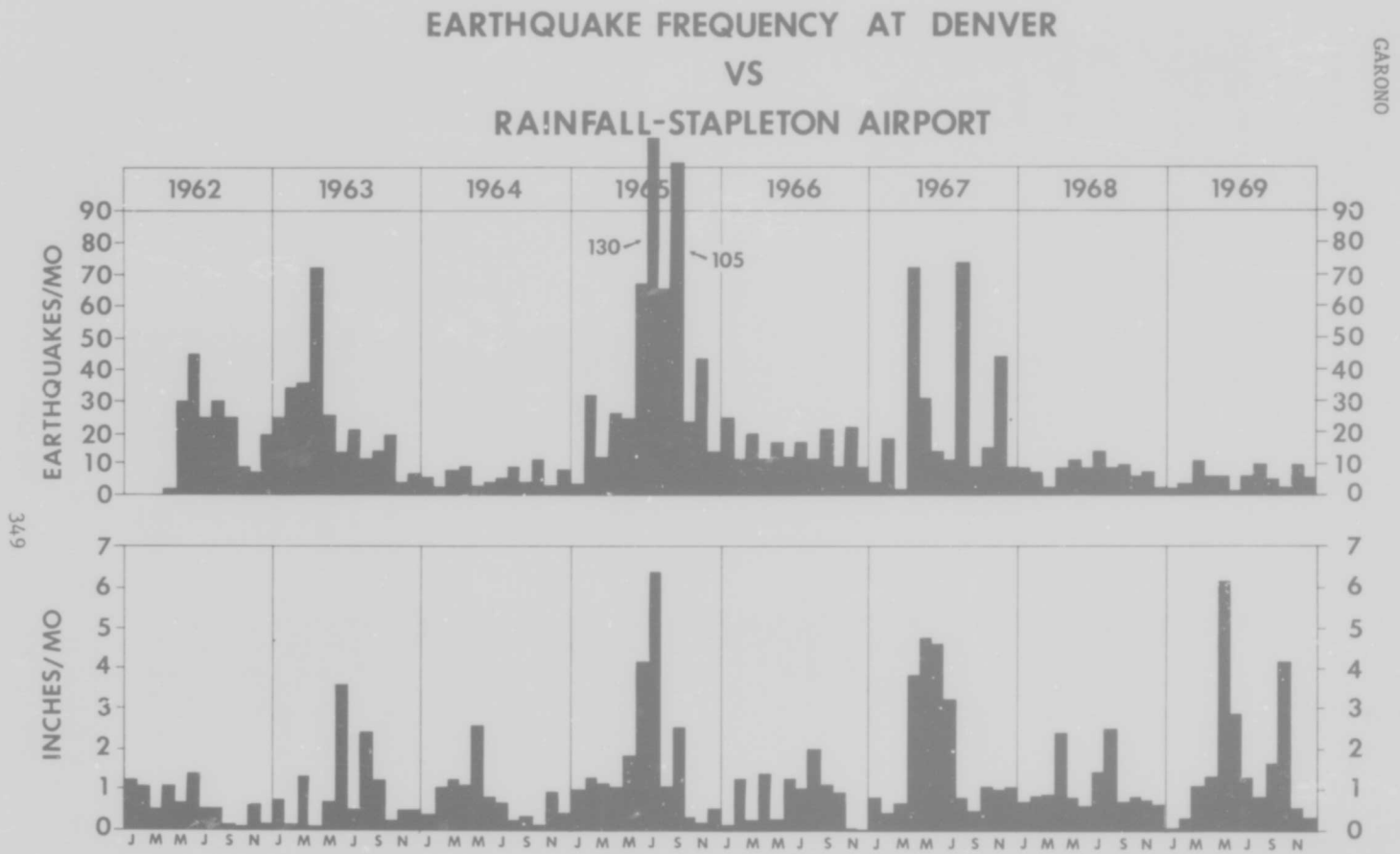

EARTHQUAKE FREQUENCY AT DENVER
VS
RAINFALL-STAPLETON AIRPORT
1962
1963
1964
1965
1966
1967
1968
1969
EARTHQUAKES/MO
130
105
INCHES/MO
J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N

FIGURE 4

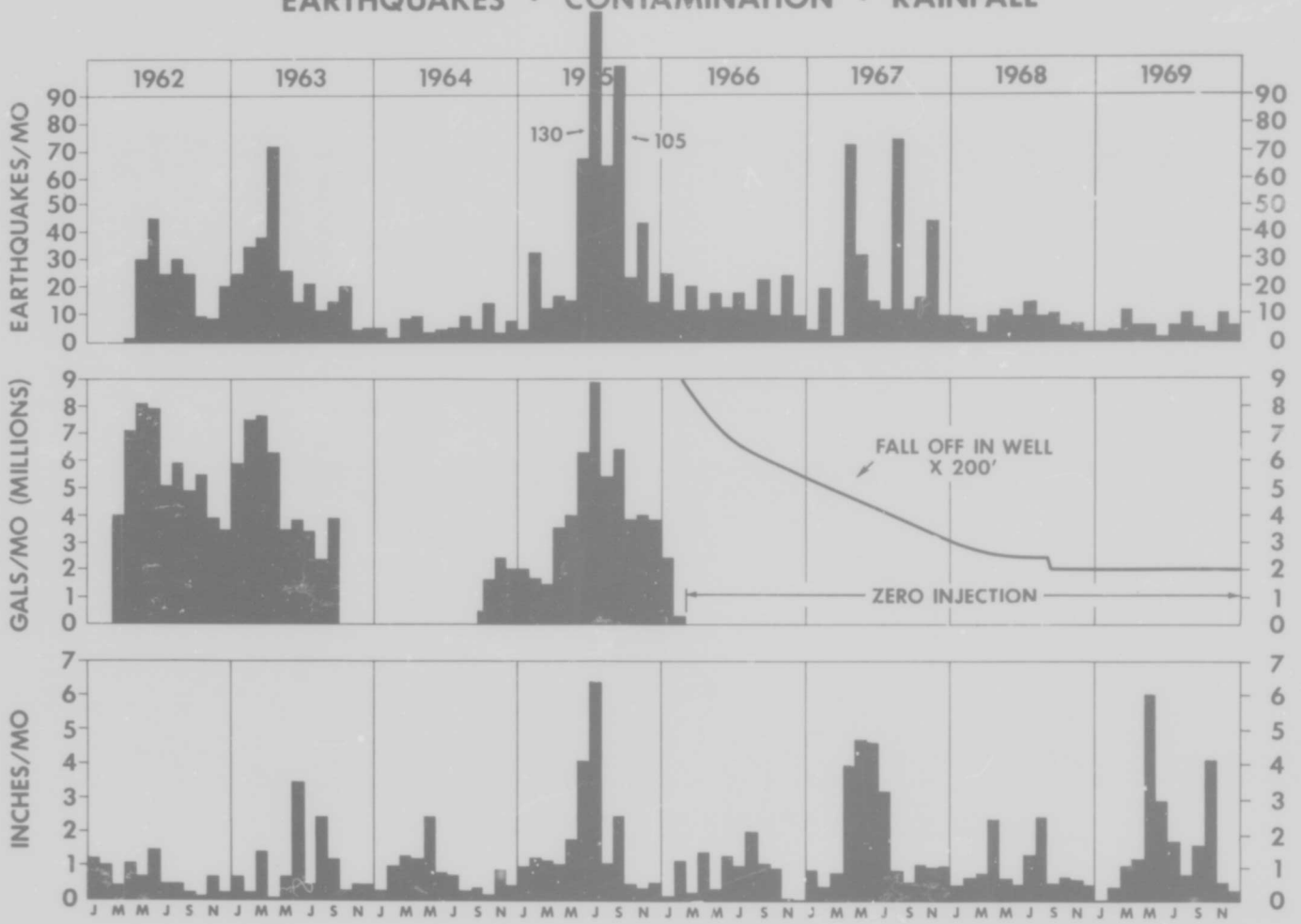

EARTHQUAKES · CONTAMINATION · RAINFALL
1962
1963
1964
1965
1966
1967
1968
1969
EARTHQUAKES/MO
130
105
GALS/MO (MILLIONS)
FALL OFF IN WELL
X 200'
ZERO INJECTION
INCHES/MO
J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N

FIGURE 5

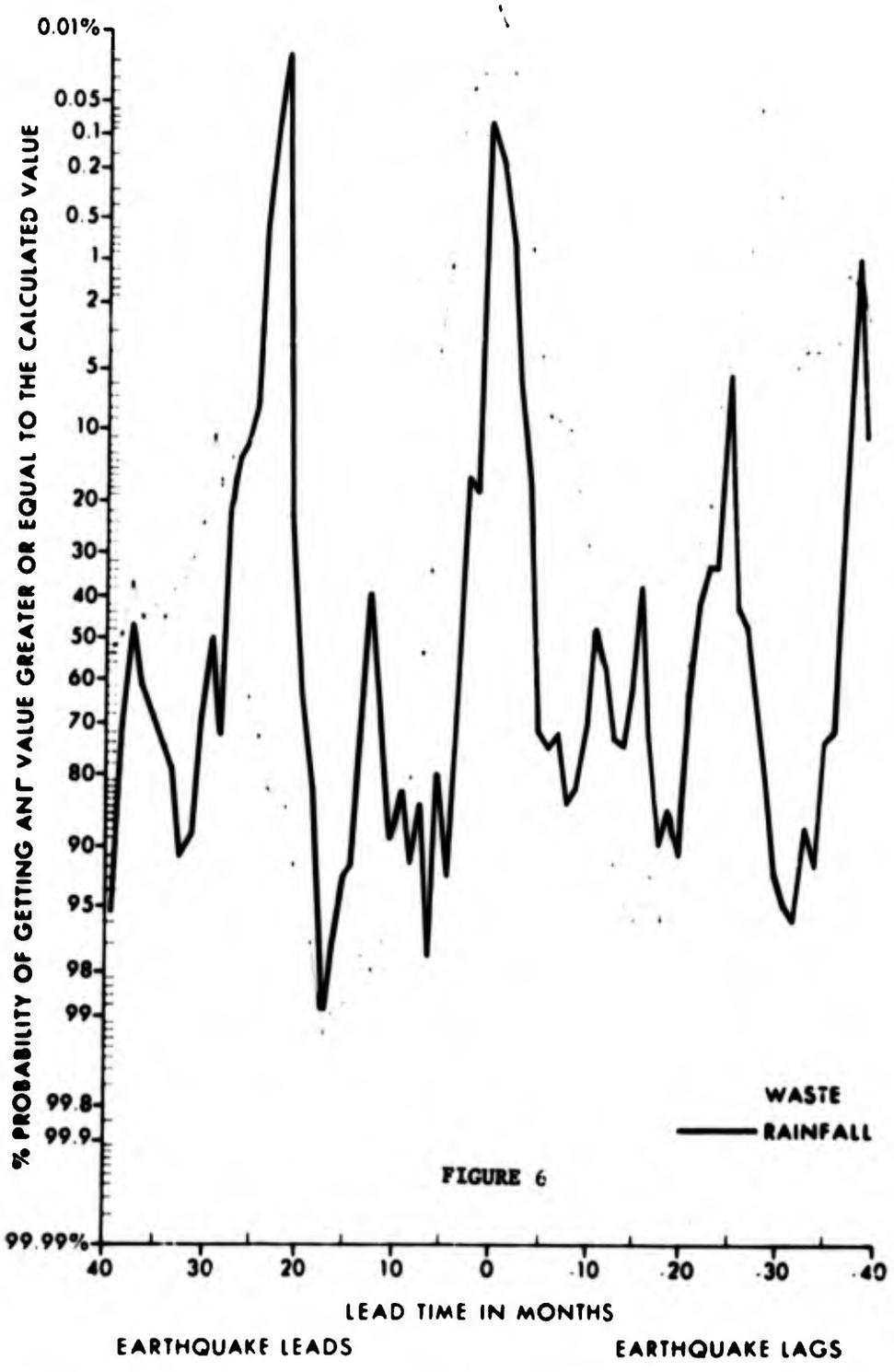

0.01%
0.05
0.1
0.2
0.5
1
2
5
10
20
30
40
50
60
70
80
90
95
98
99
99.8
99.9
99.99%
% PROBABILITY OF GETTING ANY VALUE GREATER OR EQUAL TO THE CALCULATED VALUE
WASTE
RAINFALL
40
30
20
10
0
10
20
30
40
LEAD TIME IN MONTHS
EARTHQUAKE LEADS
EARTHQUAKE LAGS

FIGURE 6

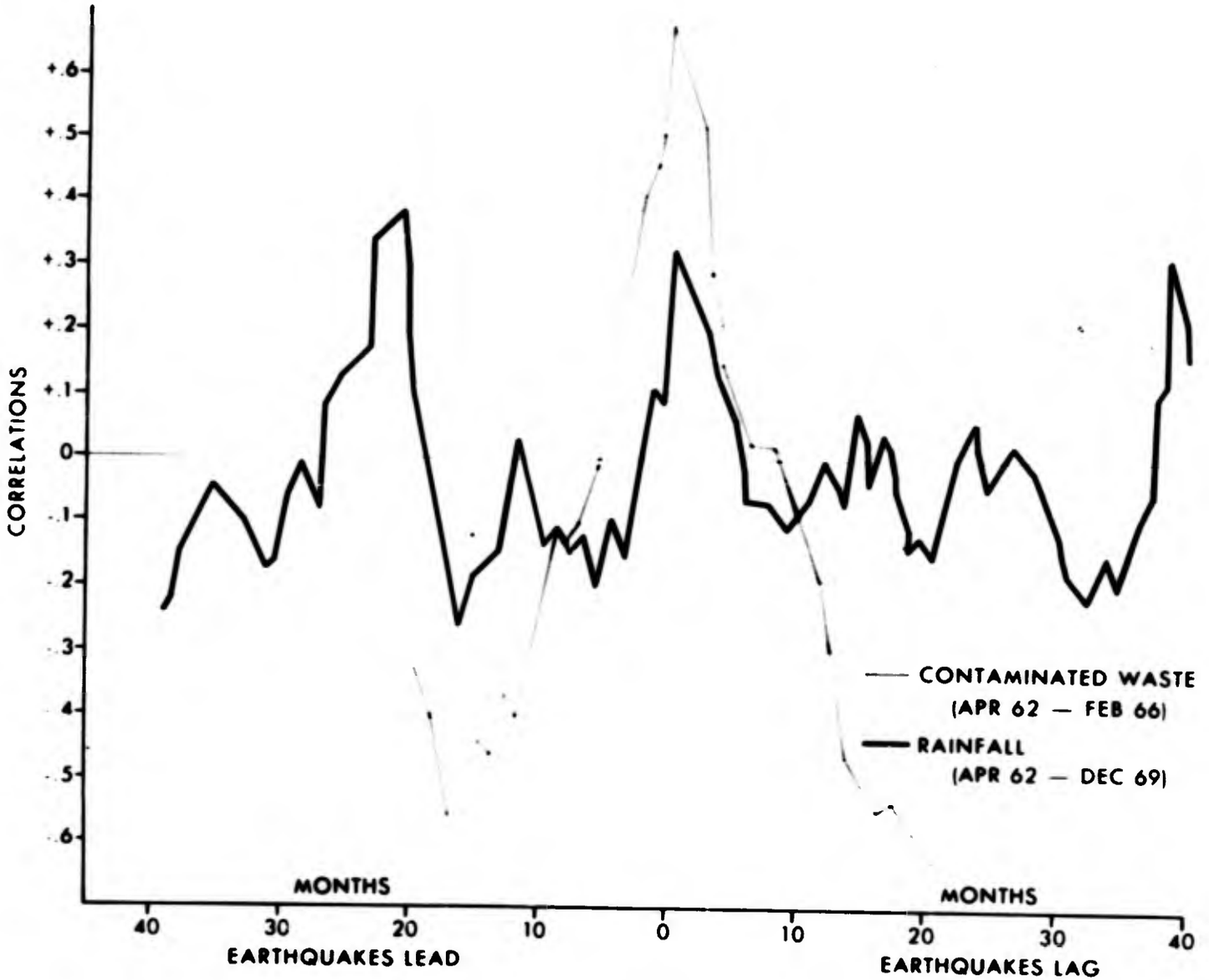

CORRELATIONS
+.6
+.5
+.4
+.3
+.2
+.1
0
-.1
-.2
-.3
-.4
-.5
-.6
CONTAMINATED WASTE
(APR 62 — FEB 66)
RAINFALL
(APR 62 — DEC 69)
MONTHS
40
30
20
10
0
EARTHQUAKES LEAD
MONTHS
10
20
30
40
EARTHQUAKES LAG

FIGURE 7

AD 713 527

THE APPLICATION OF A SOLID STATE HELIUM-NEON GAS LASER TO MISSILE GUIDANCE

WILBER H. GIBSON
U. S. ARMY ADVANCED BALLISTIC MISSILE DEFENSE AGENCY
HUNTSVILLE, ALABAMA

Although the laser has many applications in the fields of communication, range finding, medicine and as a weapon; it is in the area of the application of a gas laser to the guidance of Army missiles that this paper shall be devoted. The inherent advantages that the laser can offer have already been shown in past and present developmental programs. This solid state device offers advantages over conventional inertial systems such as smaller size, less weight, less power, no warm up or ready time, no alignment, no maintenance, higher reliability, much higher environmental threshold, lower susceptibility to electronic countermeasure techniques, and much lower cost. To provide a clear understanding of the impact of the foregoing statements, a description of the guidance systems in question shall be given.

A. MISSILE GUIDANCE AND CONTROL

Conventional missile guidance systems use as the basic sensor system a platform, which is an electromechanical device with three (3) or four (4) gimbals mechanically constrained to rotate in the azimuth, pitch and roll axes of the missile. The gimbals are supported on mechanical bearings and are servo controlled by means of electrical pickoffs and torquers (microsyns). The inertial sensors are gyroscopes (gyros) aligned so as to sense the missile movements in the three (3) coordinate axes of azimuth, pitch and roll.

The gyro is a precision instrument which is highly sensitive to the most minute movement, thereby providing a signal through the servo system to the respective platform gimbal to maintain the gimbals in their proper alignment in inertial space regardless of the vehicular movements.

The gyro contains a spinning wheel which is supported on mechanical bearings and is typically rotating at 24,000 R.P.M. to

provide a high inertia. This spinning wheel and its supporting structure obeys the physical laws of gyrodynamics, thereby providing the basis of the so called conventional inertial system.

The gyro and the platform, by virtue of their mechanical structure, are highly susceptible to linear and angular accelerations, long term operating errors, vibratory environments, shock and temperature effects. The mechanical gimbals and bearing support structures deform under environmental conditions, and therefore produce errors. Due to the thermal effects, those systems usually require heaters to meet performance goals. The need for the gyro wheel to reach its speed of 24,000 R.P.M., and the thermal stability level required contribute to the warm up or "ready" time of the platform as well as the requirement to align the gimbals to a precise angle prior to missile launching. A constant source of error in such a system is the "random" drift of the gyro. This error is caused by the physical limits of accuracy and materials control that can be applied to the design and construction of the gyro. It should also be noted that the multitude of parts contributes to the higher cost and lower reliability of the conventional inertial system as compared to a laser system.

B. LASER SYSTEMS

The laser gyro system is a solid state device with only one (1) structural element: a quartz block. This block does not move, and contains no moving parts unless a beam of light can be considered a moving part. With such a physical characteristic the environmental susceptibility of the device appears to be nil. Voluminous test data has shown that a well integrated design, for a particular application; is truly free of the long tolerated environmental effects that have plagued the conventional inertial sensors. The secret of this success is embodied in several years of research and development, and will be discussed to some depth to provide the basis for the claims previously made.

1. Laser Technology

Briefly, the basic laser operation will be outlined to provide a clear understanding of the application of a gas laser to the role of a gyro obeying the laws of gyrodynamics. The acronym"Laser" stands for "Light Amplification by Stimulated Emission of Radiation". The first condition that must be created to obtain the lasing action is to have a gain medium (amplification) to overcome losses, which in this case, is a tube filled with a mixture of helium and neon gases at a low pressure. A voltage is applied across a metallic anode and a cathode which ionizes the gas. Excited helium atoms collide with neon atoms transferring energy to the neon atoms and raising them to a higher energy level. The neon atoms, following the basic laws of physics, drop back to their normal energy level, but in doing so they release a photon (visible or invisible light). The photons strike

other excited neon atoms, thereby stimulating the release of new photons, and so the procedure by which the "lasing" takes place is begun. To continue this action a reasonant cavity is required, which is selected in a triangular form with mirrors at each corner so aligned as to reflect the light along the legs of the triangle. When a resonant cavity receives gain which is greater than its losses, the radiation intensity will increase until a steady state condition has been reached. The laser is now operating.

2. Laser Gyro

The laser gyro (Fig 1) operates in the same manner as outlined above, but certain other refinements are added. The mirrors are accurately aligned to insure that the laser beam is in an enclosed triangular path. We now have a beam of light moving in a clockwise (CW) direction, and another beam moving in a counter clockwise (CCW) direction. It is evident at this time that both beams have the same frequency. If the base upon which the triangle rests is now rotated in the plane of the triangle, then one beam will have a greater apparent distance to travel, while the other beam will have a shorter apparent distance to travel in making one revolution around the triangle. An examination of the output of this device shows that when the triangle is rotating in a given direction at a given rate, the two beams have different frequencies, and further; that this frequency difference is directly proportional to the rate of the base rotation. This is the basis on which the device described above can be used as a gyro. The conventional inertial gyro containing a spinning wheel performs a similar task, but the output of this instrument is a mechanical movement which is transformed into an electrical signal, ergo more error. The output of the laser gyro is the frequency difference of two light beams which is directly translatable to a digital electrical output. The history of the laser gyro shows a very rapid development over the past ten years. The Army, Navy, Air Force and NASA devoted small amounts of R & D funding to this area over this period of time, but the most notable advance in the laser gyro state-of-the-art was achieved on the LARS program; an Army funded study initiated in April 1966. The purpose of this program was to perform the basic and advanced research necessary to solve the then current operating problems, and produce a three axis system capable of demonstrating operation in severe environments in an exhaustive test program. Acceleration tests to 500 g's, shock to 800 g's, and random vibration at an energy level of 20 G2/CPS (20 to 2000 GPS) were successfully passed. No acceleration or vibration (g, g^2, and g^3) effects could be detected, thereby heralding the success of the LARS program. A refinement of this program is being used on the ABMDA UpStage program (See Fig 2).

Laser gyros manufactured to date exceed a hundred deliverable items, not counting test and evaluation units made on development contracts. These gyros were and are being produced for the

following agencies:

Agency	Application
NOTS	shipboard fire control
ABMDA	LARS (R&D)
AECOM	airborne navigation
NASL	navigation
NASA-ERC	navigation
NWC-CL	ATIGS-missile
ABMDA	UpStage G. & C.
NASA-Goddard	sounding rocket

Laser gyros have been successfully designed and manufactured by Honeywell, Inc., Autonetics, Sperry and at the A.I.G. & C. Laboratory at the U. S. Army Missile Command. Based on this activity, it is felt that adequate data has been gathered to permit a comparison of conventional and laser inertial systems. Using average figures for comparable systems applications, the following pertinent facts are outlined: (note: L=Laser, C=Conventional)

Rate Capability:

C - Present design capability - 200°/sec. Increase possible with gyro and electronics redesign, and an adverse effect on dynamic range, linearity, and threshold. Limited by electro-mechanical capability to about 400°/sec.

L - Present design at 2000°/sec. Increase by readout electronics bandwidth modification.

Reaction Time:

C - Requires warm up, gyro wheel spin up, thermal stability level reached, and platform gimbal alignment.

L - Instantaneous

Reliability:

C - 5,000 hours MTBF

L - 17,000 hours MTBF

Power Requirements:

C - 90 watts. Power changes as missile rates change.

L - 30 watts. Power constant regardless of missile rates.

Acceleration (G) Capability:

C - 40 g as a strap down unit without special suspension or

isolation design.

L - 500 g as a strap down unit. No limit as yet foreseen. (within materials capability)

System Complexity:

C - Platform (including gyros) 280. Electronics 494 parts-total 774 parts.

L - Laser platform 34 parts - most optically bonded. Electronics 288 parts - total 322 parts. (Note: The electronics noted above are strictly platform ancillary needs, i.e., servo systems, heater control, etc. No computer hardware is included.

Size and Weight:

C - 6" x 8" x 9" long (432 in^3) at 18 lbs. (new subminiature design)

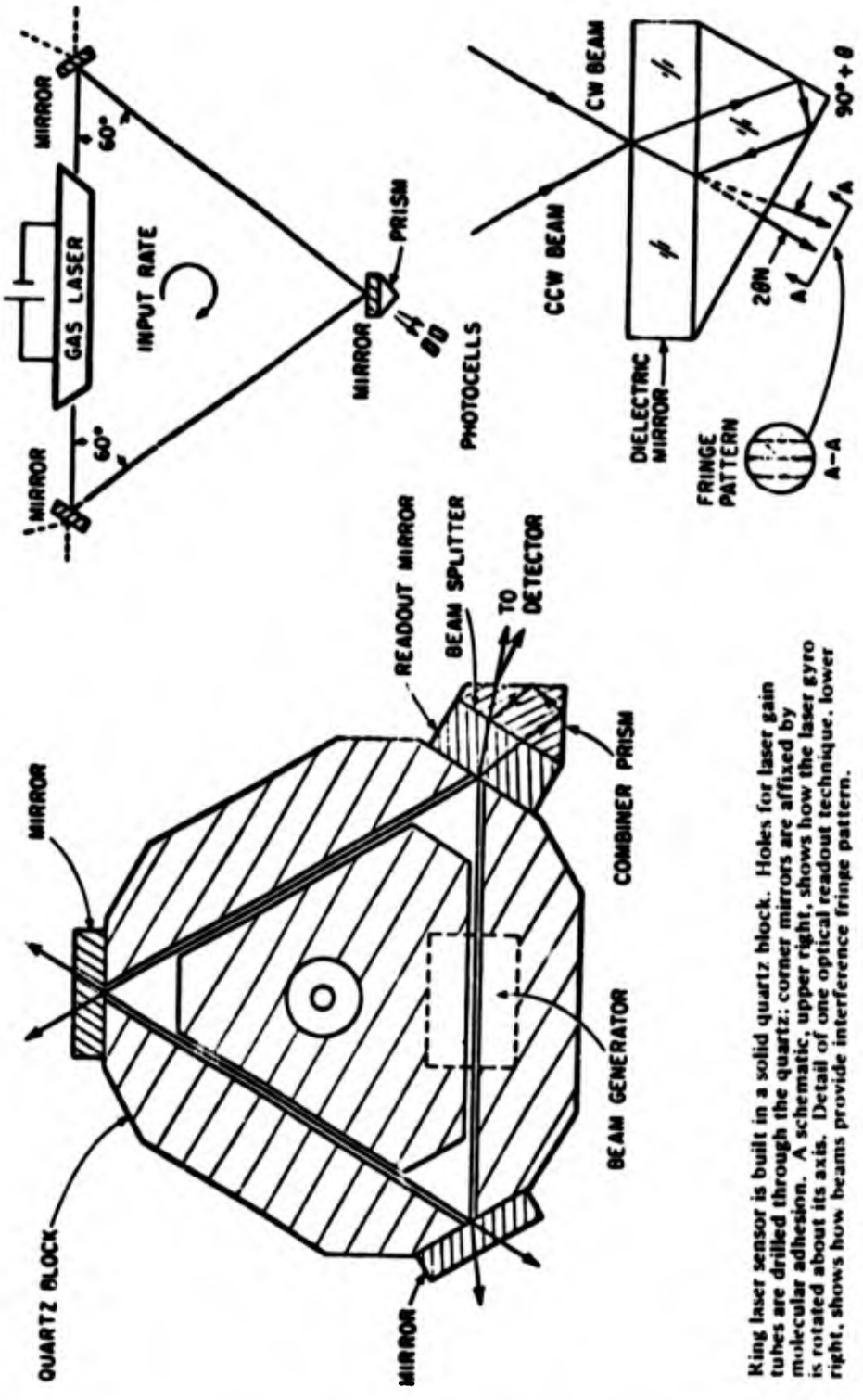
L - 3.7" x 5.8" x 8" long (171 in^3) at 5 lbs. (no attempt to miniaturize)

Platform costs have been examined to determine if a cost advantage is evident. Production costs of the platform for the Sergeant missile are about $72,000 per system, and for the Pershing missile platform about $96,000 per system. These costs reflect the platform and its ancillary electronics, which is the same level at which the laser system was examined for the same applications with the resultant costs of $3,800 for Sergeant, and $4,100 for Pershing. With potential savings of this nature available in addition to the other advantages, the future consideration of laser systems appears to be certain. Such a system is shown in figure 3 in block diagram form.

C. CONCLUSIONS

The laser gyro system has been developed and tested to the point where it can be applied to Army, Navy, and Air Force missile systems with several inherent advantages such as lower cost, lower error budgets, lower power requirements, smaller size, less weight, and higher reliability than conventional inertial systems. Although the Air Force is utilizing the system in an aircraft program, and ABMDA is incorporating a laser system on its UpStage experiment; no consideration has yet been given to using this outstanding advance in the state-of-the-art for a tactical missile system. Tomorrow's missile systems must be developed in such a manner that the prime goals of low cost, no maintenance, and high reliability can be a reality. Further, consider the fact that the performance levels required for each new missile system developed over the past years has resulted in added time and cost to most of these programs to

provide for new inertial components and systems development, or the redesign of an existing system or component to meet increased rates, faster reaction times, or higher environmental regimes. Prime examples of this can be found by choosing almost any missile system that comes to mind: Sergeant, new platform; Pershing, new gyros and accelerometers; Lance, new gyros and redesigned accelerometer; Minuteman, new platform and gyros; Thor, new platform and gyros; Hibex, redesigned gyros; Polaris, new platform and gyros; Spartan, new platform and gyros; Sprint, new and redesigned gyros - - - and so on. The laser gyro in single axis or triad forms is the basis on which a variety of system requirements can be built <u>without developing a new sensor!</u> One need only to assemble the electronics required for the application.



Ring laser sensor is built in a solid quartz block. Holes for laser gain tubes are drilled through the quartz; corner mirrors are affixed by molecular adhesion. A schematic, upper right, shows how the laser gyro is rotated about its axis. Detail of one optical readout technique, lower right, shows how beams provide interference fringe pattern.

FIGURE 1. RING LASER SENSOR

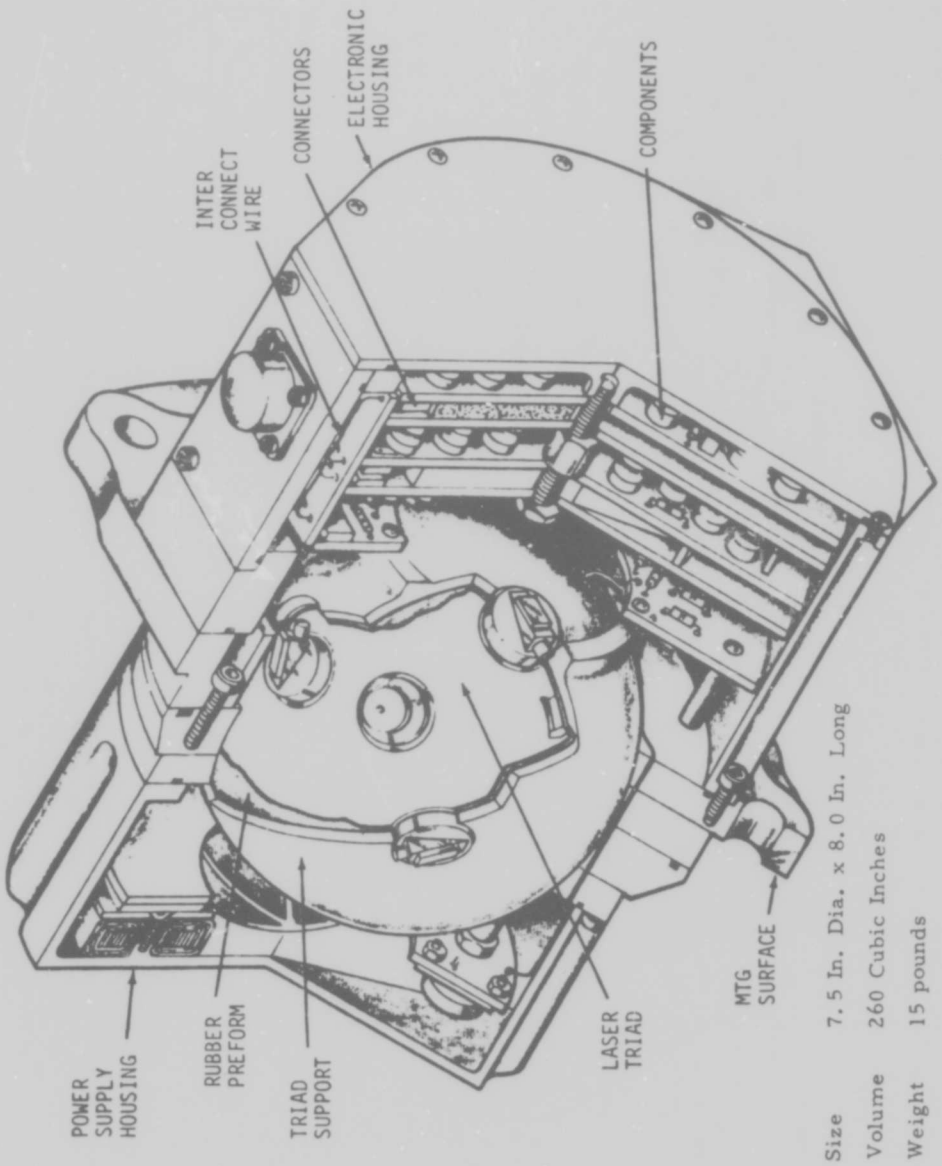


FIGURE 2. LASER TRIAD GYRO

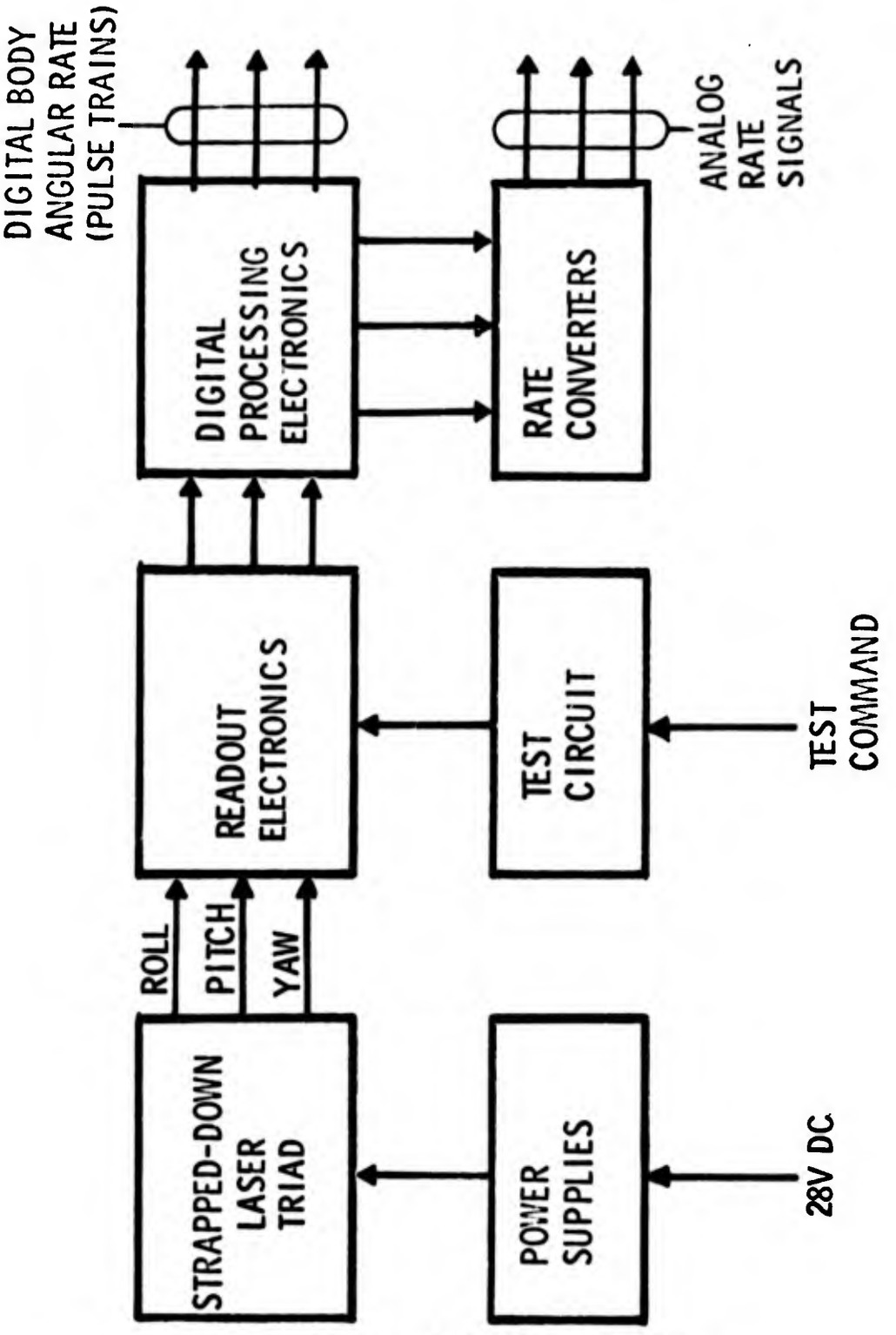


FIGURE 3. LASER RATE GYRO UNIT

AD 713 528

OPEN CYCLE HYDROCARBON-AIR FUEL CELL POWER PLANT

EDWARD A. GILLIS, O. FRED KEZER, JR., WALTER G. TASCHEK
U.S. ARMY MOBILITY EQUIPMENT RESEARCH & DEVELOPMENT CENTER
FORT BELVOIR, VIRGINIA

A new generation of highly reliable, quiet and mobile electrical power sources are required for tactical use. During the past decade many new aids to extend the capabilities of combat forces have been deployed to front line positions. Communication links have been expanded, and devices to detect personnel, vehicles, and projectiles have been perfected. Most of this new equipment has a relatively low power drain, which unfortunately matches the rating of the most unreliable generator sets in the supply system, the 0.5 KW to 10 KW Spark Ignition sets. Improvements in mean time between failure and mean time to repair are absolutely necessary in power plants for front line use. In addition, a low noise signature is desirable to limit detectability of the equipment and to prohibit interference with local personnel.

Army research and development programs in several technological fields have been aimed at solving the tactical power plant problem. The technologies pursued include electrothermal systems such as thermionics and thermoelectrics, thermodynamic systems using various open and closed cycle engines, and the electrochemical fuel cell system. The fuel cell is the only one of these systems which directly converts the fuel's chemical energy to electrical energy.

The operation of a fuel cell is shown in Figure 1. A fuel is electrochemically oxidized, converting the chemical energy of the fuel directly into electrical energy. It is exactly the opposite of classical electrolysis which generates hydrogen and oxygen. In the fuel cell, hydrogen and oxygen provided to the electrodes combine to produce water and electrical energy. This electrochemical energy conversion has several most attractive characteristics: As it is a direct conversion process the efficiency of the process can be very high. As the electrodes are stationary and invariant life is theoretically unlimited, and the process is inherently silent.

Hydrogen and oxygen are not suitable reactants for use in front line positions. The ideal fuel cell for tactical use would

substitute electrodes capable of electro-oxidizing logistic hydrocarbon fuels with oxygen derived from ambient air. Air breathing cathodes with high performance levels have been developed. Though research has found anode catalysts which electro-oxidize specific long chain hydrocarbons, a versatile catalyst capable of economically electro-oxidizing all the hydrocarbon forms in the polyglot military fuels and tolerant to fuel additives or impurities has not yet been found. Consequently direct oxidation hydrocarbon-air fuel cells are not sufficiently advanced to permit consideration for near term use.

To meet nearer term needs for tactical power, a device can be added to the fuel cell which converts the hydrocarbon fuel to hydrogen for immediate consumption by the fuel cell. This is termed an indirect hydrocarbon-air fuel cell system. Adding a hydrogen generating device compromises the fuel cells basic simplicity and efficiency. Development efforts, therefore, have focused on minimizing the cost and complexity of the indirect fuel cell system.

Early development model indirect hydrocarbon-air fuel cell systems used alkaline electrolyte (potassium hydroxide) fuel cells as technology had been advanced under aerospace funding for hydrogen-oxygen power plants. Potassium hydroxide electrolyte was a logical choice for hydrogen-oxygen systems as it exhibits the highest ionic conductivity of candidate electrolytes, and is not overly corrosive. The electrochemistry of the limiting performance electrode, the oxygen consuming cathode, was well documented from prior research on air depolarized batteries, and the oxygen reaction proceeds most rapidly in base electrolytes. However, the adaptation of this technology to hydrocarbon-air fuel cell systems was difficult for two reasons. First, the potassium hydroxide electrolyte will react chemically with carbon dioxide forming a salt which greatly reduces the ionic conductivity and cathode performance. Carbon dioxide which is present in air must be removed before air can be used in the cell. Second, the concentration of the electrolyte must be maintained constant in the cell to prevent either dilution or concentration which would severely reduce cell performance. The excess water produced in the cell can be evaporated into the process air stream at the correct rate only if the cell temperature, and air flow rate are properly matched with the power level. As military power plants have to operate in ambients from -65° to 125°F, in dry and wet climates, moisture control is complex. Subsystems must be added to the power plant for cell thermal control and for purification and climatic conditioning of the process air.

The electrolyte reaction with carbon dioxide similarly places constraints on the hydrogen generator. All feasible processes to convert the hydrocarbon fuel to hydrogen also produce carbon dioxide and other compounds harmful to the cell. The separation of hydrogen from the process stream adds to the complexity of the power source.

In all, the operational process for the indirect alkaline

electrolyte hydrocarbon-air fuel cell is very complex and requires highly sophisticated control. Though alkaline electrolyte power sources have been built and their technical feasibility established, it does not appear to be a practical or economical solution for tactical power plant requirements.

Recent advances in material technology have allowed consideration of electrolytes for fuel cells which are more corrosive than potassium hydroxide but which are unaffected by air and fuel stream impurities. Phosphoric acid in particular has two most desirable features for indirect system use:

1. It is thermally stable and non-reactive with any component in air or a hydrogen product stream derived from logistic fuels.
2. It is usable at modest temperatures of 250° to 300°F, temperatures where carbon monoxide is not strongly adsorbed on the anode catalyst, and where the fuel cell subsystem waste heat may be removed by process air.

Substituting a phosphoric acid electrolyte fuel cell for the alkaline electrolyte system allows major system simplification. A process schematic of this system is shown in Figure 2. The fuel cell process air requires only removal of particulate matter. Similarly, the product stream of the hydrogen generator may be passed directly into the fuel cell where the hydrogen is selectively electro-oxidized.

The hydrogen generator shown in Figure 2 utilizes a steam reforming reaction. This is a highly efficient process where fuel vapor and steam react over a catalyst to produce primarily carbon dioxide and hydrogen. It is particularly efficient for the fuel cell because hydrogen is made available from both the fuel and the water feed stock though some of this additional hydrogen must be burned to sustain the endothermic reforming reaction. Because the reformer product gas stream does not have to be purified, the reaction can be accomplished over a catalyst bed at ambient pressure and temperatures of 1100° to 1300°F. Structural material requirements are not severe for these conditions.

A water recovery subsystem is also shown in Figure 2, with the product water of the fuel cell and the burner condensed from the flue gas. The reforming reaction requires two to three times as much water as fuel. Since clean water is scarce in many military positions, water recovery is mandatory.

Privately sponsored research directed toward commercial hydrocarbon-air fuel cell applications has concentrated on the reformer - phosphoric acid indirect system for the past several years. Both the reformer and the fuel cell subsystems have been brought to an economically and technically advanced level.

During 1968 a special study was conducted at USAMERDC to determine preferred Power Plant approaches for silent tactical applications.[1] This study concludes that the steam reformer-phosphoric

acid electrolyte fuel cell system has the highest cost effectiveness potential of any known candidate power source in the 1980 time frame. Further, the study concludes no technological breakthroughs are needed to develop this power plant, but the militarization requires a comprehensive development effort in design and control optimization of each subsystem and each interface. This encompasses Advanced Development and Engineering Development category efforts.

Paraphrasing the above, adapting a commercial indirect fuel cell system to a military system requires extensive development due primarily to the military requirements of operating on low grade fuels over a wide climatic range. The subsystem on this power plant which basically limits military potential is the water recovery loop; it forces closing of two process loops and forces concern over cold ambient operation.

In 1967 USAMERDC began an in-house investigation of alternate hydrogen generation schemes which would not be water dependent and would use referee grade fuels. The phosphoric acid fuel cell was retained as the hydrogen consumer because of its ability to operate on impure fuels. A phosphoric acid electrolyte fuel cell was built and tested to determine minimum quality fuel requirements and control requirements for steady state and transient conditions. This data was used to determine the minimum hydrogen generator performance requirements.

Regarding control requirements, tests on a twelve cell phosphoric acid electrolyte stack proved that a fixed air flow rate over the cathodes, in excess of the fuel cell's oxygen requirements was acceptable because the nearly anhydrous electrolyte could not be dried out. Therefore, by using only cell temperature as the control parameter the reactant air supply stream may be employed for moisture and thermal control of the cell stack. In other words, as long as sufficient air is provided for thermal control of the cell stack, oxygen and moisture balance requirements are fulfilled. The cells entire thermal, oxygen, and moisture control system could consist of a fixed speed fan blowing air over the cathodes with air throughput controlled by a thermostatic damper. A simple recycle of hot process air mixed with some ambient air could reduce the cooling air temperature rise to a low level, reducing the thermal gradient in the cell and maximizing electrochemical performance.

Figure 3 shows the calculated humidity of the recycle air for minimum and maximum load over all ambient conditions. The corresponding equilibrium electrolyte concentration variation is only about two percent.

Similarly, on the fuel electrode side, it was found that a stream grossly inferior to that supplied by a reformer was acceptable to this cell. Fuel streams with twice the CO content and one-third the hydrogen content of a reformer stream were used efficiently. Figure 4 shows the change in cell performance as a function of hydrogen concentration. A poor quality fuel can also be used in the cell during start-up, so the cell can bootstrap itself to operating temperature. Figure 5 shows the fuel cell's performance on impure

fuel as a function of cell temperature during start-up.

During the course of testing this fuel cell stack over a 14 month period, it was operated under very abusive conditions, both intentionally and inadvertently. It endured 100°F over-temperature excursions of several hours duration. Fuel and/or air flow was stopped with the stack under load and at operating temperature. It was stored hot, without reactants, for a two day period. Little or no performance loss was experienced by these and other abusive test conditions which indicates the stack will not contribute a domino failure effect to a power plant. One or all of these excursions would permanently disable, sometimes catastrophically, any other candidate fuel cell. Few if any other candidate power plant technologies could tolerate such excursions without suffering other malfunctions.

Exploratory development of alternate hydrogen generating processes has very recently proven the feasibility of the power plant shown schematically in Figure 6. The hydrogen generating process shown is regenerative thermal cracking. In this process, two metal catalyzed ceramic packed reactors are cycled between modes where in one reactor, fuel directed through the hot packing cracks to hydrogen with the carbon depositing on the ceramic. Concurrently, the second reactor is provided air which burns the deposited carbon and reheats the ceramic. On present laboratory model reactors, the process streams are switched at one to three minute intervals. Reactor temperatures vary during the cycle between 1500° and 2000°F.

The power plant schematic of Figure 6 shows major simplifications over a reformer type of hydrogen generator:

1. No water loop as the process is not water dependent.
2. All open loop processes and the hydrogen generator can operate independent of the fuel cell. Feedback from the fuel cell will control the fuel pump rate to electrical load for system throttling. The title given this power plant, Open Cycle, is derived from the open loop process.

The product gas compositions and flow rates for a complete cracking and burn-out cycle are shown in Figure 7. The fuel in this case was Combat Gasoline. The quantity of hydrogen produced during the cycle is 88% of that contained in the fuel feed stock. The remainder of the hydrogen is found primarily in methane plus trace quantities of ethane, benzene, and water. The product composition and flow rate for kerosene type fuels is equivalent.

Reactors built to investigate engineering aspects of the process show agreement with the above data. These reactors routinely yield a slightly higher hydrogen concentration, 93%, because of less dilution by the combustion gas hold-up in instrumentation lines. Some hydrogen, however, is lost in purging combustion products from the reactor, thereby reducing hydrogen yield to about 80% of that available in the feed stock.

Control requirements of the reactor to prevent carbon

plugging have been investigated. The heat of combustion of the fuel's carbon to carbon dioxide is approximately six times the energy required to crack the equivalent amount of fuel. Even with no heat exchange from combustion products to inlet air, over twice the heat energy required for cracking is available. The combustion gas compostion of Figure 7 and other test results show that sufficient carbon will be removed from the reactor if air is provided at a minimum rate equal to an average 50% CO, 50% CO_2 combustion. Higher air rates produce a higher percentage of CO_2 which releases more heat and offsets the cooling of the higher air flow rates. Reactors have operated at a stable temperature with air flow rates up to twice the minimum rate for carbon removal. The reactors are also self-sustaining in temperature over a wide fuel input rate. Test reactors have operated with stable temperatures and product quality on fuel input rates varied from .15 to 1.0 pound per hour of fuel. These tests prove that the regenerative thermal cracker's control requirements to prevent carbon build-up and maintain temperature level are not strict, and that this process operates efficiently at part load conditions.

The feasibility of using low grade fuels has been proven. Combat Gasoline has been the routine fuel for laboratory tests. A reformer system cannot use leaded fuels as lead removal devices have not been found. In the regenerative thermal cracker, lead is removed from the packing during the burn-out much as it is in an engine. Sulfur in the fuel is similarly retained on the packing during the hydrogen generation portion of the cycle and then burned off by the air. This is to be expected as the nickel catalyst used favors reduction of hydrogen sulfide in the reducing atmosphere of the cracking cycle, while the formation of sulfur dioxide is favored if oxygen is present.

The use of low grade fuels has also been demonstrated on a laboratory model power source. A phosphoric acid electrolyte fuel cell was operated for over 200 hours on the thermal cracker hydrogen product. The fuel feed stock was Combat Gasoline, with no pre- or post purification of the fuel except for a paper "air" filter to remove the trace of carbon dust carry over. The fuel cell did not degrade in performance during this test.

The disadvantage of the Open Cycle system is that the fuel cracker has a higher fuel consumption than a reformer type process. Practical efficiency limits for the reformer-fuel cell system is 30%, while the Open Cycle is 20%. This is still approximately twice the efficiency of small engine-generator sets. Figure 8 shows the effect of efficiency on Life Cycle Cost[2] for a 3 KW reformer-fuel cell system. The decrease from 30% to 20% efficiency only raises the life cycle cost by a couple of percent. Also shown on Figure 8, however, is the effect of reducing just the power plant weight. As the water loop (condensers, fan, reservoir, etc) has been deleted a weight savings of 20% or more is possible with the Open Cycle system so this alone shows almost a 10% reduction in life cycle cost due to a decrease in transportation and handling

costs. Factoring the lower first cost of the Open Cycle power plant, and lower maintenance costs because of the system's reduced complexity promises a further significant decrease in life cycle costs of tactical electrical power.

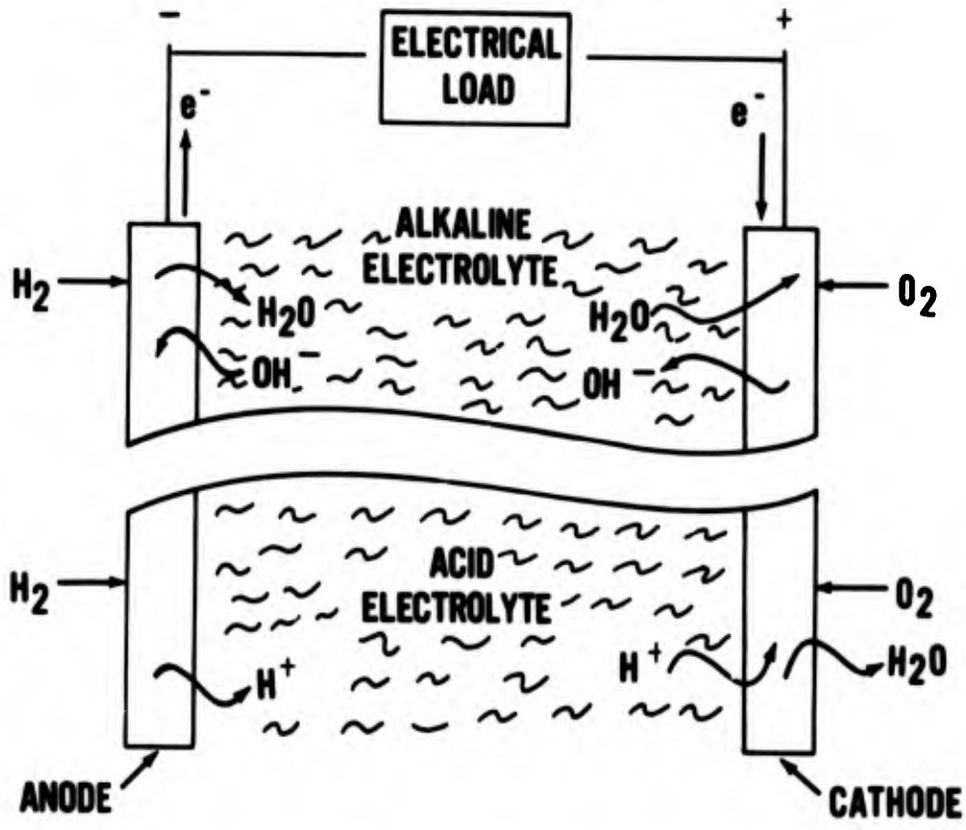
The Open Cycle fuel cell system has the potential to meet the requirements for tactical use. Based on feasibility model regenerative thermal crackers, and advanced development phosphoric acid cell subsystems, a weight reduction of 20-25% over the comparable engine-generators appears possible. The audible noise signature promises to be very low as the only noise producing components are two air blowers and the fuel pump. The reliability level of the Open Cycle is not verified. However, phosphoric acid fuel cell subsystems in this power range, complete with a cooling air blower and thermal controls have exceeded 4500 hours of unattended operation. Test reactors have exceeded 400 hours of satisfactory operation and the temperature levels involved are within permissible limits of ceramics and alloys common to the chemical processing industry and in gas turbine practice. As the system process components are few in number and unsophisticated the overall power plant reliability is potentially very high. In addition, modular design offers the possibility of field repair by modular replacement.

Advanced development model 1.5 KW subsystems, both the fuel cell and regenerative thermal cracker, are presently being fabricated. Military potential testing of each subsystem will begin at the start of Fiscal Year 1971. The advanced development program for a family of Open Cycle power plants of 0.5, 1.5, 3.0 and 5.0 KW sizes has begun. Introduction of these power plants to field personnel is scheduled to begin in 1976.

REFERENCES:

(1) "Silent, Electric Power Generators for Tactical Applications." USAMERDC Report 1935, December 1968.

(2) "Cost Effectiveness Analysis of Military Hydrocarbon Fuel Cell Power Plants." Contract DAAK02-68-C-0524, Final Report, March 1970.



HALF CELL REACTIONS: ALKALINE ELECTROLYTE

ANODE: $H_2 + 2OH^- \longrightarrow 2H_2O + 2e^-$

CATHODE: $\frac{1}{2}O_2 + H_2O + 2e^- \longrightarrow 2OH^-$

HALF CELL REACTIONS: ACID ELECTROLYTE

ANODE: $H_2 \longrightarrow 2H^+ + 2e^-$

CATHODE: $\frac{1}{2}O_2 + 2H^+ + 2e^- \longrightarrow H_2O$

NET REACTION: ALKALINE OR ACID ELECTROLYTE

$H_2 + \frac{1}{2}O_2 \longrightarrow H_2O$ + ELECTRIC POWER

FIGURE 1 - Fuel Cell Reactions.

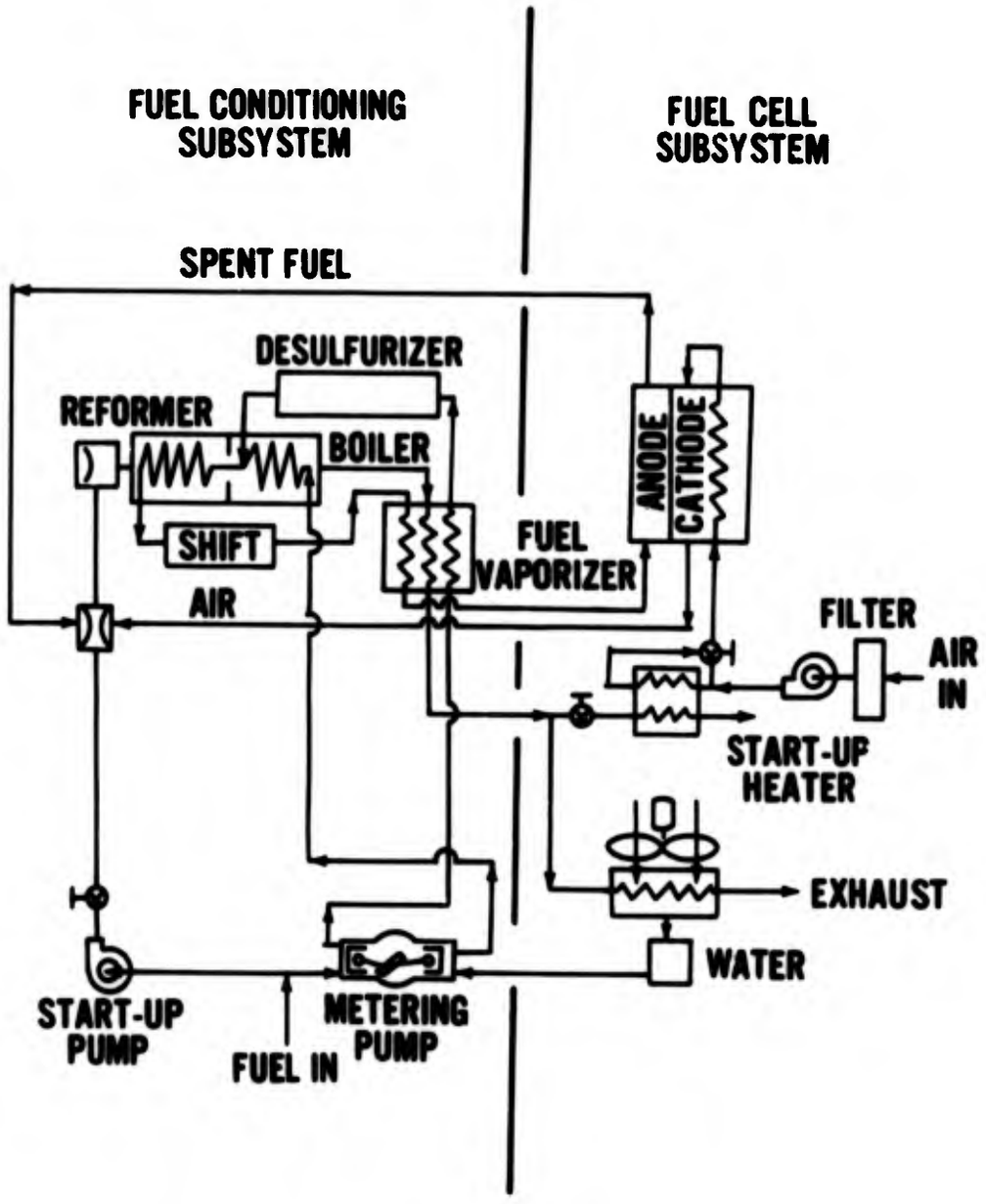


FIGURE 2 - Reformer/Phosphoric Acid Electolyte Fuel Cell Process Schematic.

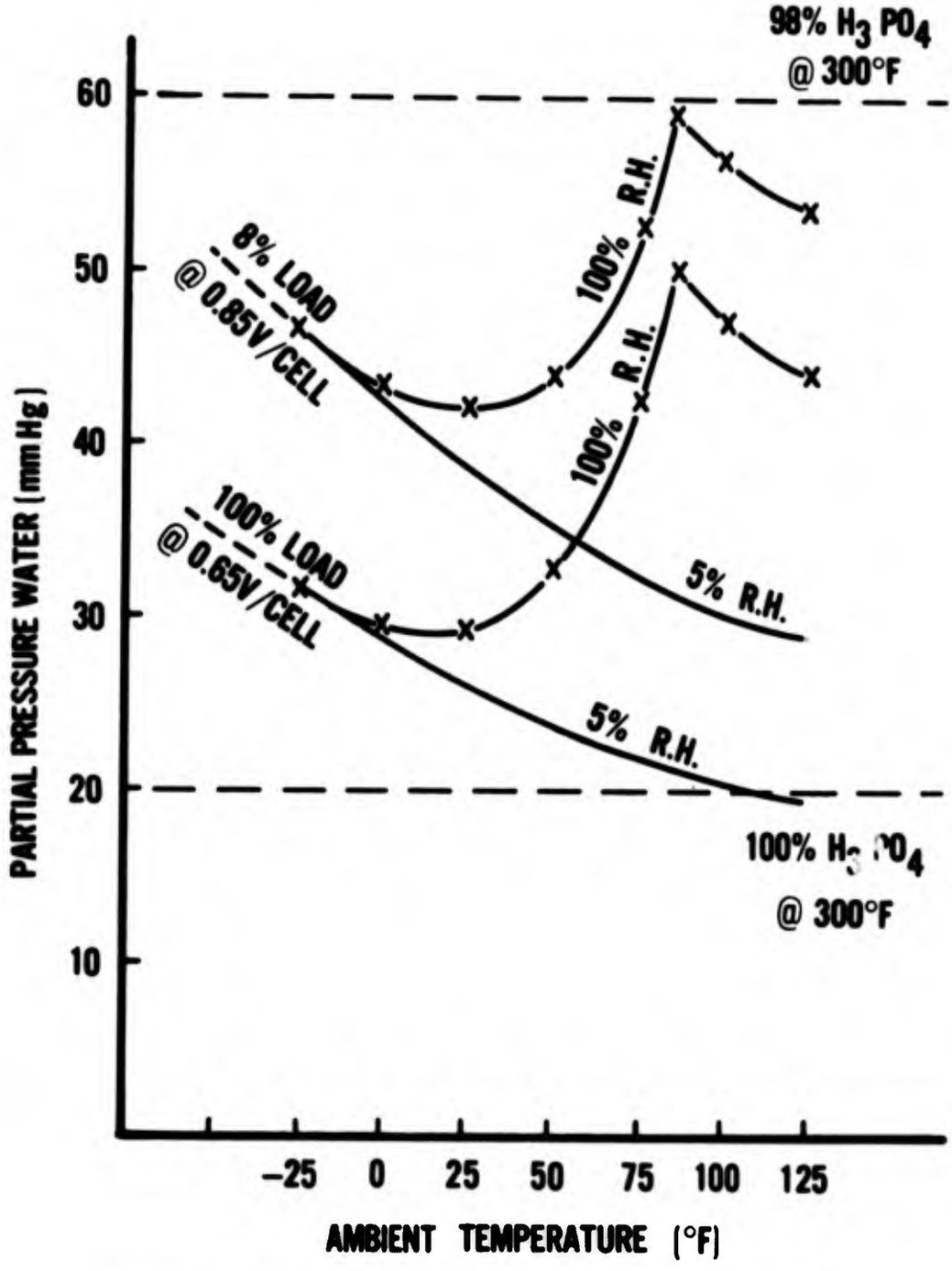


FIGURE 3 - Moisture Balance of Air Cooled Phosphoric Acid Electrolyte Fuel Cell.

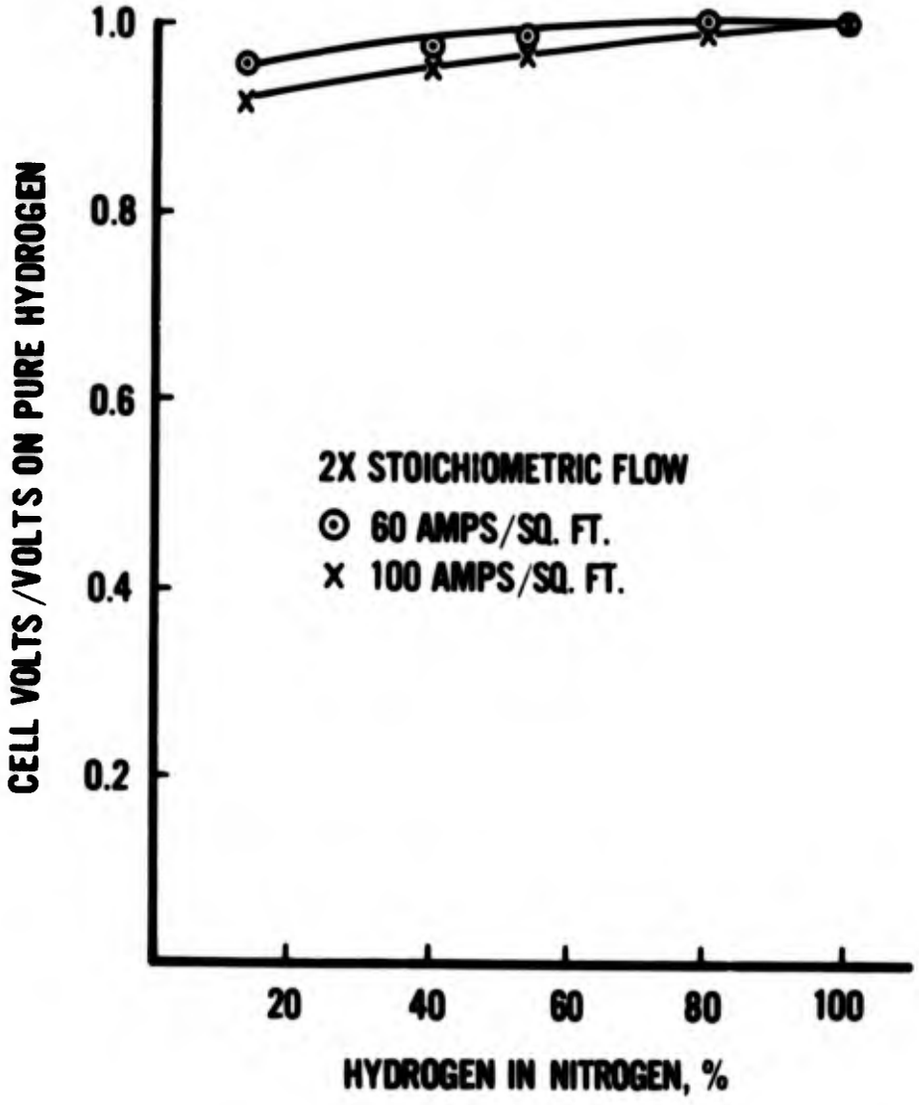


FIGURE 4 - Fuel Cell Performance on Dilute Fuel.

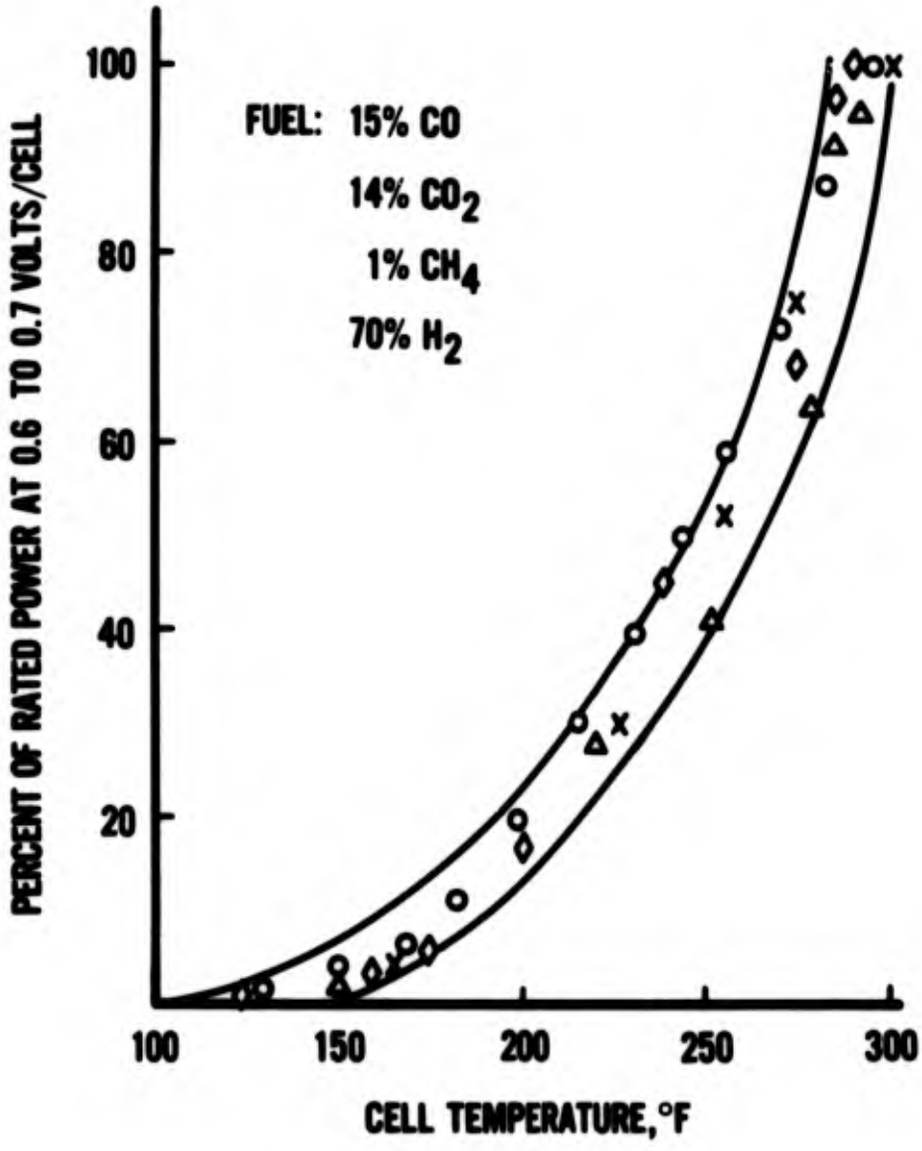


FIGURE 5 - Fuel Cell Performance on Impure Fuel During Start-up.

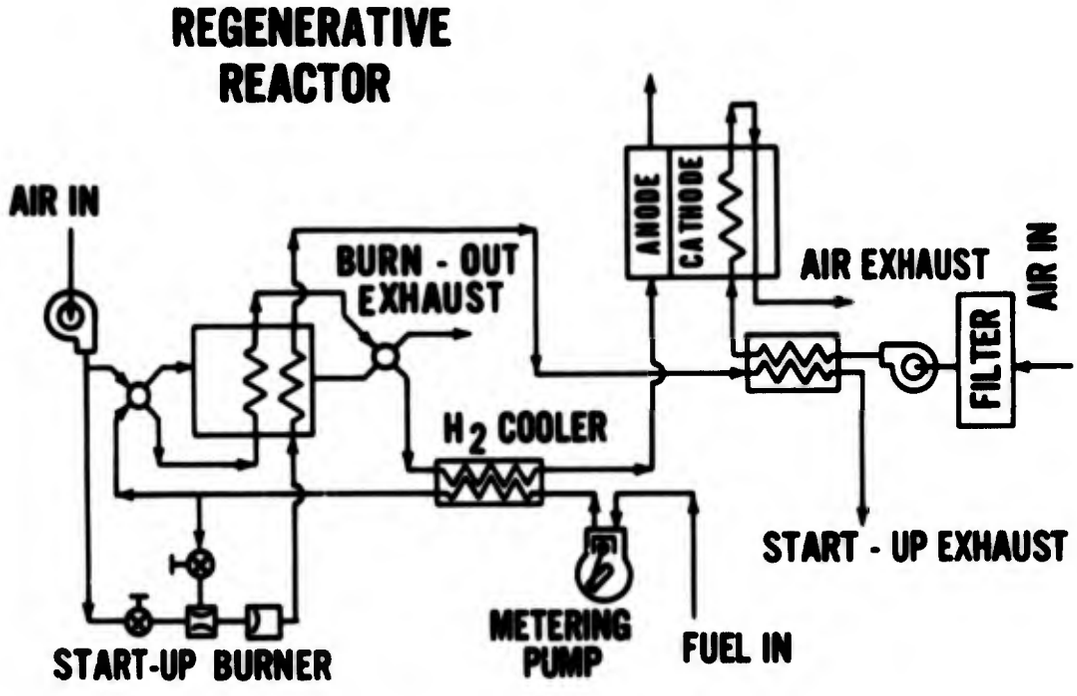


FIGURE 6 - Open Cycle Hydrocarbon-Air Fuel Cell Process Schematic.

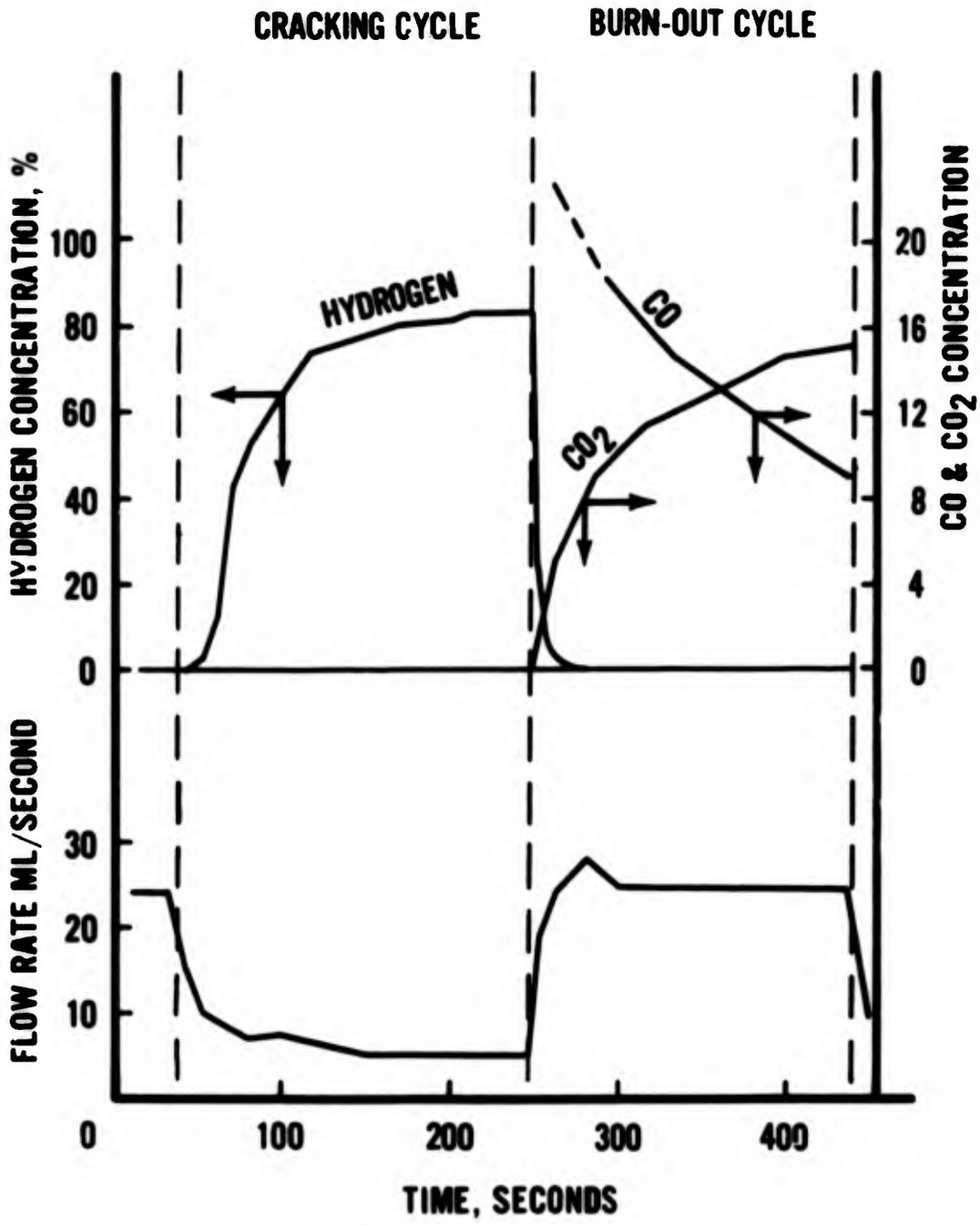


FIGURE 7 - Regenerative Thermal Cracker Performance.

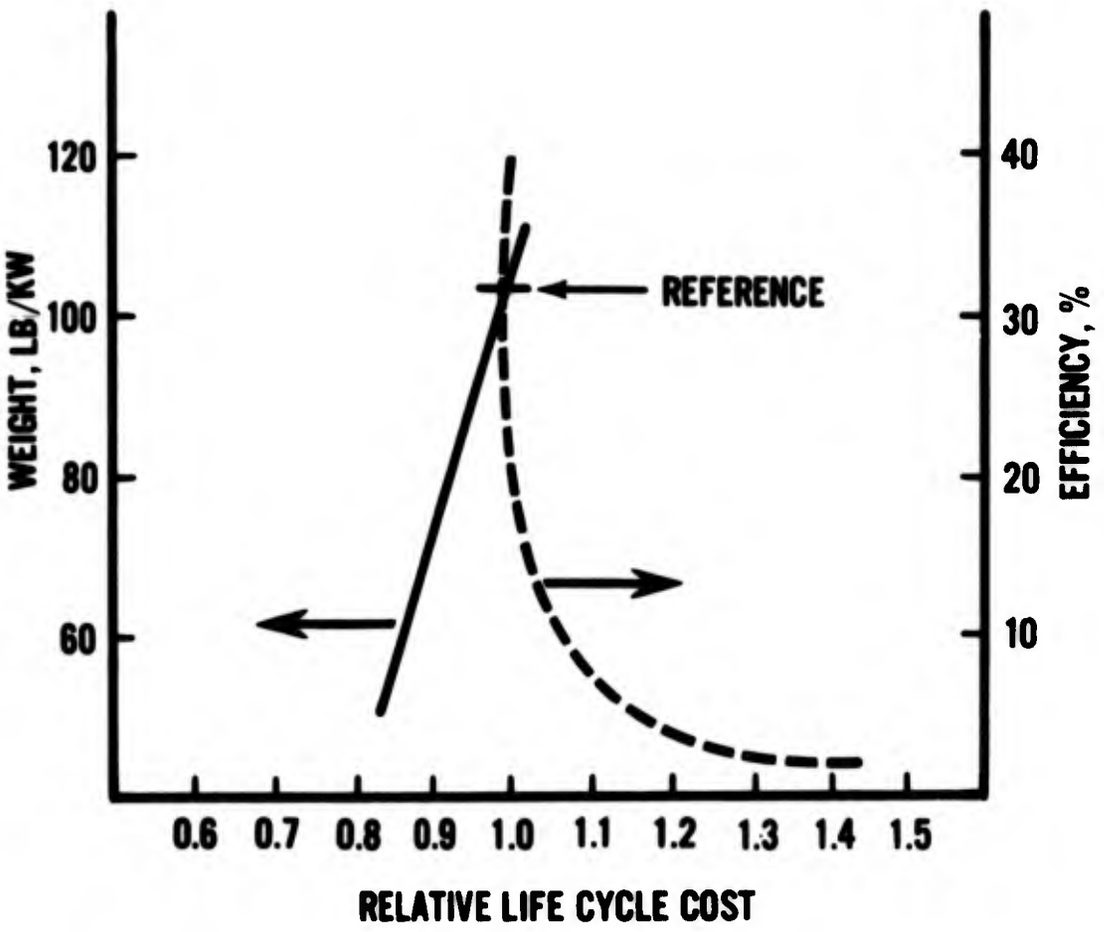


FIGURE 8 - Relative Life Cycle Cost Versus Power Plant Weight and Efficiency.

AD 713 529

MATHEMATICAL MODEL FOR PROJECTILE BODY

LEONARD M. GOLD AND FRANK SHINALY
FRANKFORD ARSENAL
PHILADELPHIA, PENNSYLVANIA

Computer simulation of small arms rifle/ ammunition systems appears to be one of the most promising techniques for analyzing a variety of problems. Simulation is utilized in practice by forming within the computer a structure analogous to the real system. Modules consisting of mathematical expressions, recursive processes, and analytical or statistical procedures make up the system. The projectile module is obtained by applying appropriate stress analysis approaches amenable to programming in an efficient form for accepting specifications reflecting direct relationships between components.

The design and analysis of small arms projectile-rifle bore systems is presently limited by the lack of mathematical techniques available for the solution of three dimensional elastic equations of motion. This deficiency has led to the use of experimentally based intuitive solutions and of quasi-dynamic governing equations.

The governing differential equations can be obtained by either an elasticity approach (1) using

$$\operatorname{div} \bar{\bar{T}} + \bar{F} = \rho \ddot{\bar{u}} \tag{1}$$

as the three equations of motion and

$$\bar{\bar{T}} = 2G\bar{\bar{E}} + \bar{\bar{\alpha}}\lambda\epsilon \tag{2}$$

to specify stress boundary conditions, or by using the governing equations of shell theory such as presented by Flügge (2), Love and others.

Present techniques for modeling consist of breaking the projectile into a number of sections (see Figure 1a.) consisting of thin circular cylinders or cones (3,4). Although this technique

is conservative, no small arms type projectile at present consists of thin wall sections or truly cylindrical or conical segments of uniform thickness (see Figure 1b.). The model shown in Figure 1b. is applied by assuming an average thickness and an average radius for each section and calculating the discontinuity stresses to determine the complete boundary stress conditions on each segment.

The forces acting on the projectile body are a resultant of projectile and filler inertia, rotation and engraving of the rotating band due to rifling, aerodynamic forces, filler charging pressure, residual stresses and propellant gas pressure. These loadings are applied in the governing shell equations as boundary conditions or meridional surface loads. Present analytical capabilities exclude consideration of transient effects; however, experimental determination of the transient effects is in progress and the results will be used as correction factors in determining maximum stress levels.

To improve the model the projectile must be mathematically described in a manner which more closely represents a true projectile geometry. Figure 1c. shows that a better representation can be obtained by having the major portion of the projectile cut into segments which have curvature and variable thickness determined from proposed projectile dimensions and geometrical or loading restraints.

The distance from the axis of symmetry to the midsurface of the shell body is given as

$$r = r(s) \tag{3}$$

where s is the meridional distance along the midsurface and

$$t = t(s) \tag{4}$$

is the thickness of the shell wall as a function of the meridional distance. The model is a shell of revolution with dependent variables and loadings which are functions of the meridional distance only. This assumes that loads which vary in the circumferential direction, such as those produced by a rotating band, are uniformly transmitted to the projectile body.

The geometry of a typical element from a shell of revolution is shown in Figure 2. The midsurface forces and moments are shown in Figure 3.

Taking $\partial/\partial\theta = 0$, $Q_\theta = 0$ in Flügge's governing equations for bending of a shell of revolution and using a comma to indicate differentiation with respect to the variable following the comma,

these governing equations take the form

$$(rN_s)_{,s} - N_\theta \cos\phi - \frac{1}{r_1}(rQ_s) = -rp_s \qquad (a)$$

$$(rN_{s\theta})_{,s} + N_{\theta s}\cos\phi = -rp_\theta \qquad (b)$$

$$N_\theta \sin\phi + \frac{1}{r_1}(rN_s) + (rQ_s)_{,s} = rp_\xi \qquad (c)$$

$$(rM_s)_{,s} - M_\theta \cos\phi = rQ_s \qquad (d) \qquad (5)$$

$$(rM_{s\theta})_{,s} + M_{\theta s}\cos\phi = 0 \qquad (e)$$

$$rN_{\theta s} - rN_{s\theta} - M_{\theta s}\sin\phi + \frac{1}{r_1}(rM_{s\theta}) = 0 \qquad (f)$$

Using the relationships between the displacements and the shell forces and moments, three governing equations can be obtained in terms of the three displacement components u, v and w. Equilibrium in the ξ direction gives

$$\frac{d^4w}{ds^4}F_0 + \frac{d^3w}{ds^3}F_1 + \frac{d^2w}{ds^2}F_2 + \frac{dw}{ds}F_3 + wF_4 + \frac{d^3v}{ds^3}G_0 + \frac{d^2v}{ds^2}G_1 + \frac{dv}{ds}G_2 + vG_3 = p\left(\frac{1-\nu^2}{E}\right) \qquad (6)$$

Equilibrium in the s direction gives

$$\frac{d^3w}{ds^3}F_5 + \frac{d^2w}{ds^2}F_6 + \frac{dw}{ds}F_7 + wF_8 + \frac{d^2v}{ds^2}G_4 + \frac{dv}{ds}G_5 + vG_6 = -p\left(\frac{1-\nu^2}{E}\right) \qquad (7)$$

The coefficients F_i and G_i are functions of the distance from the axis of symmetry to the meridional line, r(s), and its derivatives and the thickness of the shell t(s) and its derivatives. The expressions for F_i and G_i are given in the appendix.

Finally, taking $p_\theta = 0$, equilibrium in the θ direction is given as

$$\frac{d}{ds}(UH) = 0 \qquad (8)$$

where

$$U = \frac{du}{ds} - \frac{\cos\phi}{r} u$$

and

$$H = \frac{t^3}{12}\left(\frac{3r\sin\phi}{r_1} - \frac{r^2}{r_1^2} - 3\sin^2\phi\right) - r^2 t$$

Since equation (8) is an exact differential it may be integrated directly giving

$$U = \frac{\alpha}{H}\,; \qquad \alpha = \text{constant}$$

From the stress strain relationships given by Flügge, the shear stress $\tau_{s\theta}$ can be written as

$$\tau_{s\theta} = \frac{E}{2(1+\nu)}\left(\frac{r + \xi\sin\phi}{r[1 + \xi/r_1]}\right)\left(\frac{du}{ds} - \frac{\cos\phi}{r} u\right)$$

which can be rewritten as

$$\tau_{s\theta} = \left(\frac{E[r + \xi\sin\phi]}{2r[1+\nu][1 + \xi/r_1]}\right)\left(\frac{\alpha}{H}\right) \tag{9}$$

and α can be evaluated from the given torsional loading. If p_θ is not zero, equation (9) becomes

$$\tau_{s\theta} = \left(\frac{r + \xi\sin\phi}{Hr[1 + \xi/r_1]}\right)\int_{s_o}^{s} r^2 p_\theta ds \tag{10}$$

With a solution to equations (6) and (7) and evaluation of equation (9) or equation (10) the solution is complete for axisymmetric loading. A final point in the analysis is to determine if the material remains elastic during loading. After the equilibrium equations have been solved for all imposed loads, the stresses at each point are determined by superposition. The von Mises yield criterion is given by Hill (5) in the form

$$(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2 = 2\sigma_o^2 \tag{11}$$

where σ_1, σ_2 and σ_3 are the calculated principal stresses and σ_o is the yield strength in uniaxial tension. As long as the left side of equation (11) remains less than $2\sigma_o^2$, the material is considered to remain elastic.

It is assumed in applying the procedure that geometrical constraints and system specifications will provide sufficient information for determining the equation for the midsurface of each segment and the variation of thickness in the meridional direction for each segment. In addition it may be necessary to add a solution for the base plate problem if the base of the projectile cannot be simply described as a shell of revolution to include the effect of all setback forces and to determine all loads and moments applied to the first shell segment. The rotation of stress equations should be used to insure that the loads are applied to the first shell segment in the proper orientation.

The setback forces, rotational forces and pressure loadings must be determined as functions of meridional distance and specified in terms of p_ξ, p_s and p_θ. Loads which are not specified in terms of p_ξ, p_s and p_θ are utilized as boundary conditions. The differential equations (6) and (7) are then solved using the Runge-Kutta technique considering boundary conditions on one boundary for each solution run of the governing equations. At the end of each segment, existing shell forces and moments must be determined to find the boundary conditions for the next segment. Stresses may be utilized in place of forces and moments. In either procedure proper orientation of the loading must be determined for application to the next segment.

The above procedure is continued through all the shell segments and repeated until all loadings have been applied. With the displacements and their derivatives determined at all points of computation, the stresses from each loading may be determined and summed in each orientation. The resulting stresses are then utilized to determine is yielding of the projectile material has occurred. If no yielding has occurred the present model is sufficient for the stress analysis of a projectile subjected to quasi-dynamic firing conditions. If yielding har occurred the projectile design must be modified (geometry or material) to relieve the stress levels and prevent yielding.

It is recognized that the given analysis does not incorporate transient dynamic effects; however, the transient effects are being studied as a separate problem with the eventual goal of incorporation into the total stress analysis. This does not render the above analysis invalid due to the duration of the loading being on the order of fifty (50) times the period of consideration for the initial stress wave to propagate along the length of a small projectile such as the 20 mm shell. A further refinement of the analysis provides for non-axisymmetric loadings and geometry.

An experimental program has been initiated for the purpose of verification of the existing analytical techniques. Actual

projectiles and scaled models of projectiles are subjected to firing and/or simulated firing conditions and the recorded strains are compared with those determined by the existing computer analysis.

The program is presently divided into three phases. The first phase, which is presently in operation, is the application of static compression and internal pressure to scaled projectile models. The second phase of the program is the simulation of firing conditions on the actual projectiles and scaled models. This will be accomplished by the application of a pressure pulse to the specimen which simulates the firing pressure history. The third phase of the program is the measurement of strains during actual firings.

Phase one of the test program has been initiated with the application of compressive loads and internal pressure to a 20 mm shell scaled to 81 mm size Figure 4. shows the projectile model used. The projectile geometry has been idealized for initial verification. The dimensions of the model were used as input data for the computer program and the program was run for the same loadings that were applied to the specimen. Figure 5. shows a comparison of strains determined from the computer stress analysis with the recorded strains during compression tests for various axial locations on the outer surface of the specimen. The tests were run on a closed loop electro-hydraulic test machine and strains were monitored with 120 ohm foil strain gages bonded to the specimen. The computer analysis reasonably predicted strain levels in the region of the model between the base and the rotating band seat. A large discrepancy exists near the nose of the model which is probably due to the modification of the quasi-dynamic computer program to handle static loadings; however, the computer results are conservative in that portion of the projectile. Figure 6. shows a similar comparison for the internal pressure tests. It is clear from the comparisons that the computer results could lead to overdesign throughout the projectile. Phase two and phase three testing will determine how significant is the conservatism in the existing analysis.

The scope of the information presented here may be summarized in the following manner. A technique has been presented which will effectively improve the modeling of a small projectile and the analysis of stresses. Although the applied theory generally refers to thin shells, no assumptions have been made concerning the order of magnitude of shell thickness to radius and the ratios have been retained in the equations. The basic technique consists of breaking the projectile into segments which can be described as shells of revolution with varying curvature and thickness and then solving the governing differential equations with the Runge-Kutta technique adapted to the boundary value problem.

The existing computer analysis models the small projectile

as a series of cylindrical shell segments or conical shell segments. While this technique may be appropriate for large projectiles which can more accurately be called thin-walled, it is questionable for small projectiles which are relatively thick-walled. Variation of curvature and thickness are accounted for in an averaging procedure. There is no doubt that conservative design data can be obtained utilizing the present analysis; however, a better modeling of the projectile can lead to savings in design time and a smaller probability of malfunction or material failure. The proposed analysis will also readily enable the systems engineer to include the projectile as part of a system analysis and determine the interaction between projectile and weapon.

From a quasi-dynamic point of view, the general technique can be applied to the analysis of certain phases of cartridge case loading and the analysis of stresses in a rifle barrel or mortar tube as well as the stress analysis of small and large projectiles. The firing stresses in a bullet jacket can also be obtained using the described procedure.

It is expected that the proposed analysis presented here and the incorporation of the dynamic transient analysis will lead to a simplified and more efficient projectile development program as well as more meaningful systems simulations.

APPENDIX

The listing below contains the coefficients for the governing differential equations.

$$\psi = \frac{1}{r_1} - \frac{\sin\phi}{r}$$

$$F_0 = \frac{t^3}{12}$$

$$F_1 = \frac{1}{6r}\frac{d}{ds}(rt^3)$$

$$F_2 = \frac{1}{r}\left\{\frac{d}{ds}\left(\frac{t^3\cos\phi}{12} + \frac{d}{ds}\left[\frac{rt^3}{12}\right]\right) + \frac{t^3}{12}\left(\frac{2\psi}{r_1} + \frac{\cos^2\phi}{r}\right)\right\}$$

$$F_3 = \frac{1}{r}\frac{d}{ds}\left\{\frac{t^3}{12}\left(\frac{2r}{r_1}\psi - \frac{\cos^2\phi}{r}\right) + \frac{d}{ds}\left(\frac{\nu t^3\cos\phi}{12}\right)\right\}$$

$$F_4 = \frac{1}{r}\frac{d}{ds}\left\{\frac{t^3\psi\sin\phi\cos\phi}{12r} + \frac{d}{ds}\left(\frac{t^3 r}{12 r_1}\psi\right)\right\} + \frac{t^3}{12}\left(\frac{\psi}{r_1^3} - \frac{\sin^3\phi}{r^3}\right) + t\left\{\frac{\sin\phi}{r}\left(\frac{\sin\phi}{r} + \frac{2\psi}{r_1}\right) + \frac{1}{r_1^2}\right\}$$

$$F_5 = -\frac{t^3 \sin\phi}{12r}$$

$$F_6 = \frac{1}{r}\frac{d}{ds}\left(\frac{t^3 r}{12}\psi\right) - \frac{1}{rr_1}\frac{d}{ds}\left(\frac{rt^3}{12}\right)$$

$$F_7 = -\frac{1}{rr_1}\frac{d}{ds}\left(\frac{\nu t^3 \cos\phi}{12}\right) + \frac{t^3\cos^2\phi}{12r^2}\left(\psi + \frac{1}{r_1}\right) + \frac{t}{rr_1}\left(r + \nu r_1 \sin\phi\right)$$

$$F_8 = \frac{1}{r}\frac{d}{ds}\left\{\frac{t^3 r}{12 r_1^2}\psi + tr\left(\frac{1}{r_1} + \frac{\nu\sin\phi}{r}\right)\right\} + \frac{1}{rr_1}\frac{d}{ds}\left(\frac{rt^3}{12r_1}\psi\right)$$

$$- \frac{t^3 \sin\phi\cos\phi}{12r^2}\psi^2 - \frac{t\cos\phi}{r}\left(\frac{\sin\phi}{r} + \frac{\nu}{r_1}\right)$$

$$G_0 = F_5$$

$$G_1 = -\frac{2}{r}\frac{d}{ds}\left(\frac{t^3\sin\phi}{12}\right) + \frac{t^3}{12r_1^2}\frac{dr_1}{ds}$$

$$G_2 = \frac{1}{r}\left\{\frac{d}{ds}\left[\frac{t^3}{12}\left(\frac{2r}{r_1^2}\frac{dr_1}{ds} - \frac{\nu\cos\phi}{r_1}\right) + \frac{d}{ds}\left(-\frac{t^3\sin\phi}{12}\right)\right]\right.$$

$$\left. - \frac{t^3\cos\phi}{12}\left[\frac{\nu}{r_1^2}\frac{dr_1}{ds} - \frac{\cos\phi}{r}\left(\psi + \frac{1}{r_1}\right)\right] + t\left[\left(\frac{r}{r_1} + \nu\sin\phi\right)\right]\right.$$

$$G_3 = \frac{1}{r}\left\{\frac{d}{ds}\left[-\frac{t^3\cos\phi}{12}\left\{\frac{\nu}{r_1^2}\frac{dr_1}{ds} - \frac{\cos\phi}{r}\left(\psi + \frac{1}{r_1}\right)\right\}\right.\right.$$

$$\left.\left. + \frac{d}{ds}\left\{\frac{t^3}{12}\left(\frac{r}{r_1^2}\frac{dr_1}{ds} - \frac{\nu\cos\phi}{r_1}\right)\right\}\right] + \frac{t^3}{12}\psi\left(\psi\frac{\sin\phi\cos\phi}{r}\right.\right.$$

$$\left.\left. + \frac{r}{r_1^3}\frac{dr_1}{ds}\right) + t\cos\phi\left(\frac{\sin\phi}{r} + \frac{\nu}{r_1}\right)\right\}$$

$$G_4 = t + \frac{t^3\sin\phi}{12rr_1}$$

$$G_5 = \frac{1}{r}\frac{d}{ds}(rt) + \frac{1}{rr_1}\frac{d}{ds}\left(\frac{t^3\sin\phi}{12}\right) + \frac{t^3}{12r_1^2}\frac{dr_1}{ds}\left(\psi - \frac{1}{r_1}\right)$$

$$G_6 = \frac{1}{r}\frac{d}{ds}\left(\gamma t\cos\phi + \frac{t^3 r}{12 r_1^2}\frac{dr_1}{ds}\gamma\right) - \frac{1}{rr_1}\frac{d}{ds}\left(\frac{t^3}{12 r_1}\left[\frac{r}{r_1}\frac{dr_1}{ds} - \nu\cos\phi\right]\right)$$

$$- \frac{t^3\cos\phi}{12r}\frac{\cos\phi}{r}\left(\frac{1}{r_1^2} + \frac{\nu}{r_1} + \nu^2\right) - \frac{\nu}{r_1^3}\frac{dr_1}{ds} - \frac{t\cos^2\phi}{r^2}$$

NOMENCLATURE

$\bar{\bar{T}}$ - Stress Tensor

$\bar{\bar{E}}$ - Strain Tensor

$\bar{\bar{\delta}}$ - Identity Dyadic

$\bar{u}$ - Displacement Vector

$\bar{F}$ - Body Force

λ - Lame' Parameter

ϵ - div $\bar{u}$

ρ - Mass Density

ν - Poisson's ratio

s - Meridional distance

θ - Circumferential coordinate

ξ - Distance along normal to shell midsurface

N_s - Normal force (s direction)

N_θ - Normal force (θ direction)

M_s - Bending moment (s plane)

M_θ - Bending moment (θ plane)

$N_{s\theta}$ - Shear force (s plane, θ direction)

$N_{\theta s}$ - Shear force (θ plane, s direction)

$M_{s\theta}$ - Twisting moment (s plane)

$M_{\theta s}$ - Twisting moment (θ plane)

Q_s - Shear force (s plane, ξ direction)

Q_θ - Shear force (θ plane, ξ direction)

p_ξ - ξ component of unit area load

p_θ - θ component of unit area load

p_s - s component of unit area load

r_1 - Principal radius of curvature in plane containing axis of symmetry

ϕ - Angle between r_1 and axis of symmetry

u - Displacement in the θ direction

v - Displacement in the s direction

w - Displacement in the ξ direction

REFERENCES

(1) Chou, P. C.; and Pagano, N. J.; Elasticity. Princeton: D. Van Nostrand Company, Inc., 1967.

(2) Flügge, W.; Stresses in Shells. Berlin/Göttingen/Heidelberg Springer-Verlag OHG. 1960.

(3) Larson, C. W.; Knutelsky, B. G.; and Salomon, R. G.; "Stress and Deflection Equations for Analyzing Shell with Axially Symmetric Geometry and Loading-Theory and Computer Program". Tech. Rep. 3762 Picatinny Arsenal, Dover, N.J., 1969.

(4) U. S. Army Material Command 1964. Engineering Design Handbook, Ammunition Series, Section 4, Design for Projection. AMCP 706-247. U. S. Army Material Command, Washington, D.C.

(5) Hill, R.; Plasticity. 1st ed. London: Oxford At The Clarendon Press, 1950.

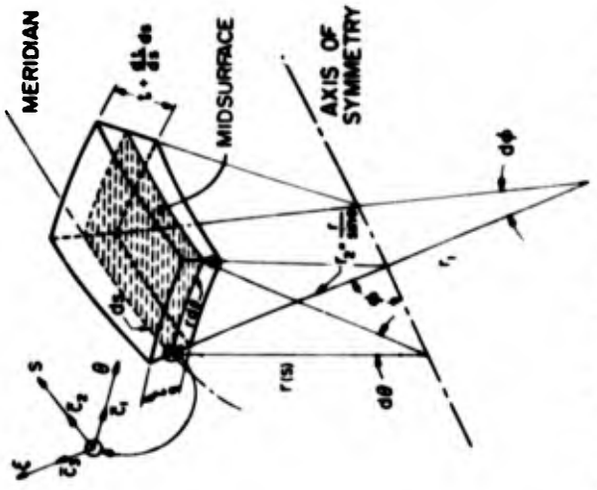


FIGURE 2. TYPICAL ELEMENT OF A SHELL OF REVOLUTION

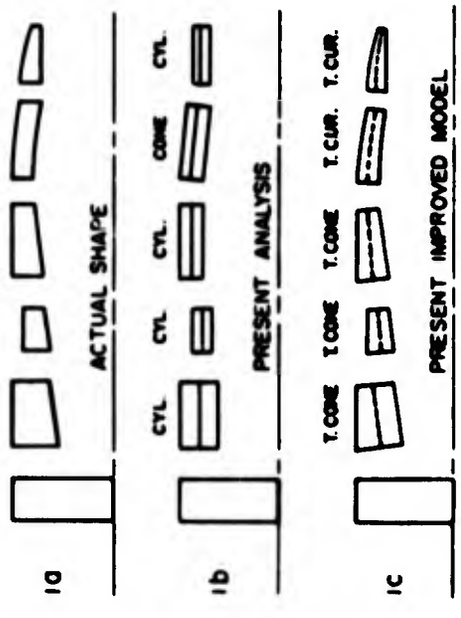


FIGURE 1. BREAKDOWN FOR STRESS ANALYSIS OF 20MM PROJECTILE

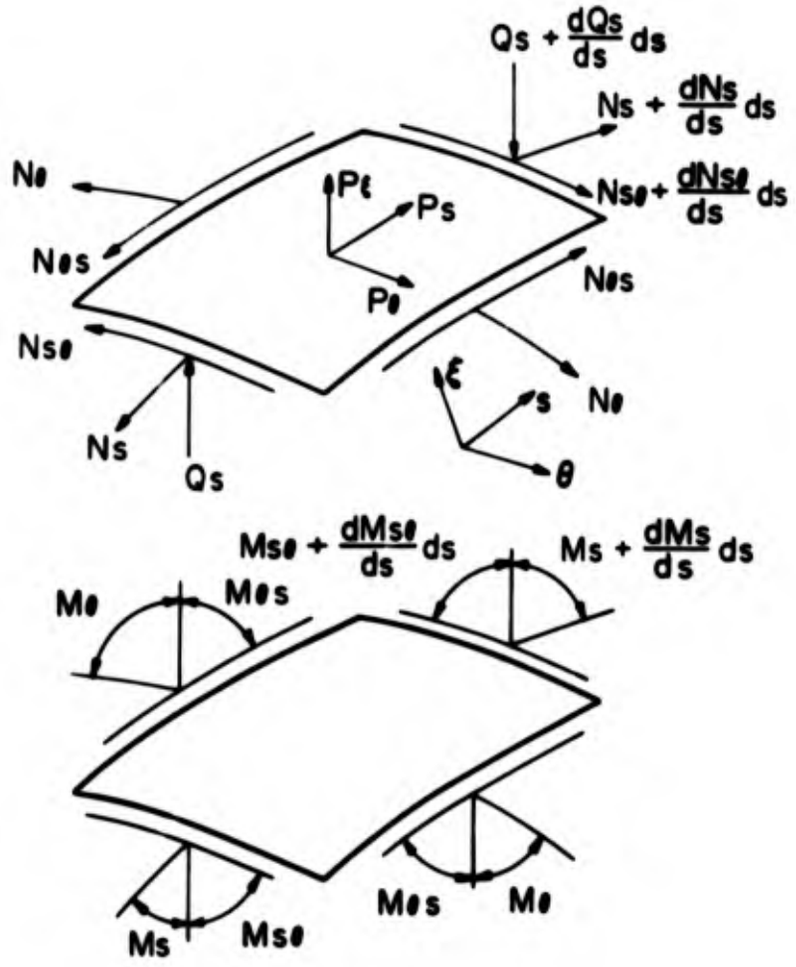


FIGURE 3. SHELL FORCES AND MOMENTS ON TYPICAL MIDSURFACE ELEMENT

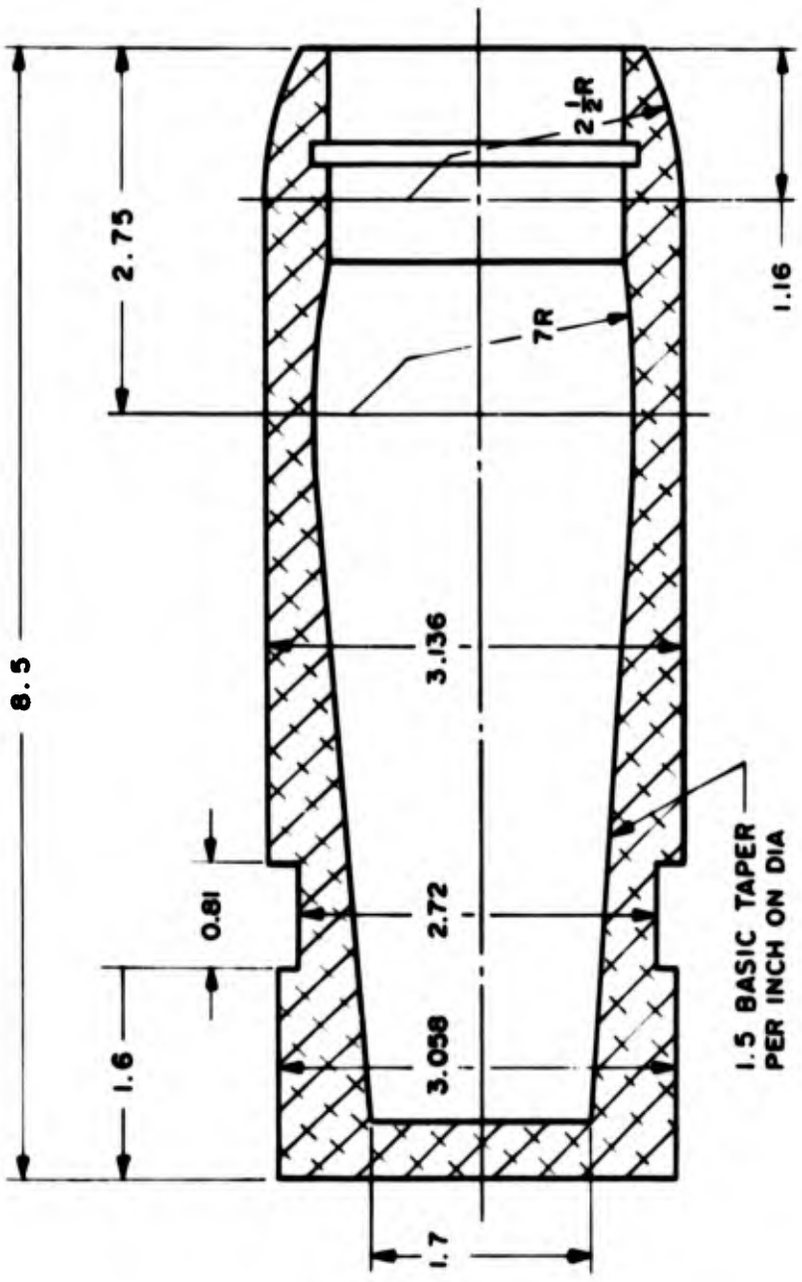


FIGURE 4. EXPERIMENTAL TEST MODEL

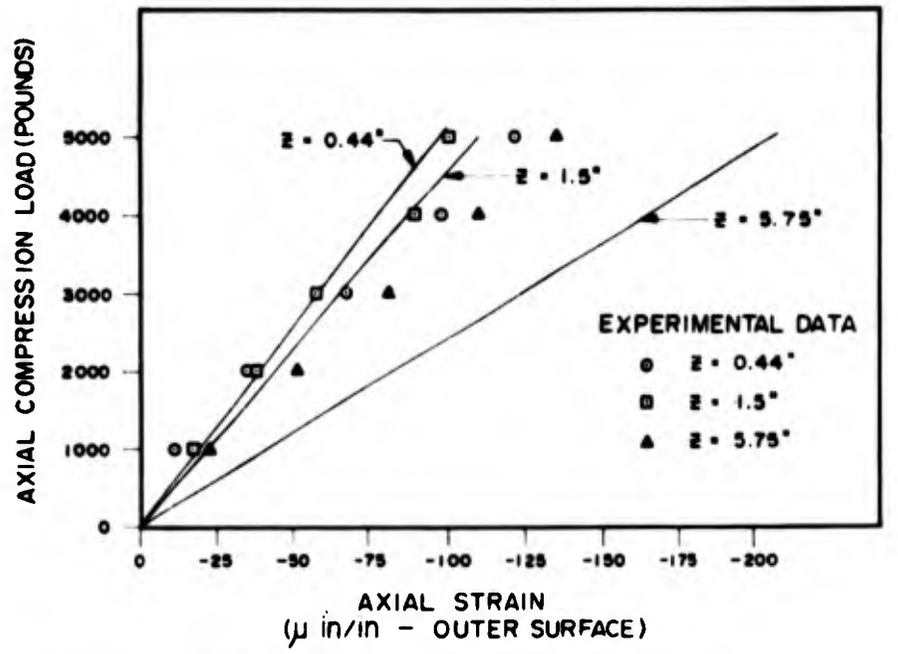


FIGURE 5. EXPERIMENTAL AND COMPUTER RESULTS
COMPRESSION TEST

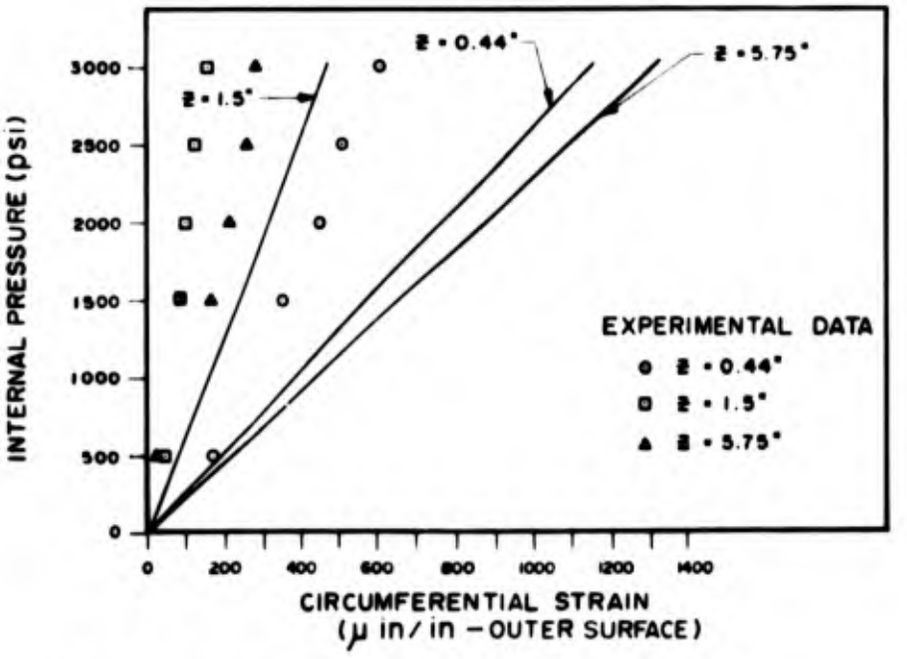


FIGURE 6. EXPERIMENTAL AND COMPUTER RESULTS
PRESSURE TEST

AD 713 530

TACTICAL IMPLICATIONS OF THE PHYSIOLOGICAL STRESS IMPOSED BY CHEMICAL PROTECTIVE CLOTHING SYSTEMS

RALPH F. GOLDMAN, Ph.D.
US ARMY RESEARCH INSTITUTE OF ENVIRONMENTAL MEDICINE
NATICK, MASSACHUSETTS 01760

At the beginning of WW I, Fritz Haber (later famous for the "Haber Process"), a German chemist and subordinate of Prof. Nernst at the University of Berlin, initiated a program to utilize chemicals as anti-personnel weapons. Despite discouragement from the German General Staff, who considered the idea unworkable, by April of 1915 about 170 tons of Chlorine gas were emplaced in 5,700 cylinders along a 3½ mile front, opposite a junction where British lines joined those of French Algerian troops. Early on 22 April, Allied troops noted a green cloud being blown toward them. Within a few hours, this initial gas attack had killed 5,000 and gassed 15,000 Allied troops, while the Germans had captured 6,000 men and completely breached the Allied lines at this point. Fortunately for the Allies, the German army was completely unprepared for such an effect and could not exploit the breach. The Allies immediately appointed a French Army battalion surgeon, Andre Mayer, to organize defensive measures. His invention of the first military gas mask initiated the 55 years of collaboration on protective equipment between physiologists and chemical warfare experts which led to this report.

Although gas warfare was widely used by both sides in WW I, it caused less than 1.4% of the total fatalities and only about 5.7% of the casualties; however, while only about 4% of the U.K., French and German casualties were from chemical agents, this figure rose to almost 12% for Russian troops, who lacked protective equipment. Protection was primarily provided by respirators and normal clothing, although special decontamination troops were issued two layer cotton uniforms whose outer layer had been impregnated with 45% rosin in 55% rosin oil; worn with oiled, cotton gloves, this system provided about 40 minutes of chemical protection. The users of this early protective equipment had physiologic problems and in December 1918, the following guidance appeared in AEF #1433 "Defense Against Gas" over General Pershing's signature:

a) troops need "practice in wearing the respirator for longer periods"..."6-8 hours of wartime may be necessary."

b) "thorough training and drill of troops in use of protective appliances to enable them to adjust them and to perform normal duty while wearing them" is required.

In 1922, the U.S. General Staff restricted the Chemical Warfare Service to improving defensive techniques and by 1924, research had developed an impregnite for clothing, "CC-2", a chloramide deposited in a chlorinated parrafin binder. A Herringbone Twill uniform, impregnated with CC-2 was accepted in 1928 and remained as the Standard Chemical Protective uniform until 1943, when, because of poor storage life, "CC-3" was produced. Gas mask technology had been preserved and advanced at Edgewood Arsenal, Md., although from 1927 through 1938 less than 20,000 masks per year were produced. A rubberized hood (M-5) was developed to be worn over the mask for protection against bacterial agents. A gas mask with the M-5 hood and with long cotton gloves, with impregnated ankle length underwear and sox worn under the impregnated chemical protective uniform, provided about 90% protection against bacterial agents. In the late 1930s, China had accused Japan of using chemical agents and, after reassuring Russia of Allied retaliation if the Germans initiated gas warfare, by the close of 1942 development of protective clothing was given a higher priority; the U.S. Army was to be provided with improved protective equipment, including CC-3 impregnated clothing, by June 1943. Despite this official assessment of the threat of chemical warfare, after two years of training during WW II, 30% of all U.S. units were judged deficient in defense against chemical warfare. Chemical protective clothing was worn by various units during training and by troops during the Normandy invasion. Some heat exhaustion probably occurred in troops wearing chemical protective clothing during WW II but there are few reports of it. In the summer of 1944, the Allies became alerted to the potential danger of low level aerial spray attacks and by 1945 were aware of the German development of nerve gases, but this threat ended with the war.

During the Berlin crisis of 1958, awareness of Russian capability reinitiated concern about U.S. chemical protective systems. As a result of reported difficulties incurred in training with protective clothing, field trials were scheduled at Camp Pickett, Virginia, to identify whether or not the reported problems were physiological and to compare the then Standard A, 2 layer impregnated chemical protective uniform with a new system in which a button in liner, impregnated with CC-3, was worn as part of a multiple layer system which provided wide range climatic and other protection by appropriate selection of layers. The Camp Pickett trial with troops wearing these systems began with a road march on a hot, humid morning in 1959. It ended 30 minutes later when, after the early collapse of three heat casualties, it became obvious the trials could not be completed and that there was a high risk of heat stroke, which carries a 50% fatality incidence. This paper details the subsequent laboratory studies, delineates the physiological problems, describes subsequent, successful field evaluations and suggests some of the tactical limitations imposed by wearing current chemical protective systems.

METHODS

A multidisciplinary approach has been evolved to assess the interactions between the environment, the uniform worn, the man and his military mission. Studies are conducted at five different levels of analysis, with each level providing information which can be related to the others, as follows: 1) the physical heat transfer characteristics of the uniform materials are measured using classical heated flat plate theory and also a unique "sweating" flat plate; 2) complete protective clothing ensembles, with and without such additional items as gas masks, hoods, gloves, helmets or packs, are evaluated on a "sweating" copper manikin for the heat transfer characteristics of the uniform system; the values obtained are used in biophysical calculations in a programmed computer model to predict the soldier's tolerance limits; 3) carefully controlled physiological trials are carried out in climatic chambers with military volunteer subjects dressed in these uniforms, to validate or refine the computer predicted tolerance limits; 4) controlled, small scale studies are conducted in the field, with military units wearing specified clothing systems and carrying specified loads under conditions of environment, terrain and march rate where physiological problems are anticipated based upon experience in climatic chamber trials; 5) studies with these clothing systems are carried out, collaboratively, during field maneuvers scheduled by CDC or others. Specific details of the methodology for the laboratory studies (i.e. physical plate material studies, biophysical copper man evaluations and tolerance predictions and physiological chamber validation experiments) are presented below. The methods used in field studies are adapted for each problem, are therefore difficult to generalize and will be included in discussing the results.

The "clo" unit:

Some years ago physiologists working in the field of clothing and the associated heat transfer from a man, developed a technique to determine how much heat would pass through a garment by thermal radiation and convection from the skin (3). The difference between a man's skin temperature and the ambient temperature was taken as a gradient across which, to avoid a change in body temperature, he had to eliminate the difference between his metabolic heat production and the heat he could lose by evaporation of sweat from his skin or of water from his lungs. The non-evaporative component was assumed to go through the clothing by normal radiation and convection heat transfer. They then defined the insulation of a clothing system, plus the overlying still air layer, in terms of a "clo" unit and derived a system of units such that the dry heat transfer (i.e. convective plus radiant) per unit gradient from skin to ambient temperature (^{o}C) would be 5½ kilocalories per meter square of surface per hour, per clo; that is:

$$H_{Dry} = \frac{5.55 \text{ kcal } m^2 \text{ hr}}{\text{clo}} \text{ per } {}^{o}C \qquad \text{EQUATION 1}$$

This physical equation states that heat flow equals the driving force, in this case a temperature gradient, divided by the resistance, expressed here in "clo". This basic approach for radiation-convection

heat loss yields a quantitative assessment of how good a given uniform is for a resting man in cold weather since radiation and convection are major avenues of heat loss in the cold. However, for a working man in the cold, evaporation of sweat becomes an important avenue of heat loss. Furthermore, radiation and convection heat loss decrease with increasing ambient temperature while evaporative cooling rises. Thus, the "clo" value alone is insufficient in the heat.

The permeability index, "i_m":

A similar form of equation can be used to predict evaporative heat transfer:

$$H_{Evap.} = \frac{5.55\ S\ (p_s - p_a)}{clo} \qquad \text{EQUATION 2}$$

Eq.(2) is a form of the psychrometric (i.e. slung) wet bulb thermometer equation, where the "clo" value is the insulation of the air layer around the thermometer. The gradient for evaporative transfer is the difference between the vapor pressure at the surface (p_s) and the ambient vapor pressure (p_a) in millimeters of mercury (mm Hg). Using the slope (S) of the wet bulb lines on a psychrometric chart, which is about 2°C per mm Hg, a vapor pressure difference can be converted to an equivalent temperature gradient. One can then determine the evaporative heat loss from a square meter of surface with a given water vapor pressure; e.g. at 35°C (the skin temperature of a sweating man) there would be a vapor pressure of 42 mm Hg at the skin and the gradient will thus be 42 mm minus the ambient air vapor pressure. The late Dr. Alan Woodcock of this laboratory proposed that the evaporative heat transfer for a nude man, or for any clothing system, could be expressed as the ratio of the actual evaporative heat loss, as hindered by the clothing, to that of a wet bulb with equivalent "clo" insulation (15). He suggested expanding Eq. (2) to include this "permeability ratio index (i_m)" so that:

$$H_{Evap.} = \frac{5.55\ i_m\ S\ (p_s - p_a)}{clo} \qquad \text{EQUATION 3}$$

The index of evaporative loss (i_m) could range from 0, for a system with no evaporative transfer, to 1 for a system which had no more impedance to evaporative heat transfer than the usual wet bulb thermometer. The conventional wet bulb, of course, is a slung (i.e. rapidly moving) wet bulb where the still air barrier is greatly reduced. Since a soldier is surrounded by a relatively undisturbed air layer, "i_m" seldom approaches 1.0 for a man, but tends to be limited in still air to about 0.5.

Determination of "clo" and "i_m":

Figure 1A includes a diagram of the flat plate apparatus used in measurement of the "clo" insulation value. The apparatus consists of a test section (A), surrounded on all four sides by a guard section (B) with another guard section (C) beneath the entire upper plate (A+B). All three sections are instrumented with plate temperature sensors, heating elements and thermostats. The sample to be tested (D) is placed on the surface and the entire assembly is placed in a constant temperature cabinet. In operation, power to the guard sections is controlled so that their surface temperature is identical to

that of the test section. Thus there is no gradient for heat loss from the bottom or edges of the test section (A). After equilibrium is established, the power required by the test section equals the heat lost through the insulation and can be expressed as kcal/m^2-hr per degree of gradient from plate surface to ambient air temperature. This can be converted to the corresponding "clo" insulating value for the sample plus the adhering air layer, using Eq.(1). If a thin cotton "skin" is placed on the plate surface and a water level is maintained at the surface of some small holes drilled in sections A and B, then the "skin" wicks out enough water to maintain a constant saturated surface pressure. A constant ambient vapor pressure is maintained in the measuring chamber and power requirements measured just as for the dry plate. The permeability index value (i_m) can be determined for a given sample plus its adhering air layer by means of Eq.3. Figure 1A also shows the "sweating" flat plate and its water supply in the constant temperature and humidity chamber with its control and recorder panel. The flat plate determinations of clo and i_m are primarily of use in selection of the fabrics to be used in a clothing system. The effects on heat transfers of different weaves, perforations, different finishes or treatments, the effects and best arrangement of multiple layers, etc. all can be established using the sweating flat plate (1). Heated, dry and "sweating" cylinders have been developed to mimic the cylindrical shape of the body. These are useful in studying wind penetration through clothing, and effects of spacer materials, but there are factors of drape, fit, and shape which are difficult to simulate even on a cylinder. Also, a complete uniform is made up of a number of different components, protecting various parts of the body, so that evaluation of a complete clothing system requires a more sophisticated model than a cylinder (2).

The Copper Man:

The solution has been the development of life sized, heated copper manikins. Figure 1A shows a manikin with his "sweating" cotton skin. The heat provided to the manikin to maintain a constant skin temperature can be measured and the ambient temperature and vapor pressure of the test chamber can be controlled; skin and air temperature and vapor pressure are measured. Thus the radiant-convective heat loss and the evaporative heat loss caused by a given gradient of temperature and vapor pressure can be calculated for any clothing system worn. This technique has been in use for the last five years. Using the insulation and evaporative transfer indices, "clo" and "i_m", with some physiological knowledge, tolerance times can be predicted, for a given task, for men in the chambers and in the field. One solves the equation that heat stored by the body must be the difference between the heat produced at work, and that lost by evaporation and by radiation and/or convection through the clothing system. Since the average soldier has 1.8 m^2 of surface area, by estimating his skin temperature (T_s), total non-evaporative heat transfer can be calculated for any given ambient dry bulb temperature (T_a in ^oC) as:

$$H_{Dry} = 10\,(T_s - T_a)/clo \qquad \text{(kcal-man-hr)} \qquad \text{EQUATION 4}$$

One can similarly calculate the maximum evaporative heat transfer through clothing for any given ambient vapor pressure (mm Hg) as:

$$H_{Evap} = 20\ (42-P_a)\ i_m/clo \quad \text{(kcal-man-hr)} \quad \text{EQUATION 5}$$

where a 35°C skin temperature has been assumed for the sweating man. If heat production is known, after allowing for respiratory heat loss and any solar heat load gain, one can calculate whether the man can eliminate all the heat he is producing or whether some of it will be stored in his body. Indeed, using the specific heat of human tissue (0.83 kcal/kg-°C), it can be calculated that the body temperature of a 70 kg (154 lb) soldier will increase by 1°C (1.8°F) for each 58 kcal that must be stored (i.e. 0.83 x 70). This allows prediction of tolerance as the time to reach a given body temperature. A computer program has been devised which incorporates many of the significant physiological, physical and environmental factors involved in human heat transfer. If the appropriate values for clothing, environment and metabolic heat production are supplied, the model will predict the body temperature response of an individual under the chosen conditions. However, the predicted responses are frequently checked by actual environmental chamber exposures of men.

Physiological Chamber Trials:

Standard protocols have been developed for these trials; one requires 2 fifty-minute walks, separated by a ten-minute break, followed by a one-hour rest. Figure 1C shows subjects in an environmental chamber, during a rest period. The men are seated on benches placed on one of the large 4 man treadmills. Each subject is wearing a different garment since, as usual, a Latin square randomization of garment, day and subject was used. Each subject is wearing a rectal catheter, to measure deep body temperature, and a 3 point skin temperature harness. Two connecting cables from each man are led outside the chamber to the instrumentation shown in Fig. 1D. Each subject's rectal temperature is indicated continuously on one of the eight meters at the base of the master timer. Skin temperatures are recorded sequentially on the recorder, along with the rectal and chamber temperatures. Each point printed is simultaneously encoded and punched on the digital punch tape system shown at the right. This tape is used in computer data reduction and analysis and the agreement with predicted responses can be checked.

RESULTS AND DISCUSSION:

Values obtained for the fatigue and overgarment materials with the flat plate technique are present in Table I; it has so far been impossible to run the impregnated 2 layer system. The plate technique, which is a pure measure of fabric thermal properties (without flexibility or weight effects on drape)exhibits the greatest differences in thermal properties between materials. These large differences in material "clo" and "i_m" are considerably damped when comparing uniforms made from the materials, as seen in the copper manikin values also included in Table I. Thus, while flat plate values are useful in selection of materials for clothing development, it is difficult to predict uniform system characteristics solely on the basis of flat plate measurements.

TABLE I. Flat plate values of protective materials and the corresponding sweating copper manikin values for the Fatigue, the 2 layer CB (US2CB) and overgarment (USOG) systems.

Plate Materials Values:

	U.S.Stnd Fatigues	US2CB	USOG
clo	0.66	--	1.19
i_m	0.66	--	0.48
i_m/clo	1.00	--	0.41

Manikin Values w/ and w/o mask, hood and gloves:

	Fatigues		US2CB		USOG	
	w/o	w/	w/o	w/	w/o	w/
clo	1.33	1.48	1.50	1.87	1.60	1.90
i_m	0.50	0.38	0.40	0.38	0.42	0.36
i_m/clo	0.37	0.26	0.27	0.20	0.27	0.19
95°F 50% RH						
Loss *(kcal/hr)	150	108	112	86	111	80
Tol Time *(hrs)	1.3	.9	.9	.8	.9	.8

*prediction for worst possible case of still air, subject doing hard work (450 kcal/hr) and clothing not becoming wetted by sweat.

The copper manikin "clo" values reflect both material thickness and the trapped air insulation between the skin and clothing as a function of material stiffness, i.e. drape. Because the skin to air temperature gradient for non-evaporative heat loss is small in the heat, the important value is the i_m/clo; this indicates the obtainable fraction of the maximum sweat evaporative cooling possible in a given environment without wind, since these manikin values are still air determinations. Thus, a man wearing just the fatigue uniform can obtain about 37% of the maximum cooling possible (i_m/clo = 0.37), while with either of the two protective uniforms, only 27% is available.

Predicted heat loss (dry plus evaporative, Eq. 4 and 5) is included in Table I, with the estimated time to a 50% heat casualty level for troops doing hard work under conditions of 95°F, 50% relative humidity. These predictions are derived from chamber studies of the physiology of unclothed men, resting (10) or working (12) in the heat where body heat storage of about 80 kcal was enough to make a number of these volunteers stop; a heat storage of about 160 kcal resulted in heat exhaustion in 50% of those who continued while almost no one could tolerate 220 kcal. Hard work in a 95°F, 50% RH environment is so severe a combination that even wearing just the usual fatigue uniform, the predicted tolerance time is less than 1½ hours; in reality, the fatigue uniform would rapidly become sweat soaked and this, in combination with any natural ambient air motion, would result in a considerably longer tolerance time than this predicted value. The protective uniforms, however, can not readily wet out with sweat, and while the air permeability of the overgarment is high enough to allow additional cooling from any wind, the 2 layer uniform is relatively impermeable to wind.

Figure 2A presents the mean body heat storage (ΔS) for 7 volunteers walking at 3.5 mph in a 1965 study at 95°F, 50% RH (6). Five different uniform ensembles were worn as specified in the legend,

which indicates the corresponding clo and i_m/clo values; subjects wore underwear, socks and combat boots, but not packs, gas masks etc. Heat storage is almost identical when wearing the fatigue or tropical combat uniform, both of which have a copper manikin i_m/clo value of 0.34. Heat storage when the men wore either of these uniforms under the overgarment (i_m/clo = 0.22 and 0.23) is also nearly identical, while when the overgarment is worn alone (i_m/clo =0.27), the mean heat storage clearly is greater than that in the 0.34 i_m/clo uniforms and less than that with the 0.22 i_m/clo systems. The circled numbers represent the subjects remaining, as individual volunteers stopped; the general reliability of 80 kcal of heat storage as a critical lower level for subjective tolerance is well supported. It is also obvious that the relative ranking of physiologic stress suggested from the simple physical measurements with the sweating copper manikin, is valid. Figure 2B presents results of the latest chamber evaluation in which the thermal effects of the gas mask, hood and gloves were studied. Three clothing ensembles were studied: the utility fatigue uniform, the latest protective overgarment (recommended for standardization in March 1970) and the now Standard B, 2 layer impregnated protective system. The copper manikin values included in the legend, suggest that there has been little real progress in developing clothing systems with respect to the heat problem. However, the mean physiologic responses (8 Ss) of skin temperature ($\bar{T}_s$) and rectal temperature (T_r) do show the significant advantage afforded by the air permeability available in the overgarment; the skin and rectal temperatures of subjects wearing the overgarment with mask, hood and gloves are identical to their responses in the 2 layer system without these items.

Any clothing item which blocks passage of liquid or vapor from the outside to the skin surface must also block evaporative cooling from the skin. However, the thermal effect of the mask, hood and gloves per se can be seen to be negligible in the fatigue uniform, with other areas of the body apparently compensating for the loss of evaporative cooling from the hands and head; there is a significant thermal effect when these items are worn with the overgarment and a still greater effect with the 2 layer system since, with these two uniforms, little or no additional cooling can be obtained from other body areas to compensate. When men wore the fatigue uniform, adding the mask, hood and gloves significantly increased sweat production, and the men were able to evaporate enough of the extra sweat so that total evaporative cooling was increased. When either of the protective ensembles was worn, the production of additional sweat upon addition of the mask, hood and gloves was limited, perhaps because of a depression of sweating produced by skin wetness, as well as because the sweat production rates without mask, hood and gloves were already nearly maximal. There was about a 7% drop in sweat evaporation when the mask, hood and gloves were added to the overgarment and a 20% drop when these items were worn with the 2 layer protective system. Thus, while the copper man based predictions are useful in ranking clothing systems, they obviously should be supplemented by physiological chamber trials to assess the effects of body motion,

wind and the efficiency of sweating as shown above.

Men walking on treadmills under controlled conditions cannot indicate all the problems that will occur in troops attempting to perform their military missions in the field. While thermal effects resulting from the addition of packs, body armor, helmets, weapons etc. have been measured both on copper manikins and in physiologic chamber studies (5,9) the problems associated with military tasks other than marching, and with solar heat load, terrain and wind variation are impossible to reproduce indoors. Results of the initial chamber studies conducted after the aborted Camp Pickett field study, indicated that when wearing chemical protective uniforms, "soldiers should be heat-acclimatized and physically conditioned if they are to be required to do moderately heavy work for one hour" and that "to test the clothing out-of-doors where solar radiation must be considered, the temperature should be within a range from 70° to 90°F with humidity between 30 and 75%" (4). Accordingly, a controlled field trial of the 2 layer protective ensemble, an integrated (protective liner) ensemble and the (non-protective) utility fatigue uniform was conducted at Ft. Lee, Va. The results supported the chamber findings as to the advisability of heat acclimatizing individuals wearing protective clothing, and confirmed the finding that soldiers doing moderately heavy work in a fully encapsulated condition had a severe tolerance limitation at temperatures above approximately 75°F. The study concluded that "inability to dissipate the heat generated, limited tolerance time to about 30 minutes for hard work in the sun and this is the price that must be paid by the wearer for CBR protection as presently developed" (4).

When the CDC Experimentation Center (CDEC) at Fort Ord, Calif. was directed, in 1962, to investigate operational decrements that might occur with encapsulation, the USARIEM findings cited above were presented to indicate that a summer field study at Fort Ord would not involve long term exposure but could result in significant heat casualties during the first few hours. A continuing program was initiated by CDC to obtain physiologically significant meteorological measurements (especially the WBGT index calculated as 0.7 x naturally convected wet bulb temperature + 0.2 x black globe radiant temperature + 0.1 x shaded dry bulb temperature (13) and a carefully planned field research study was carried out by USARIEM, in collaboration with the CDEC, at Ft. Clayton, Panama in January 1963. The tactical exercises studied were chosen by the CDEC Infantry staff to represent the more vigorous tasks performed by combat arms elements and included: Rifle platoon in the Attack; 81 mm Mortar in the Attack; 4.2 inch Mortar in the Attack; Infantry squads in the Defense; 105 mm Howitzer in Fire Support; these exercises involved between 5 and 30 men and were replicated from 4 to 8 times each. Specialized tasks involving 2 to 4 men were also studied, including a Radar section, a Battalion Maintenance Section, a Rifle Company Headquarters, a Communications Section and an Engineer mine removal squad. Tactical scenarios, prepared by CDC experts in the appropriate combat arms, were adjusted, collaboratively with USARIEM experts, so that they would yield maximum information on physiological tolerance and yet

retain as much tactical realism as possible. Figure 2C represents the results of one day of the Rifle Platoon in the Attack problem; the tactical play and WBGT conditions are given across the top of the graph of the mean rectal temperature of the men remaining at each period. Frank heat exhaustion casualties are indicated at the time each occurred, by open circles, while closed circles represent men removed from action at a rectal temperature of 103°F by medical monitors. This level has been established as one which would result in a significant incidence of heat exhaustion but minimal risk of fatality. The 50% unit survival time for an encapsulated rifle platoon was 90 ± 10 minutes in this typical infantry scenario. However, Figure 2C reveals little rise in rectal temperature during the first 50 minutes of light activity, but that it began to rise sharply from the beginning of the approach march. The 50% casualty level occurred after 30 to 40 minutes of moderately heavy work, with massive heat casualties being incurred during the actual assault.

The overall experience during these field maneuvers is summarized in Figure 2D, where the predicted time to 50% heat casualties is presented as a function of work rate, for men a) wearing the 2 layer protective uniform as open as possible, with only the gas mask ("open suit") or b) wearing mask, hood and gloves, with the uniform completely buttoned up ("closed suit") (13). During hard work at these WBGTs the additional cooling possible in the open suit condition has little effect; the internal heat generated by the men is too great for these minor differences. Moderate work poses little problem at WBGTs below about 75°F, while light activity can be performed readily at WBGTs below about 85°F. It appears that simply adding 10 degrees to the WBGT as an allowance for complete encapsulation, permits use of the usual WBGT guidelines that apply to troops training in regular uniforms; TB Med 175 sets 85°F WBGT as the initial danger level for unseasoned troops during training in conventional uniforms.

Men carrying heavier loads or covering more ground during an exercise were shown to have a greater risk of becoming heat casualties, as were men in poorer physical condition, particularly fatter individuals. Apart from the better overall physical condition of men who had been acclimatized by working in the heat, in field studies with protective clothing, heat acclimatization *per se* appeared to be of little benefit. A major benefit of heat acclimatization results from evaporation of the extra sweat produced by acclimatized men; extra sweat production appears to be of limited benefit with these protective systems, which limit sweat evaporation. Further, unlike chamber or controlled field trials where the work rate is established by the investigator, troops on maneuvers adjust their work to compensate for poor physical condition or lack of acclimatization.

These findings were detailed in military medical journals (13), are being incorporated in the appropriate field manual (FM 21-40) and are available as a USARIEM technical report (11). A recent U.S. Marine Corps assessment of the latest protective uniforms, worn during an amphibious assault in a tropical environment (in which USARIEM collaborated) is essentially in agreement with these earlier findings (16); so are U.K. and Canadian studies as well as reported Russian

work (14). Guidance from USARIEM and adherence to the guidelines (11) derived from these studies on heat effects (4,5,13,16), has provided safety for engineering and service tests of these items by USATECOM at Ft Lee and Ft Benning, and has contributed to the successful completion of a number of classified studies on non-thermal problems experienced by troops wearing protective clothing during military operations. The conclusions of these studies on the physiological problems of soldiers attempting to conduct their military operations in a toxic environment, where protective encapsulation must be worn are presented below. However, it should be noted that the systematic method outlined for analyzing the interaction between a man, his environment, his work and his clothing was not developed solely to study chemical protective systems, but has had more general objectives and applications. It has been used to assess a wide variety of military problems involving clothing such as air crew uniforms (6), naval wet and dry diving suits (8), body armor (9) and thermal flash protective ensembles (5), as well as such civilian items as football uniforms (7), raincoats (6) and conventional clothing (8); it is currently being applied in studies on improving the human environment by heating and/or air conditioning.

CONCLUSIONS:

The tactical implications of wearing current standard chemical protective ensembles are clear. They apply as well to all foreseeable protective uniforms except those incorporating forced ventilation with filtered ambient or conditioned air. Even at low temperatures, there are obviously performance decrements for tasks requiring delicate work as a result of the loss of manual dexterity and tactile sensation wearing gloves. Vision in the mask, particularly downward vision, is limited producing problems when moving in jungles or through marked areas and in identification, while some optical sighting devices will require redesign to be used at all. Auditory as well as oral communication is also impeded by gas mask wear. These problems can be only partially overcome by training or redesign. A more serious non-thermal problem is the respiratory impedance of the gas mask; in each study, some of the troops report difficulty in breathing during hard work. This reoccurred in the latest field study, in addition to a new problem, seasickness primarily in those individuals wearing the gas mask during the amphibious landing. While the limited field data obtained (16), suggested that these problems may be both psychological and physiological, they can be reinvestigated in the laboratory.

Tolerance time for troops fighting in protective clothing even in temperate climates will be severely limited by heat stress at WBGT conditions above 75°F. The heat problems that will occur at WBGT above 75°F will have greatest impact on troops doing hard work. Infantry on an approach march or in an assault will have about 30 minutes of tolerance before a long rest break in the shade is necessary. Soldiers with unusually heavy loads, e.g. mortar men, radio telephone operators, recoiless riflemen or any individual carrying a load of more than 40% of his body weight, will seldom complete an approach march of 2 miles or more and will straggle far to the rear.

Such individuals, as well as those men in less fit physical condition (particularly if overweight or with a hangover) will have the greatest difficulty completing their mission and are more apt to be heat casualties. Unit leaders, who cover more ground in checking on the progress of their men, are also at greater risk unless transport is provided.

Conserving the soldier by transporting him, or at least as much of his load as possible, until he must go into action is far more important when protective clothing is worn than usual. Mechanized infantry will be far more effective than regular infantry. Mechanized armor and artillery crewmen may have some heat problems during sustained missions, e.g. when manhandling heavy ammunition at high rates of fire, but even then rotating the heavier jobs between crew members can help solve the problem. Engineer companies can also usually rotate personnel between the heavier tasks. The troops will suffer the most severe degredation of their tactical performance when fully encapsulated. For such units, the rest time required for recovery after work rapidly becomes longer than the time available for their military mission, as ambient temperature rises above 75°F WBGT (11); indeed fully encapsulated troops in full sunlight or in a desert environment will have heat problems even at rest.

The addition of a drinking device to the M17A1 mask can alleviate the problems of dehydration if drinking is vigorously encouraged by commanders during any prolonged encapsulation, so these problems have not been discussed. The risk of chemical attacks and of not having protection already in place before the onset of such an attack, has also been ignored in this presentation. It seems reasonable to conclude that in the heat, although the threat of chemical attack may force a unit leader to have his troops wear protective uniforms, donning mask, hood and gloves should be delayed until a chemical attack is imminent.

RECOMMENDATIONS:

Short of developing and standardizing ventilated chemical protective clothing, the only solutions in a hot climate are tactical. When possible:

a) avoid operations during daytime heat
b) maintain a high level of physical condition and eliminate the unfit troops from action
c) transport the soldier and his load by vehicle
d) rotate the heaviest tasks
e) avoid masking until the threat is imminent
f) assure adequate water intake by command control
g) allow 30 minute or longer rest periods, in the shade whenever possible

Above all, be aware that men wearing chemical protective clothing will have a significant performance decrement in the heat.

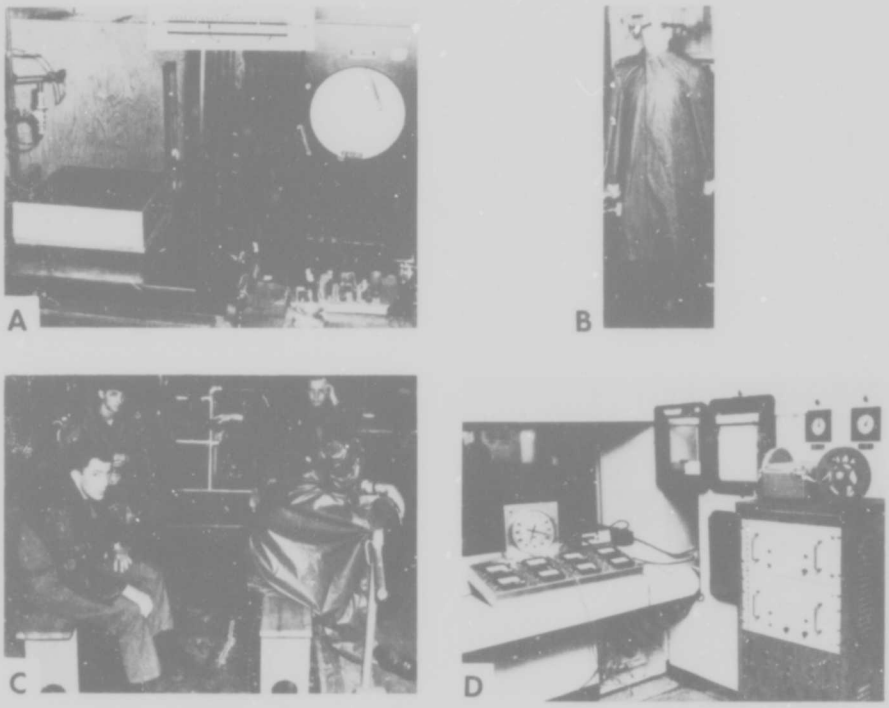


FIGURE 1. Laboratory Test Methodology

A. The "sweating" flat plate apparatus for assessing materials characteristics; the insert in the upper center, diagrams the test section (a), upper (b) and lower (c) guard sections and the position of the material (d) to be evaluated.

B. The sweating copper manikin used to assess non-evaporative ("clo") and evaporative ("i_m") heat transfer characteristics of a complete uniform ensemble.

C. Volunteer subjects, each with a different clothing ensemble, seated in the climatic chambers on benches placed on the 4-man treadmills during a rest break. The cables leading overhead from each subject carry rectal and skin temperature information.

D. The physiologic data collection instrumentation just outside the test chamber includes the individual meters for continuous safety monitoring of rectal temperature (just below the timer), a multipoint recording system for skin, rectal and environmental temperatures and a paper tape punch system to convert the data to a form for subsequent computer reduction and analysis.

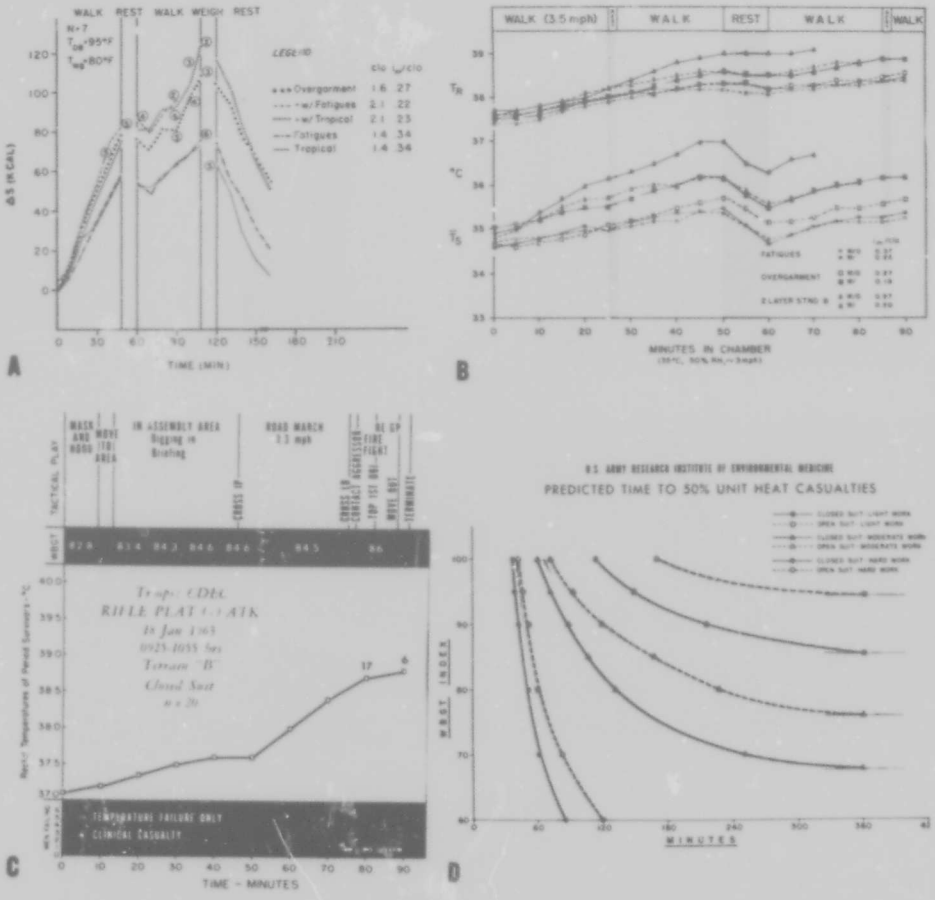


FIGURE 2. Typical physiological responses of men wearing protective clothing and predicted time limits for military operations.

A. Body heat storage of men in a chamber study. Copper manikin clo and i_m/clo values are given for each ensemble. Circled numbers represent men remaining at each point.

B. Mean rectal (T_r) and skin (T_s) temperatures of 8 subjects at 95°F, 50% RH wearing Fatigues or the chemical protective 2 layer or overgarment system, with or without mask, hood and gloves.

C. Mean T_r of soldiers during field maneuvers. Tactical play is detailed across the top, above the WBGT environmental conditions. Heat casualties (•) or men removed for unsafe T_r levels (o) are also shown.

D. Predicted tolerance for men at 3 levels of work, wearing protective clothing systems "OPEN" with just the gas mask, or "CLOSED", tightly buttoned up, with mask, hood and gloves.

REFERENCES

1. Fonseca, G.F. Moisture transfer through perforated impermeable foam insulations. Textile Res. J. 37: 1072-1076, 1967.
2. Fonseca, G.F. and J.R. Breckenridge. Wind penetration through fabric systems. Textile Res. J. 35: 95-103, 1965.
3. Gagge, A.P., A.C. Burton and H.C. Bazett. A practical system of units for the description of heat exchange of man with his environment. Science 94: 428-430, 1941.
4. Goldman, R.F. Tolerance time for work in the heat when wearing CBR protective clothing. Mil. Med. 128: 776-786, 1963.
5. Goldman, R.F. Physiological effects of wearing a rubberized net overgarment in the heat. Text. Res. J. 33: 764-766, 1963.
6. Goldman, R.F. Systematic Evaluation of Thermal Aspects of Air Crew Protective Systems. In: Proc. of the 24th annual meeting Aerospace Research & Development (AGARD) Brussels, Oct. 1967.
7. Goldman, R.F. The effects of football equipment on heat transfer. In: Physiological Aspects of Sports and Physical Fitness (ed. B. Balke). Amer. Coll. of Sports Medicine, 1968.
8. Goldman, R.F. Design and use of clothing. In: Proc. NATO Advanced Study Institute on Human Factors, Palermo, Oct. 1969.
9. Goldman, R.F. Physiological costs of body armor. Mil. Med. 134: 204-209, 1969.
10. Goldman, R.F., E. Green and P.F. Iampietro. Tolerance of hot wet environments by resting men. J. Appl. Physiol. 20: 271-277,1965.
11. Goldman, R.F. and R.J.T. Joy. Prevention of heat casualties in men wearing chemical-biological protective clothing. Tech. Rpt. U.S. Army Res. Inst.of Environmental Med.,Natick, July 1967.
12. Iampietro, P.F. and R.F. Goldman. Tolerance of men working in hot, humid environments. J. Appl. Physiol. 20: 73-76, 1965.
13. Joy, R.J.T. and R.F. Goldman. A method of relating physiology and military performance. Mil. Med. 33: 458-470, 1968.
14. Manets, F.I., P.F. Sevostyanov, A.F. Dudnikov and A.A.Kondrashov. Defense against weapons of mass destruction. Moscow, 1967.
15. Woodcock, A.H. Moisture transfer in textile system. Textile Res. J. 32: 628-723, 1962.
16. Yarger, W.E., P.L. Schwartz and R.F. Goldman. An assessment of CBR protective uniforms during an amphibious assault in a tropical environment. U.S. Navy Medical Field Research Lab. Rpt. XIX No. 17, Camp LeJeune, No. Carol. Nov. 1969.

ACKNOWLEDGEMENTS

While it is impossible to acknowledge all the professional collaborators in these years, LTC R.J.T. Joy, M.C., U.S.A. LCDR W.E. Yarger, M.C., U.S.N.R., Maj H. Martins, R.A.M.C., U.K. and Mr. J.R. Breckenridge, Mr. H. Hanson and Mr. G. Newcomb of USARIEM have at times been coauthors; without them, this much could not have been accomplished. Significant support was provided by other DOD elements including: AMC, particularly its Natick Labs. and Edgewood Arsenal; CONARC and its troops, and TECOM at Forts Lee, Benning, Clayton, and Carson; the CDC-CDEC, CBRA and CSG; the US Navy and its Medical Field Research Lab; the US Marine Corps FMF, Atlantic.

AD 713 531

SIMULATION OF HIGH ALTITUDE IONIZED AIR WITH AN ARTIFICIAL DIELECTRIC

WILLIAM GRAY AND GEORGE MERKEL
ELECTROMAGNETIC EFFECTS LABORATORY, U. S. ARMY MERDC
FORT BELVOIR, VIRGINIA

1. INTRODUCTION

A number of nuclear electromagnetic pulse simulators have been constructed to simulate various nuclear electromagnetic pulse environments. These simulators yield a variety of pulse magnitudes, pulse shapes, polarizations, angles of arrival, etc. Present techniques for the simulation of the interaction between an in flight missile and a nuclear electromagnetic pulse do not include the effect of ionized air that can be produced by the Compton-recoil electrons arising from the interaction of gamma-rays with the atmosphere. This paper describes a relatively modest program at the Electromagnetic Effects Laboratory at MERDC in which we are investigating the possibility of developing simulation methods that include some of the effects of the presence of an ionized air plasma when a nuclear EMP interacts with a missile or satellite.

During the past twenty-five years (1) so-called artificial dielectrics consisting of regularly spaced rods, parallel plates, metal spheres, etc. have been devised to reproduce the essential macroscopic properties of a dielectric. The ordinary dielectric consists of molecular particles of microscopic size, whereas the artificial dielectric consists of discrete metallic or dielectric particles or lattices of macroscopic size. These artificial dielectrics were first actually conceived as large scale macroscopic models of microscopic crystal lattices. The practical motivation for the development of the first artificial dielectrics was the desire to

obtain relatively inexpensive lightweight materials that could be used for microwave and radar lenses. Several of the artificial dielectrics proposed for microwave lenses have the macroscopic electromagnetic properties of a plasma.

In this work we investigate the feasibility of using an artificial dielectric consisting of a cubic resistive rod lattice structure (2)(3)(4) to simulate the macroscopic electromagnetic properties of pre-ionized air at high altitudes. As we shall indicate the interaction between a missile or satellite and an incident electromagnetic pulse might be appreciably affected by an ambient plasma of ionized air.

The interaction of a nuclear EMP with a missile can be very complicated. For example, at relatively low altitudes the slowing down of the high energy electrons produced by Compton scattering of γ-rays results in a collision dominated plasma (ν = plasma collision frequency > ω_p = plasma frequency) with a time varying conductivity that varies at a rate comparable to the rise time of the nuclear electromagnetic pulse. The situation is complicated by the fact that the current sources which produce the nuclear electromagnetic pulse field are the electrons that are ionizing the air. The coupled interaction between a missile, the Compton current, the ionized air with its varying conductivity, and the nuclear electromagnetic field would be very difficult to simulate with an artificial dielectric.

In this work we consider a simpler simulation problem, i.e., simulation of the interaction of a missile or satellite and a nuclear EMP when the satellite or missile is immersed in pre-ionized air. For example, the ionized air might be the result of a prior detonation of another nuclear device.

In short, our objective is to produce a homogeneous isotropic solid material which obeys the macroscopic constitutive relations of a simple plasma. The macroscopic properties of a simple non-collision dominated plasma are described by a collision frequency ν which depends on altitude roughly as $\nu = 4\times10^{12}\,\rho/\rho_o$ 1/

1/ ρ and ρ_o refer to the atmospheric density at altitude H and at sea level respectively. (5)

and an electron density N. Then with the plasma frequency given by

$$\omega_p^2 = \frac{Ne^2}{\epsilon_o m}$$

where m is the mass of the electron, e is the electron charge, and ϵ_o is the permittivity of free space, the ionized air (in the absence of magnetic fields) can be described by a complex dielectric constant (6)

$$\epsilon_p = \epsilon_o \left(1 - \frac{\omega_p^2}{\nu^2 + \omega^2} + j \frac{\omega_p^2 \nu/\omega}{\nu^2 + \omega^2} \right), \qquad (1)$$

and by a magnetic permeability, μ_o, equal to that of free space.

Rotman (3) suggested that a medium which demonstrates behavior similar to that predicted by Eq. (1) can be constructed with a lossy cubic grid of rods such as is shown in Fig. (1). In a number of experimental studies the properties of one and two dimensional grids have been measured and have been found to exhibit many of the properties we would expect for a simple lossy plasma. (2)(3)(4)(7) We are presently measuring the electromagnetic properties of 3-dimensional cubic grids immersed in a plastic matrix and are examining the possibility of developing an isotropic homogeneous artificial dielectric that behaves according to Eq. (1). A missile or satellite might then be immersed in this isotropic (fine meshed) plasma simulating structure and be subjected to an EMP that corresponded to a possible tactical situation.

Before describing our effort, it would be worthwhile to explain why ionized air may influence the interaction of a satellite or missile with a propagating electromagnetic pulse produced by a nuclear explosion. We can look upon the skin of a satellite or missile as the interface between an impinging nuclear electromagnetic pulse and the interior of the missile. The resonances of a conducting sphere or prolate spheroid (e.g., a missile) are quite pronounced. (8)(9) For example, Fig. (2) shows the radar cross sections for a sphere as a function of wave-length. The same sort of behavior holds for a prolate spheroid or cigar shaped figure. The maxima in the radar cross sections correspond to large skin currents. If these surface currents cause electromagnetic fields to leak into the

interior of the missile and if the skin resonance frequencies correspond to critical damage inflicting frequencies, we could have a bad situation.

In general, a missile of length L will resonate at frequencies given by (10)

$$f = \frac{c}{L(N/2 + 1)} ,$$

where c = speed of light and N is an integer. When N=0 we have the first resonance, this corresponds to the first peak in Fig. (2). For simplicity of argument, let us consider what happens at an altitude high enough so that we can set the plasma collision frequency ν equal to zero. Then from Eq. (1)

$$\epsilon_p = \epsilon_o \left(1 - \frac{\omega_p^2}{\omega^2}\right),$$

and the index of refraction is given by

$$n = \sqrt{\mu_o \epsilon_p} = \left[1 - \left(\frac{\omega_p}{\omega}\right)^2\right]^{\frac{1}{2}} .$$

The missile now resonates when $(\lambda/n)(\frac{1}{2}+N) = L$ or at a new set of frequencies,

$$f = \frac{c}{\left[1 - \left(\frac{\omega_p}{\omega}\right)^2\right]^{\frac{1}{2}}} \quad \frac{1}{L\left(\frac{N}{2} + 1\right)} .$$

In other words, the change in the index of refraction of the medium surrounding a missile or satellite changes the effective electrical length or size of the missile or satellite. The changes in effective electrical length caused by the changes in the effective ambient dielectric constant change skin resonance frequencies. The response of individual components to γ or neutron radiation is unlike the response of a resonating missile skin to a nuclear EMP in that an entire system or subsystem responds to the currents developed by the nuclear EMP. If the changes in missile skin resonance frequencies correspond to critical resonant frequencies in subsystem electronic circuits, a very deleterious interaction may

arise because of the increased coupling between the nuclear EMP and the internal missile subsystems.

2. THEORETICAL DISCUSSION

From an intuitive point of view it might be useful to discuss the rodded media in terms of simple lumped element transmission lines. (11) In the simple one dimensional harmonic case Maxwell's Equations

$$\text{curl } E = -\mu_o \frac{\partial H}{\partial t}$$

$$\text{and curl } H = \epsilon_o \frac{\partial E}{\partial t}$$

can be simplified to

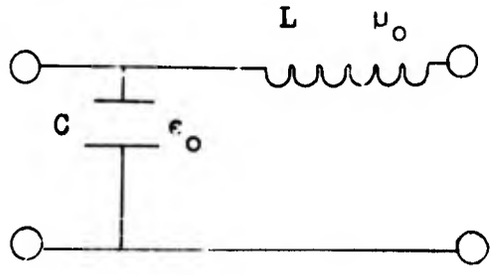
$$\frac{\partial E}{\partial x} = -j\omega\mu_o H \tag{2}$$

$$\text{and } -\frac{\partial H}{\partial x} = j\omega\epsilon_o E \ . \tag{3}$$

These two equations are analogous to the voltage and current equations for the lumped element transmission line formed of sections as shown in Diagram (A):

$$\frac{\partial V}{\partial x} = -j\omega LI \tag{4}$$

$$\text{and } -\frac{\partial I}{\partial x} = j\omega CV \tag{5}$$



(A)

Table I summarizes the correspondence between some of the free space and transmission line parameters.

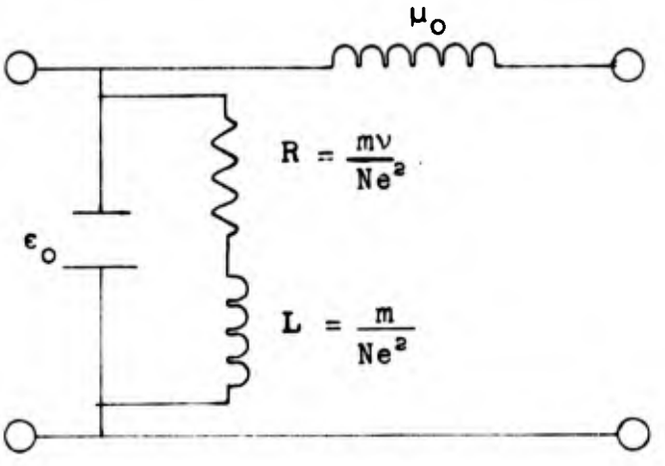
TABLE I

free space	circuit
E	V
H	I
ϵ_o	C
μ_o	L

If we consider the passage of a monochromatic electromagnetic plane wave through a plasma with electron density N, electron charge e, and collision frequency ν, Eq. (3) is modified to

$$-\frac{\delta H}{\delta x} = j\omega\epsilon_o E + \frac{Ne^2 E}{m\nu + j\omega m}, \qquad (6)$$

whereas the equation for $\delta E/\delta x$ remains the same as (2). A transmission line which is described by an equation similar to Eq. (6) can be constructed with sections as shown in Diagram (B):



(B)

In this more complicated section the equation for $\delta V/\delta x$ remains as Eq. (4), but the equation for $\delta I/\delta x$ becomes

$$-\frac{\delta I}{\delta x} = \left(j\omega C + \frac{1}{R + j\omega L_D}\right) V.$$

The relation between the plasma electron density, N, the plasma collision frequency, ν , and the corresponding resistance, R, and inductance, L_p, in the filter section is indicated in Diagram (B).

A method of modifying free space so that an electromagnetic wave passing through it obeys an expression similar to Eq. (6) is to fill the free space with a cubic grid of inductive lossy wires. (2)(3)(4)(12) Although not investigated with thorough mathematical rigor, the suggestion seems plausible. The plausibility argument goes as follows: In free space, the ratio of the E field to an H field in a propagating wave is Z_o = 377 Ω ; i.e., E/H = 377 Ω. When a plane wave interacts with a grid of equally spaced lossy wires, the electromagnetic field is partially transmitted through the grid, and the ratio of E to H along the plane of the wires can be shown to be given by (13)(14)(15)

$$Z_t \approx \frac{1}{2r}\left(\frac{\omega\mu_o}{\pi\lambda_o\sigma}\right)^{\frac{1}{2}} (1+j) + j\frac{Z_o d}{\lambda_o}\log_e\left(\frac{d}{2\pi r}\right) \qquad (7)$$

where d is the grid spacing, r is the wire radius, λ_o is the wavelength in free space, and σ is the wire conductivity. In other words, the wire grid acts as an inductance in series with a resistor. As shown in Eq. (7) the values of the effective resistance and the inductance of the wire grid can be controlled by varying the wire spacing, the wire conductivity, and the wire diameter.

Rotman (3) who suggested a cubic lossy inductive wire medium as a plasma simulator, gives the following expression for the complex propagation constant, $\gamma_r = \alpha_r + i\beta_r$ (α_r = attenuation constant, $\beta_r = 2\pi n/\lambda_o$ = phase coefficient) for a medium consisting of a cubic grid of rods embedded in a material with a relative dielectric constant, ϵ_m:

$$\cos\gamma_r d = \cos\left(\frac{2\pi\sqrt{\epsilon_m}\, d}{\lambda_o}\right) + j\frac{Z_o}{2Z_t\sqrt{\epsilon_m}}\sin\left(\frac{2\pi\sqrt{\epsilon_m}\, d}{\lambda_o}\right) \qquad (8)$$

where d is the grid spacing, ϵ_m is the relative dielectric constant of the lossless embedding medium, λ_o is the free space wavelength, Z_o = 377 Ω , Z_t is the grid impedance (Eq. 7), and n is the index of refraction.

What we attempt to do in the artificial dielectric is to replace the lossy inductive plasma electrons with a symmetric lossy inductive grid structure. Of course, the wires are lumped elements, whereas the plasma is continuous. We anticipate that we will eventually be able to make the grid structure fine enough so that it appears continuous to a missile skin. As can be seen from Eq. (7), the inductive impedance of the wire grid, Z_t, can remain constant as the distance between wires, d, is decreased provided that the wire radius, r, becomes very small.

3. EXPERIMENTAL RESULTS

We have constructed two blocks of artificial dielectrics by embedding 2 mil copper wires in 6.5" x 6.5" x 3.25" polystyrene blocks. In one block the basic cube measured 4.13 cm^3; in the other the basic cube was 2.75 cm^3. Figure (1) shows a typical cubic grid structure and Fig. (3) shows an orthographic presentation of one of our artificial dielectric blocks. The two blocks were inserted into an L band waveguide and their relative dielectric constants, ϵ , were measured for a number of frequencies. The measurement procedure is summarized in Fig. (4). (16) The results for one of the blocks are given in Table II.

TABLE II

ARTIFICIAL DIELECTRIC MADE BY EMBEDDING 2 MIL CU WIRE IN POLYSTYRENE ϵ=2.6.

Grid Cube	4.13 cm^3	
f (kmc/sec)	calc ϵ	measured ϵ
1.7	2.16	1.8
1.6	2.10	1.2
1.5	2.03	1.2
1.4	1.94	1.9
1.3	1.84	2.1
1.2	1.70	1.2
1.1	1.53	1.5

We feel that divergence between calculation and waveguide measurements is due to lack of effective homogeneity of the artificial dielectric. (7) We would expect better agreement if the size of the cubic cell were much less than the wavelength. (See Sec. 4) The reduction of the effective dielectric constant of the polystyrene blocks by the embedded cubic grid is encouraging in that it

indicates that we may be able to achieve a workable plasma simulating material by embedding a cubic grid of wire in a dielectric foam that has a dielectric constant of approximately 1. (See caption of Fig. 1)

In a second set of experiments we are attempting to measure the homogeneity of our artificial dielectric by utilizing the nodal shift or tangent method described by Weissfloch. (17) Weissfloch showed that (for a lossless junction) a plot of the voltage minimum at the input of a discontinuity in a guide as a function of the position of a shorting plunger at the output side resulted in an S curve, as shown in Fig. (5), which could be interpreted to yield information about the discontinuity. For example, curve B shows the expected behavior of a nodal shift curve right before the plasma cutoff frequency. In our artificial dielectric we embedded wires in a polystyrene block with a relative dielectric constant of 2.6. The effect of the wire grid is to reduce the dielectric constant. According to Eq. (8) there is a frequency at which the wires embedded in the plastic block should produce an artificial dielectric that has an average dielectric constant equal to one. If the artificial dielectric were completely homogeneous it should essentially disappear when $\epsilon_{av} = 1$, and the Weissfloch S curve should become a straight line. Up to the present, most probably because of our rather gross cell structure compared to wavelength, we have not been able to obtain such a straight line.

There is a deviation from the well behaved periodic pattern of the S curve when the plunger actually contacts the artificial dielectric. This deviation emphasizes another aspect of the inhomogeneity that must be considered. As indicated, the wires embedded in our plastic matrix act as lumped inductive elements and are not homogeneously distributed. A strictly accurate description of our artificial dielectric would have to explicitly consider evanescent non-propagating modes. As A. M. Model (12) has indicated in his discussion of the propagation of waves in a space filled with plane parallel grids, when we consider a wave normally incident toward a plane of a grid the incident wave excites currents in the grid which in turn produce a field which is propagated in both directions from the grid. A basic assumption in the derivation of Rotman's formula (Eq. 8) is that the secondary wave is a plane wave. The inaccuracy of our derivations is contained in this assumption. We anticipate that the closer the grid spacing the more reasonable the plane wave assumption. Brown (18)

has discussed the importance of non-propagating evanescent modes in a cubic grid lattice and found them to be relatively small. Our approach in this project is to go to as small a grid spacing as we can. Of course, if we want to go to smaller wire spacings we have to use thinner wires, and with wire much thinner than 0.0001" the physical problem of handling a very fragile gossamer-like wire grid arises.

4. EXPERIMENTAL WORK IN PROGRESS

We are now in the process of extending our experimental measurements down to approximately 300 mc/sec. This work involves the use of rather large type WR2300 waveguide equipment, the dimensions of which are 23.00 x 11.5 inches. We are assembling an artificial dielectric with 0.00012" tungsten wire. The advantages of the relatively low frequency in combination with the thin wire should result in a close mesh in which the ratio of our basic cubic cell dimensions to wavelength will be smaller. "Common sense" more or less indicates that as the ratio of the artificial dielectric cell compared to the wavelength becomes small enough, the result should be an artificial dielectric that behaves as if it were homogeneous. We hope that we can make a step toward achieving such an artificial dielectric in the 400 Megahertz region.

What we would really like to be able to do is to immerse a round cylindrical object or a spherical object in our plasma simulator with its basic cubic grid structure and be able to describe the boundary conditions between the round missile surface and the artificial dielectric as if the artificial dielectric were completely homogeneous.

5. CONCLUSIONS

As indicated in Table II, our measurements so far are in qualitative agreement with Rotman's formula describing the average constitutive properties of a cubic grid or rods embedded in a dielectric matrix (Eq. 8). When subjected to measurements as described in Sec. (3) the material does not act as if it were completely homogeneous. For example, our artificial dielectric samples acted as if they were made up of a lattice of cubic cells with different dielectric constants. The dielectric constants, however, tended to average to values relatively

close to those predicted by Eq. (8). By decreasing the ratio of the artificial dielectric cell size to wavelength we hope to achieve a more homogeneous artificial dielectric. Then by embedding the wires in a matrix such as styrofoam which has a dielectric constant equal to one we may be able to achieve a versatile plasma simulating material.

ACKNOWLEDGMENT

The authors wish to thank D. H. Petley for his technical assistance, and A. Renner, R. Bostak, and W. Haas for their interest and encouragement.

REFERENCES

1. W. E. Kock, Bell System Tech. J., 27, p. 58 (1948). H. Jasik, "Antenna Engineering Handbook", McGraw-Hill, New York, Chapt. 14 (1961).
2. J. Brown, "Microwave Lenses", Methuen & Co., Ltd, London (1953).
3. Walter Rotman, IRE Trans. on Antennas and Propagation, Vol. AP-10, p. 82, Jan. (1962).
4. K. E. Golden, IEEE Trans. on Antennas and Propagation, Vol. AP-13, p. 587, July (1965).
5. R. A. Nielsen and N. E. Bradbury, Phys. Rev. 51, 69 (1937).
6. R. F. Whitmer, Microwave J., 2, p. 17, Febr. (1959).
7. M. M. Z. El-Kharadly, Proc. IEE (London), Paper No. 1700R, p. 17, Jan. (1955).
8. H. Weil, M. L. Barasch, and T. A. Kaplan, Univ. Mich. Eng. Research Inst. Rept. 2255-20, July (1956).
9. K. M. Siegel, et al, IRE Trans., Vol. AP-4, p. 266, July (1956).
10. E. C. Jordan, "Electromagnetic Waves and Radiating Systems", Prentice-Hall, New York, p. 550 (1950). (Note 1950 Edition)
11. R. N. Bracewell, Wireless Engineer, 31, p. 320, Dec. (1954).
12. A. M. Model, Radiotekhnika, 10, p. 52, June (1955).
13. G. G. McFarlane, J. Inst. Elec. Engrs. (London) 93, 1523 (1946).
14. R. Honerjäger, Ann. Physik, 4, p. 25 (1948).
15. J. P. Casey, Jr. and E. A. Lewis, Jour. Opt. Soc. of Am., 42, p. 971, Dec. (1952).
16. A. L. Lance, "Introduction to Microtheory and Measurements", McGraw-Hill, New York, p. 289 (1964).
17. A. Weissfloch, "Hochfrequenz und Electroakustic", 60, p. 67 (1942).
E. Feenberg, J. Appl. Phys., 17, 530 (1946).
18. J. Brown, Proc. IEE, Monograph No. 62R, 100, 51, May (1953).

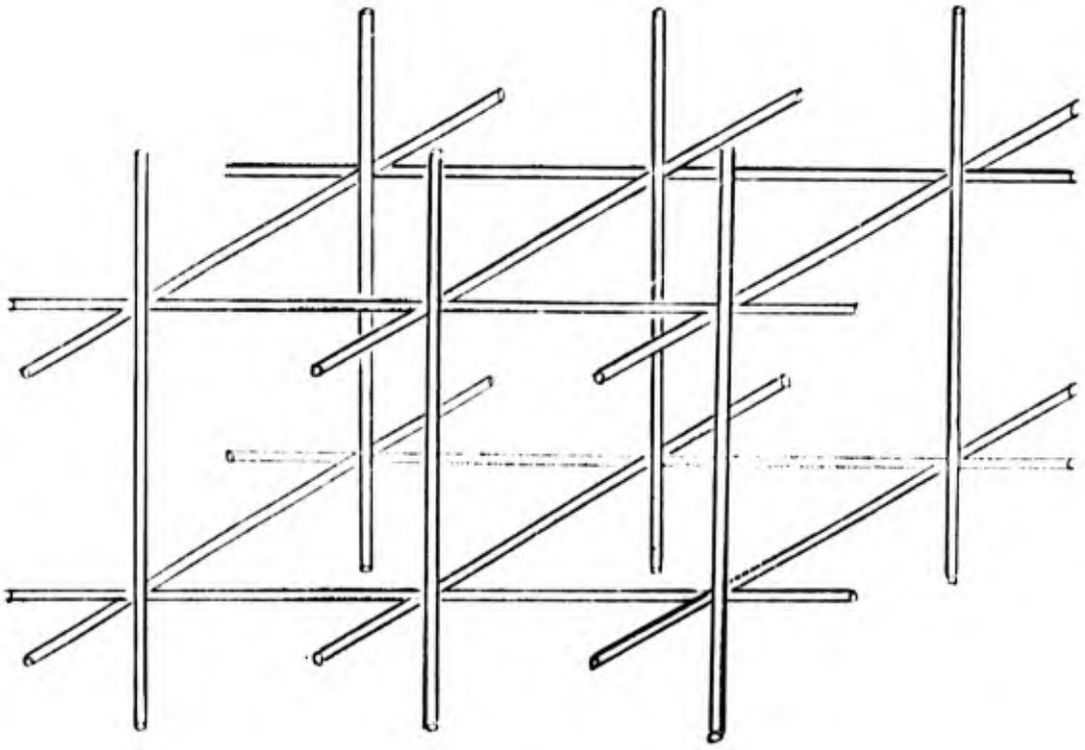


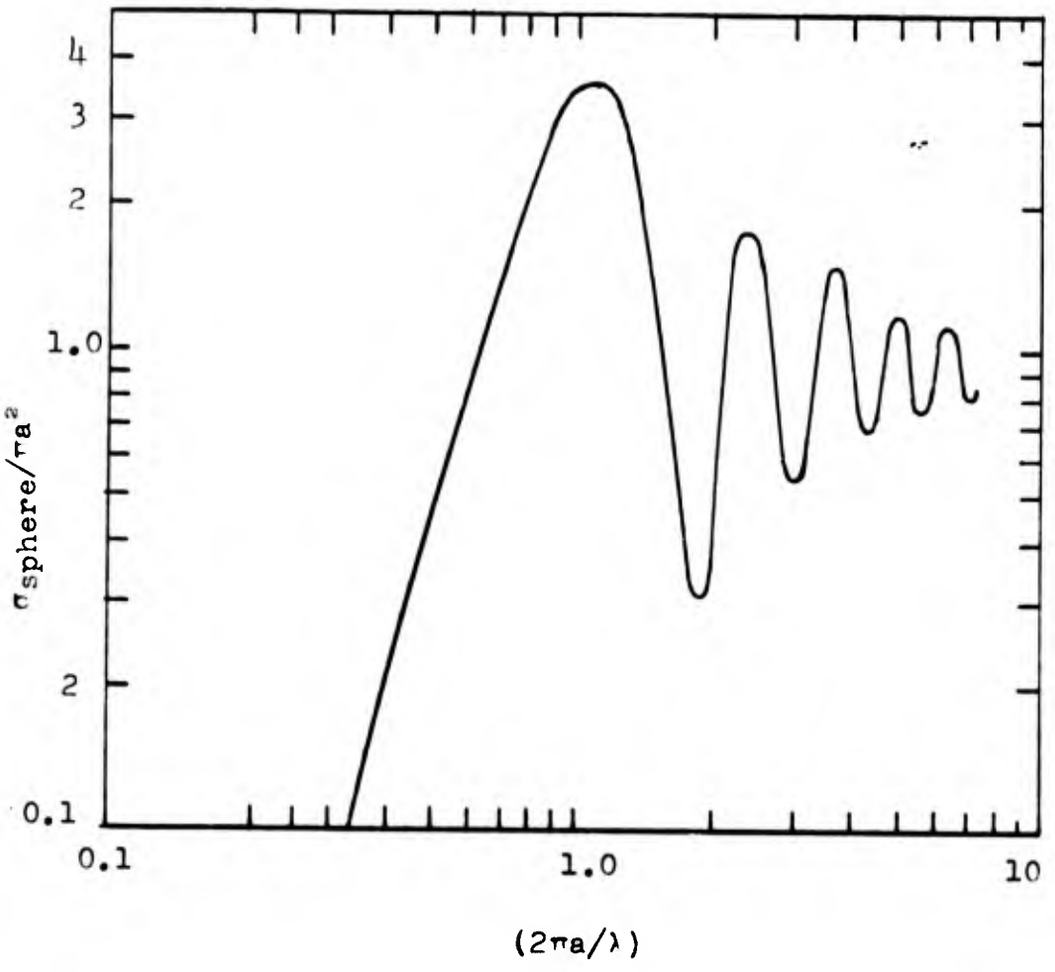
Figure 1

MEDIUM CONSTRUCTED WITH LOSSY RODS:
THREE-DIMENSIONAL CUBIC LATTICE

n = index of refraction of cubic lattice
$\approx (1 - \lambda_o^2/\lambda_p^2)^{\frac{1}{2}}$, where $\lambda_p^2 \approx 2\pi d^2 \ln(d/2\pi r)$,
d = rod spacing, and r = rod radius.

The index of refraction of a simple collisionless plasma is governed by the same form of expression as that for a cubic lattice of lossless rods. Note that the value of λ_p can remain constant as the rod spacing or grid mesh is decreased if the rod radius becomes very small. When the grid is embedded in a plastic with dielectric constant ϵ, the index of refraction n, becomes $n \approx (\epsilon - \lambda_o^2/\lambda_p^2)^{\frac{1}{2}}$.

Figure 2



where λ = wave-length, which is a function of electron density, N; and a = radius of sphere, a constant.

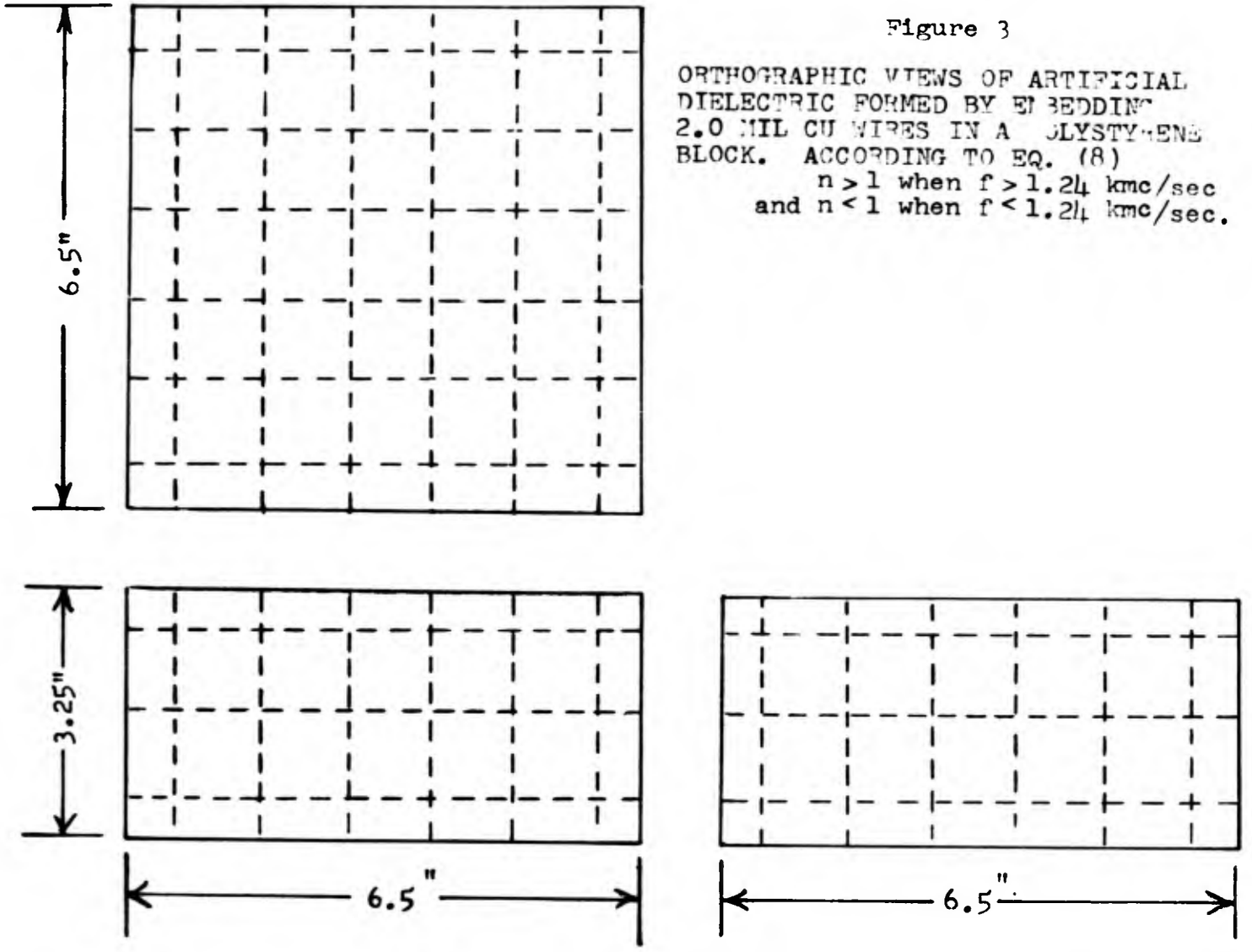


Figure 3

ORTHOGRAPHIC VIEWS OF ARTIFICIAL DIELECTRIC FORMED BY EMBEDDING 2.0 MIL CU WIRES IN A POLYSTYRENE BLOCK. ACCORDING TO EQ. (8)
$n > 1$ when $f > 1.24$ kmc/sec
and $n < 1$ when $f < 1.24$ kmc/sec.

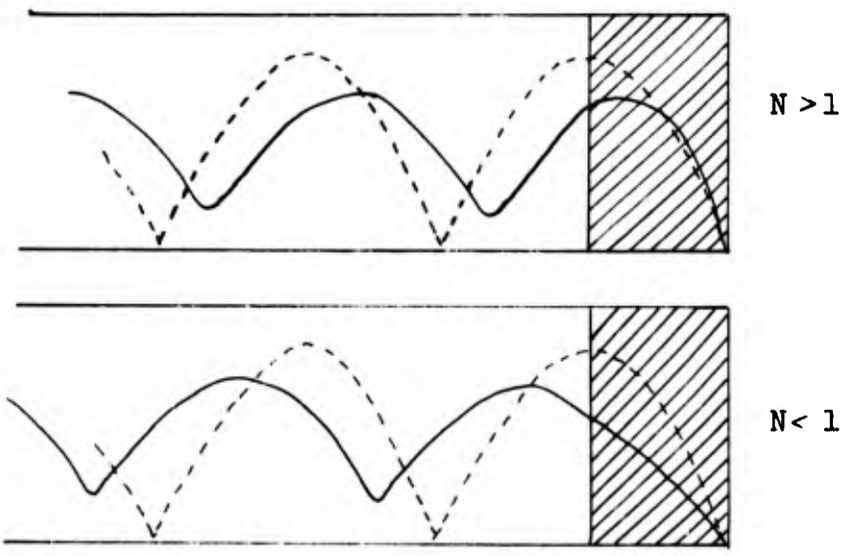


Figure 4

Change in position of minimum yields value of N. Value of VSWR yields value of complex component of the dielectric constant.

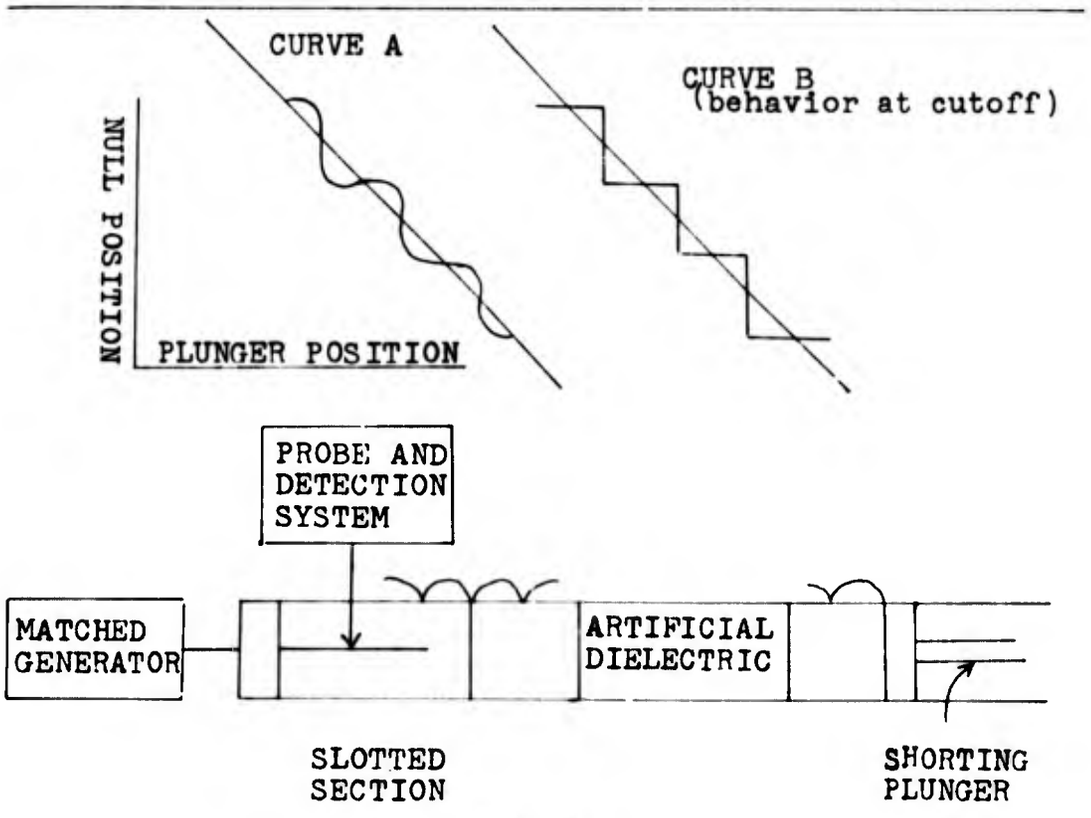


Figure 5

Nodal Shift Measurement Setup

AD 713 532

M16 RIFLE/AMMUNITION MALFUNCTION MODELING

HARRY A. GREVERIS
DEPARTMENT OF THE ARMY, FRANKFORD ARSENAL
PHILADELPHIA, PENNSYLVANIA 19137

In some respects, the design of ammunition and rifle systems is an art relying heavily on past experiences and inventive know-how. Often, abbreviated research and development leads to the neglect of interfaces and interactions. Malfunctions are not understood and meaningful acceptance/rejection criteria remain unidentified. For example, customary testing and large accumulation of data did not immediately provide explanations for malfunctioning of the M16 rifle system because the system was sensitive to rifle-to-rifle and ammunition-to-ammunition variation. It became necessary to augment research and development by utilizing simulation techniques for uncovering interactions and evaluate matching criteria. The approach consisted of developing mathematical models describing the interfaces between ammunition and rifle and diagnostic techniques to measure these interactions.

Figure 1 contains a summary of some of the types of malfunctions experienced during a series of M16 rifle system tests in terms of occurrences per 1000 rounds distributed according to cyclic rate. A malfunction model must readily simulate this type of experience.

The M16 rifle system provides an appropriate basis for a discussion of the procedure. Figure 2 shows the basic construction of the gas operated cycling mechanism of the M16 rifle. The primer ignites the propellant. The high pressure gases drive the projectile down the barrel. Near the muzzle, the projectile passes a porthole which diverts a portion of these gases through a gas tube to the bolt mechanism. The pressure pulse drives the carrier rearward as a unit. The rearward movement depresses the hammer, separates the gas tube from the carrier, extracts and ejects the spent case, and completes compression of the driving spring. After reaching its most rearward position, the carrier assembly rebounds and moves forward, picks up a new round from the magazine, chambers it, and locks the bolt. The

firing pin is struck by the hammer and the cycle is repeated. It is this interrelationship between the driving force produced by the ammunition and the associated response of cycling mechanism which provides a basis for defining a malfunction model. In practice, the model involves computer programming of the governing differential equations combined with measurements and data derived from diagnostic procedures.

For single shot operation, a non-linear differential equation describing the motion of the bolt carrier gives a good mathematical model. That is, cycling mechanism response in terms of ammunition impulse and other interaction can be nearly represented by:

$$M_i\ddot{x} + B_i\dot{x} = \sum F_i(x,t)$$

where M_i = masses

x = displacements

B_i = damping

$F(x,t)$ = accelerating and retarding forces

Since the forces are not only functions of displacement but also of time, the motion of the bolt carrier is separated into continuous motion phases where forces are added or deleted depending upon cycling mechanism, ammunition, and magazine component and effects relative locations in time and displacement. All interactions for single firing by mathematical expressions derived from analysis or measurements are accounted for in the force functions. Additions to the model and operational as separate elements include cartridge case extraction, ejected cartridge trajectory, gas transmission, cartridge feeding and chambering, part inertia and geometry, and dynamic spring response (surging).

Dynamic response of the driving spring or surging is an example of the detail worked into the model. This phenomenon is included in the model because the cyclic rate of the M16 rifle and the natural frequency of the spring are of the same order. This contributes to weapon cyclic rate and feeding variations. Two mathematical models have been developed and are operational; one a coupled mass model and the other a continuous model which is expressed by a modified wave equation:

$$\rho_0\frac{\partial^2 u}{\partial t^2} + \alpha\frac{\partial u}{\partial t} = \beta\frac{\partial^2 u}{\partial x^2} + \eta\frac{\partial^3 u}{\partial x^2 \partial t}$$

where P_o = initial linear mass density

u = displacement measured from equilibrium

α, β, η = positive constant moduli describing respectively the elastic and viscous properties of driving spring

The solution to the above equation coupled with the generalized bolt motion model gives the coil displacement and spring force as a function of time. Figure 3 shows an actual measurement of spring surging during rifle operation. Figures 4 and 5 show agreement between the model and experience.

Automatic fire and burst fire present complications in modeling because of the random characteristic of feedbacks and ammunition pressure characteristics. Recursive models for these characteristics are needed before full automatic burst fire can be simulated.

The mathematical model accommodates malfunction events in the form of dependency. Exercise of the model requires specification of propellant gas impulse characteristics, some material properties, part geometry and weight, some tabulated and measured data, and identification of sequencing events. Model response takes the form of system kinematics; for example, displacement, velocity, and acceleration of the bolt carrier. Comparison criteria immediately identifies the type of malfunction and the point of occurrence. The subtle effects of heat transfer, friction, and part fits do not have convenient models and are best handled by real system inputs. It is for this reason that the malfunction modeling system has evolved into mathematical simulations combined with diagnostic procedures to form a small arms diagnostic system.

The small arms diagnostic system provides a capability for evaluating and predicting malfunctions of production or fielded weapon/ammunition systems. Rifles and/or ammunition are non-destructively "plugged" into a facility which measures characteristic parameters related to rifle and ammunition populations having the capability for predicting malfunctions.

One such parameter useful for the M16 rifle system is the rifle signature curve which relates firing rate or actually cycle time to the initial velocity of the bolt carrier or in terms of ammunition, the imparted impulse produced by the propellant gases. As shown in Figure 6 for small initial velocities or low output ammunition energies, the bolt carrier does not travel rearward to its maximum allowable distance. The resulting short travel corresponds to a small time interval and an associated high rate of fire. For greater initial velocities, increasing ammunition energy, the travel distance increases and the firing rate decreases. As the energy increases, the initial velocity also increases until it is just sufficient to

drive the bolt carrier to its maximum rearward displacement; this is the point of minimum firing rate. For still greater velocities, the bolt carrier will rebound thereby reducing the cyclic time and conversely increasing the firing rate.

The initial rearward velocity of the bolt determines the time intervals spent by the bolt between different locations in the rifle. These time intervals are critical to the functioning of the rifle. For example, failure-to-feed malfunction is illustrated schematically in Figure 7. One cause of this failure is that the magazine does not move the round into place before the bolt is in position to pick it up. This type as well as other malfunctions can be related to the firing rate. This correlation between malfunction criteria and firing rate is obtained by establishing malfunction lines on the rifle characteristic curve. For example, Figure 8 shows the malfunction lines for the failure-to-feed criteria. Malfunction lines for this malfunction as well as for other malfunctions require identification either indirectly from the mathematical model or directly from the diagnostic device.

Experience indicates that each rifle has its own signature curve and that the curve is distinguishable among rifles. A series of curves can be generated to represent the population of rifles. The limiting velocity of the bolt carrier and the malfunction derived firing rate range determine acceptable rifle and/or ammunition distribution. Any rifle whose signature curve lies outside the accepted distribution will be unacceptable.

The ammunition signature curve relates the statistical distribution of the ammunition energy in a specific lot of ammunition. Each lot has such a signature curve. In particular, the ammunition signature establishes the initial velocity that the ammunition is capable of imparting to the bolt carrier; that is, the initial velocity available from the ammunition. This signature can be used directly to obtain malfunction probability. For example, in Figure 9, the rifle signature is plotted along with its failure-to-feed malfunction line; included in this figure is the ammunition signature. The malfunction line divides the function and malfunction regimes. If the ammunition provides an initial bolt carrier velocity to the right of this line, then this type of malfunction will occur. Therefore, the crossed/hatched area under the ammunition signature curve represents the probability of a failure-to-feed malfunction. In a similar manner, the malfunction probability for other types of malfunctions can be obtained by suitably combining the rifle and ammunition signature curves. Extending this single rifle/ammunition group performance evaluation, malfunction probabilities can be obtained for population of rifles and ammunition by using such methods as Monte Carlo techniques. These effectively utilize random samples from population of rifles (production lots or from a squad in field use) and a population of ammunition (production lot) and by

this statistical sampling provide a meaningful malfunction expectation for the real ammunition/rifle system, under real conditions.

A key element in the small arms diagnostic system is availability of devices for non-destructive measuring of characteristic parameters in a "plug-in" fixture. Two such fixtures have been designed and built. Figure 10 shows a quasi-steady test fixture equipped with pneumatic actuators and transducers which operate and measure the motion of the carrier and bolt and the related forces during the firing cycle. The specific measurement as shown in Figure 11 is force and displacement. Utilization of these data in the mathematical model produces the rifle signature. An example of this signature, taken from the computer run, is shown in Figure 13 with displacement-type malfunction lines superimposed. All displacement-type malfunctions occur at values of initial bolt velocity which are less than that required to just impact the buffer.

The mean and one sigma rifle signature distribution for a test sample of 44 rifles is shown in Figure 13. Here it can be seen that 23 of the rifles tested have a value of initial bolt velocity for buffer impact falling between 200 in/sec and 205 in/sec. In addition, from this test sample of 44 rifles, it is important to note that the one sigma characteristic curve spread is considerable to the left of the buffer impact, and very tight to the right. This dramatic change in the spread of the rifle characteristic amplifies the displacement malfunction associated with rifles operating to the left of the minimum point.

Figure 14 shows an alternate and more direct procedure consisting of a dynamic test fixture designed to cycle the rifle at any specified rate by varying the magnitude of the impulse applied to the bolt carrier. The fixture is instrumented to measure bolt carrier velocity, displacement and critical times directly. The rifle signature curve is generated directly by varying the impulse through a wide range of values.

Plans call for comparing the quasi-steady static procedure with the direct dynamic impulse procedure to determine if indeed each rifle has a unique signature when it is tested in both ways. Successful testing will lead to the design, fabrication, and evaluation of an on-line diagnostic facility for small arms weapon systems. The system will take the form of diagnostic devices and computer software. Rifle and ammunition producers as well as testing facilities and system developers will have a powerful tool for maintaining reliability and evaluating new concepts.

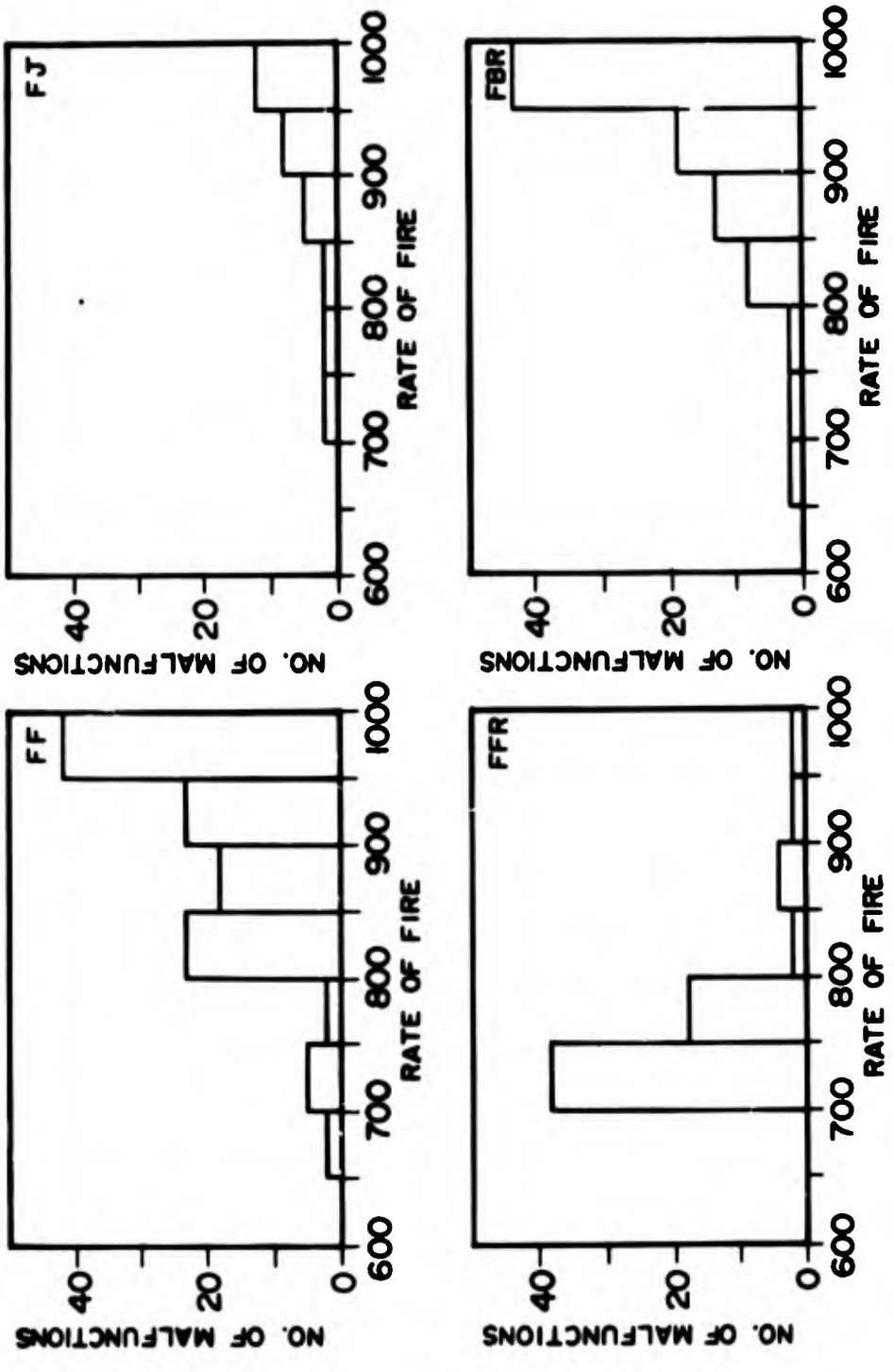


FIGURE 1: TYPICAL M16 MALFUNCTION DISTRIBUTION

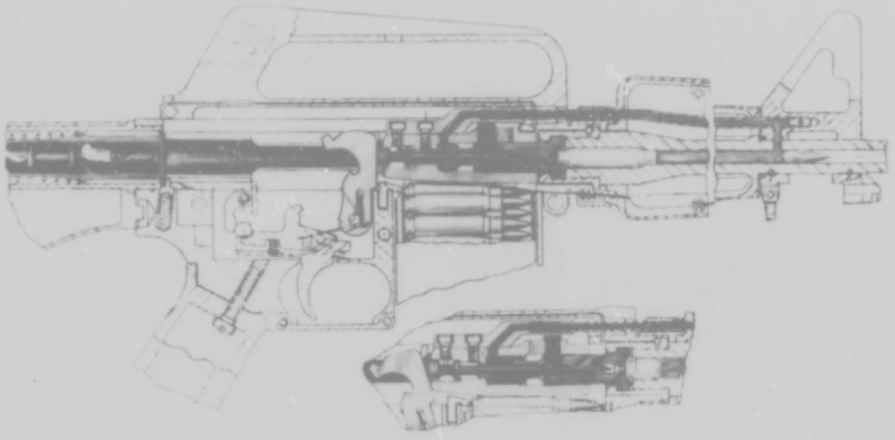


FIGURE 2: RIFLE CYCLING MECHANISM

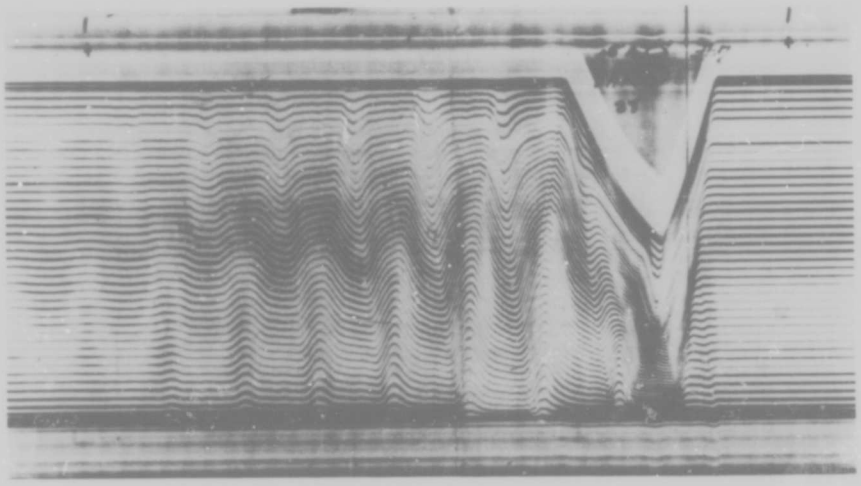


FIGURE 3: DYNAMIC SPRING RESPONSE

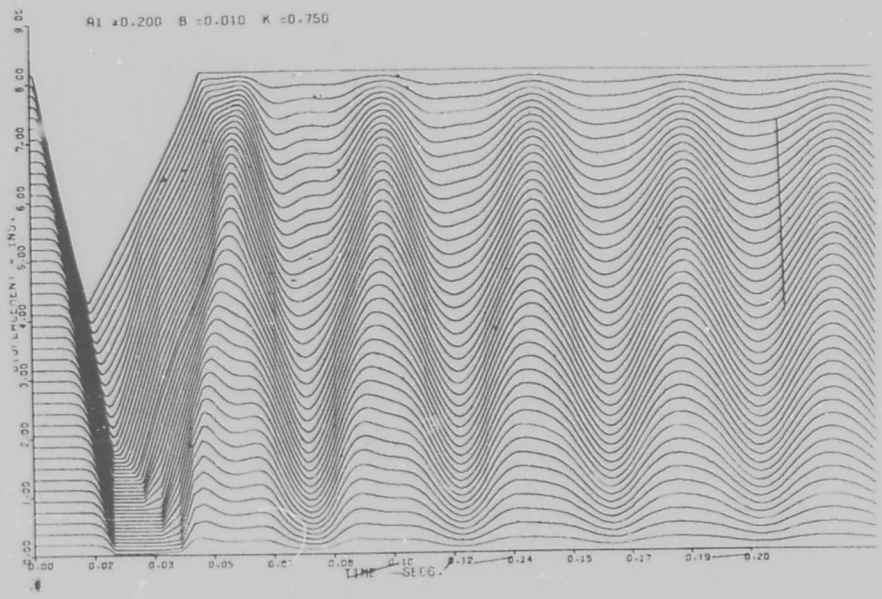


FIGURE 4: SPRING SURGE MODEL

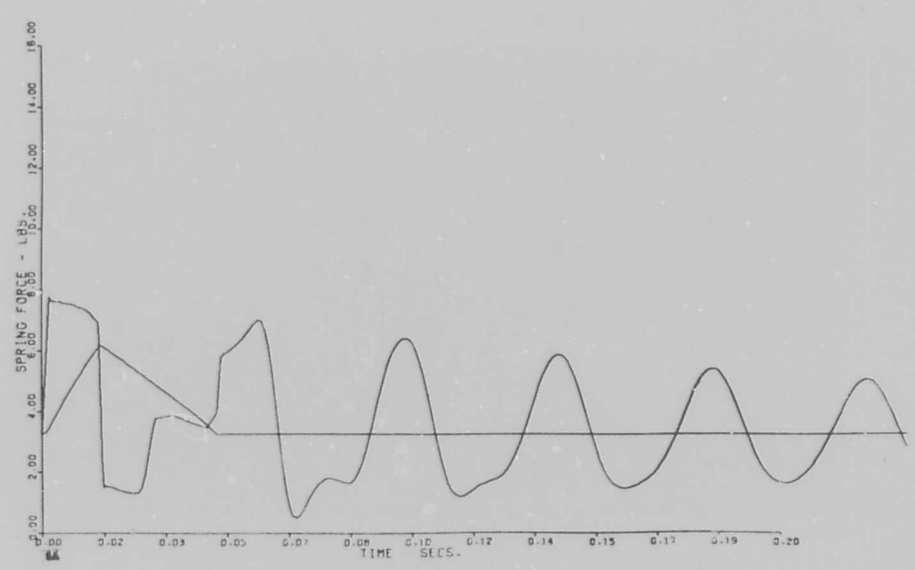


FIGURE 5: DYNAMIC SPRING FORCE

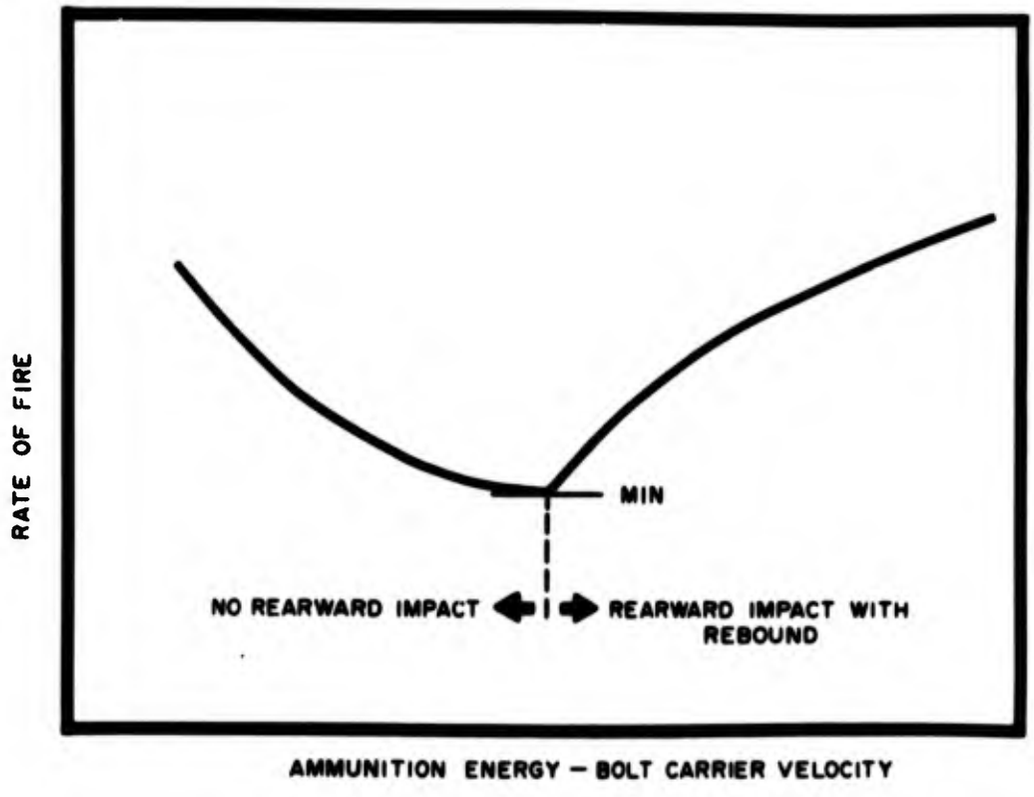


FIGURE 6: RIFLE SIGNATURE CURVE

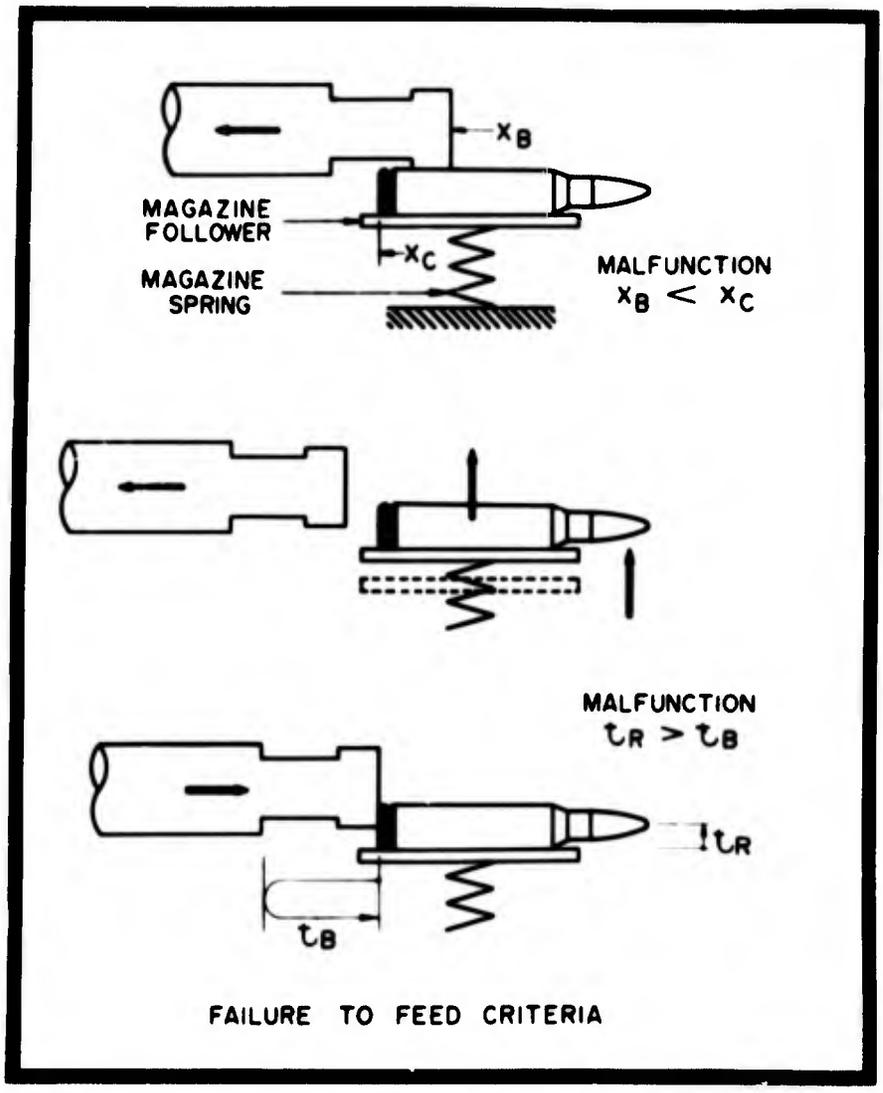


FIGURE 7: FAILURE TO FEED MALFUNCTION

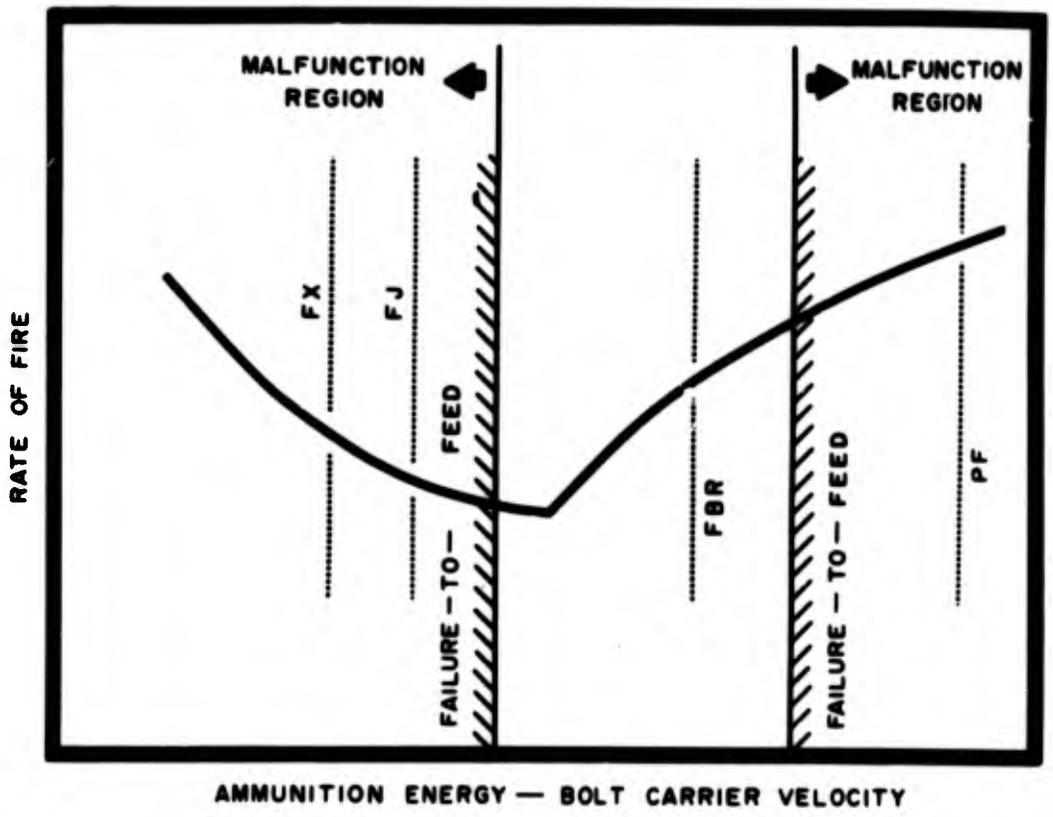


FIGURE 8: RIFLE SIGNATURE/MALFUNCTION CURVE

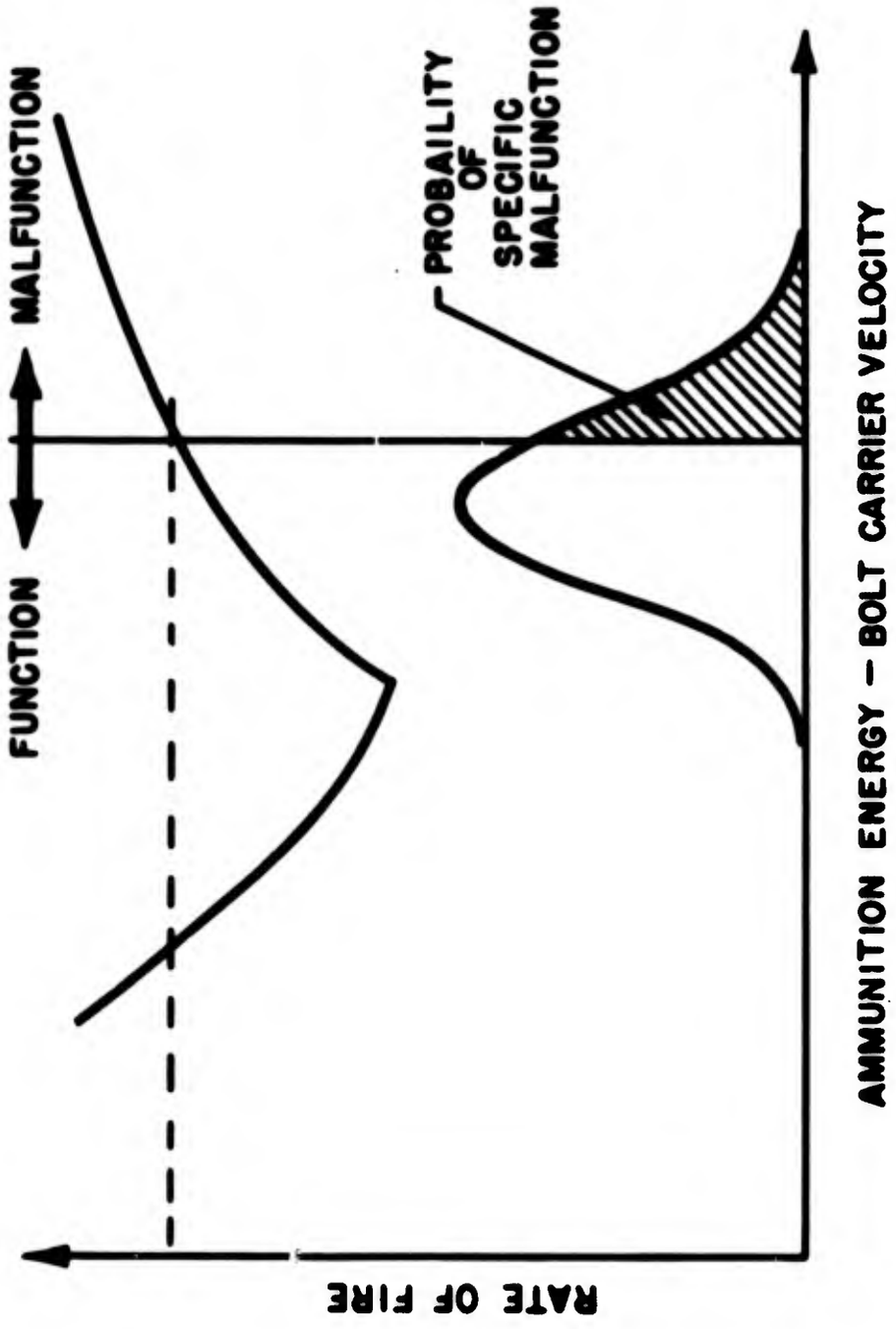


FIGURE 9: SYSTEM MALFUNCTION PROBABILITY

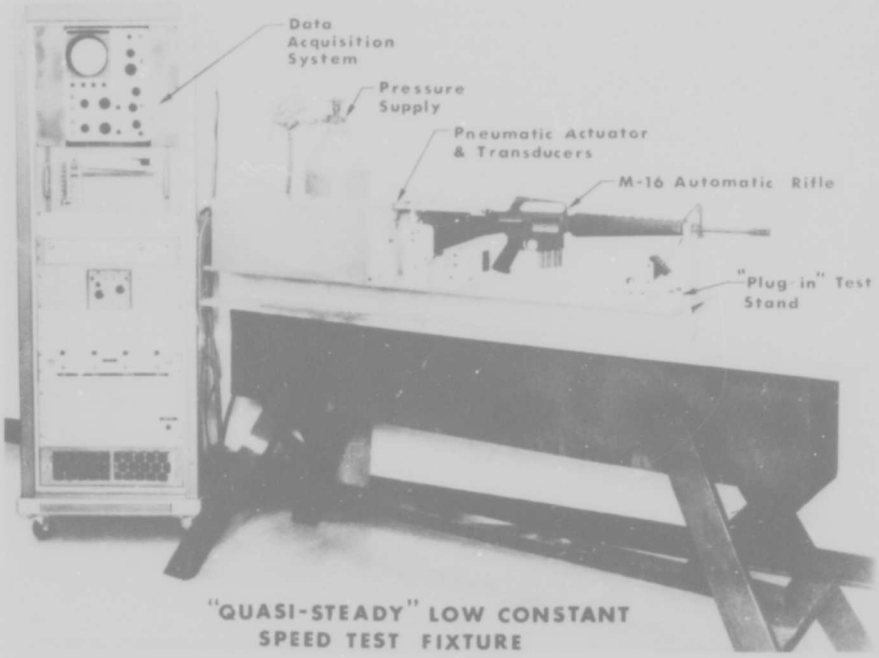


FIGURE 10: STATIC TEST FIXTURE

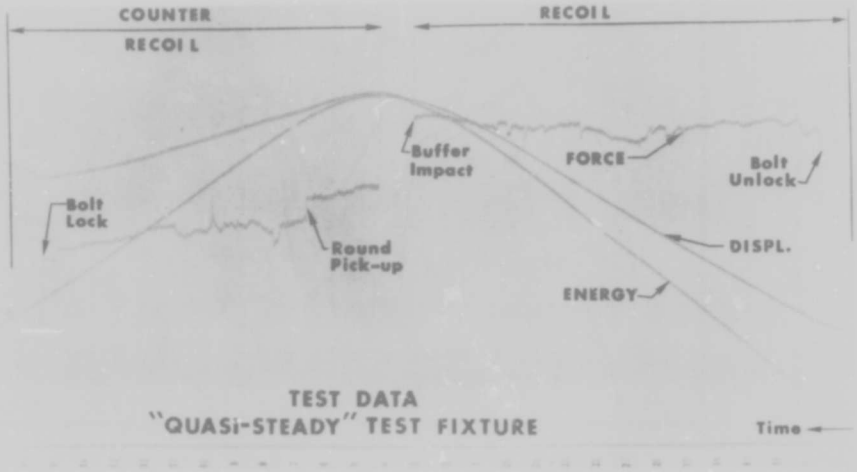


FIGURE 11: FORCE DISPLACEMENT CURVE

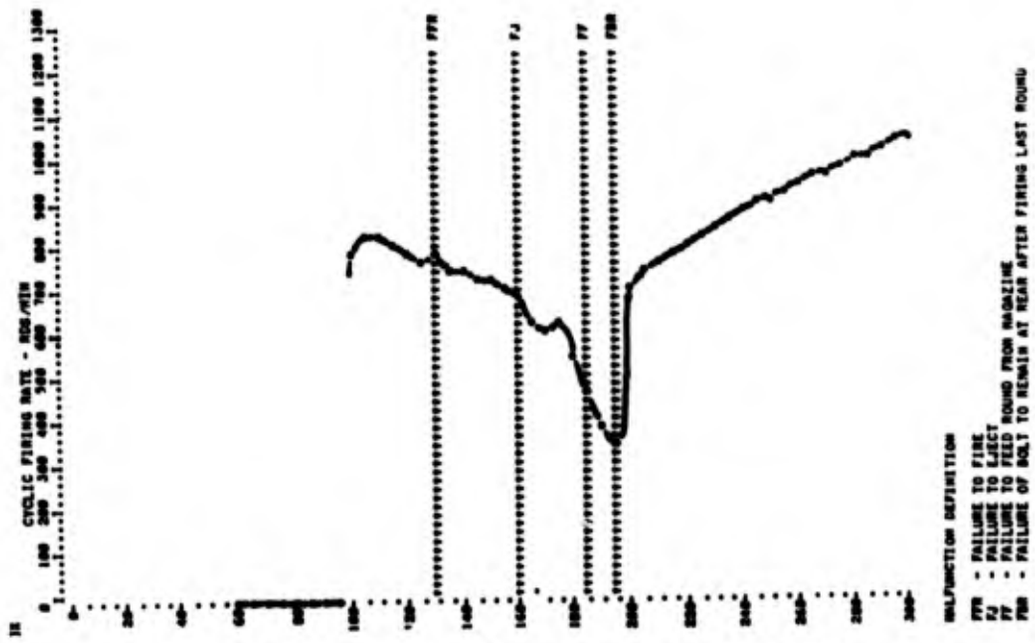


FIGURE 12: TYPICAL RIFLE SIGNATURE CURVE

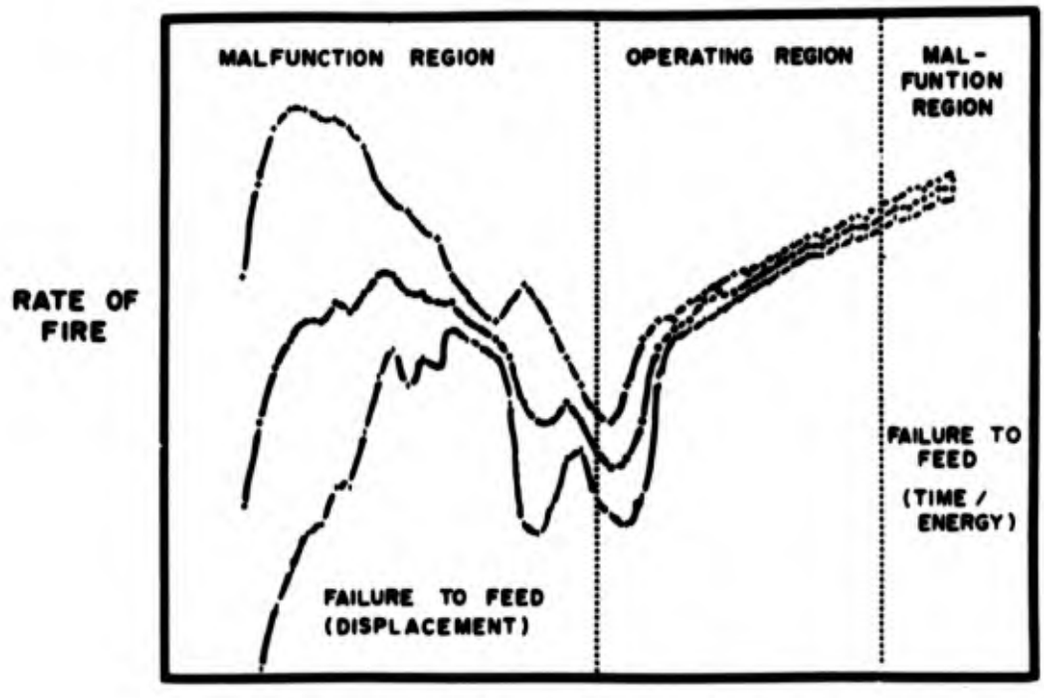
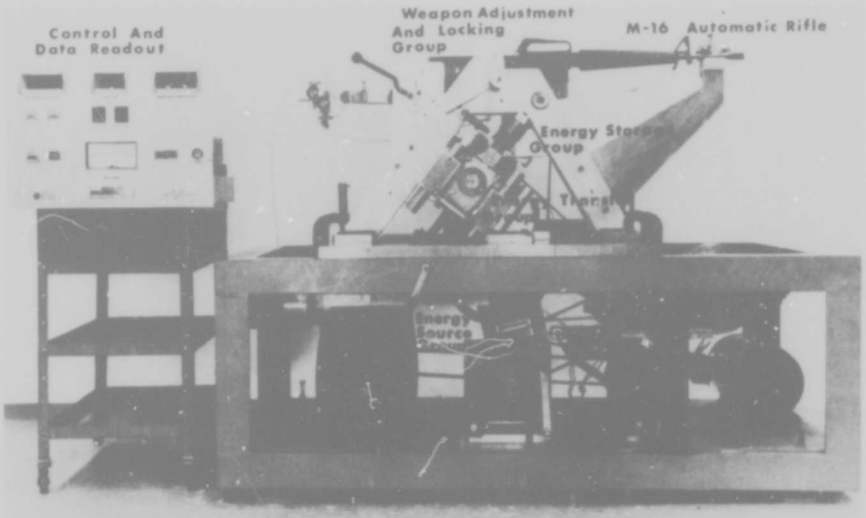


FIGURE 13: SIGNATURE CURVE FOR RIFLE POPULATION



SMALL ARMS
DYNAMIC DATA ACQUISTION SYSTEM

FIXTURE 14: DYNAMIC TEST FIXTURE