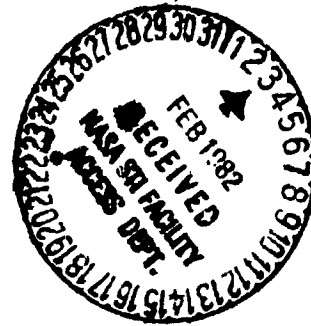STIK (M37)

NASA CONTRACTOR
REPORT

NASA CR-161855

RESEARCH REPCRTS -
1981 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

By Gerald R. Karr, James B. Dozier, Marion I. Kent and B. F. Barfield
University of Alabama in Huntsville
Huntsville, Alabama

and

University of Alabama
University, Alabama

Final Report

January 1982

Prepared for

NASA - George C. Marshall Space Flight Center
Marshall Space Flight Center, Alabama 35812

RESEARCH REPORTS

1981 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

George C. Marshall Space Flight Center
The University of Alabama in Huntsville
and
The University of Alabama

EDITORS:

Dr. Gerald R. Karr
Associate Professor of Mechanical Engineering
The University of Alabama in Huntsville

Dr. James B. Dozier
Director, Research and Technology Office
Marshall Space Flight Center

Mr. Marion I. Kent
Office of University Affairs
Marshall Space Flight Center

Dr. Bob F. Barfield
Professor of Mechanical Engineering
The University of Alabama

1. REPORT NO. NASA CR-161855	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
4. TITLE AND SUBTITLE Research Reports - 1981 NASA/ASEE Summer Faculty Fellowship Program		5. REPORT DATE January 1982
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Dr. Gerald R. Karr, Dr. James B. Dozier, Mr. Marion I. Kent, Dr. B. F. Barfield		8. PERFORMING ORGANIZATION REPORT #
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Alabama in Huntsville, Huntsville, Alabama and University of Alabama, University, Alabama		10. WORK UNIT NO.
		11. CONTRACT OR GRANT NO. NGT01-008-021
12. SPONSORING AGENCY NAME AND ADDRESS National Aeronautics and Space Administration		13. TYPE OF REPORT & PERIOD COVERED Contractor Report June 1, 1981 - August 7, 1981
		14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

This document is a collection of reports on the research conducted by the participants in the 1981 NASA/ASEE Summer Faculty Fellowship Program which was held at the George C. Marshall Space Flight Center (MSFC) for the seventeenth consecutive year. The 1981 program ran from June 1, 1981 through August 7, 1981, and was conducted by the University of Alabama in Huntsville in cooperation with MSFC and the University of Alabama. As in the past, the program was operated under the auspices of the American Society for Engineering Education (ASEE) with sponsorship and funding by the University Affairs Office, NASA Headquarters, Washington, D. C. The University of Alabama in Huntsville and Marshall Space Flight Center conducted one of the seven Aeronautical and Space Research Programs which were funded nationally during the summer of 1981. The other programs were conducted at six other NASA centers throughout the country in cooperation with other universities.

The fellows spent ten weeks working with MSFC personnel on an individual research problem of mutual interest to the fellow and MSFC.

17 KEY WORDS	18. DISTRIBUTION STATEMENT Unclassified - Unlimited J. A. BETHAY Director, Executive Staff, MSFC

19. SECURITY CLASSIF. (of this report) Unclassified	20. SECURITY CLASSIF. (of this page) Unclassified	21. NO. OF PAGES 576	22. PRICE NTIS

Preface

This document is a collection of technical reports on the research conducted by the participants in the 1981 NASA/ASEE Summer Faculty Fellowship Program at Marshall Space Flight Center. This was the seventeenth consecutive year the program has been conducted at MSFC. The 1981 program was administered ι·· the University of Alabama in Huntsville (UAH) in cooperation witι MSFC and the University of Alabama. The program was operated under the auspices of the American Society for Engineering Education (ASEE) with sponsorship and funding by the University Affairs Office, NASA Headquarters, Washington, D.C. The program conducted by the University of Alabama in Huntsville and Marshall Space Flight Center was one of seven such programs which were funded, the other six programs were conducted at the six other NASA centers. The basic objectives of all the programs were:

a. To further the professional knowledge of qualified engineering educators.

b. To stimulate an exchange of ideas between participants and NASA.

c. To enrich and refresh the research and teaching activities of participants and,

d. To contribute to the research objectives of the NASA centers.

The Faculty Fellows spent ten weeks (June 1 through August 7, 1981) at MSFC working with MSFC technical personnel on individual research problems of mutual interest to the fellow and MSFC. This document is a compilation of the work performed. A separate report (UAH Report No. 289, September, 1981) reports on the administrative aspects of the program. Further information on the 1981 program can be obtained by contacting any of the editors.

TABLE OF CONTENTS

First-Year Fellows

XVII. Payne, James E.: "The Design and Evaluation of Superconducting Connectors"

XVIII. Payne, Linda L.: "A Study of Optically Contacted Quartz at Cryogenic Temperatures"

XIX. Ray, John R.: "Matter in General Relativity"

XX. Saylor, Annie V.: "Surface Contour of a Biaxially Strained, Doubly Curved Antenna Reflective Mesh"

XXI. Six, N. Frank: "Ray Tracing Studies of Jupiter's Magnetosphere"

XXII. Tan, Arjun: "Thunderstorm Electric Field Modelling"

XXIII. Williams, Alton C.: "Determination of the Efficiency of A Proportional Counter for Detecting X-Rays"

Second-Year Fellows

XXIV. Antar, Basil N.: "Solidification of a Binary Mixture"

XXV. Brindley, Thomas A.: "Assessment of MSFC's Supervisory Training Programs and Courses"

XXVI. Brown, Richard C.: "Digital Compensation Techniques for the Effects of Time Lag in Closed-Loop Simulation Using the 6 DOF Motion System"

XXVII. Chan, Chia-Hwa: "A New SPS Antenna Design to Reduce Grating Lobes"

XXVIII. Davis, Jack: "Aluminum and Chromium Ion Plating Studies for Enhancement of Surface Properties"

XXIX. Doss, Deva C.: "Statistical Analysis of Daily Mean Temperatures"

XXX. Foster, Larry M.: "Surface Control Temperatures for the Bridgman-Stockbarger and Float Zone Techniques"

XXXI. Jandebeur, Thomas S.: "Design Consideration of a Thermally Stabilized Continuous Flow Electrophoresis Chamber II"

XXXII. Mathis, Frank H.: "Mathematical Programming Techniques for Scheduling Spacelab Crew Activities and Experiment Operations"

XXXIII. Shand, Julian B, Jr.: "A Study of the Relativistic Rise In Ionization As Measured by Ion Chambers of Different Design"

XXXIV. Temple, Enoch: "An Investigation of the Uniform Random Number Generator"

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

DEVELOPMENT OF TOUGHENED EPOXY POLYMERS FOR HIGH
PERFORMANCE COMPOSITE AND ABLATIVE APPLICATIONS

Prepared by: Vernon R. Allen, Ph. D.

Academic Rank: Professor

University and Department: Tennessee Technalogical University
                                       Department of Chemistry

NASA/MSFC
      Division: Non-Metallic Materials
      Branch: Polymers and Composites

MSFC Counterpart: Dr. W. J. Patterson

Date: August 21, 1981

Contract No.:   NGT 01-008-021
                                       The University of Alabama in Huntsville

DEVELOPMENT OF TOUGHENED EPOXY POLYMERS
FOR HIGH PERFORMANCE AND ABLATIVE APPLICATION

BY

VERNON R. ALLEN
PROFESSOR OF CHEMISTRY
Tennessee Technological University
Cookeville, Tennessee 38501

ABSTRACT

To facilitate a cause-effect study of the chemical modification
of epoxy polymers, a survey was extended of current procedures for
the assessment of state of cure in epoxy polymers and for the evaluation
of polymer toughness as related to nature of the crosslinking agent.
Various conformations of sample morphology were examined to identify
testing variables and to establish optimum conditions for the selected
physical test methods. Dynamic viscoelasticity testing was examined
in conjunction with chemical analyses to allow observation of the extent
of the curing reaction with size of the crosslinking agent the primary
variable.

I-ii

INTRODUCTION

Space applications of load-bearing composite structures require high strength, high modulus polymers capable of withstanding the impr··c extremes in mechanical stresses and thermal variations. Similarly, adhesives designed for ablative bonding must be expected to resist higher thermal gradients than are normal for non-space utilization. Although much effort has been directed toward producing exotic specialty resins capable of retaining high mechanical strength properties to very high temperatures, such characteristics quite often are accompanied by undesirable preparation and processing extremes.

The epoxy resins, by contrast, are notable for processing simplicity and a broad range of fabrication conditions. Though unmodified epoxy systems are relatively brittle compared with the bonded (metal) substrates, high impact strength elastomer-modified epoxy resins have been part of adhesives technology for two decades. In such systems the resistance to crack propagation yielding high fracture strength characteristics, up to twenty times higher than values for the unmodified brittle resin, has been shown (1) to result from phase separation during network formation with particle size of the elastomer phase (2) being the primary determinant of strength variation. An even greater increase in resin toughness was reported by Riew, Rowe and Siebert (3) on addition of bisphenol A to a standard rubber-modified epoxy system. This increase in fracture surface work was attributed to formation of a binodal distribution of heterophase particles dispersed throughout the glassy epoxy matrix.

The higher impact strength of such rubber-modified epoxy resins is not achieved without reduction in other properties, such as modulus and heat distortion temperature (4), obviating the more desirable characteristics when considered in light of the harsh environmental conditions concomitant with space application design. It was, therefore, proposed (5) to modify the basic epoxy resin system by varying the structure of the hardener in an homologous fashion. It is expected that, by using a series of poly-functional chain extenders in which an increase in chain length between the bifunctional amine groups is afforded by incorporation of biphenyl links (starting with phenylene diamine), aromatic interaction may generate (heterophase) regions of high density. Such nodular formation will not only enhance the specimen toughness but should also offset any reduction in thermal properties resulting from network dilatency as the more bulky crosslinking agents are employed.

Structure modifications of this type necessarily involve problems associated with meaningful interpretation of comparison data sets which establishes the basis of the sub-project as reported herein. Specifically the aims of this project are twofold: (1) to consider the experimental variables associated with development of "extent of cure" analysis, and (2) to assess methodology of fracture energy determination and to prescribe a meaningful and reproducible procedure. The following is separated into these two categories for ease of presentation.

A DETERMINATION OF THE EXTENT OF CURE

All efforts to establish an unambiguous procedure for identification of the extent of reaction have been required to treat independently the reaction kinetics preceding and following sample gelation. As reported by Erath and Spurr (6) using electron microscopy, dense aggregates (micelles) are formed during gelationof epoxide polymers. These aggregates, which are the active centers for continued polymerization, are embedded in the matrix of lower crosslink density. Further polymerization in these micellar regions becomes diffusion controlled with an expected higher energy of activation. Indeed, Warfield and Petree (7) have attributed the high rates of polymerization exhibited by an Epon 828 epoxy resin hardened with diethylene triamine to the existence of a high concentration of micelles. Thus, in the rapidly polymerizing systems, the larger number of micelles present would necessarily result in a solid (gelled) network characterized by a higher micelle-to-matrix ratio and, hence, higher activation energies for ion conduction.

An alternative explanation for the observed higher activation energies for ionic transport may be deduced in part from the derived relationship between the electrical resistivity and the microviscosity of polymeric systems (8). The higher microviscosity of the polymerizing medium at the lower temperatures employed in epoxy resins hardened with (more reactive) aliphatic amines would lower diffusion rates such that adiabatic regions (micelles) could result in which the exothermic reactions would be self-catalyzing and also self-propagating, given the low thermal conductivity of the system.

Micellar micro-dense regions in crosslinked epoxy systems may be associated with the extent of epoxide conversion, as measured by infrared analysis (9, 10) and by chemical methods (10), being indicative of less than complete amine-epoxide reaction, even for extended cure times. It is probable that the catalytic activity (11) of the hydroxyl group - produced b, epoxide reaction -

$$R\text{-CH-CH}_2 + R'NH_2 \longrightarrow R\text{-CH-CH}_2\text{-N-R}' \quad (1)$$
$$\quad\quad O \quad\quad\quad\quad\quad\quad\quad\quad OH \quad\quad H$$

would be decreased as the hydroxyl group is increasingly "locked up" in the dense micellar region.

As the curing temperature is increased, the higher specific volume (lower microviscosity), the decreased effect of the adiabatic regions, and the higher diffusion rates should reduce the tendency to micellar formation, producing a more homogeneous network. By using a time-temperature shift to superimpose the (pregelation) maximum reaction rates, Lunak, Vladyka and Dusek (12) demonstrated clearly that (micellar-induced) diffusion-controlled retardation occurred at lower extents of conversion and concurrently limited the maximum extent of conversion at the lower polymerization temperatures. The glass-transition temperature (T_g) has also been reported to be directly associated with the reaction temperature (T_c) by Fisch, Hofmann and Schnid (12). They found that the difference between the glass temperature and the cure temperature, (T_g - T_c) was about thirty degrees, over the range of 60 - 120 C for T_c, and decreasing asymtotically toward zero as T_c approached the T_g (maximum) of 165 C for an epoxy system crosslinked with an aromatic diamine.

Aspbury and Wake (13) confirmed the existence of supermolecular structures in cured epoxy resins and proposed that the small heterophase regions 1.3 to 2.1 nm diameter - were probably due to stacked residues of bisphenol A. Larger nodular regions - ca. 5 nm - were observed to increase in size to 40 nm on postcuring. The particle size distribution was noticeably shifted to smaller particle size at higher temperature of cure. It was also observed that microtomy of the smaller particle size samples (higher T_c) was considerably more difficult. Since microtomy is a fracture process, this observation is in agreement with the results presented by Gledhill and Kinlock (14) that post-cured samples exhibited a higher fracture energy than samples having the same tensile strength and T_g, which were not post-cured. From this, Aspbury and Wake proposed that fracture strength may be directly related to aggregate size. Similarly, Sultan and McGarry (15) have published data on rubber-modified epoxy resins that show increased fracture strength with increased particle size.

Based on the preceding interplay of curing conditions and mechanical properties, it became evident that this extent of reaction study should be primarily concerned with establishing the onset of diffusion control as well as the cure time required to produce reproducible mechanical properties, insensitive to short-term measurement conditions. To this end this aspect of the project has been concentrated on electrical resistivity measurements (16-18), on differential scanning calorimetry (18-25), and on dynamic mechanical analysis (26-30).

B. MEASUREMENT OF FRACTURE STRENGTH

An indepth review of fracture surface work theory and methods is included in the paper by Broutman and McGarry (31) in which several methods for fracture of thermoplastic materials is included in (their) Table I along with comparative fracture surface energy data. According to Bucknall and Yoshii (32) the "state-of-the-art" revealed many factors which influence fracture strength of toughened epoxy resins:

   i)   Solubility parameter and content of rubber.
   ii)   Nature and content of the hardener.
   iii) Content of resin extender.
   iv)   Nature of the end-group in the rubber phase.
   v)   Curing temperature and extent of post-cure.
   vi)   Rubber phase particle size and particle size distribution.

All of these except factor iv) were discussed in the preceding section and will not be considered again here. Another factor which was not included by Bucknall and Yoshii and which is probably more relevant to the current project is the method of measurement of fracture stress. Marshall, Culver and Williams (33), in a study of the fracture characteristics of PMMA, compared three standard test methods, namely notched tension, parallel cantilever beam cleavage, and tapered cantilever beam cleavage. They concluded that the tapered beam analysis is superior to the other two because the measurement of crack speed can be achieved with greater precision and over a broader range of test condiions. As a result of this comparative study they showed that the critical stress field intensity factor for all three methods was single-valued within experimental error over about six decades of testing speed. Likewise, Meeks (34) used both the tapered cantilever cleavage and the single -edge-notch tension fracture toughness tests and found them to yield similar results for several epoxy resin systems. A thorough analysis of the methods of calculations of fracture energy for these test methods is included.

A somewhat different approach has been used by Tattersall and Tappin and by Davidge and Tappin (35, 36). They examined the fracture behavior of several materials, aluminum, PMMA, glass and graphite using a notched retangular bar as test specimen stressed in a bending mode with three pressure points. The data analysis proved relatively straight forward if the work of fracture was assumed to be totally involved in formation of the crack. Thus the work of fracture was calculated directly from the area within the load-deflection curve. Based on the ease of sample preparation and on the direct nature of the calculation, this method had beensdected as the procedure of choice for this project. Initial efforts at measurement had resulted in scatter in the data which questioned the correctness of this choice. Further

examination of current methodology revealed, in a paper by Mostovoy, Ripling and Bersch (37), a test method which does not require knowledge of crack length or of sample compliance. The particular specimen is designed such that only the load required for fracture and the tensile modulus of the sample be known to allow calculation of fracture energy. Even though this tapered double cantilever beam requires considerable machining to reduce the variation in sample compliance to the point that crack length is not a factor, the advantage of being able to obtain several data points from a single sample made this method appear worthy of experimental effort. Thus, the second aim of this project, the development of a fracture method suitable for comparative analysis, has been directed toward increased precision in the 3-point beam loading method and examination of the modified tapered cantilever beam cleavage method.

EXPERIMENTAL

In this section there is described the materials used in this work together with the methods of analysis and of fracture strength determinations.

Materials : The resin used throughout the work was Shell Epon 828, a low mo lecular adduct of bisphenol A with diglycidyl ether having an epoxide equivalent of 5.1 equivalent per kilogram. The hardener, Shell Z, a mixture of m-Phenylene diamine and methylene dianiline, was added in stoichiometric amount. Both reactants were preheated to 60 - 70 ^{o}C, mixed vigorously by stirring, evacuated several minutes under fore pump vacuum, and poured into preheated aluminum molds which had been sprayed with Teflon mold-release agents. Samples for cure studies were poured into aluminum pans. The resin was allowed to set at room temperature for around sixteen hours, and then cured for variable times with a standard cure for mechanical analysis of 2 hours at 100° C and 2 hours at 150° C. Samples were removed from the molds after cooling, machined or ground to the desired form and polished with 600 sand paper.

A rubber modified sample was prepared by mixing into the above system 20 phr of WC-8006 (Wilmington Chemical Co.), an epoxy-terminated acrylonitrile - buta diene rubber, and given the same cure cycle.

Equipment: For calorimetry data the Perkin-Elmer DSC-2 equipped with a computer-assisted data analyzer and printout system. A blank sample holder was used in the reference pan. Electrical resistivity analysis was determined using a Hewlett-Packard 4329 A High Resistance Meter with logarithmic ohm readout. The resin was cast from a 30 w/o solution in t etrahydrofuran onto a sample cell consisting of one inch long copper leads 0.030 inch in width separated by a 0.035 inch gap and mounted on a teflon-coated circuit board. After air-drying, the cell was suspended in a Fisher Isotemp Model 350 circulating air oven. Dynamic mechanical analysis data were collected using a Toyo Industries Rheovibr onModel DDV-III-C with temperature regulated forced-air oven. Glass transition temperature data were also obtained using a Tinuis-Olson thermal expansion apparatus. Fracture data were collected on a floor-model Instron Tensile Testing machine.

RESULTS AND DISCUSSION

Evaluation of the curing reaction by resistivity was limited by time restrictions but the data plotted in Figure 1 show clearly that the technique gives a very direct means of in situ examination of the crosslinking reaction. Kinetic analysis of such measurements, as made by Warfield and Petree (17), reveal that this method of cure study can obtain both an activation energy for the conversion reaction as well as a separate activation energy for ion transport. The latter could be quite meaningful in correlation with phase separation studies.

Based on the evidence as reported in the literature, it is indicated that the calorimetric analysis using the DSC is the ideal method for study of the crosslinking reaction. Thus, Prime and Sacher (20) compared resistivity and infrared data with the dynamic trace of rate of heat evolution and concluded that the DSC trace could be broken down and associated with the predominant reaction of each stage. The differential power output of the DSC was plotted versus the scan temperature (his figures 9 & 10) by Fava (19) showing clearly the disappearance of the post-cure exotherm followed by a slight decrease in T_g, which may reflect subtle changes in the heterophase structure (or may simply be a result of the non-equilibrum nature of the measurement). Efforts to reproduce this set of data for our epoxy system was constantly stymied by equipment failure.

Better results were found in the measure of the glass-transition temperature on one-inch long rectangular rods ($\sim 1/4$ inch square) by measure of the coefficient of linear expansion versus temperature. Typical data, shown in Figure 2, were obtained by heating the glass-cylinder enclosed sample using an oil-bath with a heating rate of 2- 3 degrees per minute. The difference in T_g between the two samples is as expected for straight epoxy versus the rubber modified since the rubber serves as a diluent yielding the lower cross-link density. Further effort along these lines were discontinued in light of the successes achieved with the dynamic mechanical studies.

The samples used in determination of the dynamic loss angle, tan δ, were cut into the rectangular shape of 3.5 inches by 0.5 inches by 0.125 inches and polished with sand paper to obtain a dull finish. Since absolute values of the sample modulus was not needed, no effort was made to make each sample the same. Samples of resin cast into the aluminum pans were cured for one and for two hours at 100°C and for one and two hours at 150°C.

The plots of tan δ for these cured specimen are presented in Figure 3 for the unmodified resin and in Figure 4 for the rubber extended system. After heating to 170° C the samples were maintained at 172 - 174°C for ten minutes and then cooled slowly to approximately 100°C - both heating and cooling rates were about 3 degrees per minute. The enclosed area should represent the continued curing at the elevated temperature for the heating and the cooling. At a selected value of the tan δ it can be used to represent

I-8

continued cure. Such a determination of $\triangle$ (tan δ) can be plotted versus cure temperature and would asymtotically approach zero much as the % residual heat of reaction versus cure temperature as reported by Fisch, Hofmann and Schnid (12) from DSC analysis (their Figure 6).

A study of the fracture strength of the two exoxy systems was made according to the procedures as reported Davidge and Tappin (36) and by Mostovy, Ripling and Bersch (37). The three-point beam loading of pre notched (but not precracked) rectangular samples were examined for both the pure epoxy - 828/Z - and the rubber-modified system - 828/Z/8006. Although there is scatter in the value of the fracture surface energy obtained by this method, the higher fracture strength of the rubber-toughened system is clearly obtained. The results using the tapered double cantilever beam were not as promising because the crack could not be retained in the grooved region even for samples in which the stress thickness was only one-third of the total beam thickness. This was expecially true for the rubber-modified samples, making suspect the obviously higher fracture strength results. Similar difficulties had been found using parallel-sided double cantilever beams.

CONCLUSIONS AND RECOMMENDATIONS

Even though the research program failed to yield either of the objectives originally sought, the specific goals are now clearly identified. It seems safe to conclude that a combination of in situ cure reaction rate studies of electrical resistivity and infrared analysis would yield sufficient information to design a workable cure schedule. Either calorimetry or dynamic viscosity can be used to reinforce the resistivity and IR interpretation.

It is quite obvious that a thorough examination of the toughening of epoxy systems should include not only fracture analysis data - the three - point loading system seems adequate - but also an effort should be made to include phase - separation studies using the transmission or scanning electron microscopy.

LIST OF REFERENCES

1. McGarry, F., J. Proc. Roy. Soc. London A-319. 59 (1970).
2. Sayre, J.A., Assink, R.A. and Lagasse, R.R., Polymer 22, 87 (1981).
3. Riew, C.K., Rowe, E.H. and Siebert, A.R., p. 326 in "Toughness and Brittleness of Plastics" Deavin and Dregnola, Ed., Adv. Chem. Series #154, ACS.
4. Selby, K. and Miller, L.E., J. Mat. Sci. 10, 12 (1975).
5. Patterson, W. J., Private communication, June, 1981.
6. Erath, G.H. and Sperr, R. A., J. Polymer Sci.. 35, 391 (1959).
7. Warfield, R.W., and Petree, M.C., SPE Trans. 1. 3 (1961).
8. Purdon, J.R. and Morton, M., J. Polymer S . 57, 453 (1962).
9. Lewis, A.F. and Ramsey, W.B., Adhesives Age 9, 20 (1966)/
10. Dannenberg, H. and Harp-Jr., W.R., Anal. Chem. 28, 86 (1956).
11. Shechter, L., Wynstra, J. and Kurkjy, P., IEC 48, 94 (1956).
12. Fisch, W., Hofmann, W. and Schmid, R. J. Appl. Poly. Sci. 13, 295 (1969).
13. Aspburg, P.J. and Wake, W.C. Brit. Poly. J. 11, 17 (1979).
14. Gledhill, R. A. and Kinloch, A. J., J. Mat. Sci. 10, 1261 (1975).
15. Sultan, J.N. and McCarry, F.J., Poly. Eng. Sci., 13, 27 (1973).
16. Adamec, V., J. Appl. Poly. Sci. 10, 1277 (1972).
17. Warfield, R. W. and Petree, M.C., Makromol. Chem. 58, 139 (1962).
18. Acitelli, M.A., Prime, R.B. and Sacher, E., Polymer, 12, 335 (1971).
19. Fava, R.A., Polymer 9, 137 (1968).
20. Prime, R. B. and Sacher, E., Polymer 13, 455 (1972).
21. Barton, J.M., Makromol. Chem. 171, 247 (1973).
22. Kamal, M.R., Polym. Eng. Sci., 14, 231 (1974).
23. Sourbour, S. and Kamal, M.R., Thermochim. Acta 14, 41 (1976).
24. Richardson, M.J. and Savill, N.G., Brit. Polymer J., 11, 123 (1979).
25. Barton, J. M., Polymer 21, 603 (1980).
26. Kaeble, D. H., SPE Journal 15, 1071 (1959).
27. May, C.A. and Weir, C. A., SPE Trans., 2, 207 (1962).
28. Kwei, T.K., J. Poly. Sci. A-2, 1, 943 (1966).
29. Babayevsky, P. G. and Gillham, J.K., J. Appl. Poly. Sci,, 17, 2067 (1973).
30. Pogany, G.A., Eur. Poly. J., 6, 343 (1970).
31. Broutman, L.F. and McGarry, F. J., J. Appl. Poly. Sci., 9, 589, (1965).
32. Bucknoll, C. B. and Yoshii, I., Brit. Polymer J., 10,53 (1978).
33. Marshall, G. P., Culver, L. E. and Williams, J.G., Plastics and Polymers, (Feb.) 75 (1969).
34. Meeks, A.C., Brit. Polymer J. 7, 1 (1975).
35. Tuttersall, H. G. and Tappin, G., J. Mat. Sci., 1, 296 (1966).
36. Davidge, R.W. and Tappir, G., J. Mat. Sci., 3, 165 (1968).
37. Mostovoy, S., Ripling, E.J. and Bersch, C.F., J/ Adhesion 3, 125 (1971).
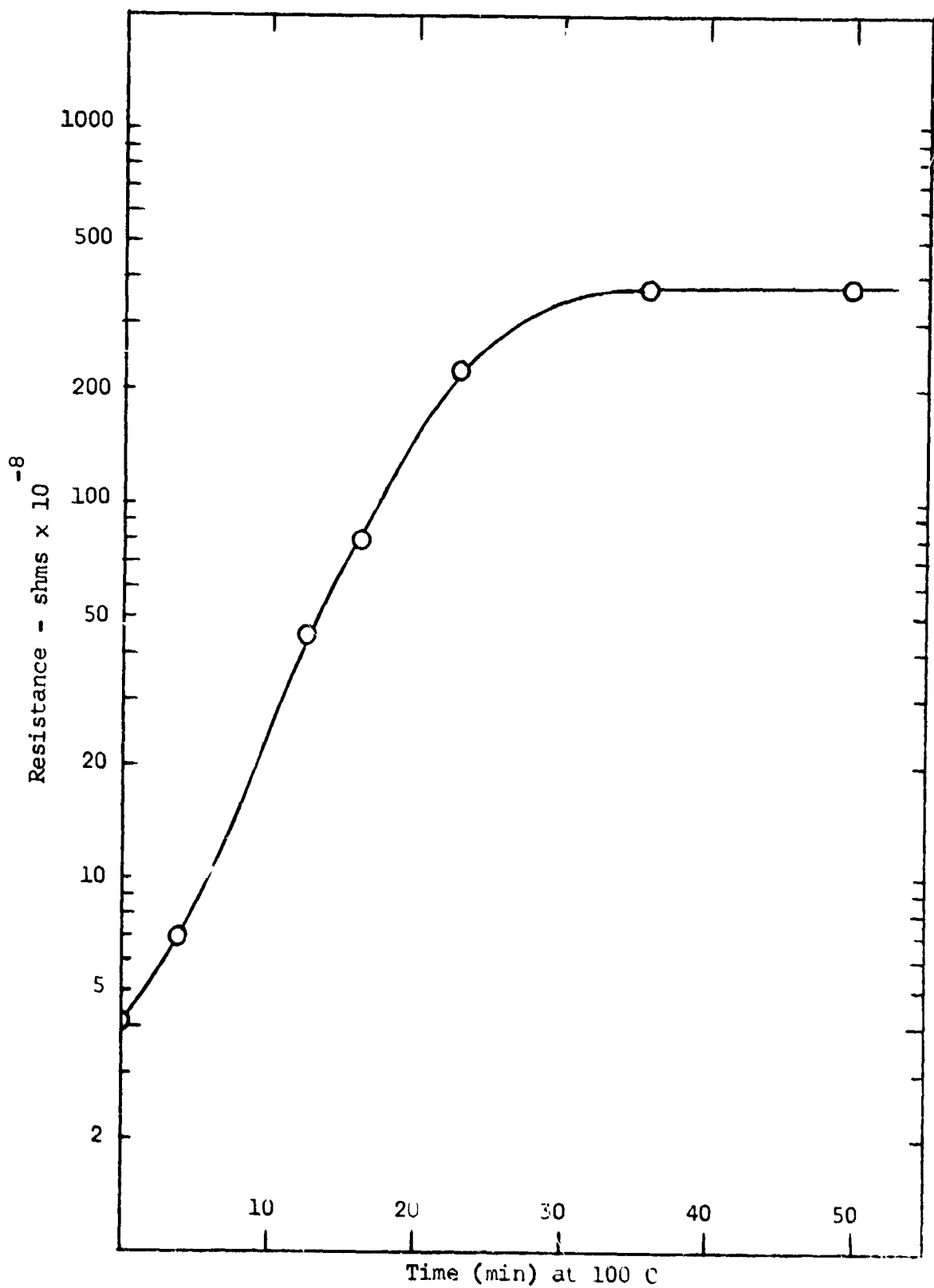
Figure 1. The dependence of the electrical resistivity (at 500 volts) on the cure time at 100 C for a sample of Epon 828/Shell Z.
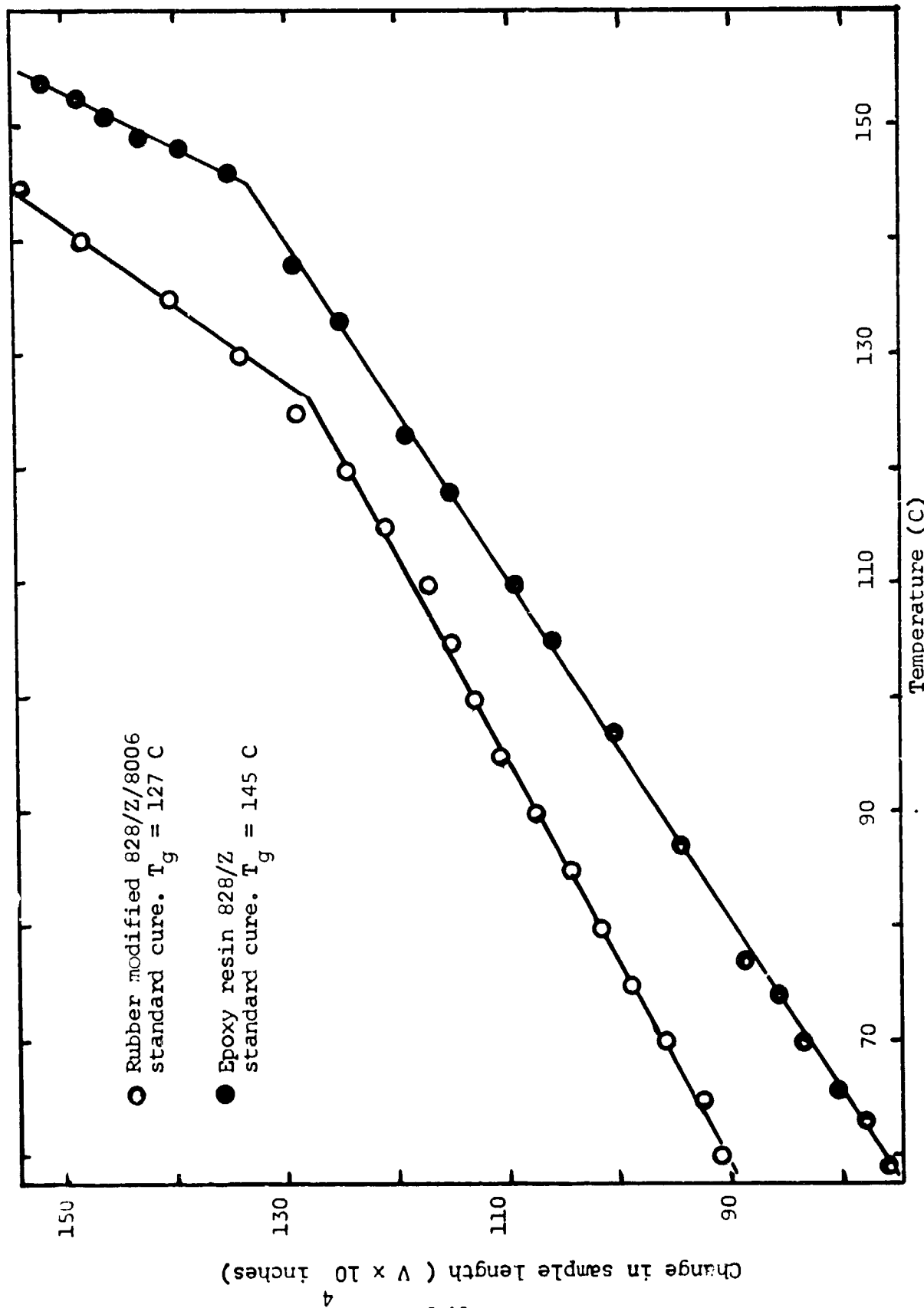
Figure 2. Linear expansion versus temperature with the glass transition temperature taken as the intercept of the glassy and rubbery slopes.

I-13

● Cure 2 hours at 100 C. Cooled immediately.

● Cure 2hours at 100 C and 2 hours at 150 C. Held at maximum temperature 10 minutes.

—— Heating curve.

---- Cooling curve.

Temperature (C)

tan δ

I-14

Fig ... Exter... cure of unmodified resin 828/Z determined by hysteresis in tan δ with the heating cycle of measurement.
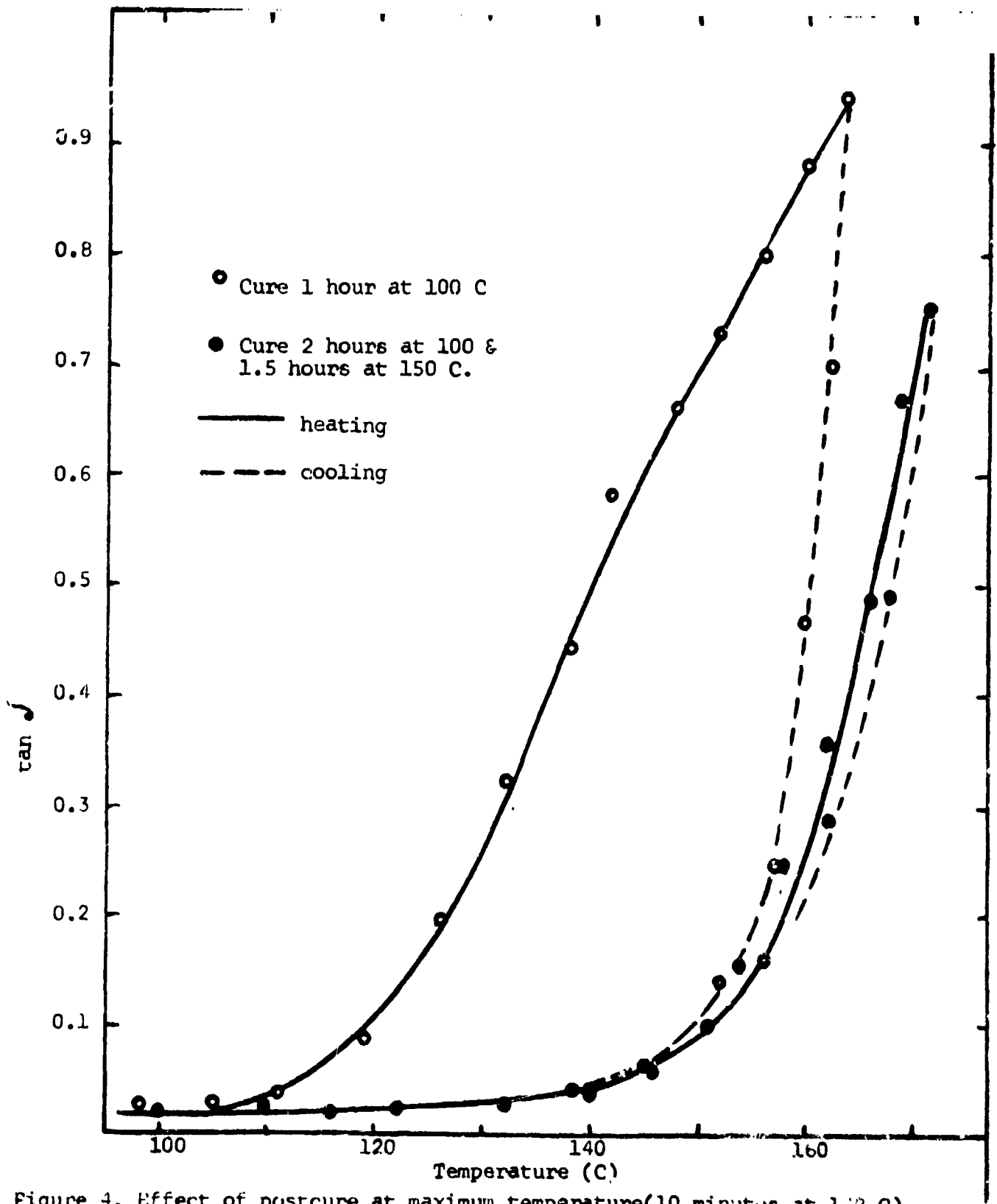
Figure 4. Effect of postcure at maximum temperature(10 minutes at 1'2 C) for unmodified epoxy resin 828/Z.

I-15

NASA-ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

A DETERMINATION OF THE CURRENT DENSITY IN ELECTRON BEAMS

Prepared By: Robert J. Beil, Ph.D.

Academic Rank: Associate Professor

University and Department: Vanderbilt University
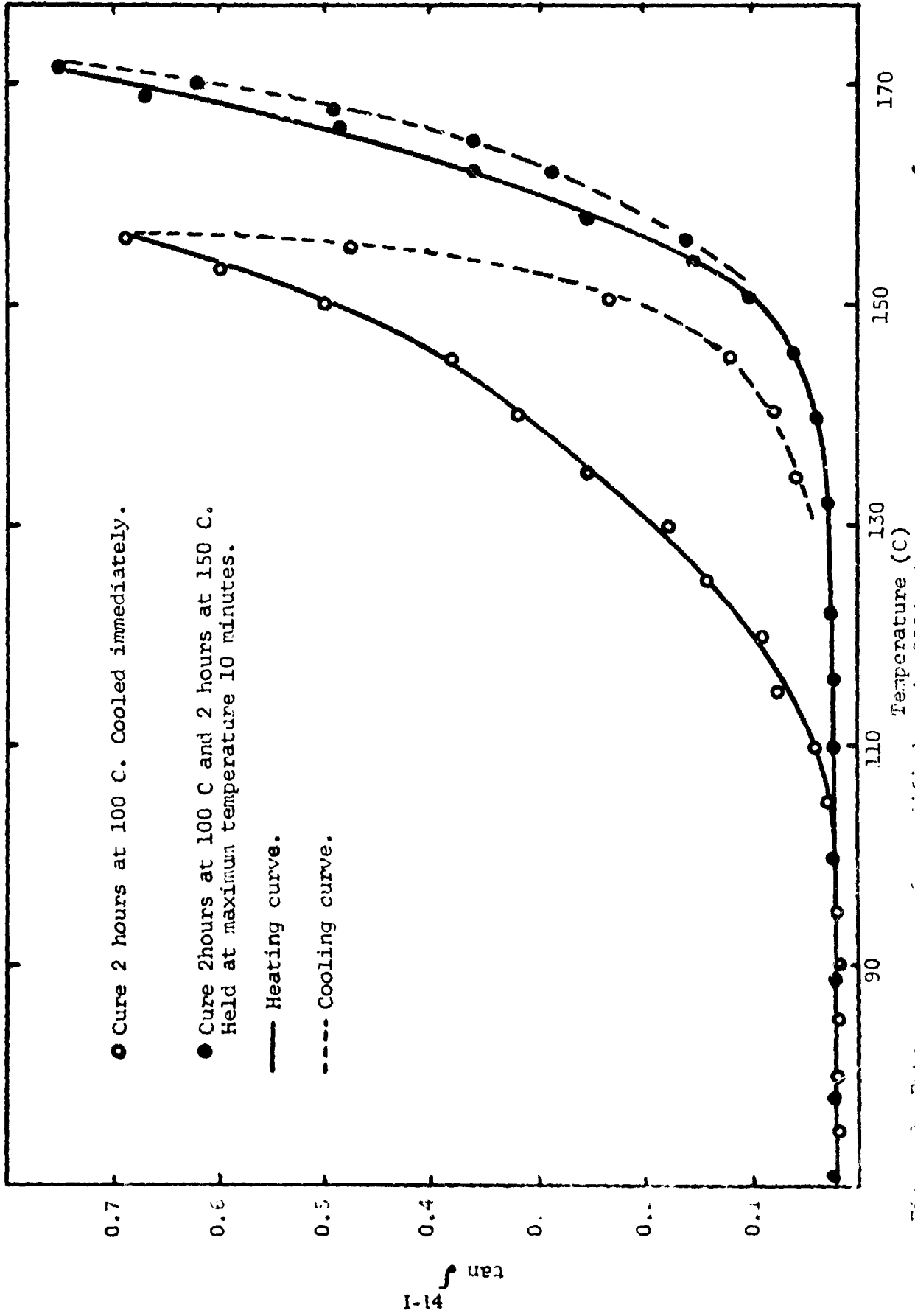   Department of Civil and Environmental
                                        Engineering

NASA/MSFC:
    Laboratory: Materials and Processes
    Division: Process Engineering
    Branch: Metals Processes

NASA Counterpart: Arthur C. Nunes, Ph.D.

Date: August 7, 1981

Contract No.:   NGT 01-008-020
                                      The University of Alabama in Huntsville

ACKNOWLEDGEMENTS

A DETERMINATION OF THE CURRENT DENSITY IN ELECTRON BEAMS

BY

Robert J. Eeil
Associate Professor of Engineering Mechanics
Vanderbilt University
Nashville, Tennessee

ABSTRACT

Current gathering rotating probe techniques can be used to examine the envelope shape and power density profile of electron beams used in electron beam welding devices. Electron beam power density contours determine the shape of the weld vapor cavity and thus control penetration, local heat distribution, etc. In short, the structure of electron beam welds is very sensitive to the power/current density contours of the beam, and a knowledge of these contours is often needed by both researchers and practical welders.

This paper provides the mathematical analysis, consistent with a rotating probe technique, necessary to determine the current density distribution (assumed symmetrically radial) in a cross-section of the beam. An explanation of the experimental technique for obtaining data, a BASIC language computer program to determine the current density from the data, and a study indicating the level of confidence to be associated with results obtained are provided. An example of the application of the analysis to some experimental electron beam data is included.

INTRODUCTION

Weld speed of and power input to the welding process contribute to the temperature history in the heat affected zone surrounding the weld. They are factors in determining the microstructure remaining after the welding process. Hence, they are factors in determining the strength of the welded work.

How speed and power input contribute in electron beam welding can be visualized by considering Figure 1.

Deep penetration is enhanced, in high power well focused electron beam welding, by vapor pressure which forces a deep depression in the liquid surface. Depth of penetration and the slope of the advancing interface of the weld cavity are dependent upon the local current density of the electron beam which impinges on the cavity wall.

Thus, a predictive model which allows calculation of the geometrical shape and size of the weld and calculation of the history of the temperature distribution in the heat affected zone and, consequently the resulting microstructure, depends on knowing the current density distribution within the electron beam. Conversely, characteristics of a desired microstructure in the heat affected zone may be used to dictate the operating characteristics of the welding process.

The information presented in the following pages is an expansion and continuation of the work done by Eugene Edwards; 1980 NASA/ASEE Fellow (Alabama A&M University), MSFC and Dr. A. C. Nunes, NASA, MSFC. It provides a method to determine the current density within an electron beam when the current density is radially symmetric. This assumption appears to be practical for electron beam cathodes which are circular.

The whirling wire probe technique, used here and to be described later, was used by A. Sanderson[2] for determining the diameter and envelope shape of the electron beam. His paper clearly indicates the strong dependence of weld shape and penetration on electron beam characteristics; and photographs in his paper show the correspondence between probe traces and welding cross-sections. However, he does not provide the method to obtain quantitative results for local current densities. The following paragraphs describe the mathematical basis for such calculations and a computer program for making the calculations is provided in Appendix 1.
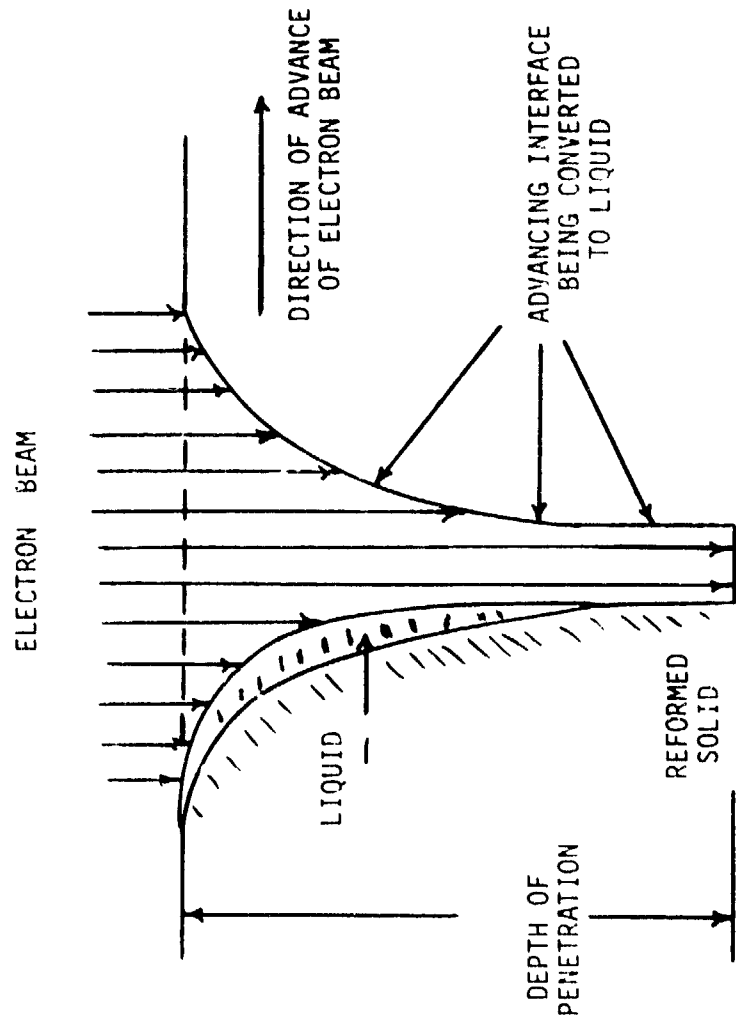
FIGURE 1. THE WELD CAVITY

DATA ACQUISITION

The whirling wire probe shown in Figure 2 rotates through a cross-section of the electron beam. When the probe intercepts the electron beam, current is conducted through the probe to ground. A measure of the probe current versus time is the data to be recorded either by trace followed by digitization or recorded directly in digitized form. Types of traces to be expected are also shown in Figure 2 for probes of width wider than or narrower than the electron beam The computer program shown in Appendix 1 analyzes the data for any width probe.

Traces should have an axis of symmetry, if the beam is radially symmetric, corresponding to the time when the axis of the probe passes through the center of the beam. Ribbon probes made of tungsten and of reasonable dimensions serve satisfactorily in this simple device since the intercept time through an electron beam, typically of diameter .050 inches, during which the probe heats up is short compared to the cooling time outside the electron beam.

FIGURE 2. EXPERIMENTAL ACQUISITION OF DATA

MATHEMATICAL ANALYSIS

Users of the computer program in Appendix 1 must determine the data points corresponding to the probe positions when the leading edge of the probe (Figure 3):

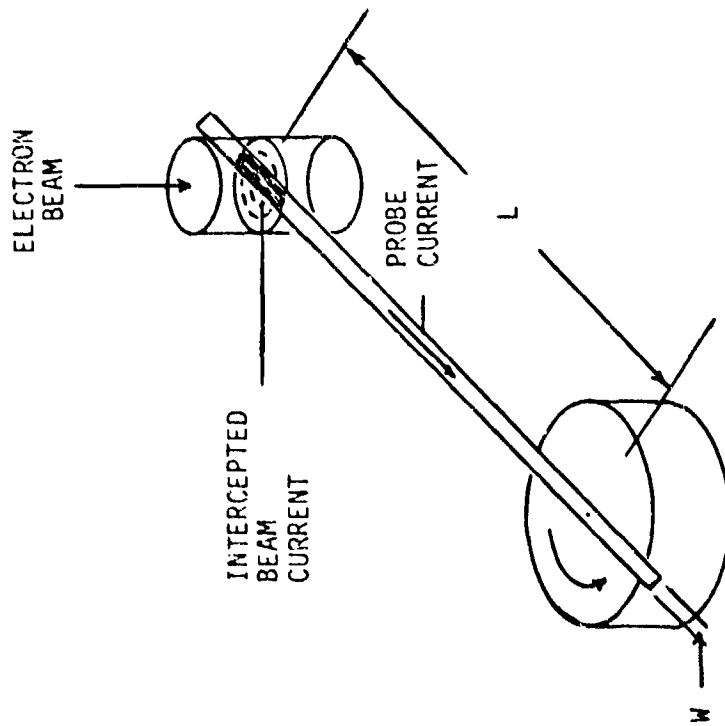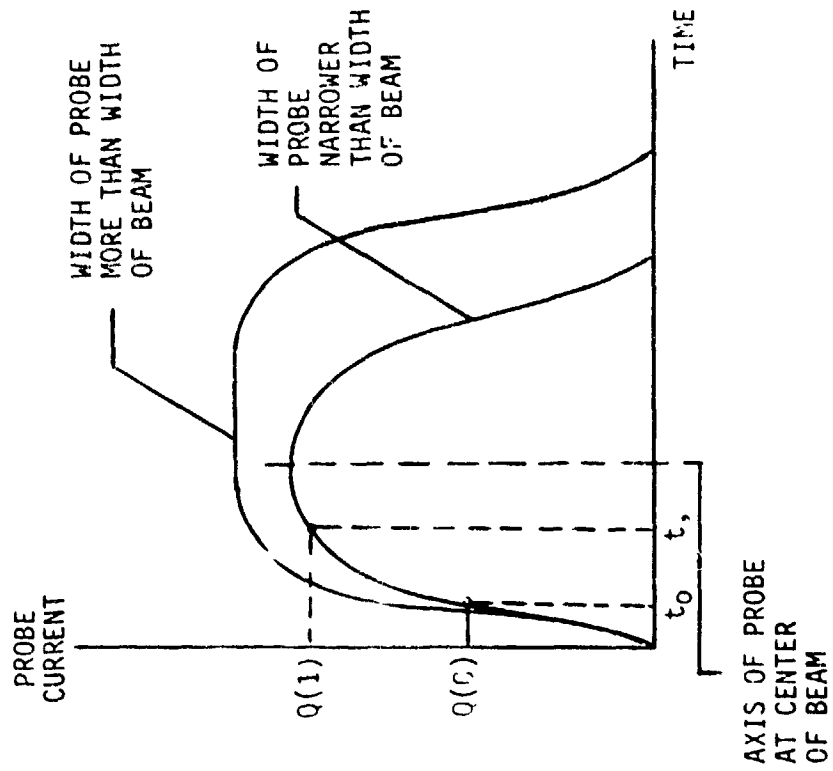a) is at the center of the electron beam.

b) is just tangent to the outer bounding circle of the beam after passing through it (Position N).

All other data correspond to positions of the probe and must be entered in the program sequentially as the probe would move through them passing from zero position to position N.

For a particular data point, when the probe is in position I, it is possible to theoretically construct, in the cross-section of the beam cut by the probe, a circle with center at the center of the beam and tangent to the leading edge of the probe (Figure 3). Hence, for N+1 data points (I=0,1,2,···,N) there are constructed N circles dividing the cross-sectional area into N rings, ring 1 being the innermost ring and ring N the outermost.

Angle theta, T(I), is the angle measured counterclockwise from the leading edge of the probe in zero position to the leading edge of the probe in position I. The angle, theta, through which the probe rotates from the zero position to the symmetry position, can be shown to be $\sin^{-1}W/2L$ where W is the width of the probe and L is the distance between the center of rotation of the probe and the center of the cross-section of the electron beam.

It can be shown that the relationships between the time, t,t=o at zero position, the rate of rotation of the probe, ω, and the radius of the constructed circle at position 1 are:

$$T = theta = \omega t$$

$$theta \text{ at position } I = T(I) = \sin^{-1}(W/2L) - \sin^{-1}((W/2 - R(I))/L)$$

Hence, the required data for entry into the program in Appendix 1 can be calculated.

The cross-section of the electron beam divided into four rings with the probe in position 1 is shown in Figure 4. Probe current, Q(1), for position 1 is the sum of the products which are obtained by multiplying the current density of the electron beam in ring J,CD(J), by the area of ring J,A(I,J) intercepted by the probe.

The intercepted areas A(I,J) can be calculated once the data points are known, and the four equations obtained in this case for the probe in positions 1, 2, 3, and 4 can be solved for the four unknown current densities.
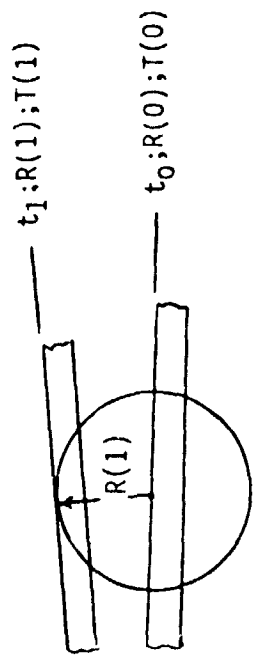
II-8
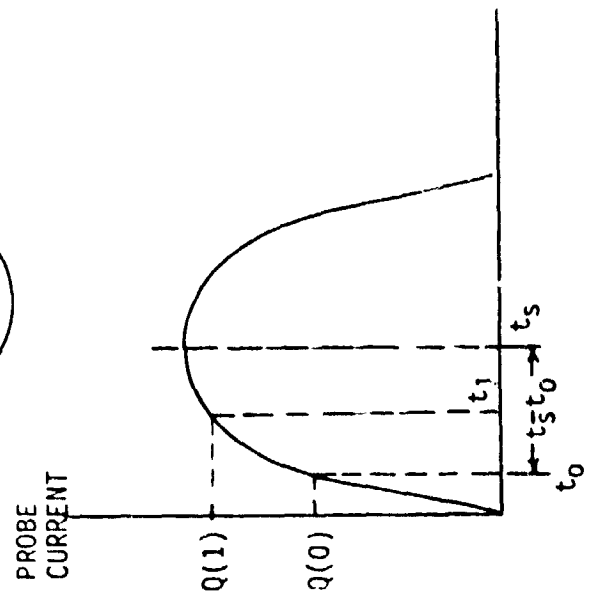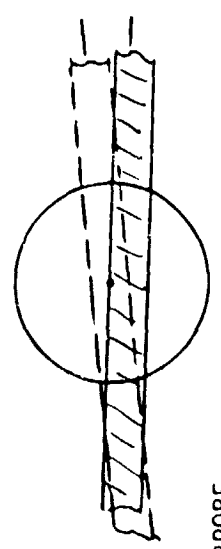
FIGURE 3. MATHEMATICAL ANALYSIS

PROBE LEADING
EDGE SHOWN AT
POSITION I=1

$Q(1) = A(1,1) \; CD(1) +$
$A(1,2) CD(2) +$
$A(1,3) CD(3) +$
$A(1,4) CD(4)$

A(1,1)

A(1,2)

A(1,3)
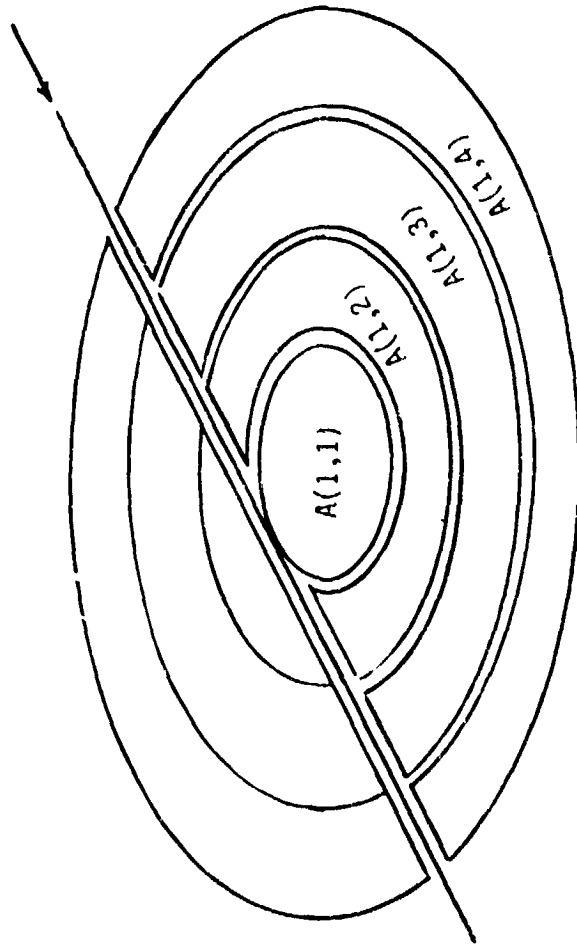
A(1,4)

$Q(I)$ = CURRENT INTERCEPTED BY WIDE PROBE AT POSITION I.

$CD(J)$ = CURRENT DENSITY IN RING J.

$A(I,J)$ = AREA OF THE Jth RING INTERCEPTED BY PROBE AT POSITION I.

$Q(I)$ = $A(I, J) CD(J)$ or $\underline{Q} = \underline{A} \; \underline{CD}$ (MATRIX NOTATION)

HENCE: $\underline{CD} = \underline{A}^{-1} \; \underline{Q}$ ⟵ PROBE DATA

Dt.SITY DISTRIBUTION ⟋

NOTE: BEIL ANALYSIS INCORPORATES TRAILING AS WELL AS LEADING EDGE.

FIGURE 4. MATHEMATICAL BASIS FOR CALCULATION OF RADIAL CURRENT
DENSITY DISTRIBUTION FROM WHIRLING PROBE DATA

SENSITIVITY OF SOLUTION

If the current density is assumed to be 1 MA/in^2 across the cross-section of the electron beam and zero elsewhere, the probe current at any position will equal the sum of the intercepted areas. Calculation of the sum of these areas by computer program results in an output to six significant figures. Calculations of five such probe currents were made, the results re-entered in the computer program as probe currents and the resulting calculated current density distribution compared with the assumed uniform current density within the cross-section of the beam. Figure 5(a) shows that the computer program returned the uniform current density satisfactorily.

The remainder of Figure 5 and Figure 6 indicate that, as the calculated probe currents were rounded to fewer and fewer significant figures by the even digit convention, the deviation of the calculated current densities from the assumed uniform current densities increased. Relative current error is relative to the six significant digit current values. The result of zero current density in ring 5 shown in Figure 6(c) is due to obtaining the same value for Q(4) and Q(5) in the rounding off process. This sensitivity check indicated that:

a) the computer program did not degrade the data collected

b) the process is sensitive to accuracy of input, and, in this case .4% maximum relative current error would return satisfactory results. That is, for the range of currents expected, precision of the current measurement must be at least to 100 microamps.

Results from the same calculations for ten rings, shown in Figure 7 and Figure 8, for the same bounding circle of the cross-sectional area, indicate that precision to 10 microamperes would be needed for the calculated current densities to satisfactorily approximate a uniform current density. Thus, approximating the current density profile over a finer net is accompanied by the need for greater precision in acquiring data.

(c) CUPPENT DENSITY

DATA TO FOUR SIGNIFICANT FIGURES

.0^1% MAXIMUM RELATIVE CURRENT ERROR

1.4% MAXIMUM RELATIVE DENSITY ERROR

1 2 3 4 5

RADIUS

(b) CURRENT DENSITY

DATA TO FIVE SIGNIFICANT FIGURES

.005% MAXIMUM RELATIVE CURRENT ERROR

.2% MAXIMUM RELATIVE DENSITY ERROR

1 2 3 4 5

RADIUS

(a) CURRENT DENSITY

DATA TO SIX SIGNIFICANT FIGURES

0% MAXIMUM RELATIVE CURRENT ERROR

.0027% MAXIMUM RELATIVE DENSITY ERROR

1 2 3 4 5

RADIUS

MATRIX QUALITY NUMBER 151

FIGURE 5. SENSITIVITY ANALYSIS ILLUSTRATION OF EFFECT OF DATA PRECISION

(c)

DATA TO ONE SIGNIFICANT
FIGURE

8%   MAXIMUM RELATIVE
      CURRENT ERROR

100% MAXIMUM RELATIVE
      DENSITY ERROR

CURRENT
DENSITY

1

2

3

4

5

RADIUS

(b)

DATA TO TWO SIGNIFICANT
FIGURES

2.9% MAXIMUM RELATIVE
      CURRENT ERROR

130% MAXIMUM RELATIVE
      DENSITY ERROR

CURRENT
DENSITY

1

2

3

4

5

RADIUS

(a)

DATA TO THREE SIGNIFICANT
FIGURES

.4%   MAXIMUM RELATIVE
      CURRENT ERROR

13%   MAXIMUM RELATIVE
      DENSITY ERROR

CURRENT
DENSITY

1

2

3

4

5
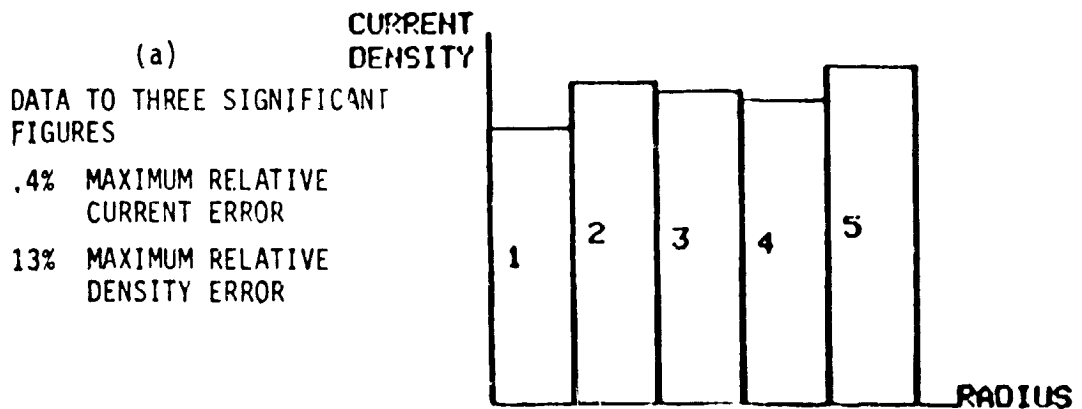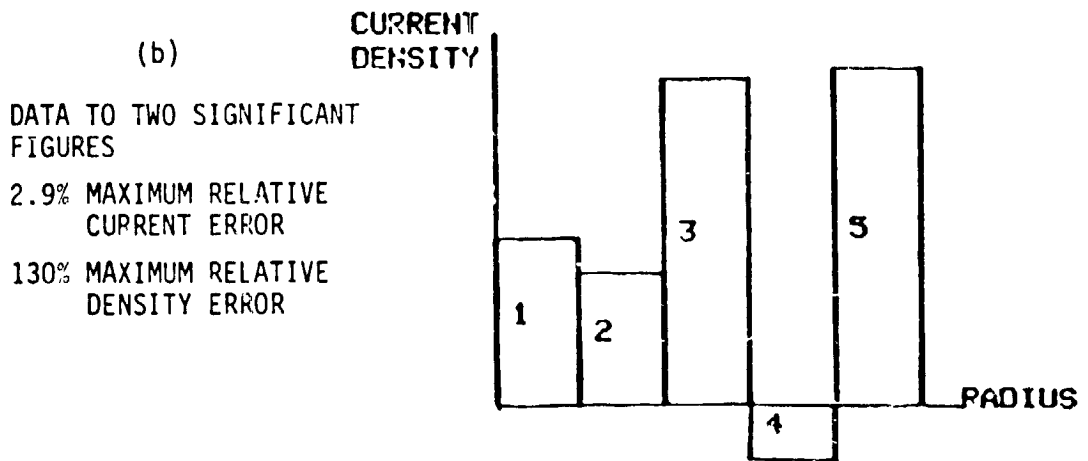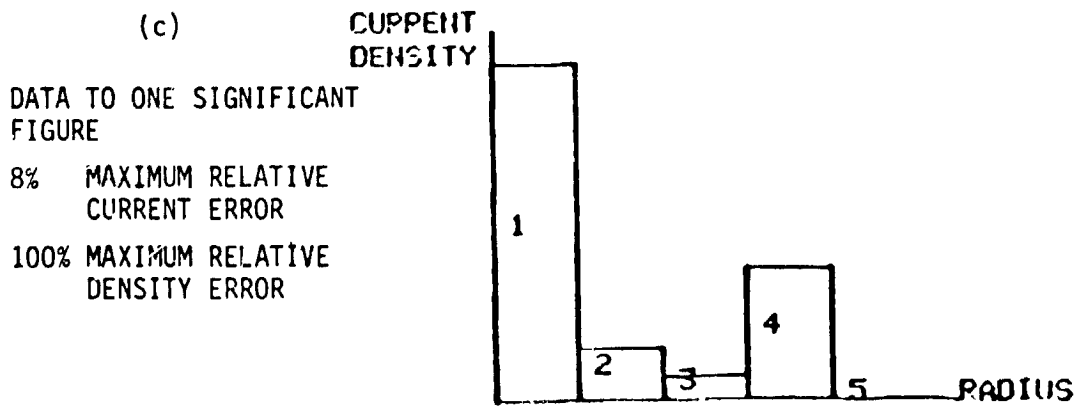
RADIUS

MATRIX QUALITY NUMBER 151

FIGURE 6.   SENSITIVITY ANALYSIS ILLUSTRATION
            OF EFFECT OF DATA PRECISION

II-13

(c)

DATA TO FOUR
SIGNIFICANT FIGURES

.041% MAXIMUM RELATIVE
CURRENT ERROR

5.2% MAXIMUM RELATIVE
DENSITY ERROR

(b)

DATA TO FIVE
SIGNIFICANT FIGURES

.005% MAXIMUM RELATIVE
CURRENT ERROR

.53% MAXIMUM RELATIVE
DENSITY ERROR

(a)

DATA TO SIX
SIGNIFICANT FIGURES

0 % MAXIMUM RELATIVE
CURRENT ERROR

.023% MAXIMUM RELATIVE
DENSITY ERROR
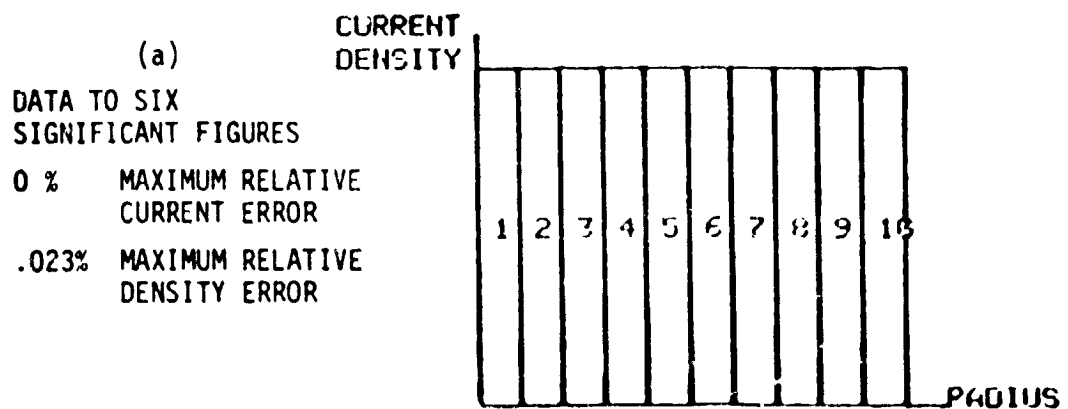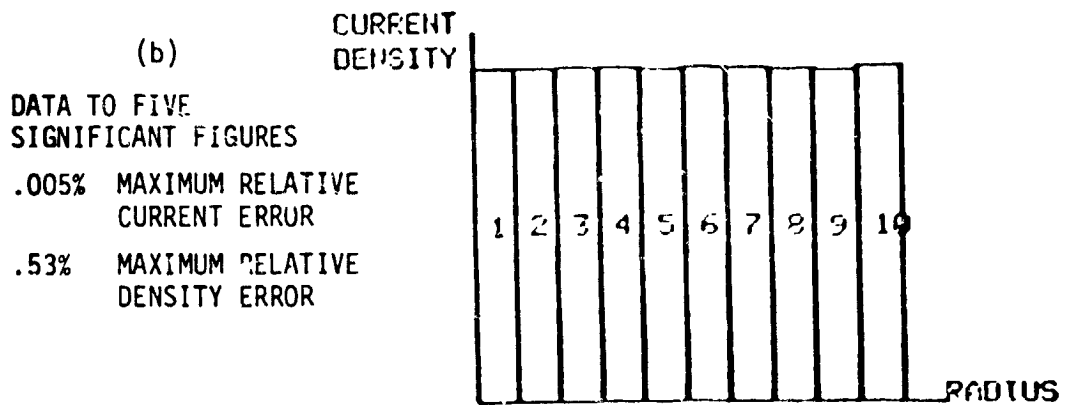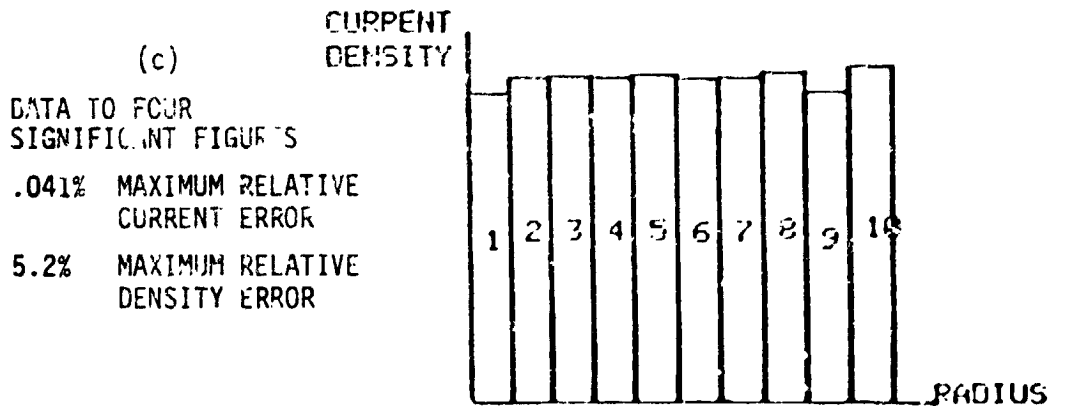
MATRIX QUALITY NUMBER 666

FIGURE 7.   SENSITIVITY ANALYSIS
OF EFFECT OF DATA PRECISION

II-14

(c)

DATA TO ONE
SIGNIFICANT FIGURE

6.0% MAXIMUM RELATIVE
      CURRENT ERROR

114% MAXIMUM RELATIVE
      DENSITY ERROR

(b)

DATA TO TWO
SIGNIFICANT FIGURES

3.3% MAXIMUM RELATIVE
      CURRENT ERROR

521% MAXIMUM RELATIVE
      DENSITY ERROR

(a)

DATA TO THREE
SIGNIFICANT FIGURES

.45% MAXIMUM RELATIVE
      CURRENT ERROR

32.5% MAXIMUM RELATIVE
      DENSITY ERROR

MATRIX QUALITY NUMBER 666
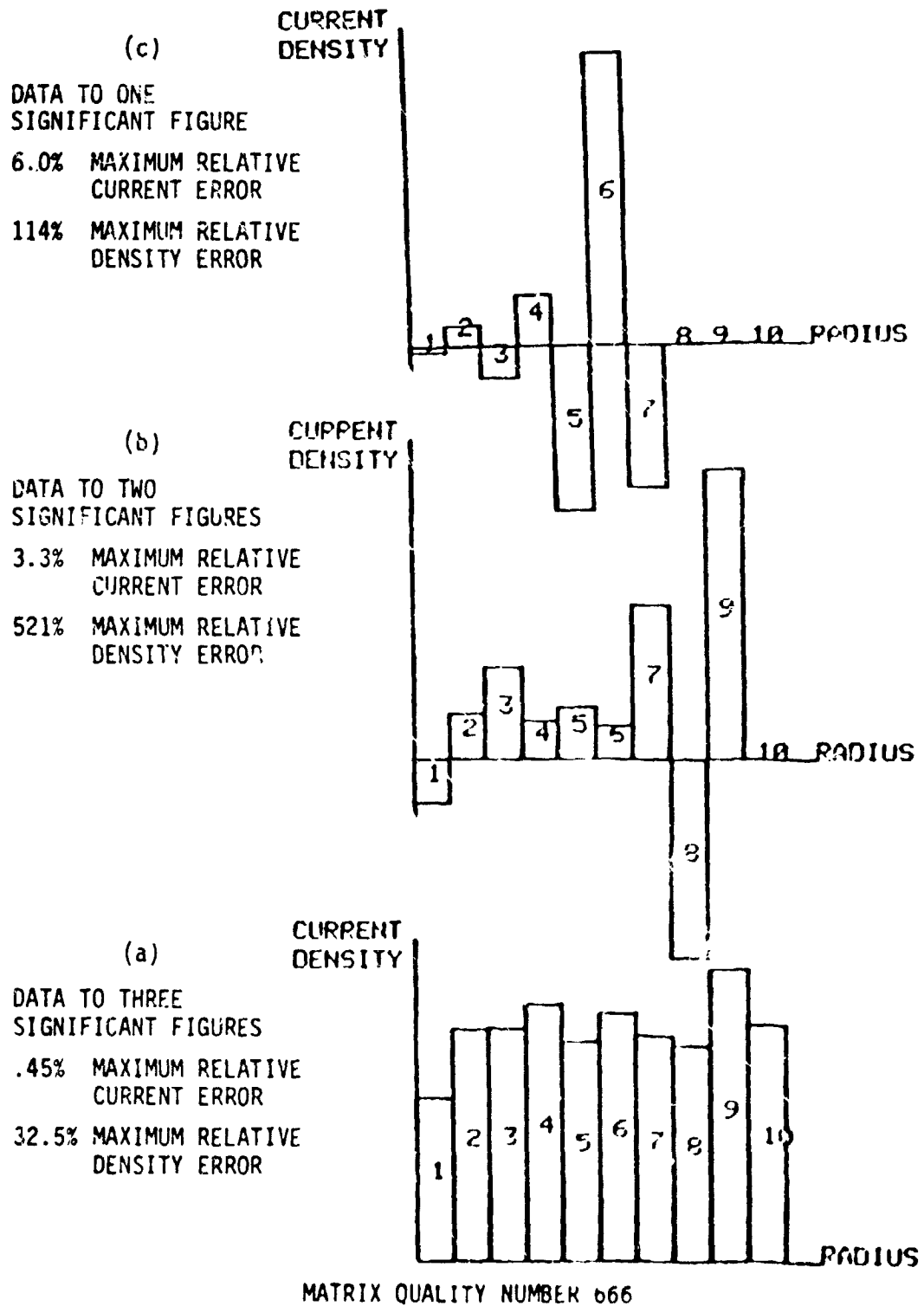
FIGURE 8.   SENSITIVITY ANALYSIS ILLUSTRATION
            OF EFFECT OF DATA PRECISION

II-15

ERROR ANALYSIS

The equations that are to be solved for current densities can be written in matrix form. For instance, for three rings they are:

$$A(1,1)CD(1)+A(1,2)CD(2)+A(1,2)CD(3)=Q(1)$$
$$A(2,1)CD(1)+A(2,2)CD(2)+A(2,3)CD(3)=Q(2)$$
$$A(3,1)CD(1)+A(3,2)CD(2)+A(3,3)CD(3)=Q(3)$$

or $\quad \underline{A}\ \underline{CD} = \underline{Q}$

in matrix form. The current densities are then

$$\underline{CD} = \underline{A}^{-1}\underline{Q}$$

If $\underline{Q}$ is subjected to a small error $\underline{dQ}$, then

$$\underline{dCD} = \underline{A}^{-1}\underline{dQ}\ ,$$

and $\quad \left\| \underline{dCD} \right\|_E = \left\| \underline{A}^{-1}\ \underline{dQ} \right\|_E \leq \left\| \underline{A}^{-1} \right\|_E \left\| \underline{dQ} \right\|_E$

where $\left\| \underline{CD} \right\|_E = SQR\left[CD(1)^2 + CD(2)^2 + CD(3)^2 \right]$

and $\quad \left\| \underline{A} \right\|_E = SQR\left[\text{sum of the squares of elements of } \underline{A} \right]$,

where SQR means "square root of". Use of norms, as above, in error analysis is discussed in Wilkinson's[3] text. Continuing,

$$\left\| \underline{Q} \right\|_E \leq \left\| \underline{A} \right\|_E \left\| \underline{CD} \right\|_E$$

Hence, $\quad \dfrac{\left\| \underline{dCD} \right\|_E}{\left\| \underline{CD} \right\|_E} \leq \left\| \underline{A} \right\|_E \left\| \underline{A}^{-1} \right\|_E \dfrac{\left\| \underline{dQ} \right\|_E}{\left\| \underline{Q} \right\|_E}$

Thus, the matrix quality number $\left\| \underline{A} \right\|_E \left\| \underline{A}^{-1} \right\|_E$ indicates the relative response of the current density vector to a small relative error in the current vector. It has been found that the matrix quality number is a pessimistic bound. It is, however, useful in error analysis since smaller quality numbers generally indicate less sensitivity of the mathematical analysis to small errors in the data required.

The area influence coefficients, $A(I,J)$, have the property that, for I greater than or equal to J, all of the $A(I,J)$ are equal. For instance,

$$A(1,1)=A(2,1)=A(3,1)=\pi R(1)^2.$$

Hence, it is possible to rewrite the equations as

$$A(1,1)CD(1) \cdot \qquad A(1,2) \; CD(2)+ \qquad A(1,3) \; CD(3)=Q(1)$$
$$(A(2,2)-A(1,2))CD(2)+(A(2,3)-A(1,3))CD(3)=Q(2)-Q(1)$$
$$(A(3,3)-A(2,3))CD(3)=Q(3)-Q(2)$$

The last equation gives

$$CD(3)=(Q(3)-Q(2))/(A(3,3)-A(2,3)).$$

An error of $\pm$ B in both $Q(3)$ and $Q(2)$ would give

$$\text{maximum relative error in } CD(3)=2B/(Q(3)-Q(2))$$

which would be large if $Q(3)-Q(2)$ is small. The uncertainty of results, due to algebraic operations, such as that above is discussed by Baird[4].

If it is assumed that the only current in error is $Q(3)$, and that error is $+ 1$, then how the error propagates into the calculation of current densities can be found directly. For this example,

$$dCD(3)=1/(A(3,3)-A(2,3))$$

$$dCD(2)=-dCD(3)(A(2,3)-A(1,2))/(A(2,2)-A(1,2))$$

and $dCD(1)$ can be found by substituting these values in the first equation. Calculation of the error distribution for fourteen rings on a beam of radius 1 using the computer program given in Appendix 2 is shown in Figure 10.

From another viewpoint, a systematic error which increases each current reading by 1 will only affect the current density in the inner ring. Conversely, raising the current density in ring 1 so that $Q(1)$ is increased by 1 must, from physical considerations, raise all $Q(I)$ by 1. Only a brief discussion of how errors may propagate and the effect their magnitudes have on calculated current densities has been presented here. However, even this brief discussion shows that the analysis is extremely sensitive to the precision of the data acquired.

Finally, the thin wire analysis, for a small probe of width W, for two rings is shown with the dimensions assumed in Figure 9. The graph of $CD(1)$ versus $CD(2)$ shows clearly how it is possible, with $Q(0)$ fixed, to obtain a large change in calculated current densities for a small change in $Q(1)$.
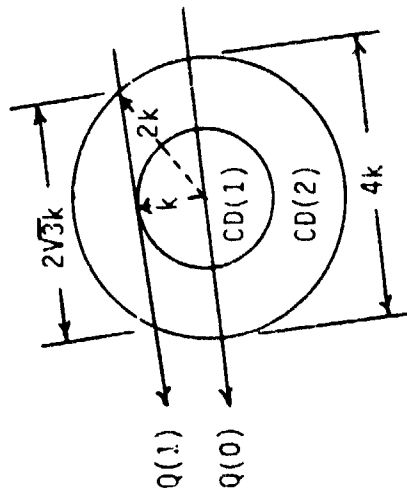
$$Q(1) - 2\sqrt{3}\ k\ w\ CD(2)$$
$$Q(0) = 2\ kw\ CD(1) + 2\ kw\ CD(2)$$

Q(1) = .866 Q(0)

CD(1)=CD(2)

Q(1)

Q(0)

Q(1) = 1.732 Q(0)

CD(1)=0

Q(1)

Q(0)

2√3k

2V3k

2k

k

CD(1)

CD(2)

4k

Q(1)

Q(0)

CD(1)

Q(0)/2kw

Q(0)/2kw

CD(1)=CD(2)

CD(2)

Q(1)/√3 2kw

Q(0)/2kw
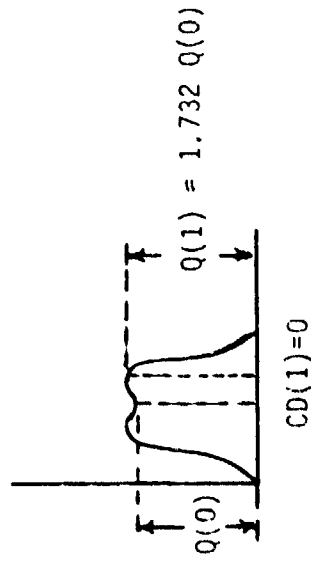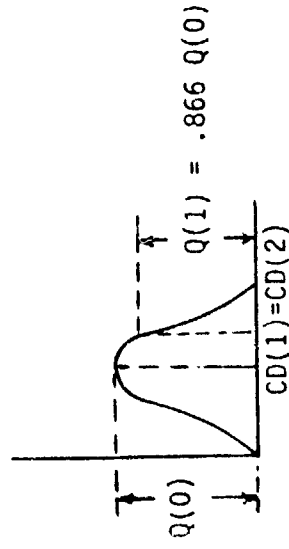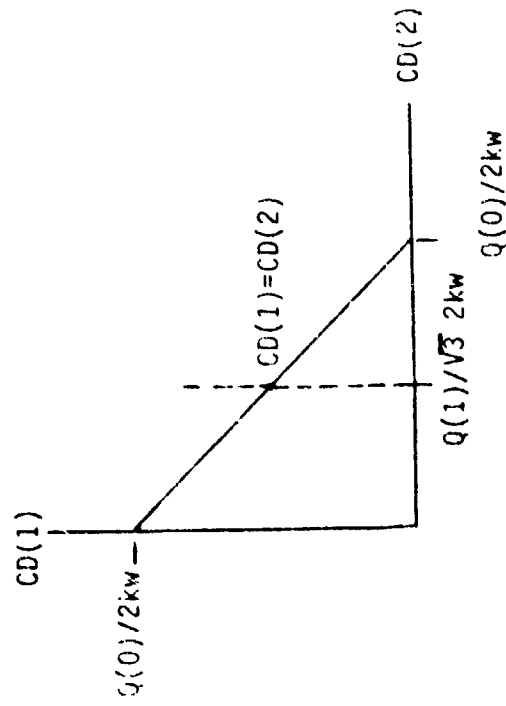
FIGURE 9. ACCURACY OF RESULTS THIN WIRE PROBE

TOTAL ERROR IN RING N-J

J	ERROR IN RING N-J,RING N IS THE OUTER RING
0	28.0333
1	-24.3771
2	4.29094
3	-3.56956
4	-.178739
5	-.983959
6	-.50146
7	-.53669
8	-.43702
9	-.410864
10	-.377691
11	-.368335
12	-.34836
13	-62.7302

FIGURE 10. CURRENT DENSITY ERROR COMPUTATION FOR A UNIT ERROR
IN Q(N) FOR RINGS OF EQUAL WIDTH WHEN P(N) = 1

CONCLUSIONS

A computer analysis capable of reducing whirling wire current probe data to a radial ring current density distribution has been constructed and run successfully. For a five ring density distribution with a 13 percent maximum relative current density error on flat distribution data, precision of three figures or $\pm$ 0.4 percent relative current error is needed. Data precision requirements increase with increase in the number of rings included in the analysis.

Finally, precision levels of rough data obtained to date with the MSFC whirling wire probe (about $\pm$ 1.MA) are shown to be in Figure 11 not sufficient to warrant using more than one ring.

RECOMMENDATIONS

The Metals Processing Branch will acquire soon equipment which will record data in digitized form, thus eliminating the step of digitizing traces to obtain data. Then, a study to determine the data precision limits attainable with the MSFC whirling wire probe should be undertaken. If adequate precision is not attainable using the whirling wire probe, a study of other candidate methods for determining the electron beam current density should be made.

● THE PRECISON LEVELS OF ROUGH DATA OBTAINED TO DATE
WITH THE MSFC WHIRLING WIRE PROBE (ABOUT ± 1 MA)
YIELDS A SINGLE RING ANALYSIS:
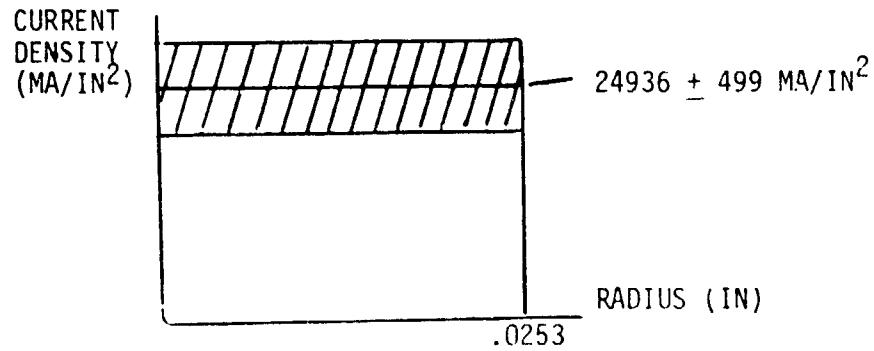
AND A DOUBLE RING ANALYSIS:
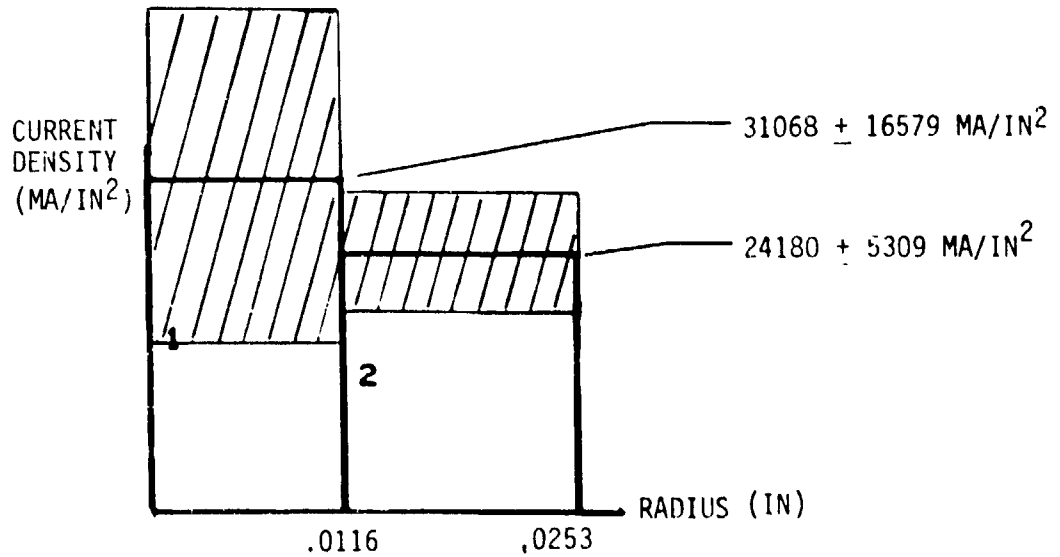
WITH INSUFFICIENT PRECISION FOR MORE RINGS.

FIGURE 11. ROUGH DATA PRECISION LEVELS

REFERENCES

1. Edwards, Eugene, "Power Distribution for Electron Beam Welding", Report, NASA/ASEE Summer Faculty Research Fellowship Program, MSFC/ University of Alabama, 1980.

2. Sanderson, A., "Electron Beam Monitoring Techniques", Proceedings of the Conference on Advances in Welding Processes, 14-16 April 1970, The Welding Institute, Abingdon Hall, Cambridge, 1971, pp 161-166.

3. Wilkinson, J.H., Rounding Errors in Algebraic Processes, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963.

4. Baird, D.C., Experimentation: An Introduction to Measurement Theory and Experiment Design, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962.

```
10 REM ***ELECTRON BEAM RADIAL CURRENT DISTRIBUTION COMPUTATION***
20 DELETE T,Q,R,A,AO,CO,AC\CLEAR
30 PRINT "INPUT:RADIUS TO BEAM(INCHES),WIDTH OF RECEPTOR(INCHES)."
40 INPUT L,W
50 PRINT "INPUT: TOTAL BEAM CURRENT ( MILLIAMPS)."
60 INPUT Q0
70 PRINT "INPUT: TOTAL NUMBER OF DATA POINTS."
80 INPUT M\LET N=M-1
82 REM ***DIM(N) DIMENSIONS AN ARRAY WITH N+1 ELEMENTS,A(0),A(1),***
83 REM ***...,A(N)
84 REM ***T(I) IS THE ANGLE AT WHICH THE CURRENT Q(I) IS RECORDED.***
85 REM ***Q(0) IS THE CURRENT RECORDED WHEN THE *LEADING EDGE* OF THE***
86 REM ***RECEPTOR PASSES THROUGH THE CENTER OF THE ELECTRON BEAM.***
87 REM ***THE CORRESPONDING ANGLE T(0)=9.***
90 DIM T(N),Q(N),R(N),A(N,N),AO(N,N),CO(N)
100 T=0\Q=0
110 PRINT "INPUT DATA: ANGLE(RADIANS),CURRENT(MILLIAMPS)."
120 PRINT "NOTE:THE ANGLE T(I) IS THE ANGLE MEASURED COUNTRCLOCKWISE"
130 PRINT "FROM THE LEADING EDGE OF THR RECEPTOR AT T(0) TO THE"
140 PRINT "LEADING EDGE OF THE RECEPTOR AT T(I). ANGLES ARE TO BE TAKEN"
141 PRINT "POSITIVE AND LESS THAN PI/2 FOR THIS PROGRAM."
150 FOR I=0 TO N
160 INPUT T(I),Q(I)
170 NEXT I
210 PI=4*ATN(1)
220 R=0
230 AL=ATN(W/2/SQR(L*L-W*W/4))
240 FOR I=1 TO N
250 R(I)=W/2+L*SIN(T(I)-AL)
260 NEXT I
270 REM ***NOTE THAT R(0)=0***
310 REM ***CALCULATION OF AREA INFLUENCE COEFFICIENT MATRIX***
311 A2=0
320 FOR I=1 TO N
```

```
325 IF W>R(I) THEN 400
330 FOR K=1 TO N
335 A(I,K)=0
345 IF R(I)>=R(K) THEN 355
350 A(I,K)=-R(K)^2*ATN(SQR(R(K)^2-R(I)^2)/R(I))+R(I)*SQR(R(K)^2-R(I)^2)
352 A(I,K)=A(I,K)-R(I)*SQR(R(K-1)^2-R(I)^2)
353 A(I,K)=A(I,K)+R(K-1)^2*ATN(SQR(R(K-1)^2-R(I)^2)/R(I))
355 IF (R(I)-W)>=R(K-1) THEN 370
360 A(I,K)=A(I,K)-R(K-1)^2*ATN(SQR(R(K-1)^2-(R(I)-W)^2)/(R(I)-W))
365 A(I,K)=A(I,K)+(R(I)-W)*SQR(R(K-1)^2-(R(I)-W)^2)
370 IF (R(I)-W)>=R(K) THEN 385
375 A(I,K)=A(I,K)+(R(K)^2)*ATN(SQR(R(K)^2-(R(I)-W)^2)/(R(I)-W))
380 A(I,K)=A(I,K)-(R(I)-W)*SQR(R(K)^2-(R(I)-W)^2)
385 NEXT K
390 GOTO 460
400 FOR K=1 TO N
405 A(I,K)=PI*(R(K)^2-R(K-1)^2)
410 IF R(I)>=R(K) THEN 425
415 A(I,K)=A(I,K)-R(K)^2*ATN(SQR(R(K)^2-R(I)^2)/R(I))
416 A(I,K)=A(I,K)-R(I)*SQR(R(K-1)^2-R(I)^2)
417 A(I,K)=A(I,K)+R(K-1)^2*ATN(SQR(R(K-1)^2-R(I)^2)/R(I))
420 A(I,K)=A(I,K)+R(I)*SQR(R(K)^2-R(I)^2)
425 IF (W-R(I))>=R(K) THEN 440
430 A(I,K)=A(I,K)-R(K)^2*ATN(SQR(R(K)^2-(W-R(I))^2)/(W-R(I)))
435 A(I,K)=A(I,K)+(W-R(I))*SQR(R(K)^2-(W-R(I))^2)
440 IF (W-R(I))>=R(K-1) THEN 455
445 A(I,K)=A(I,K)+R(K-1)^2*ATN(SQR(R(K-1)^2-(W-R(I))^2)/(W-R(I)))
450 A(I,K)=A(I,K)-(W-R(I))*SQR(R(K-1)^2-(W-R(I))^2)
455 NEXT K
460 NEXT I
461 FOR I=1 TO N
462 FOR J=1 TO N
463 A2=A2+A(I,J)*A(I,J)
464 NEXT J
465 NEXT I
```

```
466 PAGE
467 PRINT " AREA INFLUENCE COEFFICIENT MATRIX"
468 FOR I=1 TO N
469 PRINT
470 FOR J=1 TO N
472 PRINT A(I,J);TAB(10*J);
474 NEXT J
476 NEXT I
478 PRINT "^[^H^[^L"
480 WAIT 10000
490 REM ***MATRIX INVERSION***
500 AD=0
510 FOR I=1 TO N
520 FOR J=1 TO N
530 IF I=J THEN AD(I,J)=1
540 NEXT J
550 NEXT I
560 FOR I=1 TO N
570 C=A(I,I)
580 FOR J=1 TO N
590 A(I,J)=A(I,J)/C
600 AD(I,J)=AD(I,J)/C
610 NEXT J
620 FOR K=1 TO N
630 IF K=I THEN 690
640 B=A(K,I)
650 FOR J=1 TO N
660 A(K,J)=A(K,J)-B*A(I,J)
670 AD(K,J)=AD(K,J)-B*AD(I,J)
680 NEXT J
690 NEXT K
700 NEXT I
701 A3=0
702 FOR I=1 TO N
703 FOR J=1 TO N
```

```
704 A3=A3+AD(I,J)*AD(I,J)
705 NEXT J
706 NEXT I
707 CN=SQR(A2)*SQR(A3)
710 REM ***CURRENT DENSITY COMPUTATION***
720 CD=0
730 FOR I=1 TO N
740 FOR J=1 TO N
750 DC=AD(I,J)*Q(J)
760 CD(I)=CD(I)+DC
770 NEXT J
780 NEXT I
782 REM ***TOTAL CURRENT INSIDE R(N)***
790 Q1=0
800 FOR I=1 TO N
810 Q1=Q1+PI*(R(I)^2-R(I-1)^2)*CD(I)
820 NEXT I
860 REM ***PRINTOUT***
870 PAGE
880 PRINT "ANGLE";TAB(10);"CURRENT";TAB(20);"RADIUS";TAB(40);
930 PRINT "CURRENT DENSITY"
900 PRINT "(RAD)";TAB(10);"(MA)";TAB(20);"(IN)";TAB(40);
910 PRINT "(AMPS/SQ IN)"
920 PRINT\PRINT\PRINT
930 FOR I=1 TO N
940 PRINT T(I);TAB(10);Q(I);TAB(20);R(I);TAB(40);CD(I)/1000
950 NEXT I
960 PRINT\PRINT\PRINT
970 PRINT "RATIO OF TOTAL CURRENT CALCULATED TO TOTAL BEAM CURRENT"
980 PRINT\PRINT Q1/Q0
990 PRINT "^E^H^E^L"
1000 WAIT 10000
1010 PRINT "SENSITIVITY MATRIX;AD=INV(A)"
1020 FOR I=1 TO N
1030 PRINT
```

```
1040 FOR J=1 TO N
1050 PRINT A0(I,J);TAB(10*J);
1060 NEXT J
1070 NEXT I
1080 PRINT "^C^H^C^L"
1090 WAIT 10000
2000 INITG
2010 DIM AC(N)
2011 AC=0
2020 PRINT
2021 PRINT
2022 PRINT
2023 PRINT
2024 PRINT
2030 PRINT "UNIT";TAB(6);"OUTER";TAB(20);"CURRENT"
2040 PRINT TAB(6);"RADIUS";TAB(20);"DENSITY"
2050 PRINT TAB(7);"(IN)";TAB(21);"(MA)"
2060 PRINT
2070 FOR I=1 TO N
2080 PRINT TAB(1);I;TAB(6);R(I);TAB(20);C0(I)
2093 NEXT I
2100 PRINT
2101 PRINT
2110 PRINT "LENGTH OF RECEPTOR=";L;"INCHES"
2120 PRINT
2130 PRINT "WIDTH OF RECEPTOR=";W,"INCHES"
2140 PRINT
2150 PRINT "CONDITION NUMBER OF AICM=";CN
2160 SMOVE 902,389
2170 PRINT "RADIUS"
2180 SMOVE 439,667
2190 PRINT "CURRENT"
2200 SMOVE 438,642
2210 PRINT "DENSITY"
2220 FOR I=1 TO N
```

II-27

```
2230 AC(I)=ABS(CD(I))
2240 NEXT I
2250 YR=1.1*MAX(AC)
2260 WINDOW -.1*R(N),1.1*R(N),-YR,YR
2270 VIEWPORT 512,902,111,667
2280 FOR I=1 TO N
2290 MOVE R(I-1),0
2300 DRAW R(I-1),CD(I),R(I),CD(I),R(I),0
2301 MOVE R(I-1),CD(I)/2
2302 PRINT I
2310 NEXT I
2320 MOVE 0,-YR
2330 DRAW 0,YR
2340 MOVE 0,0
2350 DRAW 1.1*R(N),0
2360 PRINT "^[^W^[^L"
2370 END

READY
*
```

```
200 REM *** CURRENT DENSITY ERROR CALCULATION FOR A UNIT***
210 REM *** ERROR IN Q(N) WHERE R(N)=1***
250 REM *** INPUT ***
251 PRINT "INPUT THE NUMBER OF RINGS"
252 DELETE W,M,E,P,Z\CLEAR
253 DELETE X\CLEAR
255 INPUT N
256 DIM W(N,N),M(N,N),E(N),P(N,N),Z(N)
257 DIM X(1,N)
260 W=0
264 PI=4*ATN(1)
270 M=0
280 E=0
290 P=0
291 L=0
300 Z=0
340 REM *** CALCULATION OF MULTIPLIERS AND ERRORS***
350 FOR I=1 TO N-2
360 FOR J=0 TO I-1
370 A=N-J
380 B=N-J-1
390 C=N-I
400 D=N-I-1
410 U=(A^2-B^2)*(ATN(SQR(A^2-D^2)/U)-ATN(SQR(A^2-C^2)/C))
420 F=U-D*SQR(A^2-D^2)+D*SQR(B^2-D^2)
430 G=F+(B^2)*(ATN(SQR(A^2-D^2)/D)-ATN(SQR(B^2-D^2)/D))
440 H=G+C*SQR(A^2-C^2)-C*SQR(B^2-C^2)
450 P(N-I,N-J)=H-(B^2)*(ATN(SQR(A^2-C^2)/C)-ATN(SQR(B^2-C^2)/C))
460 Z(N-I)=(C^2)*ATN(SQR(C^2-D^2)/D)-D*SQR(C^2-D^2)
470 M(N-I,N-J)=P(N-I,N-J)/Z(N-I)
480 NEXT J
490 NEXT I
500 U=ATN(SQR(N^2-(N-1)^2)/(N-1))-(N-1)/N*SQR(1-((N-1)/N)^2)
510 E(0)=1/U
```

```
520 FOR I=1 TO N-2
530 FOR J=0 TO I-1
540 W(N-I,N-J)=-M(N-I,N-J)*E(J)
550 E(I)=E(I)+W(N-I,N-J)
560 NEXT J
570 NEXT I
571 X=0
572 FOR K=2 TO N
573 X(1,K)=PI*(K^2-(K-1)^2)-K^2*ATN(SQR(K^2-1))
574 X(1,K)=X(1,K)-SQR((K-1)^2-1)+SQR(K^2-1)
575 X(1,K)=X(1,K)+(K-1)^2*ATN(SQR((K-1)^2-1))
576 L=L+X(1,K)*E(N-K)
577 W(1,K)=-X(1,K)/PI
578 NEXT K
579 E(N-1)=-L/PI
580 REM ***OUTPUT***
585 PAGE
586 PRINT
590 FOR J=0 TO N-2
610 PRINT "CO ERROR IN RING N-I DUE TO CALCULATED CO ERROR IN RING N-J"
620 PRINT
630 PRINT "W(N-I,N-J)","J=";J
640 PRINT
650 FOR I=1 TO N-1
660 PRINT W(N-I,N-J),"I=";I
670 NEXT I
680 NEXT J
690 PRINT
700 PRINT "TOTAL ERROR IN RING N-J"
710 PRINT
720 PRINT "J","ERROR IN RING N-J,RING N IS THE OUTER RING"
730 PRINT
740 FOR J=0 TO N-1
750 PRINT J,E(J)
760 NEXT J
```

```
700 END

READY
*
```

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

FUEL OPTIMAL MANEUVERS FOR SPACECRAFT
WITH FIXED THRUSTERS

Prepared by: Thomas E. Carter, Ph.D.

Academic Rank: Associate Professor

University and Department: Eastern Connecticut State College
                                      Mathematical Sciences

NASA/MSFC:
     Division: Control Systems Division
     Branch: Dynamics & Trajectory Analysis Branch

MSFC Counterpart: Harry J. Buchanan

Date: August 20, 1981

Contract No.   NGT 01-008-021
                                      The University of Alabama
                                      in Huntsville

FUEL OPTIMAL MANEUVERS FOR SPACECRAFT
WITH FIXED THRUSTERS

by

Thomas E. Carter
Associate Professor of Mathematical Sciences
Eastern Connecticut State College
Willimantic, Connecticut

ABSTRACT

Several models are used for the qualitative investigation of fuel optimal maneuvers for spacecraft with fixed thrusters. With the exception of singular solutions, all solutions are found to consist of intervals of "full thrust" and "coast" indicating that thrusters do not need to be designed as "throttleable" for fuel optimal performance.

For the primary model under consideration, singular solutions can occur only if the optimal solution is "pure translation." For this case, the state space can be separated into a region where singular solutions cannot occur and a region where each point admits infinitely many singular solutions for sufficiently large flight times. Where singular solutions exist, a "time optimal" singular solution can always be found which consists of intervals of "coast" and "full thrust." The shape of the optimal fuel consumption curve as a function of flight time is found to depend on whether or not the initial state is in the region admitting singular solutions.

Comparison of fuel optimal maneuvers in deep space with those relative to a point in circular orbit indicate that qualitative differences in the solutions can occur.

A minimum integral square criterion problem is considered for comparison with the minimum fuel problem. The solutions consist of intervals of "linear" and "full thrust." Computation of fuel consumption for certain "pure translation" cases indicates that considerable savings in fuel can result from the fuel optimal maneuvers.

TABLE OF CONTENTS

LIST OF FIGURES

I. INTRODUCTION

With the successful operation of the Space Shuttle, a wide variety of payloads can economically be placed in orbit. Teleoperator Maneuverable Systems can be used to travel back and forth from the Shuttle to these payloads, docking with them, and pushing them in such a way as to satisfy a wide range of requirements. This establishes some motivation to examine again some fuel optimal control problems of spacecraft even though little has been published in this area since the mid-sixties.

The mathematical framework for minimum fuel problems of spacecraft is well established. In 1965, Neustadt published necessary conditions for a class of minimal fuel space trajectories in a general gravitational field. The more specialized problems we consider here are all of the form presented in [2, Sec. 5.2, Theorem 5], hence the minimum principle provides both necessary and sufficient conditions for a minimum. Since we deal with fixed time problems in which the set of admissible control values is compact and convex and the integrand of the cost function is convex in the controls, we are assured existence of optimal solutions if the target set is reachable in the specified fixed time. Even if the control thrusters are not "throttleable", that is, if they can take only discrete values, an optimal "relaxed" solution exists and the optimal "relaxed" trajectory can be approximated to an arbitrary degree of accuracy [21]. ("Relaxing the controls" effectively changes the set of admissible control values to its convex hull and the integrand of the cost function becomes convex in the control values. See the comments in the Introduction of Reference 22.)

A large amount of early work has been done on the minimal fuel rocket trajectory problem in which the initial-to-final mass ratio is minimized. The system dynamics for rockets in free space were typically described by equations similar to

$$\ddot{x} = c \frac{\dot{m}(t)}{m(t)} e_u(t)$$

where t denotes time, the dot represents differentiation with respect to t, x is the position vector, c the exhaust speed, m the rocket mass, and $e_u(t)$ is a unit steering vector [3, Sec. 3.2], [4].

In general, the mass decrease for teleoperator maneuverable systems is much smaller than for conventional rockets and we shall regard the peak value of $c\,\dot{m}(t)/m(t)$ as constant over the closed time interval $[0, t_f]$. After normalizing, the system dynamics is given by

$$\ddot{x} = u(t) e_u(t)$$

where $u(t) \in [0, 1]$ for each $t \in [0, t_f]$ and we seek to minimize $\int_0^{t_f} u(t)\,dt$.

This idealization does not take into account the fuel loss required to steer the spacecraft, but it can be used to establish bounds on the amount of fuel used during an optimal maneuver. Formulations of the problem which include the steering losses and also which consider motion relative to a circular orbit are presented in Section V. A closed form solution for these cases is not presented as is for the more idealized case. Additionally, a closed form solution is found for the case where the boundary conditions on position and velocity are replaced by boundary conditions on the distance and speed along the trajectory. The solution of the idealized problem in Section II is found to consist of intervals of full thrust ($u(t) = 1$) or coast intervals ($u(t) = 0$) where at most one coast interval can occur. Two mutually orthogonal vectors occur naturally in the solution of the problem and the changes in position and velocity (i.e., $x(t) - x(0)$ and $\dot{x}(t) - \dot{x}(0)$) are confined to the plane of these vectors although the actual position and velocity (i.e., $x(t)$ and $\dot{x}(t)$) may not be planar .

An important special case arises when one of the orthogonal vectors is zero. In this case, and only in this case (with the exception of the trivial case where the trajectory is "coast" on the entire interval $[0, t_f]$), the vectors $\dot{x}(t) - \dot{x}(0)$ and $x(t) - x(0)$ are colinear in the same direction for the entire interval $[0, t_f]$. This is the only situation in which the optimal steering vector $e_u(t)$ can be discontinuous on the open interval $(0, t_f)$. In this situation, the steering vector is constant except at one possible discontinuity where it jumps to the opposite direction and remains. (This occurs also for the rocket problem discussed by Leitmann in [3, Sec. 3.2, p. 77].) This case is also the only case in which singular solutions can arise. This special case can be modeled by replacing $u(t)e_u(t)$ by $u(t)e_u(0)$ where $u(t) \in [-1, 1]$ for each $t \in [0, t_f]$ and replacing the original cost function by

$$\int_0^{t_f} |u(t)|\,dt .$$

It thus becomes a problem of minimum fuel translation and is discussed in Section III.

This problem or its mathematical equivalent, minimum fuel rotation, has been studied by several authors, including Flugge-Lotz and Marbach, Meditch, Swartz, Conrad, and Heller [5] - [9], and is also a special case of the second order linear systems of Ladd and Friedland [10] and Snow [11]. The singular solutions are not of the type to be described by higher order necessary conditions such as, for example, in [22] but can be found as degenerate cases of the class considered in [11]. In view of the sufficiency condition to which we have already alluded in [2], however, one can see that any singular extremal which satisfies the boundary conditions is optimal.

Other linear systems and aspects of more general linear systems were studied by Athans, Craig and Flugge-Lotz, ..., ..., and Athans [12] - [16].

Linear Systems in which

$$\int_0^{t_f} u(t)^2 dt$$

is minimized have been widely studied. See [14], [18] - [20], for example.
In Section IV, we briefly examine this integral square criterion for our
spacecraft dynamics model and find certain similarities to the minimum
fuel problem. The degenerate case in which the motion is pure translation
is found in Leitmann's book [3, Sec. 3.5]. See also [17]. We compare
this solution with the fuel optimum solution and find that it requires
as much as fifty percent more fuel for certain initial conditions.

II. PROBLEM FORMULATION AND SOLUTION

The minimum fuel problem in free space is posed as follows.

Let $\mathcal{U}$ be the set of all Lebesque measurable functions which map the closed interval $[0, t_f]$ into the three-dimensional unit ball. We seek $r \in \mathcal{U}$ to minimize $x^0(t_f)$ subject to the following conditions which hold a.e. on $[0, t_f]$:

$$\dot{x}^0(t) = |r(t)|$$
$$\dot{x}(t) = v(t) \qquad x(0) = x_0 \qquad x(t_f) = x_f$$
$$\dot{v}(t) = r(t) \qquad v(0) = v_0 \qquad v(t_f) = v_f$$

where x_0, v_0, x_f, and v_f are specified points in $\mathbb{R}^3$ and $| \ |$ denotes the Euclidean norm or magnitude.

For this problem, the minimum principle provides both necessary and sufficient conditions for an optimal solution. The Hamiltonian is

$$\mathcal{H} = q^0 |r(t)| + p(t) \cdot v(t) + q(t) \cdot r(t)$$

where the functions p and q which map $[0, t_f]$ into $\mathbb{R}^3$ are absolutely continuous solutions of the adjoint system

$$\dot{p}(t) = 0$$
$$\dot{q}(t) = -p(t)$$

and q^0 is either 1 or 0. If $q^0 = 1$, the solution is called __normal__; otherwise, it is referred to as __abnormal__. In either case, q is described by the following linear function of t

$$q(t) = -p_0 t + q_0 \qquad (0 \le t \le t_f)$$

where $p_0, q_0 \in \mathbb{R}^3$.

The abnormal optimal control, which is actually a time optimal control, is characterized a.e. on $[0, t_f]$ by

$$r(t) = -\frac{q(t)}{|q(t)|}$$

which is clearly continuous unless p_0 and q_0 are colinear. If p_0 and q_0 are colinear, then the locus of the relative trajectory

$$\left\{ (x(t), v(t)) - (x_0, v_0) \mid t \in [0, t_f] \right\}$$

is a straight line segment and r is constant with the exception of a single discontinuity at the time t_d where $p_0 t_d = q_0$ and where r reverses direction. The abnormal solution is a degenerate minimal fuel solution which occurs in the situation where the time optimal trajectory is the only trajectory in which the target point (x_f, v_f) can be reached in the time t_f. Since this solution is of the same form as a normal

III-4

solution over an interval of "full thrust", we shall restrict our attention to normal solutions.

We shall show that the normal solutions are nonsingular with the exception of the "pure translation" case presented in Section III. A normal nonsingular fuel optimal control function is characterized a.e. on $[0, t_f]$ by

$$r(t) = 0 \quad , \quad |q(t)| < 1$$

$$r(t) = -\frac{q(t)}{|q(t)|} , \quad |q(t)| > 1$$

and the set

$$\{ t \varepsilon [0, t_f] \mid |q(t)| = 1 \}$$

has Legesque measure zero. An optimal control function having a singular regime of positive Lebesque measure is characterized a.e. on that regime by

$$|q(t)| = 1 .$$

This implies that $p_0 = 0$ and $|q_0| = 1$ which shows that the solution must be singular a.e. on $[0, t_f]$. Since singular solutions can exist only in the case where $p_0 = 0$, they cannot occur in the general situation where p_0 and q_0 are not colinear.

We can therefore conclude that, if p_0 and q_0 are not colinear, the optimal solution is nonsingular, consists entirely of intervals of full-thrust and coast with, at most, one coast interval. The situation is depicted by Figure 1 in which the straightline segment defined by q punctures the unit ball. If the line segment does not puncture the unit ball, then the optimal solution has not coast interval.

FUEL OPTIMAL CONTROL
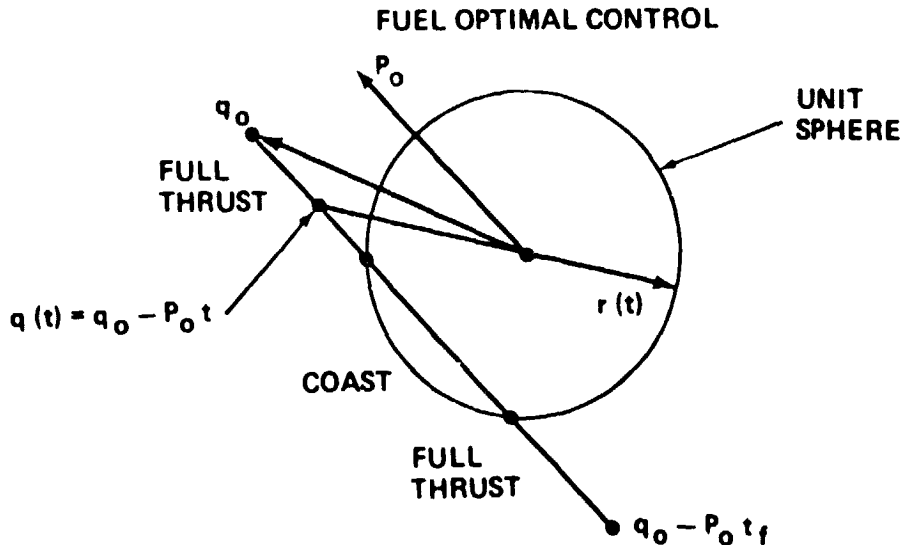
FIGURE I

If the flight begins with a thrust interval, then for each t on that interval, the velocity difference $\Delta v(t) = v(t) - v_0$ is given by

$$\Delta v(t) = \int_0^{t_f} \frac{(q_0 - p_0 t)\,dt}{|q_0 - p_0 t|} \quad .$$

This can be integrated to obtain

$$\Delta v(t) = \left[\frac{|q(t)| - |q_0|}{|p_0|^2} \right] p_0 - \frac{1}{|p_0|} \left[\sinh^{-1}\left(\frac{|p_0|}{|q_0|}\left(t - \frac{p_0 \cdot q_0}{|p_0|^2}\right)\right) + \sinh^{-1}\left(\frac{p_0 \cdot q_0}{|r_0||p_0|}\right) \right] r_0$$

where

$$r_0 = q_0 - \frac{p_0 \cdot q_0}{|p_0|^2} p_0 \quad .$$

It should be noted that r_0 and p_0 are orthogonal. There is a natural resolution of the velocity difference in the directions of p_0 and r_0. In general, we shall denote the components of vectors in the directions of p_0 and r_0 by subscripts p and r. We can integrate the expressions $\Delta v_p(t)$ and $\Delta v_r(t)$ to find the respective position differences $\Delta x_p(t)$ and $\Delta x_r(t)$.

The switch times which define the boundaries of the coast ...terval can be calculated as the roots of the equation

$$|q(t)| = 1 \quad .$$

These roots are found to be

$$k \pm \tau_0$$

where

$$k = \frac{p_0 \cdot q_0}{|p_0|^2} \qquad \text{and} \qquad \tau_0 = \frac{(1 - |r_0|^2)^{1/2}}{|p_0|}$$

This shows how p_0 and q_0 determine the type of motion which occurs. For example, if $0 < k - \tau_0$ and $k + \tau_0 < t_f$, then the trajectory is of the "thrust-coast-thrust" type, whereas if $k - \tau_0 < 0$ and $k + \tau_0 > t_f$ no thrust occurs at all. If $|r_0| \geq 1$ a coast interval cannot occur. Other similar observations can easily be made.

We now present the velocity and position differences in terms of the parameters k, τ_0, and $a = |r_0|/|p_0|$.

For compactness of notation, we define $b = (a^2 + k^2)^{1/2}$ and $c = (a^2 + \tau_0^2)^{1/2}$.

$$0 \le t \le k - \tau_0 \qquad (\text{thrust interval})$$

$$\Delta v_p(t) = (t^2 - 2kt + b^2)^{1/2} - b$$

$$\Delta v_r(t) = -a \left[\sinh^{-1}\left(\frac{t-k}{a}\right) + \sinh^{-1}\left(\frac{k}{a}\right) \right]$$

$$\Delta x_p(t) = \frac{1}{2} \left\{ \Delta v_p(t)(t-k) - a \Delta v_r(t) + (2 v_p(0) - \;)t \right\}$$

$$\Delta x_r(t) = \Delta v_r(t)(t-k) + a \Delta v_p(t) + v_r(0) t$$

$$k - \tau_0 < t < k - \tau_0 \qquad (\text{coast interval})$$

$$\Delta v_p(t) = \Delta v_p(k - \tau_0)$$

$$\Delta v_r(t) = \Delta v_r(k - \tau_0)$$

$$\Delta x_p(t) = \Delta v_p(t)(t - k + \tau_0) + \Delta x_p(k - \tau_0)$$

$$\Delta x_r(t) = \Delta v_r(t)(t - k + \tau_0) + \Delta x_r(k - \tau_0)$$

$$k - \tau_0 \le t \le t_f \qquad (\text{thrust interval})$$

$$\Delta v_p(t) = (t^2 - 2kt + b^2)^{1/2} - c + \Delta v_p(k + \tau_0)$$

$$\Delta v_r(t) = -a \left[\sinh^{-1}\left(\frac{t-k}{a}\right) - \sinh^{-1}\left(\frac{\tau_0}{a}\right) \right] + \Delta v_r(k + \tau_0)$$

$$\Delta x_p(t) = \frac{1}{2} \left\{ \Delta v_p(t)(t-k) - a \Delta v_r(t) + \left[2(v_p(0) + \Delta v_p(k + \tau_0)) - c\right](t - k - \tau_0)\right\} + \Delta x_p(k + \tau_0)$$

$$\Delta x_r(t) = \Delta v_r(t)(t-k) + a \Delta v_p(t) + \left[v_r(0) + \Delta v_r(k + \tau_0)\right](t - k - \tau_0) + \Delta x_r(k + \tau_0).$$

Sketches of flight path shapes for various initial and final conditions is presented in Figure 2. The optimum fuel consumption can also be found in terms of the above parameters by integrating $\dot{x}^0(t)$. We find

$$x^0(t_f) = \max(k - \tau_0, 0) + \min(t_f - k - \tau_0, t_f),$$

and in the case where $0 \leq k - t_0$ and $k + t_0 \leq t_f$, this becomes

$$x^0(t_f) = t_f - 2\tau_0.$$

A SKETCH OF FUEL OPTIMAL FLICHT PATH SHAPES
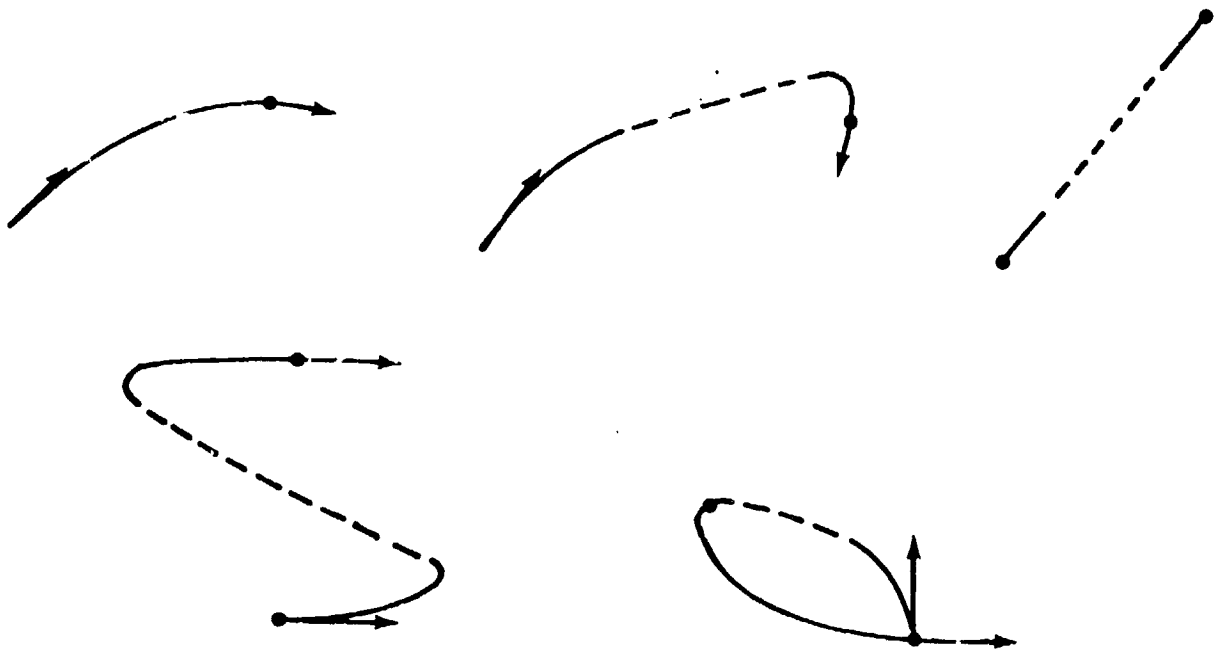
FIGURE 2

III. MINIMUM FUEL TRANSLATION

The minimum fuel translation problem is formulated exactly as in Section II with the exception that the three-dimensional unit ball which is the range of the functions in $\mathcal{U}$ is replaced by the interval $[-1, 1]$ and x_0, v_0, x_f, and v_f become points in $\mathbb{R}$. The constants p_0 and q_0 are also in $\mathbb{R}$ so that q is a linear scalar valued function of t.

The abnormal optimal control function is defined a.e. on $[0, t_f]$ by

$$r(t) = -\, \text{sgn}\, q(t)$$

whereas the normal optimal control is given a.e. on $[0, t_f]$ by

$$
\begin{aligned}
r(t) &= 0 & |q(t)| &< 1 \\
r(t) &= -\,\text{sgn}\, q(t) & |q(t)| &> 1 \\
r(t) &\in [-1, 0] & q(t) &= 1 \\
r(t) &\in [0, 1] & q(t) &= -1 \, .
\end{aligned}
$$

The later two expressions characterize the regimes where the optimal control function is singular. The existence of singular regimes of positive measure implies that $p_0 = 0$ and $q_0 = 1$ or $p_0 = 0$ and $q_0 = -1$. This shows that an optimal solution cannot have a singular regime without the solution being entirely singular, nor can the two types of singularities occur on the same solution.

Typical nonsingular normal solutions are of the "thrust-coast-thrust" type with switching times

$$\frac{q_0}{p_0} \pm \frac{1}{p_0}$$

which are the roots of the equation

$$|q(t)| = 1 \, .$$

In the phase plane the thrust portions of the trajectory have the form of parabolic segments and the coast phases are horizontal line segments. The abnormal solutions have the parabolic segments only. A phase plane sketch of the trajectories in the region where the origin is reachable in a specified time is presented in Figure 3.
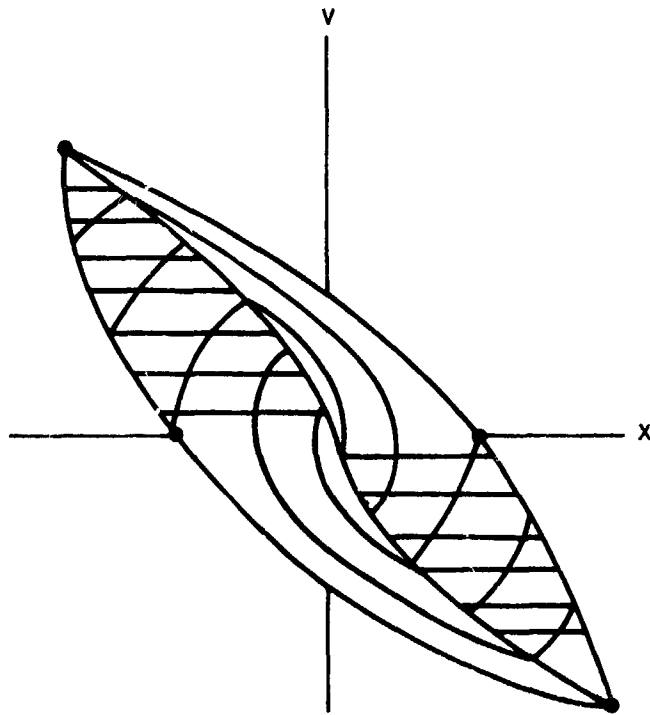
FIGURE 3.   PHASE PLANE SKETCH OF FUEL OPTIMAL
             TRAJECTORIES IN REGION WHERE ORIGIN
             IS REACHABLE IN A SPECIFIED TIME

There are four basic types of optimal trajectories. We illustrate these in Figure 4 for the case where

$$0 < v_0 < \sqrt{-x_0} \quad \text{and} \quad v_f = x_f = 0 .$$

For two curves having these end conditions, it can be proved that if one is entirely below the other in the plane, then it takes a longer time to reach the terminal point. For this reason, the trajectories labeled I IV require increasingly more time to reach the origin. The abnormal arc I is time optimal. For a smaller value of t_f the origin is not reachable. Arc II represents a typical "thrust-coast-thrust" trajectory. Arc III occurs if the smaller of the two switching times is zero. If this switching time is negative, then infinitely many singular solutions IV occur. These consist of any trajectory in which

$$u(t) \in [-1 , 0] \text{ a.e. on } [0, t_f] \text{ and } x(t_f) = v(t_f) = 0.$$

Included, in particular, is the situation where the position and velocity follow Arc III, arriving at the origin at time $t_1 < t_f$ and remaining there for the interval $t_1 \leq t \leq t_f$. Since this singular arc is above any other singular arc in the region of the phase plane under consideration, it arrives at the origin in the shortest time. The argument that singular arcs must always lie between Type III arcs and the x-axis can be used to show that singular solutions are possible only in the region

$$S = \{ (x, v) \in \mathbb{R}^2 \mid 0 < v < \sqrt{-x} \text{ or } -\sqrt{x} < v < 0 \} .$$

However, singular solutions exist for any initial point in this region if t_f is sufficiently large. If one agrees to use only the "time optimal" singular arc as representative of the singular arcs, then for any initial point the optimal trajectory to the origin consists of full thrust and coast intervals with at most two full thrust intervals.

By computing $x^0(t_f)$ for the piecewise constant control function r we can show that the optimal fuel consumption decr es with increasing t_f and can be made arbitrarily close to the value I if t_f is taken sufficiently large. The shape of the optimal fuel   sumption curve depends on whether or not $(x_0, v_0) \in S$, the region where singular solutions occur for sufficiently large values of t_f. The two basic shapes are presented in Figure 5.
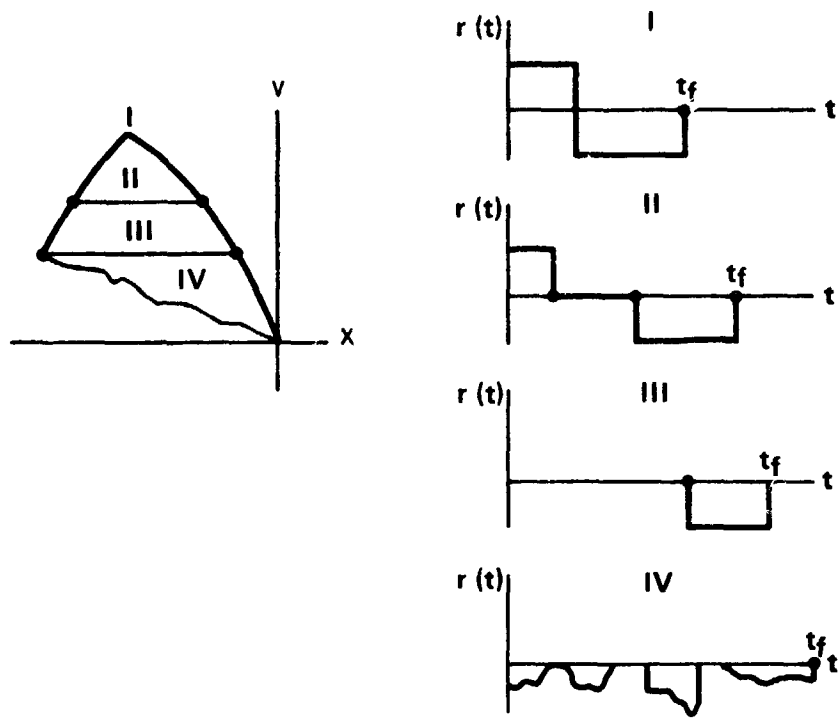
FIGURE 4. TYPES OF FUEL OPTIMAL SOLUTIONS
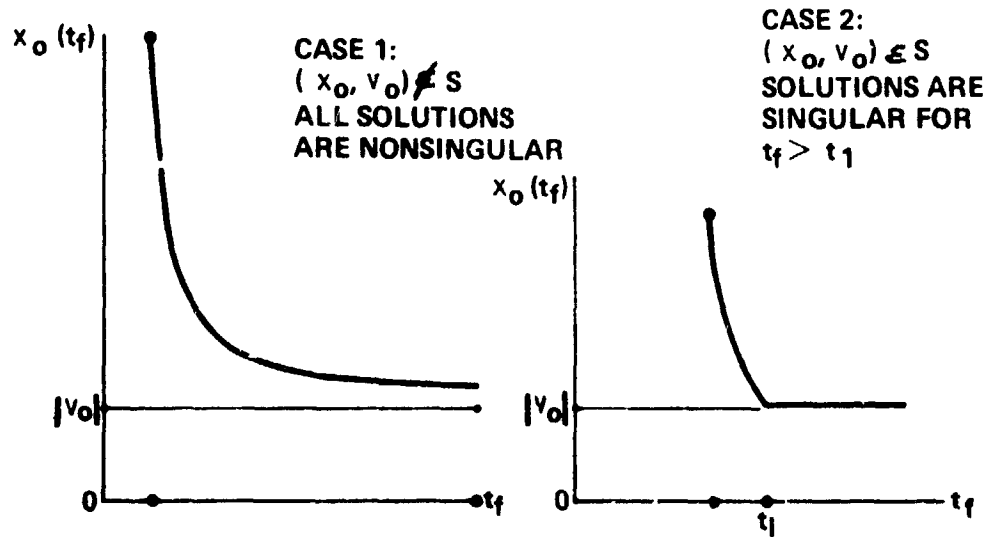
FIGURE 5. OPTIMAL FUEL CONSUMPTION CURVES

IV. COMPARISON WITH THE INTEGRAL SQUARE CRITERION

We formulate the minimum integral square criterion problem exactly as the minimum fuel problem of Section II, except that the criterion $x^0(t_f)$ to be minimized is defined by

$$\dot{x}^0(t) = \frac{1}{2} r(t) \cdot r(t)$$

a.e. on $[0, t_f]$.

The Hamiltonian is

$$\mathcal{H} = \frac{1}{2} q^0 r(t) \cdot r(t) + p(t) \cdot v(t) + q(t) \cdot r(t)$$

where the adjoint system and the abnormal optimal solutions are exactly the same as in Section II.

We see that the normal optimal solutions are all nonsingular and the optimal control functions are described a.e. on $[0, t_f]$ by

$$r(t) = - q(t) \qquad |q(t)| < 1$$

$$r(t) = - \frac{q(t)}{|q(t)|} \qquad |q(t)| > 1 \ .$$

The solution is of the "full thrust-linear-full thrust" type as depicted by Figure 6. The similarities with Figure 1 are apparent.

FIGURE 6. INTEGRAL SQUARE CRITERION OPTIMAL CONTROL
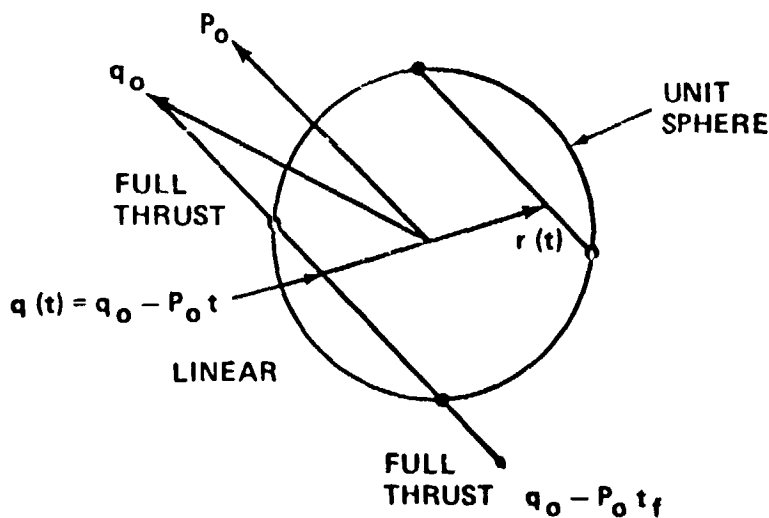
The solutions in the "full thrust" intervals are of the same form as those presented in Section II. Letting t_0 be the earliest time in which the "linear" interval is entered, the solution in this interval is

$$v(t) = P_0 \frac{(t - t_0)^2}{2} - q_0 (t - t_0) + v(t_0)$$

$$x(t) = P_0 \frac{(t - t_0)^3}{6} - q_0 \frac{(t - t_0)^2}{2} + v(t_0)(t - t_0) + x(t_0).$$

The "optimal translation" case in which p_0 and q_0 are colinear is mentioned briefly for comparison with the fuel optimal case. A solution where $v_0 = v_f = 0$ and where the control function was linear ($|q(t)| < 1$ for all $t \in [0, t_f]$) was compared with a fuel optimal solution having the same t_f and initial and final conditions. A phase plane sketch of the trajectories is presented in Figure 7. As t_f increases, the maximum value of $x_2(t)$ for the "integral square optimum" approaches one and one-half times that of the "minimum fuel optimum" and the respective values of $x^0(t_f)$ approach the same ratio.
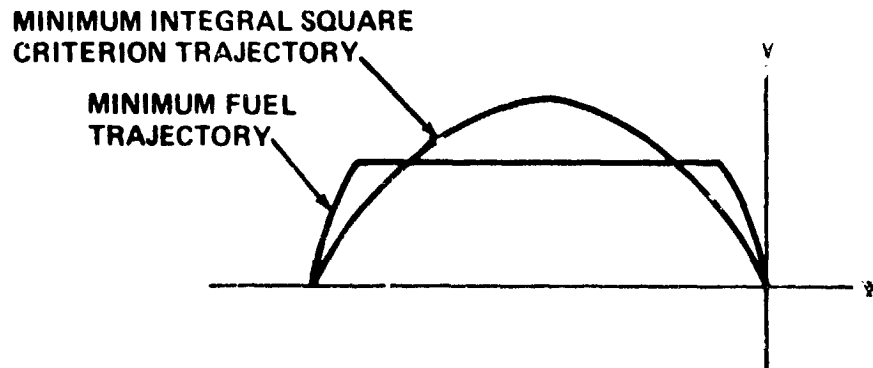
FIGURE 7. PHASE PLANE COMPARISON OF OPTIMAL TRAJECTORIES HAVING DIFFERENT PERFORMANCE CRITERIA

III-14

V. OTHER MODELS

In this section we consider some generalizations of the problem in Section II.

First, we consider the modifications which appear when the motion is relative to a point moving in orbit. The state variable equations take the following form:

$$\dot{x}^0(t) = |r(t)|$$
$$\dot{x}(t) = v(t)$$
$$\dot{v}(t) = A x(t) + B v(t) + r(t)$$

where A and B are 3x3 matrices. (For a circular orbit, A and B represent the centripetal and coriolis effects, respectively, e.g.,

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\Omega^2 & 0 \\ 0 & 0 & \Omega^2 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 & 2\Omega \\ 0 & 0 & 0 \\ -2\Omega & 0 & 0 \end{pmatrix}$$

where Ω is the orbital angular speed.)

For this problem, the adjoint system is defined as follows:

$$\dot{p}(t) = -A^T q(t) \qquad\qquad p(0) = p_0$$
$$\dot{q}(t) = -B^T q(t) - p(t) \qquad\qquad q(0) = q_0 .$$

The optimal controls are characterized exactly as in Section II. The form of the function q, however, leads to some qualitative differences in the solutions. Although the normal nonsingular optimal solutions still consist of full thrust and coast intervals, there is no longer the restriction of at most one coast interval. The restriction that the thrust vector can turn through no more than half a circle is likewise removed since q(t) no longer moves in a straight line. As in Section II, singular regimes must include the entire interval $[0, t_f]$. Since singular solutions are characterized by

$$|q(t)| = 1 \qquad (t \in [0, t_f])$$

a necessary condition for a singular solution is

$$p(t) \cdot q(t) + q(t) \cdot B^T q(t) = 0 \qquad (t \in [0, t_f]).$$

III-15

For the special coriolis matrix B considered previously, this implies that p(t) and q(t) must be orthogonal on $[0, t_f]$.

The previous models have all represented spacecraft in which the effects of thrusts could be represented by a single thruster acting through the center of gravity. Fuel consumed to apply torques to turn the vehicle was not considered. The diagram in Figure 8 depicts a planar model relative to a circular orbit with fixed axial and lateral thrusters which do not act through the center of gravity.
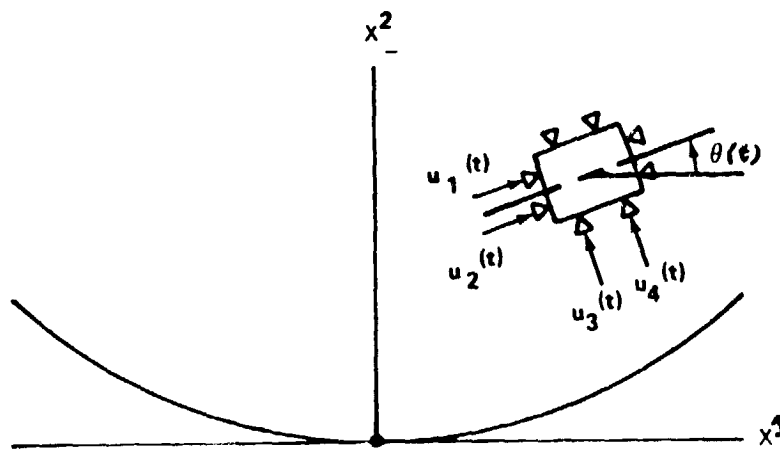
FIGURE 8. SPACECRAFT WITH FIXED THRUSTERS

This model is described by the following dynamical system:

$$\dot{X}^0(t) = \sum_{i=1}^{4} |u_i(t)|$$

$$\dot{X}^1(t) = V^1(t)$$

$$\dot{X}^2(t) = V^2(t)$$

$$\dot{V}^1(t) = (u_1(t) + u_2(t))\cos\theta(t) - (u_3(t) + u_4(t))\sin\theta(t) + 2\Omega V^2(t)$$

$$\dot{V}^2(t) = (u_1(t) + u_2(t))\sin\theta(t) + (u_3(t) + u_4(t))\cos\theta(t) - 2\Omega V^1(t) + 3\Omega^2 X^2(t)$$

$$\dot{\theta}(t) = \omega(t)$$

$$\dot{\omega}(t) = \sum_{i=1}^{4} \ell_i u_i(t)$$

where $u_i(t) \in [-1, 1]$ a.e. on $[0, t_f]$ ($i = 1,\ldots,4$). For this model also it can be seen that the normal nonsingular optimal control functions are of the "full thrust--coast—full thrust" type and the abnormal solutions are "full thrust with switching." Additional qualitative information about the solutions of this problem are difficult to obtain. Numerical solutions may be found by iterative methods for solution of the two-point boundary value problem.

For a problem in deep space ($\Omega = 0$) we consider the following transformation of variables:

$$\dot{y}^1(t) = w^1(t)$$
$$\dot{y}^2(t) = w^2(t)$$
$$\dot{w}^1(t) = \dot{v}^1(t) \cos\theta(t) + \dot{v}^2(t) \sin\theta(t)$$
$$\dot{w}^2(t) = -\dot{v}^1(t) \sin\theta(t) + \dot{v}^2(t) \cos\theta(t).$$

In terms of the transformed variables the model becomes

$$\dot{x}^0(t) = \sum_{i=1}^{4} |u_i(t)|$$
$$\dot{y}^1(t) = w^1(t)$$
$$\dot{y}^2(t) = w^2(t)$$
$$\dot{w}^1(t) = u_1(t) + u_2(t)$$
$$\dot{w}^2(t) = u_3(t) + u_4(t)$$
$$\dot{\theta}(t) = \omega(t)$$
$$\dot{\omega}(t) = \sum_{i=1}^{4} \gamma_i \, u_i(t).$$

If one has the initial and final conditions for these variables, one has a considerably simpler problem than the original. This problem is <u>not</u> equivalent to the original problem, however, and is of no value unless one is interested in the variables, y^1, y^2, w^1, w^2. An optimal control function for the original system is not necessarily optimal for the corresponding boundary values of y and w; however, an optimal control function for the transformed problem is optimal for the corresponding boundary values of x and v.

We shall consider the simplification that occurs if

$$u_3(t) = u_4(t) = 0$$

a.e. on $[0, t_f]$. In this case y^1 and v^1 can be interpreted respectively as distance traveled and speed. If one imposes boundary conditions only on the distance traveled and speed, one has (for $\gamma_1 = 1$ and $\gamma_2 = 1$) the following system:

$$\dot{x}^0(t) = |u_1(t)| + |u_2(t)|$$
$$\dot{y}^1(t) = w^1(t)$$
$$\dot{v}^1(t) = u_1(t) + u_2(t)$$
$$\dot{\theta}(t) = \omega(t)$$
$$\dot{\omega}(t) = -u_1(t) + u_2(t) .$$

This problem is easily solved by the transformation

$$z^1(t) = \tfrac{1}{2}\left(y^1(t) - \theta(t) \right)$$
$$z^2(t) = \tfrac{1}{2}\left(y^1(t) + \theta(t) \right)$$

which decouples the problem into an equivalent set of two problems of the horizontal translation type,

$$\dot{x}_i^0(t) = |u_i(t)|$$
$$\dot{z}^i(t) = s^i(t)$$
$$\dot{s}^i(t) = u_i(t) \qquad\qquad (i = 1, 2) .$$

A sketch of a solution of this problem can be presented in the form of a "double trajectory" in the phase plane. This is indicated in Figure 9 with a sketch of the corresponding flight path shape.
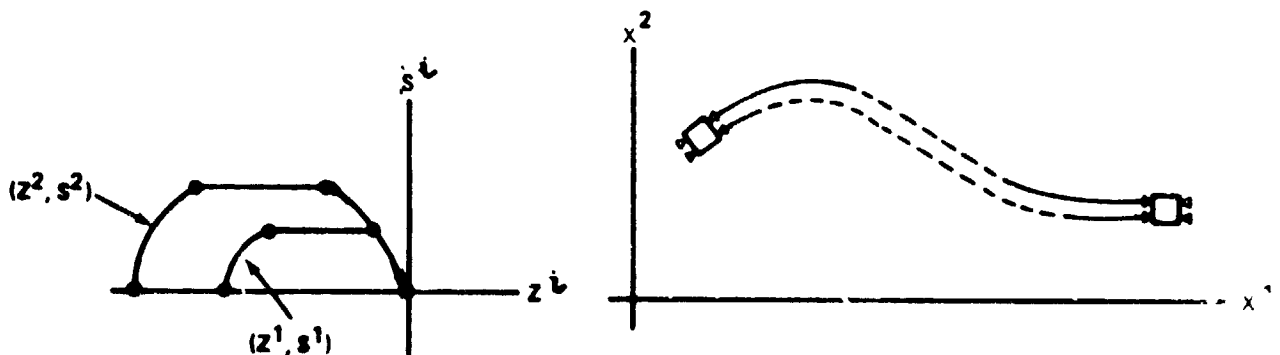
FIGURE 9. DOUBLE TRAJECTORY AND FLIGHT PATH SHAPE

VI. CONCLUSIONS

In each of the minimum fuel problems investigated in this study, the optimal solution consisted of intervals of full thrust and coast with the exceptions of abnormal solutions (which contained only full thrust and reverse full thrust) and singular solutions. For this reason the design of "throttleable" thrusters may not be required for fuel-optimal performance.

Optimal controllers derived under deep space assumptions may have some validity in terms of motion relative to a point in orbit if the spacecraft is near the point and if the time interval is short, but qualitative differences regarding the number of coast intervals and the limit on the total "angle of turn" do exist.

Cases in which the vehicle is stationary at the initial and terminal time may be of the most practical importance for rendezvous and docking problems. According to the model used in Section II, the solution for these cases is "pure translation" where the vehicle thrusts, coasts, then thrusts in the reverse direction. This model, which does not consider fuel losses due to turning, admits a "jump" discontinuity in the attitude angle at each end of the flight.

For vehicles which use separate thrusters for translation and steering, a sub-optimal approximation to the above solution could consist of a "turning maneuver" which is used to point the vehicle to the target point followed by a "thrust-coast-reverse thrust" translation to the target point and finally, another "turning maneuver" to reach the specified final attitude. The two "turning maneuvers" could be fuel-optimal rotation which is mathematically equivalent to the problem discussed in Section III and which is of the form "torque-coast-reverse torque." The difference between the fuel cost of this approximation and the absolute optimum program is less than or equal to the fuel consumed during the two "turning maneuvers." Spacecraft which do not use separate thrusters for translation and steering may obtain a better result with a thrusting program of the type indicated in Figure 9.

REFERENCES

1. L. W. Neustadt,"A General Theory of Minimum-Fuel Space Trajectories,"
   SIAM J. Control, (1965), pp. 317-356.

2. E. B. Lee and L. Markus, Foundations cf Optimal Control Theory,
   Wiley, 1967.

3. G. Leitmann, An Introduction to Optimal Control, McGraw-Hill, 1966.

4. G. M. Ewing,"A Fundamental Problem of Navigation in Free Space,"
   Quart. Appl. Math. 18 (1961), pp. 355-362.

5. I. Flügge-Lotz and H. Marbach,"The Optimal Control of Some At-itude
   Control Systems for Different Performance Criteria," Trans. ASME J.
   Appl. Mech. (1964), pp. 107-115.

6. J. S. Medicch,"On Minimal Fuel Satellite Attitude Controls," Preprints
   o; the 1963 JACC.

7. L. Schwartz,"Minimum Energy Attitude Control for a Class of Electric
   Propulsion Devices," Proceedings, 1964 JACC.

8. D. A. Conrad,"Minimum Fuel Closed-Loop Translation," AIAA J. (May 1965).

9. W. F. Keller,"Study of Spacecraft Hover and Translation Modes above
   the Lunar Surface," AIAA Preprint 64-341.

10. H. O. Ladd and R. Friedland,"Minimum Fuel Control of a Second Order
    Linear Process with a Constraint on Time-to-Run," Trans. ASME J.
    Appl. Mech. (1964) pp 165-176.

11. D. R. Snow,"Singular Optimal Controls for a Class of Minimum Effort
    Problems," SIAM J. Control, (1964) pp. 203-219.

12. M. Athans,"Minimum-Fuel Control of Second-Order Systems with Real
    Poles," 1963 JACC, pp. 232-240.

13. A. J. Craig and I. Flügge-'otz,"Investigation of Optimal Control
    with Minimum-Fuel Consumption Criteria for a Fourth-Order Plant
    with Two Control Inputs; Synthesis of an Efficient Sub-Optimal
    Control," 1963 JACC.

14. M. Athans,"On Optimal Control of Self Adjoint Systems," 1963 JACC.

15. W. C. Grimmell,"The Existence of Piecewise Continuous Fuel Optimal
    Controls," SIAM J. Control (1967), pp. 515-519.

REFERENCES (Cont'd)

16. D. L. Grey and M. Athans,"Computation of Fuel-Optimal Control via Newton's Method - Theory,"1968 JACC.

17. A.E. Bryson,"Linear Feedback Solutions for Minimum Effort Interception, Rendezvous, and Soft Landing,"AIAA J. (Aug. 1965), pp. 1542-1544.

18. L. W. Neustadt,"Minimum Effort Control Systems,"SIAM J. Control (1962) pp. 16-31.

19. J. S. Meditch,"Synthesis of a Class of Linear Feedback Minimum Energy Controls,"IEEE Trans. Automatic Control, (1963), pp. 376-378.

20. B. H. Billik, 'ome Optimal Low-Acceleration Rendezvous Maneuvers," AIAA J., (1964), pp. 510-516.

21. J. Warga, Optimal Control of Differential and Functional Equations, Academic Press, 1972.

22. T. E. Carter,"Singular Regimes in Certain Classes of Relaxed Control Problem,, SIAM J on Control and Optimization," (1976), pp. 359-386.

NASA/ASEF SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

SOLIDIFICATION PHENOMENA
OF BINARY ORGANIC MIXTURES

Prepared by: Kichoon Chang, Ph.D.

Academic Rank: Associate Professor

University and Department: Alabama A&M University
Department of Chemistry

NASA/MSFC:
    Division: Space Processing
    Branch: Solid State

MSFC Counterpart: Donald O. Frazier, Ph.D.

Date: September 30, 1981

Contract No.:   NGT 01-008-021
The University of Alabama in Huntsville

SOLIDIFICATION PHENOMENA
OF BINARY ORGANIC MIXTURES

BY

Kichoon Chang
Associate Professor of Chemistry
Alabama A&M University
Normal, Alabama

ABSTRACT

The immediate concern in this project is to study the coalescence rates and motion of liquid bubbles in binary organic mixtures.

It is believed that several factors such as temperature gradient, composition gradient, interfacial tension, and densities of the two phases play important roles in separation of phases of immiscible liquids.

Thus, an attempt will be made to study the effect of initial compositions on separation rates of well-dispersed organic mixtures at different temperatures and, ultimately on the homogeneity of solidification of the immiscible binary organic liquids. These organic mixtures will serve as models for metallic pseudo binary systems currently under study.

Currently two systems are under investigation in this laboratory: ethyl salicylate - diethyl glycol and succinonitrile - water.

FIGURE   LIST OF FIGURES

LIST OF TABLES

I. Introduction

Immiscible binary liquid mixtures will invariably separa-
te into more than one homogeneous phases below the convolu-
tion temperature within certain composition ranges. During a
vigorous shaking the mixture becomes apparently homogeneous.
Upon standing, however, the single phase solution immediate-
ly breaks into two or more phases by coalescence of tiny
droplets.

Several factors are well known to influence the separa-
tion of phases. The important factors are temperature, com-
position of the mixture, densities of the phases and inter-
facial tensions between the phases in addition to the chem-
ical characteristics of the substances.

Temperature is one important factor in achieving the ho-
mogeneous single phase of the binary system. The heterogen-
eous mixture will eventually turn into a homogeneous phase
above the convolution temperature.

The compositions of the mixture will define the phase
boundaries within which each phase is homogeneous. A com-
plete phase diagram is required to study the immiscible
mixtures extensively.

Densities of the different phases in the immiscible
mixture play an important role in separation of the phases.
A substantial difference in density between the phases will

make the separation rapid and stable.

On the other hand, the function of the interfacial tensions between the separate phases appears to be most complicated and least understood in relation to other factors. Since the interfacial tension is always affected by the composition, temperature and impurities present in the phases, our work in this project has been concentrated to the study of the variation of the interfacial tension as a function of temperature, composition and density of the phases.

For the investigation of solidification phenomena of binary organic mixtures the binary system of succinonitrile and water (SN-H$_2$O) was assigned for detailed study partly because succinonitrile freezes like metals.[1] Other compounds that freeze like metals are also known[1]; a few examples are t-butyl bromide, t-butyl chloride, carbontetrachloride, cyclohexane, and cyclohexanol.

One of the first relatively detailed studies of SN-H$_2$O system was published in 1897 in Germany.[2] The binary system is a monotectic. The solidification behavior of a monotectic system is known to be related t the height of the miscibility gap in the system.[3] However, the SN-H$_2$O system had not been extensively investigated as to the role of interfacial tensions and densities of the phases.

Prior to the experimental measurement of this system, a considerable time had been spent for literature survey on motion and coalescence of bubbles, temperature coefficient

of surface tension and solification of other systems.[4]

Succinonitrile exists as both liquid and solid in the convenient temperature range to work with, the melting point being 54.5°C. The system of succinonitrile and water has the monotectic point of 18°C near 28 mol% water.[2] For the understanding of solidification phenomena of the binary systems of succinonitrile and water, the coalescence and motion of the liquid droplets and the Marangoni effect of the system are yet to be investigated.

II. Objectives

The objectives of this project are two-fold: the immediate objective is to determine how the interfacial tension is related to the overall composition and densities of the phases in immiscible binary organic mixtures at various temperatures. The other objective is to explore the causes of composition gradients during solidification of the binary dispersions. During this project period an attempt was made first to determine the temperature coefficient of the interfacial tensions of the phases and the density variation with respect to the composition and temperature.

III. Measurements of Physical Quantities

The motion and coalescence of bubbles (or droplets) and separation of phases in the immiscible mixtures are affected by several factors. For the first task of measurement of these factors, two most prominent ones were selected: density of the phases and interfacial tension at several compositions and temperatures.

It is relatively easy to prolong the state of dispersion of the immiscible liquid solutions of succinonitrile and water after thorough mixing by closely equalizing the densities of both phases through the adjustment of H_2O/D_2O ratio in the mixture. It is also simple to reverse the upper and lower phases by controlling the densities of the two phases.

(a) Sample Preparation

This project involves purification of succinonitrile. During this project period, however, it was not possible to obtain sufficient amount of ultrapurified succinonitrile in time for the sample preparation. Hence succinonitrile used throughout this project was purified by vacuum distillation. Water was also purified by distillation in this laboratory. Both of these were not subjected to any further purification and they were not analyzed for any impurities. Each binary

mixture was prepared by weighing accurately solid succino-
nitrile and mixing with a measured volume of water.

(b) Measurement of Densities of Phases

Densities of both phases of SN-H_2O systems were measured
at several temperatures for the following compositions:
1. 19.99 mol% SN/H_2O
2. 12.57 mol% SN/8.00 wt.% D_2O/H_2O
3. 19.99 mol% SN/D_2O
4. 19.99 mol% SN/50.0 mol% D_2O/H_2O

Some of the apparently conclusive trends are shown in
Tables I and II and in Figs. 1, 2, and 3.

A Cahn electrobalance and glass bobs were used for the
density measurement. The Cahn balance was calibrated with a
few calibration weights for the range of the anticipated
weights of the bobs in the liquids. Each bob was weighed in
the air with a Metler balance within 0.01 mg. The volume of
the bob was then measured in distilled water at a known tem-
perature using the Cahn balance.

For the measurement of the density of each phase, the bob
must be placed completely in the liquid phase. This requires
that the depth of the phase be deep enough to eliminate any
interference from the interface. This was not a small pro-
blem when the interface is shaped into a convex or concave.
Only a much larger amount of the sample solution in a larger
cell will minimize   the interference from the interface

Table I. Apparent Densities of Water-
Succinonitrile Mixtures

Mixture	$T^{o}C$	Density, g/cm^3 SN-rich Layer	Density, g/cm^3 Water-rich Layer
19.99 mol% SN* in H_2O	20.0	1.0955	1.0842
	22.0	1.0510	1.0384
	24.0	1.0390	1.0295
	26.0	1.0291	1.0206
12.57 mol% SN in 8.00 wt.% D_2O/H_2O	19.8	1.0405	1.0330
	22.0	1.0360	1.0259
	24.0	1.0337	1.0232
	24.9	1.0290	1.0233
	26.0	1.0223	1.0217

Table II. Apparent Densities of Water-
Succinonitrile Mixtures

Mixture	$T^{o}C$	Density, g/cm^3 SN-rich Layer	Density, g/cm^3 Water-rich Layer
19.99 mol% SN in D_2O	22.0	1.0373	1.1149
	24.0	1.0913	1.1709
	2?.0	1.1260	1.2097
19.99 mol% SN in 50.0 mol% D_2O/H_2O	24.3	1.0462	1.0797
	25.1	1.0462	1.0825
	26.0	1.0323	1.0688

* Succinonitrile.

in the cell. The interference was indicated very sensitively in the density measurement.

Fig. 1 clearly shows that densities of the binary solutions with water or a mixture of water and D_2O increase with decreasing temperature as might be anticipated. Only the variation with temperature is far less with a mixture of water and D_2O. It is seen in the figure that the neutral density (equal density) of both phases in the binary systems is approached at a low temperature (near 20°C) with water and at a high temperature (near 26°C) with H_2O-D_2O mixtures. From this observation it may be said that the temperature at which the neutral density of the two phases is obtained can be adjusted by varying the ratio of D_2O to H_2O for a given composition of the binary solution.

As the D_2O to H_2O ratio increases the density of water-rich phase becomes greater than that of SN-rich phase as shown in Fig. 2. This is expected in view of the relative densities of D_2O and H_2O. The composition of the water-rich phase in 19.99 mol% SN/H_2O and 19.99 mol% SN/50.0 mol% D_2O/ H_2O were not analyzed for comparison. The comparison of the solubility of succinonitrile in water and in a mixture of H_2O and D_2O would be interesting.

The effect of D_2O on the density of the phases shown most remarkably with respect to temperature in Fig. 3. It is obvious that more data should be obtained in order to see

Fig. 1. Densities of Phases

————— 19.99 mol % SN/H_2O

- - - - 12.57 mol % SN/8.0 wt % D_2O/H_2O
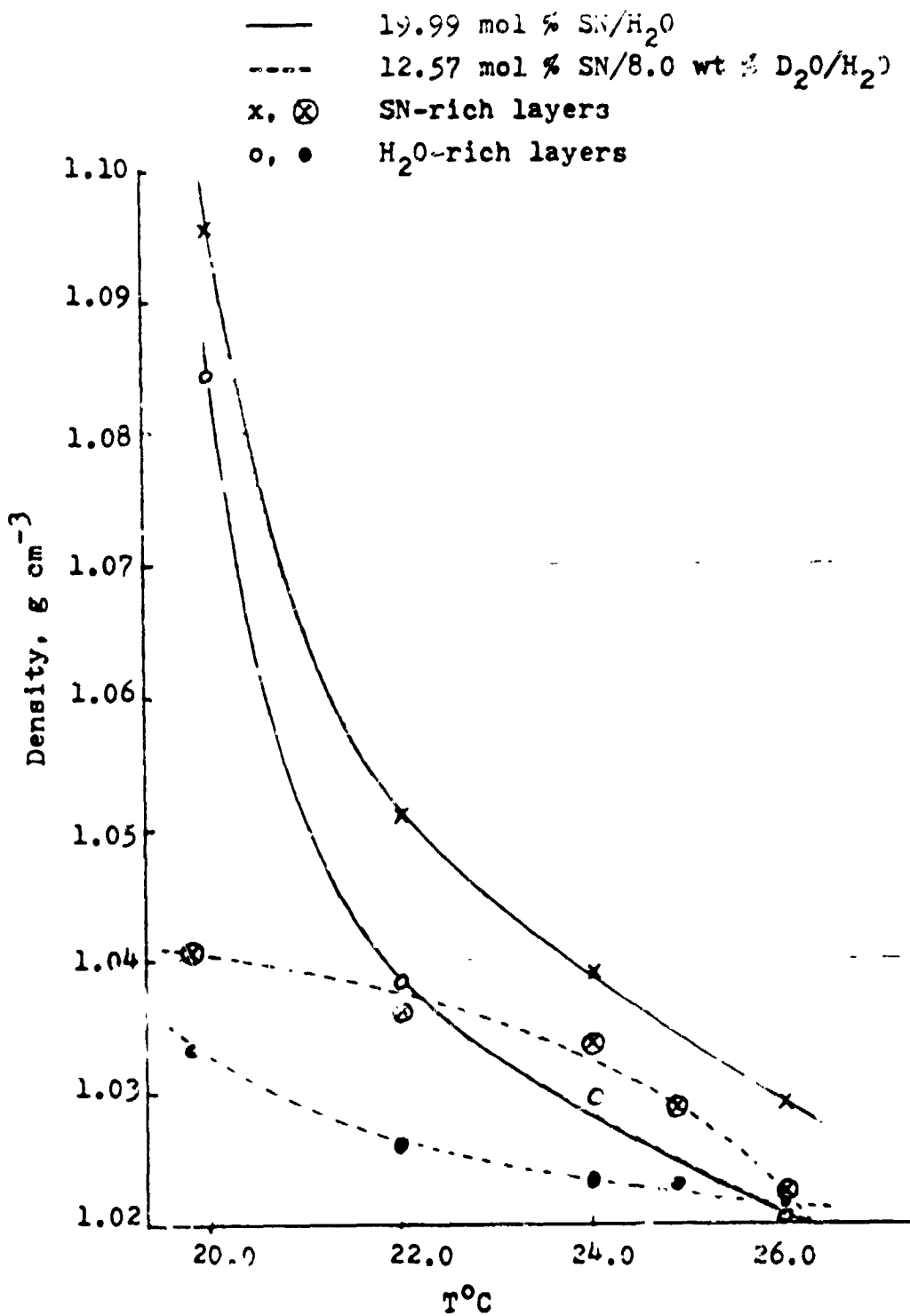
x, ⊗ SN-rich layers

o, ● H_2O-rich layers

Fig. 2. Densities of Phases for 19.99
mol % SN in 50.0 mol % D_2O/H_2O

x   SN-rich layer
o D_2O/H_2O-rich layer
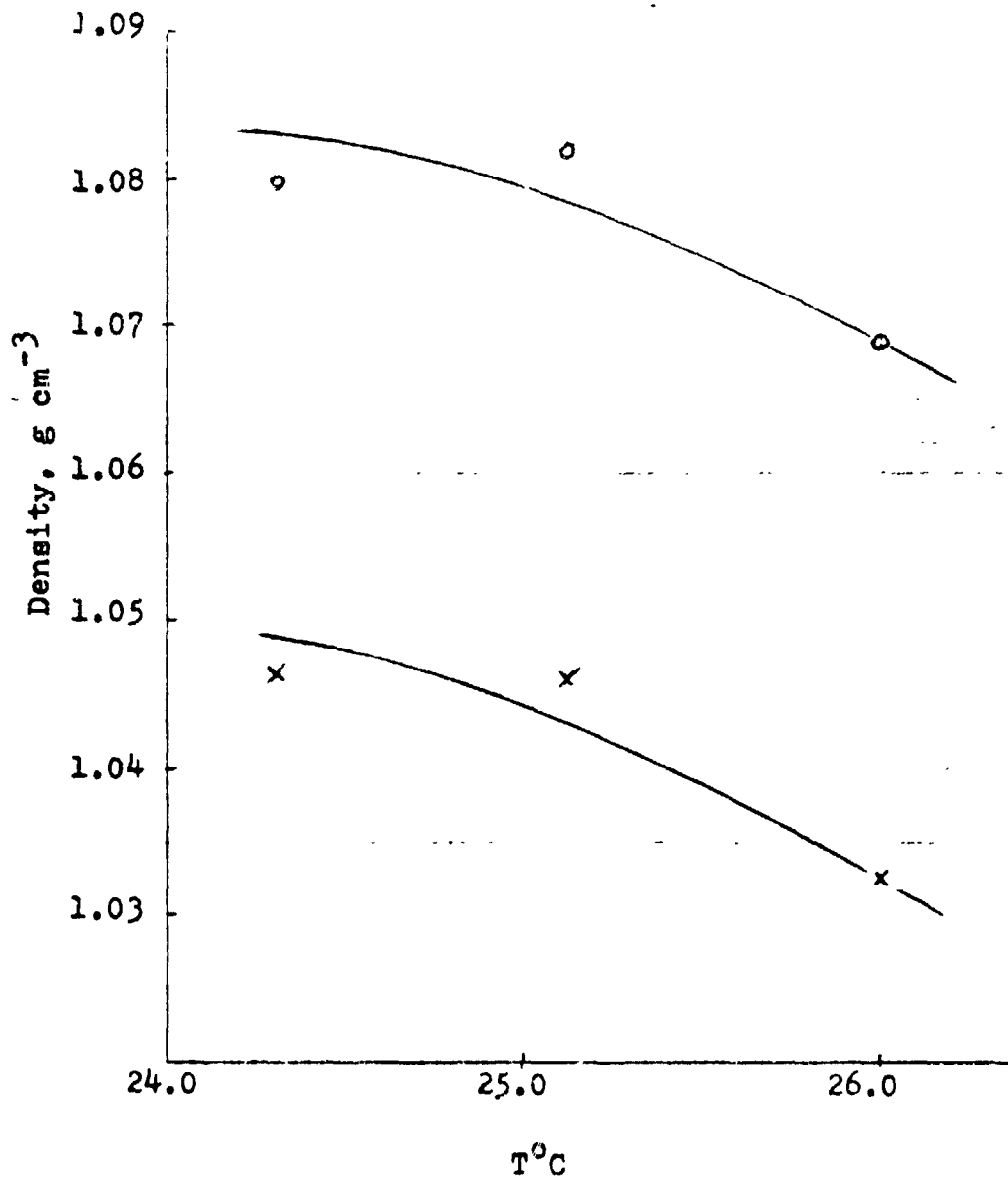
Fig. 3. Densities of Phases for
19.99 mol % SN/D$_2$O

x   SN-rich layer
o   D$_2$O-rich layer
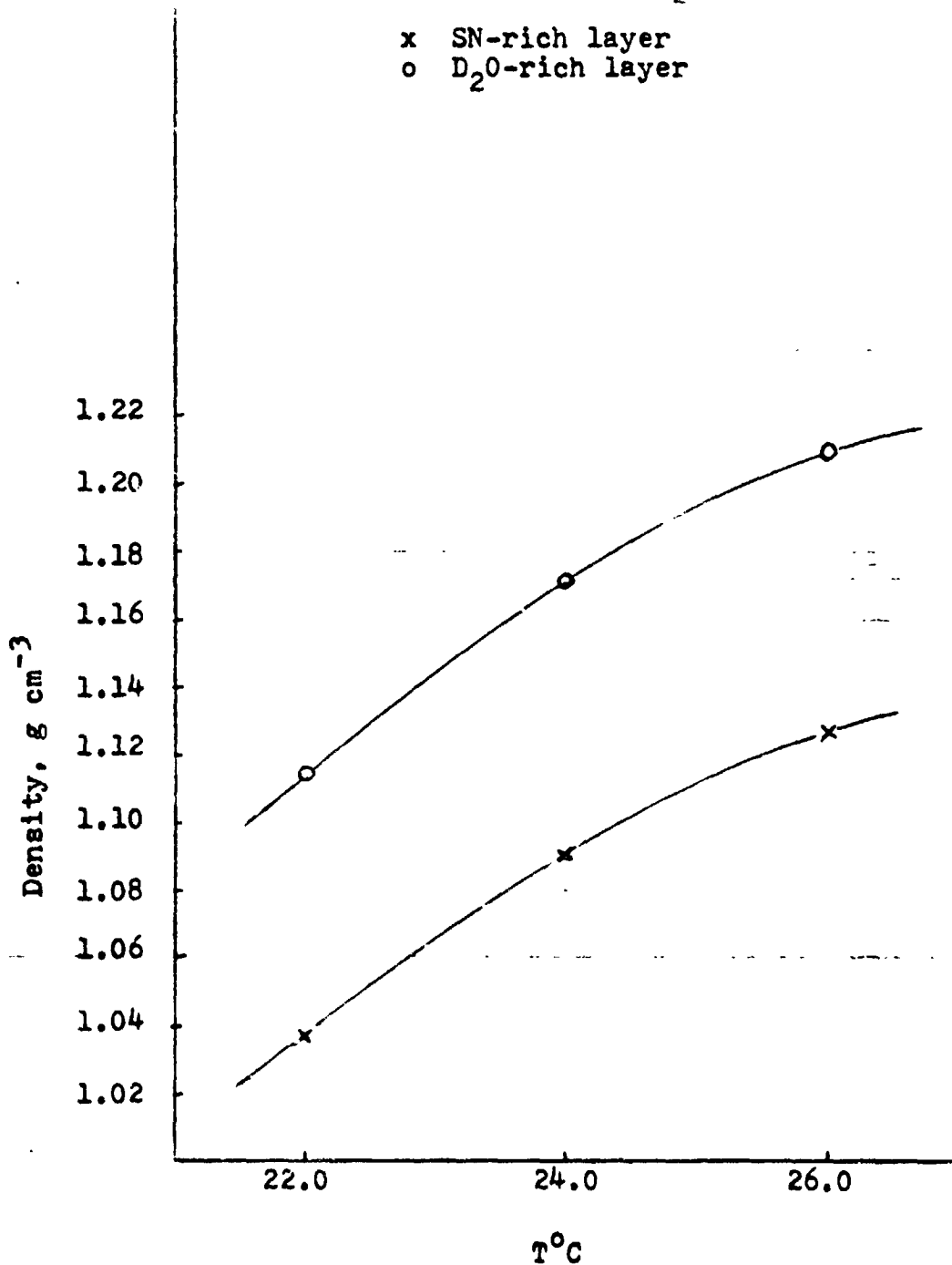
more significant variation of the density with temperature.
Yet, the increasing density of the phases with temperature
has not been anticipated. The higher density in D_2O-rich
layer is simply due to the much greater density of D_2O than
that of succinonitrile in the absence of water.

(c) Measurement of Interfacial Tension

A systematic measurement of the interfacial tensions be-
tween the two phases as functions of temperature, the overall
composition and density of both phases was attempted. How-
ever, several difficulties in the measurement with the Cahn
electrobalance were encountered. As in the density measure-
ment, the phase equilibria could be reached in many cases
only after many hours, sometimes more than a day of stand-
ing. This was especially true with the system of nearly
neutral density of the phases. The coalescence of the drop-
lets and separation into phases was extremely slow after
vigorous mixing of the solution in the cell. Temperature
raise of a few degrees would not speed up the equilibrium
appreciably.

The size of the cell itself was not a small problem in
some mixtures. When the entire interface is shaped into a
convex or concave, the thickness of one phase becomes too
thin and the available area too narrow in the cell for the
ring to be placed for the interfacial measurement. In this
case it was very difficult or sometimes impossible to keep

the ring in the center of the phase due to the strong attraction toward the glass wall.

Apparent interfacial tensions were measured for two systems at several temperatures and they are listed in Table III. Figs. 4 and 5 show the variation of the interfacial tension with temperature for 19.99 mol% succinonitrile in water and 19.99 mol% succinonitrile in D_2O, respectively. In both systems, the solution froze near $19^{\circ}C$ when the ring was immersed in the upper layer. The monotectic temperature of the system in water is about $18^{\circ}C$.

The apparent interfacial tensions were calculated according to the formula, $$\gamma_{appar.} = \frac{Mg}{2L}$$ where M is the mass in gram in the Cahn balance required to detach the ring from the interface, g the gravitational acceleration (979.656 cm s^{-2}) and L the mean circumference of the ring (6.005 cm).

For the estimation of the true interfacial tension, the densities of both phases of a system are required to calculate the correction factor. However, since it was not always possible to measure both densities as described in (b) in many cases, the true interfacial tensions were not calculated. When it was possible, the true tension calculated was greater than the apparent tension by about 5-6%. This small difference indicates that the apparent tensions can still illustrate the general tendency of the true interfacial tensions

Table III. Apparent Interfacial Tensions

Mixture	T^{o}C	$\gamma_{appar.}$
19.99 mol% SN* in Water	20.1	3.2073
	20.9	3.1910
	22.0	2.9994
	22.4	3.0752
	24.4	2.4895
	26.1	2.4840
19.99 mol% SN in D_2O	19.0	4.5581
	19.9	4.5288
	20.0	4.5190
	21.0	4.4146
	22.0	4.2286
	24.0	3.8273
	25.4	3.6217
	26.0	3.5412

* Succinonitrile.

Fig. 4. Apparent Interfacial Tension
for 19.99 mol % SN in Water.
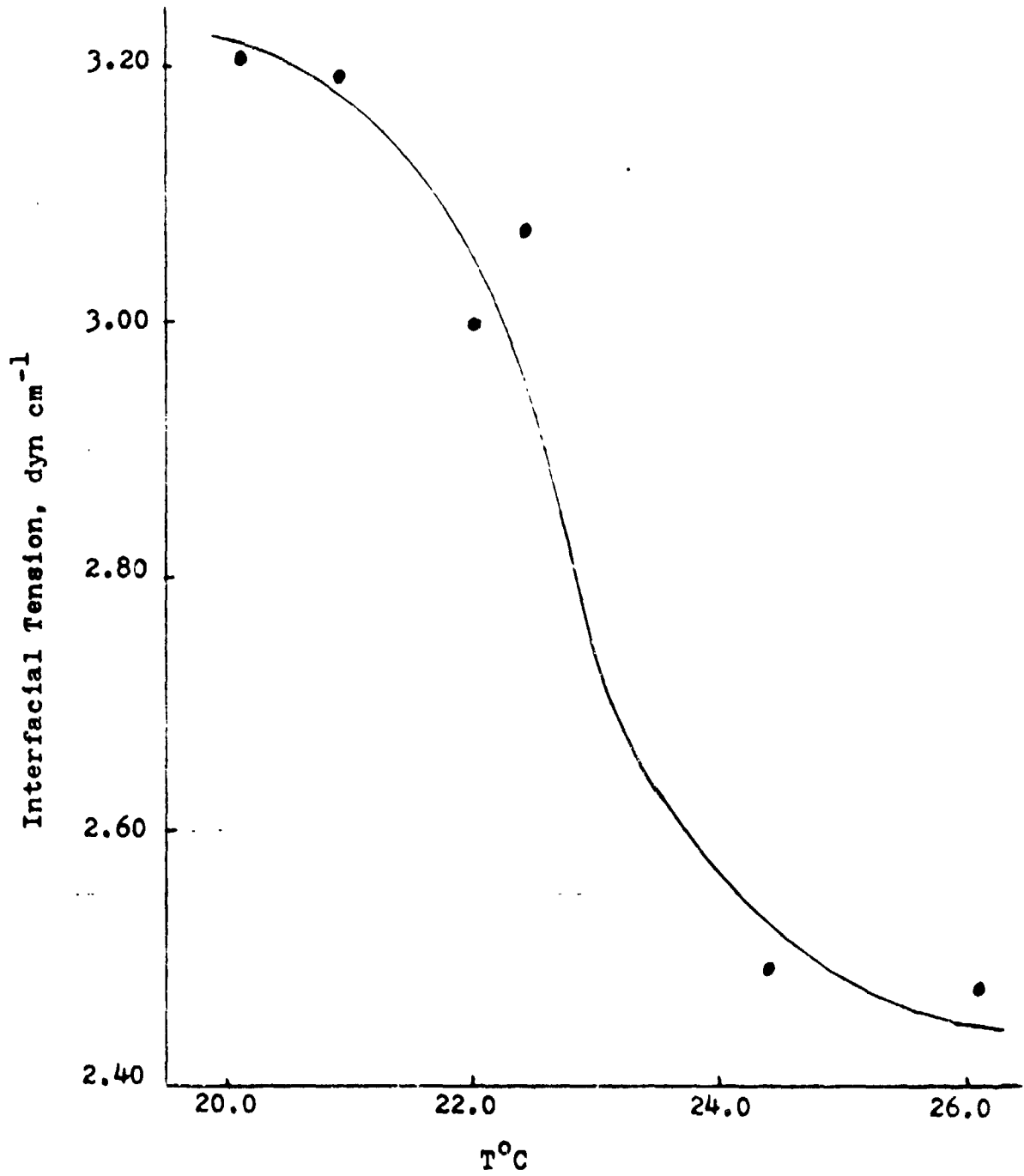
Fig. 5. Apparent Interfacial Tension
for 19.99 mol % SN in D_2O
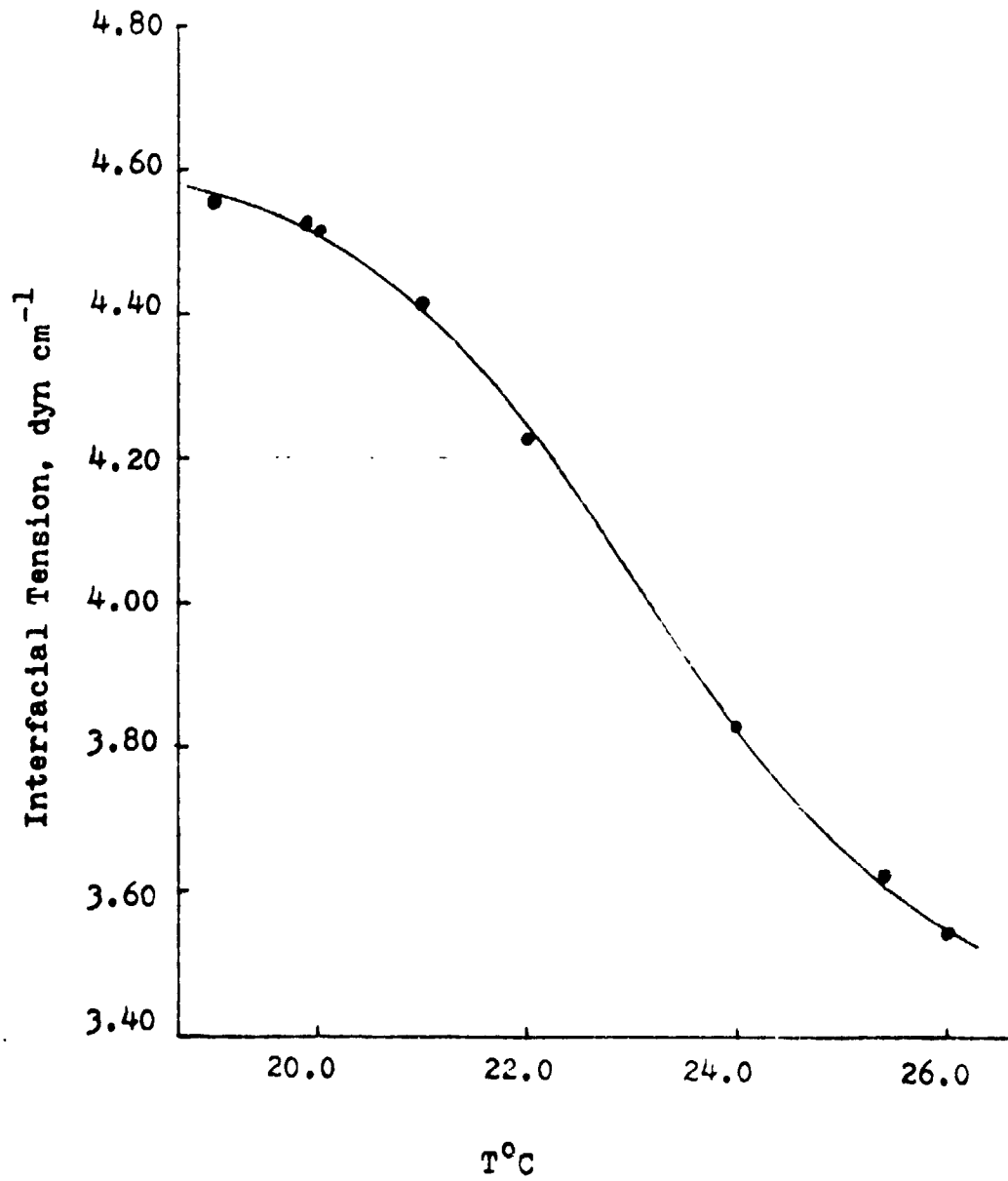
of the binary systems within certain temperature range.

It can be seen in Fig. 4 that the interfacial tension varies most drastically between 21° and $24^{\circ}C$ for the given concentration. As might be expected, the interfacial tension in the SN-H_2O system increases very rapidly as the temperature approaches the monotectic point of the mixture. Beyond $24^{\circ}C$ the tension decreases far more slowly with temperature. The same general behavior of the interfacial tension is also seen in Fig. 5 for the SN-D_2O system for the same temperature range. However, the interfacial tensions are much greater (up to 40%) for this system than the SN-H_2O even though the system has greater densities (up to about 20%).

IV. Conclusions and Suggestions

The current experimental data of densities and interfacial tensions for the immiscible binary systems of succinonitrile with water, with a mixture of water and D_2O, and with D_2O are not sufficient to draw any general conclusion.

Nevertheless, the data do indicate that the neutral density of the two phases in a system of a given overall composition with a mixture of water and D_2O may be achieved by varying the temperature within certain composition ranges. The neutral density of both phases in a system will minimize the gravity field effect in motion and coalescence of liquid droplets in studying the Marangoni effect.

Density behavior with temperature of the two phases in a system is reversed at a certain concentration of D_2O in the mixture of water and D_2O. Determination of this concentration requires more experimental data to be obtained as to the overall composition of the binary system and also as to the concentration of D_2O in the mixture of water and D_2O.

During the measurement of densities and interfacial tensions a number of difficulties have been encountered. Hence, a few suggestions on experimental work are made.

1. Density measurement should be made after each phase is physically separated from the equilibrium mixture in order to minimize the effect of the interface.

2. The size of the bob should be chosen to minimize the effect of the surface tension of the liquid. This may be done by measuring the density of distilled water using the bob with the Cahn balance.

3. For the measurement of interfacial tension using the ring method, relatively large cells are desirable in order to minimize the glass wall effect. When the cell size is too small the phases do not provide enough space for the ring to be placed if the interface is convexed or concaved.

4. It is also suggested that a system of certain overall composition with a mixture of water and D_2O be prepared to determine a temperature at which the density behavior with respect to temperature is reversed. This particular system may be used for continuous determination of density and interfacial tension as a function of temperature. These data may be used to calculate the interfacial tension gradient and to study the Marangoni effect.

REFERENCES

1. Jackson, K. A. and Hunt, J. D., Acta Met., $\underline{13}$, 1212 (1965).

2. Schreinemakers, F. A. H., Zeitschr. f. Physik Chemie, Leipzig, $\underline{23}$, 417 (1897).

3. Grugel, R. and Hellawell, A., Met. Trans., 12A, 669 (1981).

4. Young, N., Goldstein, J. and Block, M., J. Fluid Mech., $\underline{6}$, 350 (1959).

   Hardy, S. C., J. Coll. & Interface Sci., $\underline{69}$, 157 (1979).

   Krishnamurthy, S. and Goldburg, W. I., Phys. Rev. A, $\underline{22}$, 2147 (1980).

   Skapski, A. S., J. Chem. Pnys., $\underline{16}$, 386 (1948).

N82-17048

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

TOLERANCE ANALYSIS OF OPTICAL
TELESCOPES USING COHERENT ADDITION OF
WAVEFRONT ERRORS

Prepared By:	John W. Davenport, Ph.D.
Academic Rank:	Associate Professor
University and Department:	James Madison University Department of Mathematics and Computer Science
NASA/MSFC: Division: Branch:	Optical and R. F. Systems Optics
MSFC Counterpart:	Donald B. Griner
Date:	August 17, 1981
Contract No.:	NGT 01-003-021 The University of Alabama in Huntsville

ACKNOWLEDGEMENTS

The author would like to express his gratitude to Dr. Gerald Karr for directing a successful summer program. I also wish to express thanks to my MSFC counterpart, Don Griner, for his consultation and guidance during the summer, and to Gary Hunt for his assistance on this project.

LIST OF FIGURES

TOLERANCE ANALYSIS OF OPTICAL TELESCOPES
USING COHERENT ADDITION OF WAVEFRONT ERRORS

BY

John W. Davenport
Associate Professor of Mathematics
James Madison University
Harrisonburg, Virginia

ABSTRACT

A near diffraction-limited telescope requires that tolerance
analysis be done on the basis of system wavefront error.   One
method of analyzing the wavefront error is to represent the wave-
front error function in terms of its Zernike polynomial expansion.
The Ramsey-Korsch ray trace package has been expanded to include
the Zernike polynomial expansion up through the fifth-order spherical
term.   An option to determine a 3-dimensional plot of the wavefront
error function has also been included in the Ramsey-Korsch package.
Several assimulation runs have been analyzed to determine the particular
set of coefficients in the Zernike expansion that are effected by
various errors such as tilt, decenter and despace.   We have also
included a 3-dimensional plot of each error up through the fifth-order
spherical term.

Introduction

The Ramsey-Korsch computer program is a FORTRAN IV program that simulates the tracing of rays through an optical telescope system. Various options are available. The user can introduce tilt, decenter and despace upon each surface as well as compute a best focus correction and a centroid correction for a given surface in the optical system. Also, the program calculates the coefficients of the first few terms in the Zernike polynomial expansion of the wavefront error function. Part of our work this summer was involved in extending the Zernike polynomial expansion from ten terms to twenty-two terms (see Table 1).

Also, an additional option was included in the Ramsey-Korsch package that would allow one to obtain a plot of the wavefront error function. After this option was implemented, several trial runs were obtained (Figs. 1 - 5) as well as a separate plot of the wavefront error for each individual term of the Zernike polynomial expansion (Figs. 6 - 11).

Analysis of Data

Figure 1 is obtained from a tilt of .75 arc minute about the x-axis of the secondary mirror. Figure 2 is obtained from a decenter of 135 μm from the x-axis of the secondary mirror while Figure 3 is a combination of a tilt about the x-axis coupled with a decenter from the y-axis. Figures 4 and 5 result from a combination of errors due to tilt and decenter. Figure 4 results from a tilt of .85 arc minute about the x-axis combined with a decenter about the x-axis of 150 μm whereas Figure 5 differs from Figure 4 only in that the amount of decenter is in the y-axis direction instead of the x-axis direction.

In the case of tilt, the significant terms in the polynomial expansion are those that correspond to $R(1,1)$, $R(2,2)$, $R(1,3)$ and $R(1,5)$ with the sign of $R(1,1)$ and $R(1,3)$ being the same as the direction of tilt. In the case of decenter, the significant terms are those that correspond to $R(1,1)$, $R(2,2)$, $R(1,3)$ and $R(1,5)$ with the sign of $R(1,1)$ and $R(1,3)$ being opposite that of the motion of decenter. The significant terms in the presence of despace are $R(0,0)$, $R(0,2)$, $R(0,4)$ and $R(0,6)$ with the sign of $R(0,0)$ and $R(0,2)$ being opposite the motion of despace.

Derivation of the Zernike Polynomial Coefficients

A focusing system attempts to concentrate the impinging rays of a wave motion into a single point, referred to as the focus. However, it is impossible to achieve a point focus. There are limitations in practice such as the refraction properties of the system as well as the inevitable imperfections of the optical surfaces; in either case, the result is a "bright spot" (actually a 3-dimensional region) of finite dimension. The deviations of

the "bright spot" from the ideal spot is described in terms of
aberrations. In principle, an ideal system would produce
spherical waves converging into a point focus. In fact, however,
the phase fronts of the waves is not a (Gaussian) sphere; owing
to the nature of the system, the phase front will deviate from a
sphere as shown in Figure i. The aberration at a certain point
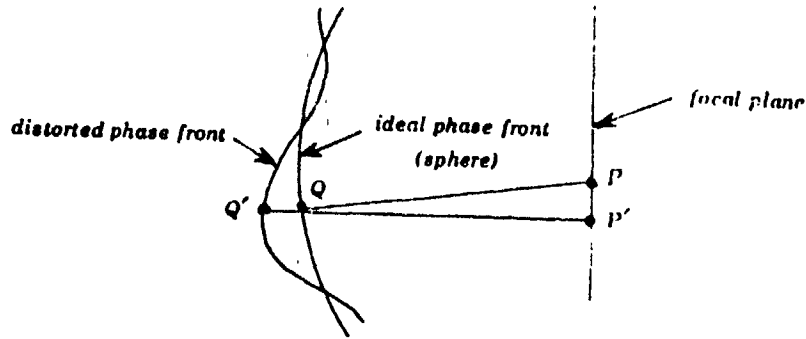
Aberration function.

Figure i

of the wavefront is the optical path difference between the
points Q and Q' and we note (2) that the aberration functions,
$\Phi(r,\theta)$, can be expanded in orthogonal functions f and g, i.e.,

$$(1) \qquad \Phi(r,\theta) = \sum_{n=0} \sum_{m=0} c_{mn} f_n^m (r) g_m(\theta).$$

It is, however, particularly convenient to make (1) an orthogonal
expansion in functions simply related to elementary aberrations
such as spherical aberrations, coma and astigmatism. These
aberrations are characterized as those that are generated by
linear combinations of the functions

$$r^n \cos^m \theta$$

('elementary' or Seidel aberrations have n < 4, m < 3). If, in
addition, we require that the polynomials be orthogonal over a
circular or annular region, then the aberration function can be
expanded in orthogonal polynomials of the type

$$R_n^m(r) \{\cos m \theta + \sin m \theta\}$$

where the R's are the Zernike polynomials. The polynomials R_n^m are obtained by orthogonalizing the sequence $r^m, r^{m+2}, \ldots, r^n$ over $\{\epsilon, 1\}$ using the Gram-Schmidt orthogonalization process with inner product

$$< f , g > = \frac{1}{1 - \epsilon^2} \int_\epsilon^1 f(r) \cdot g(r) \; r \; d r.$$

Let us consider the derivation of each Zernike polynomial.

1. R_0^0 - orthogonalize $\{1\}$.

     $R_0^0 \; = \; 1$

2. R_1^1 - orthogonalize $\{r\}$.

     $R_1^1 \; = \; r$

3. R_2^0 - orthogonalize $\{1, r^2\}$.

     $R_2^0 \; = \; r^2 - < 1, r^2 > 2 = r^2 - \dfrac{2}{1-\epsilon^2} \int_\epsilon^1 r^3 \; dr = r^2 - (1+\epsilon^2)/2$

     NOTE: $\quad ||1||^2 = 2.$

4. R_2^2 - orthogonalize $\{r^2\}$.

     $R_2^2 \; = \; r^2$

5. R_3^1 - orthogonalize $\{r, r^3\}$.

     $R_3^1 \; = \; r^3 - < r^3, r > \dfrac{4}{1-\epsilon^2} r = r^3 - \dfrac{4r}{1-\epsilon^4} \int_\epsilon^1 r^5 \; dr$

         $= \; r^3 - 2r(1-\epsilon^6)/3(1-\epsilon^4)$

6. R_3^3 - orthogonalize $\{r^3\}$.

     $R_3^3 \; = \; r^3$

7. R_4^4 - orthogonalize $\{r^4\}$.

     $R_4^4 \; = \; r^4$

8. R_4^2 — orthogonalize $\{r^2, r^4\}$.

Using $||r^2||^2 = \langle r^2, r^2 \rangle$ we have

$$R_4^2 = r^4 - \langle r^4, r^2 \rangle 6(1-\epsilon^2)r^2/(1-\epsilon^6)$$

$$= r^4 - 3(1-\epsilon^8)r^2/4(1-\epsilon^6)$$

9. R_4^0 — orthogonalize $\{1, r^2, r^4\}$.

$$R_4^0 = r^4 - \langle r^4, 1 \rangle 2 - \frac{\langle r^4, 2r^2 - (1+\epsilon^2) \rangle \cdot 6}{(1-\epsilon^2)^2} \cdot$$

$$\{2r^2 - (1+\epsilon^2)\}$$

$$= r^4 - (1+\epsilon^2)r^2 + (1+4\epsilon^2+\epsilon^4)/6$$

10. R_5^5 — orthogonalize $\{r^5\}$.

$$R_5^5 = r^5$$

11. R_5^3 — orthogonalize $\{r^3, r^5\}$.

Since $||r^3||^2 = \frac{1}{1-\epsilon^2} \int_{\mathcal{E}}^{1} r^7 dr = (1-\epsilon^8)/8(1-\epsilon^2)$

Then $R_5^3 = r^5 - \langle r^5, r^3 \rangle 8(1-\epsilon^2) \cdot r^3/(1-\epsilon^8)$

$$= r^5 - 4(1-\epsilon^{10}) r^3/5(1-\epsilon^8)$$

12. R_5^1 — orthogonalize $\{r, r^3, r^5\}$.

First, normalize r, thus r becomes $2r/\sqrt{1+\epsilon^2}$. Second, orthogonalize the set $\{r, r^3\}$ to obtain $\dfrac{\sqrt{8} \, (3(1+\epsilon^2) \cdot r^3 - 2r(1+\epsilon^2+\epsilon^4)}{\sqrt{(1-\epsilon^4)(1+3\epsilon^2-3\epsilon^4-\epsilon^6)}}$

Thus, $R_5^1 = r^5 - < r^5, r > 4r/(1+\varepsilon^2)$

$- < r^5, 3(1+\varepsilon^2)r^3 - 2r(1+\varepsilon^2+\varepsilon^4) >,$

$8\{3'1+\varepsilon^2)r^3 - 2r(1+\varepsilon^2+\varepsilon^4)\}/(1-\varepsilon^4)(1+3\varepsilon^2-3\varepsilon^{4-}\varepsilon^6)$

$= r^5 - 6(1+\varepsilon^2)(\varepsilon^8+\varepsilon^6-4\varepsilon^4+\varepsilon^2+1)r^3/$

$5(\varepsilon^8+2\varepsilon^6-6\varepsilon^4+2\varepsilon^2+1)+3(1+4\varepsilon^2+10\varepsilon^4+4\varepsilon^6+\varepsilon^8).r/10(1+4\varepsilon^2+\varepsilon^4)$

13. R_6^0 - orthogonalize $\{1,r^2,r^4,r^6\}$.

$R_6^0 = r^6 - < r^5,6r^4-6(1+\varepsilon^2)r^2+(1+4\varepsilon^2+\varepsilon^4) > .$

$10(6r^4-6(1+\varepsilon^2)r^2+(1+4\varepsilon^2+\varepsilon^4)/(1-\varepsilon^2)^4$

$- <r^6,2r^2-(1+\varepsilon^2)> . 6(2r^2-(1+\varepsilon^2)/(1-\varepsilon^2)^2 - <r^6,1> 2$

$= r^6-3(1+\varepsilon^2)r^4/2+3(1+3\varepsilon^2+\varepsilon^4)r^2/5-(1+9\varepsilon^2+9\varepsilon^4+\varepsilon^6)/20$

Table 1 gives a listing of the above vectors after they have been normalized to 1, i.e., $||v||_{max.} = 1$.

Procedure for the Computer Program

A sample run of the computer program named JDSPOT is shown below. In this run, we requested that a best focus correction be made (IFOC=1) using a grid size of 31 (NRAS=31) with a printout of the summary data and Zernike coefficients (PRSM=1). A plot file is created on FILE1 (HCOP=1). Note that if a plot file is required, the command ! SET F ! TT/(FILE NAME) must be entered before running the program. See reference (6) for a further explanation of the options available in the SPOT program.

'SET F:TT/FILE1

!START JDSPOT
ENTER NEW DATA
?IFOC
?1
?MRMS
?31
?GX(2)
?.75
?MMSR
?1
?MCOP
?1
?END

*PAUSE 30
?
ENTER HEADING
?(HEADING FOR PLOT FILE)
PLOT COMPLETED
ENTER NEW DATA
?DONE
*STOP: THATS ALL FOLKS
--

ST 8.4M CONFIGURATION, FIELD PT.   .000,   .000
     BEST FOCUS
     TILT: X .750; Y
     820 RAYS

RMS OPD   .2962
STREHL   -.6708
CENTROID:
     X   .2774
     Y   -.5035E-21
     Z   -.5046E-04

CHIEF RAY:
     X   .2703
     Y   .0000

ZERNIKE POLYNOMIAL COEFFICIENTS

R(0,0).......... -.45692E-04   R(1,1)COSA.....   3.6508
R(1,1)SINA..... -.67768E-07   R(0,2).........   .22078E-04
R(2,2)COS2A....   .34539E-02   R(2,2)SIN2A....   .34704E-07
R(1,3)COSA.....   1.0385   R(1,3)SINA..... -.43245E-07
R(3,3)COS3A....   .18647E-06   R(3,3)SIN3A.... -.12570E-06
R(0,4).........   .24914E-04   R(2,4)COS2A....   .24017E-05
R(2,4)SIN2A....   .45571E-07   R(4,4)COS4A.... -.24480E-07
R(4,4)SIN4A.... -.53790E-08   R(1,5)COSA..... -.59984E-02
R(1,5)SINA..... -.37962E-07   R(3,5)COS3A....   .23204E-06
R(3,5)SIN3A.... -.17036E-06   R(5,5)COS5A....   .14900E-06
R(5,5)SIN5A.... -.17078E-07   R(0,6).........   .61713E-04

ZERNIKE POLYNOMIALS

Table 1. $R_n^m(r)$ for $m,n \leq 6$

R_0^0	1
R_1^1	r
R_2^0	$(2r^2 - (1+\epsilon^2))/(1-\epsilon^2)$
R_2^2	r^2
R_3^1	$r \cdot \{3(1+\epsilon^2)r^2 - 2(1+\epsilon^2+\epsilon^4)\}/(1+\epsilon^2-2\epsilon^4)$
R_3^3	r^3
R_4^0	$\{6r^4-6(1+\epsilon^2)r^2+(1+4\epsilon^2+\epsilon^4)\}/(1-\epsilon^2)^2$
R_4^2	$r^2\{4(1-\epsilon^6)r^2 - 3(1-\epsilon^8)\}/(1-4\epsilon^6+3\epsilon^8)$
R_4^4	r^4
R_5^1	$10r\{r^4(1+4\epsilon^2+\epsilon^4) - 1.2(1+\epsilon^2)^2(1+3\epsilon^2+\epsilon^4)r^2+.3(1+4\epsilon^2+10\epsilon^4+4\epsilon^6+\epsilon^8)\}/(1-\epsilon^2)^2(1+6\epsilon^2+3\epsilon^4)$
R_5^3	$\{5(1-\epsilon^8)r^2 - 4(1-\epsilon^{10})\}r^3/(1-3\epsilon^8+4\epsilon^{10})$
R_5^5	r^5
R_6^0	$\{20r^6-30(1+\epsilon^2)r^4 + 12(1+3\epsilon^2+\epsilon^4)r^2 - (1+9\epsilon^2+9\epsilon^4+\epsilon^6)\}/(1-\epsilon^2)^3$

Figure 1

TILT   X=.75
SPOT DIAGRAM AT SURFACE #   4
RMS OPD = .20621 -.67875

ENCIRCLED ENERGY DISTRIBUTION
100%......   .97701132E-01
90%......   .69398980E-01
80%......   .55111699E-01
70%......   .48138693E-01

CENTROID AT   X =   .27740895E 00
              Y = -.52249409E-21
              Z = -.59455558E-04

BOX IS   .00800000 CM SQUARE CENTERED ON CENTROID
?PAUSE?◻
?
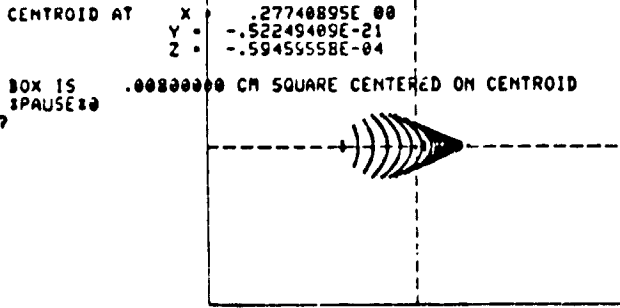
ZERNIKE POLYNOMIAL COEFFICIENTS

R(0,0)........	-.456928E-04	R(1,1)COSA....	3.05000
R(1,1)SINA....	-.677089E-07	R(0,2)........	.22070BE-04
R(2,2)COS2A...	.345394E-02	R(2,2)SIN2A...	.341704E-07
R(1,3)COSA....	1.03255	R(1,3)SINA....	-.432461E-07
R(3,3)COS3A...	.18647×E-06	R(3,3)SIN3A...	-.125707E-06
R(0,4)........	.249143E-04	R(2,4)COS2A...	.240172E-05
R(2,4)SIN2A...	.455718E-07	R(4,4)COS4A...	-.244303E-07
R(4,4)SIN4A...	-.537997E-08	R(1,5)COSA....	-.599864E-02
R(1,5)SINA....	-.373520E-07	R(3,5)COS3A...	.232048E-06
R(3,5)SIN3A...	-.170366E-96	R(5,5)COS5A...	.149007E-06
R(5,5)SIN5A...	-.179789E-97	R(0,6)........	.617139E-04

?PAUSE?◻
?
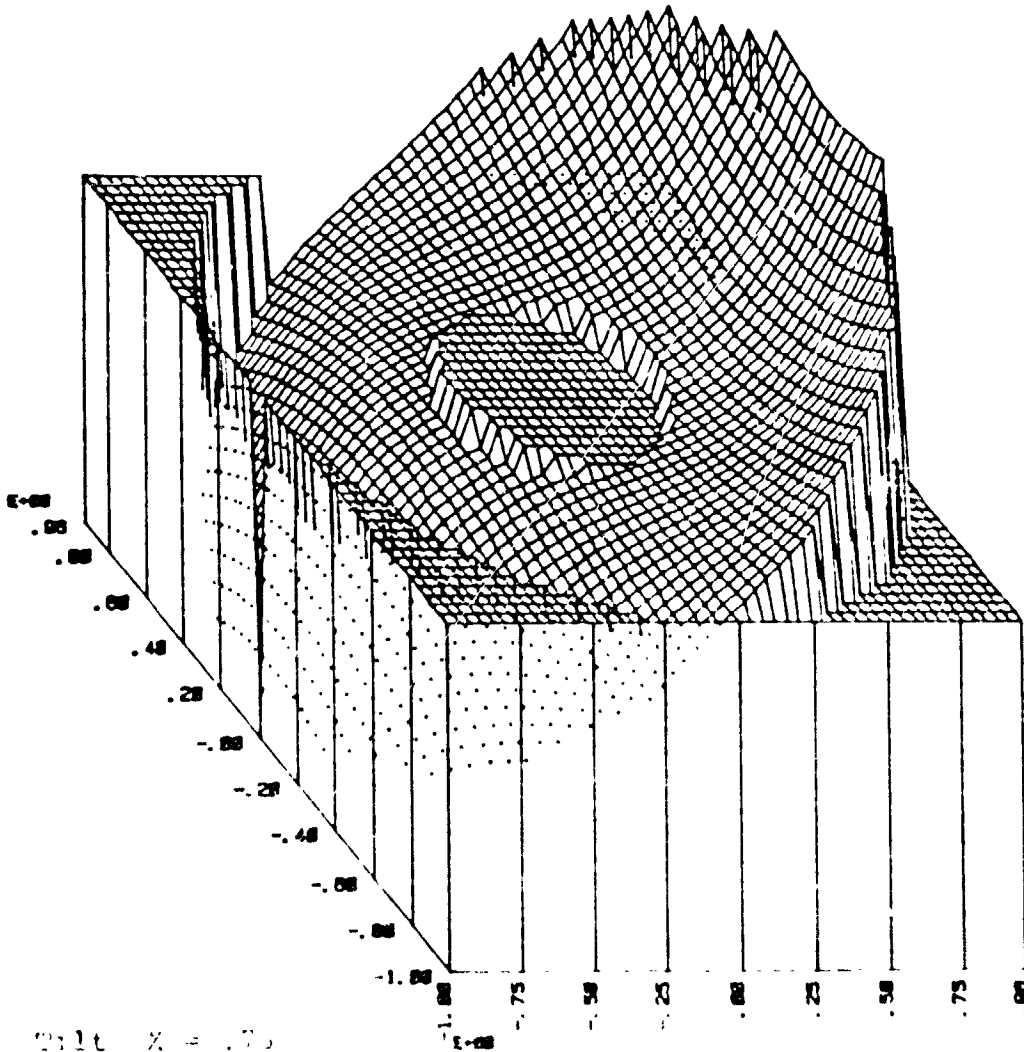
Tilt X = .75

Figure 2   ORIGINAL PAGE IS
OF POOR QUALITY

TILT   8-05-81

SPOT DIAGRAM AT SURFACE 8   4
RMS OPD = .21051 -.74943

ENCIRCLED ENERGY DISTRIBUTION
100%......   .99657536E-01
90%......   .71453691E-01
80%......   .56191619E-01
70%......   .49175344E-01

CENTROID AT   X =   -.12567903E 00
               Y =   .37274836E-20
               Z =   -.17290785E-03

BOX IS   .03800000 CM SQUARE CENTERED ON CENTROID
:PAUSE:0
?
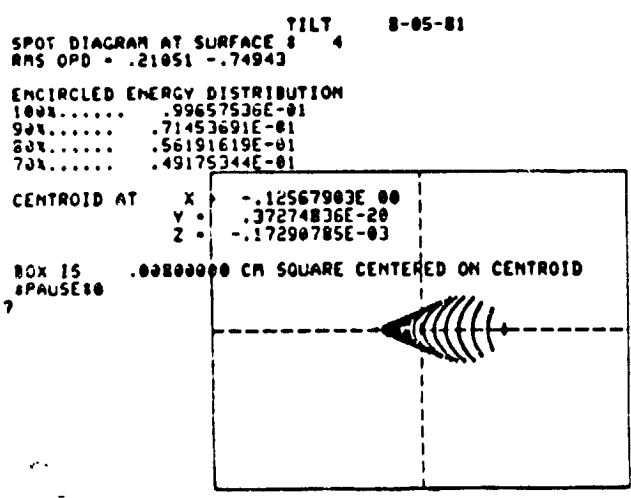
ZERNIKE POLYNOMIAL COEFFICIENTS

R(0,0)........	-.753513E-04	R(1,1)COSA....	-3.11364
R(1,1)SINA....	.186271E-06	R(0,2)........	.199349E-04
R(2,2)COS2A...	-.347387E-03	R(2,2)SIN2A...	-.388937E-07
R(1,3)COSA....	-1.05410	R(1,3)SINA....	.770492E-07
R(3,3)COS3A...	-.137990E-06	R(3,3)SIN3A...	.765711E-07
R(0,4)........	.309234E-04	R(2,4)COS2A...	.252765E-05
R(2,4)SIN2A...	-.206548E-07	R(4,4)COS4A...	.676988E-08
R(4,4)SIN4A...	.141741E-08	R(1,5)COSA....	.609210E-02
R(1,5)SINA....	.274056E-06	R(3,5)COS3A...	-.280703E-06
R(3,5)SIN3A...	.224658E-06	R(5,5)COS5A...	-.115481E-06
R(5,5)SINSA...	.189959E-07	R(0,6)........	.604598E-04

:PAUSE:0
?

DECENTER   PLOT 1   X=135
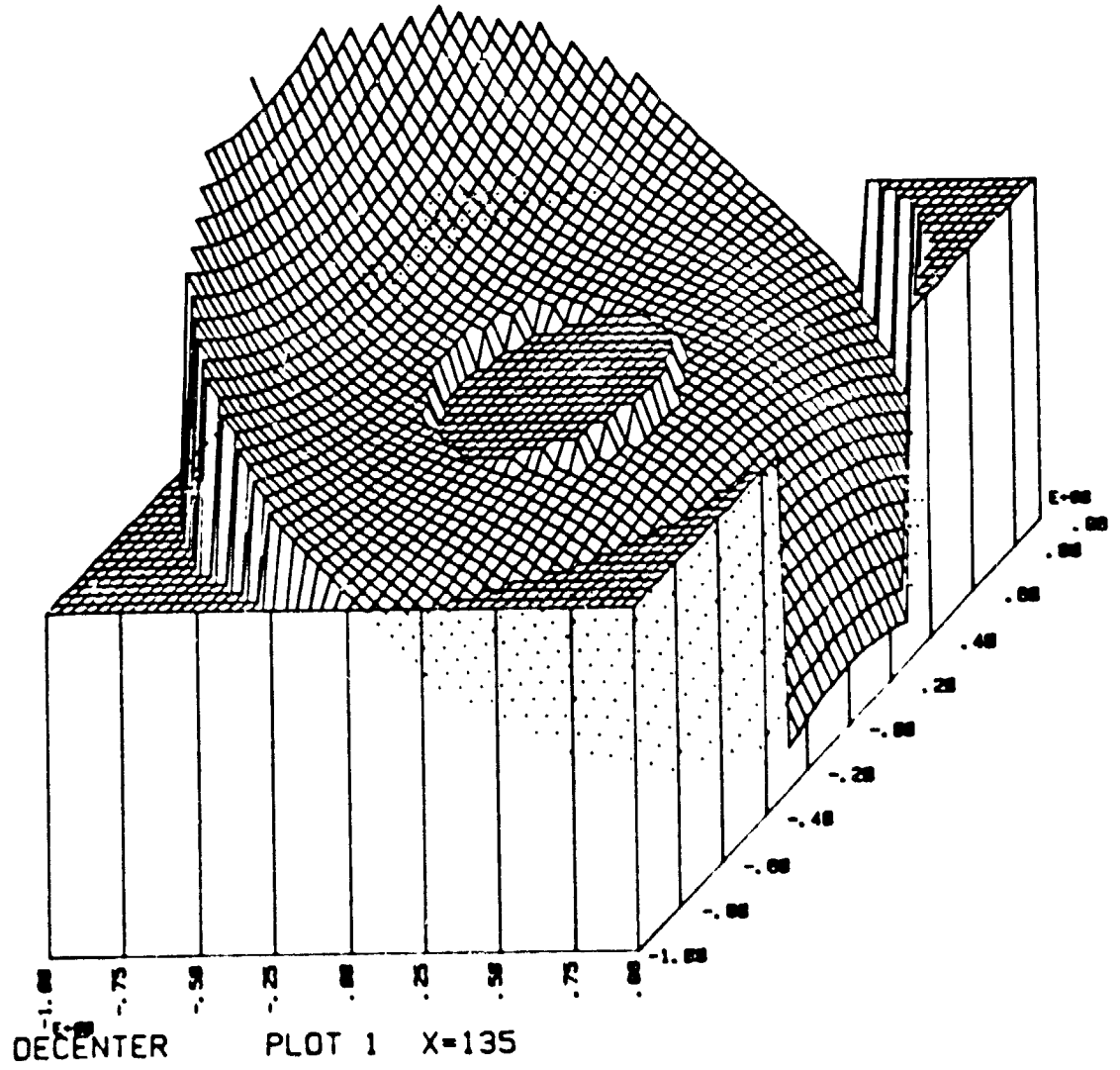
Figure 3

TILT X=.5 + DECENTER Y=80
SPOT DIAGRAM AT SURFACE S 4
RMS OPD = .18564 -.36345

ENCIRCLED ENERGY DISTRIBUTION
100%...... .92881680E-01
95%...... .65766096E-01
50%...... .49041633E-01
70%...... .42659262E-01

CENTROID AT   X =   .18493929E 00
              Y = -.74476475E-01
              Z = -.85464402E-04

BOX IS   .00800000 CM SQUARE CENTERED ON CENTROID
2PAUSE20
?
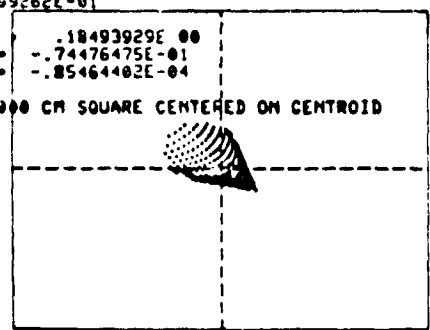
ZERNIKE POLYNOMIAL COEFFICIENTS

R(0,0)........	-.936563E-04	R(1,1)COSA....	2.03339
R(1,1)SINA....	1.84512	R(0,2)........	.548048E-05
R(2,2)COS2A...	.165706E-02	R(2,2)SIN2A...	.126204E-02
R(1,3)COSA....	.688370	R(1,3)SINA....	.624651
R(3,3)COS3A...	-.857778E-07	R(3,3)SIN3A...	-.447536E-07
R(0,4)........	.277484E-04	R(2,4)COS2A...	.105674E-06
R(2,4)SIN2A...	.208051E-05	R(4,4)COS4A...	-.266313E-07
R(4,4)SIN4A...	-.519276E-07	R(1,5)COSA....	-.399973E-02
R(1,5)SINA....	-.361067E-02	R(3,5)COS3A...	.108539E-06
R(3,5)SIN3A...	-.748542E-07	R(5,5)COS5A...	.745636E-07
R(5,5)SIN5A...	.558777E-07	R(0,6)........	.617785E-04

2PAUSE20
?

TILT X=.5 + DECENTER Y=80
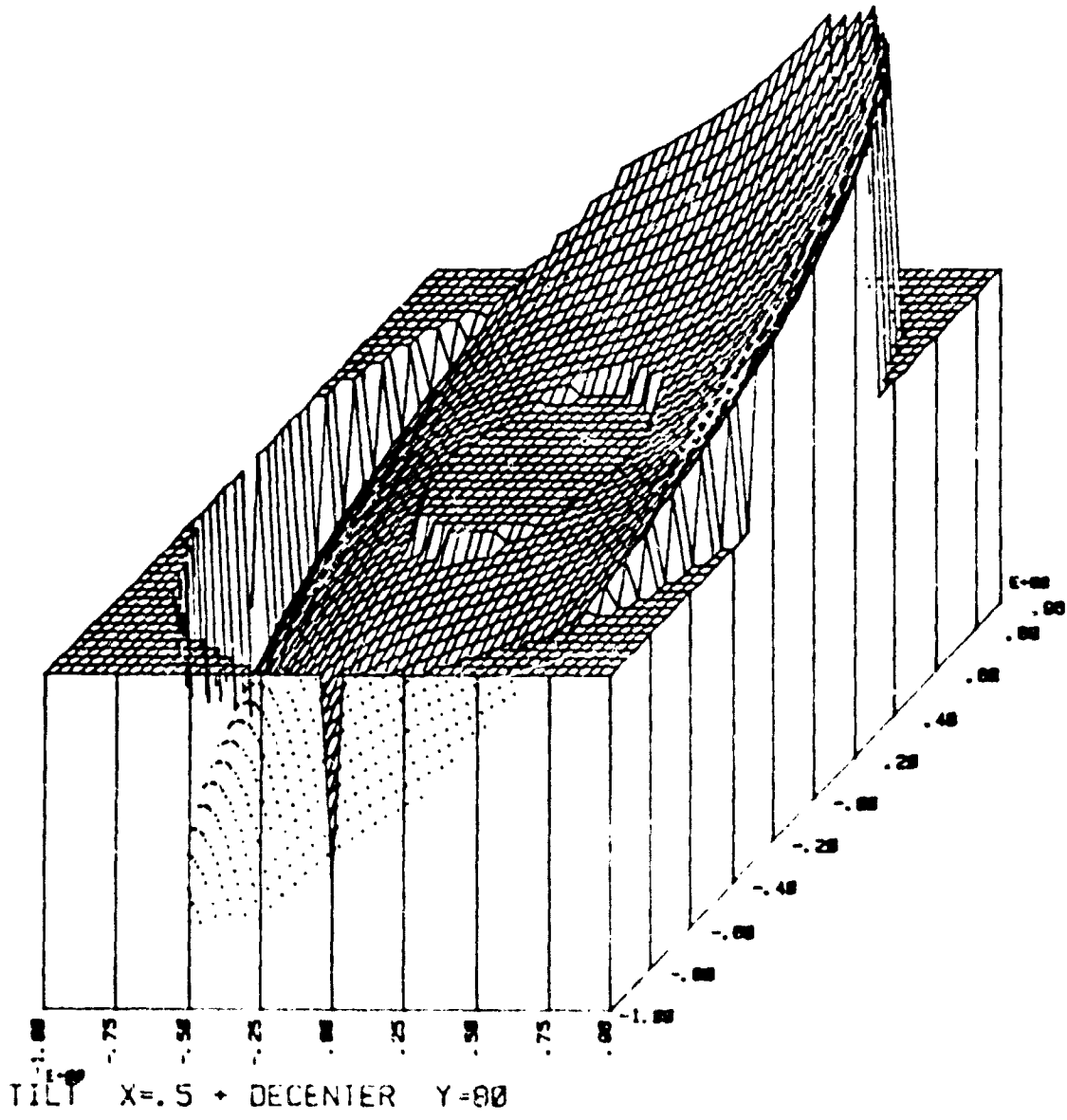
Figure 4

```
TILT   X=.85   DECENTER   X=150
SPOT DIAGRAM AT SURFACE 8   4
RMS OPD = .00019 1.00000

ENCIRCLED ENERGY DISTRIBUTION
100%  .....   .10330659E-03
92%   ......  .69376401E-04
80%   ......  .53297932E-04
70%   ......  .45925743E-04

CENTROID AT   X    .17475347E 00
              Y    .00000000E 00
              Z    .18356914E-03

BOX IS   .00000800 CM SQUARE CENTERED ON CENTROID
:PAUSE:0
?
```
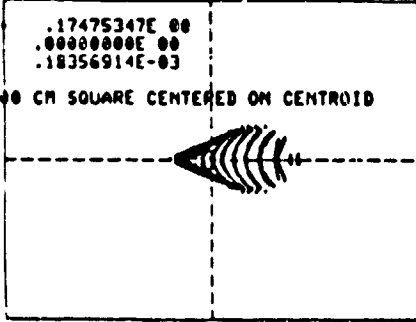
ZERNIKE POLYNOMIAL COEFFICIENTS

R(0,0)........	-.175324E-03	R(1,1)COSA....	-.284270E-02
R(1,1)SINA....	.280314E-09	R(0,2)........	-.457959E-04
R(2,2)COS2A...	.212250E-04	R(2,2)SIN2A...	.270989E-10
R(1,3)COSA....	-.997547E-03	R(1,3)SINA....	.109886E-09
R(3,3)COS3A...	-.504061E-08	R(3,3)SIN3A...	.742572E-11
R(0,4)........	.206002E-04	R(2,4)COS2A...	-.586874E-07
R(2,4)SIN2A...	.563772E-11	R(4,4)COS4A...	.110683E-08
R(4,4)SIN4A...	-.193075E-10	R(1,5)COSA....	-.298442E-04
R(1,5)SINA....	.258245E-09	R(3,5)COS3A...	.150860E-08
R(3,5)SIN3A...	.109911E-10	R(5,5)COS5A...	.185169E-08
R(5,5)SIN5A...	.316731E-11	R(0,6)........	.607103E-04

```
:PAUSE:0
?
```
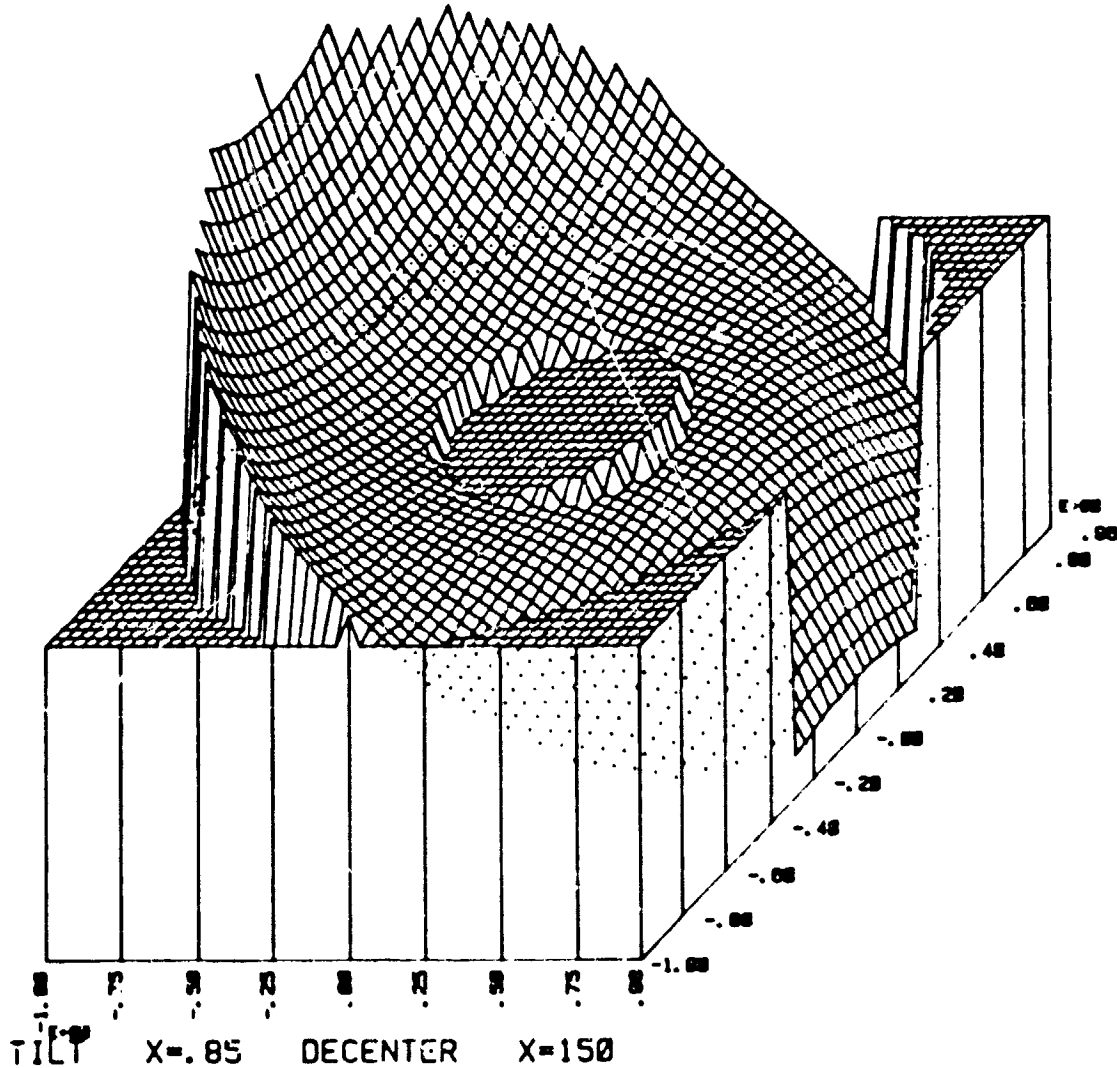

TILT   X=.85   DECENTER   X=150

Figure 5

TILT X=.85   DECENTER   Y=150
SPOT DIAGRAM AT SURFACE 4
RMS OPD = .33065 -3.31606

ENCIRCLED ENERGY DISTRIBUTION
100%......   .16579163E 00
90%......   .11394823E 00
60%......   .88358581E-01
70%......   .76761007E-01

CENTROID AT   X :   .31439671E 00
               Y =   -.13964335E 00
               Z =   -.30233527E-03

BOX IS   .02000000 CM SQUARE CENTERED ON CENTROID
*PAUSE*0
?
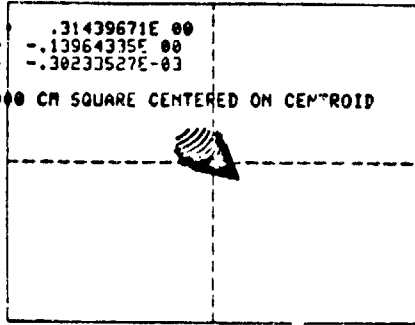
ZERNIKE POLYNOMIAL COEFFICIENTS

F(0,0)........	.188155E-03	R(1,1)COSA....	3.45676
R(1,1)SINA....	3.45959	R(0,2)........	.134556E-03
R(2,2)COS2A...	.486516E-02	R(2,2)SIN2A...	.402284E-02
R(1,3)COSA....	1.17023	R(1,3)SINA....	1.17122
R(3,3)COS3A...	-.134261E-06	R(3,3)SIN3A...	-.522126E-07
R(0,4)........	.273415E-04	R(2,4)COS2A...	-.119985E-06
R(2,4)SIN2A...	.654611E-05	R(4,4)COS4A...	-.519700E-07
R(4,4)SIN4A...	-.105210E-07	R(1,5)COSA....	-.679783E-02
R(1,5)SINA....	-.676928E-02	R(3,5)COS3A...	.823396E-07
R(3,5)SIN3A...	-.179650E-06	R(5,5)COS5A...	.152733E-06
R(5,5)SIN5A...	.894043E-07	R(0,6)........	.656847E-04

*PAUSE*0
?
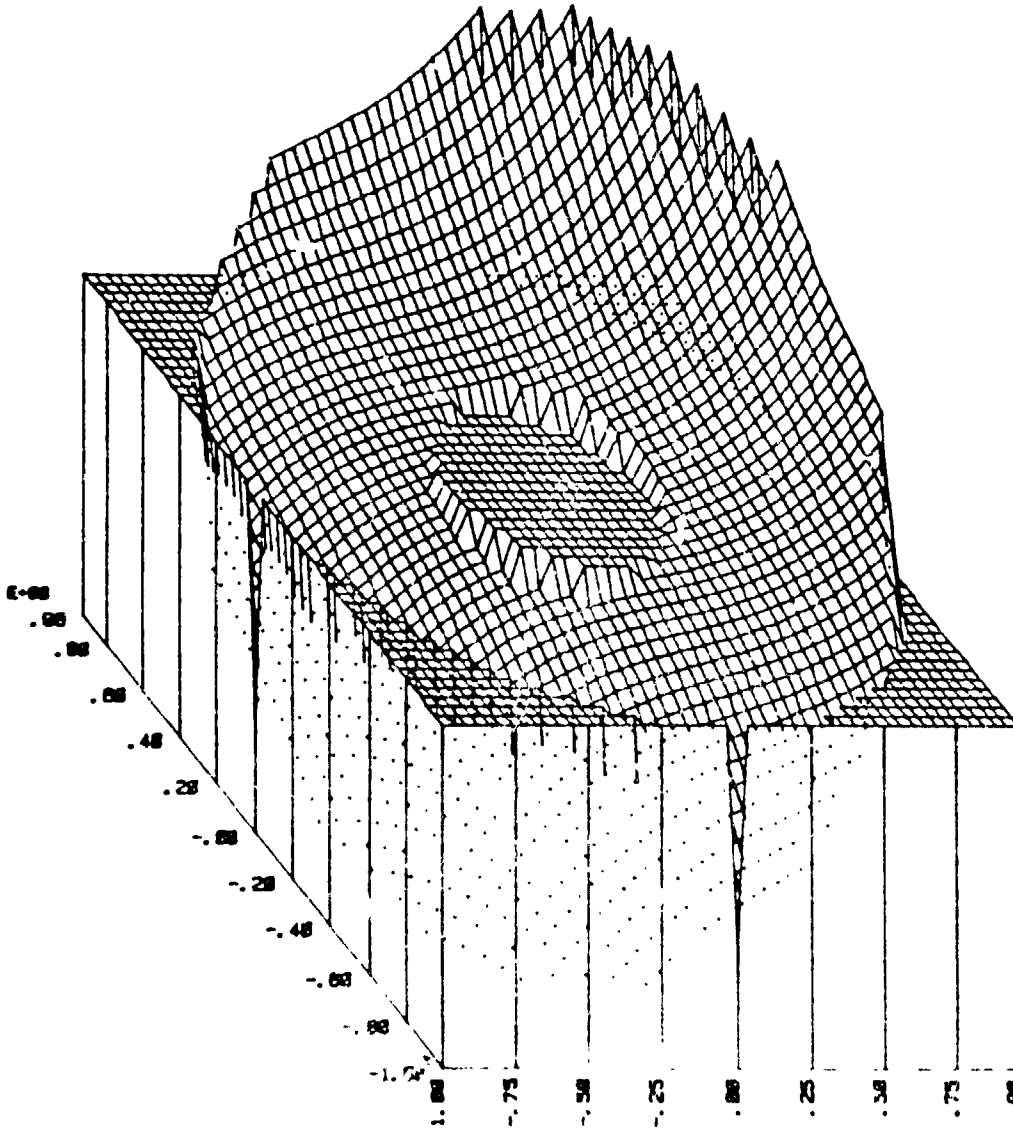
TILT   X=.85   DECENTER   Y=150

Figure 6

ORIGINAL PAGE IS
OF POOR QUALITY
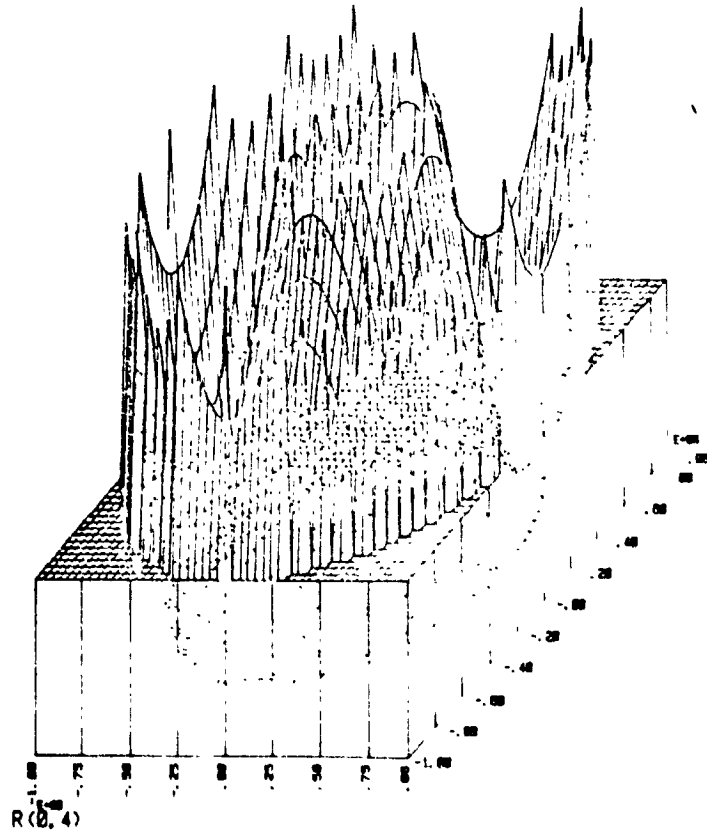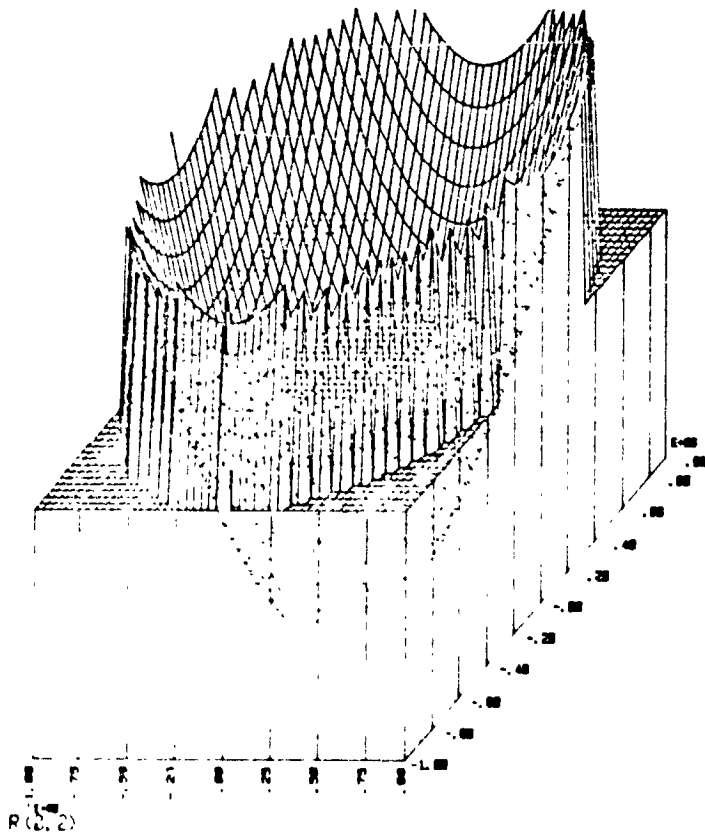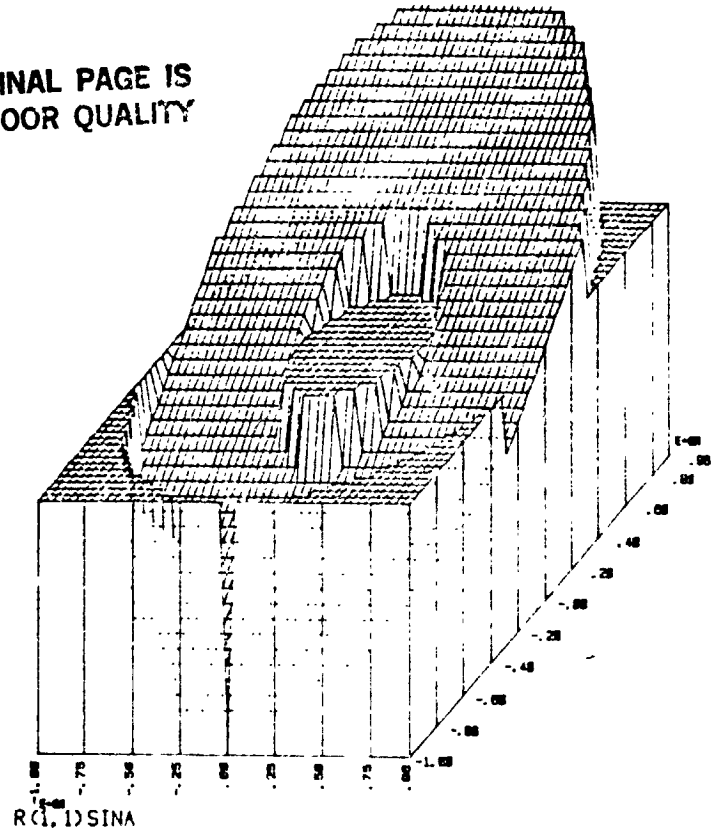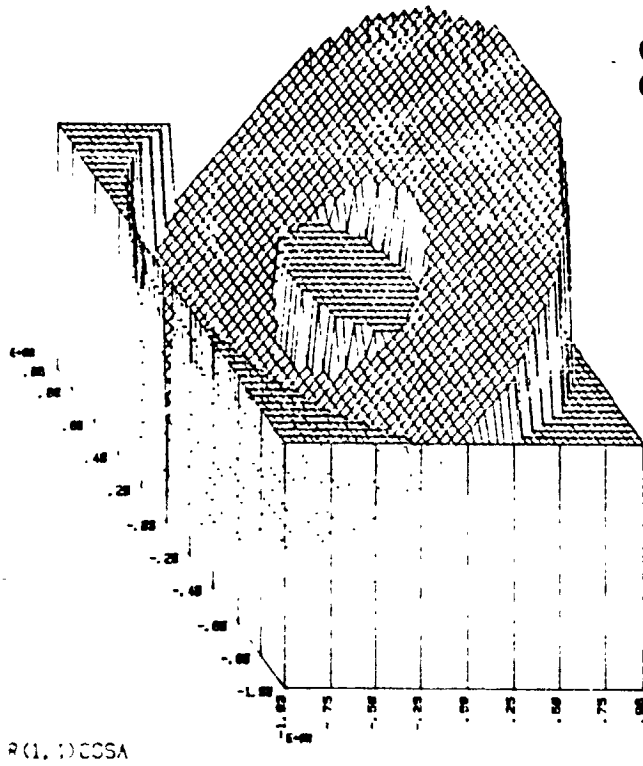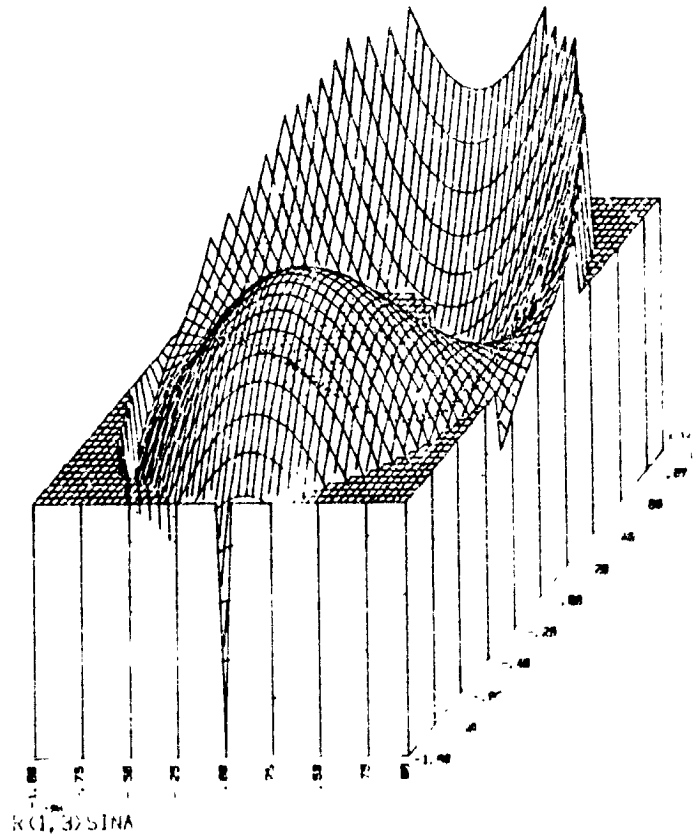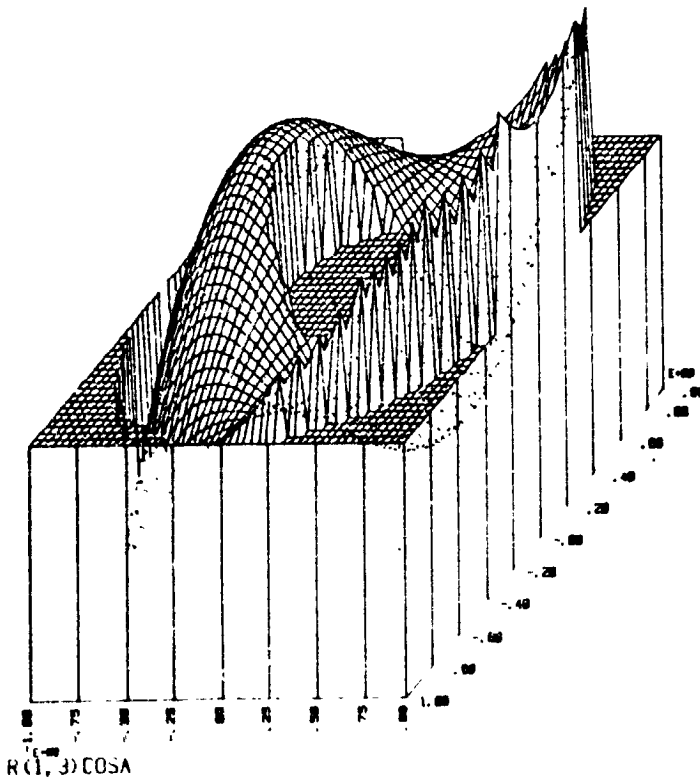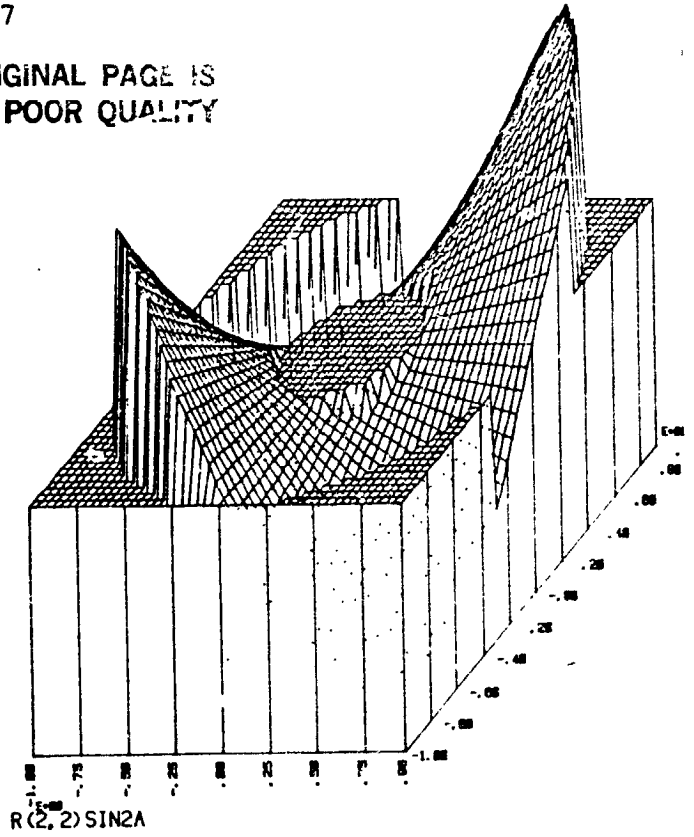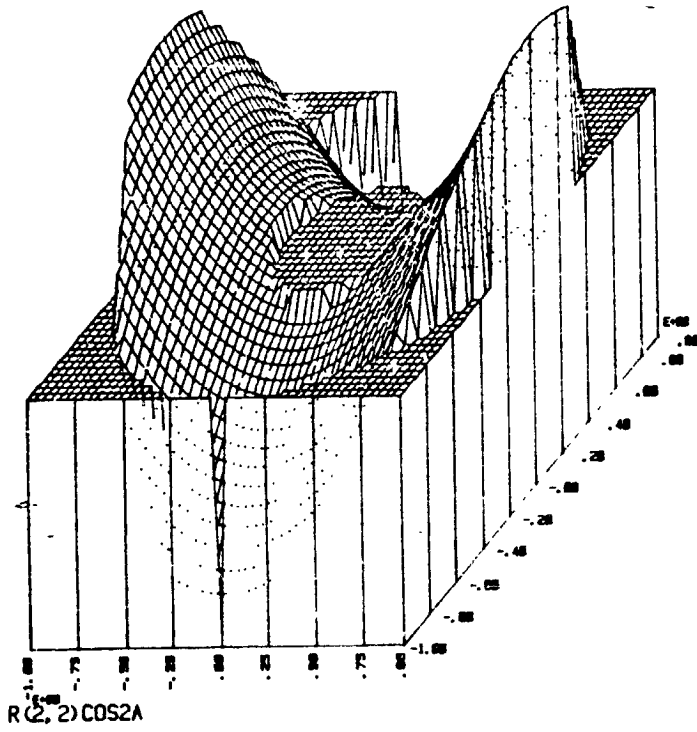
R(1,1)COSA

R(1,1)SINA

R(0,2)

R(0,4)

Figure 7

R (2, 2) COS2A

R (2, 2) SIN2A

R (1, 3) COSA

R (1, 3) SINA

V-14

Figure 8

R(3, 3)COS3A

R(3, 3)SIN3A

R(3, 4)COS2A

R(3, 4)SIN2A

Figure 9

R(4, 4)COS4A

R(4, 4)SIN4A

R(1, 5)COSA

R(1, 5)SINA
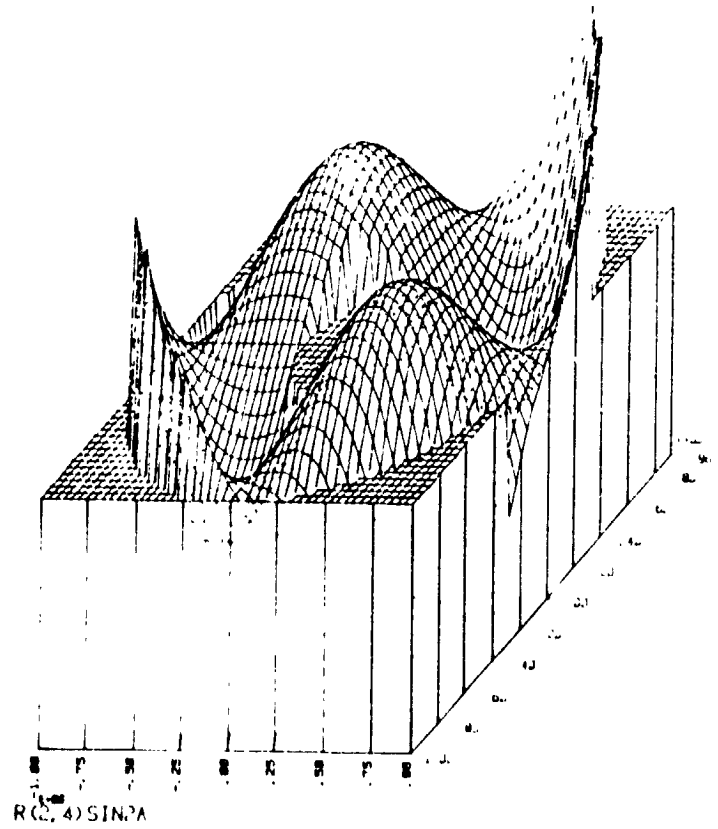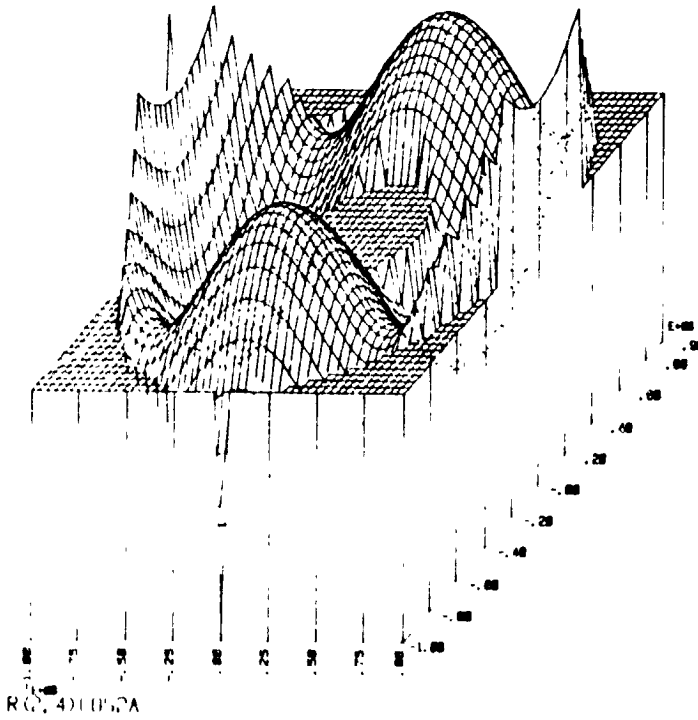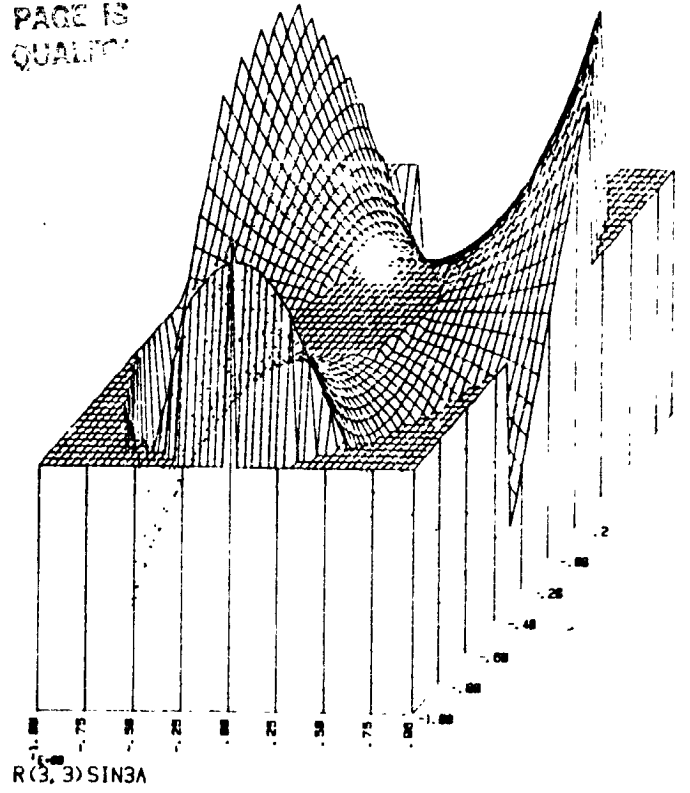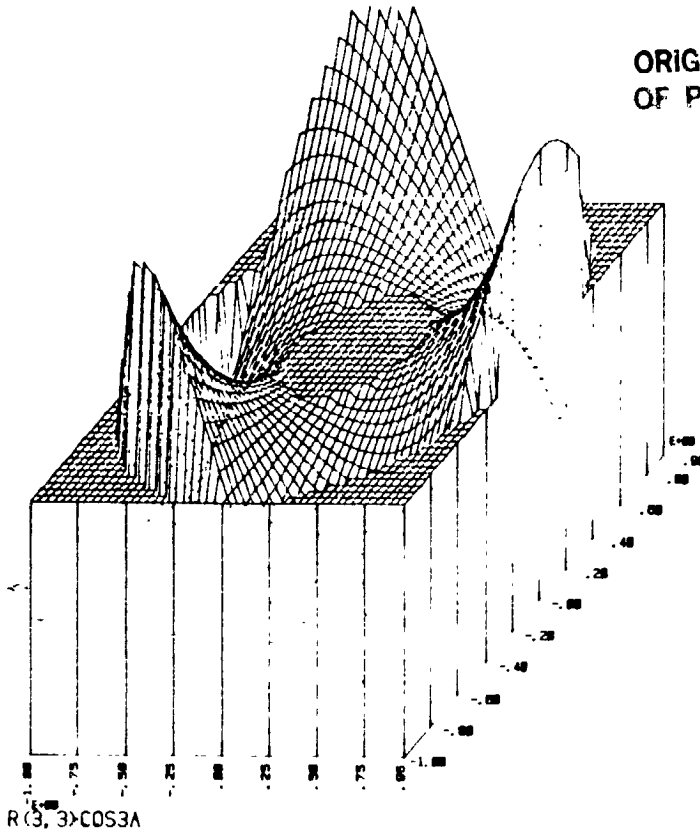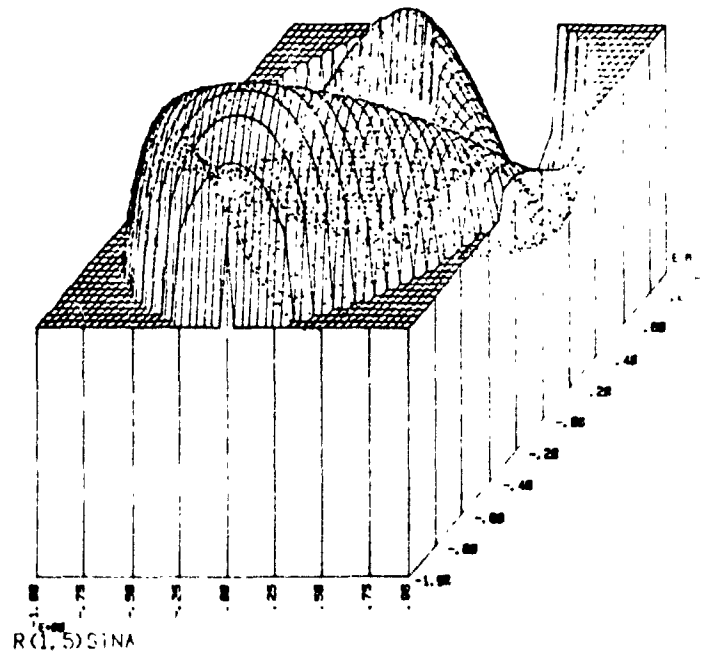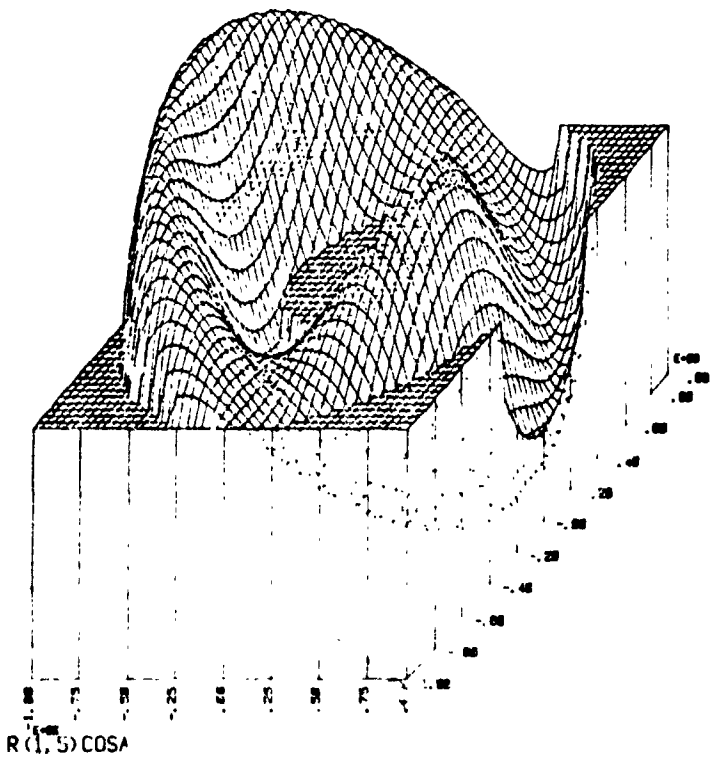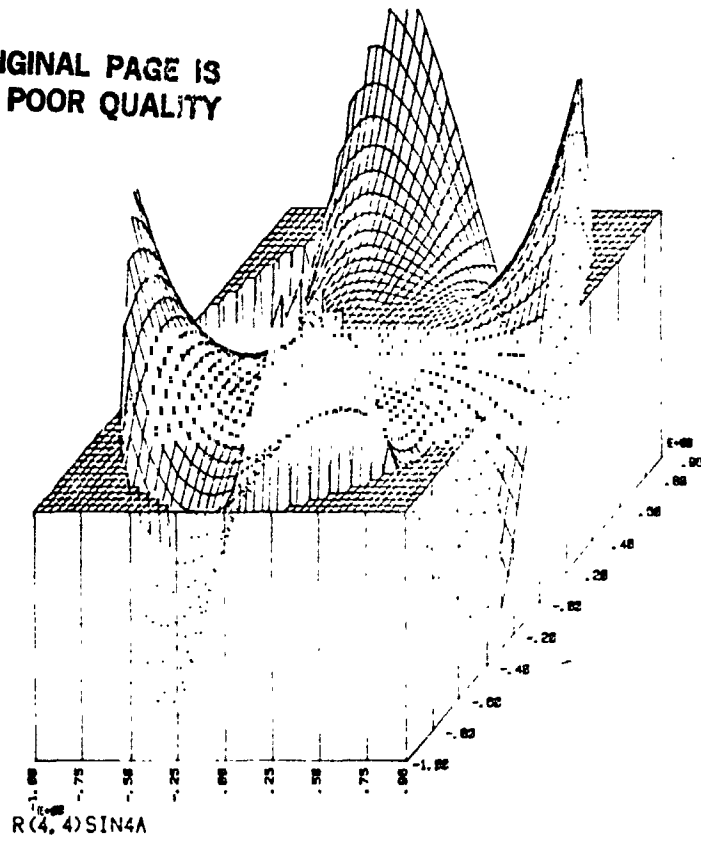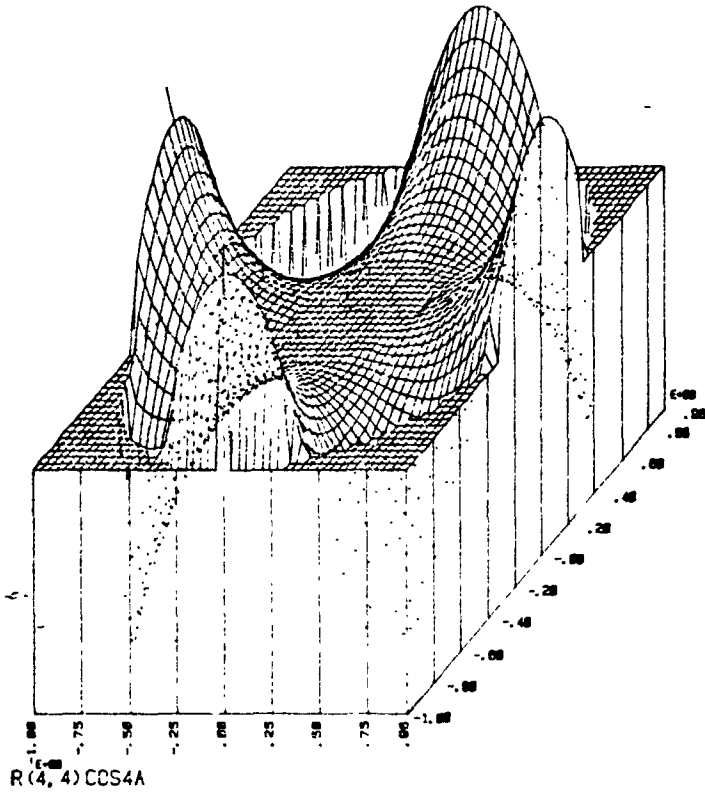
Figure 10

ORIGINAL PAGE IS
OF POOR QUALITY

R(3,5)COS3A

R(3,5)SIN3A

R(5,5)COS5A

R(5,5)SIN5A

Figure 11

R(b, 6)
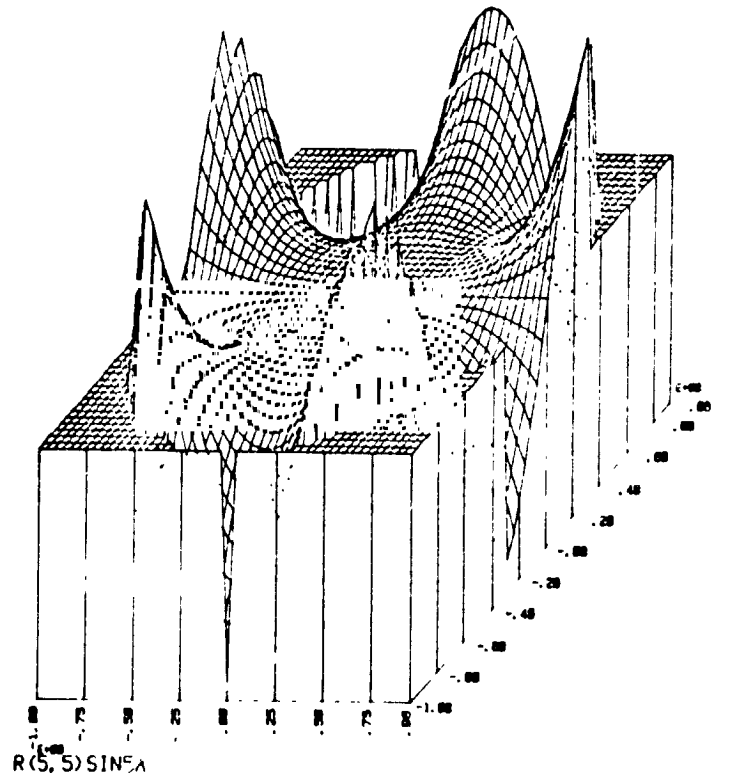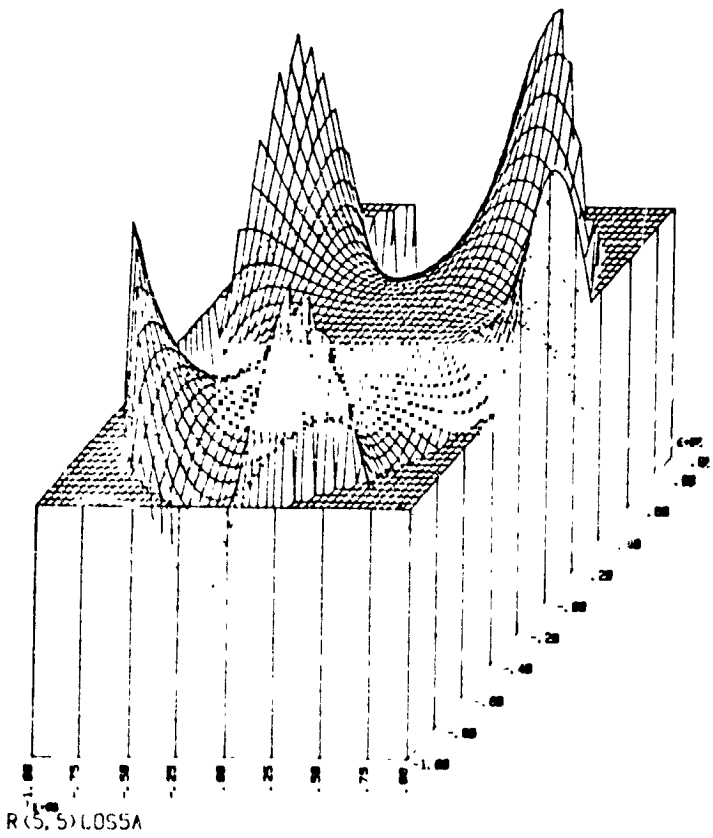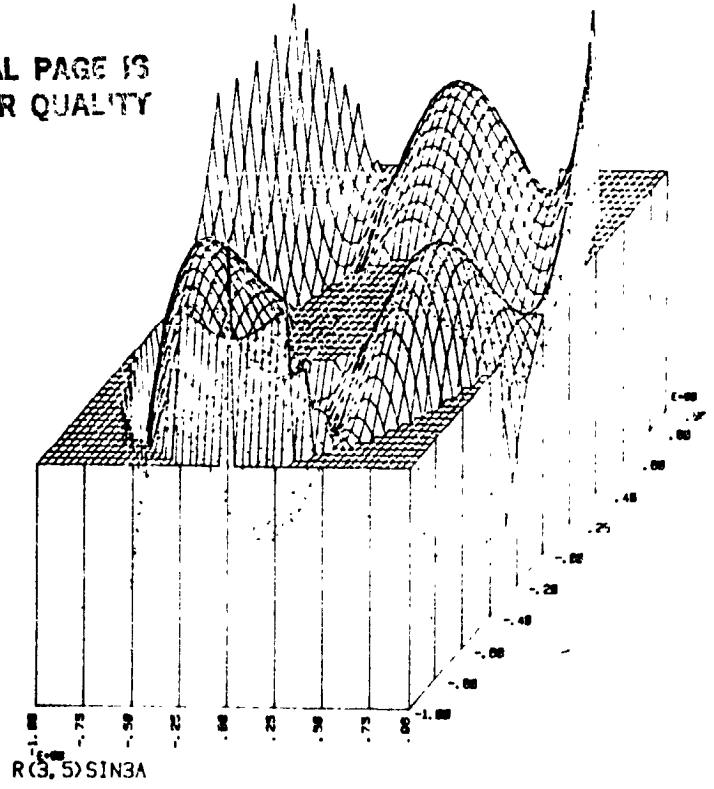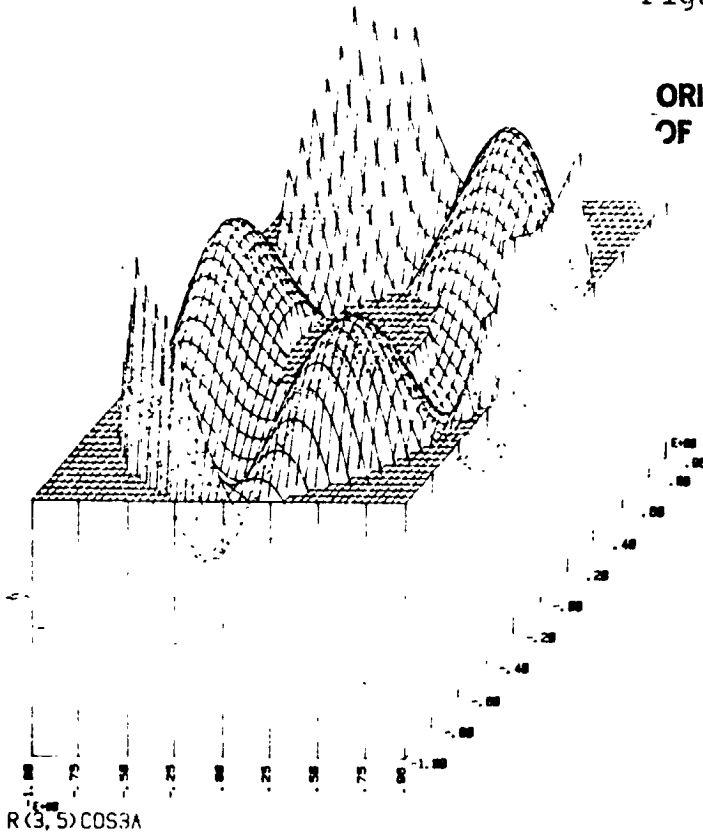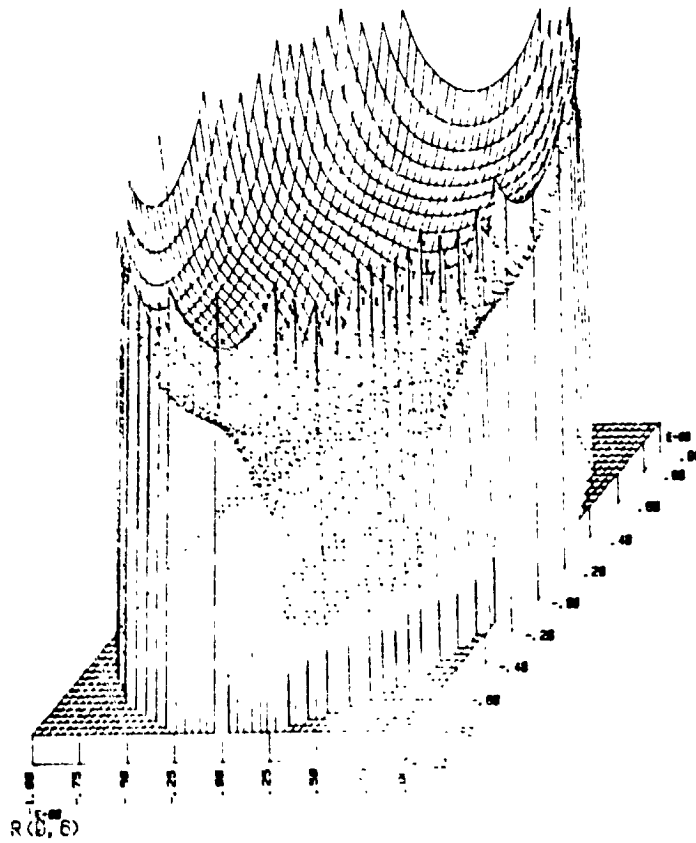
REFERENCES

1. Richard Barakat, Optimum Balanced Wavefront Aberrations for Radially Symmetric Amplitude Disturbances: Generalizations of Zernike Polynomials, J. Opt. Soc. Am. 70, (1980).

2. Bhatia, A. B. and E. Wolf, On the Circle Polynomials of Zernike and Related Orthogonal Sets, Proc. Camb. Phil. Soc., 50, (1954).

3. M. Born and E. Wolf, Principles of Optics, Pergamon Press, (1959).

4. B. R. A. Nijboer, The Diffraction Theory of Aberrations, Physica (Utrecht) 10, (1947).

5. R. J. Noll, Zernike Polynomials and Atmospheric Turbulence, J. Opt. Soc. Am. 66, (1976).

6. Ramsey, Timothy, Optical System Analysis Program, Sperry Report, No. SP-212-0991, (1976).

7. Berge Tatian, Aberration Balancing in Rotationally Symmetric Lenses, J. Opt. Soc. Am. 64, (1974).

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

DEVELOPMENT OF PHOTOMULTIPLIER ELECTRONICS AND COMPUTER INTERFACING
SOFTWARE FOR A HIGH SPEED PHOTOELECTRIC PHOTOMETER-STOKE'S METER

Prepared By: Terry R. flesch, Ph.D

Academic Rank: Assistant Professor

University and Department: North Georgia College
                                    Department of Physics

NASA/MSFC:
  Division: Space Physics
  Branch: Astrophysics

MSFC Counterpart: Dr. Allen Gary and
                                    Dr. Walter Fountain

Date: June 1 -- August 14, 1981

DEVELOPMENT OF PHOTOMULTIPLIER ELECTRONICS AND COMPUTER INTERFACING SOFTWARE FOR A HIGH SPEED PHOTOELECTRIC PHOTOMETER-STOKE'S METER

BY

Terry R. Flesch
Assistant Professor of Physics
North Georgia College
Dahlongega, Georgia

ABSTRACT

The Astrophysics Branch of the Marshall Space Flight Center has been actively involved in the study of high energy astronomical objects using ground-based instrumentation specifically developed at Marshall for this purpose. In particular, work has been carried out using a photoelectric photometer interfaced with an IMSAI 8080 microcomputer. This system allows photoelectric observations of from 10 msec to 10 sec duration through each of several filters.

Work has been carried out this summer involving the initial development of a second system which will allow not only the rapid photoelectric observations, but will have the added capability of obtaining information on the polarization of the radiation emitted by the source under study. The new system will employ a rapidly rotating polarizing element which will allow information to be obtained on the Stoke's parameters in as short a period of time as 1 minute. The polarization measurements and the rotation of the polarizing element will also be controlled by the microcomputer.

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge the assistance which I have received in the realization of this project. All those scientists and technicians with whom I interacted were always more than willing to give of their time and talents to help me out of my innumerable blind alleys.

A special thanks goes to Dr. Allen Gary and Dr. Walter Fountain, who had to put up with my shortcomings daily, and Bob Austin, without whom the electronics circuits would never have been debugged.

I would like to express my appreciations to the National Aeronautics and Space Administration and to the University of Alabama in Huntsville for making such a worthwhile program available to faculty members across the country.

INTRODUCTION

Since the application of microcomputers to the technology available to the astronomer, new kinds of observing equipment and techniques have expanded the observational capabilities of even modest-sized telescopes. This project deals with the preliminary stages in the design and construction of a photoelectric photometer system modeled after a system built at the Marshall Space Flight Center in the late 1970's and currently in use at the NASA/University of Arizona's Mt. Lemmon Observatory.

Of particular interest is the capability of measuring both the intensity and polarization of the radiation emitted by objects which show very rapid light variations. Such objects include pulsars and the optical counterparts of X-ray sources. The rapid nature of the variations in these types of objects necessitates the ability of the instrument to make observations with very high time resolution. In addition, the study of the polarization of the radiation from such objects will lead to a better understanding of the magnetic fields and interstellar material which characterize the immediate environment of these stellar systems. With these factors in mind, the instrumentation being designed involves the interfacing of a microcomputer with the photometer to allow data to be taken with time resolution as high as 10 msec. In addition, the instrument will include magnetic tape and magnetic disk storage capabilities to enable rapid data storage.

OBJECTIVES

The goal of this project dealt with the construction of electronic
circuitry to enable the photoelectric photometer to detect stellar
radiation, to allow communication between the photometer and the micro-
computer, and the application of existing cor uter hardware and software,
to enable a teletype terminal to relay comma. s to the computer-photo-
meter system. The design and construction of the photometer itself,
including the filter box and offset guider is the aim of future work.

THE PHOTOMETRIC SYSTEM

A photoelectric photometer system was designed and constructed at MSFC
in the late 1970's. The instrument allows high time-resolution observa-
tion of rapidly varying astronomical objects and is now being used in
observing programs at the NASA/University of Arizona's 60" telescope at
Mt. Lemmon, Arizona. The main components of the system are schematically
shown in Figure 1.

The current project is a modification of this basic design and the con-
struction of the modified version. Many of the characteristics of the
operational system at M.. Lemmon are being incorporated into this
second generation instrument. The planned modification will allow the
acquisition of polarization measurements of the stellar radiation. This
will be accomplished by including a rotating 0.35 wave plate into the
optical path of the incoming light beam. A 3-element Wollaston prison
will then divide the radiation into two separate beams. Each of these
will be 99.9 percent linearly polarized with their axes of polarization
mutually perpendicular. A separate EMI 9658B photomultiplier tube will
monitor each beam. The relative signal strengths from the two tubes
will enable the four Stoke's parameters and, from these, the polariza-
tion of the incoming radiation to be calculated. A schematic of the
proposed system is shown in Figure 2.

The filter box will contain a set of UBVrI filters which will allow for
photometric observations to be made in a standard astronomical photo-
metric system. The 0.35 waveplate will also be included for use in the
polarization measurement.

The cold box will house two EMI 9658B photomultiplier tubes. Work has
been completed on the wiring and interfacing of a single tube-cold box
assembly with the computer. Current plans call for the construction
of a single-channel photometric system with later expansion to the
dual-channel simultaneous system necessary for obtaining the polariza-
tion information. The tube possesses an S-20 photocathode surface to
enhance its red sensitivity, and will be operated at 1230 volts to
maximize the anode sensitivity. The base of the tube has been wired
in a grounded-cathode configuration for use in the pulse counting mode
(see Figure 3).

Pulse preamplifier, height discriminator and line driver-pulse inverter
circuits will insure a pulse height of at least +1.8V being fed to the
computer. This voltage is necessary to activate the optical isolator
light emitting diode in the computer interface electronics used to
initialize the pulse counting procedure.

The preamp circuit is included in the end bell can of the cold box to
minimize R F noise pick up from the dome and platform electronics.

The remaining electronics are housed in a separate unit. The circuit diagrams for these circuits are shown in Figures 4 and 5. The circuitry has a pulse pair resolution of approximately 600 nsec. This results in a 6 percent pulse coincidence loss at an apparent count rate of 10^5 counts/sec. At an operating voltage of +1230 volts, this count rate corresponds to a star of apparent magnitude of $\approx$ 10.7 when observed with a 60" telescope.

The pulse signal is carried to the computer by a shielded cable and is fed through a (6N137=HP5082-4360) electrically isolated light emitting diode on the computer interface board. The pulses are summed for a length of time previously determined by the observer. The final readings are then sent to a magnetic tape and/or a magnetic disk recording system.

COMPUTER SOFTWARE/HARDWARE

The microcomputer used is an IMSAI-8080 system. Attempts to activate the system at the beginning of the summer were unsuccessful due to a defective MIO board in the computer hardware. The problem led to the replacement of the MIO board with a SIO board. This allowed the inter- facing of the computer with the teletype and magnetic disk drives, but communications between the computer and magnetic tape units will require the MIO board.

Since the development of most of the computer software and the testing of the majority of the necessary computer interfaces and accompanying electronic circuitry were still possible with the SIO board, the project was pursued with this change incorporated into the computer hardware. The correction of the MIO board is being effected through a NASA contract with the Computer Sciences Corporation.

The computer software for the actual data-taking uses the URTH 2.1 language developed at the University of Rochester. This language was chosen due to its widespread use at major astronomical observatories and its flexibility in allowing interaction between operator and computer for defining the various parameters necessary to control the data-taking procedures. Most of the basic software was developed along with the prototype of the photometric system mentioned earlier. However, some modifications in the software will be necessary as the project continues to allow for magnetic tape storage of data and the dual-channel readings to be accepted by the computer. A listing of the current URTH software for data acquisition can be found in appendix I.

TESTING OF THE PHOTOMULTIPLIER ELECTRONICS

The photomultiplier and associated electronics were tested for sensi- tivity by monitoring dark current pulses produced by the tube under a variety of high voltages. The tube was placed inside the cold box and operated at room temperature. Under actual observing conditions, the photocathode will be kept at approximately $-56^\circ C$ by filling the cold

box with dry ice. The dark slide was kept closed during the tests to prevent external light from striking the cathode surface. The resulting pulses were monitored on an oscilloscope and fed into the microcomputer for pulse counting. The number of counts increased with increasing applied voltage as expected (see Table 1). This confirmed the correct performance of the photomultiplier tube and the electronic circuitry discussed earlier. This also proved the reliability of the computer interfacing electronics contained in the computer hardware. A listing of the URTH commands used in the testing procedure is contained in appendix TI.

TABLE 1

Dark current counts (relative scale) vs. voltage applied to photo-multiplier tubes.

Voltage (V)	Counts	Voltage (V)	Counts
1000	1	1300	1600
1050	3	1350	2700
1100	16	1400	4000
1150	55	1450	5500
1200	200	1500	7000
1250	600		

FIGURE 1   MAIN COMPONENTS OF THE OPERATIONAL PHOTOELECTRIC
PHOTOMETER DEVELOPED AT MSFC
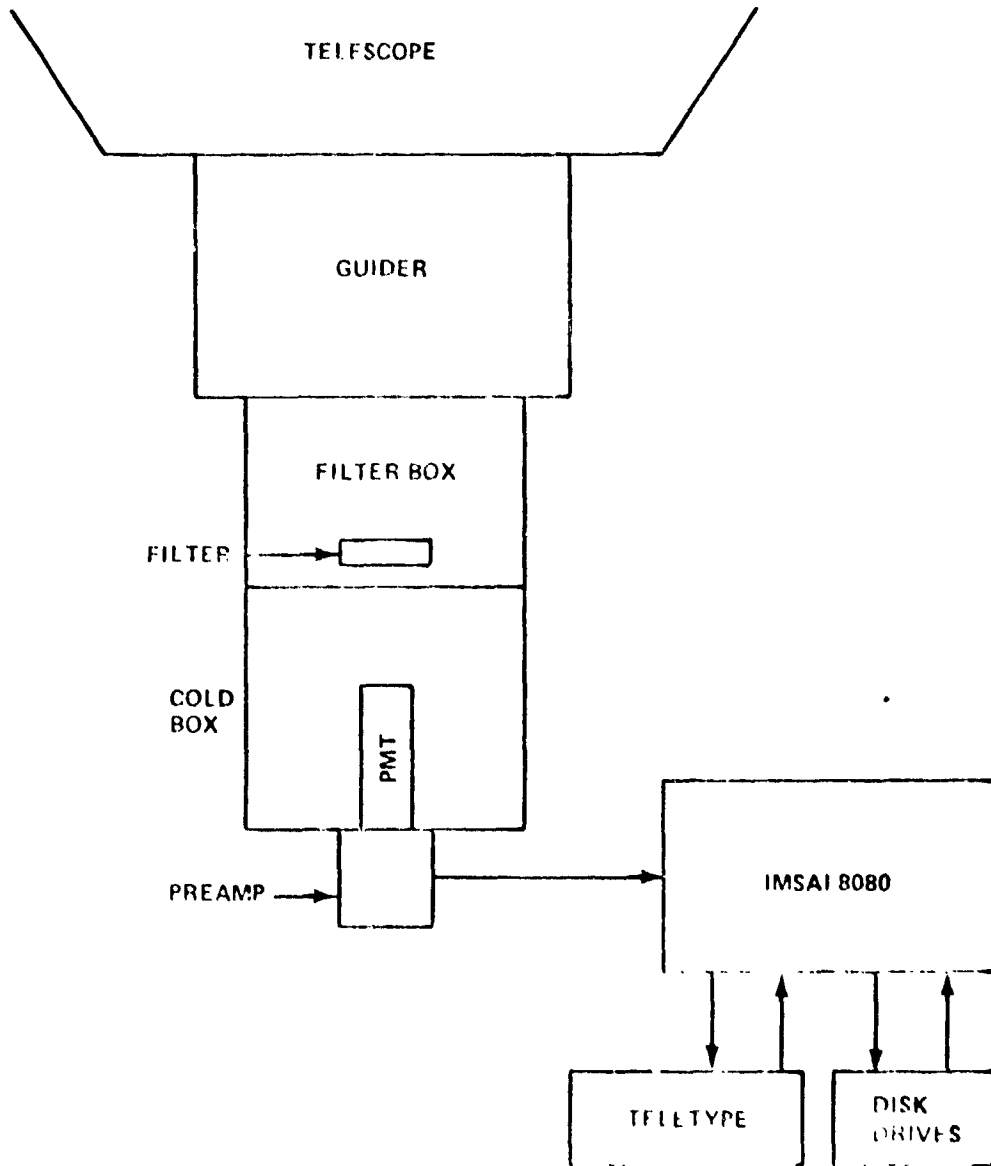
FIGURE 2   MAIN COMPONENTS OF PROPOSED PHOTOELECTRIC PHOTOMETER - STOKE'S METER CURRENTLY UNDER DEVELOPMENT AT MSFC
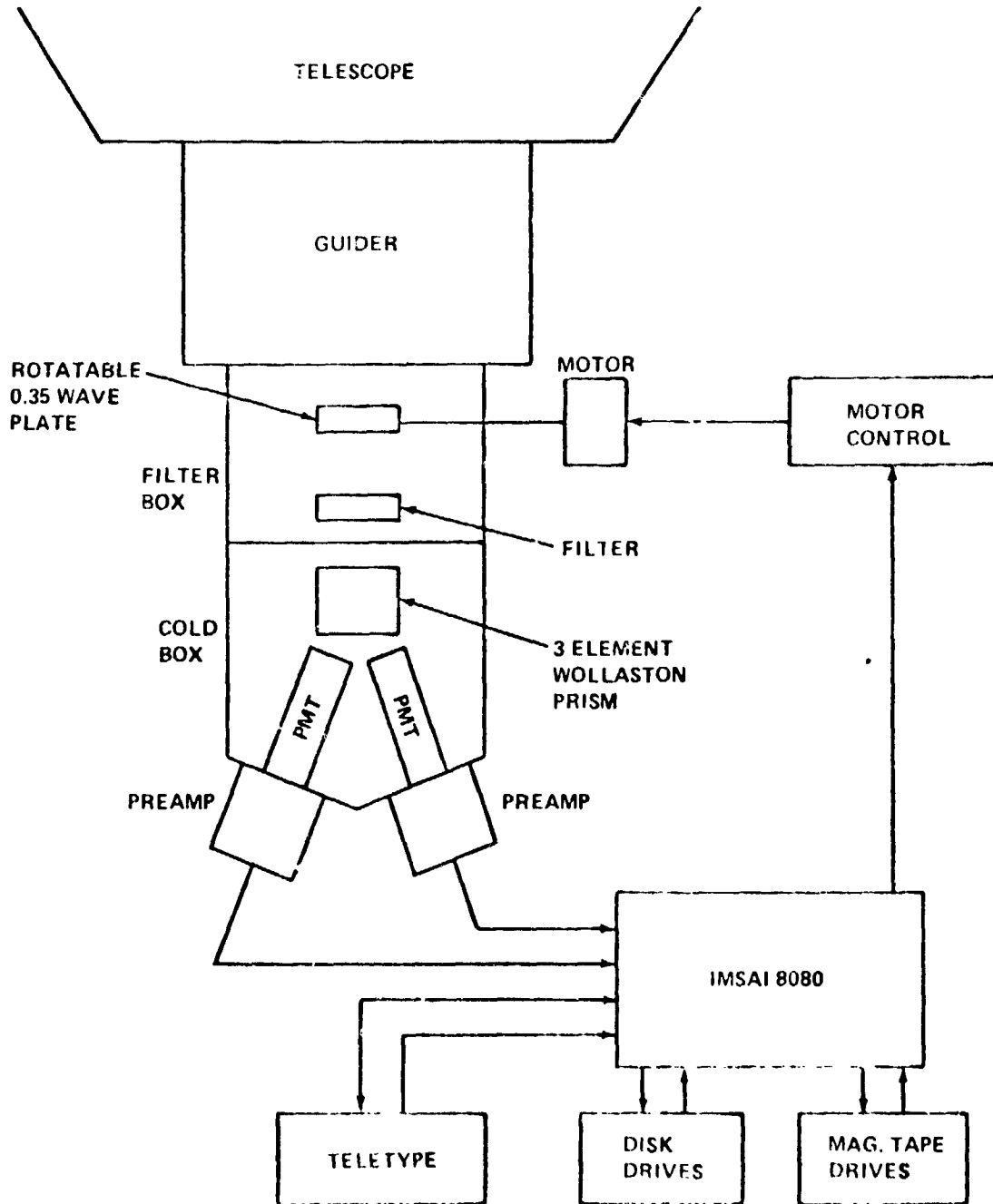
FIGURE 3   PHOTOMULTIPLIER SOCKET VOLTAGE DIVIDER NETWORK FOR THE EMI
9658 PHOTOMULTIPLIER TUBE

VI-11
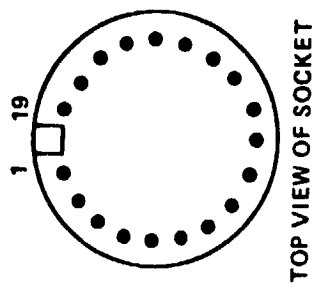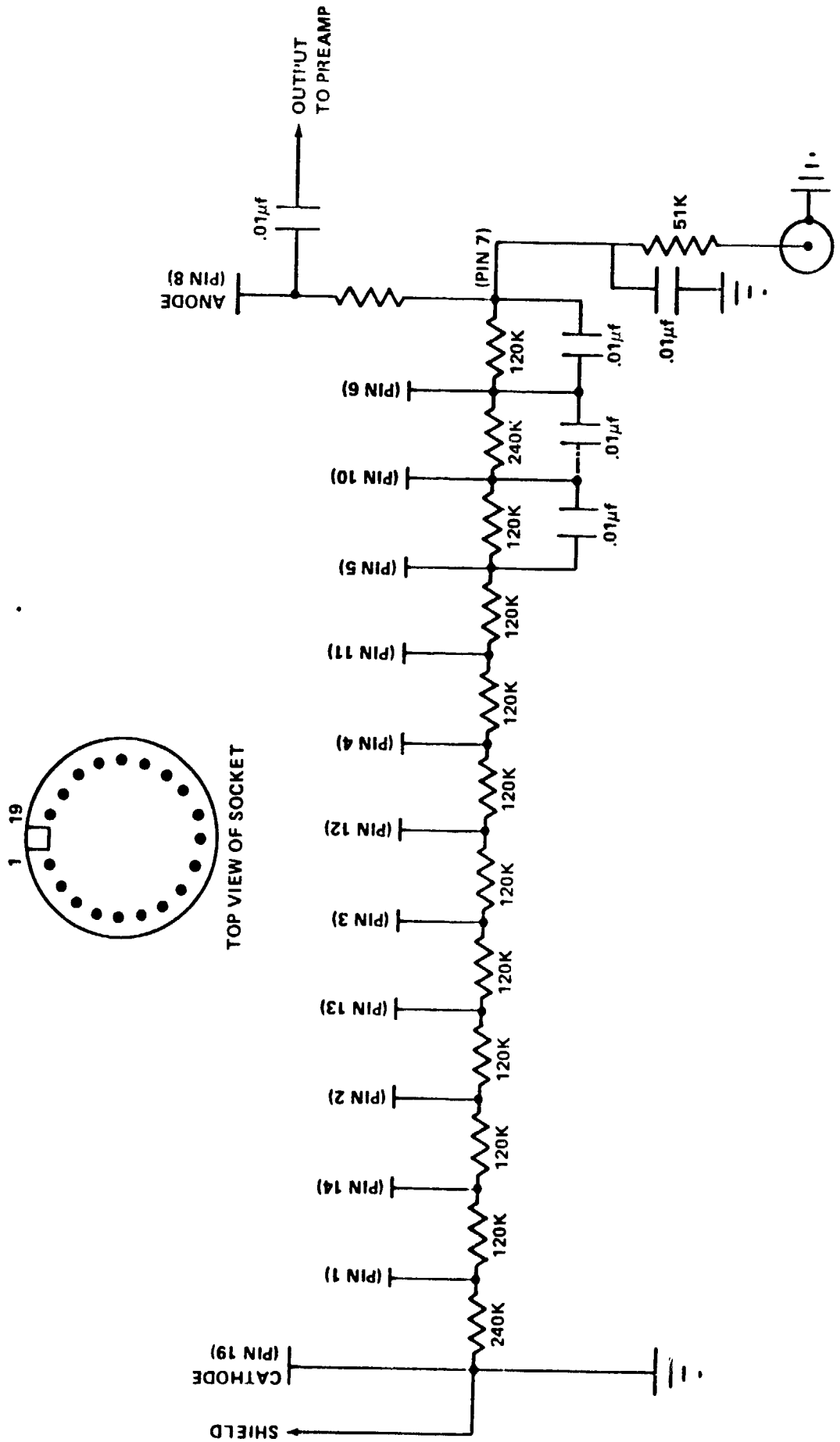
FIGURE 4   CIRCUIT DIAGRAM OF PREAMP AND DISCRIMINATOR ELECTRONICS
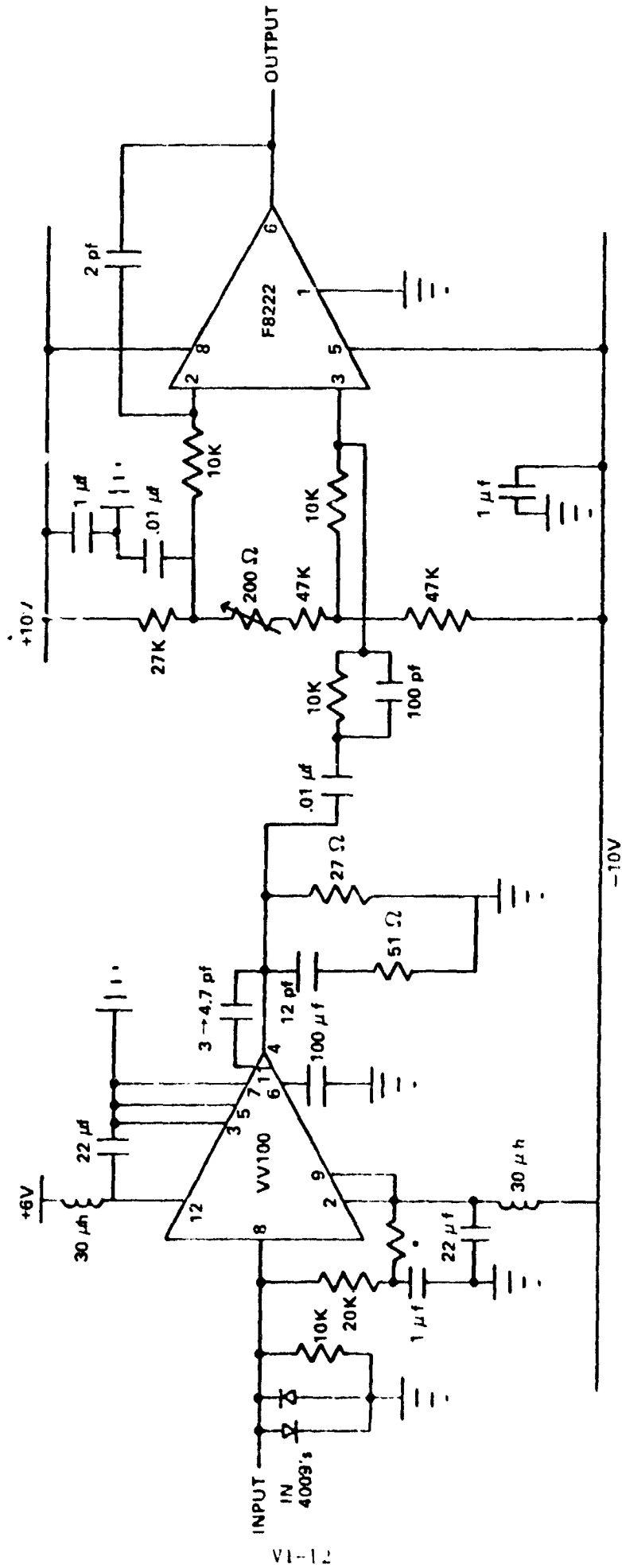
* 68K → 290K ADJUSTABLE
  TO ZERO OUTPUT

FIGURE 5   SCHEMATIC OF LINE DRIVER - PULSE INVERTER CIRCUIT
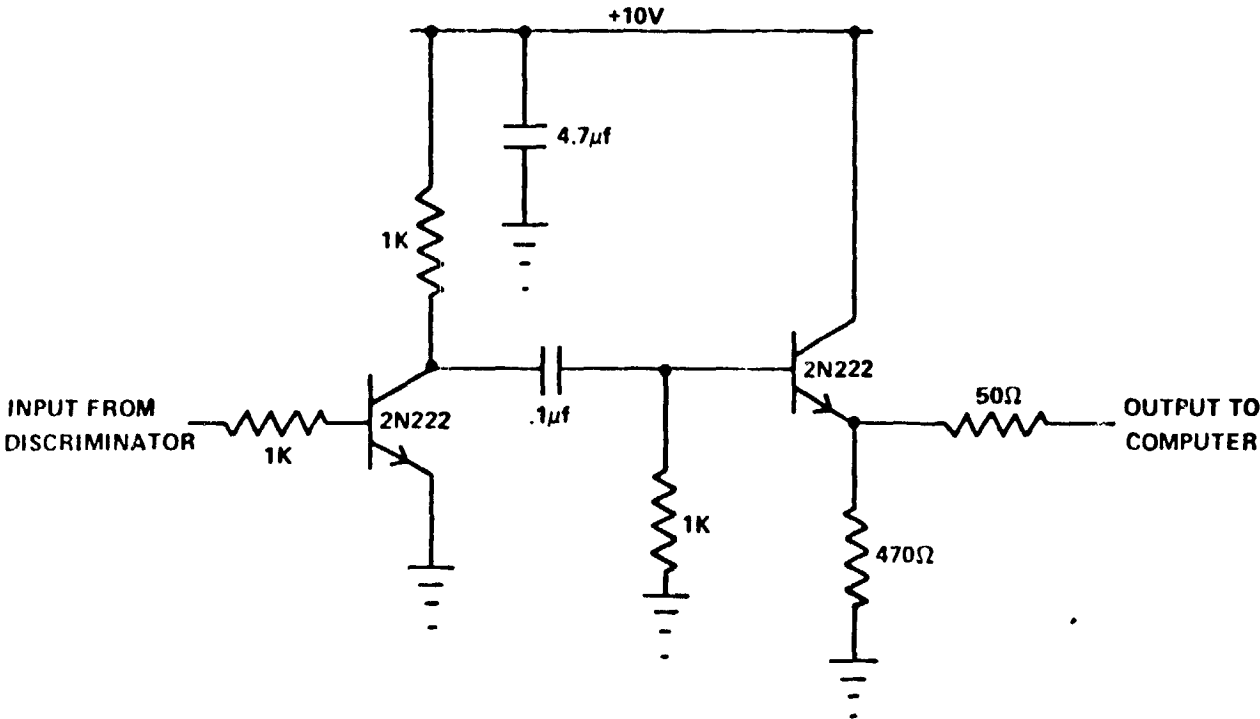
CONCLUSIONS AND RECOMMEDATIONS

This project has seen the successful completion of the first steps in
the development of a photometric-polarimetric system which will prove
invaluable for the study of all types of variable star systems, partic-
ulary systems showing variations on very short (seconds to tens of
milliseconds) timescale. Further work is needed in the development of
the software and interfaces necessary to handle the dual-channel simul-
taneous monitoring capabilities of the final system, but these advances
require only an extension of procedures already shown to be feasible.
In addition, design of the system hardware (guider, filter box, two-
channel cold box) should be possible now that the electronics and
computer interfacing problems are being solved.

There is some doubt by the author that one more summer by a single
faculty member will be able to realize the completion of the system.
With this in mind, arrangements are being made for various components
of the system to be taken by the author to his home institution of
North Georgia College with the hope that more progress can be made
toward the final construction and testing of the instrument.

REFERENCES

1. Gary, G. Allen, et al., "Measurement of the Polarization of Visible
    Light from Optically Identified X-Ray Sources", NASA Discretionary
    Fund Proposed, Nov. 1978.

2. Fishman, Gerald J., et al., "Rapid Photometry and Polarimetry of
    High Energy Objects", NASA Technical Report, 1979.

3. Austin, Robert W., et al., "Characteristics and Operation of an
    Astronomical Photometer for Rapid Time - Series Observations",
    NASA Report, August, 1979.

APPENDIX I

COMPUTER SOFTWARE PREVIOUSLY DEVELOPED AT

MSFC FOR RAPID PHOTOELECTRIC PHOTOMETRY

```
BLOCK 69
0   ( MAGTAPE WRITE ROUTINE -- 04OCT78 BY JGD ) BASE @ HEX
1   200 CONSTANT BLKSIZE  0 CONSTANT #BY  : 0#B 0 ' #BY ! ;
2   : DTP, ASSEMBLER A XRA,  3 OUT, ;  : CTP, ASSEMBLER 80 A MVI,
3     3 OUT, ;  : PULSE, ASSEMBLER  1 OUT, A XRA, A XRA, A XRA,  1
4   OUT, ;  CODE \FULL? CTP,  1 IN, 80 ANI, A C MOV, 0 B MVI,
5   B PUSH, NEXT JMP,  CODE \RDY? CTP, BEGIN,  1 IN, 40 ANI, 0=,
6   END, NEXT JMP,  CODE \WEOF CTP, 4 A MVI, PULSE, NEXT JMP,
7   CODE \REW CTP, 8 A MVI, PULSE, NEXT JMP,  CODE \IRG CTP, 2 A
8   MVI, PULSE, NEXT JMP, CODE \WB CTP, 1 A MVI, PULSE, NEXT JMP,
9   CODE CPREP DTP, B POP, C A MOV, 1 OUT, NEXT JMP,
10  CODE BPREP DTP, B POP, B A MOV, 1 OUT, NEXT JMP,
11  : FULL?  \FULL?  #BY BLKSIZE  = ! ;  : REWIND \RDY? \REW 0#B ;
12  : IRG  10 MSEC \RDY? \IRG 0#B ;  : EOF 10 MSEC \RDY? \WEOF 0#B ;
13  : WRITE \RDY? CPREP \WB ' #BY 1+! FULL? IF IRG THEN ;
14  : DWRITE DUP \RDY? BPREP \WB \RDY? CPREP \WB ' #BY DUP
15  1+! 1+! FULL? IF IRG THEN ;  BASE ! 51 LOAD    ;S


BLOCK 31
0   ( INTERRUPT DEFINITIONS -- 30NOV77 BY RAB ) BASE @ HEX
1   CODE EI EI, NEXT JMP, CODE DI DI, NEXT JMP, 0 VARIABLE VIMASK
2   CODE TXMASK B POP, C A MOV, CMA, 37 OUT, FF OUT, NEXT JMP,
3   : 2** DUP 0- IF DROP 1 ELSE 1 SWAP 0 DO 2 * LOOP THEN ;
4   : ION 2** VIMASK @ ! DUP VIMASK ! TXMASK ;  : 1OFF 2** VIMASK
5     @ X! DUP VIMASK ! TXMASK ;  : ALLOFF 0 TXMASK 0 VIMASK ! ;
6   : INTRLINK SWAP 7 SWAP - 3 <-L 1+ ' ION ;
7   ( COLON INTERRUPT DEFINITIONS )
8   CODE RINTX PSW POP, B POP, D POP, H POP, EI, RET,
9   : ;I ' RINIX 3 - , 0 STATE ! CONCUR INTRLINK ;
10  IMP ;I  : !: DUP OCD 1B1 C! : 0C3 1B1 C! HERE 3 - ;CODE XTHL,
11          D PUSH, B PUSH, PSW PUSH, NEXTHL JMP,
12  ( CODE INTERRUPT DEFINITIONS )
13  : 'CODE DUP CODE HERE ;
14  : ;C! [ ASSEMBLER ] RET, INTRLINK ;        IMP ;C+
15   ALLOFF  DI  BASE !         ;S


BLOCK 51
0       ( PHOTON COUNTING ROUTINES BY JGD  OCTOBER 1978 )
1   0 VARIABLE #DAY  0 VARIABLE #MON  0 VARIABLE #Y'R
2   0 VARIABLE #STR  0 VARIABLE #RUN  0 VARIABLE ?NEW
3   0 VARIABLE ?ON?  0 VARIABLE ?GO? 0 VARIABLE LOC
4   0 VARIABLE LGTH  0 VARIABLE BYTS  0 VARIABLE NEW*
5   0 VARIABLE 1FLAG 0 VARIABLE TYPE  0 VARIABLE SPEED
6   0 VARIABLE HOURS 0 VARIABLE MINS  0 VARIABLE STATUS
7   0 VARIABLE OLOC  0 VARIABLE SECS  0 VARIABLE NWTYP
8   0 VARIABLE 1SECS 0 VARIABLE 1MIN  0 VARIABLE 1HOUR
9   0 VARIABLE 2SECS  0 VARIABLE 2MIN  0 VARIABLE 2HOUR
10  0 VARIABLE NEWLGTH  0 VARIABLE RMODE 0 VARIABLE 2MODE
11  2050 VARIABLE OBLON  2050 VARIABLE ENDBLOCK 2050 VARIABLE BLON
12  0 VARIABLE TALLY  0 VARIABLE #BLOCKS
13  0 VARIABLE RATE  0 VARIABLE .1SEC 0 VARIABLE .01SEC
14  1 VARIABLE EXTEND
15  31 LOAD 52 LOAD         ;S
```

```
BLOCK 52
 0  : SASK GCH 48 - ; : 2SASK SASK 10 * 2SK + ; : ECHO T"      "
 1   DUP . ; : 4SASK 2SASK 100 * 2SASK + ; : TEST DUP @ . 2SASK DUP
 2   176 + 0 IF 2DROP ELSE SWAP ' THEN ;       HEX
 3   CODE 1GET FO IN, A L MOV, O H MVI, PUSH JMP,
 4   CODE 2GET FO IN, A L MOV, F1 IN, A H MOV, PUSH JMP,
 5   CODE 3GET F2 IN, A L MOV, O H MVI, H PUSH, FO IN, A L MOV,
 6          F1 IN, A H MOV, PUSH JMP,
 7   CODE OSET 7 A MVI, F4 OUT,  1 A MVI, FF OUT, NEXT JMP,
 8   CODE 1SET 6 A MVI, F4 OUT,  3 A MVI, FF OUT, NEXT JMP,
 9   CODE 2SET 5 A MVI, F4 OUT,  7 A MVI, FF OUT, NEXT JMP,
10   CODE 3SET 4 A MVI, F4 OUT,  F A MVI, FF OUT, NEXT JMP,
11   CODE 4SET 3 A MVI, F4 OUT, 1F A MVI, FF OUT, NEXT JMP,
12   CODE 5SET 2 A MVI, F4 OUT, 3F A MVI, FF OUT, NEXT JMP,
13   CODE 6SET 1 A MVI, F4 OUT, 7F A MVI, FF OUT, NEXT JMP,
14   CODE 7SET O A MVI, F4 OUT, FF A MVI, FF OUT, NEXT JMP,
15   DECIMAL 53 LOAD ;S


BLOCK 53
 0  : OLIN T" DAY # ?    (2#'S) " #DAY TEST T"    " #DAY @ . ;
 1  : 1LIN T" MONTH # ? (2#'S) " #MON TEST T"    " #MON @ . ;
 2  : 2LIN T" YEAR # ?  (2#'S) " #Y'R TEST T"    " #Y'R @ . ;
 3  : 3LIN T" STAR # ?  (2#'S) " #STR TEST T"    " #STR @ . ;
 4  : 4LIN T" RUN # ?   (2#'S) " #RUN TEST T"    " #RUN @ . ;
 5  : 5LIN T" # DATA PTS ? ( 4 DIGITS ) " 4SASK ECHO LGTH ! ;
 6  : 6LIN T" RECORDING MODE( 01 = TAPE , 02 = DISC , 03 = BOTH ) "
 7     "         RMODE TEST T"   " RMODE @ . ;
 8  : 7LIN T" BYTS/PT. ? (1,2 OR 3 ) " SASK ECHO BYTS ! ;
 9  : 8LIN T" SPEED ? ( 1 FAST 2 SLOW ) " SASK ECHO SPEED ! ;
10  : 9LIN T" TYPE ( 0 PHOTOM 1 POLARM ) ? " SASK ECHO TYPE ! ;
11  : 10LIN T" PHOTOMETRY RUN " ; : 11LIN ." POLARIZATION RUN "
12  CR T" NOT YET PROGRAMMED " ;
13  : 12LIN T" ILLEGAL VALUE OF RMODE " CR 6LIN ;    "
14  : 13LIN T" RATE (1,2,3 FOR1,.1,.01 SEC)" SASK ECHO RATE ! ;
15   54 LOAD INSTALL NEW.ABORT.DEFINITIONS          ;S


BLOCK 54
 0  : INCBLON BLON 1+! ; : UPRUN #RUN 1+! ; : +FL 1FLAG 1+! ;
 1  : 2+! DUP 1+! 1+! ; : INCLOC LOC @ 1023 - 0= IF +FL 0 LOC !
 2  INCBLON ELSE LOC 2+! THEN ;
 3  : SENDEM BLON @ 'BLOCK LOC @ + ' INCLOC ;
 4  : 2TUCK RMODE @ DUP 1 - 0= IF DROP DWRITE ELSE DUP 2 - 0= IF
 5  DROP SENDEM ELSE 3 - 0= IF DUP DWRITE SENDEM ELSE 12LIN
 6  THEN THEN THEN ; : 1GRAB 1GET 2TUCK ;
 7  : 2GRAB 2GET 2TUCK ; : 3GRAB 3GET 2TUCK 2TUCK ;
 8  : +HOURS HOUR 23 - 0= IF 0 MINS !
 9  +HOURS ELSE MINS 1+! THEN ; : +SECS SECS @ 59 - 0= IF 0 SECS !
10  +MINS ELSE SECS 1+! THEN ; : NOW SECS @ MINS @ HOURS @ ;
11  : 10 CR T" HOURS(2#'S) " 2SASK HOURS ! CR T" MINS(2#'S) " 2SASK
12  MINS ! CR T" SECS (2#'S) "2SASK SECS ! 4SET E! ;
13  : TIME NOW . . . ; : 1START NOW 1HOUR ! 1MIN ! 1SECS ! ;
14  : TEND NOW 2HOUR ! 2MIN ! 2SECS ! ;
15   55 LOAD          ;S
```

```
BLOCK 55
 0  : 1HOUSE CR OLIN CR 1LIN CR 2LIN CR TO ;
 1  : 2HOUSE CR 3LIN CR 4LIN CR 9LIN CR 7LIN CR 5LIN CR 8LIN
 2  CR 6LIN CR 13LIN ;
 3  : CHECK BEGIN CR T" PARAMETERS OK ? ( 1 YES 0 NO ) " SASK
 4  DUP 0= IF 2HOUSE ELSE THEN 1 - 0= END ;
 5  : MODETEST TYPE @ DUP 0= IF DROP 10LIN ELSE 1- 0= IF 11LIN
 6  ELSE T" UPROGRAMMED TYPE " CR 9LIN  THEN THEN ;
 7  : BYTETEST BYTS @ DUP 1 - 0= IF DROP ELSE DUP 2 - 0= IF
 8  DROP ELSE 3 - 0= IF ELSE T" BYTE ERROR " CR   7LIN
 9  THEN THEN THEN ;
10  : MODESWITCH 1FLAG @ 10 : IF 1 RMODE ! THEN ;
11  : DATAGET BYTS @ DUP 1 - 0= IF 1GRAB DROP ELSE DUP 2 -
12  0= IF 2GRAB DROP ELSE 3 - 0= IF 3GRAB ELSE T" ERROR "
13  THEN THEN THEN MODESWITCH TALLY 1+! ;
14  : RMODETEST RMODE @ 1   IF CR  12LIN THEN RMODE @ 3   IF CR
15  12LIN THEN ; 56 LOAD        ;S
```

```
BLOCK 56
 0  : DISKPREP 2131 2050 DO 0 I BLOCK ! LOOP ;
 1  : FIRSTRUN CR CR CR 1HOUSE CR CR CHECK MODETEST CR BYTETEST
 2  RMODETEST ;  10 BUFALLOC
 3  : NBLOCKS BYTS @ 2 - 0 = IF LGTH @ 512 / 1BLOCKS ! ELSE
 4  LGTH @ 256 / 1BLOCKS ! THEN ;
 5  : CHKBLK 1BLOCKS @ 10   IF 10 1BLOCKS ! THEN OBLON @
 6  1BLOCKS @ + ENDBLOCK !  0 TALLY ! ;
 7  : 1TEXT T" THIS IS RUN # " 1RUN @ . ;
 8  : 2TEXT T" IT STARTS AT BLOCK # " OBLON @ . ;
 9  : 3TEXT T" IT ENDS AT BLOCK # " ENDBLOCK @ 1 - . ;
10  : 4TEXT T" IT HAS TOTAL # OF BYTES " LGTH @ . ;
11  : TPHSK 37 1 DO I 1 - 11 * 1DAY + @ DWRITE 10 MSEC LOOP ;
12  : PREPDISC NBLOCKS ENDBLOCK @ DUP OBLON ! BLON ! ;
13  : 1WAY RMODE @ 2MODE ! ; : 2WAY 2MODE @ RMODE ! ;
14  57 LOAD        ;S
15
```

```
BLOCK 57
 0  : RFIX 2MODE @ RMODE ! ;
 1  : TEXT STATUS @ 1 - 0= IF T" END IF RUN  " CR ELSE THEN ;
 2  : ABORT 0 STATUS ! OBLON @ DUP ENDBLOCK ! BLON ! 0 LOC ! ;
 3  HEX CODE SENSE 03 IN, 10 ANI, A L MOV, 0 A MVI, A H MOV,
 4  PUSH JMP, DECIMAL
 5  : STALL BEGIN SENSE 16 - 0=  END ;
 6  : TPCLR RMODE @ 2 = IF ELSE TPHSK IKG 100 MSEC EOF 5000 MSEC
 7  EOF  THEN ; : TFLUSH RMODE @ 1   IF FLUSH THEN ;
 8  : RESET MINS DUP 1+! 14! 0 SECS ! 0 .1SEC ! 0 .01SEC ! ;
 9  : NSET STATUS @ 0= IF 4SET ELSE 2SET THEN ;
10  : FINIS T" START" .SEC @ 1MIN @ 1HOUR @ . . . CR T" END"
11  2SECS @ 2MIN @ 2HOUR @ . . . ;
12  : ENDTEST TALLY @ LGTH @   IF EXTEND @ 0= IF LGTH @ 256
13  - TALLY ! 1WAY 0 STATUS ! TEND STALL  SET 2WAY TEXT FINIS
14  TPCLR TFLUSH STALL 4SET UPRUN CR TEXT THEN PSET ;
15  : CONTINUE 0 EXTEND ! ; : EXIT 1 EXTEND ! ; 58 LOAD ;S
```

```
BLOCK 58
 0  : PPREP NBLOCKS ENDBLOCK @ OBLON ! OBLON @ BLON ! CHKBLK CR
 1  0 2134 !BLOCK ! OBLON @ DUP 10 + SWAP DO 0 I !BLOCK ! LOOP ;
 2  3 !: ZAPPO OSET DATAGET ENDTEST NSET ;!
 3  : RESTART #RUN 1+!    PPREP ;
 4  2 !: CLOCKER OSET +SECS NSET ;!
 5  : NEXTRUN 2WAY 0 1FLAG ! 0 TALLY ! CHECK CR
 6  MODETEST CR BYTETEST CR RMODETEST ;
 7  : READY 1WAY 7SET STALL RMODE @ 2 = IF ELSE TPHSK IRG
 8  THEN RMODE @ 1 = IF ELSE PPREP THEN 0 TALLY !
 9  0 1FLAG ! STALL RESET 1 STATUS ! TSTART 2GET DROP 2SET ;
10    ;S
11
12
13
14
15
```

COMPUTER SOFTWARE USED IN THE TESTING OF
PHOTOMULTIPLIER TUBE ELECTRONICS

```
BLOCK 69
0    0 VARIABLE DLAST 0 VARIABLE BN 0 VARIABLE LOC
1    0 VARIABLE SUM : AD DUP SUM @ + SUM ! ;
2    : PDD 31 1 DO DATAGET LOOP !FLUSH ;
3    : RDD 6 1 DO BN @ @ DUP DLAST @ - . DLAST ! BN @ 2 + BN !
4    LOOP CR ; : TSUM CR T" SUM OF COUNTS = " SUM @ . ;
5    : GETIT 2GET ; : 1WDI 6 1 DO GETIT AD . LOOP CR ;
6    : 1GDI 0 SUM ! 7 1 DO 1WDI LOOP TSUM ;
7    : 2WDI 6 1 DO GETIT DUP DLAST @ SWAP DLAST ! - . LOOP CR ;
8    : 2GDI 7 1 DO 2WDI LOOP ;
9    : RDO 2050 BLOCK BN ! 0 DLAST ! 7 1 DO RDD LOOP ;
10   : RD1 2050 BLOCK 1 + BN ! 0 DLAST ! 7 1 DO RDD LOOP ;
11   : RDC 7 1 DO RDD LOOP ;   ;S
12
13
14
15
```

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

DEVELOPMENT OF HIGH STRENGTH,
HIGH TEMPERATURE CERAMICS

Prepared By: William B. Hall, Ph.D.

Academic Rank: Professor

University and Department: Mississippi State University
                                      Department of Chemical Engineering

NASA/MSFC:
    Division: Non-Metallic
    Branch: Ceramics

MSFC Counterpart: Marshall King

Date: August 7, 1981

Contract No.:   NGT 01-008-021
                                      The University of Alabama
                                      in Huntsville

A PRELIMINARY STUDY OF THE DEVELOPMENT OF HIGH STRENGTH, HIGH TEMPERATURE CERAMICS

BY

William B. Hall, Ph.D.
Professor of Chemical Engineering
Mississippi State University
Mississippi State, Mississippi

ABSTRACT

The operation of rocket engine turbine pumps is limited by the temperature restrictions of metallic components used in the systems. Mechanical strength of these metal components typically decreased drastically at high temperatures. Ceramic materials that retain strength at high temperatures appear to be promising candidates for use as turbine blades and impellers. These high strength materials are sensitive to many related processing parameters such as impurties, sintering aids, reaction aids, particle size, processing temperature, and post thermal treatment. The objectives of this research obtained are:

1. Identified and defined the processing parameters that effect the properties of Si_3N_4 ceramic materials.

2. Specified methods required for processing high strength ceramics.

3. Suggested test for evaulating the high temperature properties of Si_3N_4.

INTRODUCTION

The Space Shuttle Orbiter is the heart of NASA's space transportation system during the decade of the eighties. This reusable vehicle trims the cost of space travel while increasing NASA's capabilities in space. The system can carry a crew of three, four scientist, and a sixty five thousand pound load into orbit. The Orbiter returns to earth like an airplane and it is anticipated that it can make cne hundred round trips into space and back. Although the Shuttle System utilizes the most advanced aerospace technology available to date to achieve economical and useful space flight, it is anticipated various improvements will be made on the system as experience is gained and new technologies emerge. Marshall Space Flight Center has the main responsibility for the Space Shuttle main engine and solid rocket boosters used in this system.

One area in which improvement is needed is in the high-pressure turbopumps, both fuel and oxidizer, in the Space Shuttle main engine. The operation of these pumps is limited by temperature restrictions of the metallic components used in these pumps. Spot melting, oxidation, and erosion-corrosion are some of the problems encountered in the turbopumps. Figures 1 & 2 are flow diagrams of the turbopumps showing approiate temperatures and pressures. The maximum use temperature shown is 1731°R which is in the liquid hydrogen pump. The maximum use temperature is the liquid oxygen pump is 1471°R. These temperatures are on the upper edge of permissable use temperatures for the available metals utilized in the pumps. This study will specifically Si_3N_4, in some applications in the turbopumps.

Figure 1. SSME Proplellant Flow Schematic with temperatures
and pressures indicated.

VII-2
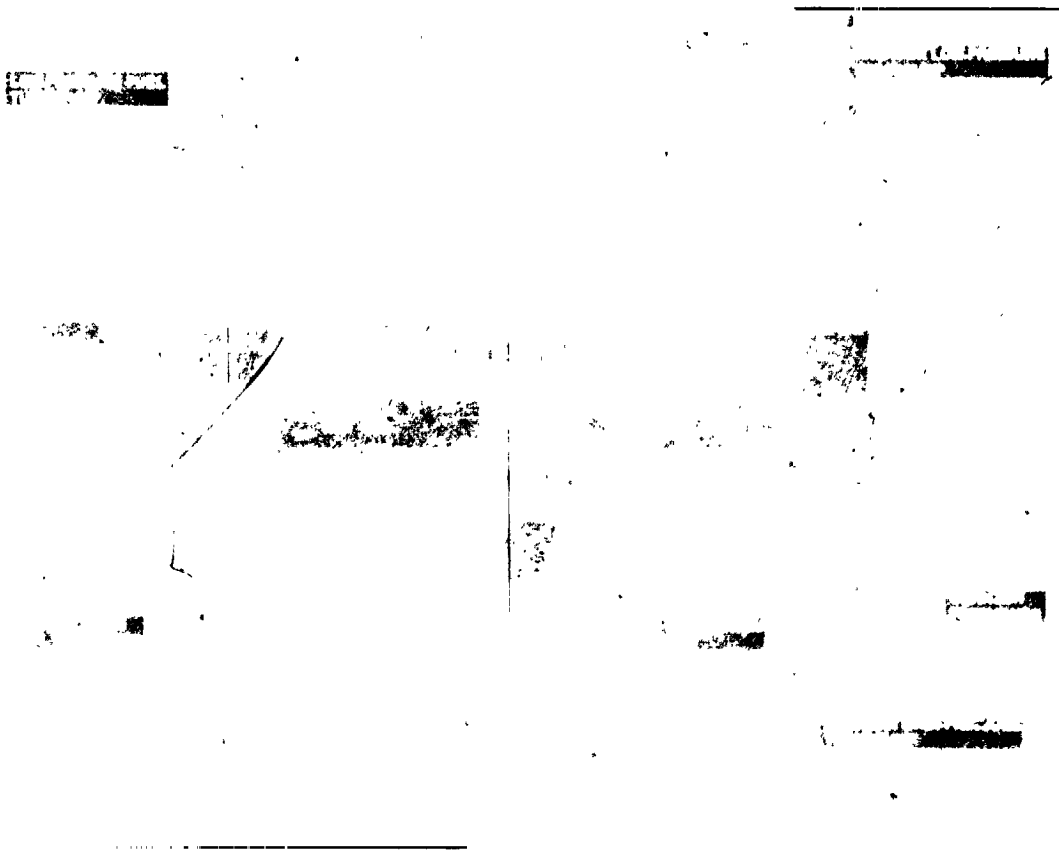
Figure 2. SSME Propellant Flow Schematic with enlarged turbopumps.

OBJECTIVES

The objectives of this two year study are:

1. Identify and define the processing parameters that affect the properties of Si_3N_4 ceramic materials.

2. Design and assemble equipment required for processing high strength ceramics.

3. Design and assemble test apparatus for evaluating the high temperature properties of Si_3N_4.

4. Conduct a research program of manufacturing and evaluating Si_3N_4 materials as applicable to rocket engine applications.

ENGINEERING ANALYSIS

Silicon nitride has the desired properties for utilization as a high temperature structural material. These properties include high strength, low coefficient of friction, high decomposition temperature, good corrosion resistance, good oxidation resistance, high wear resistance and good thermal-shock resistance. These positive features have been known for two decades and slowly silicon nitride is approaching the fulfillment of its potential. Two problems have slowed the use of silicon nitride. The first is the fabricating of suitable or useful shapes with the desirable properties. The second is the brittleness of silicon nitride, and indeed, all ceramic materials. Traditionally, all engineering design compromises on the selection of a material to use in an application, taking into consideration the total cost of a material and the use properties of that material. The evaluation of these compromises in design generally show that a metallic material has a higher net profit. Because of this, there is very little design experience with structural applications of silicon nitride or other brittle ceramics. It is being considered at this time because of the great need for the potential of Si_3N_4 and the development of computer capability to define the stresses as required for brittle material design.

There are two hexagonal crystal structures of silicon nitride, α-Si_3N_4 and β-Si_3N_4, with the α-Structure unit cell being approximately twice the size as the β-Structure unit cell. Since either structure can be produced from the other by rotation of two basal plones, requiring breaking and forming of primary bonds, the $\alpha \rightarrow \beta$ transformation requires time and energy. Two possible methods are solutioning-precipitation and volatilization-condensation. Both structures can be formed over a wide range of temperatures and both are relatively stable, although β-Si_3N_4 appears to be more stable at temperatures in excess of 1500°C. The formation of Si_3N_4 through a vapor-phase reaction favours the formation of the α-Structure, while precipitation from the liquid state favours the β-Structure fromation. The morphology of the formed crystals is important with the fibrous β-Structure preferred because it enhances the strength and toughness of the resulting bulk material.

There are three basic methods of forming bulk Si_3N_4. The first of these three is called "reaction-bonded" Si_3N_4 whereby formed piece of pressed silicon powders is nitrided in nitrogen gas in the range of 1300-1400°C. During this nitridation process, the silicon powders are reacted with N_2 gas to form a mixture of α- and β-Si_3N_4. The second method is called "hot-pressed" Si_3N_4. This process entails the nitriding of Si powders to form α-Si_3N_4 powders. These Si_3N_4 powders are mixed with desirable additives and then pressed into a compact in a graphite mold, under a pressure of 1-2 tons/in^2, at a temperature of 1700-1800°C. This results in a high strength, high density β-Si_3N_4 product. The third process is called "sintered" Si_3N_4. In the sintered product, α- and/or β Si_3N_4 powders are mixed with desirable additives and then pressed into the desired shape. This pressed bulk is then sintered in a controlled atmosphere at 1800-1911°C for a considerable length of time. The product of this sintering operation is essentially β-Si_3N_4. Typical physical properties of these three different processed Si_3N_4 products are given in Table I as published by K_aT_z and Rocketdyne is given in Table II. While reaction-bonded Si_3N_4 increases in strength with an increase in use temperature. This produces a cross-over in highest strength of the three types. While reaction-bonded is considerably weaker at room temperature than hot pressed or sintered Si_3N_4, it is the strongest at a use temperature of 1375°C. These typical properties are very dependent on the starting materials and processing parameters. Some of the parameters that are controlled in the manufacturing of bulk Si_3N_4 are as follows:

Table I. Typical properties of silicon nitride.

Type	Bend strength (four-point) (MPa)			Young's modulus (GPa)	Coefficient of thermal expansion, or $(10^{-6}°C^{-1})$
	RT*	1000°C	1375°C		
Hot-pressed (MgO additive)	690	620	330	317	3.0
Sintered (Y$_2$O$_3$ additive)	655	585	275	276	3.2
Reaction-bonded	210	345	380	165	2.8

Table II. Characteristic Flexural Strength of Silicon Nitride/KSI

TEMP (°F)	Rocketdyne Injection Molded/Sintered	Competitive products Hot Presses	Reaction Bonded
70	70	82	35
1000	70	81	45
1800	70	80	49
2100	67	70	55
2500	35		

a. Initial particle size and size distribution
b. Impurities
c. Milling and Mixing procedures
d. Sintering aids
e. Reaction (sintering) temperature
f. Reaction (sintering) pressure
g. Reaction (sintering) time
h. Reaction (sintering) atmosphere
i. Post heat treatment

Not all of these parameters apply to each of the three types of bulk solid Si_3N_4. Normal impurities present in Si_3N_4 regardless of type are iron, luminum, calium, magnesium and oxygen. These elements are usually present in the combined state such as oxides, silicides or silicates. A typical chemical analysis of a silicon nitride powder as reported by Vasilos Cannon and Wvensch is given in Table II.

The preceeding varying parameters causes a wide range in phase equilibrium and microstructures in the final bulk Si_3N_4 which in turn causes a wide range in properties such as fracture toughness, strength, creep, and oxidation resistance. For example, Edington et al gave K_{ic} values for Si_3N_4 (with varying starting powders and processing parameters) of 2.11 MN/m$^{3/2}$. The other properties varied in values in a similar way.

VII-7

TABLE III

CHEMICAL ANALYSIS OF AS-RECEIVED SILICON NITRIDE POWDER

ELEMENT	CONCENTRATION DETERMINED (wt. %)
Calium	0.03
Iron	0.39 ± .04
Aluminum	0.085 ± .009
Magnesium	.0008
Sodium	0.001
Potassium	0.001
Carbon	0.28 ± .03
Antimony	0.003
Manganese	0.001
Nickel	0.01
Chromium	0.02
Cobalt	0.01
Copper	0.02
Oxygen	1.4

CONCLUSIONS AND RECOMMENDATION:

Si_3N_4 is a viable material candidate for utilization in high temperature structural applications. However, it must be fully evaluated under use conditions and environments before utilization. The tv main areas of concern for turbopump application would be thermal shock and exposure to fuel/fuel products. It is recommended that these two areas be studied and evaluated under actual environments existing in the turbopumps early in the program. These evaluation should continue throughout the life of the program as long as any change is made in the Si_3N_4 processing parameters. Additional studies could include:

a. Creep and strength improvement based on different sintering aids

b. Forming techniques

c. New processing techniques such as reaction-bonding followed by sintering

d. Basic studies on the glass bond formed during sintering.

e. Effects of varying furnace environments.

REFERENCES

1. Edington, J. W., Rowcliffe, D. J., and Henshall, J. L., The Mechanical Properties of Silicon Nitride and Silicon Carbide, Powder Metallurgy International, Vol. 7, No. 2, 1975,

2. Some Exempt Vol. 7, No. 3, 1975

3. Jack, K. H., Nitrogen Ceramics, 17th Mellor Memorial Lecture, 1973.

4. Guha, J. P., Goursat, P., and Billy, M., Hot Pressing and Oxidatition Behavior of Silicon Nitride with Ceria Additive, Journal of the American Ceramic Society, Vol. 63, No. 1-2, 1980.

5. Jennings, H. M., Danforth, S. C., and Richman, M. H., On a Growth Mechanism for B phase Silicon Nitride, Journal of Materials Science Vol. 14, 1979

6. Dervisbeqovic, H., and Riley, F. L., The Influence of Iron and Hydrogen in the Nitridation of Silicon, Journal of materials Science, Vol. 14, 1979

7. Lindley, M. W., Elias, D. P., Jones, B. F. and Pitman, K. C., The Influence of Hydrogen in The Nitriding Gas on the Strength, Structure and Composition of Reaction-Sintered Silicon Nitride Journal of Materials Science, No. 14, 1979.

8. Elias, D. P., and Lindley, M. W., Reaction Sintered Silicon Nitride, Part 1, Journal of Materials Science, No. 14, 1979

9. Jones, B. F., and Lindley, M. W., Reaction Sintered Silicon Nitride Part 2, Journal of Materials Science, No. 14, 1979

10. Katz, R. Nathan, High Temperature Structral Ceramics, Science, Vol. 208, 1980.

11. Greskovich, C., and Rosolowski, J.H., Sintering of Covalent Solids, Journal of the American Ceramic Society, Vol. 59, No. 7-8, 1976.

12. Vasilos, T., Cannon, R. M., and Wvench, B. J., Improving the Stress Rupture and Creep of Silicon Nitride, Final Report NASA Contract NAS 3-20088, March, 1979.

13. Richman, M. H., Effect of Processing Parameters on Reaction Bonding of Silicon Nitride, Final Report, NASA Contract NSG 3118, October, 1980.

14. Jack, K. H., Sialcns and Related Nitrogen Ceramics, Journal of Materials Science, No. 11, 1976.

15. Longe, F. F., Silicon Nitride Polyphase Systems; Fabrication, Microstructure, and Properties, International Metals Reviews, No. 1, 1980.

16. Vasilso, T., and Cannon, R. M. Jr., Improving the Toughness of Refractory Compounds, Final Report, NASA Cpntract NAS3-17768, November, 1975.

17 Palm,J. A., and Greskovich C. D., Silicon Nitride for Airborne Turbine Applications, Final Report, Naval Air Systems Command Contract No 0019-77-C-0259 July, 1978.

18. Sikara, P. F. and Yeh, H. C., Consolidation of Silicon Nitride Without Additives, NASA Technical Memorandum NASATM 73693, 1976.

19. Williams, R. M., Linear Thermal Expansim of Hot-Pressed Si_3N_4, The Journal of the American Ceramic Society, Vol. 63, No. 1-2, 1980.

20. Moulson, A. J., Reaction-Bonded Silicon Nitride: Its Formation and Properties, Journal of Materials Science, No. 14, 1979.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

A PRELIMINARY STUDY OF THE PARTICULATE FALLOUT
FROM THE SPACE SHUTTLE (STS-1)

Prepared By: G. Martin Hudson, Ph.D

Academic Rank: Assistant Professor

University and Department: University of Central Florida
                                    Department of Physics

NASA/MSFC:
    Division: Atmospheric Sciences
    Branch: Environmental Applications

MSFC Counterpart: Dr. J. B. Stephens

DATE: 14 August 1981

Contract No.:   NGT 01-008-021
                                    The University of Alabama
                                       in H sville

A PRELIMINARY STUDY OF THE PARTICULATE FALLOUT
FROM THE SPACE SHUTTLE (STS-1)

BY

G. MARTIN HUDSON
ASSISTANT PROFESSOR OF PHYSICS
UNIVERSITY OF CENTRAL FLORIDA
ORLANDO, FLORIDA

ABSTRACT

The objective of this investigation in retrospect was to be able to predict acid fallout and the dry deposition of Al_2O_3. The models available were not appropriate and the data available was negligible. Thus, a bimodal particle distribution was assumed normalized to the few existing data points and used as a foundation for a crude zeroth order approximation for the acid fallout. In addition, a settling spectrum for the Al_2O_3 particles was devised as a table look-up since the graphs in the literature at first pass could not be fitted with reasonable analytic functions. Consulting services were rendered to researchers both here and at Kennedy Space Center. Special emphasis was placed on improving current techniques and adding LIDAR (Light Radar).

I. INTRODUCTION

As an applied nuclear physicist working mainly in applications to air particulate pollution, I was awarded a NASA contract to do air particulate sampling at Kennedy Space Center, Florida. This contract can be broken down into two separate parts—background sampling and launch related sampling. The launch related samples could be used to validate the particulate portion of NASA's Multilayer Diffusion Model (MDM). I applied for and received a NASA/ASEE summer fellowship to do just that. Finally, on April 12, the Space Shuttle went up. Now everything was ready. I would soon have my data and NASA had its model. Delays in funding part of my contract led to no data, there was large particle fallout from the exhaust cloud and the model was not ready to handle particles.

II. THE PROBLEM

The problem is really to decide what to do next to the model (MDM) so it can handle particles as separate from gases. Several jobs/tasks can be listed:

(1) Calculate settling velocities.
(2) Calculate particle growth.
(3) Measure particle size, number and concentration S\S-2.
(4) Model results and modify code.
(5) Verify against reasonable data. Note the data taken during STS-1 was not good enough (except for mine) even to make a first pass.

III. TOWARDS A SOLUTION

The first step was to study and calculate the settling velocities. The theory is well summarized in Dawbarn, et al. report[2] and a copy of their Table 4 is included as Table 1. My Figure 1 and their Figure 72, can be used to estimate an average time of fall $t\frac{1}{2}$ (given in Table 2). This number was generated using the equation:

$$t\frac{1}{2} = \frac{500\ m}{V_t}$$

where V_t is the terminal or settling velocity and 500 m is an average cloud height. In STS-1, the top of the cloud was at about 1000 m. Now looking at Table 2, and since we are only worried about the first few hours of the cloud, it appears that everything 10 μm and smaller can be treated as a gas and the model will have to be modified for the larger particles. Going back to Dawbarn's source, The Mechanics of Aerosol, by Fuchs[3], and lifting Figure 2, the experimental data is readily available in a log-log plot and is given in Table 3. We cannot use Klyachko's harmless looking formulae in Table 1 because Re, the Reynolds Number, also depends on V_t and so the equation is indeterminate. Simple minded[4] attempts to fit the data with a curve failed with a ln distribution being the closest. The equation:

$$V_t = a + b\ \ln D \quad Eq\ (2)$$

Table 1. Settling Velocities of Particles

Particle Size Range, μm	Flow Regime	Terminal Velocity
0.001 to 0.1	Molecular	$V_T = \dfrac{d\rho g\,\alpha\lambda}{9z}$
0.1 to 1.0	Cunningham	$V_T = \dfrac{d^2\rho g\left(1 + \dfrac{2\beta}{d}\right)}{18z}$
1.0 to 10.0	Stokes	$V_T = \dfrac{d^2\rho g}{18z}$
10.0 to 1,000.0	Klyachko	$V_T^2 = \dfrac{1.3\,d\rho g}{\rho_a\left(\dfrac{6}{Re} + \dfrac{1}{\sqrt[3]{Re}}\right)}$

where d = diameter of particle (cm)

ρ = density of particle (gm/cm^3)

ρ_a = density of air (gm/cm^3)

z = viscosity of air (dyne-sec/cm^2)

λ = mean free path of air molecules (cm)

α = coefficient determined by the degree of accommodation of the gas molecules to the particle surface:

    for diffuse reflection with conservation of gas molecule velocities = 1.09
    for specular reflection = 1.175
    for diffuse reflection with gas molecules accommodating to particle surface temperature = 1.131

β = coefficient determined by the ratio of gas molecules which are reflected diffusely to those reflected specularly:

    for glass spheres β = 0.82 x 10^{-5}
    for oil droplets β = 0.813 x 10^{-5}
    for brass cylinders β = 0.66 x 10^{-5}
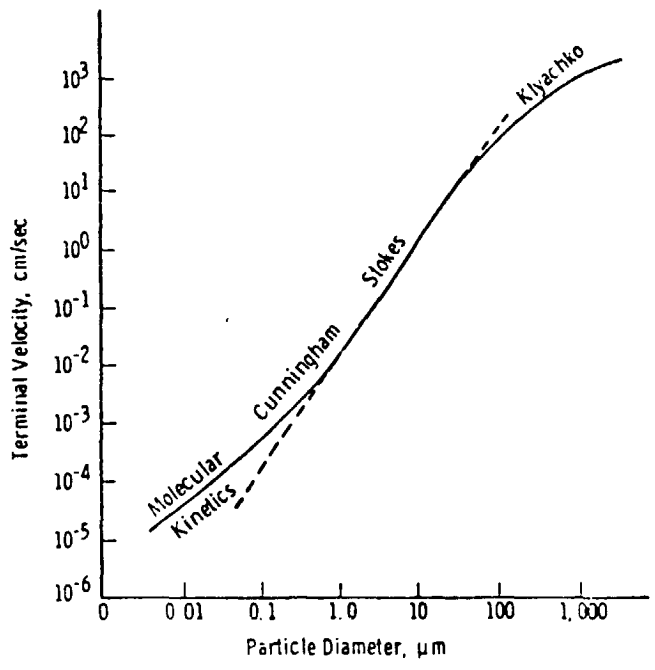
Figure 1. Terminal velocity of Al_2O_3 spheres.
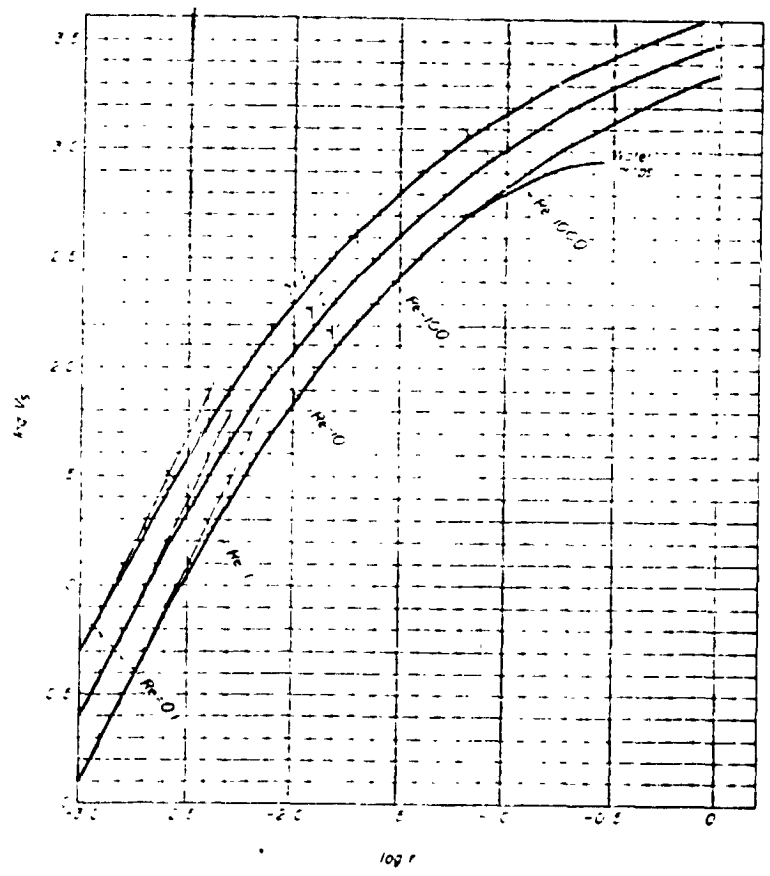
FIG. 2 Rate of settling of aerosol particles in air.

TABLE 2. AVERAGE FALL TIMES

Particle Diameter (μm)	V_t (m/s)	$t\frac{1}{2}$ (sec)	$t\frac{1}{2}$ (Appropriate)
0.1	8 E - 6	6.3 E 7	2 years
1.0	2 E - 4	2.5 E 6	1 month
10.0	2 E - 2	2.5 E 4	7 hours
100.0	1	5.0 E 2	8 minutes
1,000.0	10	5.0 E 1	1 minute

where a and b are least squares fitting parameters and D is the diameter. The parameters were found to be a = 2.29 E 3 and 6 = 4.12 E 2. Unfortunately, r^2 = .75 (regression coefficient) and should be 0.9 or greater for a "good" fit. Thus, Table 3, which is a composite of all the "good" data put together by Fuchs, will have to be used in the table look-up mode in the program (especially since each terminal velocity varies by more than 10% from the next; at least that level of precision should be maintained.) Having got a handle on job (1), we can now turn to particle growth.

Particle growth is reasonably well understood and a good reference is Evaporation and Droplet Growth in Gaseous Media by Fuchs[5]. However, the model begins to operate after cloud stabilization which takes 4-15[6] minutes (about 6 minutes for STS-1), but some of the particles will have fallen out by then. Droplet growth is only well understood for non-turbulent atmospheres, but the cloud before stabilization is turbulent. (Opinion: not enough library research).

According to calculations made by the H. E. Cramer Company, Inc.,[7] the cloud arrived at PAD-39B about 4 minutes (range:2,100 m) and was gone by 8½ minutes. In this time period, 100-200 μm would be reaching the ground. The density of visible particles was 23,000 per square meter.[8] This would imply, assuming the distribution taken from Dawb rn, et al., (see Fig. 3) a peak count of over 250,000/m^2 at 16,000 m. Further data points will be available in the future[8] to better indicate the shape of the agglomerate distribution as seen on STS-1. Thus, I don't feel it is reasonable to estimate the amount of Al and Cl that was lost from the cloud at this time. However, such estimates are needed to run NASA's Multilayer Diffusion Model (MDM) because currently all the matter is assumed transported as small particles and gas.

Due to the lack of data, tasks (3) - (5) cannot be accomplished at this time.

IV. SUMMARY/RECOMMENDATIONS

In order to be concise, I will enumerate this section.

(1) LIDAR should be used to map the location and concentration of the cloud.
(2) The acid fallout must be measured so that the initial concentrations for the MDM can be adjusted.
(3) The acid fallout occurs before, during, and after stabilization and should, therefore, be treated separately from the MDM.
(4) One person should be made responsible for obtaining and analyzing acid fallout samples to include particle size, concentration, and chemistry.

TABLE 3. TERMINAL VELOCITIES OF Al_2O_3 PARTICLES

Diameter (cm)	Velocity (cm/s)	Diameter (cm)	Velocity (cm/s)
9.57 E-4	1.26	1.03 E-2	79.4
1.07 E-3	1.58	1.21 E-2	100
1.21 E-3	2.00	1.45 E-2	126
1.38 E-3	2.51	1.66 E-2	159
1.55 E-3	3.16	2.00 E-2	200
1.78 E-3	3.98	4.91 E-2	251
2.00 E-3	5.01	3.10 E-2	316
2.30 E-3	6.31	3.81 E-2	398
2.58 E-3	7.94	4.80 E-2	501
2.96 E-3	10.0	6.32 E-2	631
3.32 E-3	12.6	8.15 E-2	794
3.81 E-3	15.9	1.05 E-1	1000
4.28 E-3	20.0	1.52 E-1	1260
4.91 E-3	25.1	2.14 E-1	1580
5.64 E-3	31.6	3.24 E-1	2000
6.47 E-3	39.8	5.14 E-1	2510
7.60 E-3	50.1	7.10 E-1	3160
8.73 E-3	63.1	1.63 E-0	3980
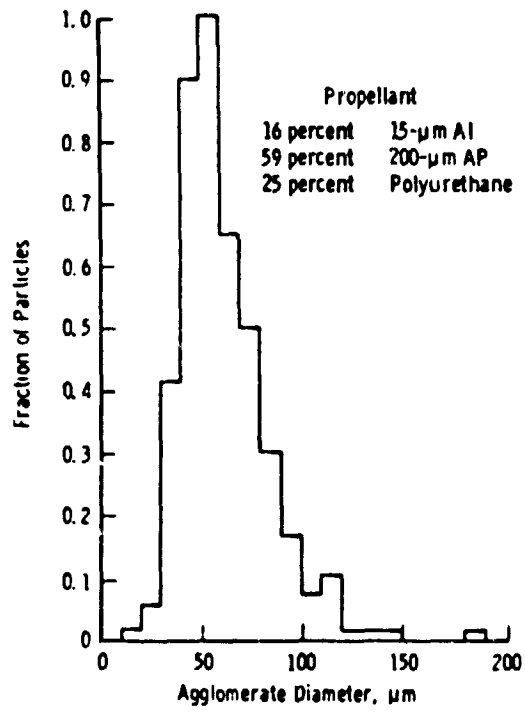
Figure 3.. Histogram of agglomerated Al droplet sizes.

REFERENCES

1. Stephens, J. Briscoe and Roger B. Stewart, "Rocket Exhaust Effluent Modeling for Tropospheric Air Quality and Environmental Assessments," NASA TR R-473, Washington, D.C., June 1977.

2. Dawbarn, R., M. Kinslow and D. J. Watson, "Analysis of the Measured Effects of the Principal Exhaust Effluents From Solid Rocket Motors," NASA Contractor Report, 3136, January 1980.

3. Fuchs, N. A., The Mechanics of Aerosols, The Macmillan Company, New York, 1964

4. HP-33 E Statistics Application, February 1978.

5. Fuchs, N. A., Evaporation and Droplet Growth in Gaseous Media, Pergamon Press, 1959.

6. Stephens, J. Briscoe, NASA/MSFC Private Communications.

7. Dumbauld, R. K., H. E. Cramer Company, Private Communications.

8. Bowie, William J., Technology, Inc., Private Comminications.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

DETECTION AND ANALYSIS
OF RADIO FREQUENCY LIGHTNING EMISSIONS

Prepared by: Fereydoun Jalali, Ph.D.

Academic Rank: Associate Professor

University and   Fort Valley State College
  Department: Electronic Engineering Technology

NASA/MSFC:

  Labor tory: Electronics and Control
  Division: Optical and R.F. Systems

MSFC Counterpart: James Warren Harper

Date: August 1981

Contract No.:   NGT 01-008-021
                                The University of Alabama -
                                Huntsville

IX

ACKNOWLEDGEMENT

DETECTION AND ANALYSIS OF
RADIO FREQUENCY LIGHTNING EMISSIONS

by

Fereydoun Jalali
Associate Professor of Electronic Engineering Technology
Fort Valley State College
Fort Valley, Georgia

ABSTRACT

The feasibility study of detection of lightning discharges
from a geosynchronous s ellite requires adequate ground-
based information regai ing emission characteristics. In
this investigation, a measurement system for collection of
S-band emission data is set up and calibrated, and the
operations procedures for rapid data collection during a
storm activity developed. The system collects emission
data in two modes: a digitized, high-resolution, short
duration record stored in solid-state memory; a continuous
long-duration record on magnetic tape. Representative
lightning flash data are shown. Preliminary results
indicate appreciable RF emissions at 2 gHz from both the
leader and return strokes portions of the cloud-to-ground
discharge with strong peaks associated with the return
strokes.

INTRODUCTION

Interest in the detection and location of the lightning discharges has increased as a result of the need of many agencies that are affected by lightning. The type of information desired covers a wide spectrum. Statistical information is needed in order to estimate the expected direct and induced surges in the communication networks. Very accurate real time location of cloud-to-ground and cloud-to-cloud discharges are among the requirements of the utility, aviation and land management agencies.[4] In the past several years, a few satellite-born sensors have been used to detect lightning from space.[2] However, these have limited capacity and cannot provide the sampling levels and the degrees of spatial and temporal resolutions required by the potential users of lightning information. Ground-based systems capable of providing the needed information would necessitate large numbers of sensors and data collection stations.

Recognizing the usefulness of the satellite-acquired lightning data, NASA has established a research and development program, an objective of which is to develop a satellite-based lightning mapping system that can optimally serve the needs of the users of lightning information. The system is envisaged to be based upon sensing of the electromagnetic radiation produced by lightning in the RF and/or optical regions. Investigation of detection possibilities in the two frequency ranges are being carried out by two groups. This report describes the measurement system developed by the RF group and the preliminary results.

OBJECTIVES

Most of the currently available radio frequency data from lightning discharges are in the VLF to VHF range. At these frequencies, a satellite-based detection system is not practical due to the ionospheric effect on propagation and the required large antennas. To determine the feasibility of a system operating at higher frequencies, more complete emission data in and above the VHF are needed.

The objectives of this investigation, therefore, are collection of the pertinent lightning data and its analysis to determine the characteristics of the emissions at the higher frequencies. To achieve this, a measurement system planned for data collection[3] was completed, installed, and calibrated. Although this process took most of the summer, it was possible to collect a limited number of lightning emission data.

MEASUREMENT SYSTEM AND TECHNIQUE

Measurement Approach

The system developed by the RF group is designed to collect
RF emission data from lightning primarily at 2000 MHz.
Reception at the VHF is also incorporated in the system for
data comparison. For each frequency, vertically- and
horizontally-polarized antennas with associated receivers are
employed and the output of the receivers are recorded on
separate tracks of a magnetic tape to produce a "long"
record of the lightning activity. Due to its bandwidth
limitation (80 KHz at maximum speed of 120 ips), the
recorder effectively filters out the high-frequency
components of the signal.

To obtain a higher-resolution record of the emission, one
receiver output is selected and fed to a high-speed analog-
to-digital converter and the digitized output stored
in a solid-state memory. The digitizer/memory unit can
produce instant display of its memory content on a CRT for
visual inspection. If significant emission data is observed,
it is stored in a long-term memory and the solid-state
memory is cleared and made available for the next emission
event. With the digitizer's maximum sampling rate of
2×10^7 bytes/second and the present capacity of the memory
unit at 128 kilobyte, high resolution records of 6.4 milli-
second duration can be obtained. Comparison of the "long"
record of the magnetic tape with the high-resolution record
of the digitizer unit makes it possible to locate the latter
within the former.

Other channels of the magnetic tape recorder are used to
record IRIG time signal and the "enable" signal from
digitizer marking the initiation of the solid-state memory
fill-up. Additional concurrent parameters such as the
outputs of an acoustic sensor, an optic sensor, and two
electrostatic field sensors are planned to be added for
recording on separate channels of the tape recorder.

Equipment

Figure 1 shows the block diagram of the measurement equipment
which presently are used (see Appendix A for specification of
major equipment). Plans call for addition of optic and acous-
tic sensors for determination of flash distance and two
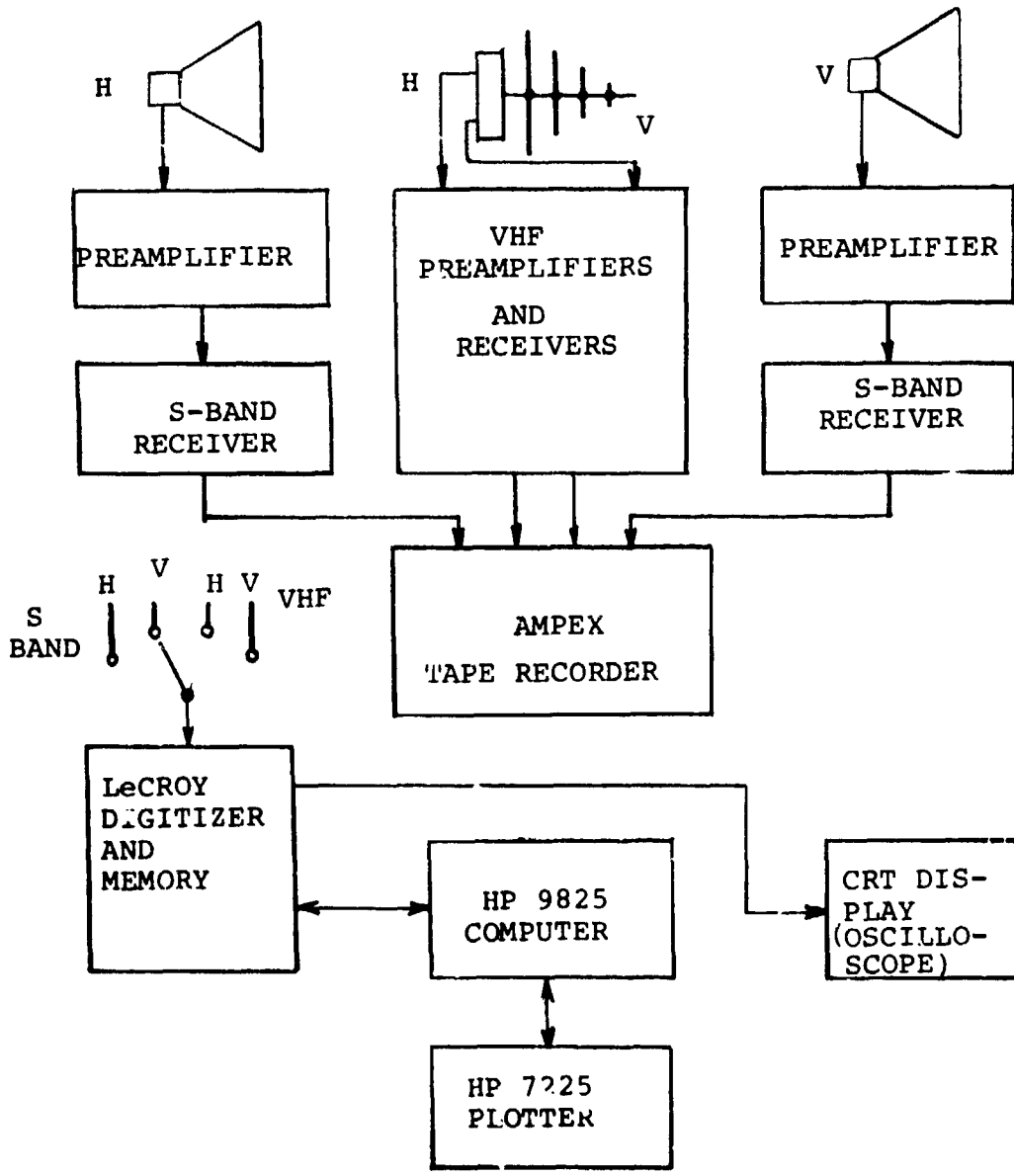
FIGURE 1.   LIGHTNING EMISSION RECORDING SYSTEM

IX-4
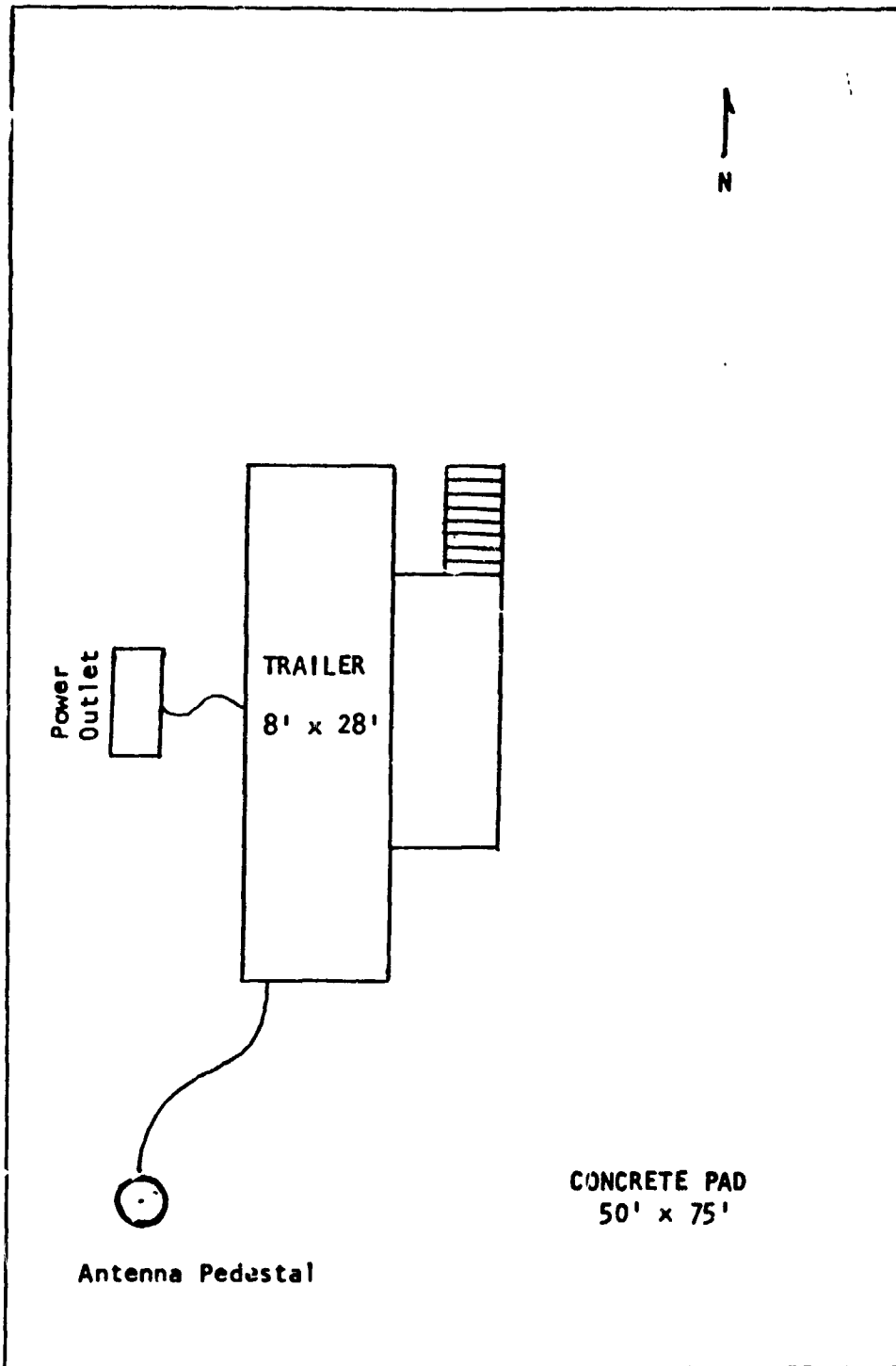
Fig. 2 LIGHTNING EMISSION RECORDING FACILITY

LAYOUT

IX-5

sensors for measurement of electrostatic field. Also, the
HP 9825 computer will be replaced by an HP 1000L with a hard
disc drive, and the HP 7225 plotter with an HP 9872C.

The equipment is housed in an 8 x 28 ft. trailer and the
antennas are mounted on a pedestal. Figure 2 shows the layout
of the facility. In order to reduce the multipath problem
associated with reflections from buildings and other struc-
tures, the trailer is located in an open field in the western
part of the Redstone. Topography of the field makes it
possible to aim the main axis of the S-band horns above the
local skyline with only a 10-15 degree elevation of the
antenna assembly.

Calibration

The system was calibrated so the actual value of the signal
strength at the receiving horn could be determined from the
values of each of the two record types produced. The cali-
bration procedure consisted of transmitting a known amount of
CW power through a standard horn placed 20 ft. from the
receiving horn. The output of the receiver was stored both
on the magnetic tape and in the solid-state memory. The
process was repeated for the vertically-polarized channel
at a range of 40 ft. Table 1 shows the data obtained.
Comparison of the calibration data at 20 ft. and 40 ft.
(see figures 3 and 4) indicates excellent agreement with
inverse-square law for all of the output levels down to
about 160.

The power density and electric field intensity at the
receiving horn may be computed using the relations[4]

$$P_t = P_{in} + Gh \qquad \text{(dBm)}$$

$$P = P_t/(4\pi R^2) \qquad \text{(watts/meter}^2)$$

$$E = (120\pi p)^{\frac{1}{2}} \qquad \text{(volts/meter)}$$

where P_t = the equivalent power transmitted by an isotropic
radiator at the location of the transmitting horn, P_{in} = the
input power to the horn, and Gh = the horn gain .

Figure 5 shows the power density at the receiving horn
versus output levels indicated directly by LeCroy digitizer
and through Ampex recorder.

Data Collection Procedures

The system in its present form requires at least two persons
for data collection during a thunderstorm: one to operate

TABLE 1

CALIBRATION DATA (2 gHz)
RF LIGHTNING EMISSION MEASUREMENT FACILITY

POLARI-ZATION	RANGE FT.	TRANSMITTED POWER dBm	OUTPUT LEVELS LECROY	AMPEX
V	20	-50		251.9
		-55	245.2	237.2
		-60	229.9	221.4
		-65	210.8	202.0
		-66	205.8	197.5
		-70	191.4	185.5
		-75	172.1	167.8
		-80	158.7	155.0
		-85		149.5
		-00	130 0	127.9
	40	-45		247.2
		-50	242.4	233.0
		-55	226.2	216.0
		-60	206.8	198.2
		-65	188.2	181.1
		-70	169.8	163.7
		-75	157.4	151.3
		-80	152.1	147.6
		-00	130.0	128.5
H	20	-50	252.8	
		-55	225.2	
		-60	199.7	
		-65	175.3	
		-70	155.6	
		-75	145.0	
		-80	140.7	
		-00	130.0	

260 ⌐

250 ▪

240 ▪ (b)   (a)

230 ▪

220 ▪

DIGITIZED   210 ▪
OUTPUT
LEVEL   200 ▪

190 ▪

180 ▪

170 ▪

160 ▪

150 ▪

140 ▪

    -40   -50   -60   -70   -80   -90

TRANSMITTED POWER (dBm)
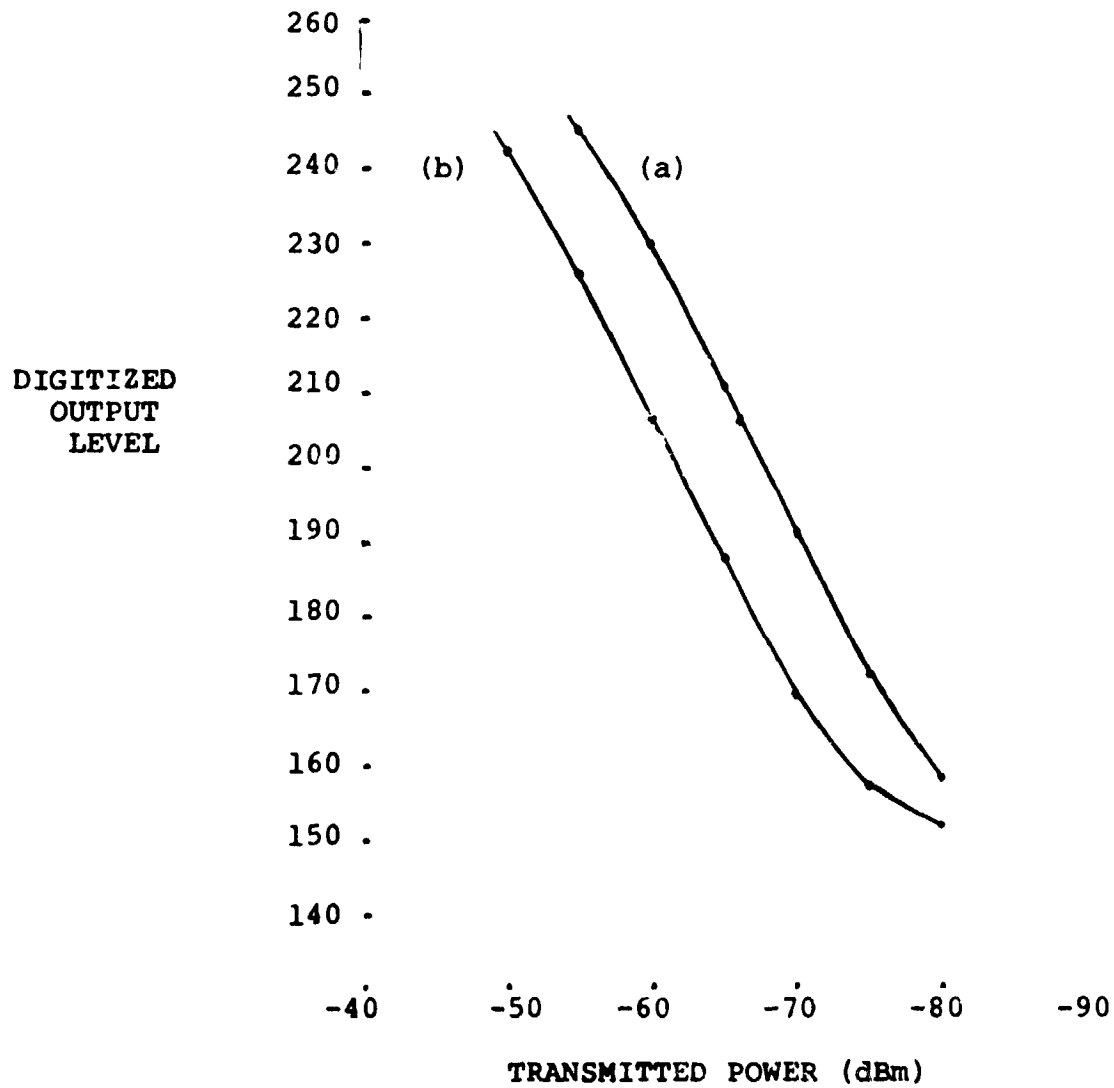
FIGURE 3.   CALIBRATION CURVE FOR LECROY DIGITIZER
             (a) 20 ft RANGE, (b) 40 ft. RANGE
             S-BAND, VERTICAL POLARIZATION

IX-8
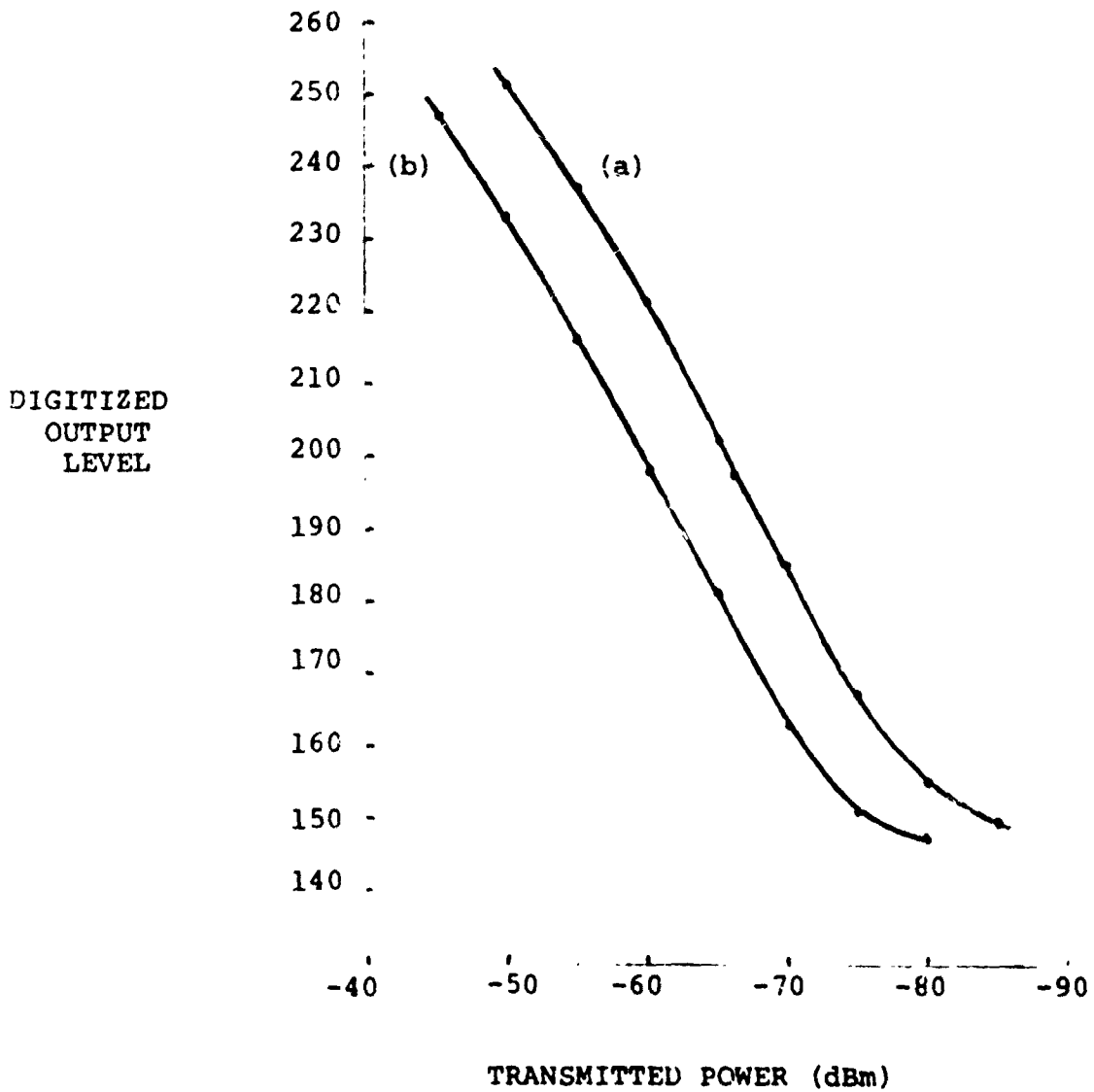
FIGURE 4.   CALIBRATION CURVE FOR AMPEX RECORDER,
             (a) 20 ft. RANGE, (b) 49 ft. RANGE
             S-BAND, VERTICAL POLARIZATION

IX-9
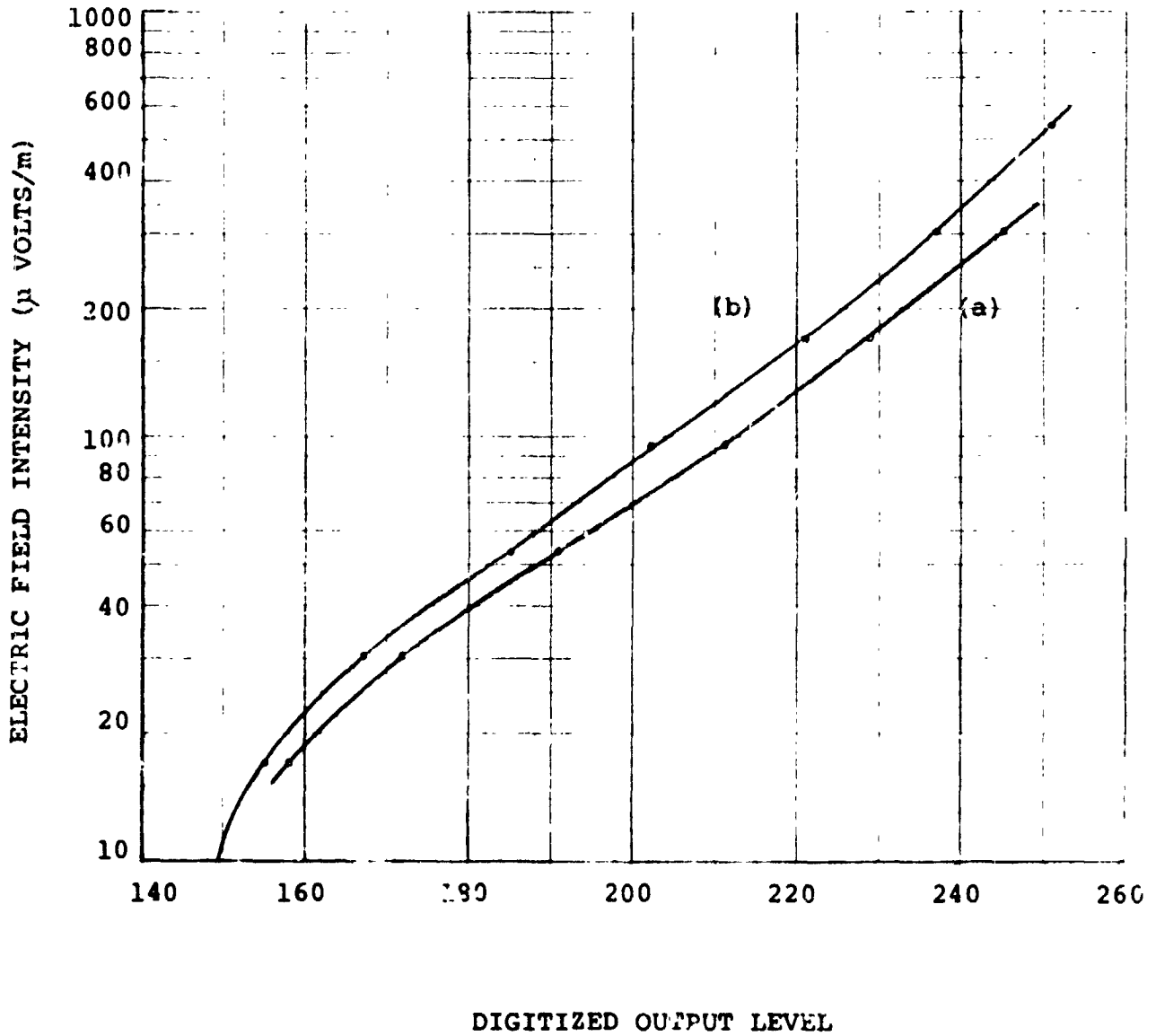
FIGURE 5.   RF LIGHTNING EMISSION MEASUREMENT SYSTEM
            CALIBRATION GRAPH.   (a) LeCroy, (b) Ampex
            S-BAND, VERTICAL POLARIZATION

the instruments and the second to measure the elasped time
between the flash and the thunder, observe the type of
discharge (cloud-to-cloud or cloud-to-ground), and approxi-
mate the relative position of the flash to the anten'
direction. The steps involved in the data collection pro-
cess are as follows:

a. The antenna cluster is aimed at the region of the
storm clouds exhibiting (or likely to produce)
lightning activity.

b. The output of the receiver selected for the high-
resolution record is connected to the digitizer.

c. The digitizer is reset, a sampling rate chosen,
and the trigger threshold is adjusted to ignore
backgr⎯d noise and very weak signals.

d. Recording is started on the Ampex recorder.

e. Upon triggering by a lightning emission, the UTC
time on the IRIG decoder read-out is ⎯ted and the
memory content displayed on the CRT is examined.
If no significant change from noise is present,
the digitizer is reset. When the displayed record
indicates appreciable level changes from noise and
the associated data on flash type and elapsed time
are taken, the content of the memory unit is stored
in the long-term storage (see Appendix B) and the
pertinent information is entered in a log sheet.

f. The digitizer is reset.

EMISSION DATA

A few lightning emission records were obtained after the completion of the measurement facility and its calibration. Most of these were cloud-to-ground discharges at a range of 10 to 30 kilometers. Determination of any general and statistical characteristics of the emissions will have to be made after a larger body of data is collected. However, inspection of the present data leads to a few tentative conclusions.

Figures 6, 7, and 8 show the plot of the Ampex record of three cloud-to-ground discharge data (see Appendix C for plotting procedures). The length of the active region prior to the large spike shown in figure 6 falls within the range of stepped leader duration as indicated by photographic data [5]. The strong peak following the leader region is, therefore, due to the first return stroke. Leader and return strokes are also clearly distinguishable in the graph of figure 7. Intervals between the first and subsequent strokes as measured on these two graphs are within the accepted range. The data shows that the return strokes generate higher levels of radiation at 2 GHz than the leader mechanism. Figure 8 is the record of a twin ground flash. The leader activity is not as pronounced as those of the other two records. However, strong peaks are again associated with the return strokes.
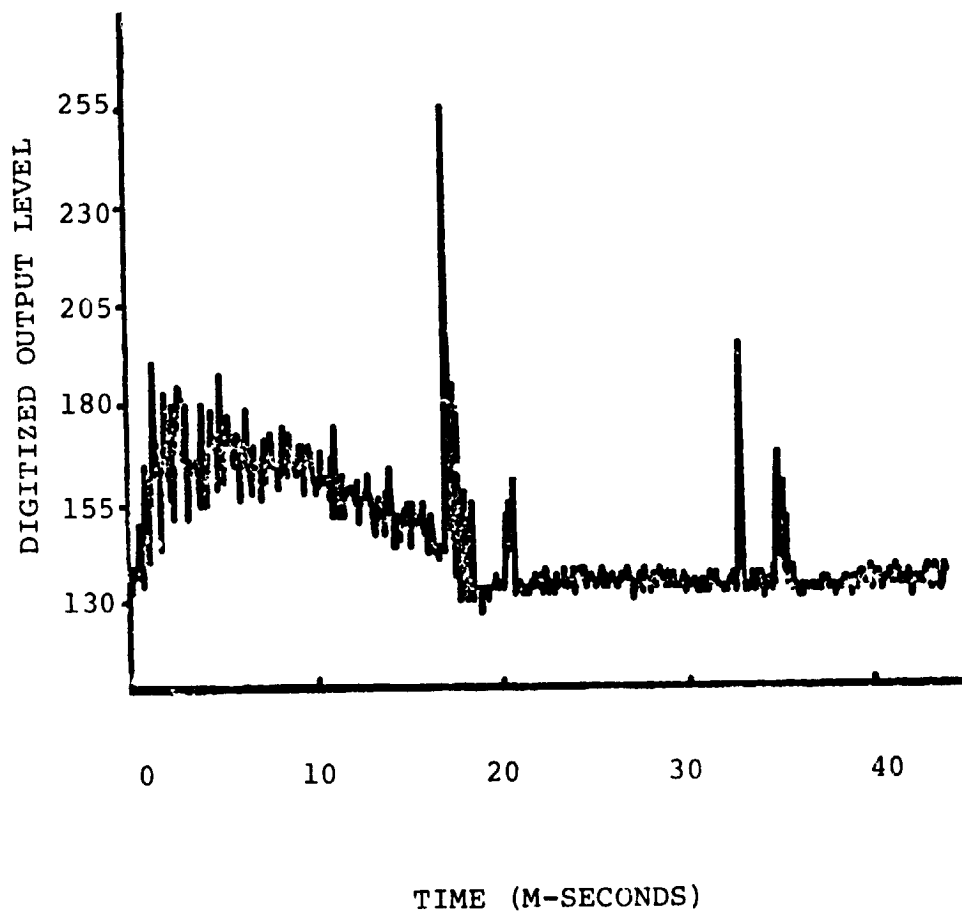
FIGURE 6. CLOUD-TO-GROUND DISCHARGE. RANGE ≅14
KILOMETERS. AMPEX RECORDER
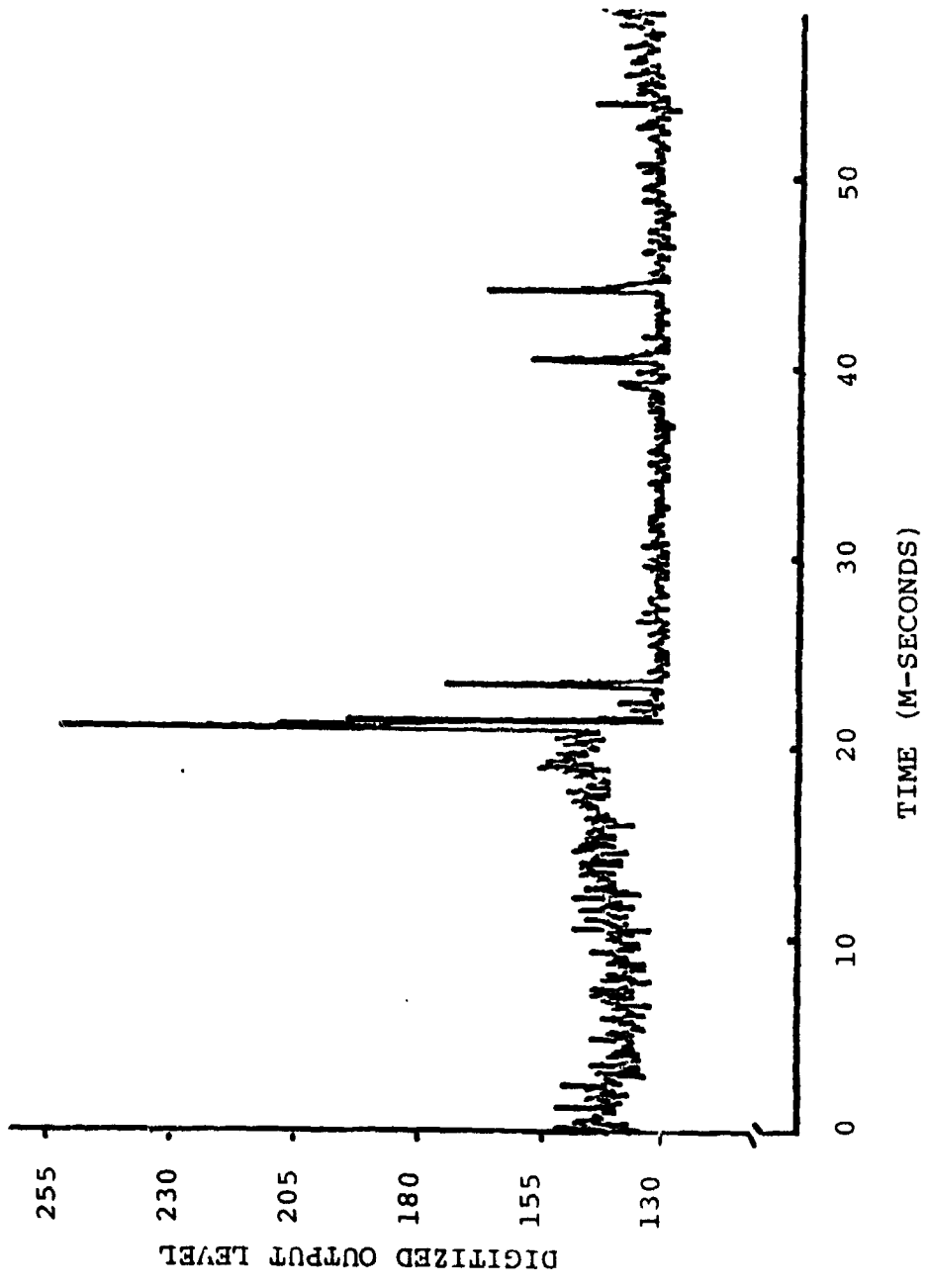
FIGURE 7. CLOUD-TO-GROUND DISCHARGE. RANGE ≅ 15 KILOMETERS.

AMPEX RECORDER
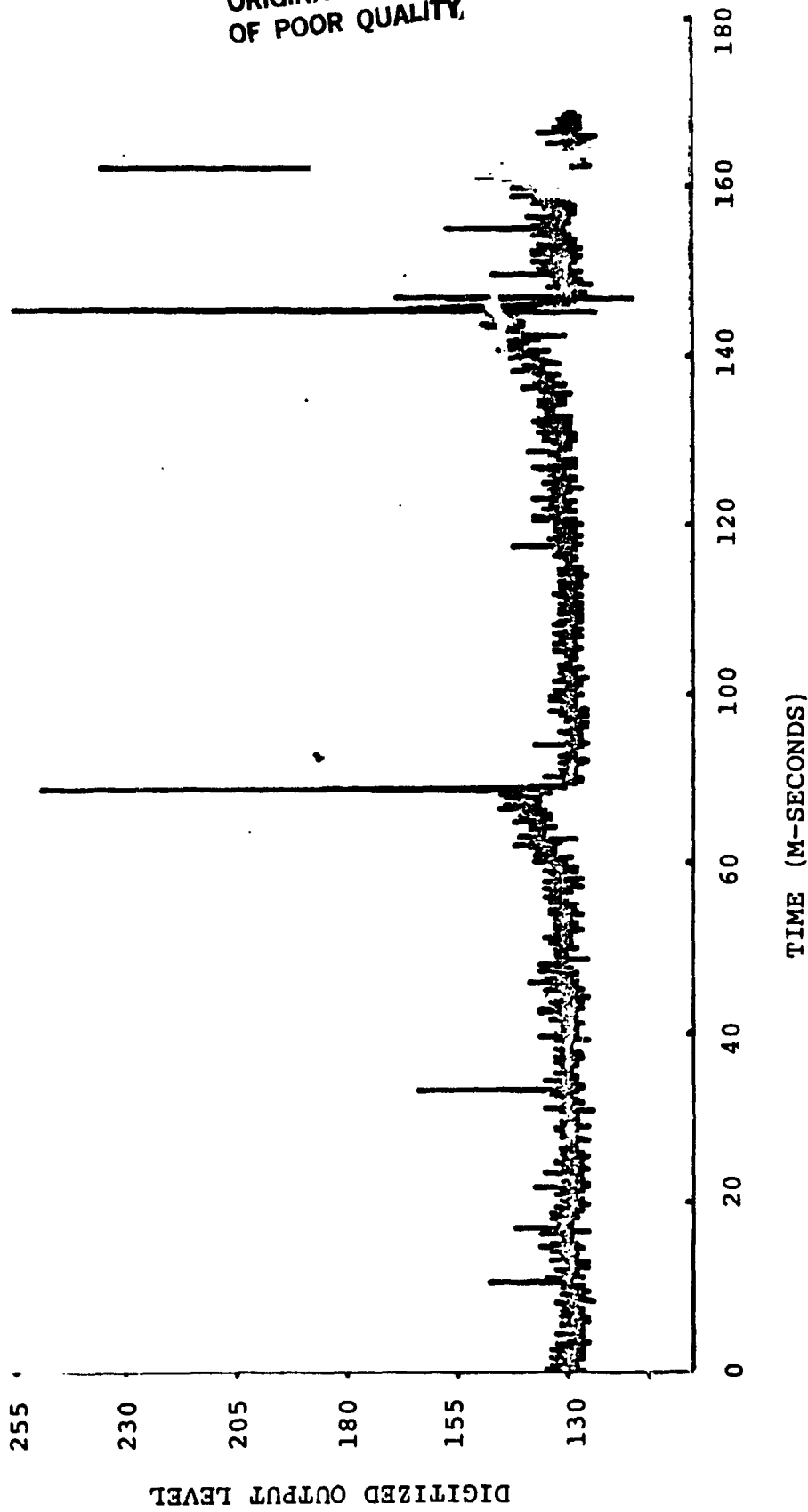
FIGURE 8. CLOUD-TO-GROUND TWIN DISCHARGES. RANGE ≅ 12 KILOMETERS. AMPEX RECORDER.

CONCLUSIONS AND RECOMMENDATIONS

A measurement system for obtaining ground-based emission data from lightning discharges is developed and calibrated. Emissions are received at 2 GHz, and two record modes are produced. A low-resolution, long record which gives information about the levels of various portions of the lightning event; a high resolution short record usable for signature analysis. A set of graphs based on the calibration data is obtained to show the value of the field intensity at the receiving horn for any digitized output level.

A procedure to expedite data collection is devised. Data collection required two persons, an instrument operator and an observer to record the flash characteristic and time. The system operates very well and is excellent for obtaining emission data.

Several emission events were recorded. The number is too small to lend to a general analysis. Tentative results show significant S-band emissions associated with the return strokes. The data indicates cloud-to-ground discharges at a range of 12 to 15 kilometers produce peak field intensities of 0.5 - 1 millivolt/meter at the receiving horn.

The following recommendations are offered:

1. Since the peak levels of some of the emission records are at, or close to, the maximum value of the LeCroy digitizer output, it may be useful to introduce a precision calibrated attenuator at the input of the receiver and obtain some of the data with a lower system gain. This should make it possible to compute the peak intensity of the discharge more accurately.

2. In order to choose any portion of the "long" record for high-sampling digitization, place a circuit in the ENABLE line of the digitizer to delay the ENABLE signal by an adjustable amount. The delay circuit may be designed using two One-Shots and a Flip-Flop.

3. Addition of a remote integrated RESET/ARM switch for the digitizer will save time during data collection when several parameters need to be observed by the operator

APPENDIX A

EQUIPMENT LIST

LIGHTNING EMISSION RECORDING FACILITY

ANTENNAS -

    S-band - Assortment of linearly-polarized horns
    VHF - Log-periodic, 150 mHz to 1 gHz, Antenna Corp. of
    America

PREAMPLIFIERS -

    S-band - Watkins Johnson Minipac CR21-010
    VHF - AV2406-1, Aydin Vector

RECEIVERS -

    S-band - Modified RHG CR305/V83B
    VHF - In-house design

TAPE RECORDER - Ampex FR 2000, 14 channel, 3-3/4|7-1/2|
    15|30|60|120 ips, 80 KHz bandwidth @120 ips

DIGITIZER - LeCroy Model 8800/8, four modules, comprising
    a total capacity of 128 kilobytes

COMPUTER -

    HP 9825

TIME SOURCE -

    Local time code generator (cesium clock) transmitted
    to facility by radio link

PLOTTER -

    HP 7225A

ANTENNA PEDESTAL -

    Scientific Atlanta, Model 3102/J329

APPENDIX B

PROCEDURE FOR STORAGE OF DIGITIZED DATA

Since the HP9825 used in the system has a maximum capacity
of 20 kilobytes, the content of the solid-state memory
(128 kilobytes) cannot be transferred in one step. Further-
more, the entire record of every discharge data may not be
useful. The following program is developed to facilitate
quick transfer of the content of any portion of the solid-
state memory.

```
0: beep;0+S
1: dim A$[15430]
   ,C$[3],D$[2]
2: char(0)&char
   (0)&char(7)+C$
3: wtb 702,char(
   97)
4: wtb 703,char(
   26)&char(0)&
   char(7)
5: red 703,D$;
   dsp num(D$[2])
6: wtb 703,char(
   8)
7: red 703,D$
8: dsp num(D$[2]
   )+0;if 0 jsto
   6
9: buf "IN",A$,3
10: char(2)&char
    (0)&char(7)+C$
11: wtb 703,char
    (106)
12: wtb 703,C$
13: tfr 703,"IN"
    ,15430
14: rds("IN")+Z;
    if 2<0;gto 14
15: beep
16: if S jsto
17: ent L
18: rcf L,A$
    +28?
```

The LeCroy digitizer scans the 128 kilobytes of memory at
a fixed rate and produces the display of the record on the
CRT. The record covers 8.3 centimeter (at 2 millisecond/cm
sweep rate). The program is designed to transfer, in each
run, a block of 15500 bytes correspondingly to slightly
over 1 centimeter of the displayed record. The block may
be stored (on the cassette tape) by entering a file number
or be dumped by transferring the next block.

APPENDIX C

The following program is used for producing the graph of
data contained in the LeCroy memory modules.

```
0: beep:0→S
1: dim A$[2001,6]
  ,t$[3],H$[2]
2: char'26'.char
  '6'.char'7'→C$
3: wtb 703,char'
  97'
4: wtb 703,char'
  26'.char'6'.
  char'7'
5: red 703,0$;
  dsp t$.C$[2];
6: wtb 703,char'
  5'
7: red 703,H$
8: dsp num(0$[2]
  )-0;i+0;1;9to
  6
9: wtb 703,A$;3
10: ttr 703;
  0;char'7'→t$
11: wtb 703,char
  '10'
```

```
12: wtb 703,C$
13: ttr 703,"IN"
  ,20000
14: rds("IN")→Z$
  it Z'0;9to 14
15: beep
16: if S 1;9to
17: ent M
18: ent N
19: (N-M)/500→K
20: lcl 703
21: scl 0,500,0,
  260
22: for I=1 to
  500
23: n+I+K→J
24: plt I,num(A$
  [J],-1;9;9dsp V
25: nex I
26: pen
27: end
*19317
```

The program transfers to the computer consecutive 20-kilo-
byte blocks in each run. If the transferred block is of
interest, integers M and N are entered to designate the
beginning and end of a segment within the block to be
plotted. The program then gives the plot of 500 equally
spaced points of the designated segment.

REFERENCES

1. Bent, Rodney B., "Operational Applications Committee," Proceedings of the 1979 Workshop on the Need for Lightning Observations from Space, NASA CP-2095, Washington, D.C. , 1979.

2. Turman, Bobby N., "A Review of Satellite Lightning Experiments," Proceedings of the 1979 Workshop on the Need for Lightning Observations from Space, NASA CP-2095, Washington, D.C., 1979.

3. Harper, James W., "Experimental Measurement Program in R.F. Lightning Emissions at Marshall Space Flight Center," NASA Interdepartmental Report, 1981.

4. Reference Data for Radio Engineers, Howard W. Sams and Co., Inc., New York, Sixth Edition, 1975.

5. Uman, Martin A. , Lightning, McGraw-Hill Book Company, New York, New York, 1969.

MATHEMATICAL MODELING OF LASER-GAS
INTERACTIONS FOR LASER PROPULSION SYSTEM

Prepared by: Helen H James, Ph.D

Academic Rank: Assistant Professor

University and   University of Alabama in Huntsville
Department: Department of Mathematics

NASA/MSFC
Laboratory: Structures and Propulsion Laboaratory

MSFC Counterpart: Mr. Lee W. Jones

Date: August, 1981

Contract No.:   NGT 01-008-021
                         The University of Alabama in Huntsville

X

MATHEMATICAL MODELING OF LASER-GAS INTERACTIONS FOR LASER PROPULSION SYSTEM

By

Dr. Helen H. James
Department of Mathematics
University of Alabama in Huntsville
Huntsville, Alabama 35899

Abstract

A mathematical model is being developed which describes the absorption of laser energy by a working fluid to produce a very ')t, dence plasma in hydrogen gas. This process takes place in the absorption chamber of a laser rocket engine. Working with investigators at MSFC, the University of Tennessee Space Institute, and Lockheed Huntsville Research & Engineering Center, the numerical techniques are examined that are required to extend the existing one-dimensional laser-supported wave model to include two-dimensional effects, a converging laser beam, and possibly the non-equilibrium effects of photoionization of cold hydrogen gas ahead of the laser-supported combustion wave by vacuum ultra-violet radiation from the hot plasma. Possible simplifications of the full three-dimensional solutions to the Navier-Stokes equations are among the approaches considered, so that the size and complexity of the numerical calculations may be made more amenable to solution using the MSFC 1100/82 computer.

NASA/MSFC Structures and Propulsion Laboratory

Lee W. Jones - NASA/MSFC Counterpart

NASA/ASEE FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

THE UNIVERSITY OF ALABAMA

A STUDY OF THE VENT PIPE FOR THE GRAVITY PROBE-B

PREPARED BY: RONALD N. KNOSHAUG, PH.D.

ACADEMIC RANK: ASSOCIATE PROFESSOR

UNIVERSITY AND   CARROLL COLLEGE, MONTANA
DEPARTMENT: DEPARTMENT OF MATH/PHYSICS

NASA/MSFC:

DIVISION: AEROPHYSICS

BRANCH: UNSTEADY AERODYNAMIC AND
                        THERMAL ENVIRONMENT

MSFC COUNTERPART: JOSEPH L. SIMS

DATE: AUGUST 7, 1981

CONTRACT NO.:   307671
                        THE UNIVERSITY OF ALABAMA
                        IN HUNTSVILLE

A STUDY OF THE VENT PIPE

FOR THE GRAVITY PROBE B

By

Ronald N. Knoshaug
Associate Professor
Department of Mathematics/Physics
Carroll College
Helena, Montana

ABSTRACT

The Gravity Probe B experiment to test Einstein's relatively theory is a possible Space Shuttle project. Cooling of the gyroscopes and thrusting of the Probe into a zero-drag environment is accomplished by vaporizing liquid helium and venting the gas from the Probe through a nozzle system. In order to avoid malfunctioning of this cooling-thruster system, the pressure difference across the vent pipe-thruster system must be maintained to be less than the vapor pressure of liquid helium at the given temperature.

This study investigated helium gas flow through the vent pipe for various configurations, mass flow rates, and pipe diameters. In addition, pertinent properties of both liquid and gaseous helium were investigated.

It was observed that helical coiling of the pipe about the dewar caused less frictional drag and less pressure change than did sharp pipe bending. Also, for various flow rates and nozzle throat area sizes, the system would malfunction for a given bath temperature.

ACKNOWLEDGEMENTS

NOMENCLATURE

a   radius of tube, in.

A   known function of N_R and δ, $= (0.5 N_R)/(1 - 2\delta/3 + \delta^2/6)$

A*   throat area of nozzle, cm^2

B_1 known function of N_R and δ, $= 2N_R/\delta)/(1 - 2\delta/3 + \delta^2/6)$

B_2 known function of N_P. $= 0.9656 \sqrt{N_D} + 1.6500$

c_p specific heat of fluid at constant pressure

C   coiling effect factor, $= \Delta p/\Delta p_s$

C_1 slip-flow coefficient

C_2 slip-flow coefficient

D   diameter of tube, $= 2a$, in.

E   enhancement factor, $= N_{Nu_c}/N_{Nu_s}$

$\bar{E}$ θ-averaged enhancement factor

h   heat transfer coefficient, $cal./(sec\ cm^2\ °C)$

K   thermal conductivity, $cal/(sec\ cm\ °C)$

L   tube length, m

L_v latent heat of vaporization, cal/mole

$\dot{m}$ mass flow rate, g/sec

$\dot{M}$ evaporation rate, g/sec. By continuity equation, $\dot{M} = \dot{m}$

N_D Dean number, $= N_R\ (a/R)^{\frac{1}{2}}$

N_K Knudsen number, $= \lambda/D$

N_{NU} Nusselt number, $= 2\ h\ a/K$

N_R Reynolds number, $= \rho u\ D/\mu$

N_{RC} critical Reynolds number for transition from laminar to turbulent flow

N_p	Prandtl number, $= c_p \mu /K$
P	gas pressure, torr
q	heat flux at wall to the fluid per unit area and unit time
Q_L	heat leak in dewar, mW
R	radius of curvature of the coil, cm
T	temperature, °K
u	mean axial velocity, cm/sec
v_s	sonic velocity of helium, m/sec
z	axial distance, cm

GREEK LETTERS

α	slip-flow parameter
β	entrance parameter which depends upon N_K Range of β: $.15 < \beta < .20$
δ	dimensionless hydrodynamic boundary - layer thickness, $= \dfrac{4.63}{N_D^{\frac{1}{2}}} - \dfrac{0.766}{N_D}$
Δ	finite - difference increment
$\mathcal{L}$	thickness ratio, $= 1/5\,[2 + (^{10}/N_p^2 - 1)]$
η	temperature where He^4 changes from a super fluid to a fluid
σ	polar angle in radians
λ	resistance coefficient
μ	viscosity, μ-poise
ρ	density, g/cm^3
χ	slip-flow factor function of N_K
τ	mean-free paths
Ω	ratio of downstream pressure to reference pressure

SUBSCRIPTS

o	position immediately downstream from ground plug
1	axial reference position within vent-tube
b	helium bath
c	curved pipe
e	entrance
m	mean fluid value
s	straight pipe
w	tube wall

SUPERSCRIPTS

·	average value

I. INTRODUCTION

In 1916, Einstein proposed a theory which helps in the modeling and understanding of our Universe. This theory, which has established the incompleteness of Newton's theory, has been very difficult to test. The Gravity Probe-B is a program which will utilize the Space Shuttle to launch an experiment which measures the very small relativisitic effects due to the Earth's gravitational field. The experiment, first proposed in 1963, was considered as a possible Shuttle project in 1974.

The experiment consists of two gyroscopes, which when orbiting as massive a body as the Earth, undergoes a relativistic precession with respect to the fixed stars. One gyroscope will undergo geodetic precession, which will be about 7 arc sec per year. The other gyro will undergo spin-spin coupling between the gyroscope and the Earth's rotation, which will undergo a motion of about 0.05 arc sec per year.

In order to maintain a controlled environment for the gyroscope, they will be placed in a dewar which is maintained at a low temperature (between 1.6° and 2.0°K). This temperature will be maintained by boiling off liquid helium at the necessary rate. In addition, the helium gas will be vented into 8 thruster nozzles which will be used to produce a zero-drag environment for the probe. To help vent the incoming heat out of the dewar, the vent pipe will be wrapped in some manner about the dewar.

In this paper, the flow of the helium gas from the porous plug containing the helium bath to the plenum is examined. Part II of this paper examines the theoretical considerations in hydrodynamics; part III investigates the thermodynamic considerations; and part IV contains the results and conclusions pertaining to the vent system.

An appendix is also included which contains important properties of helium pertinent to the experiment.

The ultimate goal of this paper is to discuss which configurations of the vent pipe system will allow the experiment to be performed properly.

II. THEORETICAL CONSIDERATIONS IN HYDRODYNAMICS

The tube used to vent the helium during cryogenic cooling of the dewar will be coiled through the dewar insulation and wrapped in some manner about the dewar. Since the pressure gradient is small and the vent pipe long compared to its diameter, three flow regions must be considered: the developing low in the entrance region; the region with continuous flow; and the slip-flow region[1]. In all cases, laminar flow will be discussed.[1]

Analysis made by Hanks [7] on the velocity development in the entrance region of a straight tube showed that the distance Z_e necessary for the centerline velocity to become 0.99 of the fully-developed value to be

$$z_e = 1/4 \ \beta \ N_r \ D = \dot{m} \ \beta \ * \ 10^6/\pi\mu \qquad (2.1)$$

where the parameter β depends upon the Knudsen number and has a range $.15 < \beta < .20$.

When fully-developed straight pipe laminar flow enters a curved pipe, the primary (axial) velocity flow field is accompanied by a secondary flow field which acts in a plane perpendicular to the tube axis and which is symmetrical about the plane of curvature of the tube. (See Figure 2.1)

Center of
curvature

R
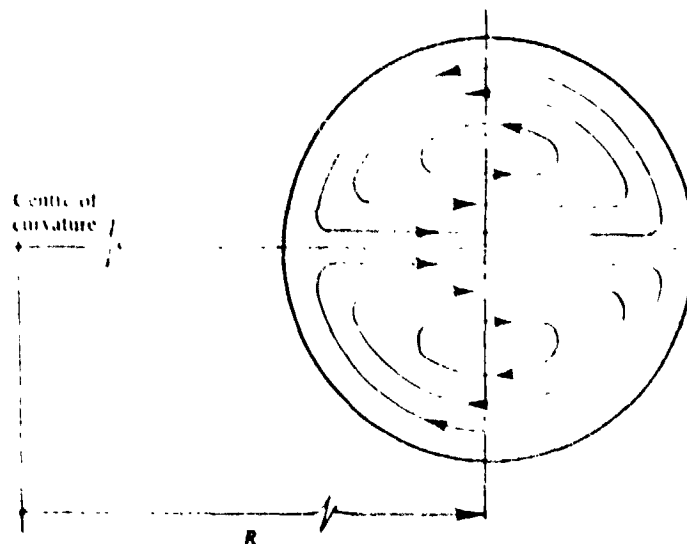
Fig. 2.1.   Flow pattern of the secondary flow of curved pipes.
(Patankar, [14])

---

1.   See Part IV.

Dean [4] showed that the parameter $N_D = N_R (A/R)^{\frac{1}{2}}$ is the unique dynamic similarity parameter governing the fluid motion in a curved tube.

At the onset of the flow entering the curve, the flow is in a transition region until the curved flow eventually becomes fully developed. Squire [23] found that for air in a gradual pipe bend (R/A = 24), the curved flow did not become fully-developed for about 120° from the beginning of the curve. Percival [13] showed that, in the transition region in a straight section following the bend, the pressure loss is greatly affected by the presence of residual secondary flow at the bend exit. Rowe [18] showed that in 180° bend, the amount of secondary flow is greatest at about 30° from the bend entrance; after that the secondary flow is reduced. He also showed that full-development of flow occurs about 90° from the bend entrance. Downstream of the bend, the secondary flow persists for a considerable distance, as high as 29 pipe diameters. Figure 2.2 illustrates the effect of the Dean number on the axial velocity profiles.
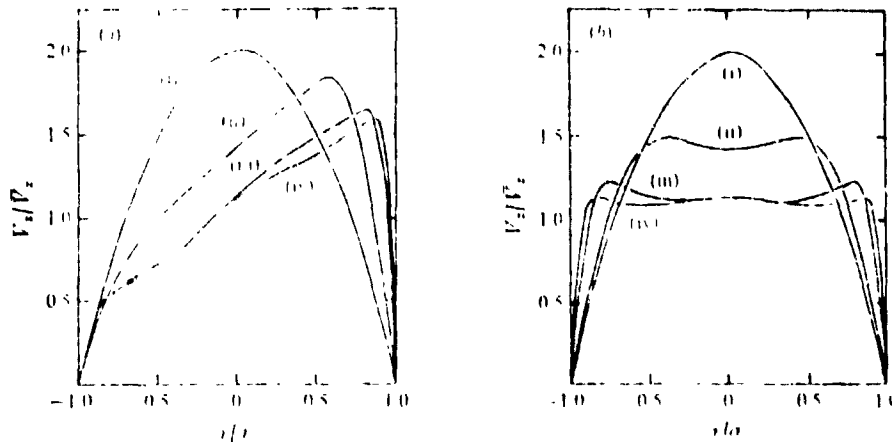
Fig. 2.2. Effect of the Dean number on the axial profiles in (a) the plane along radius of curvature, (b) along diameter plane perpendicular to radius of curvature, (i) straight tube, (ii) N_D = 60, (iii) N_D = 500.

For fully-developed curved pipe flow, many experiments have shown that there is an increase in the critical Reynolds number for transition from laminar to turbulent flow. The measured results are approximated by the empirical equation $N_{RC} = 2 * (\frac{a}{r})^{0.32} * 10^4$, (2.2) valid in the range 15 < R/a < 860. For R/a 860, the critical Reynolds number for a curved pipe approximates that of a straight pipe. (ITO, [8])

The resistance coefficient ratio of a curved pipe, defined by

$$\lambda_c = \left(\frac{\Delta Pc}{L}\right) \frac{D}{1/2 \, \rho \, u^2} \quad ,$$ (2.3)

to that of a straight pipe,

$$\lambda_s = 64/N_R \quad ,$$ (2.4)

is given by

$$\frac{\lambda_c}{\lambda_s} = 0.1033 \, N_D^{\frac{1}{2}} \left[\left(1 + \frac{1.729}{N_D}\right)^{\frac{1}{2}} - \frac{1.315}{(N_D)^{\frac{1}{2}}}\right]^{-3} \quad .$$ (2.5)

This equation agrees closely with experimental results for $N_D > 30$. (ITO [8]). Many other researchers have obtained formulas for resistance coefficient ratios (ITO [8]; Mori and Nakayama, [12]). For N_D small. ($N_D = 0$ for straight pipes), $\lambda_c/\lambda_s \approx 1$.

Sharp bends in the tube causes a fairly large increase in the resistance coefficient as compared to the straight pipe. For large helical coils, the ratio λ_c/λ_s remains close to unity. The flow in the helical coils reaches full-development by 120° from the entrance. Once full-development of the velocity field is reached, the resultant combined primary and secondary flow patterns cause a fluid element to have a screw-like motion through the tube. At one instant of time, a fluid element may be traveling near the center of the tube section; after a short axial distance downstream, the same fluid element may be found near the outside wall of the tube. This type of flow causes a higher axial-pressure gradient, a diametral-pressure gradient, and higher mass-transfer ratio per unit axial pressure drop.

For coils of a small radius a/R, White [27] obtained the coiling effect factor C, defined $\frac{1}{C} = \frac{\Delta P_s}{\Delta P_c}$, by experimental results to be

$$C^{-1} = 1 - [1 - (11.6/N_D)^{0.45}]^{(1/0.45)} \quad (N_D > 11.6) \quad .$$ (2.6)

For $N_{De} < 16.0$, Larrain [10] obtained values for C in terms of the Dean number and radius ratio a/R.

An estimate of the pressure difference in a straight tube for continuous flow is obtained by using Poisewille's formula successively over short pipe lengths (each length short enough so that the viscosity and density can be assumed to be constant). Then adding the results over the entire pipe length, the contribution of the pressure difference due to the curved pipe becomes

$$\Delta P_c^2 = C^2 (2.30 * 10^{-3} \frac{\dot{m}L}{D^4} \int_{1.7}^{T} r \, dt)^{(1)} . \qquad (2.7)$$

The pressure charge in the orifice is

$$\Delta P_o = \frac{.789 \, \dot{m} \sqrt{T_s}}{A^*} \quad (1) \qquad . \qquad (2.8)$$

Then the pressure down stream from the porous plug becomes

$$P = \left[\left(\frac{.781 \, \dot{m} \cdot \sqrt{T}}{A^*} \right)^2 + \Delta p_c^2 \right]^{\frac{1}{2}} . \qquad (2.9)$$

The flow of rarefield gas through tubes with large pressure ratios permit all or some of the flow regimes (continuous, slip, transition, and free molecule) to occur as the degree of rarefaction increases along the length of the tube due to pressure gradient. For Knudson number N_K in the range $.01 < N_K < .1$, the mathematical description of the motion of the fluid by use of Navier-Stokes equation with boundary conditions of zero velocity and equality of temperature between gas and wall becomes inadequate. In this range of Knudson numbers, the gas should be treated as an assemblage of discrete particles. Based on the assumptions that the flow is isothermal and locally fully-developed, the pressure ratio $\omega = P/P_1$ is

$$\omega^2 - 1 + .16 \, C_1 \, (\omega - 1) + .0032 \, C_2 \log \omega + 2 \, N_R^2 \, \alpha \, \chi \, \{.08 \, C_1 \, (\omega$$

$$- 1) + .0008 \, C_2 \, (\omega^2 - 1) - \log \omega\} = -64 \, (N_R) \, \alpha \, \frac{z - z_1}{D} \qquad (2.11)$$

---

1. Gravity Probe-B, [6].   XI-5

with

$$\chi = \frac{1/3 + 4\,C_1\,N_K + 16\,C_1^{\,2}\,N_K^{\,2} + 64\,C_2^{\,2}\,N_K^{\,4} + 64\,C_1\,C_2\,N_K^{\,3}}{(1/2 + 4\,C_1\,N_K + 8\,C_2\,N_K^{\,2})^2}$$

and

$$\alpha = \frac{\mu^2 R\,T}{P^2\,D^2} \quad . \qquad \text{(Sreekanth [24])}$$

Experimental values show that the best value for the slip coefficients are $C_1 = 1.1466$ and $C_2 = 0.14$. The solution to equation (2.11) shows that the gas compressibility significantly affects the pressure drop through changes in both momentum and viscous shear. As the pressure level decreases, the contribution of momentum change plays a minor role and the pressure drop is affected primarily through an increase in viscous shear rather than through an increase in momentum flux.

III. THEORETICAL CONSIDERATIONS IN THERMODYNAMICS

Very near the tube inlet the thermal boundary layer is very thin, and circumferential conduction and both circumferential and radial convection can be neglected compared with radial conduction. In the entrance region, the ratio of the local heat transfer coefficient in a coiled tube to that in a straight tube, defined as the enhancement factor E, is given by

$$E = \left[\frac{1}{2N_{R^-}} \left\{ \frac{2}{8} \left(A + \frac{B_1}{B_2} \sin \theta \right) - \frac{B_1}{B_2} \sin \theta \right\} \right]^{1/3} \qquad (3.1)$$

(Dravid, [5]). For values obtained by Mori and Nakayama [12], $\bar{E}$ (averaged over θ) is equal to $\bar{E} = 0.447 \, N_{D4}^{1/6}$. At $N_D = 125$, $\bar{E} = 1$; at $N_{De} = 1000$, $\bar{E} = 1.41$. Hence, in the entrance region, which is fairly short in length, the secondary flow does not improve heat transfer coefficients significantly because the temperature field has not sufficiently penetrated the secondary field.

Within the developing temperature field, the axial temperature or heat flux profiles show large amplitude oscillations which decay and damp out in the fully-developed region (Seban and McLaughlin [20], Dravid [5]).

Once the thermal region is fully-developed, the secondary flow may increase the heat transfer ratios. Mori and Nakayama [12] have shown that for high Dean number, the heat transfer coefficients are higher in a coiled tube than in a straight tube by a factor which varies as $\sqrt{N_D}$. The enhancement factor in fully-developed flow is the ratio N_{NU_c}/N_{NU_s}. Since the Nusselt number for a straight pipe is 48/11, then

$$N_{NU_c}/N_{NU_s} = \frac{11}{8} \cdot \frac{2 \, aq}{K[Tw-Tm]} \qquad (3.2)$$

For $Np \leq 1$,

$$E = \frac{N_{NU_c}}{N_{NU_s}} = \frac{0.1979 \, N_D^{1/2}}{4 - 37.05 \left[\frac{\ell^2}{12} + \frac{1}{24} - \frac{1}{1203} - \left(\frac{4J}{3} - \frac{1}{33} + \frac{1}{15J^2} \right) \frac{1}{20} \, Np_1 \right] N_D^{-1/2}} \qquad (3.3)$$

XI-7

where

$$\mathcal{L} = \frac{1}{5} \left[2 + \left(\frac{10}{Np_r^2} - 1 \right)^{1/2} \right] \quad . \quad \text{(Mori, [12])}$$

and

$$N_D > 15 \quad .$$

It should also be observed that N_{NU_c}/N_{NU_s} varies with the polar angle θ, and increases with increasing Deam number. (Patankar [14]).

IV. RESULTS AND CONCLUSIONS

It has been estimated that the total heat-leak into the dewar to be at least 70 mW. Later estimates indicate this value could be as high as 120 mW. In order to fulfill the requirement of a drag-free environment, the mass flow rate may be much higher. In light of this, the analysis uses as a minimum flow rate that is required to utilize 70 mW of energy and as a maximum, the rate is taken as six times minimum value.

Since either He^3 or He^4 is a cryogenic agent, both are considered. The following properties are examined to decide which helium is preferred.

Temperature of bath: $1.6° K < T_b < 1.8° K$ [1]

Latent heat of vaporization at 1.6° K [2]

He^3 ... $L_v = 10.6$ cal/mole,

He^4 ... $L_v = 21.6$ cal/mole;

Evaporation rate: [3]

He^3 ... .0063 g/sec $< \dot{M} <$ .0378 g/sec,

He^4 ... .0031 g/sec $< \dot{M} <$ .0186 g/sec;

Dewar loading requirements: [4]

He^3 ... 199 Kg $<$ Load $<$ 1192 Kg,

He^4 ... 98 Kg $<$ Load $<$ 587 Kg;

Because of the high load requirement using He^3, it is eliminated from further consideration.

---

1. Gravity Probe-B [6].
2. See Appendix.
3. $\dot{M} = 9.5622 * 10^{-4} Q_L / L_v$.
4. Based upon the probe being required to function for one year.

The helium vapor as it exhausts from the dewar through the vent pipe is assumed to increase in temperature from 1.7° K to 220° K.[1] Without additional information, the temperature increase is expected to be linear:

$$T = .0809 \, z + 1.7 \quad . \tag{4.1}$$

The vapor is expected to have a mean velocity within the range 0 cm/sec $< u < 87860$ cm/sec.[2]

Due to increasing temperature of the helium vapor, by use of equation (4.1), the viscosity coefficient μ [3] varies with the axial length z, as

$$\mu = \begin{cases} .2244 \, Z + 4.588 \, \mu^P & Z < 153 \text{ cm} \\ (.0809 \, Z + 1.7)^{.647} & Z > 153 \text{ cm} \end{cases} \tag{4.2}$$

The Reynold's number N_R [4] for the flow considered has a range $9 < N_R < 2709$. Based on the scale due to Rotty [17] for straight pipe: $N_R < 2000$, laminar flow; $2000 < N_R < 3500$, laminar or turbulent flow depending on well roughness; and $N_R < 3500$, turbulent flow; and the fact that for curved pipe the critical Reynold's number is in general higher, only laminar flow is considered.

To determine the flow regime, based on the Knudson number N_K [5] the following scale is used: $N_K < .01$, continuous flow; $0.01 < N_K < .1$, slip flow; $.1 < N_K < 3.0$, transition flow; $N_K < 3.0$, free molecule flow. (Schaaf, [19]). The flow regimes for different range of parameters are presented in table 4.1.

---

1. Gravity Probe - B [6]
2. Sonic speed is attained as the upper limit of the velocity in the vent tube. The sonic speed of helium varies considerably with temperature but very little with pressure. (Cook, [2]). By linear regression, the sonic speed of helium gas is $v_s = 1.641 \, T + 617.56$ with correlation coefficient $= .998$.
3. See Appendix.
4. $N_R = \rho \, Du/\mu = 5.01 \, 2 \, 8 \, \dot{m}/D\mu$
5. $N_K = \lambda/D = 1.5546 \times 10^{-9} \, uD/\dot{m}.$

Table 4.1. Knudsen numbers and corresponding flow regimes for different ranges of parameters $N_K = 1.5546 * 10^{-9} \; uD/\dot{m}$.

D (in.)	$\dot{m}$ (g/sec)	M = U/87860	N_K	Flow Regime
3/4	.0031	M < .30	N_K < .01	Continuous
		.30 < M < 1	.01 < N_K < .033	Slip-flow
	.0186	M < .72	N_K < .01	Continuous
		.72 < M < 1	.01 < N_K < .014	Slip-flow
1	.0031	M < .09	N_K < .01	Continuous
		.09 < M < .89	.01 < N_K < .1	Slip-flow
		.89 < M < 1	.1 < N_K < .11	Transition
	.0186	M < .54	N_K < .01	Continuous
		.54 < M < 1	.01 < N_K < .019	Slip-flow

The vent pipe configuration was considered in two extreme cases: one configuration considered was with very sharp bends with the radius of curvature taken as the pipe diameter; the other configuration was that of helical coils with the radius of curvature taken as the radius of the dewar, R = 65 cm. In either case, the Dean number N_D [1] decreases with increasing z since the viscosity increases with increasing axial length z. Table 4.2 shows the ranges of Dean numbers for different radius of curvatures and different parameter values.

By use of equation 2.1 and 4.1, the hydrodynamic entrance region is determined by

$$Z_e = (104.53 + 28.37 * 10^4 \; \dot{m})^{\frac{1}{2}} - 10.22 \quad . \tag{4.3}$$

For the mass-flow rates considered, the maximum length of the developing velocity field would be Z_e = 63 cm = .02 L, which is considered an insignificant part of the total flow.

---

1. $N_D = N_R (a/R)^{\frac{1}{2}} = 5.649 * 10^5 \; \dot{m}/\mu(RD)^{\frac{1}{2}}$

Table 4.2. Dean's numbers for different radius of curvature and different parameters. $N_D = 5.649 * 10^5 \, \dot{m}/\mu \, (RD)^{\frac{1}{2}}$.

D (in.)	m (g/sec)	R (cm)	N_D
3/4	.0031	1.9	$8.9 < N_D < 318.5$
		6.5	$1.5 < N_D < 54.6$
	.0186	1.9	$53.4 < N_D < 1911$
		6.5	$9.0 < N_D < 360.4$
1	.0031	2.5	$6.8 < N_D < 243.3$
		6.5	$1.3 < N_D < 47.3$
	.0186	2.5	$40.7 < N_D < 1459.8$
		6.5	$7.9 < N_D < 284.1$

The frictional drag in curved pipes as compared to straight pipes is greater and depends on the Dean number, as indicated by the ratio λ_c/λ_s in equation (2.5). Using the worst case analysis with D = 3/4 in/ and $\dot{m}$ = 0.0186 g/sec, λ_c/λ_s = 4.9 for R = 1.9 cm and λ_c/λ_s = 2.4 for R = 65 cm. To reduce the frictional drag due to pipe curvature, the larger radius of curvature is preferred.

The effect of curvature in helical coiled pipes on the pressure difference in curved pipes as compared to straight pipes is indicated by $\Delta P_c = C \, \Delta P_s$, where C is determined by equation (2.6) for $N_D > 11.6$ or by coiling-effect factor tables. (Larrain, [10]). Table 4.3 utilizes equations (2.7), (2.8), and (2.9) to give the pressure P immediately below the porous plug due to back pressure in the vent tube and orifice.

The vapor pressure of He^4 at 1.6°K is 5.69 torr (see Appendix). Table 4.3 shows that for many orifice areas and mass flow rates, this value is exceeded. If the bath temperature is taken at 1.8°K, then the vapor pressure of He^4 is 12.47 torr. The pressure P_o for high mass flow rates will be less than this critical pressure for larger orifice areas.

Table 4.3. Pressure P_o below porous plug is obtained for various configurations. The pipe length is taken as 27 m and temperature immediately below porous plug is 1.7°K. The viscosity is averaged over the temperature [1] to give the average viscosity $\bar{\mu} = 100.8\ \mu P$ [1].

D (in.)	A (cm^2)	m (g/sec)	N_{Dp}	C	ΔP_c (Torr)	ΔP_o (Torr)	P_o (Torr)
3/4	.02875	.0031	2.5	1.00011	3.66	1.26	3.81
		.0093	7.5	1.00206	6.35	3.79	7.40
		.0186	15.0	1.01946	9.14	7.57	9.89
	.00875	.0031	2.5	1.00011	3.66	4.15	5.53
		.0093	7.5	1.00206	6.35	12.44	13.97
		.0186	15.0	1.01946	9.14	24.88	26.51
1	.02875	.0031	2.2	1.00011	2.06	1.26	2.41
		.0093	6.6	1.00206	3.57	3.79	5.21
		.0186	13.2	1.01946	5.11	7.57	9.13
	.00875	.0031	2.2	1.00011	2.06	4.15	4.63
		.0093	6.6	1.00206	3.57	12.44	12.94
		.0186	13.2	1.01946	5.11	24.88	25.40

For low mass flow rates, the flow transcends from continuous to slip-flow fairly quickly. Since back pressure with low flow rates does not appear to be a problem, research for slip-flow regime need not be further undertaken.

The heat-transfer rates in fully-developed thermal flow increase due to the secondary flow in curved pipes. The enhancement factor E for coiled pipes is given by equation (3.3). For helium, $N_p = .68$ and $\mathcal{L} = 1.308$. Then equation (3.3) becomes

$$E = \frac{N_D}{6.6094\ N_D^{\frac{1}{2}} - 12.2708} \tag{4.4}$$

1. $\bar{\mu} = \frac{1}{T - 1.7} \int_{1.7}^{T} \mu\ dT.$ XI-13

for N_D > 15. Table 4.4 gives the enhancement factors for various flow rates using helical coiled pipe with radius of curvature R = 65 m.

Table 4.4. Enhancement factors for various Dean numbers.

D (in.)	m (g/sec)	N_D	E
3/4	.0031	$1.5 < N_D < 54.6$	$1 < E < 1.49$
	.0186	$9 \quad < N_D < 360.4$	$1 < E < 3.18$
1	.0031	$1.3 < N_D < 47.3$	$1 < E < 1.43$
	.0186	$7.9 < N_D < 284.1$	$1 < E < 2.87$

APPENDIX

PROPERTIES OF HELIUM

Heat Capacity of Liquid Helium
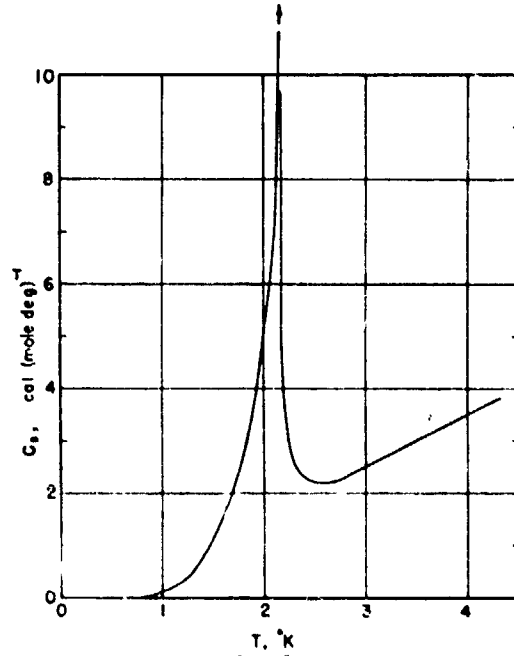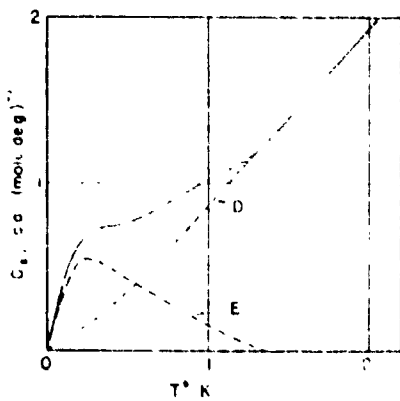
Heat Capacity of LHe^3 (Weinstock, [28])

Heat Capacity of LHe^4 (Van Dijr, [25])

$T_\eta = 2.173°K$

For T close to T_η,

$$C_v = \omega - 3.0 \log_{10} |T - T_\eta| \quad \text{cal/(mole °K)}$$

where

$$\omega = \begin{cases} 3.0 \text{ cal/(mole °K)}, & T < T_\eta \\ -1.6 \text{ cal/(mole °K)}, & T > T_\eta \end{cases}$$
(Atkins, [1])

Latent Heat of Vaporization
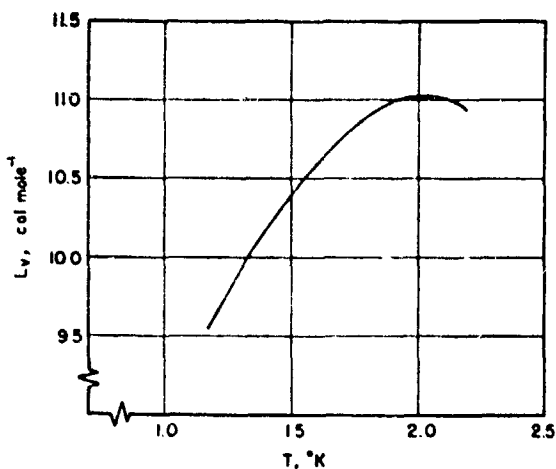
Latent Heat of Vaporization
of LHe^3. (Weinstock, [28])
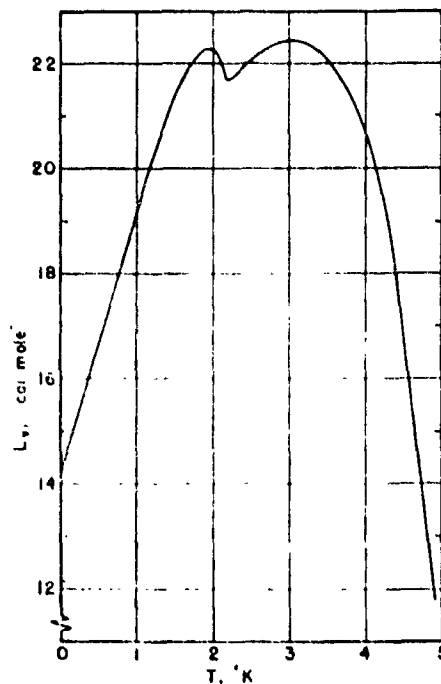
Latent Heat of Vaporization
of LHe^4. (Van Dijk, [25])

Vapor Pressure

T (°K)	He^3 (1) P (Torr)	He^4 (2) P (Torr)	T (°K)	He^3 P (Torr)	He^4 P (Torr)
1.4	38.03	2.15	2.4	288.60	63.30
1.6	64.91	5.69	2.6	381.02	93.73
1.8	101.93	12.47	2.8	491.00	132.95
2.0	150.55	23.77	3.0	619.92	182.07
2.2	212.28	40.47	3.2	769.04	242.27

Vapor pressure of He^3 and He^4.
(1) Roberts, [16].
(2) Van Dijk, [25].

Viscosity of Helium Gas (He^4)

$$\mu = \begin{cases} (1) & 2.7744\ T\ +\ .1498\ \mu P\ , & T\ <\ 14^\circ K \\ (2) & 5.023\ T^{0.647}\ \mu P & ,\ \ \ 14^\circ K\ <\ T\ <\ 1000^\circ K \end{cases}$$

(1) Based upon data from Van Itterbeck [26]. Linear regression was used with correlation factor = .98.

(2) Coremans [3].

REFERENCES

1. Atkins, K. R.: Liquid Helium, Cambridge Press, New York (1959).

2. Cook, G. A.: Argon, Helium, and the Rare Gases, Vol. 1, Inter-science Publishers, New York (1961).

3. Coremans, J. M., et al.: The Viscosity of Gaseous He, Ne, H_2, and D_2 Below 80°K, Physica, Vol. 24, p. 557 (1958).

4. Dean, W. R.: Note on the Motion of Fluid in a Curved Pipe, Phil. Mag., No. 4, p. 208 (1927).

5. Dravid, A. N., et al.: Effect of Secondary Fluid Motion on Laminar Flow Heat Transfer in Helically Coiled Tubes, ALCHE J., Vcl. 17, No. 5, p. 1114 (1971).

6. Gravity Probe-B, Phase-A Report, MSFC-NASA (1980).

7. Hanks, R.: Velocity Profile Development in the Entrance Region of a Right Cylinder with Slip at the Walls, Phy. Fluids, Vol. 6, No. 11, p. 1645 (1963).

8. Ito, H. L.: Friction Fractors for Turbulent Flow in Curved Pipes, J. Basic Eng., Trans. ASME, Vol. 81, p. 123 (1959).

9. Landau, L. D. and E. M. Lifshitz: Fluid Mechanics, Addison-Wesley, Reading, Mass. (1959).

10. Larrain, J. and C. F. Bonilla: Theoretical Analysis of Pressure Drop in the Laminar Flow of Fluid in a Coiled Pipe, Trans. Soc. Rheology, No. 14:2, p. 135 (1970).

11. Lin, C. C.: The Theory of Hydrodynamic Stability, Cambridge Press (1955).

12. Mori, Y. and W. Nakayama: Study on Forced Convective Heat Transfer in Curved Pipes, Int. J. Heat Mass Transfer, Vol. 8, p. 67 (1965).

13. Pervival, P. M. E.: Some Problems in the Rotational Flow of Fluids, Ph.D. Thesis, Cambridge University (1958).

14. Patankar, S. V., et al.: Prediction of Laminar Flow and Heat Transfer in Helically Coiled Pipes, J. Fluid Mech., Vol. 62, p. 539 (1974).

15. Remfry, J.: Heat Transference and Pressure Loss for Air Flowing in Passages of Small Dimensions, Aero. Res. Council, R&M, No. 2638 (1954).

16. Roberts, T. R. and S. G. Sydoriak: Phys. Rev. Vol. 106, p. 175 (1957).

17. Rotty, R. M.: Introduction to Gas Dynamics, John Wiley & Sons, Inc., New York (1962).

18. Rowe, M.: Measurements and Computations of Flow in Pipe Bends, J. Fluid Mech., Vol. 43, p. 771 (1970).

19. Schaaf, S. A. and P. L. Chambre: Fundamentals of Gas Dynamics, ed. by H. Emmons, Princeton University Press, Princeton, N. J. (1958).

20. Seban, R. A. and E. F. Mclaughlin: Heat Transfer in Tube Coils with Laminar and Turbulent Flow, Int. J. Heat Mass Trans., Vol. 6, p. 387 (1963).

21. Shidlovskiy, V. P.: Introduction to Dynamics of Rarefied Gases, American Elsevier, New York (1967).

22. Shidovskiy. V. P.: Special Case of Viscous Gas Motion in Cylindrical Tube in Slip Flow Regime, Sixth Rarefied Gas Dynamics Symposium, Academic Press, New York, p. 215 (1969).

23. Squire, H. B.: Note on Secondary Flow in a Curved Circular Pipe, Aero. Res. Council, N. 16601 (1959).

24. Sreckanth, A. K.: Slip Flow Through Long Circular Tubes, Sixth Rarefied Gas Dynamics Symposium, Academic Press, New York, p. 667 (1969).

25. Van Dijk, H., et al.: Physics 24, S129 (1958).

26. Van Itterbeek, A., et al.: Measurements of the Viscosity of Helium Gas at Liquid Helium Temperatures as a Function of Temperature and Pressure, Physica 19, p. 1158 (1953).

27. White, C. M.: Streamline Flow Through Curved Pipes, Proc. Roy. Soc., A123, p. 645 (1929).

28. Weinstock, B., et al.: Nuovo Cinento Suppl. 9, p. 310 (1958).

29. Zucrow, M. J. and J. D. Hoffman: Gas Dynamics, Vol. 1. John Wiley & Sons. Inc., New York (1976).

NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

THE DETERMINATION OF THE STACKING FAULT
ENERGY IN COPPER-NICKEL ALLOYS

Prepared By:	H. P. Leighly, Jr.
Academic Rank:	Professor
University and Department:	University of Missouri - Rolla Department of Metallurgical Engineering

NASA/MSFC:
    Division:
    Branch:
	
	Metallic Materials
	Failure Analysis

MSFC Counterpart:	John C. McClure
Date:	14 August 1981
Contract No.:	NGT 01-008-021 The University of Alabama in Huntsville

THE DETERMINATION OF THE STACKING FAULT
ENERGY IN COPPER-NICKEL ALLOYS

BY

H. P. Leighly, Jr.
Professor of Metallurgical Engineering
University of Missouri - Rolla
Rolla, Missouri

ABSTRACT

The loss of mechanical properties of high performance alloys
tested under high pressure hydrogen cannot be predicted at the pre-
sen. time. It is obvious that hydrogen introduced into such an alloy
may interfere with dislocation movement, thus leading to a loss in
ductility and a reduction in mechanical properties. Since some high
nickel alloys show embrittlement due to hydrogen, while pure copper
does not, it has been decided to determine the stacking fault energies
of a series of nickel-copper alloys to gain an insight into the em-
brittling effect of hydrogen.

ACKNOWLEDGEMENT

INTRODUCTION

High performance alloys by their very nature often exhibit severe loss of strength and ductility as the result of the environment in which they are expected to operate. This is particularly true in the case of turbine blade alloys which are required to operate at high temperatures in high pressure hydrogen in the pumps which feed the fuel to rocket engines. At the present time, there is a dearth of understanding of the qualities that an alloy must possess which makes it a successful candidate for this type of application.

Since hydrogen may embrittle an alloy, one must be concerned with the mechanical behavior of the alloy and implicitly, with the ease of dislocation motion. In the case of complex stresses, the ability for dislocations to cross slip may be important; hence, the stacking fault energy of an alloy on which cross slip depends will become a quantity which should be known.

OBJECTIVES

McPherson (1) has compiled data on the loss of mechanical strength of 50 metals and alloys caused by high pressure hydrogen. In these measurements, the fracture stress was determined for notched specimens mechanically tested in either high pressure hydrogen or high pressure helium. The ratio of these fracture stresses was determined and are given in Table I.

In some of the alloys, one can infer that the loss of fracture strength is the result of a crystal structure which is expected to be less ductile, i.e., body centered cubic lattice as compared to face centered cubic lattice. However, this is not a firm conclusion inasmuch as some of the face centered cubic alloys perform less well than body centered cubic alloys, as can be seen in Table I.

Odegard, Brooks and West (2) have determined the loss in reduction in area for tensile specimens which were tested in air at one atmosphere or hydrogen at 10,000 psi. The alloys used in this experiment were basically stainless steels of several compositions. For the alloy, nimonic 32, with a stacking fault energy of 12 mJ/m^2, the reduction of area loss was about 55%, whereas for SAE 316 stainless steel with a stacking fault energy of 78 mJ/m^2 the loss was slightly over 1%. The nature of the fracture surface changed sharply when the stacking fault energy exceeded 35 mJ/m^2, going from intergranular fracture to a dimple rupture.

The rationale for this behavior can be understood in terms of hydrogen being transported by dislocations (3, 4). When the stacking fault energy is low, cross slip for dislocations is very difficult; hence, moving dislocations on a slip plane will transport any hydrogen to a small region of the grain boundary as they pile up. This will cause the loss of integrity in the grain boundary. Dislocations in alloys with high stacking fault energy will easily undergo cross slip; hence, any hydrogen in the dislocation core will be scattered throughout the crystal with less damage to the grain boundary integrity. One now has a clue as to the effect of hydrogen on embrittlement.

Table I

RELATIVE RESISTANCE TO HYDROGEN EMBRITTLEMENT
NOTCHED STRENGTH RATIO (H_2/He) FOR VARIOUS ALLOYS IN HYDROGEN
AT ROOM TEMPERATURE

ALLOY	K_t	PRESSURE Ksi	RATIO H_2/He
250 MARAGING	8	10	.12
410	8	10	.22
1042 (Q & T)	8	10	.22
17-7 PH (TH 1050)	8	10	.23
HP 9-4-20	8	10	.24
H-11	8	10	.25
INCONEL X-750	6.3	7	.26
RENE 41	8	10	.27
ED NICKEL	8	10	.31
4140	8	10	.40
INCONEL 718	8	10	.46
MP 35N	6.3	10	.50
440 C	8	10	.50
TI-6 AL-4 V (STA)	8	10	.58
MONEL 400	6.3	7	.65
D 979 STAINLESS	6.3	7	.69
NICKEL 270	8	10	.70
CG 27 STAINLESS	6.3	7	.72
A 515-G70	8	10	.73
HY 100	8	10	.73
A 372-IV	8	10	.74
1042 (NORMALIZED)	8	10	.75
INCONEL 625	8	5	.76
A517-F (T-1)	8	10	.77
A 533-B	8	10	.78
TI-6 AL-4 V (ANN.)	8	10	.79
1020	8	10	.79
HY 80	8	10	.80
INCONEL 706	6.3	7	.80
TI-5 AL-2.5 Sn ELI	8	10	.81
ARMCO IRON	8	10	.86
304	8	10	.87
321	8	5	.87
HASYELLOY X	8	5	.87
305	8	10	.89
ASTROLOY	8	5	.90
347	8	5	.91
HAYNES 188	6.3	7	.92
304 N	6.3	15	.93
310	8	10	.93
Be-Cu (ALLOY 25)	8	10	.93
RA 330	6.3	7	.95
A-286	8	10	.97
21-6-9	6.3	7	.97
7075-T73	8	10	.98
6061-T6	8	10	1.00
OFHC COPPER	8	10	1.00
316	8	10	1.00
INCOLOY 903	8	5	1.00

After W.B. McPherson (1)

STACKING FAULT ENERGY DETERMINATION

There are a number of methods for determining the stacking fault energy in metals. With one exception, they rely on indirect methods for determining the stacking fault energy. For example, the earliest method relies on the determination of the stacking fault probabilities using x-ray diffraction (5, 6). In this method, changes in peak separation between x-ray diffraction lines permit the calculation of the stacking fault probability. From the stacking fault probability, one can calculate the stacking fault energy. Henderson (7) has done this for pure copper, pure nickel and nine copper-nickel alloys. The values he obtained for the pure components were much too high when compared with currently accepted values. A second method relying on electron microscopy in which the stacking fault energy is calculated from measurements on the radius of curvature on dislocation nodes (8). This method is successful for pure metals and alloys in which the stacking fault energy is low. The University of Birmingham group under Smallman (9) has determined stacking fault energies in aluminum and zinc by measuring the kinetics of loop shrinkage for the pure metals and alloys using electron microscopy.

The only direct method for the determination of stacking fault energies utilizes the weak beam dark field electron microscopy (10) in which the distance between the partial dislocations in an extended dislocation are measured directly. Using the appropriate elastic constants, the stacking fault energy can be determined from the distance measurements. This method relies on high resolution electron microscopy. To achieve the necessary high resolution, small diameter objective lens pole pieces must be used which limits the ability to tilt the specimen. To observe the extended dislocation, one needs to have the specimen oriented with the [111] pole approximately parallel with the electron beam. In order to achieve this, the specimen must be cut from a single crystal with the surfaces parallel to the (111) plates.

During this summer, this author familiarized himself with the Phillips Model 300 electron microscope. In addition, methods for the preparation of electron microscopy specimens of pure copper and pure nickel were successfully developed.

CONCLUSIONS

To help understand the effect of hydrogen on the mechanical properties of metals and alloys, the determination of stacking fault energies of any candidate alloy is a necessity. Although at least two papers have been published on the stacking fault energy of copper-nickel alloys, their results are out of date and obtained by using indirect experimental techniques. The measurements need to be redone by weak beam dark field electron microscopy.

RECOMMENDATIONS

The author plans to prepare single crystals (or specimens with very large crystals) of copper, nickel and several copper-nickel alloys during the forth-coming academic year with the view of completing the experiment next summer. The alloy specimens will be analyzed chemically to determine the exact composition. The electron microscope will have been gotten into perfect condition during the winter. It should be possible to perform the same determination on polycrystalline specimens by careful selective cutting of alloys having a strong crystallographic texture.

REFERENCES

1.   This data has been accumulated from many research reports by W. B. McPherson, Metallic Materials Division, Marshall Space Flight Center.

2.   B. C. Odegard, J. A. Brooks, and A. J. West, "Effect of Hydrogen on Behavior of Materials," Eds. A. W. Thompson and I. M. Bernstein, The Metallurgical Society of AIME, New York, 1975, p. 116.

3.   A. Q. Thompson, "Hydrogen in Metals," Eds. I. M. Bernstein and A. W. Thompson, ASM, Metals Park, Ohio, 1974, p. 91.

4.   A. J. West and M. R. Louthan, Jr., Met. Trans, Vol 10A, p. 1675.

5.   M. S. Paterson, JAP Vol 23 (1952) p. 805.

6.   B. E. Warren and E. P. Warekois, Acta Met, Vol 3 (1955) p. 473.

7.   B. Henderson and J. Just, Met, Vol 92 (1963-64) p. 55.

8.   A. Howie and P. R. Swann, Phil Mag, Vol 6 (1961) p. 1215.

9.   J. W. Edington and R. E. Smallman, Phil Mag, Vol 11 (1965) p. 1109.

10.   D. J. H. Cockayne, L. L. Ray and M. J. Whelan, Phil Mag, Vol 20 (1969) p. 1265.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

SYNTHESIS OF AMINOSILANE AND AROMATIC
DIAMINE INTERMEDIATES FOR POLYMERIZATION STUDIES

Prepared By: Samuel P. McManus, Ph.D.

Academic Rank: Professor

University and Department   The University of Alabama in Huntsville
                                        Department of Chemistry

NASA/MSFC:
    Laboratory: Materials and Processes
    Division: Non-metallic Materials
    Branch: Polymers and Composites

MSFC Counterpart: W. J. Patterson, Ph.D.

Date: 7 August 1981

Contract No.:   NGT 01-008-021
                                        The University of Alabama in Huntsville

SYNTHESIS OF AMINOSILANE AND AROMATIC
DIAMINE INTERMEDIATES FOR POLYMERIZATION STUDIES

BY

Samuel P. McManus
Professor of Chemistry
The University of Alabama in Huntsville
Huntsville, Alabama

ABSTRACT

The objective of this study was to prepare monomers for use in a study to improve properties of epoxy polymers and silicone elastomers. Various approaches to the solution of the syntheses were taken and the results are discussed herein.

Project A. Toughened Epoxy Polymers.

Introduction

The typical epoxy polymer used in high performance adhesives is a derivative of bisphenol A (I). This raw material can be converted to a glycidyl ether (II) by

(I)
bis-phenol A

(II)
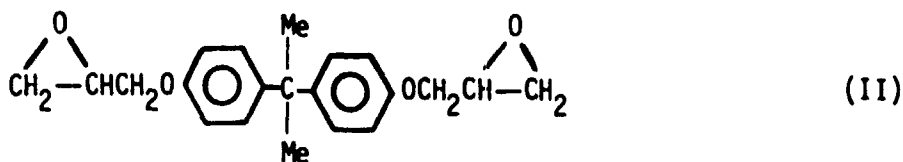
reaction with epichlorohydrin. These two reactants also produce an epoxy-containing prepolymer (III) of variable molecular weight (i.e., n is variable but generally is <7).

(III)
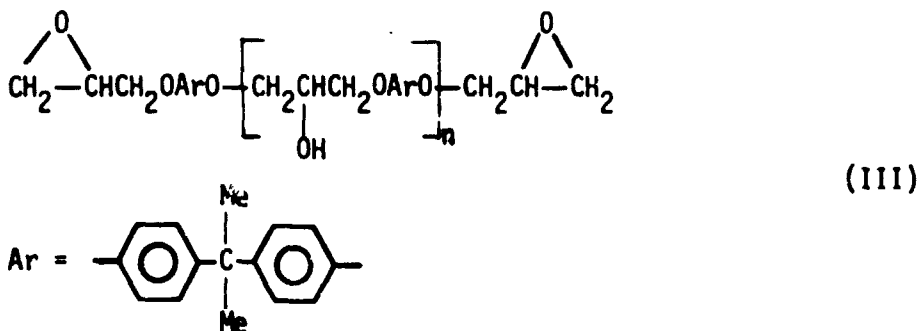
The resin III can be cured into a hard cross-linked polymer with various di- or polyamines, anhydrides or other difunctional nucleophilic compounds. Diamines give products which have the best chemical resistance. Such polymers have applications as coatings, in composite formulation, and as adhesives.[1,2] For high performance adhesives, the diamines most often used as curing agents are m-phenylene diamine (IV), 4,4'-methylene dianiline (V), and the corresponding sulfone VI. While these diamines give cured epoxies with good adhesive strength, toughness, and thermal stability, the polymer tends to be brittle.[3]

IV
V, X = CH$_2$

VI, X = SO$_2$
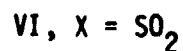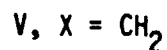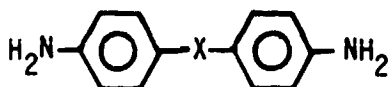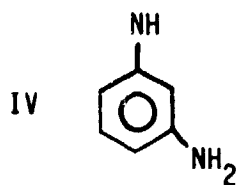
Objectives.

The purpose of this research program was to synthesize new diamine curing agents for epoxy resins with the goal of obtaining final cured polymers which have better mechanical properties.

The direction already planned by Dr. W. J. Patterson for the ongoing in-house program is well precedented. It is well known that a linear arrangement of polar groups increases intermolecular (interchain) interactions and, therefore, stiffness. Figure 1 illustrates a highly regular structure with considerable interchain interactions (dotted lines). Reducing these interactions can be achieved by reducing the regularity of the chains. Figure 2 shows a non-linear polymer with the same number of polar groups, but with fewer, and therefore weaker, interchain interactions. The goal then is to engineer a cured epoxy with fewer intermolecular interactions by reducing the polymer linearity.

Proposed Study.

A series of non-linear diamine curing agents were targeted for test comparisons with standard cure agents. Ideally, it was desired to test a series of di-amines which impart linearity to a cured polymer versus a series which impart non-linearity. The series initially chosen for possible study were ortho-, meta-para-substituted diamines of general structure VII where X is CH$_2$, C(CH$_3$)$_2$, O, S, NH or SO$_2$, and n = 0, 1, 2, and 3. Because of the availability of some key
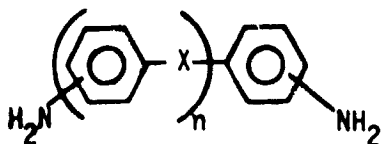
(VII)
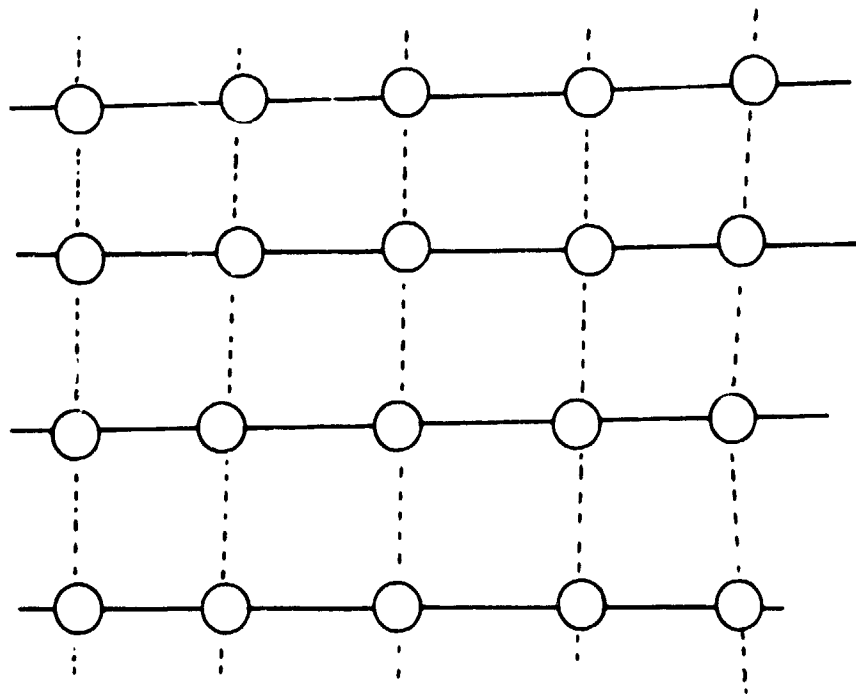
Figure 1. Representation of the interchain (dotted line) attraction between polymer chains which have a linear structure.
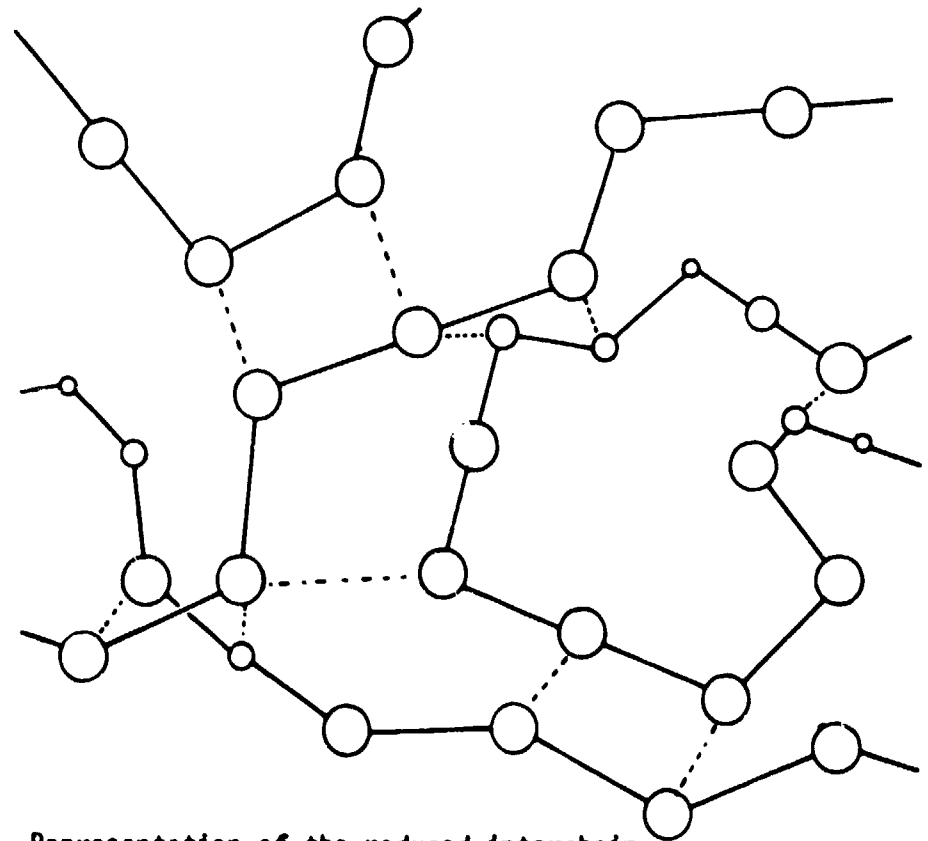
Figure 2. Representation of the reduced interchain attractions in polymers with a nonlinear structure.

members of the series when X = 0 and CH_2, these series were finally chosen for the greatest amount of effort.

Experimental Effort.

In order to get ortho substitution candidates, catechol (VIII) was converted to its sodium salt and reacted with m-chloronitrobenzene (IX).

(VIII)   (IX)
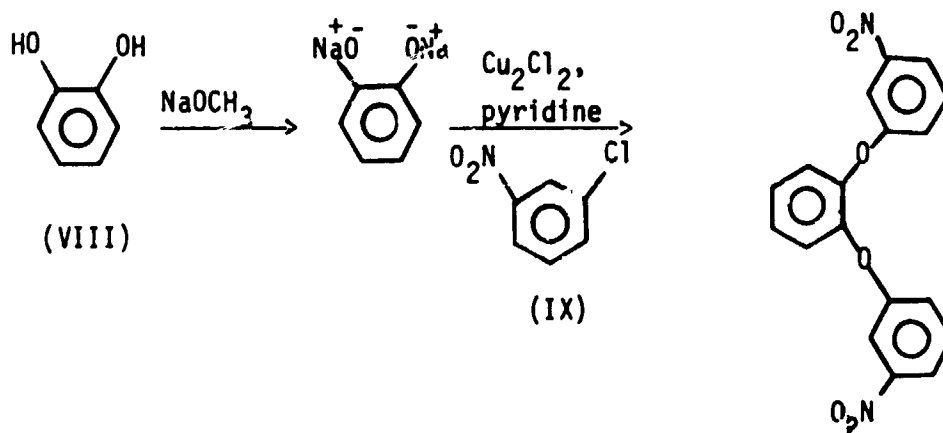
Initially, it was determined that the sodium cathecholate was extremely sensitive to oxygen, even the traces remaining in nitrogen passed over hot copper turnings. Therefore, only reactions in an argon atmosphere were useful for evaluation. Even then multiple reaction pathways were competitive and several products formed (see Figure 3). Therefore, this series was inappropriate for catechol although it works satisfactorily for resorcinol,[4] the diamine product of which was procured.

Another target molecule experimentally sought was VII (X = CH_2, n = 2). Two routes were being pursued. One route involved nitration of commercially-available m-dibenzoylbenzene followed by reduction of the resulting nitro and carbonyl groups. Preliminary results proved this to be a feasible route.[5] A second route involves the Grignard coupling of the blocked amine XI and the dibromide XII. While this procedure has yet to be successful, catalyzed processes[6-8] for this reaction offer good promise.

The most direct method of synthesis, via nucleophilic displacement, is not very applicable to meta-substituted derivatives because of very slow rates.[9] The

starting
materials

product and
biproducts
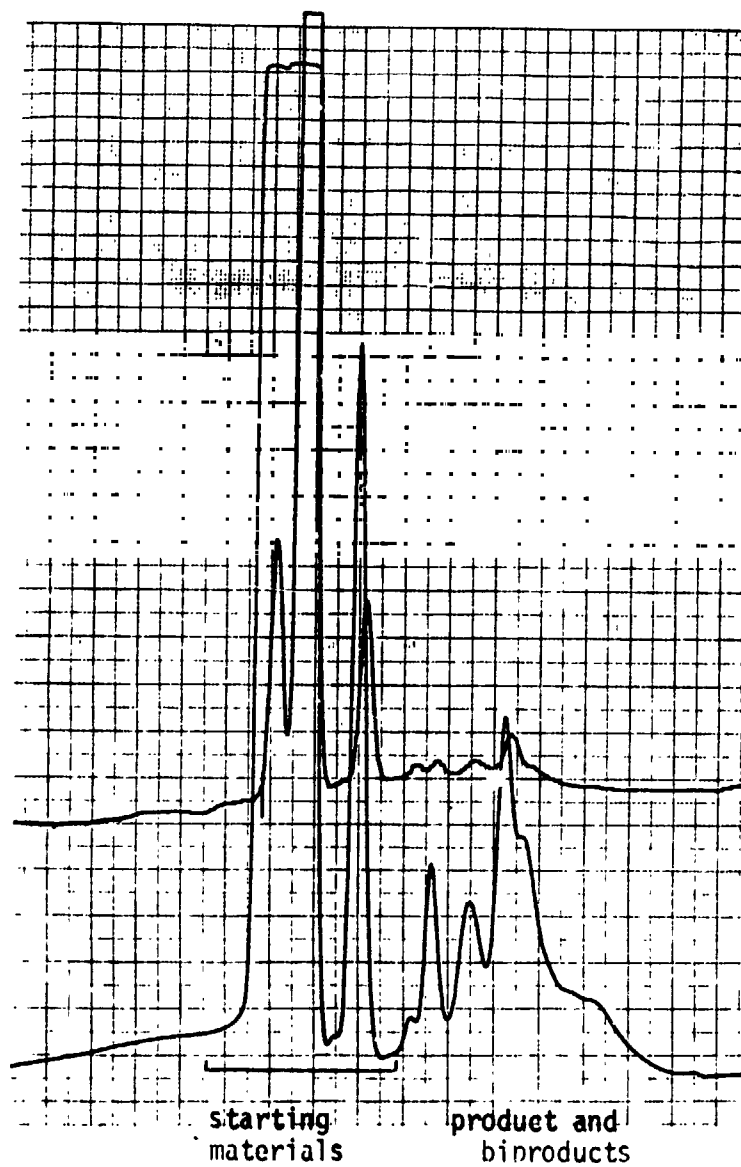
Figure 3. Liquid chromatogram of the crude product mixture from the reaction of catechol, sodium salt, with m-chloro-nitrobenzene.
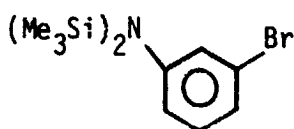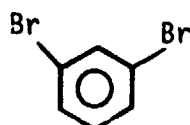
(XI) (XII)

use of aprotic dipolar solvents offers some promise,[10] yet our attempts to convert m-phenylenediamine into the three-ring tetramine, under a variety of conditions failed. The sulfur compounds, had the appropriate m-phenylene dithiol been available, would surely have given a good yield of product XIII.
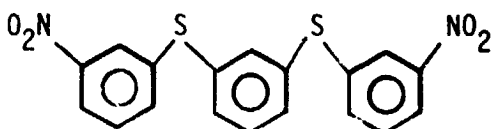
(XIII)

The target compounds VII for the $X = CH_2$ and 0, n = 0, 1, and 2 series not mentioned above have been procured or are being prepared presently. The only other target being attempted as part of this effort is VII ($X = CH_2$, n = 3). This molecule is being prepared by the route shown below:

(XIV)

(XV)

$$\xrightarrow[-AcOH]{H_3O^+} \xrightarrow[H_3PO_2]{NaNO_2} \xrightarrow{reduction} VII \quad (X = CH_2, n = 3)$$

(-NH)
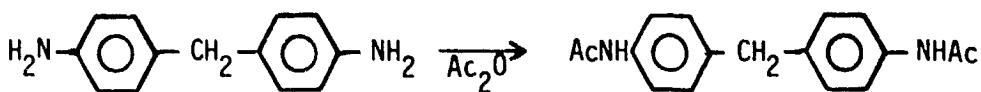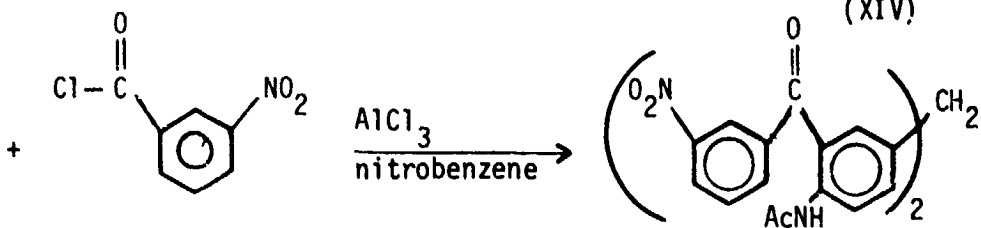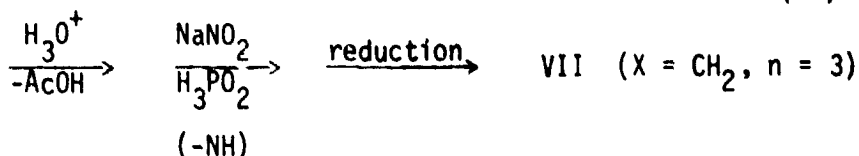
Project B. Ultrahigh Molecular Weight Silicone Elastomers.

The objective of this project is to investigate ways of preparing silphenylene-siloxane copolymers having ultrahigh molecular weights. Such polymers should be tougher elastomers than the lower molecular weight varieties. Participation in

this project has been limited to an advisory capacity in discussions of strategy. Basically, the problem is getting a coupling agent that can link together the termini of the existing chains. Dr. W. J. Patterson has procured or prepared the materials necessary for a test of the theory that highly reactive diaminosilanes may serve to give ultrahigh molecular weights with hydroxyl-terminated silphenylene-siloxane copolymers.

References

1. Lee, H. and Neville, K., Handbook of Epoxy Resins, McGraw-Hill Book Co., New York, 1967.

2. Epoxy Resins, May, C. A. and Tanaka, Y., Eds., Marcel Dekker, Inc., New York, 1973.

3. Patterson, W. J., personal communication, 1981.

4. Bilow, N. H., Boschan, R. H., and Landis, A. L., Hughes Tech Report AFML-TR-72-57, part II, 1973.

5. Semmel, M. L., unpublished results, 1981.

6. Tamura, M. and Kochi, J. Synthesis, 303 (1971).

7. Tamano, M., Sumitani, K , and Kumada, M., J. Amer. Chem. Soc., 94, June 14 (1972).

8. Tamura, M. and Kocki, J., J. Organomet. Chem., 42, 205 (1972).

9. Newton, A. B. and Rose, J. B., Polymer, 13, 465 (1972).

10. Su, T.-L., Klein, R.S., and Fox, J. J., J. Org. Chem., 46, 1790 (1981).

1981 NASA/ASEE SUMMER FACULTY RESEARCH
FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

A STUDY OF COMPONENT DESIGN LOADS DUE TO
DYNAMIC ENVIRONMENT

Prepared By: Rajinder S. Mehta

Academic Rank: Assistant Professor

University and   Alabama A&M University
Department: Department of Civil Engr./Engr. Tech.

NASA/MSFC
  Division: Systems Dynamics Laboratory
  Branch: Structural Dynamics Division

MSFC
Counterpart: R. Schock/L. Schutzenhofer/R. Ryan

Date: August, 1981

Contract No.:   NGT 01-008-021
                    The University of Alabama in Huntsville

A STUDY OF COMPONENT DESIGN LOADS DUE

TO DYNAMIC ENVIRONMENT

BY

Rajinder S. Mehta
Assistant Professor of Civil Engineering/Engr. Technology
Alabama A & M University
Normal, Alabama

ABSTRACT

The dynamic environments for Space Shuttle payloads during liftoff are the engine generated acoustics as a result of the solid rocket motors and Space Shuttle main engines exhaust streams shearing out the ambient atmosphere causing acoustic transmission through the cargo bay walls. This results in internal cargo bay acoustic impingement on the payload causing random vibration. The low frequency dynamic environment is due to the dynamics of the integrated Shuttle vehicle. The discussion in this paper is primarily on the random vibration of the components causing concern for high load factors for design evaluation. The important factors to be considered in the design of acoustically loaded structure are the prediction of acoustic environment, prediction of structural response to this environment, and to determine the effect of the response on the structural strength requirements.

This paper accounts for two basic responses. The primary structure response is analyzed first due to random vibration. The output of primary structure response is used as an input to the secondary structure to calculate the component response.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

A STUDY OF COMPONENT DESIGN LOADS DUE TO DYNAMIC ENVIRONMENT

I. INTRODUCTION

In evaluating the acoustical design requirements on support structure and its component, appropriate combined environments should be specified. The payload design loads of Space Shuttle are critical because both the high frequency random vibration and the low frequency vehicle dynamics are acting simultaneously during lift-off. A component and its support structure is thus designed to withstand these loads. However, the primary objective is that the component be able to withstand the following design criteria:

1) Applied loads -- low and high frequency

2) Environments

3) Constraints

4) Margin of safety.

This paper considers only the evaluation of the design loads of the component due to dynamic environments.

Low frequency transient loads can be derived by the normal-mode analysis techniques from which acceleration time histories are computed and thus, low frequency loads established. High frequency loads are based on the local stiffness of the component and its main properties, and are discussed in this paper.

II. RESPONSE OF COMPONENTS

The dynamic response of components in a high level acoustic environment will consist of one or more of the following:

1) Vibration of the components mounting induced by acoustic excitation of the support structure.

2) Vibration of the component as a rigid body on its mounting due to net acoustic forces on the component surface.

A hard mounted component will be more susceptible to the vibration of the mounting points rather than direct acoustic excitation. If one thinks of the noise level inside and outside the vehicle as being approximately the same, it is to be expected that the diaphragm action of the skin would result in more vigorous excitation of the component than is possible by the direct acoustic impingement.

In general the response of the primary structure to the acoustic environment causes a vibration input to a component through its attach points that is more severe than that resulting from direct acoustic excitation of the components. Components with high mass to volume ratio, small surface area, high internal damping, or those with condensed parts are not susceptible to acoustic excitation.

The structural damping ratio is an important aspect in the response of a component and its support structure. Once the structure is built it can be tested to measure the damping ratio by measuring the bandwidth at the half power points of the frequency response curve. However, in calculating the dynamic g_{peak} load a value of damping ratio must be assumed.

The response of any structure to a single point random excitation can be computed by a simple numerical procedure, provided the spectral density of the excitation and the frequency response curve of the structure are known. Consider the structure of Figure 1a whose base is subjected to a random acceleration input with the power spectral density function shown in Figure 1b. To compute the response of the point p and establish the probability of exceeding any specified acceleration.
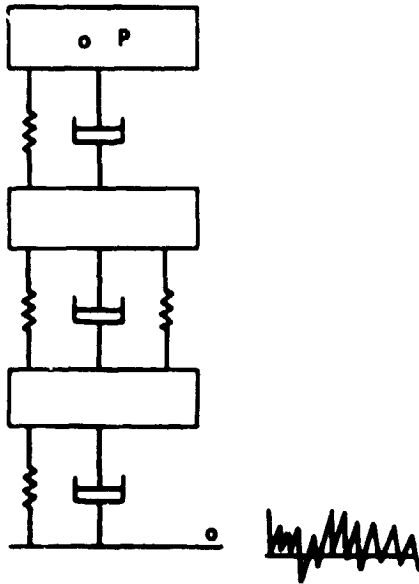
Figure 1a. Base excited structure.

Figure 1b. Input power spectral density function.
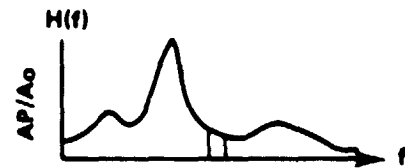
Figure 1c. The frequency response function.

The frequency response function H(f) for the point p may be obtained experimentally by applying to the base a variable frequency sinusoidal shaker with a constant acceleration input A_o and measuring the acceleration response at p. Dividing the measured acceleration by A_o, H(f) is shown in Figure 1c.

The mean square response $\bar{A}_p^2$ at p can be calculated numerically from the equation

$$\bar{A}_p^2 = \sum_i W(f_i) \, |H(f_i)|^2 \, \Delta f_i \quad . \tag{1}$$

To ensure structural integrity of the intervening structure to which a component is mounted, design loads based on estimated transmissibility (Q) have to be calculated prior to testing. Since vibration damage to structures occurs at resonances, therefore response of a resonator from random signal is investigated.

A. Response of a Resonator

In dealing with random signals we are not dealing with energy or power at any one frequency, but with the total energy in some specified band of frequencies called the acceleration density. Total summation of acceleration density over a frequency spectrum yields the mean square value of acceleration. The function obtained by allowing the bandwidth Δf to approach zero is called the acceleration power spectral density

$$W(f) \;=\; \lim_{\Delta f \to 0} \frac{a^2 \, \Delta f}{\Delta f} \quad , \tag{.2}$$

where a is the root mean square value of the random acceleration. The over-all rms value is obtained by taking the square root of the area under the power spectrum.

$$a^2 \;=\; \int_{f_1}^{f_2} W(f) \, df \tag{3}$$

where a^2 is the mean square value or the variance of the vibration from a zero mean between the frequency limits f_1 and f_2.

At a given frequency f, the power spectral density of the response $W(f)$ of the mechanical resonator is equal to the product of the power spectral density of the excitation $W_x(f)$ and the square of the absolute value of the sinusoidal transfer function $x(f)^2$ at the same frequency:

$$W(f) \;=\; W_x(f) \cdot |x(f)|^2$$

where

$$x(f) \;=\; \frac{1 + \left(\dfrac{f}{f_1}\right)^2}{\left(1 - \dfrac{f^2}{f_1^2}\right)^2 + \dfrac{f^2}{f_1^2} Q^2} \tag{4}$$

The rms response of a simple resonator consisting of m_1, c_1, and k_1 to an excitation of constant power spectral density $W(f)$ with resonant frequency f_1 and $Q = 2 \pi f_1 m_1 / c_1$ is obtained

$$a^2 = \int_0^\infty \frac{\left(1 + \dfrac{f^2}{f_1^2} Q^2\right) W(f)}{\left(1 - \dfrac{f^2}{f_1^2}\right)^2 + \dfrac{f^2}{f_1^2 Q^2}} \, df$$

$$= W(f_1) \int_0^\infty \frac{df}{\left(1 - \dfrac{f^2}{f_1^2}\right)^2 + \dfrac{f^2}{f_1^2 Q^2}} = \frac{\pi}{2} f_1 Q \setminus f_1) \ . \tag{5}$$

The above approximation is based on the assumption that the response results primarily from the fundamental mode and the resonator is not too heavily damped. However, in practical design problems many mechanical resonators are coupled together. The responses of coupled resonators will behave differently.

Two basic responses are considered in this paper. The primary structure is analyzed first for random response. The output from this response is used as input to the secondary structure to calculate the component response. This can be better explained by electric circuit analogy making use of the Norton's theorem.

B. Response of Component and Its Intervening Structure Supported on the Primary Structure Using Application of the Impedance

Consider the response of the primary structure before addition of mechanical impedance of a component mass.

1) Z_s has no impedance due to mass

2) Z_c has impedance due to mass only
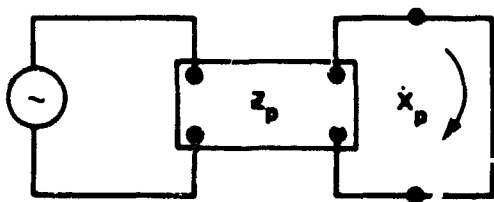
Figure 2a. Primary structure.
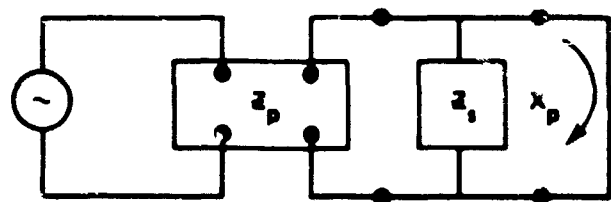
Figure 2b. Addition of component support structure.
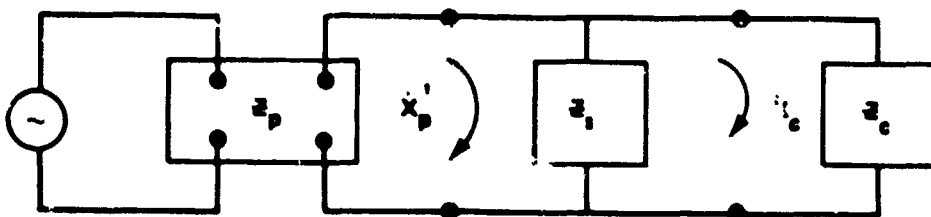
Figure 2c. Addition of component.

Assume that the response X_p at any point in the primary structure is known or is computed using a force generator or a motion generator. The response of the primary structure is not altered when Z_s, the mechanical impedance of the component support structure, is added to the primary structure (Z_s has no impedance due to mass).

When the component is added to the system, the structural response $\dot{X}_p$ is altered to $\dot{X}_p'$; the component has a response $\dot{X}_c$, while the supporting structure response to the difference of the altered primary structure response and the component response, i.e., $\dot{X}_p' - X_c$.

Using Norton's Theorem

For Component

$$\frac{\dot{X}_c(\omega)}{\dot{X}_p(\omega)} = \frac{Z(\omega)_i}{Z(\omega)_i + Z_c(\omega)} = \frac{Z(\omega)_i}{1 + Z_c(\omega)\, Z(\omega)_i^{-1}} \quad . \tag{6}$$

For Support Structure

$$\frac{\dot{X}_p'(\omega) - \dot{X}_c(\omega)}{\dot{X}_p(\omega)} = \frac{X_c(\omega)}{X_p(\omega)} \frac{Z_c(\omega)}{Z_s(\omega)} \quad . \tag{7}$$

The mechanical impedance $Z(\omega)_i$ at the ith position in the structure is defined as the impedance of the primary structure at the point of attachment plus the effect of the mechanical impedance of the component support structure.

$$Z^{-1}(\omega)_i = \left(Z_p^{-1}(\omega)_i + Z_s^{-1}(\omega)_i \right) \quad . \tag{8}$$

For any given frequency, the mechanical impedance $Z_p(\omega)_i$ at the ith position of the primary structure is determined by applying a force $F(\omega)_i$ and computing the complex ratio $Z_p(\omega)_i = F(\omega)_i / \dot{X}(\omega)_i$ and using the force generator equation

$$\{X(\omega)\} = \left[X_N \right] \left[H_N(\omega) \right] \left[X_N^T X_N \right]^{-1} = \left[X_N^T \right] \left[X_F \right] \{F\} \quad . \tag{9}$$

$$Z_p^{-1}(\omega)_i = \frac{\dot{X}(\omega)_i}{F(\omega)_i}$$

$$= j\omega \{e_i\}^T \begin{bmatrix} X_N \end{bmatrix} \begin{bmatrix} H_N(\omega) \end{bmatrix} \begin{bmatrix} X_N^T X_N \end{bmatrix}^{-1} \begin{bmatrix} X_N^T \end{bmatrix} \begin{bmatrix} X_F \end{bmatrix} \{e_i\} \qquad (10)$$

e_i = a column vector where the ith element is unity and all other elements are zero and

$$\{a\} = \begin{bmatrix} X_N^T X_N \end{bmatrix}^{-1} \{X_N^T\} \{X_{ST}\} \qquad .$$

The mechanical impedance of the component support structure is

$$Z_s^{-1}(\omega) = J\omega(K_S + J\omega C_S)^{-1} \qquad (11)$$

where

$$K_S = \text{stiffness of component support structure}$$

$$C_S = \text{damping coefficient of component support structure} \qquad .$$

The mechanical impedance of the component is

$$Z_c(\omega) = j\omega M_c \qquad . \qquad (12)$$

The normalized impedances are obtained by multiplying (11) (12) and (10) (12)

$$Z_c(\omega) \, Z_s^{-1}(\omega) = -\omega^2 M_c (K_S + J\omega C_S)^{-1} = -\left(\frac{\omega}{\omega_c}\right)^2 \left(1 + 2J \zeta_c \frac{\omega}{\omega_c}\right)^{-1}$$

$$Z_c(\omega) \, Z_p^{-1}(\omega) = -\omega^2 M_c \{e_i\}^T \begin{bmatrix} X_N \end{bmatrix} \begin{bmatrix} H_N(\omega) \end{bmatrix} \begin{bmatrix} X_N^T X_N \end{bmatrix}^{-1} \begin{bmatrix} X_N^T \begin{bmatrix} X_F \end{bmatrix} \{e_i\} \qquad .$$

To find the magnitude of

$$\frac{\dot{X}_c(\omega)}{\dot{X}_p(\omega)} = \frac{1}{1 + Z_c(\omega)\, Z(\omega)_i^{-1}}$$

we have to use theory of complex variables

$$Z_c(\omega)\, Z_p^{-1}(\omega) = A_p + jB_p$$

where

A_p = real part of normalized mechanical impedance of component mass

B_p = imaginary part of normalized mechanical impedance of component mass

$$A_p = -\omega^2 M_c \{e_i\}^T [X_N] \left[\text{Re } H_N(\omega) \right] [X_N^T X_N]^{-1} [X_N^T][X_F] \{e_i\}$$

$$B_p = -\omega^2 M_c \{e_i\}^T [X_N] \left[\text{Im } H_N(\omega) \right] [X_N^T X_N]^{-1} [X_N^T][X_F] \{e_i\} \qquad \text{(13)}$$

$$Z_c(\omega)\, Z_s^{-1}(\omega) = A_s + J B_s = -\left(\frac{\omega}{\omega_c}\right)^2 \left(1 + 2J\,\zeta_c \frac{\omega}{\omega_c}\right)^1$$

where

A_s = real part of normalized mechanical impedance of component support structure

B_s = imaginary part of normalized mechanical impedance of component support structure

$$A_s = -\left(\frac{\omega}{\omega_c}\right)^2 \left[(1)^2 + \left(2\,\zeta_c \frac{\omega}{\omega_c}\right)^2\right]^{-1}$$

$$B_s = 2\,\zeta_c \left(\frac{\omega}{\omega_c}\right)^3 \left[(1)^2 + \left(2\,\zeta_c \frac{\omega}{\omega_c}\right)^2\right]^{-1}$$

$$\left| Z_c(\omega)\, Z_s^{-1}(\omega) \right| = \left(\frac{\omega}{\omega_c} \right)^2 \left[(1)^2 + \left(2\, \zeta_c\, \frac{\omega}{\omega_c} \right)^2 \right]^{-1/2} \quad .$$

Therefore Magnitude of Responses

<u>Component</u>

$$\left| \frac{\dot{X}_c(\omega)}{\dot{X}_p(\omega)} \right| = \left\{ (1 + A_p + A_s)^2 + (B_p + B_s)^2 \right\}^{-1/2} \quad . \tag{14}$$

<u>Component Support Structure</u>

$$\left| \frac{\dot{X}_p{}'(\omega) - \dot{X}_c(\omega)}{\dot{X}_p(\omega)} \right| = \left| \frac{\dot{X}_c(\omega)}{\dot{X}_p(\omega)}\, \frac{Z_c(\omega)}{Z_s(\omega)} \right| = \left| \frac{\dot{X}_c(\omega)}{\dot{X}_p(\omega)} \right| \left| Z_c(\omega)\, Z_s^{-1}(\omega) \right|$$

$$= \left| \frac{\dot{X}_c(\omega)}{\dot{X}_p(\omega)} \right| \left(\frac{\omega}{\omega_c} \right)^2 \left[1^2 + \left(2\, \zeta_c\, \frac{\omega}{\omega_c} \right)^2 \right]^{-1/2} \quad . \tag{15}$$

C. Random Response of Two Coupled Resonators for No Loading Effect Between Resonators

The root mean square response of a simple mechanical resonator mounted on another mechanical resonator excited by a white noise is co. 'dered in this section for the case when there is no loading effect between resonators Consider the model of Figure 3 for discussion.
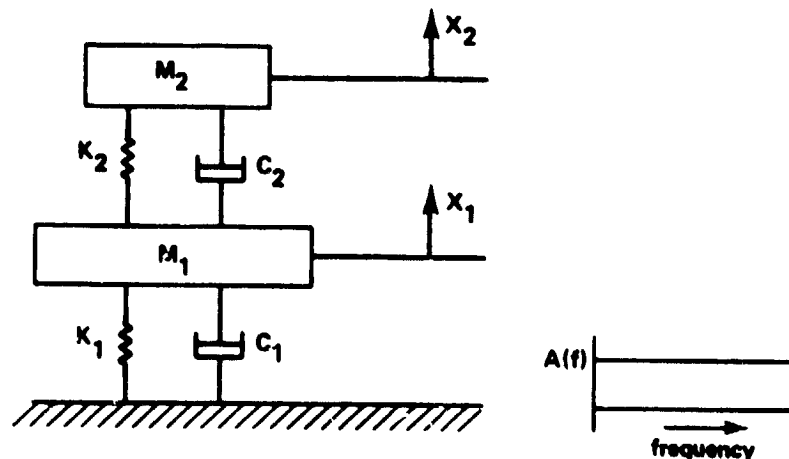
Figure 3. Two coupled mechanical resonators excited by random vibration.

We shall be concerned with only four parameters, f_1, f_2, Q_1, and Q_2, instead of m_1, c_1, K_1, m_2, c_2, and K_2 because the resonator comprising of m_2, K_2, and c_2 does not affect the motion of m_1 (case of no loading).

$$
\left.
\begin{aligned}
f_1 &= \frac{1}{2\pi}\sqrt{\frac{K_1}{m_1}} \\[2ex]
f_2 &= \frac{1}{2\pi}\sqrt{\frac{K_2}{m_2}} \\[2ex]
Q_1 &= \frac{1}{2\zeta_1} = 2\pi f_1 \frac{m_1}{c_1} \\[2ex]
Q_2 &= \frac{1}{2\zeta_2} = 2\pi f_2 \frac{m_2}{c_2}
\end{aligned}
\right\}
\tag{16}
$$

The rms response of a single resonator consisting of m_1, c_1, and K_1 excited by a constant power spectral density has already been obtained by equation (5)

$$
\sqrt{\langle X_1^2 \rangle_{avg}} = \left[\frac{\pi f_1 Q_1 W(f_1)}{2} \right]^{1/2} \quad ,
\tag{17}
$$

provided that the resonator is not heavily damped. The rms response of m_2 is obtained by taking the square root of the area under the power spectrum of the response. For a constant excitation power spectral density W, the response in partially normalized form is

$$
\lambda = \frac{\left(\bar{X}_2^2\right)^{1/2}}{W^{1/2}} = \left[\int_0^\infty \left\{ \frac{1 + \dfrac{f^2}{f_1^2 Q_1^2}}{\left(1 - \dfrac{f^2}{f_1^2}\right)^2 + \left(\dfrac{f^2}{f_1^2 Q_1^2}\right)} \right\} \left\{ \frac{1 + \dfrac{f^2}{f_2^2 Q_2^2}}{\left(1 - \dfrac{f^2}{f_2^2}\right)^2 + \dfrac{f^2}{f_2^2 Q_2^2}} \right\} df \right]^{1/2} .
\tag{18}
$$

If f_1 and f_2 are not close and Q_1, Q_2 are large, the response may be considered to come from the immediate region of the two resonances. One fraction of the integrand peaks, where the other is constant. Hence

$$\left\{ \frac{\pi}{2} f_1 Q_1 \left[\frac{1 + \dfrac{f_1^2}{f_2^2 Q_2^2}}{\left(1 - \dfrac{f_1^2}{f_2^2}\right)^2 + \dfrac{f_1^2}{f_2^2 Q_2^2}} \right] + \frac{\pi}{2} f_2 Q_2 \left[\frac{1 + \dfrac{f_2^2}{f_1^2 Q_1^2}}{\left(1 - \dfrac{f_2^2}{f_1^2}\right)^2 + \dfrac{f_2^2}{f_1^2 Q_1^2}} \right] \right\}^{1/2} \quad , \tag{19}$$

or

$$\left\{ \frac{\pi}{2} \left[\frac{f_1 Q_1}{\left(1 - \dfrac{f_1^2}{f_2^2}\right)^2} + \frac{f_2 Q_2}{\left(1 - \dfrac{f_2^2}{f_1^2}\right)^2} \right] \right\}^{1/2} \quad .$$

If $W(f)$ is not constant

$$\sqrt{\langle X_2^2 \rangle}_{\text{avg}} = \left\{ \frac{\pi}{2} \left[\frac{f_1 Q_1 W(f_1)}{\left(1 - \dfrac{f_1^2}{f_2^2}\right)^2 + \dfrac{f_1^2}{f_2^2 Q_2^2}} + \frac{f_2 Q_2 W(f_2)}{\left(1 - \dfrac{f_2^2}{f_1^2}\right)^2 + \dfrac{f_2^2}{f_1^2 Q_1^2}} \right] \right\}^{1/2} \quad . \tag{20}$$

The above concept can be extended to n resonators cascaded without interloading.

III. POSSIBLE CAUSE OF HIGH ACCELERATIONS AND HIGH STRESSES

The predicted random vibration environment defined at the interface has considerable conservatism because vibration criterion is extrapolated to predict the responses in zones where there are no flight measurements. The vibration levels at the interfaces are thus conservatively enveloped to account for all the peaks.

The low frequency coupled G_{peak} loads and high frequency random G_{peak} loads are being combined and are applied simultaneously along three mutually perpendicular axes; however, in flight the possibility of this occurrence is rare. In testing a component, however, the loading is applied in sequence rather than tri-axial.

The strength analysis is based on the G_{peak} loading applied simultaneously, whereas the peaks contribute mostly to the fatigue failure. Therefore, consider rms stress response for strength. Conservatism also lies in the factors of safety being used to perform the stress analysis.

IV. CONCLUSIONS

1) The damping of resonant members in the components may often be less than that of the primary structure and the loading of one member by another may be less. For these and other reasons the transmission of the vibration through coupled resonators may sometimes produce large accelerations. However, high accelerations do not necessarily mean high stresses. The high acceleration response may be from a higher frequency mode where, due to the inertia forces being out of phase, the stresses are not correspondingly increased.

2) Vibration damage to structures ordinarily occurs at resonances. If the resonances were staggered or highly damped, it might be possible to decrease the design accelerations.

3) Ambient acoustic noise is most likely to be damaging if there is a progression from a low impedance element of a large area to the structural members of higher impedance. On the other hand, vibration in the structural members is most likely to damage the parts if there is a progression from a high to low mechanical impedance (i.e., if the parts do not load the structural members to which they are fastened.) Controlling the mechanical impedance of the component parts may help to avoid such progression. This will lower the transmissibility at frequencies of resonance and make the need for detuning or damping less critical.

4) The dynamic characteristics of a component influences its own vibration environment during flight. This affect results in a reduction of vibration input criteria to the component at the component resonant frequencies. This implies that the response of a heavy component will tend to reduce Spacelab structure responses and thus will affect the motion of the base of the component considerably. This effect should be accounted for in estimating the transmissibility.

REFERENCES

1. Miles, J. W.: "On Structural Fatigue Under Random Loading," The Journal of Aeronautical Science, Vol. 21, No. 11, November 1954.

2. Powell, Alan: "On the Fatigue Failure of Structures Due to Vibration Excited by Random Pressure Fields," The Journal of Acoustical Society of America, Vol. 30, No. 12, December 1954.

3. Mehta, R. S.: "Acoustic Fatigue and Dynamic Response of Panels," Masters Thesis at University of Alabama in Huntsville, November 1976.

4. Morrow, C. T. and Muchmore, R. B.: "Shortcomings of Present Methods of Measuring and Simulating Vibration Environments," Journal of Applied Mechanics, September 1955.

5. Morrow, C. T., Troesch, A. B., and Spence, H. R.: "Random Response of Two Coupled Resonators Without Loading," The Journal of the Acoustical Society of America, Vol. 33, No. 1, January 1961.

6. Mains, R. M.: "Minimizing Damage from Random Vibration," The Journal of Acoustical Society of America, Vol. 30, No. 12, December 1958.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

ATTITUDE CONTROL OF THE SPACE CONSTRUCTION BASE:
A MODULAR APPROACH

Prepared By: Daniel A. O'Connor, DSc.

Academic Rank: Assistant Professor

University and Department: Clarkson College of Technology
                                    Department of Electrical and
                                    Computer Engineering

NASA/MSFC:
      Division: Control Systems Division

MSFC Counterpart: Dr. Michael T. Borelli

Date: August 8, 1981

Contract No: NGT 01-008-021
                                    The University of Alabama in Huntsville

ATTITUDE CONTROL OF THE SPACE CONSTRUCTION BASE:
A MODULAR APPROACH

Daniel A. O'Connor
Dept. of Electrical and Computer Engineering
Clarkson College of Technology
Potsdam, New York

ABSTRACT

The object of this study is to develop attitude control systems for large space structures. These structures are large vibrating systems subject to various disturbances and are designed to support experiments with stringent pointing and stabilization requirements. The construction of the structure will be accomplished by joining together modules; each module is constructed on the ground and delivered to an orbiting base by the Space Shuttle. As each new part is added to the base, the dynamics change and, hence, the overall control problem will change. Each module will have on-board control computers and actuators and must integrate with the base control system to provide overall control for the augmented structure.

In this work, we consider a planar model of a space base and one module. For this simplified system, we develop a feedback controller which is compatible with the construction method described above. The systems dynamics are decomposed into two parts corresponding to base and module. The information structure of the problem is non-classical in that not all system information is supplied to each controller. The base controller is designed to accommodate structural changes that occur as the module is added and the module controller is designed to regulate its own states and follow commands from the base. Overall stability of the system is checked by Liapunov analysis and controller effectiveness is verified by computer simulation.

TABLE OF CONTENTS

1. Introduction

In this report we describe the application of some large scale systems techniques to the attitude control of multibody satellites. Specifically, we consider the problem of how to control the satellite with actuators on each body, but restrictions of the transfer of sensor information between bodies.

To focus on a particular problem, we considered a planar model of a two-body vehicle with flexible appendages. This vehicle is a variant of the hypothetical space construction base investigated previously in the Bendix report, [1], which uses multilevel techniques for attitude control.

It is important to note the difference between the techniques we apply and multilevel techniques. The basic function of multilevel control is to subdivide a large dimension linear quadratic regulator problem into smaller dimension subproblems which can be individually solved and then to coordinate the controls of each subproblem by a higher order supervisory control. This method (as practiced in [1]), is iterative and upon convergence leads to the solution of the large dimension linear quadratic regulator problem. The resulting controller requires information sharing between each of the partioned subsystems.

Our approach is to apply the methods of decomposition and aggregation [2]. Here we also subdivide the problem into smaller dimension linear quadratic regulator problems, but because of the restrictions on information sharing, we have a limit on the degree of coordination between subsystems. This lack of coordination may result in a greater quadratic cost compared with the multilevel approach, but possesses the advantage of connective stability. That is, if the controlled bodies are disconnected from one another, each one remains a stable regulator. It is easy to see [2] that this property is not shared by centralized linear quadratic regulators.

In Section 2, we describe the decomposed control techniques and their application to the attitude control problem. In Appendix C we detail the numerical simulation results computed to date and in Appendix A, we derive the equations of planar motion for the space construction base and module.

2. Control System Analysis

The small angle equations of motion of the space base and the space base with module are described by,

$$M_1 \ddot{r}_1 + D_1 \dot{r}_1 + K_1 r_1 = T_1 v_1 \tag{2.1}$$

$$M_2 \ddot{r}_2 + D_2 \dot{r}_2 + K_2 r_2 = T_2 v_2 \tag{2.2}$$

where r_1 is the 5x1 vector of base generalized coordinates $[x, y, \theta_1, q, r]^T$ and r_2 is the augmented 6x1 vector of generalized coordinates of the base and module $r_2 = [r_1 r, \theta_2]^T$

The 3x1 vector v_1 represents central body control on the base and the 4x1 vector v_2 represents central body control on the base and module. Derivation of 2.1 - 2.2 and identification of parameters is found in Appendix A.

The systems can be placed in first order state variable form as follows:

define $\quad x_1 = \begin{vmatrix} r_1 \\ \dot{r}_1 \end{vmatrix}, \quad A_1 = \begin{vmatrix} 0 & I \\ -M_1^{-1}K_1 & pM_1^{-1}D_1 \end{vmatrix}, \quad B_1 = \begin{vmatrix} 0 \\ T_1 \end{vmatrix}$

To give $\qquad \dot{x}_1 = A_1 x_1 + B_1 v_1$ $\qquad\qquad$ (2.3)

and define $\quad z = \begin{vmatrix} r_2 \\ \dot{r}_2 \end{vmatrix}, \quad A_2 = \begin{vmatrix} 0 & I \\ -M_2^{-1}K_2 & -M_2^{-1}D_2 \end{vmatrix}, \quad B_2 = \begin{vmatrix} 0 \\ T_2 \end{vmatrix}$

To give $\qquad \dot{z} = A_2 z + B_2 v_2$ $\qquad\qquad$ (2.4)

The relationship between the two models can be clearly seen if we modify equation (2.4). First we define new coordinates by the permutation $\bar{z} = Mz$

where $\qquad \bar{z} = \begin{vmatrix} x_1 \\ x_2 \end{vmatrix}, \quad x_1 = \begin{vmatrix} r_1 \\ \end{vmatrix} \quad x_2 = \begin{vmatrix} \theta_2 \\ \dot{\theta}_2 \end{vmatrix}$

we obtain the equation in $\bar{z}$

$$\dot{\bar{z}} = MAM^{-1}\bar{z} + MB_2 v_2$$

which can be written in terms the coefficient matrices of (2.3) as

$$\begin{vmatrix} \dot{x}_1 \\ \dot{x}_2 \end{vmatrix} = \begin{vmatrix} A_1 + A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} + \begin{vmatrix} B_1 + B_{11} & B_{12} \\ B_{21} & B_{22} \end{vmatrix} \begin{vmatrix} u_1 \\ u_2 \end{vmatrix} \quad (2.5)$$

here $u_1 = v_1$ is the 3x1 vector of controls which directly effect the base and u_2 is the scalar control which directly effects the module. The values of the matrices Aij and Bij are given in Table 1.

2.1 Centralized Linear Regulator

As a first step we consider the centralized linear quadratic regulator. We take this approach to determine the structure of the undecomposed problem (i.e., controllability, stabilizability) and also to provide some comparison with the controllers designed by the decomposition aggregation approach. In this case, the distinction between the states, x_1, x_2 can be suppressed and we can rewrite (2.5) as

$$\dot{x}_1 = Ax + Bu$$

where $x = \begin{vmatrix} x_1 \\ x_2 \end{vmatrix}$ is the 12x1 composite state vector

and $u = \begin{vmatrix} u_1 \\ u_2 \end{vmatrix}$ is the 4x1 composite control vector

$$A = \begin{vmatrix} A_{11} + A_1 & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \qquad B = \begin{vmatrix} B_{11} + B_1 & B_{12} \\ B_{21} & B_{22} \end{vmatrix}$$

Optimal regulation is obtained by choosing the control law which minimizes the quadratic performance index

$$\int_0^\infty (x^T Q x + u^T R u) dt$$

where

      Q is a non-negative definite 12x12 matrix
      R is a positive definite 4x4 matrix

If a minimizing control exists, it is realized by

$$u = -R^{-1}B^T P x$$

where P is a solution of the algebraic Riccati equation

$$A^T P + PA - PBR^{-1}B^T P + Q = 0$$

Moreover, the minimizing control law will exist if the pair (A,B) is stabilizable and the associated closed loop system

$$\dot{x} = (A - BR^{-1}B^T P)x$$

will be stable is $(Q^{1/2}, A)$ is a detectable pair [3].

Stabilizability of the composite system is easily checked with the following criterion.

Proposition 1   (A,B) is a controllable pair if and only if

$$\text{rank } [M_2\lambda^2 + D_2\lambda + K_2, \; T_2] = 6 \qquad \text{for all } \lambda$$

If the rank condition holds for all λ such that $\text{Re}\lambda \geq 0$ then (A,B) is stabilizable.

Proof   Immediate application of the well-known Hautus condition [4] to mechanical systems.

Now taking advantage of the particular structure of the composite model, we can partition the characteristic matrix

$$M_2\lambda^2 + D_2\lambda + K_2$$

as

$$\begin{vmatrix} M_{11}\lambda^2 & M_{12}\lambda^2 \\ M_{21}\lambda^2 & M_{22}\lambda^2 + \overline{D}\lambda + \overline{K} \end{vmatrix}$$

where Mij are 33 matrices

and $\overline{D}$, $\overline{K}$ are 3x3 diagonal matrices (see Appendix A)

$$\overline{D} = \begin{vmatrix} 2\xi_A\omega_A M_A & 0 & 0 \\ 0 & 2\xi_B\omega_B M_B & 0 \\ 0 & 0 & D_\theta \end{vmatrix}$$

$$\overline{K} = \begin{vmatrix} \omega_A^2 M_A & 0 & 0 \\ 0 & \omega_B^2 M_B & 0 \\ 0 & 0 & K_\theta \end{vmatrix}$$

if we partition T_2 as $\begin{vmatrix} \hat{T}_1 \\ \hat{T}_2 \end{vmatrix}$ where $\hat{T}_1$, $\hat{T}_2$ are

3x4 matrices, the Proposition 1 implies that the composite system is controllable if and only if there is no nonzero pair of vectors x_1, x_2 such that for any λ

$$x_1^T M_{11} \lambda^2 + x_2^T M_{21} \lambda^2$$

$$x_1^T M_{12} \lambda^2 + x_2^T [\lambda^2 M_{22} + \lambda \overline{D} + \overline{K}] = 0 \qquad (2.6)$$

$$x_1^T \hat{T}_1 + x_2^T \hat{T}_2 = 0$$

Consideration of the rigid body mode $\lambda = 0$ implies that the system will be controllable only if $\hat{T}_1$ has full row rank, or equivalently if each rigid body coordinate is controlled. Since we assume this to be true, the system will be completely controllable if and only if there is no nonzero vector x_2 such that

$$x_2^T M_{21} = 0$$

$$x_2^T [\lambda^2 M_{22} + \lambda \overline{D} + \overline{K}] = 0$$

$$x_2^T \hat{T}_2 = 0$$

or equivalently if

$$\text{rank } [\lambda^2 M_{22} + \lambda \overline{D} + \overline{K}, \ M_{21}, \ \hat{T}_2] = 3$$

Expanding this criterion in the model parameters (Appendix A) we obtain the controllability condition;

$$\text{rank } \begin{vmatrix} M_A(\lambda^2 + 2\xi_A \omega_A + \omega_A^2) & 0 & 0 \\ 0 & M_B(\lambda^2 + 2\xi_B \omega_B + \omega_B^2) & 0 \\ 0 & 0 & \hat{I}_2 \lambda^2 + D_\theta \lambda + K_\theta \end{vmatrix}$$

$$\begin{vmatrix} 0 & M_A \tilde{\phi} & r_{11} M_A \tilde{\phi} & 0 & 0 & 0 & 0 \\ 0 & M_B \tilde{\phi} & r_{11} M_B \tilde{\phi} & 0 & 0 & 0 & 0 \\ M_2 r_{22} & 0 & \hat{I}_{21} & 0 & 0 & 0 & 0 \end{vmatrix} = 3$$

Now since we assume the two wings are identical, the rank of the above matrix is two when λ is a flexible wing mode and, hence, the system is uncontrollable. However, there is some small damping on these modes so the system is stabilizable and the linear quadratic regulator problem will have a solution.

The flexible modes cannot be controlled; therefore, we do not reflect wing deflection errors in the cost. The weighting matrices Q and R are chosen as follows.

If $\tilde{x}_i$ is the maximum allowable value of the state x_i and $\tilde{u}_i$ is the maximum allowable value of the control u_i we pick Q and R as the diagonal matrices

$$Q = \text{diag}\{\tilde{x}_i^{-2}\} , \qquad R = \text{diag}\{\tilde{u}_i^{-2}\}$$

2.2 Decomposition Aggregation Approach

We now make explicit use of the decomposition of the state x into x_1, x_2 and break (2.5) into subsystems

$$\dot{x}_1 = (A_1 + A_{11})x_1 + A_{12}x_2 + (B_1 + B_{11})u_1 + B_{12}u_2 \qquad (2.7)$$

$$\dot{x}_2 = A_{21}x_1 + A_{22}x_2 + B_{21}u_1 + B_{22}u_2 \qquad (2.8)$$

It is important to note that (2.7) - (2.8) are not in input decentralized form. That is, the controls applied to the base u_1 directly affect the module through the matrix B_{21} and similarly control u_2 applied directly to the module effects the base through B_{12}. This fact makes decomposition aggregation methods difficult to apply as we shall see. It is possible to change coordinates and transform (2.7) - (2.8) into a pair of input decentralized systems, but this seems to destroy the physical significance that the states in the first system refer to the base vehicle and the states in the second system refer to the module. In what follows, we will leave the subsystems in the form (2.7) - (2.8), but it would be interesting to investigate the method of input decentralization and its effects on information structure in future work.

If the module were not present, the system would be described by

$$\dot{x}_1 = A_1 x_1 + B_1 u_1$$

This system is stabilizable and we can construct a control law to minimize the cost

$$J = \int_o^\infty (x_1^T Q_1 x_1 + u_1^T R_1 u_1) dt$$

The control law is given by

$$u_1 = -R_1^{-1}B_1^{T}P_1 x_1 \qquad (2.9)$$

$$P_1 A_1 + A_1^{T}P_1 - P_1 B_1 R_1^{-1} B_1^{T}P_1 + Q_1 = 0$$

Now when the module is added to the base, the system dynamics become (2.7) – (2.8) with u given by (2.9). The first question we ask is: does there exist a control of the form

$$u_1 = -R_1^{-1}B_1^{T}P_1 x_1 + K_{11}x_1 \qquad (2.10)$$

such that the perturbed linear system

$$\dot{x}_1 = (A_1 + A_{11})x_1 + (B_1 + B_{11})u_1 \qquad (2.11)$$

is a good regulator? This problem has received some attention previously [5] in a different context. We have the following partial result:

Proposition 2: If A_{11} and B_{11} satisfy the matching conditions, $A_{11} = B_1 C_1$, $B_{11} = B_1 C_2$ for some 3x10 matrix C_1 and some 3x3 matrix C_2 and if all the eigenvalues of $C_2 + I$ are positive, then the system (2.11) is stabilized by the control

$$u_1 = -(1 + \gamma)R_1^{-1}B_1^{T}P_1 x_1$$

for some sufficiently large constant γ.

Proof: A straightforward extension of Theorem 3.4 in [5] to the case where B_1 as well as A_1 is perturbed.

The significance of this result is that if the perturbations A_{11}, B_{11} satisfy the conditions of the proposition, then it will be possible to use the same control law that was used to control the base to control the base and module combination. The only modification required is that the gains are increased. A control law that can adapt to perturbations by minor changes, (such as modification of gains), would be advantageous to the attitude control problem faced here, and we will continue to study this type of controller in our future work.

$$\dot{x}_1 = (A_1 + A_{11})x_1 + (B_1 + B_{11})u_1 \qquad (2.12)$$

$$u_1 = -(1 + \gamma)R_1^{-1}B_1^{T}P_1 x_1$$

and

$$\dot{x}_2 = A_{22}x_2 + B_2 u_2 \qquad (2.13)$$

Since the control u_1 has already been selected, we need only select u_2. We pick u_2 to minimize the quadratic cost

$$\int_0^\infty [x_2^T Q_2 x_2 \cdot u_2^T R_2 u_2] dt$$

and obtain

$$u_2 = -R_2^{-1} B_2^T P_2 x_2 \tag{2.14}$$

where

$$P_2 A_2 + A_2^T P_2 - P_2 B_2 R_2^{-1} B_2^T P_2 + Q_2 = 0$$

With this choice of control, the decoupled subsystems are stable regulators and no information sharing between controllers is required; however, overall system stability is at question. Overall stability can be determined by Liapunov stability analysis [2].

Since the closed loop systems (2.12), (2.13), (2.14) are stable regulators, there exist two Liapunov functions for the respective systems

$$V_1(x_1) = x_1^T P_1 x_1 \qquad\qquad V_2(x_2) = x_2^T P_2 x_2$$

both functions are positive definite and their derivatives along the trajectories of (2.12), (2.13) are negative semidefinite

$$\dot{V}_1(x_1) = -x_1^T S_1 x_1 \qquad\qquad \dot{V}_2(x_2) = x_2^T S_2 x_2$$

where

$$S_1 = Q_1 + P_1 B_1 R_1^{-1} B_1^T P_1 + P_1 B_{11} R_1^{-1} B_{11}^T P_1 + P_1 B_{11} R_1^{-1} B_1^T P_1$$

$$+ \gamma(P_1(B_1 + B_{11}) R_1^{-1} B_1 P_1 + P_1 R_1^{-1}(B_1 + B_{11}) P_1)$$

$$- A_{11}^T P_1 - P_1 A_{11}$$

and

$$S_2 = Q_2 + P_2 B_2 R_2^{-1} B_2^T P_2$$

To derive stability conditions for the overall base-module composite system, we follow the techniques of Siljack, [2, pp. 96-99]. We obtain the following criterion.

Proposition 3: Let $\lambda_m(A)$ denote the minimum eigenvalue and $\lambda^m(A)$ be the maximum eigenvalue of an arbitrary matrix A. The overall system will be stable if the 2x2 aggregate matrix W is stable

$$W = \begin{vmatrix} -1 & 2\left(\dfrac{\lambda^m(P_1)}{\lambda_m^{1/2}(P_1)}\right)\lambda_m^{1/2}(P_2)\lambda_m^{-1}(S_2)\xi_{12} \\[4ex] 2\left(\dfrac{\lambda^m(P_2)}{\lambda^{1/2}(P_2)}\right)\lambda_m^{1/2}(P_1)\lambda_m^{-1}(S_1)\xi_{21} & -1 \end{vmatrix}$$

Here ξ_{12}, ξ_{21} are defined by

$$\xi_{12} = || A_{12} - B_{12} R_2^{-1} B_2^{T} P_2 ||$$

(2.15)

$$\xi_{21} = || A_{21} - (1 + \gamma) B_{21} R_1^{-1} B_1^{T} P_1 ||$$

(2.16)

where $||\cdot||$ denotes matrix norm.

At this point we note the difficulties inherent in applying this technique. Stability of the aggregate matrix W requires $1 - W_{12} W_{21} > 0$. For this to be the case, the stability margins of the decoupled systems should be large, (i.e., $\lambda_m(S_1)$, $\lambda_m(S_2)$ are large), and the interconnection connection terms should be small, (i.e., ξ_{12}, ξ_{21} are small). To achieve large stability margins will require large feedback gains, but from (2.15), (2.16) we see that large feedback gains will tend to increase the interconnection terms. Therefore, this approach may be difficult to implement since our problem is not input decentralized. Finally, we tried another approach which takes advantage of the input structure of our problem.

2.3 Controller with Partial Feedback

In this approach we restrict the control law for the overall system to have the form

$$u_1 = K_{11} x_1$$

$$u_2 = K_{21} x_1 + K_{22} x_2$$

(2.17)

The term $K_{21} x_1$ can be thought of as signalling from the base to the module for the purpose of coordinating control action. If we apply this control law to sytem (2.7)-(2.8) we obtain the closed loop system

$$\dot{x}_1 = (A_1 + A_{11}) + (B_1 + B_{11}) K_{11} + B_{12}^{r}{}_{21}) x_1 + (A_{12} + B_{12} K_{22}) x_2$$

$$\dot{x}_2 = (A_{21} + B_{21} K_{11} + B_{22} K_{21}) x_1 + (A_{22} + B_{22} K_{22}) x_2$$

The basic idea is to pick (2.17) to force the c ed loop dynamics into triangular form, that is pick K_{11}, K_{21} so that

$$A_{21} + B_{21} K_{11} + B_{22} K_{21} = 0$$

(2.18)

Equation (2.18) has a solution K_{11}, K_{21} if and only if the columns of A_{21} are contained in the range space of the matrix $[B_{21}, B_{22}]$. For our model, this is the case and a solution of (2.18) is given by

$$\left| \begin{matrix} K_{11} \\ K_{21} \end{matrix} \right| = -[B_{21}, B_{22}]^{(1)} A_{21} + \Pi F$$

(2.19)

Here $[B_{21},B_{22}]^{(1)}$ is the generalized inverse of $[B_{21},B_{22}]$ F is an arbitrary 4x10 matrix, and Π is the projection onto the null space of $[B_{21},B_{22}]$.

With the matrices K_{11}, K_{12} given in 2.19, the closed loop dynamics are

$$\dot{x}_1 = ((A_1 + A_{11}) - [B_1 + B_{11},B_{12}][B_{21},B_{22}]A_{21} + [B_1 + B_{11},B_{12}]\Pi F)x_1$$

$$+ (A_{12} + B_{12}K_{22})x_2$$

$$\dot{x}_2 = (A_{22} + B_{22}K_{22})x_2$$

Now the two subsystems are effectively decoupled and we must determine if we have enough freedom to control the first subsystem. Thus, we must determine if the pair

$$\{(A_1 + A_{11})-[B_1 + B_{11},B_{12}][B_{21},B_{22}]A_{21}, \ [B_1 + B_{11},B_{12}]\Pi\} \quad (2.20)$$

is stabilizable. If stabilizability follows, we will select the arbitrary matrix F to stabilize the pair (2.20) and select the matrix K_{22} to stabilize the pair (A_{22}, B_{22}). The resulting overall closed loop stability is assured in that the interconnection terms may degrade the performance, but not destabilize the system.

3. Conclusion

This work has been a preliminary study of the attitude control of multibody satellites using large-scale systems techniques. Our efforts have centered on two methods that require less than full information sharing between controllers on different bodies of a two-body satellite. Both methods have disadvantages, the decomposition-aggregation approach is difficult to apply to systems which lack a decentralized input structure and the triangularization approach we used in Section 2.3 is too model specific.

In our further work, we will build on the examples in the report and try to develop a control technique more suitable to the attitude control with restricted information problem.

APPENDIX A

PLANAR MODEL OF VEHICLE

In this appendix we use Lagrangian mechanics to derive the small angle linearized equations of motion of the space construction base and logistics module. Motion is restricted to a single plane which is perpendicular to the behicle's velocity vector as shown in Figure 1. The vehicle is modelled as two rigid bodies connected by a spring hinge suspension. The flexible motion attributed to the solar wings attached to the base is modelled by superimposing the bending motion of two clamped free beams over the rigid body motion. The schematic of the physical system is shown in Figure 2. In the figure, the base structure is labelled rigid body 1 and the logistics module rigid body 2. The clamped free beams representing the solar wings are labelled A and B as shown.

To describe the motions, we use three reference frames:

c^0: (x_0, y_0) an inertial reference frame with origin at the nominal location of rigid body 1 center of mass.

c^1: (x_1, y_1) principal axis of rigid body 1

c^2: (x_2, x_2) principal axis of rigid body 2.

We denote unit vectors by i_x, i_y, etc.

We define the following reference points on the schematic.

cm_1: center of mass of rigid body 1 (solar wings are not included in calculation of this value).

cm_2: center of mass of rigid body 2.

1 : point where nominal plane of the solar wings intersects axis y.

2 : location of the suspension hinge.

and the following vectors,

$\bar{R}_o$: $xi_x + yi_y$: translational vector of cm.

$\bar{R}_{11}$: vector from cm to 1

$\bar{R}_{11_A}, \bar{R}_{11_B}$: vectors from 1 to points where wings connect to rigid body 1.

$\bar{R}_{12}$: vector from cm to 2

$\bar{R}_{22}$: vector from cm to 2

z: the distance along wings from the clamped end $0 < z < 1$

l ;   length of wings

$u(x,t)$:   bending displacement of wing A

$v(z,t)$:   bending displacement of wing B

θ_1:   angle between i_x, i_x

θ_2:   angle between i_x . i_x.

Kinematic Description

In order to refer the vector measured in reference frame e to the inertial frame we use the transformation T_1:

$$
\begin{array}{c} x_1 \\ y_1 \end{array}
\quad=\quad
\begin{array}{cc} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{array}
\quad
\begin{array}{c} x_o \\ y_o \end{array}
$$

and to refer a vector measured in reference frame c to the c frame, we use the transformation T_2:

$$
\begin{array}{c} x_2 \\ y_2 \end{array}
\quad=\quad
\begin{array}{cc} \cos\theta_2 & \sin\theta_2 \\ -\sin\theta_2 & \cos\theta_2 \end{array}
\quad
\begin{array}{c} x_1 \\ y_1 \end{array}
$$

Let the position of an arbitrary mass point in rigid body 1 referred to coordinate system c^1 be $\bar{\rho}_1$. Its position and velocity with respect to the inertial frame are

$$\bar{R}_1 = \bar{R}_0 + T_1^{T}\bar{\rho}_1$$
$$\dot{\bar{R}}_1 = \dot{\bar{R}}_0 + \dot{T}_1^{T}\bar{\rho}_1$$

Similarly, if $\bar{\rho}_2$ is an arbitrary mass point m rigid body 2 referred to coordinate system c, its position and velocity with respect to the inertial frame are

$$\bar{R}_2 = \bar{R}_0 + T_1^{T}(\bar{R}_{12})_2 + (T_2 T_1)^{T}(-\bar{R}_{22} + \bar{\rho}_2)_2$$

$$\dot{\bar{R}}_2 = \dot{\bar{R}}_0 + \dot{T}_1^{T}(\bar{R}_{12})_2 + (T_2 T_1)^{\dot{T}}(-\bar{R}_{22} + \rho_2)_2$$

At a distance z from the clamped end of solar wing A, the position and velocity of a mass point on the wing is

$$\bar{R}_A = \bar{R}_0 + T_1^{T}(\bar{R}_{11} + \bar{R}_{11A} + z i_{x_1} + u(z,t)i_{y_1})_1$$

$$\dot{\bar{R}}_B = \dot{\bar{R}}_o + \dot{T}_1^{T}(\bar{R}_{11} + \bar{R}_{11A} + z i_{x_1} + u(z,t)i_{y_1})_1 + T_1^{T}(\dot{u}(z,t)i_{y_1})_1$$

and similarly for wing B

$$\bar{R}_B = \bar{R}_0 + T_1^T(\bar{R}_{11} + \bar{R}_{11B} - zi_{x_1} + v(z,t)i_{y_1})_1$$

$$\dot{\bar{R}}_B = \dot{\bar{R}}_0 + \dot{T}_1^T(\bar{R}_{11} + \bar{R}_{11B} - zi_{x_1} + v(z,t)i_{y_1}) + T_1^T(\dot{v}(z,t)i_{y_1})_1$$

Kinetic Energy

In developing the Lagrangian equations of motion we take as our general'zed coordinates θ_1, θ_2, x, y, u, v. We assume that these quantities are small and, thus, in particular the coordinate transformations T_1 and T_2 can be approximated by

$$T_1 \doteq \begin{vmatrix} 1 & \theta_1 \\ -\theta_1 & 1 \end{vmatrix} \qquad\qquad T_2 \doteq \begin{vmatrix} 1 & \theta_2 \\ -\theta_1 & 1 \end{vmatrix}$$

The total kinetic energy of the vehicle is given by

$$\tilde{T} = \frac{1}{2}\{ \int_{Body\ 1} \dot{\bar{R}}_1 \cdot \dot{\bar{R}}_1 dm + \int_{Body\ 2} \dot{\bar{R}}_2 \cdot \dot{\bar{R}}_2 dm + \int_{Wing\ A} \dot{\bar{R}}_A \cdot \dot{\bar{R}}_A dm + \int_{Wing\ B} \dot{\bar{R}}_B \cdot \dot{\bar{R}}_B dm\}$$

Using the above kinematic expressions and assuming small deviations, we obtain

$$\tilde{T} = \frac{1}{2} m_T(\dot{x}^2 + \dot{y}^2) + \frac{1}{2} \hat{I}_1 \dot{\theta}_1^2 + \frac{1}{2} \hat{I}_2 \dot{\theta}_2^2 + \hat{I}_{12}\dot{\theta}_1\dot{\theta}_2$$

$$+ m_2 r_{22}\dot{x}\dot{\theta}_1 + m_2 r_{22}\dot{x}\dot{\theta}_2 + (r_{11}(m_B + m_A) - m_2 r_{12})\dot{\theta}_1\dot{y}$$

$$\text{(A1)}$$

$$+ \dot{y}\{ \int_{Wing\ A} \dot{u}\ dm + \int_{Wing\ B} \dot{v}\ dm\} + \dot{\theta}_1 r_{11}\{ \int_{Wing\ A} \dot{u}\ dm + \int_{Wing\ B} \dot{v}\ dm\}$$

$$+ \int_{Wing\ A} \dot{u}^2\ dm + \int_{Wing\ B} \dot{v}^2\ dm$$

here,

I_1: moment of inertia of rigid body 1

I_2: moment of inertia of rigid body 2

r_{11}, r_{12}, r_{22} r_{11A}, r_{11B}: lengths of vectors $\bar{R}_{11}$, $\bar{R}_{12}$, $\bar{R}_{22}$, $\bar{R}_{11A}$, $\bar{R}_{11B}$

m_A: mass of wing A

m_B: mass of wing B

m_1: mass of rigid body 1

m_2: mass of rigid body 2

m_T: total mass of body

$$\hat{I}_1 = I_1 + \int_{\text{Wing A}} (r_{11}^2 + (r_{11A} + z)^2) dm + \int_{\text{Wing B}} (r_{11}^2 + (r_{11B} + z)^2) dm + I_2 + m_2(r_{12} + r_{22})^2$$

$$\hat{I}_2 = I_2 + m_2 r_{22}^2$$

$$\hat{I}_{12} = I_2 + m_2(r_{22}^2 + r_{12}r_{22})$$

Potential Energy: The total potential energy of the system is the sum of the strain energy in the wings and in the hinge suspension.

$$\tilde{V}_{\text{Wings}} = \frac{1}{2} \int_{\text{Wing A}} EI \left(\frac{\partial^2 u}{\partial z^2}\right) dz + \frac{1}{2} \int_{\text{Wing B}} EI \left(\frac{\partial^2 u}{\partial z^2}\right) dz \tag{A2}$$

$$\tilde{V}_{\text{Hinge}} = \frac{1}{2} K_\theta \theta_2^2 \tag{A3}$$

here   EI is the beam stiffness coefficient

K_θ is the spring constant of the hinge

Equations of Motion

In order to obtain a finite dimensional set of dynamic equations, we will use the assumed modes method [6]. In this method we approximate the elastic displacements by a finite linear combination of mode shapes for the clamped free beam ultiplied by time dependent generalized coordinates. Thus, we have

$$u(z,t) = \sum_{j=1}^{n} \phi_{Aj}(z) q_j(t)$$

$$v(z,t) = \sum_{j=1}^{n} \phi_{Bj}(z) \rho_j(t)$$

where ϕ_{Aj}, ϕ_{Bj} are eigenfunctions of a clamped free beam, [7],

$$\phi_{.j} = A_j [(\sin\beta_j 1 - \sinh\beta_j 1)(\sin\beta_j z - \sinh\beta_j z)$$
$$+ (\cos\beta_j 1 - \cosh\beta_j 1)(\cos\beta_j z - \cosh\beta_j z)]$$

Here β_j are the eigenvalues of the clamped free beam and A_j are scale factors chosen such that

$$\int_0^1 \phi_{.j}(z)\phi_{.k}(z) dz = \begin{cases} 1 & j=k \\ 0 & j \neq k \end{cases}$$

For simplicity, we have only taken the first bending mode in our approximation. We also assume that the two solar wings are identical. Therefore,

$$u(z,t) = q(t)\phi(z)$$

$$v(z,t) = p(t)\phi(z)$$

$$\phi(z) = 1^{-1/2}[(\cosh \frac{1.875}{1} z - \cos \frac{1.875}{1} z) - 0.734(\sinh \frac{1.875}{1} z - \sin \frac{1.875}{1} z)]$$

$$0 < z < 1$$

If we substitute the expressions (A4), (A5) into the expressions for kinetic and potential energy (A1) (A2-A3), we can write $\tilde{T}$, $\tilde{V}$ as quadratic forms. If we define the 6x1 vector of generalized coordinates r_2,

$$r_2 = (x, y, \theta_1, p, q, \theta_2)^T$$

then

$$\tilde{T} = \frac{1}{2} \dot{r}_2^T M_2 \dot{r}_2 \qquad\qquad \tilde{V} = \frac{1}{2} \dot{r}_2^T K_2 r_2$$

where the 6x6 matrices M_2 and K_2 are defined by

$$M_2 = \begin{vmatrix} m_T & 0 & m_2 r_{22} & 0 & 0 & m_2 r_{22} \\ 0 & m_T & r_{11}m_B + r_{11}M_A & m_A \bar{\phi} & m_B \bar{\phi} & 0 \\ & & -m_2 r_{11} & & & \\ m_2 r_{22} & r_{11}m_B + r_{11}m_A & \hat{I}_1 & r_{11}m_A\bar{\phi} & r_{11}m_B\bar{\phi} & \hat{I}_{12} \\ & -m_2 r_{12} & & & & \\ 0 & m_A \bar{\phi} & r_{11}m_A\bar{\phi} & m_A & 0 & 0 \\ m_2 r_{22} & 0 & \hat{I}_{12} & 0 & 0 & \hat{I}_2 \end{vmatrix}$$

$$K_2 = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & m_A\omega_A^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & m_B\omega_B^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & K \end{vmatrix}$$

where
$$\tilde{\phi} = \int_0^1 \phi(z)dz = 0.783 \ 1^{-1/2}$$

$$\omega_A^2 = (\frac{1.875}{1})^4 \ EI/m_A$$

$$\omega_B^2 = (\frac{1.875}{1})^4 \ EI/m_B$$

Damping: With each flexible mode we associate some damping through the Rayleigh dissipation function $\tilde{R}$

$$\tilde{R} = \frac{1}{2} \dot{r}_2 D_2 \dot{r}_2$$

$$R_2 = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\omega_A \xi_A m_A & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\omega_B \xi_B m_B & 0 \\ 0 & 0 & 0 & 0 & 0 & \eta_\theta \end{vmatrix}$$

where ξ_A, ξ_B damping ratios of wings

D_θ damping factor of suspension

The Euler Lagrange equations of motion can now be written

$$\frac{d}{dt} \{ \frac{d}{d\dot{r}_2}(\tilde{T} - \tilde{V}) \} - \frac{d}{dr_2}(\tilde{T} - \tilde{V}) - \frac{d}{dr_2} \tilde{R} = Q$$

or in matrix form

$$M_2 \ddot{r}_2 + D_2 \dot{r}_2 + K_2 r_2 = Q$$

ere Q is the generalized force vector

$$Q = [q_1, q_2, q_3, q_4, q_5, q_6]$$

q_1:   x-resultant of external forces acting on vehicle

q_2:   y-resultant of external forces acting on vehicle

q_3:   resultant of all external torques about CM_1

q_4, q_5:   are the generalized forces acting on the first bending
modes of the wings, if

$$f^e(z,t)$$

is the external force acting on the wings, then

$$q_4 = \int_0^1 \phi_A(z) f_A{}^e(z,t) dz$$

$$q_5 = \int_0^1 \phi_B(z) f_B{}^e(z,t) dz$$

q_6:   resultant of all external torques about CM_2

In our work we will assume that the only external forces acting on the system
are control forces and torques acting directly on the central body generalized
coordinates. Thus,

$$M_2 \ddot{r}_2 + D_2 \dot{r} + K_2 r = T_2 v_2$$

where

$$T_2 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

and v_2 is a 4x1 control vector.

In Section 1 we also consider the equations of motion of the space base
alone. We obtain these equations in the same manner as (A5), simply delete
all terms corresponding to the module in the kinetic and potential energy
expressions. The equations of the base alone are:

$$M_1 \ddot{r}_1 + D_1 \dot{r}_1 + K_1 r_1 = T_1 v_1$$

C-4

where

$$M_1 = \begin{vmatrix} m_T & 0 & 0 & 0 & 0 \\ 0 & m_T & r_{11}m_B + r_{11}m_A & m_A\tilde{\phi} & m_B\tilde{\phi} \\ 0 & r_{11}m_B + r_{11}m_A & \hat{I}_1^1 & r_{11}m_A\tilde{\phi} & r_{11}m_B\tilde{\phi} \\ 0 & m_A\tilde{\phi} & r_{11}m_A\tilde{\phi} & m_A & 0 \\ 0 & m_B\tilde{\phi} & r_{11}m_B\tilde{\phi} & 0 & m_B \end{vmatrix}$$

$$K_1 = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & m_A\omega_A^2 & 0 \\ 0 & 0 & 0 & 0 & m_B\omega_B^2 \end{vmatrix}$$

$$D_1 = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\xi_A m_A\omega_A & 0 \\ 0 & 0 & 0 & 0 & 2\xi_B\omega_B m_B \end{vmatrix}$$

$$T_1 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}$$

APPENDIX B

PHYSICAL PARAMETERS

The following model parameters were calculated on the basis of information taken from reference [1]

$m_1 = 1218$ lbs.s^2/ft
$m_2 = 639$ "
$m_a = m_b = 26$ "

$r_{11} = 26.36$ ft.
$r_{12} = 21.54$ ft.

$I_1 = 163,000$ lb.s^2ft.
$I_2 = 247,000$ "
$\hat{I}_1 = 1,470,000$ "
$\hat{I}_2 = 152,000$ "
$\hat{I}_{12} = 346,000$ "

$\ell = 104$ ft.

The stiffness parameters were estimated to be:

$EI = 10^8$ lbs/in^2
$\omega_a = \omega_b = 7.79$ rad/s

$R_\Theta = 152,000$ ft lb/rad

The damping constants were estimated to be:

$\zeta_a = \zeta_b = 0.005$

$D_\Theta = 21,280$ ft lb s/rad

APPENDIX C

NUMERICAL RESULTS

In the course of this summer's work, only the decentralized linear quadratic regulator for the full model was implemented on the computer. A FORTRAN program DIRECT was written which solves the algebraic Riccati equation,

$$A^T P + PA - PBR^{-1}B^T P + Q = 0 \qquad\qquad (C1)$$

by direct integration. A FORTRAN subroutine CLOOP computes the feedback gain matrix

$$K = R^{-1}B^T P$$

and the closed loop system matrix

$$A - BR^{-1}B^T P$$

This matrix is then fed to the commerical package EISPACK which finds the eigenvalues and eigenvectors of the closed loop system.

The program DIRECT requires as input data: the system matrices, A, B, Q, R^{-1}; the number of states of the system, N; the number of controls, NB; the step size, EpS; and the number of steps, NT. The user must choose EpS and NT to yield steady state values of P. In our work, these parameters were determined by trial and error. The output of DIRECT is P, the solution of (C1); this is fed to CLOOP which, in turn, drives EISPACK. The output of the entire program is: A, B, Q, R; the value of P after some iteration N_1, (N_1 picked by the user); the final value of P after NT iterations; the closed loop system matrix and the closed loop eigenvalues and eigenvectors.

An example of the program is included in this appendix. In this example, A and B are as given in Table 1. The weighting matrices Q and R were chosen to be

$$Q = \text{diag.} \left[1, 1, 1, 0, 0, 1, 1, 1, 0 \ 0 \right]$$

$$R = 10^8 \times \text{diag} \left[1, 1, 1, 1 \right]$$

The twelve open loop eigenvalues of A are:

six zero eigenvalues (rigid body modes)

Double complex conjugate eigenvalues

at $-0.393.10^{-2} \pm j7.80$ (wing modes)

Complex conjugate eigenvalues (hinge modes)

$-0.259 \pm j1.91$

After the linear quadratic regulator control law was added to the system, eigenvalues were shifted to

- 9.0

- 5.62

- 1.01

- 1.02

- $0.755 \times 10^{-2} \pm j0.468 \times 10^{-1}$

- $0.226 \pm j1.48$

double complex conjugate roots at $-0.393.10^{-2} \pm j7.80$

Note that the uncontrollable wing modes are not shifted by feedback.

```
1.  C        THIS PROGRAM COMPUTES THE SOLUTION OF AN ALGEBRAIC
2.  C        RICCATI EQUATION BY DIRECT INTEGRATION
3.  C
4.          DIMENSION A(12,12),B(12,4),C1(12,12),C2(4,4)
5.         1,AT(12,12),BT(4,12),C2BT(4,12),T(12,12)
6.         1,D1(12,12),D2(12,12),D3(12,12),D4(12,12),P(12,12)
7.
8.  C        INITIALIZE ARRAYS
9.  C
10.         DO 1 L=1,12
11.         DO 1 M=1,12
12.         A(L,M)=0.0
13.         C1(L,M)=0.0
14.         P(L,M)=0.0
15.         AT(L,M)=0.0
16.         D1(L,M)=0.0
17.         D2(L,M)=0.0
18.         D3(L,M)=0.0
19.         D4(L,M)=0.0
20.         T(L,M)=0.0
21.  1      CONTINUE
22.         DO 2 L=1,12
23.         DO 2 M=1,4
24.         B(L,M)=0.0
25.         BT(M,L)=0.0
26.         C2BT(M,L)=0.0
27.  2      CONTINUE
28.         DO 3 L =1,4
29.         DO 3 M=1,4
30.         C2(L,M)=0.0
31.  3      CONTINUE
32.  C        READ IN DATA
33.         N=12
34.         NB=4
35.         NT=2000
36.         EPS=0.005
37.         READ(5,199) ((A(I,J),J=1,N),I=1,N)
38.         READ(5,199) ((B(I,J),J=1,NB),I=1,N)
39.         READ(5,199) ((C1(I,J),J=1,N),I=1,N)
40.         READ(5,199) ((C2(I,J),J=1,NB),I=1,NB)
41.  100   FORMAT(6E12.3)
42.  199   FORMAT(6F12.3)
43.  C
44.  C        WRITE OUT THE INPUT MATRICES
45.  C
46.         WRITE (6,99)
47.         DO 21 I=1,N
48.  21     WRITE(6,98)(A(I,J),J=1,N)
49.         WRITE (6,97)
50.         DO 22 I=1,N
51.  22     WRITE(6,98)(B(I,J),J=1,NB)
52.         WRITE(6,96)
```

```
      DO 23 I=1,N
23    WRITE (6,98) (C1(I,J),J=1,M)
      WRITE(6,95)
      DO 24 I=1,NB
24    WRITE(6,98) (C2(I,J),J=1,NB)
99    FORMAT(1°X,'THE A MATRIX',///)
98    FORMAT(5X,5(E13.5,2X))
97    FORMAT(10X,'THE B MATRIX',///)
96    FORMAT(1°X,'THE Q MATRIX',///)
95    FORMAT(1°X,'THE R MATRIX',///)

C
C     CONSTRUCT A TRANSPOSE
C
      DO 4 I=1,M
      DO 4 J=1,N
      AT(I,J)=A(J,I)
4     CONTINUE

C     CONSTRUCT B TRANSPOSE
C
      DO 5 I=1,NB
      DO 5 J=1,N
      BT(I,J)=B(J,I)
5     CONTINUE

C
C     CALCULATE C2*BT
C
      DO 6 I=1,NB
      DO 6 J=1,N
      DO 7 K=1,NB
      C2BT(I,J)=C2BT(I,J)+C2(I,K)*BT(K,J)
7     CONTINUE
6     CONTINUE

C
C     CALCULATE B*C2BT
C
      DO 8 I=1,N
      DO 8 J=1,N
      DO 9 K=1,NB
      I(I,J)=I(I,J)+B(I,K)*C2BT(K,J)
9     CONTINUE
t     CONTINUE

C
C     ITERATION OF THE MATRIX P
C
      DO 1J K=1,NT

C     INITIALIZE ALL SUMS FOR EACH K
C
      DO 3°, I=1,N
      DO 3C J=1,K
```

```
135.         D1(I,J)=2.0
136.         D2(I,J)=2.0
107.         D3(I,J)=0.0
108.         D4(I,J)=0.0
109.   30 CONTINUE
110. C
111.         DO 11 I=1,N
112.         DO 11 J=1,N
113.         DO 12 L=1,N
114.         D1(I,J)=D1(I,J)+A(L,I)*P(L,J)
115.         D2(I,J)=D2(I,J)+B(L,I)*A(L,J)
116.         D3(I,J)=D3(I,J)+T(I,L)*P(L,J)
117.   12 CONTINUE
118.   11 CONTINUE
119.         DO 13 I=1,N
120.         DO 13 J=1,N
121.         DO 14 L=1,N
122.         D4(I,J)=D4(I,J)+P(I,L)*D3(L,J)
123.   14 CONTINUE
124.         P(I,J)=P(I,J)+EPS*(C1(I,J)+D1(I,J)+D2(I,J)-D4(I,J))
125.   13 CONTINUE
126. C
127. C              WRITE THE VALUE OF P AT K=900
128. C
129.         IF (K.EQ.190D) GO TO 16
130.         GO TO 1
131.   16 WRITE(6,101) K
132.         DO 17 I=1,N
133.   17 WRITE (6,102) (P(I,J),J=1,N)
134.   1. CONTINUE
135. C
136. C              AT THE END OF ITERATION WRITE P
137. C
138.         WRITE (6,103)
139.         DO 18 I=1,N
140.   18 WRITE(6,102) (P(I,J),J=1,N)
141.  101 FORMAT(10X,"THE VALUE OF P FOR K=",I5)
142.  102 FORMAT(2X,12(E10.3,1X))
143.  103 FORMAT(10X,"THE FINAL VALUE OF P")
144.         CALL CLGCP(A,B,C2,P,N,NB,BT)
145.         STOP
146.         END
```

```
1.    C      THIS SUBROUTINE CONSTRUCTS THE FEEDBACK GAINS
2.    C      AND FINDS THE CLOSED LOOP EIGENVALUES
3.    C
4.    C
5.           SUBROUTINE CLOOP(A,B,C2,P,N,NB,BT)
6.           DIMENSION A(12,12),BT(4,12),C2(4,4),P(12,12),D(4,12)
7.          1,RBTP(12,12),ABRBTP(12,12),WR(12),WI(12),Z(12,12)
8.          1,BRBTP(12,12),B(12,4),FVI(12),IV1(12)
9.    C
10.   C      CONSTRUCT RBTP
11.   C
12.           DO 12 I=1,NB
13.           DO 12 J=1,N
14.   10      D(I,J)=0.
15.           DO 11 I=1,NB
16.           DO 11 J=1,N
17.   11      RBTP(I,J)=0.
18.           DO 12 I=1,NR
19.           DO 12 J=1,N
20.           DU 13 K=1,N
21.           D(I,J)=D(I,J)+BT(I,K)*P(K,J)
22.   13      CONTINUE
23.   12      CONTINUE
24.           DO 14 I=1,NR
25.           DU 14 J=1,N
26.           DO 15 K=1,NB
27.           RBTP(I,J)=RBTP(I,J)+C2(I,K)*D(K,J)
28.   15      CONTINUE
29.   14      CONTINUE
30.           WRITE(6,106)
31.           DU 5J J=1,NB
32.   5)      WRITE(6,102) (RBTP(I,J),J=1,K)
33.   C
34.   C      CONSTRUCT BRBTP
35.   C
36.           DO 16 I=1,N
37.           DO 16 J=1,N
38.   16      BRBTP(I,J)=0.0
39.           DU 17 I=1,N
40.           DO 17 J=1,N
41.           DO 18 K=1,NB
42.           BRBTP(I,J)=BRBTP(I,J)+B(I,K)*RBTP(K,J)
43.   18      CONTINUE
44.   17      CONTINUE
45.   C
46.   C      ADD A AND BRBTP
47.   C
48.           DO 19 I=1,N
49.           DO 19 J=1,N
50.           ABRBTP(I,J)=A(I,J)-BRBTP(I,J)
51.   19      CONTINUE
52.           WRITE(6,101)
```

```
53.          DO 49 I=1,N
54.  49      WRITE(6,102) (ABRBTP(I,J),J=1,N)
55.          NM=12
56.  C
57.  C       CALL EISPACK
58.  C
59.          CALL EIS(ABRBTP,WR,WI,Z,FV1,IV1,N,NM,IERR)
60.          WRITE(6,103)
61.          DO 25 I=1,N
62.  25      WRITE(6,104) WR(I),WI(I)
63.          WRITE(6,105)
64.          DO 30 I=1,N
65.  30      WRITE(6,102) (Z(I,J),J=1,N)
66.  101     FORMAT(5X,10X,"THE A MATRIX",//)
67.  102     FORMAT(2X,12(E10.3,1X))
68.  103     FORMAT(5X,10X,"THE EIGENVALUES OF A",//)
69.  104     FORMAT(5X,10X,E10.3,13X,E10.3)
70.  105     FORMAT(5X,"THE EIGENVECTORS OF A",//)
71.          RETURN
72.  106     FORMAT(5X,"THE FEEDBACK GAIN MATRIX",//)
73.          END
```

THE A MATRIX

THE B MATRIX

THE G MATRIX

THE P MATRIX

THE VALUE OF P FOR K= 1900

THE FINAL VALUE OF P

THE FEEDBACK GAIN MATRIX

THE A MATRIX

THE EIGENVALUES OF A

THE EIGENVECTORS OF A

.STOP

Plane of interest

Orbit plane

Velocity vector

FIGURE 1.
Orientation of Space Base.
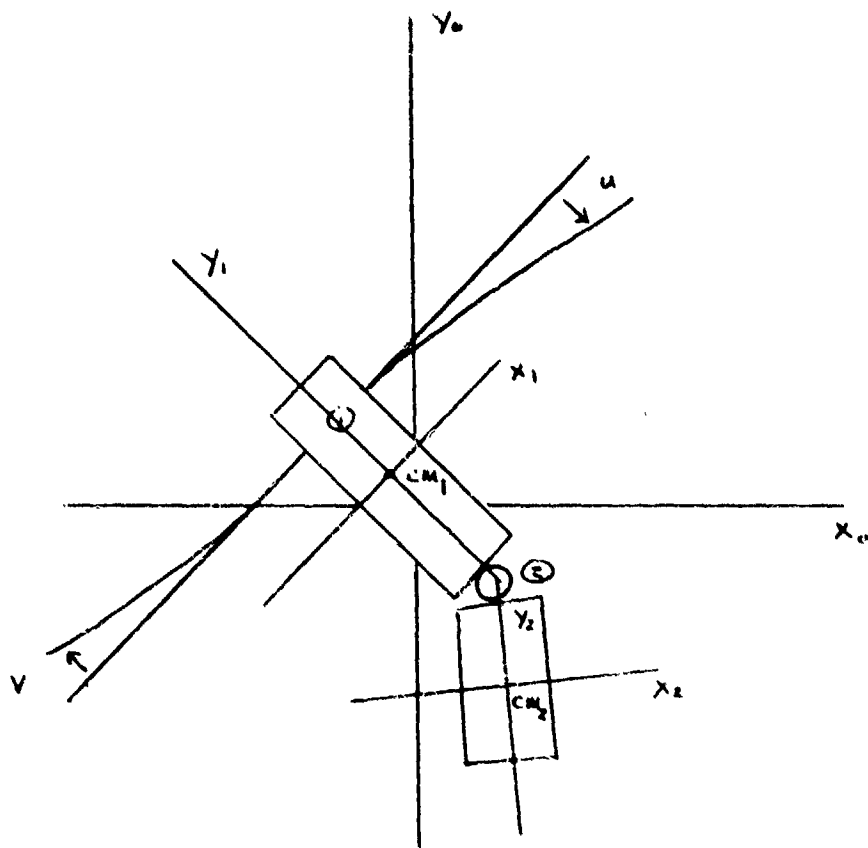
FIGURE 2.
Space Base Schematic.

THE MATRIX A (12×12)

```
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .100E 01    .0
00E 00    .000E 00    .300E 00    .000E 00    .000E 00    .000E 0
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .1
00E 01    .000E 00    .000E 00    .000E 00    .000E 00    .200E 0
      .000E 00    .000E 00    .300E 00    .000E 00    .000E 00    .000E 00    .0
00E 00    .100E 01    .000E 00    .000E 00    .300E 00    .000E 0
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .0
00E 00    .000E 00    .100E 01    .000E 00    .000E 00    .000E 0
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .0
00E 00    .000E 00    .000E 00    .100E 01    .000E 00    .000E 0
      .000E 00    .000E 00    .000E 00    .000E 00    .346E-01    .346E-01    .0
00E 00    .000E 00    .443E-05    .443E-05    .136E 02    .190E 0
      .000E 00    .000E 00    .000E 00    .601E-01    .601E-01    .000E 00    .0
00E 00    .000E 00    .707E-05    .707E-05    .545E 01    .763E 0
      .000E 00    .000E 00    .000E 00    .417E-02    .417E-02    .000E 00    .0
00E 00    .000E 00    .527E-06    .527E-06    .835E 00    .169E 0
      .000E 00    .000E 00    .000E 00   -.608E 02   -.674E-02    .000E 00    .0
00E 00    .000E 00   -.779E-02   -.865E-06   -.110E 01   -.154E 0
      .000E 00    .000E 00    .000E 00   -.674E-02   -.608E 02    .000E 00    .0
00E 00    .000E 00   -.865E-06   -.779E-02   -.110E 01   -.154E 0
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .0
00E 00    .000E 00    .000E 00    .000E 00    .000E 00    .100E 0
      .000E 00    .000E 00    .000E 00   -.114E-01   -.114E-01    .000E 00    .0
00E 00    .000E 00   -.146E-05   -.146E-05   -.371E 01   -.519E 0
```

THE MATRIX B (12×4)

```
    .000E 00    .000E 00    .000E 00    .000E 00
    .000E 00    .000E 00    .000E 00    .000E 00
    .000E 00    .000E 00    .000E 00    .000E 00
    .000E 00    .000E 00    .000E 00    .000E 00
    .000E 00    .000E 00    .000E 00    .000E 00
    .871E-03    .109E-03    .166E-04   -.895E-04
    .109E-03    .611E-03    .129E-04   -.359E-04
    .164E-04    .129E-04    .198E-05   -.549E-05
   -.219E-04   -.381E-04   -.261E-05    .723E-05
   -.219E-04   -.381E-04   -.261E-05    .723E-05
    .000E 00    .000E 00    .000E 00    .000E 00
   -.845E-04   -.359E-04   -.549E-05    .244E-04
```

STOP 0

TABLE 1
Matrix Coefficients.

THE MATRIX A1 (10×10)

```
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .100E 01    .0
00E 00    .000E 00    .000E 00    .000E 00
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .1
00E 01    .000E 00    .000E 00    .000E 00
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .0
00E 00    .120E 01    .000E 00    .000E 00
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .0
00E 00    .000E 00    .120E 01    .000E 00
      .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .000E 00    .0
00E 00    .000E 00    .000E 00    .100E 01
      .000E 00    .000E 00    -.731E-02   -.731E-02   .000E 00    .0
00E 00    .000E 00    -.937E-06   -.937E-06
      .000E 00    .000E 00    .433E-01    .433E-01    .000E 00    .0
00E 00    .000E 00    .555E-05    .555E-05
      .000E 00    .000E 00    .154E-02    .154E-02    .000E 00    .0
00E 00    .000E 00    .198E-06    .198E-06
      .000E 00    .000E 00    -.607E 02   -.336E-02   .000E 00    .0
00E 00    .000E 00    -.779E-02   -.431E-06
      .000E 00    .000E 00    -.336E-02   -.607E 02   .000E 00    .0
00E 00    .000E 00    -.431E-06   -.779E-02
```

THE MATRIX B1   (10×4)

```
      .000E 00    .000E 00    .000E 00
      .000E 00    .000E 00    .000E 00
      .000E 00    .000E 00    .000E 00
      .000E 00    .000E 00    .000E 00
      .000E 00    .000E 00    .000E 00
      .543E-03   -.229E-04   -.352E-05
     -.230E-04    .558E-03    .485E-05
     -.352E-05    .485E-05    .725E-06
      .463E-05   -.274E-04   -.977E-06
     -.463E-05   -.274E-04   -.977E-06
   %STOP% 0
```

TABLE 1 (cont.)

REFERENCES

1. Chichester, F,, et al: Space Construction Base Control System, Final Reports 1978 and 1979, NASA Contract NAS8-32660, The Bendix Corporation, Guidance Systems Division, Teterboro, New Jersey.

2. Siljak, D. D: Large Scale Dynamic Systems: Stability and Structure, Elsevier, North Holland, New York, 1978.

3. Kwakernaak, H. and Sivan, R: Linear Optimal Control Systems, Wiley, New York, 1972.

4. Kautus, M. L. J: "Stabilization, Controllability and Observability of Linear Autonomous Systems," Proc. of the Royal Academy of Science, Amsterdam, Ser. A, Vol. 1,73 (1970).

5. Sezer, M. E. and Siljak, D: "Robustness of Suboptimal Control," Proc. of the 19th IEEE Conf. on Decision and Control, 1980.

6. Meirovitch, L: Elements of Vibration Analysis, McGraw-Hill, New York, 1975.

7. Blevins, R: Formulas for Natural Frequency and Mode Shape, Nostrand Reinhold, 1979.

A STUDY OF THE GEOGRAPHIC COVERAGE
PROPERTIES OF A SATELLITE BORNE DOPPLER
LIDAR WIND VELOCITY MEASURING SYSTEM

Prepared by: Thomas H. Pate, Ph.D

Academic Rank: Assistant Professor

University and   Auburn University
Department: Department of Mathematics

NASA/MSFC
Division: Preliminary Design Office

MSFC Counterpart: Mr. J.W. Steincamp
                             Mr. R.B. Goss

Date: August, 1981

Contract No.:   NGT 01-008-021
                             The University of Alabama in Huntsville

A Study of the Geographic Coverage

Properties of a Satellite Borne Doppler

Lidar Wind Velocity Measuring System

by

Thomas H. Pate

Assistant Professor of Mathematics

Auburn University

Auburn, Alabama 36849

Abstract

The purpose of this investigation has been to measure geographic coverage frequency and geographic shot density for a satellite borne doppler lidar wind velocity measuring system. The equations of motion of the light path on the ground have been derived and a computer program has been devised to compute shot density and coverage frequency by latitude-longitude sections. Also, we have derived the equations for the coverage boundaries and devised a computer program to plot these boundaries thus making it possible, after an application of a map coloring algorithm, to actually see the areas of multiple coverage. We have also derived a theoretical cross-swath shot density function and have argued that this function gives close approximations in certain cases. This information should aid in the design of an efficient data-processing system for the doppler lidar.

0. Introduction

This report concerns a hypothetical satellite borne doppler lidar wind velocity measuring system. This system involves a satellite in circular orbit about the earth to which is attached a laser that is pointing toward the earth and rotating at a constant rate about an axis that is directed through the center of the earth. The rotation rate of the laser will be on the order of four rotations per minute. The purpose of the system is to estimate wind velocities at one kilometer intervals up to about twenty kilometers by measuring the doppler shift of the backscattered light. The firing rate of our hypothetical laser will be on the order of 8 firings per second.

As the satellite passes over an area of the earth the laser light traces a path along the ground. The boundaries of the covered region will be called swath boundaries and the region covered on a single passage will be called a swath. The path traced by the point on the earth directly below the satellite will be called the ground track of the satellite.

More specifically we have investigated the following geographic questions:

1. Given a particular section of the earth how many lasar pulses land there each day?

2. Given a particular section of the earth what is the pulse density in that section?

3. Given a particular section of the earth how many times per day is it within the coverage swath?

As a consequence of our analysis of these questions we have a computer program that answer questions (1) and (2) quickly and to a high degree of accuracy. Question (3) is somewhat tougher but the same program answers this question satisfactorily especially when the inclination angle of the orbit is between 75° and 110°. The results would be considerably improved if more memory were available.

1. Definitions and Notation

In our model we assume that the earth is a sphere of radius R where $R = 6378.15$ km. One half of the angle between the swath boundaries will be denoted by ψ (we note that ψ remains constant.) We impose upon our sphere a co-ordinate system whose origin lies at the center of the sphere and whose xy-plane contains the equator. We do not allow this co-ordinate system to rotate with the earth. Our orbital planes will always contain the x-axis. The inclination angle of the orbital plane with respect to the y-axis is denoted by β (or ZN in our computer program). The sphere is rotating counterclockwise about the z-axis at the rate of ω_0 rotations per day (counterclockwise with respect to an observer looking down from the North Pole.) The number of revolutions per day of the satellite per day is denoted by ω. In our computer program ω is assumed to be an integer such that $3 \leq \omega \leq 16$. The scan angle of the laser is denoted by α. One half the angular swath width along the equator is denoted by Γ. Clearly we have $\Gamma = \pi/\omega$. We let h denote the altitude of the satellite and let h_a denote the highest altitude at which complete coverage is desired.

We have analysed exclusively the case where the scan angle is chosen so that swaths for successive orbits are contiguous but non-overlapping at the equator at altitude $\alpha \geq 0$. In this case each point along the equator up to altitude a is covered at least twice during the course of a day. We let α_a denote the scan angle necessary to produce contiguous but non-overlapping coverage swaths at the equator up to altitude a.

2. Observations and Derivations of Simple Formulas

Consider diagram 1. We like to think of a, ω, and β as
parameters. The other quantities are determined by them. For example
we have

$$h = ((86400)^2 u/(2\pi\omega)^2)^{1/3} - R$$

where u = 398,601.6. Here R is in kilometers and ω is in
revolutions per day. The law of sines gives

$$(\sin(\alpha_a))/(R+a) = \sin(\alpha_a+\psi)/(R+h)$$

from which it follows that

$$\alpha_a = \text{Arctan}(\sin(\psi)/(((R+h)/(R+a)) - \cos(\psi))).$$

This gives α_a in terms of ψ but we do not as yet have ψ in terms
of the parameters β and ω. To obtain such a relationship consider
diagram 2. A normal to the orbital plane is (0, sin(β), - cos(β)).
This is also a normal to the plane determined by the swath tracks. A
normal to the plane containing the origon, the point (R,0,0) and
perpendicular to the orbital plane is (0, cos(β), sin(β)). The equation
of this plane is then

$$x\cdot 0 + y\cdot\cos(\beta) + z\cdot\sin(\beta) = 0$$

The plane determined by the swath boundary contains the point
(R cos(Γ), R sin(Γ.,0). The equation of this plane is therefore

$$\sin(\beta)\cdot y - \cos(\beta)\cdot z = R \sin(\Gamma)\sin(\beta) .$$

To determine the vector v we solve simultaneously the equations

$$\cos(\beta)\cdot y + \sin(\beta)\cdot z = 0$$

$$\sin(\beta)\cdot y - \cos(\beta)\cdot z = R \sin(\Gamma)\sin(\beta).$$

The result is

$$y = R \sin^2(\beta)\sin(\Gamma)$$

$$z = -R \sin(\Gamma)\sin(\beta)\cos(\beta).$$

To complete the determination of v we apply the restriction $x^2 + y^2 + z^2 = R^2$. We find that $x = R(1 - \sin^2(\Gamma)\sin^2(\beta))^{\frac{1}{2}}$. Since $\cos(\psi) = ((R,0,0) \cdot v)/R^2$ we must have

$$\psi = \text{Arccos}((1 - \sin^2(\Gamma)\sin^2(\beta))^{\frac{1}{2}}).$$

Since $\Gamma = \pi/\omega$ it is clear that ψ depends upon the altitude of the satellite. Note that $\lim_{\beta \to 0} \psi = 0$ and $\lim_{\beta \to \pi} \psi = 0$. This observation casts serious doubts upon the idea that swaths should be contiguous at the equator for small β.

We let $R(\alpha,a)$ denote the range at altitude a and scan angle α. (See diagram 1). Temporarily we let x denote $R(\alpha,a)$. The law of cosines then gives

$$(R+a)^2 = x^2 + (R+h)^2 - 2x(R+h)\cos(\alpha).$$

Solving for x we obtain

$$x = (R+h)\cos(\alpha) \pm ((R+a)^2 - (R+h)^2\sin^2(\alpha))^{\frac{1}{2}}.$$

There are two possible values for x because our triangle is not uniquely determined by the given information. After a visualization of the two possible triangles it is clear that we must have

$$R(\alpha,a) = (R+h)\cos(\alpha) - ((R+a)^2 - (R+h)^2\sin^2(\alpha))^{\frac{1}{2}}.$$

Clearly we must have $(R+h)\sin(\alpha) \leq R+a$ since otherwise the light beam would completely miss the earth.

Of interest here is the rate of change of $R(\alpha,a)$ with respect to a A computation shows that

$$\frac{dR}{da} = (-1)/(1 - (((R+h)/(R+a))^2\sin^2(\alpha)))^{\frac{1}{2}}.$$

Since $1 \leq (R+h)/(R+a) \leq 1 + (h/R)$ we have

$$\sec(\alpha) \leq |\frac{dR}{da}| \leq 1/((1 - (1+h/R)^2\sin^2(\alpha)))^{\frac{1}{2}}.$$

In case there is a maximum altitude a_{max} under consideration we have the improved inequality

$$(1 - ((R+h)/(R+a_{max}))^2\sin^2(\alpha))^{-\frac{1}{2}} \leq |\frac{dR}{da}| \leq (1-(1+h/R)^2\sin^2(\alpha))^{-\frac{1}{2}}.$$

We let $T(\alpha,a)$ denote the time required by the light to travel to altitude a and return to the satellite. Ignoring relativistic considerations we have

$$T(\alpha,a) = 2R(\alpha,a)/c$$

where c denotes the speed of light. Also, $\frac{dT}{da} = (2/c)\frac{dR}{da}$ so we have

$$(2/c)(1 - ((R+h)/(R+a_{max}))^2\sin^2(\alpha))^{-\frac{1}{2}} \leq$$

$$|\frac{dT}{da}| \leq (2/c)(1 - (1+h/R)^2\sin^2(\alpha))^{-\frac{1}{2}}.$$

This inequality is tight and seems to indicate that $\frac{dT}{da}$ is only negligably different from zero. In other words round trip times differ very little for the various altitudes.

3. Derivations of Parametric Representations for Ground Track, Swath Boundaries, and Light Path

In each of our definitions time (denoted by t and measured in days) is the parameter. The co-ordinate system described earlier is a fixed one, i.e.- it does not move with the earth. Our satellite will be directly above the point $(R,0,0)$ at time 0.

If $\beta = 0$ and the earth is not moving then the ground track would be $(R\cos(2\pi\omega t), R\sin(2\pi\omega t),0)$. To obtain the path in case $\beta \neq 0$ we simply rotate the orbital plane through angle β. The appropriate rotation matrix is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{bmatrix} .$$

The path along the Earth is then simply

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{bmatrix} \begin{bmatrix} R\cos(2\pi\omega t) \\ R\sin(2\pi\omega t) \\ 0 \end{bmatrix} = \begin{bmatrix} R\cos(2\pi\omega t) \\ R\cos(\beta)\sin(2\pi\omega t) \\ R\sin(\beta)\sin(2\pi\omega t) \end{bmatrix}$$

We let ω_0 denote the frequency of rotation of the Earth. In case time is measured in days $\omega_0 = 1$. To account for the rotation of the earth we simply multiply by the rotation matrix

$$\begin{bmatrix} \cos(2\pi\omega_0 t) & \sin(2\pi\omega_0 t) & 0 \\ -\sin(2\pi\omega_0 t) & \cos(2\pi\omega_0 t) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This rotation matrix rotates the earth clockwise (we are looking down from the North Pole) through the angle $2\pi\omega_0 t$.

The result of this multiplication is

$$x(t) = R \cos(2\pi\omega t) \cos(2\pi\omega_0 t) + R \cos(\beta)\sin(2\pi\omega t)\sin(2\pi\omega_0 t)$$

$$y(t) = -R \cos(2\pi\omega t) \sin(2\pi\omega_0 t) + R \cos(\beta)\sin(2\pi\omega t)\cos(2\pi\omega_0 t)$$

$$z(t) = R \sin(\beta)\sin(2\rho\omega t).$$

Application of some trigonometry produces the following alternate form for these equations:

$$x(t) = R \sin^2(\beta/2) \cos(2\pi(\omega + \omega_0)t) + R \cos^2(\beta/2)\cos(2\pi(\omega - \omega_0)t)$$

$$y(t) = -R(\sin^2(\beta/2) \sin(2\pi(\omega + \omega_0)t) + R \cos^2(\beta/2)\sin(2\pi(\omega - \omega_0)t)$$

$$z(t) = R \sin(\beta) \sin(2\pi\omega t).$$

We let $P(t)$ denote the vector $x(t)i + y(t)j + z(t)k$. Clearly we must have $\| P(t) \| = 1$ (here $\|\cdot\|$ denotes length) and $P(t) \cdot P'(t) = 0$ (the dot denotes dot product). This completes the derivation of the equations of motion of the point on the ground directly beneath the satellite.

To obtain a representation for the swath boundaries we proceed similarly. Let ψ denote the central swath angle. Consider the case where $\beta = \pi/2$ and pretend the Earth is not rotating. The ground track for one of the swath boundaries is then

$$(R \cos(\psi)\cos(2\pi\omega t), R \sin(\psi), R \cos(\psi)\sin(2\pi\omega t)).$$

(To obtain the second swath track we simply substitute $-\psi$ for ψ in all that follows or simply in the final result.) To obtain the track in case $\beta \neq \pi/2$ we simply rotate the plaine of the above tract (actually the shole sphere) through the angle $\beta - \pi/2$. The appropriate rotation matrix is then

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta-\pi/2) & -\sin(\beta-\pi/2) \\ 0 & \sin(\beta-\pi/2) & \cos(\beta-\pi/2) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sin(\beta) & \cos(\beta) \\ 0 & -\cos(\beta) & \sin(\beta) \end{bmatrix}$$

and the result is

$$\begin{bmatrix} 1 & 0 & \\ 0 & \sin(\beta) & \cos(\beta) \\ 0 & -\cos(\beta) & \sin(\beta) \end{bmatrix} \begin{bmatrix} R\cos(\psi)\cos(2\pi\omega t) \\ R\sin(\psi) \\ R\cos(\phi)\sin(2\pi\omega t) \end{bmatrix} =$$

$$\begin{bmatrix} R\cos(\psi)\cos(2\pi\omega t) \\ R\sin(\psi)\sin(\beta) + R\cos(\psi)\cos(\beta)\sin(2\pi\omega t) \\ -R\sin(\psi)\cos(\beta) + R\cos(\psi)\sin(\beta)\sin(2\pi\omega t) \end{bmatrix}$$

To account for the rotation of the earth we multiply by the rotation matrix

$$\begin{bmatrix} \cos(2\pi\omega_0 t) & \sin(2\pi\omega_0 t) & 0 \\ -\sin(2\pi\omega_0 t) & \cos(2\pi\omega_0 t) & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

The result is

$\hat{x}(t) = R\cos(\psi)\cos(2\pi\omega t)\cos(2\pi\omega_0 t) + R\cos(\psi)\cos(\beta)\sin(2\pi\omega_0 t)$

$\qquad\qquad\qquad\qquad + R\sin(\psi)\sin(\beta)\sin(2\pi\omega_0 t)$

$\hat{y}(t) = -R\cos(\psi)\cos(2\pi\omega t)\sin(2\pi\omega_0 t) + R\cos(\psi)\cos(\beta)\sin(2\pi\omega_0 t)$

$\qquad\qquad\qquad\qquad + R\sin(\psi)\sin(\beta)\cos(2\pi\omega_0 t)$

$\hat{z}(t) = -R\sin(\psi)\cos(\beta) + R\cos(\psi)\sin(\beta)\sin(2\pi\omega t).$

If we let $\hat{P}(t) = \hat{x}(t)\mathbf{i} + \hat{y}(t)\mathbf{j} + \hat{z}(t)\mathbf{k}$ denote the position vector for the swath boundary then an examination of the above equations reveals that

$$\hat{x}(t) = \cos(\psi)x(t) + R \sin(\psi)\sin(\beta)\sin(2\pi\omega_0 t)$$

$$\hat{y}(t) = \cos(\psi)y(t) + R \sin(\psi)\sin(\beta)\cos(2\pi\omega_0 t)$$

$$\hat{z}(t) = \cos(\psi)z(t) - R \sin(\psi)\cos(\beta).$$

We let $E(t)$ denote the vector

$$R \sin(\beta)\sin(2\pi\omega_0 t)i + R \sin(\beta)\cos(2\pi\omega_0 t)j - R \cos(\beta)k.$$

Simple computations show that $\| E(t) \| = R$ and that $P(t) \cdot E(t) = 0$. We may now write

$$\hat{P}(t) = \cos(\psi)P(t) + \sin(\psi)E(t)$$

thus expressing $\hat{P}$ as a linear combination of a pair of orthogonal vectors of length R. The other swath boundary is

$$\hat{\hat{P}}(t) = \cos(\psi)P(t) - \sin(\psi)E(t).$$ Since $E(t)$ is orthogonal to $P(t)$ it is clear that the vectors $P(t)$, $\hat{P}(t)$, and $\hat{\hat{P}}(t)$ are co-planar. In fact the points $P(t)$, $\hat{P}(t)$, and $\hat{\hat{P}}(t)$ all lie on the plane through the origon that is normal to the vector $P'(t)$.

We now turn to the problem of deriving a parametric representation for the lasar light path on the ground. Again we initially pretend that the earth is not moving. Our derivation will require two co-ordinate systems. The first of these is the standard system that we have been using. This system is regarded as being fixed. Our second co-ordinate system is fixed to the ground track and moves around the earth as the satellite moves. The x-axis of our moving coordinate system extends through the center of the earth - the positive direction being outward. The z-axis will be tangent to the direction of motion with the positive direction being in the direction of motion. The y-axis is tangent to the sphere and is oriented so that the moving co-ordinate system has the same orientation as our fixed system at the center of the earth.

At each instant of time there is a rigid motion that transforms standard co-ordinates to surface co-ordinates. (A rigid motion is a translation plus a rotation.)

Let $M(t)$ denote the orthogonal matrix

$$\begin{bmatrix} \cos(2\pi\omega t) & 0 & -\sin(2\pi\omega t) \\ \cos(\beta)\sin(2\pi\omega t) & \sin(\beta) & \cos(\beta)\cos(2\pi\omega t) \\ \sin(\beta)\sin(2\pi\omega t) & -\cos(\beta) & \sin(\beta)\cos(2\pi\omega t) \end{bmatrix}$$

This matrix rotates the non-standard co-ordinate system until its axes are parallel to and correctly oriented with the axes of the standard co-ordinate system. Our rigid motion is then

$$M(t)\begin{bmatrix} x^* \\ y^* \\ z^* \end{bmatrix} + P(t) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

In this case $P(t) =$

$$(R\cos(2\pi\omega t), R\cos(\beta)\sin(2\pi\omega t), R\sin(\beta)\sin(2\pi\omega t))$$

since the earth is not yet rotating.

We now parametrize the path of the laser on the earth with respect to the co-ordinate system fixed to the point directly below the satellite, (i.e. - the starred system). Let ω_ℓ denote the frequency of rotation of the laser. If we regard the spacecraft as being immobile then the path of the impact point (in the non-standard co-ordinate system) is

$$(R(\cos(\psi) - 1), R\sin(\psi)\cos(2\pi\omega_\ell t), R\sin(\psi)\sin(2\pi\omega_\ell t) .$$

The path of the impact point of the laser in standard co-ordinates is then

$$M(t) \begin{bmatrix} R\ (\cos(\psi) - 1) \\ R\ \sin(\psi)\cos(2\pi\omega_\ell t) \\ R\ \sin(\psi)\sin(2\pi\omega_\ell t) \end{bmatrix} + P(t)$$

To account for the rotation of the earth we multiply by the rotation matrix

$$Q(t) = \begin{bmatrix} \cos(2\pi\omega_0 t) & \sin(2\pi\omega_0 t) & 0 \\ -\sin(2\pi\omega_0 t) & \cos(2\pi\omega_0 t) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The final result is then

$$\overline{P}(t) = Q(t) \cdot M(t) \cdot \begin{bmatrix} R(\cos(\psi) - 1) \\ R\ \sin(\psi)\cos(2\pi\omega_\ell t) \\ R\ \sin(\psi)\sin(2\pi\omega_\ell t) \end{bmatrix} + Q(t)P(t) .$$

Explicit equations are

$$x(t) = R\cos(2\pi\omega_0 t)\overline{x}(t) + R\sin(2\pi\omega_0 t)\overline{y}(t)$$

$$y(t) = -R\sin(2\pi\omega_0 t)\overline{x}(t) + R\cos(2\pi\omega_0 t)\overline{y}(t)$$

$$z(t) = R\cos(\psi)\sin(\beta)\sin(2\pi\omega t) - R\sin(\psi)\cos(\beta)\cos(2\pi\omega_\ell t)$$
$$+ R\sin(\psi)\sin(\beta)\cos(2\pi\omega t)\sin(2\pi\omega_\ell t)$$

where

$$\overline{x}(t) = \cos(\psi)\cos(2\pi\omega t) - \sin(\psi)\sin(2\pi\omega t)\sin(2\pi\omega_\ell t)$$

and

$$\overline{y}(t) = \cos(\psi)\cos(\beta)\sin(2\pi\omega t) + \sin(\psi)\sin(\beta)\cos(2\pi\omega_\ell t)$$
$$+ \sin(\psi)\cos(\beta)\cos(2\pi\omega t)\sin(2\pi\omega_\ell t).$$

4. An Observation Concerning the Swath Boundaries
and Ground Tracks of the Satellite

We start with the ground track equations

$$x(t) = A \cos(2\pi(\omega+\omega_0)t) + B \cos(2\pi(\omega-\omega_0)t)$$

$$y(t) = -A \sin(2\pi(\omega+\omega_0)t) + B \sin(2\pi(\omega-\omega_0)t)$$

$$z(t) = R \sin(\beta)\sin(2\pi\omega t)$$

where $A = R \sin^2(\beta/2)$ and $B = R \cos^2(\beta/2)$. Observe that $z(t + 1/\omega)$ $= z(t)$, hence z has period $1/\omega$ (certainly no smaller number will work. We are interested in determining the response of $x(t)$ and $y(t)$ to the time increment $1/\omega$. Simple computations show that if we let M denote the rotation matrix

$$\begin{bmatrix} \cos(2\pi\omega_0/\omega) & \sin(2\pi\omega_0/\omega) \\ -\sin(2\pi\omega_0/\omega) & \cos(2\pi\omega_0/\omega) \end{bmatrix}$$

then

$$\begin{bmatrix} x(t+1/\omega) \\ y(t+1/\omega) \end{bmatrix} = M \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} .$$

Since M is the rotation matrix corresponding to a rotation through the angle $\frac{-2\pi\omega_0}{\omega}$ we see that we may obtain the entire ground track for the satellite by simply plotting a single passing and rotating the picture.

The equations for the swath boundaries are

$$x(t) = \cos(\psi)x(t) \pm R \sin(\psi)\sin(\beta)\sin(2\pi\omega_0 t)$$

$$y(t) = \cos(\psi)y(t) \pm R \sin(\psi)\sin(\beta)\cos(2\pi\omega_0 t)$$

$$z(t) = \cos(\psi)z(t) \pm R \sin(\psi)\cos(\beta).$$

We see that $\hat{z}(t+\frac{1}{\omega}) = \hat{z}(t)$. Also a consideration of the terms $\sin(2\pi\omega_0 t)$ and $\cos(2\pi\omega_0 t)$ shows that

$$\begin{bmatrix} \hat{x}(t+1/\omega) \\ \hat{y}(t+1/\omega) \end{bmatrix} = M \begin{bmatrix} \hat{x}(t) \\ \hat{y}(t) \end{bmatrix}$$

Thus swath boundaries also possess this rotational symmetry.

If we let $\overline{M}$ denote the 3×3 orthogonal matrix

$$\begin{bmatrix} \cos(2\pi\omega_0/\omega) & \sin(2\pi\omega_0/\omega) & 0 \\ -\sin(2\pi\omega_0/\omega) & \cos(2\pi\omega_0/\omega) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

then using the notation of the previous section we have

$$P(t+1/\omega) = \cos(\psi)\overline{M} \, P(t) + \sin(\psi)\overline{M} \, E(t) \ .$$

We now consider the response of the vector $\overline{P}$ to the time increment $1/\omega$. Recall that as time progresses $\overline{P}$ traces the light path along the ground. Recall also that

$$\overline{P}(t) = Q(t)\cdot M(t)\cdot V(t) + Q(t)\cdot P(t)$$

where
$$V(t) = \begin{bmatrix} R(\cos(\psi) - 1) \\ R\sin(\psi)\cos(2\pi\omega_\ell t) \\ R\sin(\psi)\sin(2\pi\omega_\ell t) \end{bmatrix}$$

and
$$P(t) = \begin{bmatrix} R\cos(2\pi\omega t) \\ R\cos(\beta)\sin(2\pi\omega t) \\ R\sin(\beta)\sin(2\pi\omega t) \end{bmatrix} \ .$$

The matrix $M(t)$ is defined in the previous section. Note that $M(t+1/\omega) = M(t)$. The function Q is a one-parameter group of rotation matrices; this implies that $Q(s)\cdot Q(t) = Q(s+t) = Q(t)\cdot Q(s)$. Since the matrix $\overline{M} = Q(1/\omega)$ we have $\overline{M}\cdot Q(t) = Q(t)\cdot\overline{M}$ and

$$\overline{P}(t+1/\omega) = Q(t+1/\omega) \cdot M(t+1/\omega)V(t+1/\omega) + Q(t+1/\omega) \cdot P(t+1/\omega)$$

$$= \overline{M} \cdot Q(t) \cdot M(t) \cdot V(t+1/\omega) + \overline{M} \cdot Q(t) \cdot P(t).$$

Note that this is nearly $\overline{M} \cdot \overline{P}(t)$. If ω_ℓ is an integral multiple of ω then $V(t+1/\omega) = V(t)$ and we have $\overline{P}(t+1/\omega) = \overline{M} \cdot \overline{P}(t)$. Thus in this case we have the same rotational symmetry that we had with the swath boundaries and the ground tracks. If we let RF denote the rotational frequency of the laser in rotations per minute then we have $\omega_\ell = 1440 \cdot RF$. Thus if $\omega = 8,9,10,12,15$ or 16 then ω_ℓ is an integral multiple of ω since each of the numbers $8,9,10,12,15$, and 16 divides 1440. (We assume RF is an integer). These observations involving symmetry are crucial to the algorithm that is used for computing shot frequency and shot density matrices. The symmetry implies that we can compute these matrices for the whole earth after gathering information from a single orbit. Even though the symmetry breaks down when ω_ℓ is not an integral multiple of ω the number of pulses is so large and the rotation rate of the laser so rapid that there is incurred a negligible loss of information by assuming this symmetry exists and in fact in the case of most importance, the case $\omega = 15$, the symmetry does exist.

5. Pulse Distribution along the South Width

It is fairly clear that the pulse distribution along the swath width is not uniform. In fact we expect greater pulse density toward the boundaries of the swath. Computer runs provide evidence of this. We claim that the distribution of pulses along the swath width is not substantially affected by any of the following:

(1) the motion of the earth,

(2) the inclination of the orbit,

(3) the motion of the satellite.

In other words, the pulse distribution along the swath width depends only upon the movement of the laser.

We consider a sphere of radius R and an angle ψ_m which is to represent the swath angle. We consider a point moving in a circle on our sphere with frequency of revolution ω_ℓ. If this point revolves about the x-axis of our standard co-ordinate system starting at the point $(R\cos(\psi_m), R\sin(\psi_m), 0)$ then the equations of motion are

$$x(t) = R\cos(\psi_m)$$
$$y(t) = R\sin(\psi_m)\cos(2\pi\omega_\ell t)$$
$$z(t) = R\sin(\psi_m)\sin(2\pi\omega_\ell t)$$

where revolution is counterclockwise with respect to a person looking down the x-axis in the negative direction. We will measure the pulse distribution in terms of the angle ψ (see diagram 3). Clearly we have $y = R\cos(\psi_m)\tan(\psi)$. Therefore we have

$$y = R\cos(\psi_m)\tan(\psi) = R\sin(\psi_m)\cos(2\pi\omega_\ell t)$$

which implies that

$$t = (2\pi\omega_\ell)^{-1} \text{Arccos}((\tan(\psi)/\tan(\psi_m)))$$

for $0 \le \psi \le \psi_m$. If we consider the time interval there we may divide $t(\psi)$ by $1/(4\omega_\ell)$ to obtain the comulative density function

$$\frac{2}{\pi} \text{Arccos}((\tan(\psi)/\tan(\psi_m))).$$

The pulse density we seek is then obtained by differentiating the above with respect to ψ. We call the resulting function L. The formula is

$$L(\psi) = \frac{2}{\pi} \sec^2(\psi)/(\tan^2(\psi_m) - \tan^2(\psi))^{\frac{1}{2}}$$

where $0 \le \psi < \psi_m$.

Inspection of this function reveals that we have the greatest pulse density when ψ is close to ψ_m - the swath angle. This is in line with the results obtained from computer runs.

This tendency for boundary clumping is significant and should be considered when the data processing system is designed.

6. Documentation for the Computer Programs

Included with this Report

There are listings for two programs included with this report. The first of these, Lasar III, was written by the author of this report while the second was written by K. Hiles of the Computer Services Corporation under the direction of the author of this report. The program language is Fortran 77 and neither program will run on a computer that understands only Fortran IV.

The plotting program causes the computer to sketch either a polar view of the swath boundaries for a single day or a polar view of the light path on the ground for a single day. The equations used are the ones presented in part 4 of this report. The user of the program must input β, the inciniation angle, ω, the number of orbits per day, and ω_ℓ the number of rotations of the laser each day. The user may either input the scan angle α or elect to have this angle computed so that swaths are contiguous at the equator. The plots produced by this program are usually quite pleasing aesthetically. Several examples are included with this report.

Each of the swath boundary plots produces a subdivision of the Northern Hemisphere into regions where the coverage number is constant. The coverage numbers for all of the regions may be obtained if all of the regions where the coverage number is least are known. This is true because two regions sharing a non-degenerate boundary arc must have coverage numbers that differ by one.

The program Laser III computes four matrices which as a group contain virtually all the raw data one could desire when studying the geographic coverage properties of a doppler lidar.

The input parameters for Lasar III are ß, ω, SF (the shot frequency of the Laser in shots per second), FRLQ (the rotational frequency of the laser in rotations per minute) and a pair of integers N amd M. The Northern Hemisphere is divided into sections by dividing the latitude range (0 to π/2) into N equal parts and the longitude range (0 to 2π) into M·ω equal parts. In each of the four output matrices rows correspond to latitude subdivisions and columns to longitude subdivisions. The upper rows correspond to the polar regions while the lower rows correspond to regions near the equator. The columns are in normal order.

The first output matrix is called the shot-frequency matrix. The ij-th entry is simply the number of lasar pulses landing in the corresponding ij-th section of the northern hemisphere.

The second matrix gives shot frequencies per unit area. The ij-th entry of this matrix is obtained by dividing the ij-th entry of the shot-frequency matrix by the area of the corresponding ij-th section. We measure area in 10,000 km^2.

The third and fourth matrices are closely related. The ij-th entry of the third matrix gives an upper bound on the number of times the corresponding ij-th region is within the coverage swath while the ij-th entry of the fourth matrix gives a lower bound on the number of times the corresponding ij-th region is within the coverage swath. In most cases the ij-th entry in the third matrix is the same as the ij-th entry in the fourth matrix. In these cases the coverage number

is determined exactly. If the ij-th entry of the third matrix differs from the ij-th entry of the fourth matrix then the swath boundary passes through the corresponding ij-th section of the northern hemisphere.

If the swath boundaries for a single passage over the Northern Hemisphere are not distinguishable in the array ARR (see program listing) then distortion will result at the latitude where this "running together" of boundaries occurs. This is evident in computer out put corresponding to $\beta \leq 50^o$. This can be corrected by incrasing the size of M and N thus making the grid network more fine. This distortion is only at lower latitudes (roughly 10° or less) and only in the third and fourth matrices. There is no distortion in the first and second matrices.

We print only the first M columns of each of our matrices simply because each of them repeats itself every M columns. Also, the write statements in the program are constructed so that if all of the columns of an output matrix will not fit on a single page then the remainder of the columns are printed on the next page and if the columns will not fit on two pages then the remainder are printed on the next page, etc. for as many pages as are necessary. If the rows will not fit on a single page they are automatically printed on the next page and these excess rows are printed before any excess columns are printed.

We do not go into the actual operation of the program itself. The algorithm is quite interesting and the speed of this program is surprising considering the amount of output obtained.
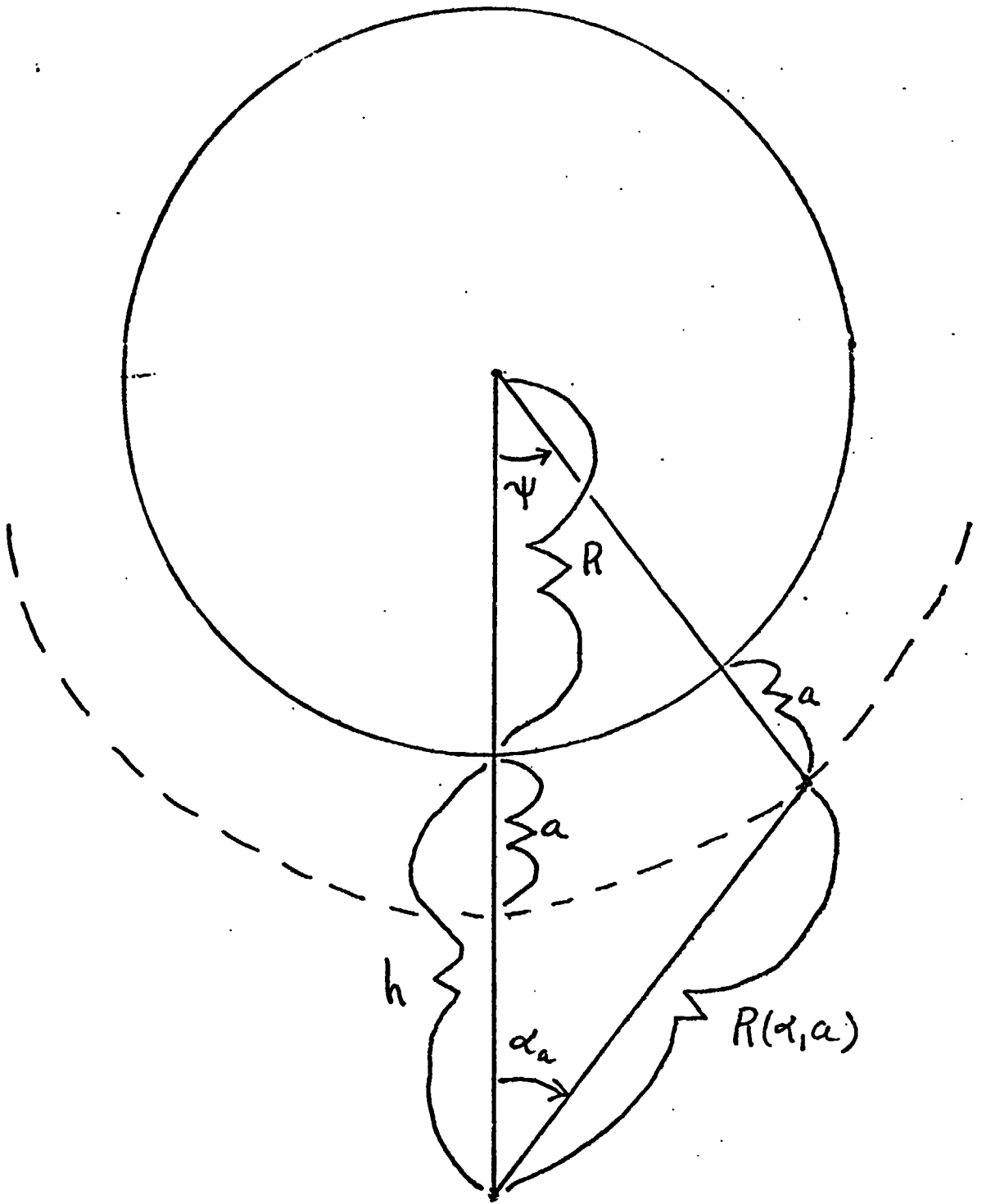
DIAGRAM 1
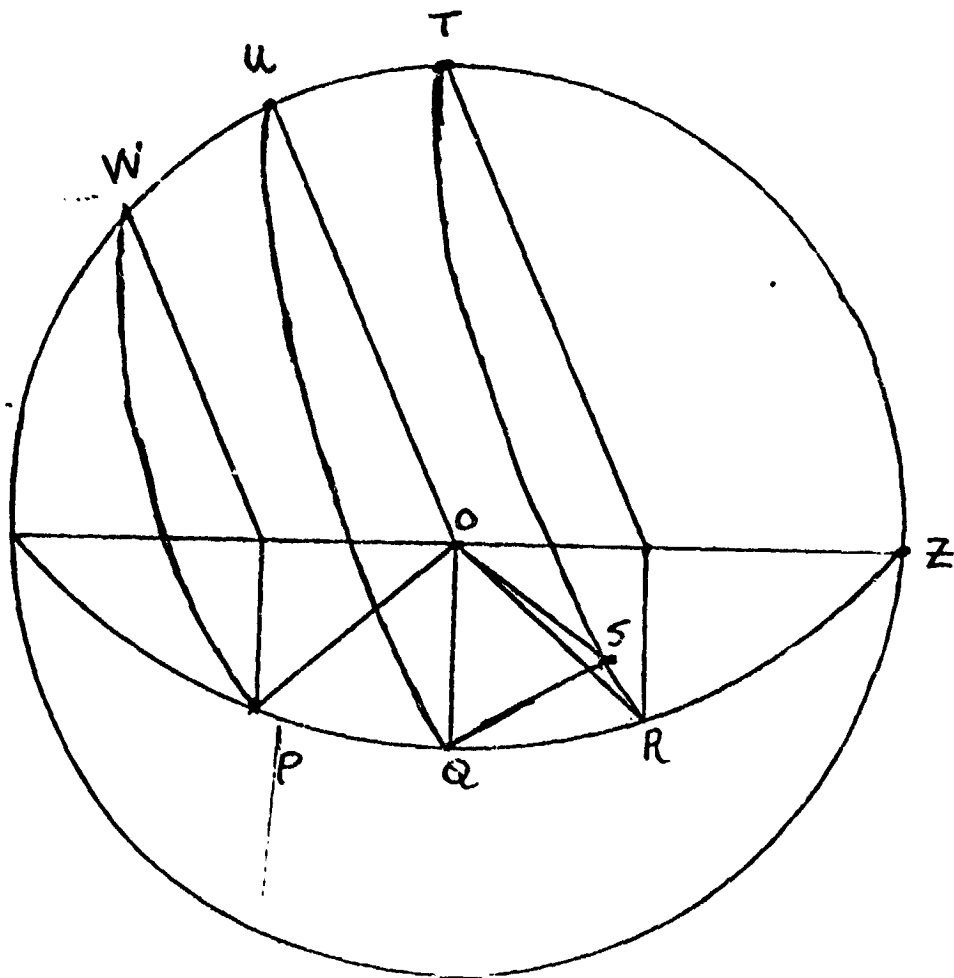
$$\psi = \angle QOS \quad , \quad \beta = \angle ZOU$$

$$\Gamma = \angle QOR = \angle QOP$$

The plane containing Q, O, and U is the orbital plane.

The vector v extends from O to S.

DIAGRAM 2

```
121.C
122.320      UE SET THE MATRICES ARR AND XY EQUAL TO 0
123.         CONTINUE
124.         DO 85 I=1,N
125.         DO 86 J=1,NSU
126. 86      ARR(I,J)=0.
127.         CONTINUE
128.         DO 87 K=1,R
129. 87      XY(I,K)=0.
130. 85      CONTINUE
131.C        UE SET EACH SWATH BOUNDARY ELEMENT EQUAL TO TWO
132.         KT = INT(1.26*U*FLOAT(M)*FLOAT(N))
133.         KS= MAX(25000,KT,ISTEP)
134.C        THE NUMBER OF TIME STEPS IS DEPENDENT UPON THE
135.CC       SIZE OF THE ARRAY AND THE SHOT FREQUENCY OF THE
136.CC       LASER BUT IS NEVER LESS THAN 25000
137.         ISTEP = KS
138.         DO 500 I=1,ISTEP
139.         T=01+02*FLOAT(I)
140.         Z1=A4*SIN(B1*T)
141.         Z=COS(SI)*Z1-SIN(SI)*A2
142.         IF(Z .LT. 0.)GO TO 510
143.         X1=-COS(B1*T)*COS(2*PI*T)+A2*SIN(B1*T)*SIN(2*PI*T)
144.         Y1=-COS(B1*T)*SIN(2*PI*T)+A2*SIN(B1*T)*COS(2*PI*T)
145.         X=COS(SI)*X1+SIN(SI)*A4*COS(2*PI*T)
146.         Y=COS(SI)*Y1+SIN(SI)*A4*SIN(2*PI*T)
147.         CALL LOCATE(I,J1,X,Y,Z,N,U,PI,M)
148.         ARR(I,J1)=2.
149.         Z=COS(SI)*Z1+SIN(SI)*A2
150. 510     IF(Z .LT. 0.) GO TO 520
151.         X=COS(SI)*X1 - SIN(SI)*A4*SIN(2*PI*T)
152.         Y=COS(SI)*Y1 - SIN(SI)*A4*COS(2*PI*T)
153.         CALL LOCATE(I,J1,X,Y,Z,N,U,PI,M)
154.         ARR(I,J1)=2.
155.520      CONTINUE
156.500      CONTINUE
157.C        THE BOUNDARIES HAVE BEEN COMPUTED
158.CC       UE NOW FILL-IN THE INTERIOR
159.         Q3 = 1./(2.*U*FLOAT(ISTEP))
160.         DO 530 I=0,ISTEP
161.         T = Q3*FLOAT(I)
162.         Z = A4*SIN(B1*T)
163.         X=-COS(B1*T)*COS(2*PI*T)+A2*SIN(B1*T)*SIN(2*PI*T)
164.         Y=-COS(B1*T)*SIN(2*PI*T)+A2*SIN(B1*T)*COS(2*PI*T)
165.         CALL LOCATE(K,L,X,Y,Z,N,U,PI,M)
166.         ARR(K,L)=1.
167.         CALL ISPY(K,L,ARR,M,N,U)
168.530      CONTINUE
169.C        UE COMPUTE THE UPPER AND LOWER BOUND COVERAGE MATRICES
170.         IQ = 1
171.         DO 531 K=1,N
172.         DO 532 L=1,M
173.         ZD=0.
174.         DO 533 L1=0,U-1
175.         L2 = L+L1*M
176.         ZE = ARR(K,L2)
177.         IF(ABS(ZE - 2.) .LT. .1) ZE = FLOAT(IQ)
178.         ZD=ZD+ZE
179.533      CONTINUE
180.         XY(K,L)=ZD
181.530      CONTINUE
182.531      CONTINUE
183.C 300    CONTINUE
184.C        THE FINAL COVERAGE MATRIX HAS BEEN COMPUTED
185.C        NOW UE WRITE THIS MATRIX
186.         WRITE(6,75)
187.         PRINT X,'COVERAGE FREQUENCY MATRIX'
188.         KLIP = INT((M-1)/24)
189.         KLAR = M-24*KLIP
190.         DO 810 I=1,KLIP
191.         K1=N-K+1
192.         WRITE(6,830)(XY(K1,L),L=(I-1)*24+1,I*24)
193.830      FORMAT(1X,24(F3.0,1X))
194.         WRITE(6,715)
195.828      CONTINUE
196.810      CONTINUE
197.         IF(KLAR .EQ. 0)GO TO 850
198.         K1=N-K+1
199.         WRITE(6,870)(XY(K1,L),L=24*KLIP+1,M)
200.870      FORMAT(1X,24(F3.0,1X))
201.860      CONTINUE
202.850      CONTINUE
203.         IF(IQ .EQ. 0)GO TO 310
204.         IQ=0
205.         GO TO 534
206.310      CONTINUE
207.         STOP
208.         END
209.SCAN:28
EOF:
```

```
SUBROUTINE LOCATE(I2,J2,X,Y,Z,M,U,PI,N)
EL=ASIN(Z)
I2=INT((EL*EMBEL)/PI)+1
IF(X .LT. 0.) THETA=ATAN(Y/X)+PI
IF(X .GT. 0. .AND. Y .GE. 0.)THETA=ATAN(Y/X)
IF(X .GT. 0. .AND. Y .LT. 0.)THETA=2*PI+ATAN(Y/X)
IF(X .EQ. 0. .AND. Y .GT. 0.)THETA=PI/2
IF(X .EQ. 0. .AND. Y .LT. 0.)THETA=(3*PI)/2
J2=INT((FLOAT(M)*THETA)/(2*PI))+1
RETURN
END
```

EOF :11
@ >.R1
NO CORRECTIONS APPLIED.
>RED.U ISPY
CASE UPPER ASSUMED
ED:ISR2-TUE-07/20/81-12:23:01-(1.2)
EDIT
@ >LMP1

```
1:      SUBROUTINE ISPY(K,L,ARR,M,N,U)
2:      DIMENSION ARR(100,361)
3:      MC = MSINT(U*.5)
4:      KNT = 0
5:      L1=L
6:      L=L+1
7: 910  KNT=KNT+1
8:      IF(KNT .GT. MC) L=L-MC
9:      IF(ABS(ARR(K,L)-Z.) .LT. .1) GO TO 920
10:     ARR(K,L)=1
11:     GO TO 910
12: 920  L=L1
13:     L=L-1
14: 930  IF(L .LE. 0) L=L+MC
15:     IF(ABS(ARR(K,L)-Z.) .LT. .1) GO TO 940
16:     ARR(K,L)=1
17:     GO TO 930
18: 940  K=K+1
19:     L=L1
20: 950  IF(K .GT. N) GO TO 960
21:     IF(ABS(ARR(K,L)-Z.) .LT. .1) GO TO 960
22:     ARR(K,L)=1
23:     GO TO 950
24: 960  RETURN
25:     END
```

EOF :88
@ >.3

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

THE DESIGN AND EVALUATION OF SUPERCONDUCTING CONNECTORS

Prepared By: James E. Payne, Ph.D.

Academic Rank: Associate Professor of Physics

University and Department: South Carolina State College
                                    Department of Natural Sciences

NASA/MSFC:
     Laboratory: Space Sciences
     Division: Space Physics
     Branch: Cryogenic Physics

MSFC Counterpart: Palmer N. Peters

Date: August 14, 1981

Contract No.:   NGT 01-008-021
                                    The University of Alabama
                                    in Huntsville

THE DESIGN AND EVALUATION OF SUPERCONDUCTING CONNECTORS

by

James E. Payne
Associate Professor of Physics
South Carolina State College
Orangeburg, South Carolina

ABSTRACT

The objective of this project is to develop a superconducting connector and determine if it will be reliable in super- conducting circuits on space flights. It is proposed that such connectors will be used between the superconducting readout loop and the SQUID magnetometer in the Gravity Probe B (GP-B) Experiment. Two types of connectors were developed. One type employs gold-plated niobium wires making pressure connections to gold-plated niobium pads. Lead-plated beryllium-copper spring contacts can replace the niobium wires. The other type is a rigid solder or weld connection between the niobium wires and the niobium pads. A description of the methods used to produce these connectors is given, as well as an analysis of their performance.

ACKNOWLEDGEMENTS

INTRODUCTION

In 1962, the late Dr. Leonard I. Schiff and Drs. Fairbank and Cannon of Stanford University proposed a new and unique test of Einstein's Theory of Gravitation, General Relativity. The experiment they proposed consisted of measuring the precession of ultra-precision gyroscopes in orbit about the Earth. Detailed descriptions of the gyroscope experiment, or Gravity Probe B Experiment as it is called, have been published in several places (1,2). The design of the experiment is such that advancements in several key technology areas have been required. The research efforts of many personnel at Marshall Space Flight Center and at Stanford University have been directed toward developing the necessary technology.

The gyroscope consists of a niobium-coated spherical quartz rotor that is floated electrostatically inside a quartz gyro housing. The superconducting niobium coating on the quartz ball establishes a magnetic field due to the rotation of the ball (3). A superconducting readout loop is deposited on the gyro housing as shown in Figure 1. Any change of alignment between the spin axis of the quartz ball and the readout loop will produce a current in the readout loop which will be monitored by a SQUID magnetometer.

OBJECTIVE

The objective of this work is to develop a method of making a superconducting connection between the SQUID magnetometer which is external to the gyro housing and the readout loops located in the housing. It was decided that two general types of connectors would be designed and undergo preliminary evaluation. A rigid connection and removable connectors would be evaluated by examining the current-voltage characteristics at liquid helium temperature.

I. RIGID CONNECTIONS

Experimental Considerations:

The most obvious method of making the connection to the read-out loops would be to either solder or weld the SQUID magnetometer leads to the readout loops. These connections must have zero resistance at the current levels involved in the GP-B experiment.

In joining the two superconductors care must be taken to prevent the formation of a tunnel junction. The superconducting material used in the experiment is niobium which readily forms niobium oxide. This niobium oxide must be removed before connecting the superconductors to prevent the formation of a superconductor-insulator-superconductor junction at the connection (4). If a solder joint is used, a superconducting solder with an acceptable transition temperature should be used.

Although an ideal method of connecting the readout loops to the SQUID magnetometer would be a weld joint, it is believed that the high currents required to produce the weld could adversely effect the niobium film and its superconducting properties.

Experimental Procedure:

The samples were prepared by soldering niobium wire (0.005-inch diameter, 99.99% pure) to a niobium film as shown in Figure 2. A 50% Sn - 50% In solder which had a transition temperature greater than 4.2° K was used for the connection. The niobium RF-sputtered onto quartz substrates for 7 to 10 minutes producing films in the 70 nm to 100 nm thickness range. These samples were produced before work on superconducting connectors was initiated. The films were returned to the sputtering chamber to sputter-etch the niobium oxide off the surface before sputtering on a 30 nm layer of gold to prevent further oxidation of the niobium. The gold layer was thin compared to the thickness of the niobium film to maintain superconductivity across the normal gold layer (4, 5, 6). The gold-plated niobium films were sputter-etched to produce the horseshoe shaped pattern shown in Figure 2. This pattern was chosen instead of a complete readout loop to simplify the experiment. The niobium films were tested at 4.2° K to insure that they were superconducting before being used in the connector test. Niobium wires (.005-inch diameter) were soldered to the niobium film using a 50% Sn - 50% In solder. A four-contact current-voltage connection was made to the sample and it was placed in liquid helium. The current-voltage characteristics of the sample were plotted on an X-Y plotter with a .05 mv/in maximum sensitivity. Although a measurement of temperature was not available, the sample could be warmed by moving it from the liquid into the vapor above. Attempts were made to produce this type of connection by welding the niobium wires to the niobium films. These efforts failed to produce satisfactory samples.

Experimental Results:

The results of three typical samples are shown below.

Sample	Critical Current at 4.2° K
1	200 mA
2	370 mA
3	410 mA

The critical current was taken as the onset of a measurable voltage (2 μV) and may vary due to wire length or location of the solder joints on the niobium film. The current-voltage curve is nonlinear after the onset of a voltage until the samples goes abruptly normal.

II. REMOVABLE CONNECTORS - NIOBIUM WIRES

Experimental Considerations:

Many of the same considerations are involved in the development of a removable connector as with the rigid connector. In addition, gold was plated onto the niobium wires that served as the other half of the connector. This meant that the effect of two gold layers and the possibility of oxide on two niobium surfaces had to be considered. With the removable connector, the problems associated with making and breaking the connection must be considered along with the effects on a pressure contact due to repeated cycling between room temperature and liquid helium temperature .

Experimental Procedure:

The same horseshoe shaped film (Figure 2) was used to replace the readout loops but niobium wires (either 0.003-inch diameter or 0.005-inch diameter) made the connection instead of the solder joint. There were two configurations for these wires. In one the wires passed through holes in a quartz plate and in the other the wires were glued to the plate edges (Figure 3). The quartz plate with the niobium wires was placed over the niobium film so that the wires contacted the tabs (Figure 2) and a pressure contact was made by securing a piece of stainless steel on each side of the sample (Figure 4). The resistance of the connection was measured at room temperature and liquid nitrogen temperature. If the sample maintained its integrity, measurements were made on the sample in liquid helium as was done on the solder joints. Each sample was then connected and disconnected several times, cycled between room temperature and liquid nitrogen temperature numerous times and allowed to remain connected for periods of several hours to two weeks before the helium temperature measurements were repeated.

Experimental Results:

The results on five samples of this type indicate that a
limited number of connections of the sample and extended
(two-week) connections had no measurable influence on the
sample characteristics. In general, the critical currents of
these connections were much lower than the solder joints
(4, 5, 6). A typical current-voltage curve would exhibit a
critical current in the microamp range, a linear region, and
then several distinct transition steps to the normal state.
The resistance of the linear region was down by three orders
of magnitude from the normal resistance. The critical currents
for four connectors using different diameter wire connected
to the same film are shown below.

Sample	Critical Current
.003" diam. Nb through holes	45 µA
.003" diam. Nb glued to edge	175 µA
.005" diam. Nb glued to edge	545 µA
.005" diam. Nb glued to edge	420 µA

These preliminary results indicate that the larger the wire
diameter and more rigidly the wire is held in place, the
better the connector.

III. REMOVABLE CONNECTOR - LEAD-PLATED STRIPS

A connector using lead plated onto beryllium-copper in place
of the niobium wires was also examined with less success than
the previous type of connection. This arrangement did produce
higher critical currents (10 mA) connectors, but they required
extremely thick layers of lead. The connectors did not maintain
their integrity during temperature cycling or over extended
periods of time. Although this method did not produce accept-
able connectors, it did produce superconducting lead films
from a plating solution (7).

CONCLUSIONS AND RECOMMENDATIONS

This preliminary study indicates that there are several methods
for making superconducting connections from the gyro readout
loops to the SQUID magnetometer. Either the rigid connector
or the removable connector could be incorporated into the
design of the housing. Before either method is used, further
investigation would be required. Whiskers (single crystals of
solder material) could cause shorts between the solder joints

unless there is su'ficient isolation. Vibrational and aging
studies should be undertaken and it would be desirable to
replace the standard four-contact current-voltage measurements
with a persistent current measurement using a superconducting
coil connected to the sample under investigation. It appears
that a rigid connection would be more reliable and satisfactory,
but if lower critical current values can be tolerated, then
the removable connector is a possibility.

EXTERNAL WIRES

SQUID MAGNETOMETER

CONNECTING PADS

READOUT COIL
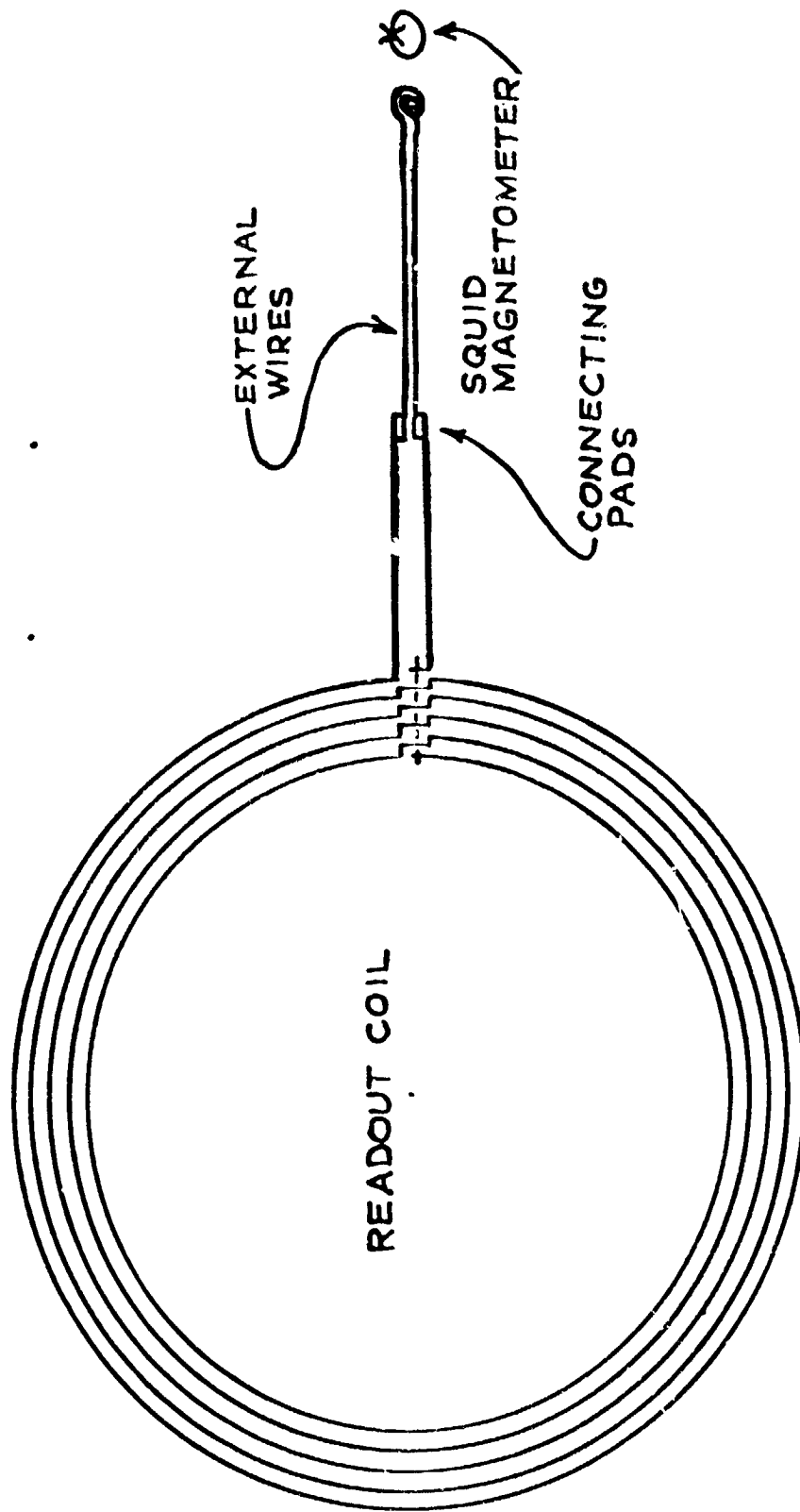
Figure 1: GP-B Readout
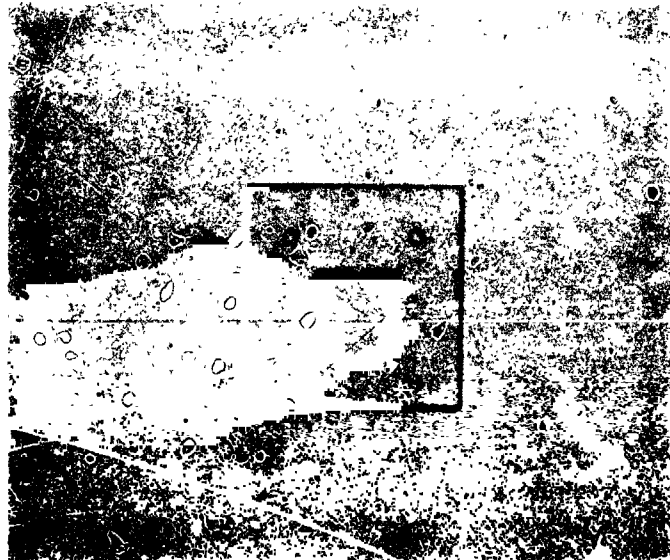
Figure 2:   Gold-Plated Niobium Film

XVII-7

Figure 3: Removable Connectors

XVII-8
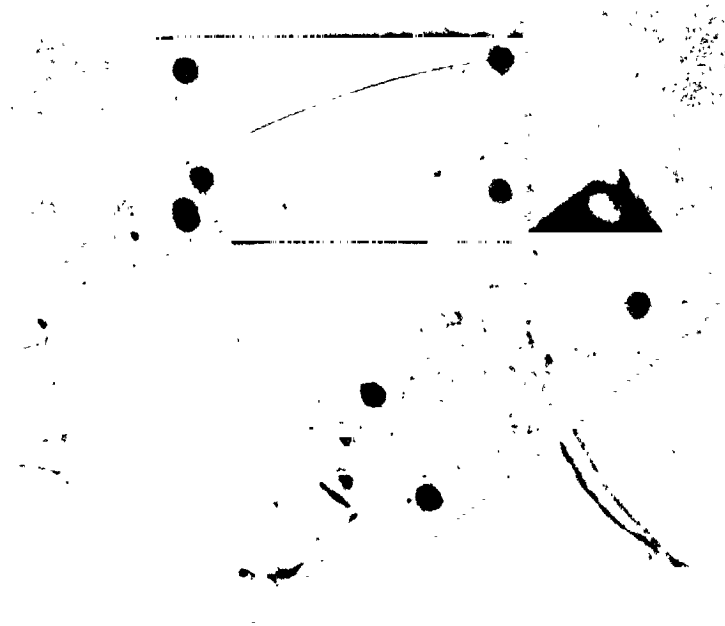
Figure 4: Pressure Plates

XVII-9

REFERENCES

1. Fairbank, W.M. et al: Final Report. NASA Grant 05-020-019, W.W. Hansen Laboratories of Physics and the Department of Aeronautics and Astronautics, Stanford University, July 1977.

2. Fairbank, W.M. et al: Report on Progress of NASA Grant 05-020-019; U.S. Air Force Contract F36615-67-C-1245 and NASA Contract NAS8-32355; W.W. Hansen Laboratories of Physics and the Department of Aeronautics and Astronautics, Stanford University, June 1980.

3. Gravity Probe B Phase A Report; George C. Marshall Space Flight Center, Marcy 1980.

4. Solymar, L. Superconducting Tunneling and Applications, Wiley-Interscience, New York, 1972.

5. Clark, J., J. Phys. Supp. au. No. 2-3, Tome 29, c2 (1968).

6. de Gennes, P.G. Superconductivity of Metals and Alloys. W.A. Benjamin, New York, 1966.

7. Momose, T., et al. J. Appl. Phys. 52(2), February 1981, p. 909.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

A STUDY OF OPTICALLY CONTACTED QUARTZ AT CRYOGENIC TEMPERATURES

Prepared By: Linda L. Payne, Ph.D

Academic Rank: Associate Professor of Physics

University and Department   South Carolina State College
                                      Department of Natural Sciences

NASA/MSFC:
    Laboratory: Space Sciences Laboratory
    Division: Space Physics Division
    Branch: Cryogenic Physics Branch

MSFC Counterpart: Palmer N. Peters

Date: August 14, 1981

Contract No.:   NGT 01-008-021
                                      The University of Alabama
                                        in Huntsville

A STUDY OF OPTICALLY CONTACTED
QUARTZ AT CRYOGENIC TEMPERATURES

by

Linda L. Payne
Associate Professor of Physics
South Carolina State College
Orangeburg, South Carolina

ABSTRACT

The Gravity Probe-B (GP-B) Experiment proposed by Stanford
University to test the theories of general relativity requires
extremely precise measurements. The quartz components of the
instruments to make these measurements must be held together
in a very stable unit. Optical contacting has been suggested
as a possible method of joining these components. The objective
of this research is to investigate the strength of the bond
between optically contacted fused quartz surfaces at liquid
helium temperature. A test apparatus has been designed which
can be used for making measurements of the tensile strength of
these bonds. Results of these tensile pull tests are given and
the reliability of such bonds analyzed.

ACKNOWLEDGEMENTS

INTRODUCTION

In 1962, the late Dr. Leonard I. Schiff and Drs. Fairbank and
Cannon of Stanford University proposed a new and unique test
of Einstein's Theory of Gravitation, General Relativity. The
experiment they proposed consisted of measuring the precision
of ultra-precision gyroscopes in orbit about the Earth. Detailed
descriptions of the gyroscope experiment, or Gravity Probe B
experiment as it is called, have been published in several
places (1,2). The design of the experiment is such that
advancements in several key technology areas have been required.
The research efforts of many personnel at Marshall Space Flight
Center and at Stanford University have been directed toward
developing the necessary technology.

The gyroscopes for the experiment and the telescope which will
be used to aim the spacecraft will be made of fused quartz.
The components of these instruments must be joined together
very rigidly and maintain their relative positions under
conditions of reduced temperature and pressure. One method
which has been proposed for joining the quartz components is
the method of optically contacting the surfaces.

OBJECTIVE

The objective of this research is to investigate the strength
of the bond between optically contacted fused quartz surfaces
at liquid helium temperature (4.2° K).

THEORY

It has been known for more than fifty years that very flat,
clean glass or silica surfaces can be brought into optical
contact, that is contact so close that reflection at the
interface is very nearly eliminated (3). When the surfaces
are contacted, they are held together by considerable force
with no adhesive or clamping. The origin of this force is not
clearly understood. It is probably the result of a number of
bonding forces which contribute in varying degrees. Some of
these may interact with one another and some may be influenced
by external parameters. Possible contributors to the strength
of the bond between optically contacted surfaces are Van der
Waals forces, chemical bonds associated with unfilled electronic
shells, bond structures in the energy levels appearing within
solids, surface states on the interfaces, possibly free and
polarization charge contributions, and capillary forces associ-
ated with surface tensions of impurity films (4).

From a study of the nature of these forces, it becomes apparent that in order for optical contacting to occur, the surfaces must be very flat (within a few hundred Angstroms) and very clean. These requirements must be considered when preparing the optical contacts for testing.

EXPERIMENTAL PROCEDURE

The fused quartz pieces which were obtained for use in this research were of two shapes; most were cylindrical pieces of quartz with a one-inch diameter and either 0.5 inch or 0.65 inch in height; a few were rectangular pieces of quartz one-inch by 0.75 inch and 0.5000 inch or 0.5125 inch in height. The surface of the quartz had been polished to within one microinch (250 Å or 25 nm).

The quartz surfaces were carefully cleaned using ethyl alcohol and a lint-free linen cloth. The surfaces were rinsed with ethyl alcohol and blown dry with air. The surfaces were brought into contact and gently rotated over one another. As the contact force between them increased, the quartz surfaces were quickly aligned. If a rigid contact occured before the final alignment, the pair was heated with a heat gun to break the bond. The cleaning procedure was then performed again and the surfaces re-contacted.

Several problems were encountered in designing a test apparatus to pull the optical contacts apart. One problem was due to the difference in thermal expansion properties of the fused quartz and stainless steel. It was found that when the quartz was glued directly to the stainless steel with DOW RTV-731 Silastic Sealant/Adhesive and the assembly cooled to liquid helium temperatue the quartz would fracture due to this difference in thermal contraction rates of the materials. To alleviate this problem, the quartz contacts were glued to pieces of quartz which served as spacers. The spacers were glued to the pull rod assembly via stainless steel braid as shown in detail in Figure 1. The RTV glue joints, although not extremely strong at room temperatures, proved to be an extremely good adhesive at 4.2° K. The best glue joints were those which employed a minimum amount of RTV on very smooth, oil free surfaces.

The purpose of the quartz rotor shown in Figure 1 was to enable the system to rotate slightly so that when the force was applied to pull the contacts apart it would not have a shear component. Figure 2 shows a picture of the sample assembly with the cover in place. This cover attaches to the outer stainless tube. The inner stainless steel tube is attached via the braid to the top of the optical contact. As

the nut shown in Figure 3 is tightened, the inner tube moves up and exerts a force to pull the contacts apart. The force applied is measured with the load cell located as shown in Figure 3. The output from the load cell is monitored with a BLH Strain Indicator, Model 120C. Figures 4 and 5 show calibration curves for two load cells using this instrument.

After a pair of optically contacted quartz flats were glued in the sample mount, the RTV joints were allowed to harden for at least 24 hours. After this time had elapsed, it was immersed in liquid nitorgen for precooling and then in liquid helium for the pull test. The load was slowly increased until the contacts or another joint failed. The braid tended to stretch and slip some causing the contacts to experience at times some sudden shocks.

A total of sixteen pull tests were conducted. Five of these tests resulted in the optical contact breaking under very small loads (less than forty pounds). Each of these contacts had some cloudy regions at the interface which indicated they lacked good optical contact at that point. The other eleven pull tests were performed on contacts which were almost free of any cloudy regions of non-contact. In these tests, more often than not, another type of failure would occur before failure of the optical contact. Table I gives a summary of the results of these pull tests. The values for load and pressure which the contact joint could withstand would exceed those shown in the table, since the values listed are the values at which another part failed.

CONCLUSIONS AND RECOMMENDATIONS

The results of this preliminary study on the strength of the optical contact bond indicates that it far exceeds the expectations, with a value above 662 psi. Indications from the work are that the optical contacts must be of very good quality, exhibiting almost no cloudy areas of reflection at the interface. The system is now capable of exerting up to 1000 pounds of tensile pull on the contact pair. Recommendations would be to obtain more quartz pieces and have extensive preliminary measurements made on the flatness and surface finish of the pieces. After the carefully matched quartz pieces are optically contacted, non-destructive tests should be performed to characterize the interface. Tensile pull tests could then be conducted on these contacts whose characteristics would have been fully documented. It is recommended that a large number of contacts be tested at liquid helium temperature to determine if the strength of the bond as measured in this research is a reliable value to assume.

Stainless Tube
To Load Cell

↑
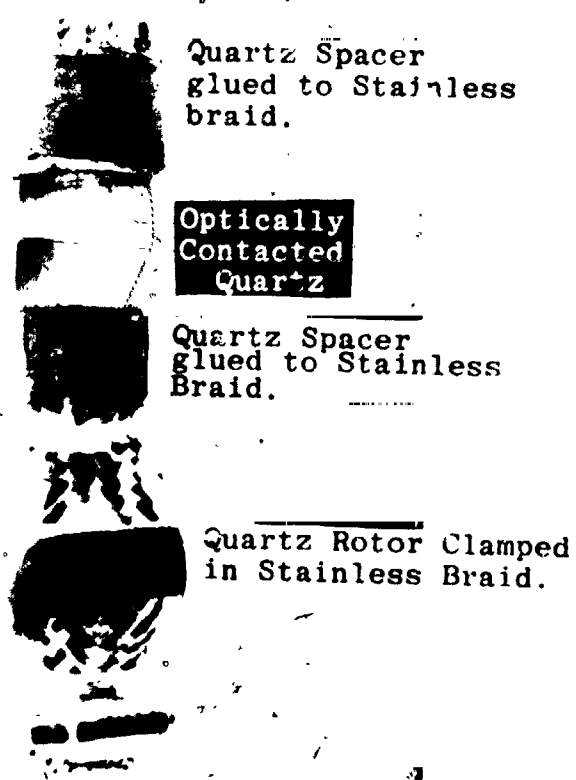
Quartz Spacer
glued to Stainless
braid.

Optically
Contacted
Quartz

Quartz Spacer
glued to Stainless
Braid.

Quartz Rotor Clamped
in Stainless Braid.

Figure 1:   Lower Pull Test Assembly

To Load Cell

Stainless Shield

Optically C ntacted
Quartz

Quartz Rotor
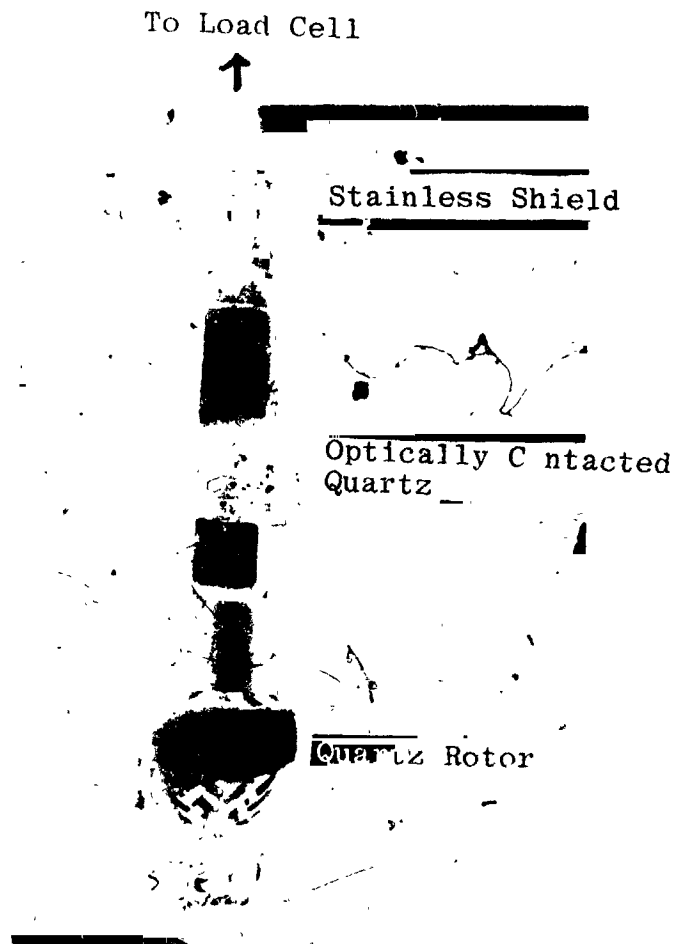
Figure 2:   Pull Test Assembly With Shield

XVIII-5
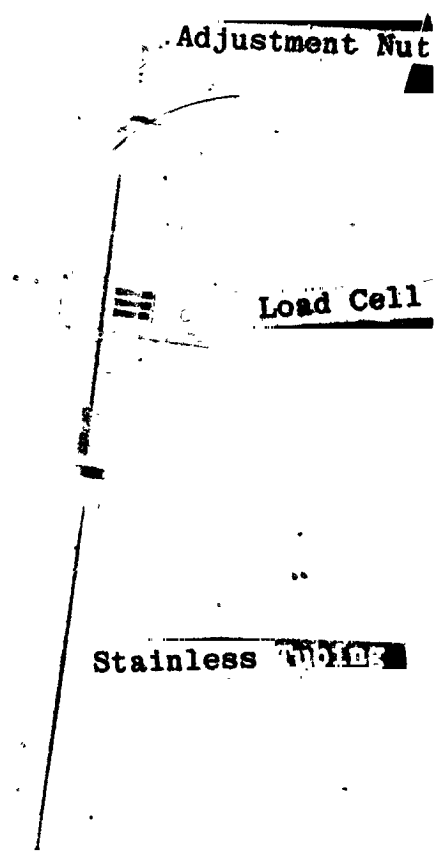
Figure 3:   Upper Pull Test Assembly

Load Cell
Toroid Corp.
Model 35-233 DJB
Range 500
Serial No. 43084

Strain Indicator
BLH Model 120-C
4-Arm Method
5-15-81
Gage Factor: 734

LOAD (lbs.)

Gauge Reading (+30 thousands)
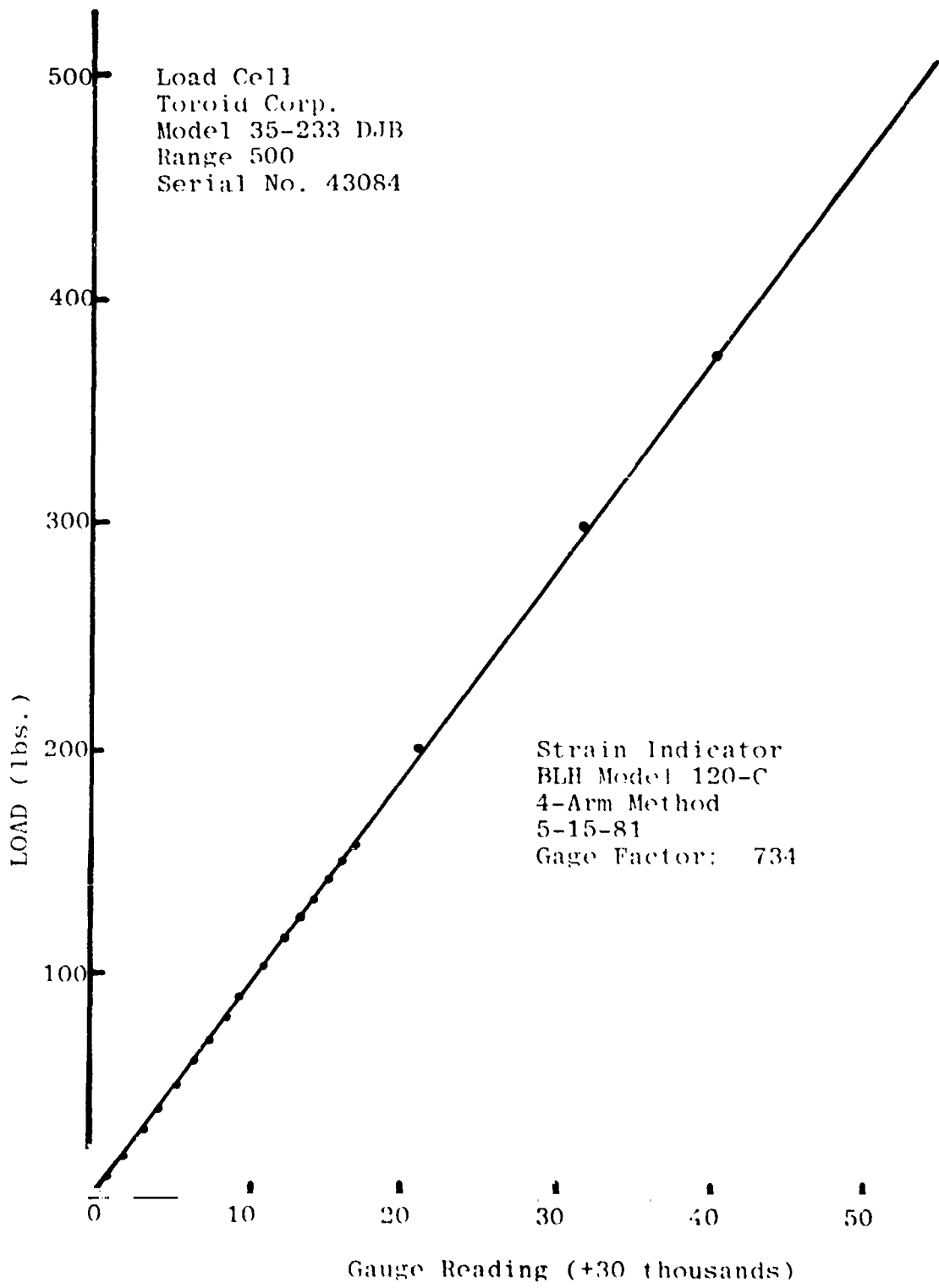
Figure 4: Load Cell Calibration Curve
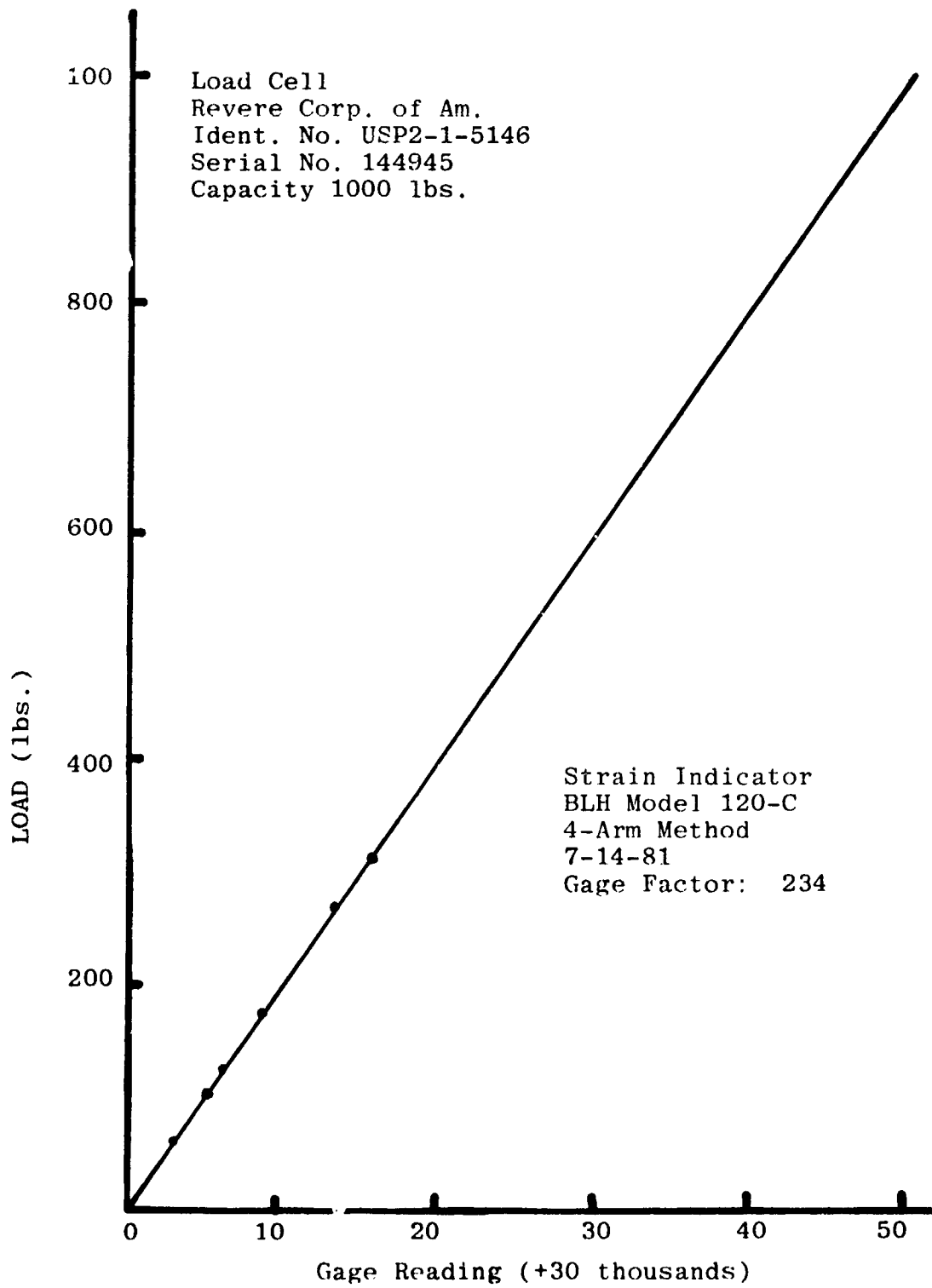
Figure 5:   Load Cell Calibration Curve

TABLE I

SUMMARY OF TENSILE PULL TEST RESULTS

Test Number	Load (lbs.)	Pressure (psi)	Failure Point
1	344	438	Braid Broke
2	287	366	Quartz Fracture
3	125	159	RTV Joint
4	110	140	RTV Joint
5	68	87	RTV Joint
6	102	130	Quartz Fracture
7	218	278	RTV Joint
8*	500	640	Load Cell Limit
9	520	662	Load Cell Limit
10	328	418	RTV Joint
11	96	122	Quartz Fracture

*Sample left under tension of 500 lbs. at 4.2° K for 4 hours.

REFERENCES

1. Fairbank, W.M., et al: Final Report.
    NASA Grant 05-020-019, W.W. Hansen Laboratories of Physics
    and the Department of Aeronautics and Astronautics,
    Stanford University, July 1977.

2. Fairbank, W.M. et al: Report on Progress.
    NASA Grant 05-020-019, USAF Contract F36615-67-C-1245,
    and NASA Contract NAS8-32355, W.W. Hansen Laboratories of
    Physics and the Department of Aeronautics and
    Astronautics, Stanford University, June 1980.

3. Lord Rayleigh, F.R.S., "Proc. Roy. Soc.", A156, (1936),
    pp. 326-349.

4. Peters, P.N.: Optical Contacting and Its Application to
    Gravity Probe B. NASA Report. unpublished. August 1981.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

MATTER IN GENERAL RELATIVITY

Prepared By: John R. Ray

Academic Rank: Professor

University and Department: Clemson University
                                        Department of Physics
                                          and Astronomy

NASA/MSFC: Space Sciences Laboratory
        Division: Space Physics Division
        Branch: Cryogenic Physics Branch

MSFC Counterpart: Peter Eby
                                        Larry Smalley

Date:

Contract No.:
                                        The University of Alabama in
                                          Huntsville

MATTER IN GENERAL RELATIVITY

by

John R. Ray

Professor of Physics and Astronomy
Clemson University
Clemson, South Carolina 29631

ABSTRACT

The Einstein field equations can be written in the schematic form

Geometry of Spacetime = Matter Content of Spacetime

The left-hand side of this equation corresponds to the geometrical description of gravitation introduced by Einstein. The right-hand side is a generalization of the mass density of Newtonian gravitational theory and is not so well defined as the geometrical term. Two important descriptions of matter that are used in general relativity are the fluid description and the kinetic theory description. Our work this summer has concentrated on studying these two theories of matter in general relativity. New results obtained include: a clarifying discussion of various methods of completing the fluid equations; a method of constructing charged general relativistic solutions in kinetic theory; and a proof and discussion of the incompatibility of perfect fluid solutions in anisotropic cosmologies. Intrepretations of NASA gravitational experiments using the above mentioned results have been started. Two papers have been prepared for publication based on this work and together make up the main part of this report.

INTRODUCTION

The Einstein field equations relate the geometry of space-time to the matter content of spacetime through the energy momentum tensor of matter. The description of matter in general relativity often uses phenomological forms for the energy momentum tensor, such as the fluid form. Even within the fluid description there are different approaches associated with multiple ways of completing the fluid equations in relativistic hydrodynamics. The first paper "Perfect Fluids in General Relativity" discusses and clarifies this subject.

Each of these papers gives a discussion of a description of matter in general relativity. Interpretations of gravitational experiments are influenced by the method chosen for describing the matter. As a result of this work Dr. A. J. Fennelly of MSFC and I are studying the possible correlation of results from experiments with the NASA ST and GP-B.

PERFECT FLUIDS IN GENERAL RELATIVITY

John R. Ray

Space Sciences Laboratory
NASA/Marshall Space Flight Center
Marshall Space Flight Center, Alabama 35812[*]

and

Department of Physics and Astronomy
Clemson University
Clemson, South Carolina 29631[**]

ABSTRACT

The equations of relativistic fluid dynamics can be completed in several different ways. We discuss and compare three distinct methods for completing these equations. We also discuss variational principles for the different fluid theories.

*NASA-ASEE Summer Faculty Fellow, MSFC, 1981
**Permanent address

I. INTRODUCTION

In general relativity one often employs a perfect fluid energy-momentum tensor to represent the source of the gravitational field. This fluid description is thought to be appropriate in cosmology where one assumes the large-scale properties of the Universe can be studied by assuming a perfect fluid description for the sources. A survey of the literature will, in fact, verify that almost all cosmological studies use the perfect fluid model. Exceptions are studies that make use of relativistic kinetic theory, which is complementary to the fluid description.[1] Also, in the astrophysics of stars, star clusters, and galaxies a fluid description is often employed.

The perfect fluid energy-momentum tensor T^{jk} can be written

$$T^{jk} = (W+P) \; U^j U^k + Pg^{jk} \quad , \tag{1.1}$$

where W is the energy density, P the pressure, U^j the normalized Eulerian four-velocity of the fluid and g^{jk} the metric tensor for space-time geometry. All thermodynamic variables W, P, ... are defined in the local rest frame of the fluid and are, therefore, scalar functions. A perfect fluid has, in general, two thermodynamic degrees of freedom, for example P and the rest mass density ρ. All other thermodynamic variables, e.g., W,...., can be written in terms of P and ρ

$$W = W(P,\rho) \quad . \tag{1.2}$$

Equation (1.2) is called the caloric equation of state and must be given in order to specify the nature of the fluid. As an example, "stiff matter" is described by the caloric equation of state W = P.

The gravitational field equations are

$$G^{jk} = 8\pi T^{jk} \quad , \tag{1.3}$$

where G^{jk} is the Einstein tensor. Equations of motion for the fluid follow from Bianchi identities

$$T^{jk}_{\;\;;k} = 0 \quad , \tag{1.4}$$

where the semicolon means covariant differentiation. The ten gravitational field equations, (1.3), together with the four fluid equations (1.4) form an incomplete system of equations for the five fluid variables P, ρ, U^i and the ten gravitational variables g^{jk}. There are several ways of completing these equations. Two often used methods are:[2]

a)   assume mass is conserved

$$(\rho U^j)_{;j} = 0 \quad , \tag{1.5}$$

b)   assume a relation between P and ρ

$$P = P(\rho) \quad . \tag{1.6}$$

Note that in case b) all the thermodynamic variables can be written in terms of one thermodynamic variable, say ρ. The first method of completing the equations is more familiar and corresponds to non-relativistic hydrodynamics where rest mass is always conserved. In order to avoid confusion we point out that it is possible to complete the equations so that both a) and b) hold. We shall discuss examples of this type of completion later in this section. Also, we point out that the results of this paper are also valid in special relativistic perfect fluid dynamics.

The combined first and second laws of thermodynamics imply

$$Tds = d(W/\rho) + Pd(1/\rho) \quad , \tag{1.7}$$

where T is the temperature and s the specific entropy. If we know the caloric equation of state, $W = W(P,\rho)$, then we may use (1.7) to solve for $s(P,\rho)$ and $T(P,\rho)$. For example, if we consider the "stiff matter" caloric equation of state $W = P$, we find

$$s = q(P/\rho^2) \quad , \tag{1.8}$$

$$T = \frac{\rho}{\dfrac{dg(P/\rho^2)}{d(P/\rho^2)}} \quad , \tag{1.9}$$

where g is an arbitrary function of P/ρ^2 such that T is well behaved and positive.

A special case of (1.7) is when $T = 0$ (cold matter) or $s = $ const (isentropic flow). For these cases

$$\frac{dW}{W+P} = \frac{d\rho}{\rho} = d\ell n\rho \quad , \tag{1.10}$$

and therefore $dW/(W+P)$ must be exact. This condition implies that we must have

$$\frac{\partial}{\partial P}\left(\frac{1}{W+P}\left(\frac{\partial W}{\partial \rho}\right)_P\right)_\rho = \frac{\partial}{\partial \rho}\left(\frac{1}{W+P} \; \partial P/\rho\right)_\rho \quad , \tag{1.11}$$

which yields

$$\left(\frac{\partial W}{\partial \rho}\right)_P = 0 \quad . \tag{1.12}$$

Hence, for $T = 0$ or $s = $ const. we find that W depends only upon P

$$W = W(P), \text{ if } T = 0 \text{ or } s = \text{const.} \tag{1.13}$$

With the results (1.13) we return to (1.10) and find the relation between P and ρ is

$$\rho = ce^{\int \frac{dW(P)}{W+P}} \quad , \text{ if } T = 0 \text{ or } s = \text{const.} \tag{1.14}$$

where c is a constant.

These results show that the requirements $T = 0$ or $s = \text{const.}$, which are often applied is cosmology and astrophysics, are an alternate way of completing the fluid equations which automatically yield the relation (1.14) between ρ and P. Thus, we have proven that the requirement $T = 0$ or $s = \text{const.}$ satisfy property b), (1.6). Next we show that $T = 0$ or $s = \text{const.}$ also imply a), (1.7). First we use T^{jk} to write

$$T^{jk}{}_{;k} = ([W(P,\rho) + P]U^j U^k + Pg^{jk})_{;k} = 0 \quad . \qquad (1.15)$$

Contracting (1.15) with U_j yields

$$(W+P)U^k{}_{;k} + W_{,k}U^k = 0 \qquad (1.16)$$

Combining (1.7) and (1.16) leads to

$$(W+P)(\rho U^j)_{;j} + \rho^2 T s_{,j} U^j = 0 \quad . \qquad (1.17)$$

From this latter important equation we see that the requirement $T = 0$ or $s = \text{const}$ implies $(\rho U^j)_{;j} = 0$ or property a), (1.5).

From this analysis we see that we should introduce a third way of completing the equations of fluid dynamics:

c) $T = 0$ or $s = \text{const.}$ (1.18)

We have proven that

c) $\rightarrow$ a) and b) . (1.19)

The main purpose of this paper is to discuss fluid motions in which method a) is used to complete the fluid equations but b) does not hold, and alternately, when b) is used to complete the fluid equations but a) does not hold. It is clear for our discussions so far that we cannot then have $T = 0$ or $s = \text{const.}$, since if either of these conditions obtain (1.19) implies that we have both a) and b).

II. COMPLETION USING a) $(\rho U^j)_{;j} = 0$

For method a) we solve the equations

$$([W(P,\rho) + P]U^j U^k + Pg^{jk})_{;k} = 0 \quad , \tag{2.1}$$

$$(\rho U^j)_{;j} = 0 \quad , \tag{2.2}$$

for the variables $(\rho(x), P(x), U^j(x))$. We, of course, must know the caloric equation of state $W = W(P,\rho)$. The temperature and entropy $T(P,\rho)$ and $s(P,\rho)$ can be found from (1.7) and then $T(x)$, $s(x)$ determined from the solutions to (2.1) and (2.2). Note that from (1.17) this solution for s automatically satisfies

$$s_{,i}U^i = 0 \quad , \tag{2.3}$$

or the specific entropy is constant along the fluid flow. Condition (2.3) does not, of course, mean that $s = $ const. In Newtonian fluid mechanics conditions (2.2) and (2.3), which are both valid, are independent equations which are always satisfied. In relativistic fluid mechanics these equations are related by (1.17) which arises from the relativistic Euler equations (2.1). It follows that the solutions which result when we complete the fluid equations using a) are homologous to the solutions in the Newtonian theory of perfect fluids. Also, we mention that the special case of isentropic flow $s = $ const., which we are not considering in this section, also obtains in the Newtonian case.

6

III. COMPLETION USING b) $P = P(\rho)$

In this case all thermodynamic variables can be written in terms of one variable, say P, $W = W(P)$. In this case our earlier proof that $\frac{dW}{W+P}$ is exact is still true since $\left(\frac{\partial W}{\partial \rho}\right)_P = 0$ is satisfied. One exceptional case is the equation of state $P = 0$, for which we must express $W = W(\rho)$. But even in this case $\frac{dW}{W+P} = \frac{dW}{W} = d \ln W$ is still exact. Since $\frac{dW}{W+P}$ is exact it is useful to introduce a new thermodynamic potential α defined by[2]

$$\frac{d\alpha}{\alpha} = \frac{dW}{W+P} \quad . \tag{3.1}$$

Using (3.1) in (1.7) we have

$$Tds = \left(\frac{W+P}{\rho}\right)d \ln(\alpha/\rho) \quad . \tag{3.2}$$

Our assumption that $T \neq 0$, $s \neq$ const implies that $\alpha \neq \rho$. Note also that $\alpha = \rho$ if and only if $T = 0$ or $s =$ const.[2] Combining (3.1) and (1.16) we obtain

$$(\alpha U^j)_{;j} = 0 \tag{3.3}$$

or the potential α is conserved. Note that since we are assuming $T \neq 0$ and $s \neq$ const. α is not equal to ρ and (3.3) does not correspond to conservation of mass. Also, since we are assuming case a) does not obtain, it follows from (1.17) that s is not constant along the flow lines, $s_{,j}U^j \neq 0$. Equation (1.17) shows that the nonconservation of entropy along the flow lines and the nonconservation of mass, balance one another and, of course, yield the conservation of α. This is a purely relativistic effect since, as mentioned previously, mass is always conserved in

Newtonian fluid theory. An interesting discussion of (1.17) in terms of particle creation has recently been presented which is valid for this case.[3]

In case b), as we have mentioned, all thermodynamic variables are expressible in terms of say P. From (3.1) we have for $\alpha(P)$

$$\alpha(P) = ce^{\int \frac{dW}{W+P}} \qquad (3.4)$$

where c is a constant. Note for $P = 0$ we obtain $\alpha = cW$. As a simple example of this case we consider the "stiff matter" equation of state

$$W = P \qquad (3.5)$$

which yields

$$s = g(P/\rho^2) \quad ,$$

$$s \neq constant \quad ,$$

$$T = \frac{\rho}{\frac{dg(\rho/P^2)}{d(P/\rho^2)}}$$

$$P = b\alpha^2 \quad , \quad b = const.$$

Here g is an arbitrary function such that T is well behaved and positive. We must next specify $P = P(\rho)$ which will allow us to determine all quantities as functions of P.

Finally, the Euler equations to be satisfied for case b) are

$$([W(P) + P]U^jU^k + Pg^{jk})_{;k} = 0 \quad , \qquad (3.6)$$

which are to be solved for $\{P(x), U^i(x)\}$. $\alpha(x)$ can then be calculated from (3.4) and automatically satisfies the conservation law (3.3).

IV. VARIATIONAL PRINCIPLES

A variational principle for a perfect fluid for which we complete the equations using conservation of mass has previously been presented.[4] The Lagrangian density is

$$L = L_1 + (-g)^{1/2}[\lambda_1(g_{jk}U^jU^k+1) + \lambda_2(\rho U^j)_{;j}$$
$$+ \lambda_3\chi_{,i}U^i + \lambda_4 s_{,i}U^i] \ , \tag{4.1}$$

where

$$L_1 = \frac{1}{2}(-g)^{1/2} R - (-g)^{1/2} W \tag{4.2}$$

is the Lagrangian for the free gravitational field R and the fluid W. The λ's in (4.1) are Lagrange multiples for the constraints that must be satisfied. Here χ is the particle identity field which must be introduced to avoid the Lin difficulty.[4] When we vary L with respect to the variables g_{jk}, ρ, s, U^j, χ, we obtain the Einstein field equations for a perfect fluid source plus the equations of motion of the fluid. This variational principle automatically gives the fluid equations in terms of the velocity potential representation[4,5]

$$U^i = \frac{\rho\lambda_{2,i}}{W+P} - \frac{\lambda_3\chi_{,i}}{W+P} - \frac{\lambda_4 s_{,i}}{W+P} \ . \tag{4.3}$$

For more details of this variational principle we refer the reader to Refs. 4,5 and references contained therein.

If we complete the fluid equations using a relation between P and ρ, we have to formulate a new variational principle. Reflection on the above variational principle leads us to construct the Lagrangian density

$$L' = L_1 + (-g)^{1/2}(\mu_1(g_{jk}U^jU^k+1) + \mu_2(\alpha U^j)_{;j}$$
$$+ \mu_3\chi_{,j}U^j) \ , \tag{4.4}$$

where the μ's are Lagrange multiples for the constraints which must be satisfied. If L' is varied with respect to g_{jk}, ρ, U^j, χ, then we obtain the Einstein equations for a perfect fluid source plus the equations of motion for the fluid. The fluid equations are again given in a velocity potential representation

$$U_j = \frac{\alpha \mu_2,_j}{W+P} - \frac{\mu_3 \chi,_j}{W+P} \quad . \tag{4.5}$$

Note that (4.5) is a representation of U_j in terms of four potentials instead of six potentials as in (4.3). Also, we point out that in varying (4.4) with respect to ρ, α is assumed to be expressed as a function of ρ, $\alpha(\rho)$ by using the assumed relation $P = P(\rho)$ and (3.4). We have, therefore, constructed a variational principle for fluid systems in the case when the fluid equations are completed by giving a relation between P and ρ.

V. CONCLUSIONS

The basic equations of relativistic fluid mechanics form an incomplete set of equations. We have discussed three distinct methods of completing the equations a), b) and c). Method a) corresponds more closely to Newtonian theory in that rest mass is assumed to be conserved. Method b) assumes the fluid is described in terms of only one thermodynamic degree of freedom by assigning a relation between, say, P and ρ. Method c) assumes cold matter or isentropic flow and implies both the properties that define methods a) and b) are satisfied. Method b) is characterized by the existence of a new thermodynamic potential

α that is conserved. Since Taub first recognized α, we call α the Taub potential. This potential is the density if and only if we assume c). The three methods of completing the fluid equations are physically distinct and correspond to different types of fluids and/or fluid flows. There is a mathematical isomorphism between b) and c) since $\alpha = \rho$ for c), but the flows corresponding to b) and c) are physically distinct since $T = 0$ or $s = $ const for c) but $T \neq 0$, $s \neq$ const for b). Finally, we have constructed a complete variational principle for b) by analogy with our earlier variational principle for a). One important aspect of this variational principle is the appearance of a four potential instead of a six potential representation of the fluid 4-velocity. This four-potential form of the fluid equations might be easier to use in numerical studies such as discussed in Reference 5.

REFERENCES

1. J. R. Ray and J. C. Zimmerman, Il Nuovo Cimento 42B 183 (1977), and references contained therein.

2. A. Taub, in Relativity and Astrophysics, ed. J. Ehlers, (Amer. Math. Soc., 1965).

3. M. A. Cirit, Lett. Nuovo Cimento, 30, 350 (1981).

4. J. R. Ray, J. Math. Phys. 13, 1451 (1972).

5. L. Smarr, C. Taubes, and J. R. Wilson, "General Relativistic Hydrodynamics: The Comoving, Eulerian, and Velocity Potential Formalisms, 1980 preprint.

KINETIC THEORY IN ASTROPHYSICS AND COSMOLOGY

John R. Ray*

Space Sciences Laboratory
Marshall Space Flight Center, ALabama 35812

and

Department of Physics and Astronomy
Clemson University
Clemson, South Carolina 29631

KINETIC THEORY IN ASTROPHYSICS AND COSMOLOGY

by

John R. Ray

ABSTRACT

We present results associated with exact solution of the Einstein-Boltzmann and Einstein-Maxwell-Boltzmann equations. An important aspect of the paper is the generalization of Ehler's killing vector approach for the distribution function to charged particles.

I. INTRODUCTION

General relativistic kinetic theory offers an alternative description of matter which in some aspects is more fundamental than the usual fluid description. Kinetic theory is appropriate for describing "dilute gases" in which correlations are not important, in other words, systems whose particles have long mean free paths and in which collisions are rare. Since large systems satisfying these conditions exist, for example, stellar clusters and the gas of galaxies now, we should employ kinetic theory for a study of such systems. Also the inclusion of gases of photons and neutrinos (zero mass fields) is straightforward using relativistic kinetic theory. The lack of a relativistic statistical mechanics for dense systems, say, along the lines of the BBYKG hierarchy is still an outstanding problem. Of course, the lack of a relativistic statistical mechanics is a reflection of the lack of a satisfactory relativistic theory of interacting particles. Relativistic kinetic theory is derived in a heuristic manner following Boltzmann's derivation of nonrelativistic kinetic theory (Hakim, 1967; Stewart, 1971; Ehlers, 1971; Israel and Stewart, 1970). The kinetic theory and fluid descriptions are complementary ways of treating matter in general relativity.

In this paper we discuss collisionless distribution functions in both the astrophysical - stellar clusters - and cosmological - gas of galaxies - context. We concentrate our discussion on two particular metrics. For the astrophysical problem we use the static spherically symmetric metric. The kinetic theory for this

metric was first presented by Fackerell (1968). For the
cosmological problem we study the Petrov G_4 VIII metric which we
have previously discussed (Ray and Zimmerman , 1977). This metric
is spatially homogeneous but anisotropic. It has been studied
by Kompaneets and Chernov (1965), Kantowski and Sachs (1966) and
Collins (1977) using a fluid description for the matter.

In order to derive the distribution functions we use a
killing vector approach suggested by Ehlers (1971), and first
applied in detail by Ray amd Zimmerman (1977). An important part
of this paper is the extension of Ehler's killing vector approach
to include charged particles and electromagnetic fields. Kinetic
theory is an important theory for studying nonrelativistic plasmas,
and it should also be important in relativistic plasmas.

In Section II of this paper we present the basic kinetic
theory that we use, in Section III we apply this theory to the
static spherically symmetric metric, and in Section IV we further
study the kinetic theory of the G_4 VIII metric. In these
examples we study both the charged and uncharged systems. Important
results we obtain are that the G_4 VIII spacetime cannot be
generated by a perfect fluid energy-momentum tensor in kinetic
theory. We must reconcile this result with the known perfect
fluid solutions for this metric. Also, we find that the G_4 VIII
model cannot be generated by charged particles. We discuss
the implications of this result for other Bianchi-type cosmological
models. All of these results are based on the distribution function

derived by using the killing vector approach outlined in Section II. Finally in the conclusion we discuss this work and present suggestions for future efforts.

II. KINETIC THEORY

We are interested in solutions to the Einstein-Boltzmann equations. The Einstein equations are

$$G_{ij} = T_{ij} \quad , \tag{2.1}$$

where G_{ij} is the Einstein tensor and T_{ij} the energy momentum tensor for the matter generating the gravitational field. In the case of neutral particles, T_{ij} is given by the kinetic theory energy-momentum tensor K^{ij}

$$T^{ij}_{(x)} = K^{ij}_{(x)} = \int P^i P^j f(x,p) \sqrt{-g} \, d^4P \quad , \tag{2.2}$$

where $f(x,p)$ is the single particle distribution function of kinetic theory, P^i is the four momentum, g is the determinant of the metric tensor, and $d^4P = dP^1 dP^2 dP^3 dP^4$. For simplicity we restrict ourselves to a one-component gas of particles having rest mass m. The integration in (2.2) is then over the mass shell $g_{ij} P^i P^j = -m^2$. This can be taken into account by introducing the delta function $\delta(P_i P^i + m^2)$ into the integrand in (2.2). The distribution function f satisfies the collisionless Boltzmann equation

$$\frac{\partial f}{\partial x^i} \frac{dx^i}{dv} + \frac{\partial f}{\partial P^i} \frac{dP^i}{dv} = 0 \quad , \tag{2.3}$$

where v is an affine parameter characterizing the particle paths which are geodesics for neutral particles

$$\frac{DP^i}{dv} = \frac{dP^i}{dv} + \Gamma^i_{jk} P^j P^k = 0 \quad , \quad P^i = \frac{dx^i}{dv} \quad . \tag{2.4}$$

Combining (2.4) and (2.3) we have the Boltzmann equation

$$\frac{\partial f}{\partial x^i}P^i - P^j P^k \Gamma^i_{jk} \frac{\partial f}{\partial P^i} = 0 \quad . \qquad (2.5)$$

The distribution function f is a function of constants of the motion. In the killing vector approach we use the fact that for a killing fector $\xi^i_{(a)}$ we have the constant of the motion $Y_a = \xi^i_{(a)} P_i$. Hence a solution to (2.5) is $f(Y_1, \text{---} Y_r)$. The index a labels the different killing vectors a = 1,2, ---, r. In the usual approach these constants are the energy, angular momentum, etc. Next we impose the killing vector symmetry on f. This yields

$$\mathcal{L}_{\xi_{(a)}} f = \frac{\partial f}{\partial x^i} \xi^i_{(a)} + \frac{\partial f}{\partial P^i} \xi^i_{(a),\ell} P^\ell = 0 \quad , \qquad (2.6)$$

where $\mathcal{L}_{\xi_{(a)}}$ stands for the Lie derivative. As Ehlers (1971) shows, (2.6) yields the conditions

$$\frac{\partial f}{\partial Y_a} C^d_{ab} Y_d = 0 \quad , \qquad (2.7)$$

on the distribution function f(Y). The equations (2.7) are r linear, coupled, partial differential equations for the distribution function f(Y). The C^d_{ab} are the structure constants for the Lie algebra of the killing vectors

$$[\xi_{(a)}, \xi_{(b)}] = C^d_{ab} \xi_{(d)} \quad . \qquad (2.8)$$

In summary, in order to determine a solution to the Einstein-Boltzmann equation we pick a spacetime of interest. From the killing vectors we calculate the Y_a's and solve (2.7) for f. Then we calculate K^{ij} from (2.2) and then solve the Einstein equations (2.1). This procedure is explained in detail in Ray and Zimmerman (1977).

We now wish to modify the above procedure for cases when the particles are charged. In this case we have the combined Einstein-Maxwell-Boltzmann equations for a charged self-gravitating gas. Physically a more interesting case would be an electrically neutral plasma. In order to treat this case we could introduce distribution functions for the electrons and ions. While this would be possible, we shall stay with a single distribution function and a charged gas for simplicity. The Einstein equations are still (2.1); however, T^{ij} now contains two terms

$$T^{ij} = K^{ij} + E^{ij} \, , \tag{2.8}$$

where E^{ij} is the Maxwell energy-momentum tensor

$$E^{ij} = F^{i\ell}F^j_{\ \ell} - \frac{1}{4} g^{ij}F_{\ell k}F^{\ell k} \, , \tag{2.9}$$

and F^{ij} is the field tensor for electromagnetic field. The particles of charge q move according to the Lorentz equation

$$\frac{DP^i}{dv} = qP_\ell F^{i\ell} \, . \tag{2.10}$$

Maxwell's equations have the form

$$F^{ij}_{\ \ ;j} = J^i \tag{2.11}$$

$$F_{[ij;k]} = 0 \, . \tag{2.12}$$

If we introduce the four-potential A_i

$$F_{ij} = 2A_{[j,i]} \, , \tag{2.13}$$

then (2.12) are satisfied identically. The current J^i in (2.11) is determined in terms of the distribution function via

$$J^i = q N^i = q \int P^i f \sqrt{-g} \, d^4P \, , \tag{2.14}$$

where N^i is the particle current. Finally, the Boltzmann equation has the form

$$\frac{\partial f}{\partial x^i} P^i + (qP_\ell F^{i\ell} - P^j P^k \Gamma^i_{jk}) \frac{\partial f}{\partial P^i} = 0 \quad . \tag{2.15}$$

Now f is a function of first integrals of the Lorentz equation (2.10). In analogy with the killing vector approach we employed for neutral particles where $\xi^i_{(a)} P_i$ were constants of the motion we try for the charged case

$$Y_a = \xi^i_{(a)} (P_i + qA_i) \quad , \tag{2.16}$$

where $\xi^i_{(a)}$ is a killing vector. It is easy to show that Y'_a is a constant of the motion if the Lie derivative of the vector potential is zero

$$Y_a = \xi^i_{(a)} (P_i + qA_i) = \text{const} \leftrightarrow \mathcal{L}_{\xi_{(a)}} A^i = 0 \quad . \tag{2.17}$$

Thus the distribution function still depends on the Y'_as but now the Y'_as are given by (2.16). We note that the Y_a are constants even if we have a multicomponent gas, and in particular in the ion-electron plasma problem mentioned earlier. Note also that the result (2.17) is not surprising since $P_i + qA_i$ is just the canonical momentum for a particle in an electromagnetic field.

Next, following the neutral particle example, we impose the symmetry on the distribution function via (2.6), with the Y'_as given by (2.16). Carrying out this calculation using (2.6) we find exactly the same result (2.7)

$$\frac{\partial f}{\partial Y_a} C^d_{ab} Y_d = 0 \quad , \tag{2.18}$$

except now the Y_a's are given by (2.16). Since the equation (2.18) is the same as (2.7) this means that the solution for f in terms of the Y'_as will have the same form. Thus, in order to find the distribution function for charged matter we only

need make the replacement $P_i \rightarrow P_i + qA_i$; that is, we replace the momentum P_i by the canonical momentum $P_i + qA_i$ in the distribution function f. This is an important result because it reduces the problem of solving for the distribution function in the charged particle case to that of the neutral particle case.

In summary, our procedure in the charged particle case is: start with a spacetime with given killing vectors $\zeta^i_{(a)}$; determine the form of the vector potential A^i such that $\mathcal{L}_{\zeta_{(a)}} A^i = 0$; solve (2.18) to find the distribution function; use the Y_a to find $f(x,P) = f(Y)$; calculate K^{ij}, E^{ij}, N^i and then solve the field equations

$$G^{ij} = K^{ij} + E^{ij} \quad , \tag{2.19}$$

$$F^{ij}{}_{;j} = qN^i = J^i \quad . \tag{2.20}$$

Note that by construction K^{ij} satisfies

$$K^{ij}{}_{;j} = J_\ell F^{r\ell} \tag{2.21}$$

and E^{ij} satisfies

$$E^{ij}{}_{;j} = -J_\ell F^{r\ell} \quad , \tag{2.22}$$

hence the Bianchi identities $G^{ij}{}_{;j} = 0$ are satisfied, which is of course necessary for self-consistency. Note also that from the definition of N^i and K^{ij} we can prove

$$\mathcal{L}_{\zeta_{(a)}} N^i = \int P^i \mathcal{L}_{\zeta_{(a)}} f \sqrt{-g} \, d^4P = 0 \tag{2.23}$$

$$\mathcal{L}_{\zeta_{(a)}} K^{ij} = \int P^i P^j \mathcal{L}_{\zeta_{(a)}} f \sqrt{-g} \, d^4P = 0 \quad . \tag{2.24}$$

Also from $\mathcal{L}_{\zeta_{(a)}} A^i = 0$ we can prove $\mathcal{L}_{\zeta_{(a)}} F^{ij} = 0$ and $\mathcal{L}_{\zeta_{(a)}} E^{ij} = 0$.

All of these Lie derivative symmetry relations are also required for self-consistency and are satisfied automatically in the killing vector approach.

III. STATIC SPHERICAL SYMMETRY

Here we use the metric in the form

$$ds^2 = e^{\lambda(r)}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) - e^{\nu(r)}dt^2 , \qquad (3.1)$$

with $(x^1, x^2, x^3, x^4) = (r, \theta, \phi, t)$. First we shall discuss the uncharged case (Fackerell, 1968), and then charge the particles. The killing vectors are

$$\xi^i_{(1)} = \cos\phi\delta^i_2 - \sin\phi\cot\theta\delta^i_3 \qquad (3.2)$$

$$\xi^i_{(2)} = \sin\phi\delta^i_2 + \cos\phi\cot\theta \, \delta^i_3 \qquad (3.3)$$

$$\xi^i_{(3)} = \delta^i_3 \qquad (3.4)$$

$$\xi^i_{(4)} = \delta^i_4 . \qquad (3.5)$$

The structure constants of the Lie algebra are

$$C^a_{bc} = \epsilon_{abc} , \quad a,b,c,.. = 1,2,3$$

$$C^a_{b4} = 0 ,$$

where ϵ_{abc} is the antisymmetric symbol. Consideration of equation (2.6) yields

$$f = f(-Y_4, (Y^2_1 + Y^2_2 + Y^3_3)^{1/2}) ; \qquad (3.6)$$

calculation of the $Y_a = \xi^i_{(a)} P^i$ leads to

$$Y^2_1 + Y^2_2 + Y^2_3 = P^2_2 + P^2_3/\sin^2\theta = J^2 , \qquad (3.7)$$

$$Y_4 = + P_4 = -E , \qquad (3.8)$$

where J is the total angular momentum and E the energy of a particle in geodesic motion. Thus, the distribution function f depends only on E and J

$$f = f(E,J) . \qquad (3.9)$$

We now need to calculate N^i and T^{ij}

$$N^i = \int P^i f \sqrt{-g}\, d^4P \quad, \tag{3.10}$$

$$T^{ij} = \int P^i P^j f \sqrt{-g}\, d^4P \quad. \tag{3.11}$$

Since we restrict ourselves to particles of a single mass $P_i P^i = -m^2$, we must introduce the delta function $\delta(P_i P^i + m^2)$ into the volume element. Picking P^2, P^3, P^4 as integration variables, the volume element becomes

$$\sqrt{-g}\, d^4P \;\to\; \sqrt{-g}\,\frac{dP^2 dP^3 dP^4}{|P_1|} \quad, \tag{3.12}$$

where

$$|P_1| = e^{\lambda/2}[e^{-\nu}E^2 - J^2/r^2 - m^2]^{1/2} \quad. \tag{3.13}$$

Changing from $dP^2 dP^3 dP^4$ to $dJ dP_3 dE$ puts the volume element in the form

$$\sqrt{-g}\, d^4P \to \frac{e^{-\nu/2}}{r^2}\; \frac{J dJ dE dP_3}{[e^{-\nu}E^2 - J^2/r^2 - m^2]^{1/2}[J^2\sin^2\theta - P_3^2]^{1/2}} \quad. \tag{3.14}$$

As an example we give the results for N_i:

$$N_\alpha = 0 \qquad \alpha = 1,2,3 \quad, \tag{3.15}$$

$$N_4 = \frac{e^{-\nu/2}}{r^2}\,\pi \int \frac{E f(E,J) J dJ dE}{[e^{-\nu}E^2 - J^2/r^2 - m^2]^{1/2}} \quad, \tag{3.16}$$

where the P_3 integration has been carried out explicitly. The components for T^i_j can also be given in a form such as (3.16). Here we only note that the energy momentum tensor corresponds to that for a <u>perfect fluid</u>

$$K^i_j = (\rho+P)U^i U_j + P\delta^i_j \tag{3.17}$$

$$P = K^1_1 = K^2_2 = K^3_3$$

$$\rho = -K^4_4$$

$$U^i = \delta^i_4$$

if the distribution function is independent of the angular momentum J, that is, $f(E,J) = f(E)$. This was first proven by Fackerell (1968) and is also true in Newtonian stellar dynamics. This is an example from the wider class of locally dynamically symmetric distribution functions of the form

$$f(x,P) - f(x,-U_i P^i) \tag{3.18}$$

which give rise to perfect fluid energy-momentum tensors, Stewart (1971).

Next we discuss the case of static, spherical symmetry when the particles are charged. The vector potential is forced by the killing vector symmetry to have the form

$$A_i = -\delta_i^4 \phi(r) \tag{3.19}$$

which gives an electrical field in the radial direction

$$F_{14} = -\frac{d\phi}{dr} = E_r \quad . \tag{3.20}$$

The only change in the distribution functions is that $Y_4 = (P_4 - q\phi)$. Thus

$$f(-Y_4, (Y_1^2 + Y_2^2 + Y_3^2)^{1/2}) = f(-P_4 + q\phi, J) \quad . \tag{3.21}$$

The integration factors are the same as in the uncharged case. As an example we give the result for N_i, which is

$$N_\alpha = 0$$

$$N_4 = \frac{-\pi e^{-\nu/2}}{r^2} \int \frac{P_4 f(-P_4 + q\phi, J) \, J dJ dP_4}{[e^{-\nu}P_4^2 - J^2/r^2 - m^2]^{1/2}} \quad . \tag{3.22}$$

The components of K_j^i can also be written down by making the replacement $E \to (-P_4 + q\phi)$ in the argument of $f(E,J)$. It is also true for the charged case that one has a perfect fluid if f depends only on $E = -U_i(P^i + qA^i)$. In this case the distribution function is still locally dynamically symmetric

$$f(x,P) = f(x,-U^i(P_i+qA_i)) = h(x,-U^iP_i) \quad . \tag{3.23}$$

The charged particle solutions just obtained represent exact solutions to the Einstein-Maxwell-Boltzmann equations for particles whose electrical repulsion is just balanced by their gravitational attraction so that the distribution generates a static field. Exact solutions of this general form are known in general relativity; however, as far as we are aware, they have not previously been discussed from a kinetic theory point of view. As mentioned previously, the more interesting case would be that for a neutral plasma. The basic theory would be the same as presented above.

One interesting result of the paper by Fackerell (1968) was his proof of the nonexistence of perfect fluid models with constant energy density $\rho = -K_4^4$. One could possibly prove this same theorem for charged matter; however, we have not yet constructed such a proof.

IV.   G. VIII COSMOLOGY

The metric in this case has the form

$$ds^2 = X^2(t)dr^2 + Y^2(t)[d\theta^2 + \sin^2\theta d\phi^2] - dt^2 \quad , \tag{4.1}$$

$$(x^1,x^2,x^3,x^4) = (r,\theta,\phi,t) \quad .$$

Note that this metric is also spherically symmetric and so has the same spherically symmetric killing vectors $\xi^i_{(1)}$, $\xi^i_{(2)}$, $\xi^i_{(3)}$. The fourth killing vector for (4.1) is $\xi^i_{(4)} = \delta^i_1$. The spacetime (4.1) is homogeneous, but anisotropic, the r direction being a preferred direction. First we shall consider the neutral particle problem. Using the same procedure as in

the previous example, we find the distribution function has the form

$$f = f(P_1, J) \quad . \tag{4.2}$$

It is convenient in this case to use the volume element in the form

$$\sqrt{-g} \, d^4P \rightarrow \sqrt{-g} \, \frac{dP^1 dP^2 dP^3}{|P_4|} \tag{4.3}$$

$$P_4 = -E = -(P_1{}^2/X^2 + J^2/Y^2 + m^2)^{1/2} \quad , \tag{4.4}$$

and change to $dP_1 dJ dP_3$ which gives

$$\sqrt{-g} \, d^4P \rightarrow \frac{1}{XY^2} \frac{J dJ dP_1 dP_3}{[P_1{}^2/X^2 + J^2/Y^2 + m^2]^{1/2} [J^2 \sin^2\theta - P_3{}^2]^{1/2}} \tag{4.5}$$

Again we can calculate N_i, K_j^i from their definitions. This yields for N_4

$$N_4 = -\frac{\pi}{XY^2} \int f(P_1, J) \, J dJ dP_1 \tag{4.6}$$

From the Einstein equations we find $G_4^1 = 0$. If we calculate K_4^1, we find that $K_4^1 = 0$ implies f must be an even function of P_1. Using this fact in N_α we find that $N_\alpha = 0$ and the particles are comoving in the coordinate system of (4.1). The components of K_j^i which are not zero have the following form

$$K_1^1 = \frac{\pi}{X^3 Y^2} \int \frac{P_1^2 f(P_1, J) J dJ dP_1}{[P_1^2/X^2 + J^2/Y^2 + m^2]^{1/2}} \quad , \tag{4.7}$$

$$K_2^2 = K_3^3 = \frac{\pi}{2XY^4} \int \frac{f J^3 dJ dP_1}{[P_1^2/X^2 + J^2/Y^2 + m^2]^{1/2}} \quad , \tag{4.8}$$

and

$$K_4^4 = - \frac{\pi}{XY^2} \int [P_1^2/X^2 + J^2/Y^2 + m^2] f \, JdJdP_1 \quad . \qquad (4.9)$$

This energy momentum tensor does not correspond to a perfect fluid $K_1^1 \neq K_2^2 = K_3^3$. Following our previous example of static spherical symmetry we would assume f is independent of J in order to obtain the perfect fluid form for K_j^i. However, from (4.9) we see that this is not possible since K_4^4 then diverges. (The limits on J are $0 \rightarrow \infty$.) We conclude that this spacetime is inconsistent with a perfect fluid energy-momentum tensor in kinetic theory. This is not surprising considering that this model has a nonzero shear

$$\sigma = \frac{1}{\sqrt{3}} \left| \frac{\dot{X}}{X} - \frac{\dot{Y}}{Y} \right| \quad . \qquad (4.10)$$

This fact was also pointed out by Ray and Zimmerman (1977) and by Kantowski (1966). As was mentioned previously, exact perfect fluid models are known for this metric and are interpreted as reasonable cosmological models. Our results using kinetic theory lead one to question the physical relevancy of the perfect fluid solutions in this spacetime. This is analogous to Fackerell's result that the constant density Schwarzschild solution is inconsistent with kinetic theory. The interest in this result for the G_4 VIII spacetime is that the same result would probably occur in other homogeneous, anisotropic Bianchi spacetimes. This is an important point since the isotropy of the microwave background radiation is often (but not always) discussed using perfect fluid solutions in anisotropic models. A possible explanation of the meaning of perfect fluid solutions in anisotropic spacetime is

that they are valid at early times when the matter is strongly interacting and kinetic theory is not valid. Of course, kinetic theory is an accurate description of the gas of galaxies at the present epoch. Treciokas and Ellis (1971) have discussed the conditions under which a distribution function yields a perfect fluid energy momentum tensor.

Next we shall consider the charged problem from the G_4 VIII spacetime. Using the killing vectors we find the vector potential is restricted by symmetry to the form

$$A_i = A(t) \, \delta^1_i \quad , \tag{4.11}$$

which again gives the radial electric field

$$F_{14} = -\dot{A} = E_1 \quad . \tag{4.12}$$

The distribution function takes the form

$$f = f(P_1 + qA(t), J) \quad . \tag{4.13}$$

Calculation of N_4 leads to the result

$$N_4 = -\frac{\pi}{XY^2} \int f(P_1, J) J dJ dP_1 \quad , \tag{4.14}$$

the same as in the uncharged case (4.6). Next we investigate Maxwell's equations

$$\frac{1}{\sqrt{-g}} \, (\sqrt{-g} \, F^{ij})_{,j} = J^i \quad . \tag{4.15}$$

For $i = 4$, the left hand side of (4.15) vanishes, which implies $J^4 = 0$. But from (4.14) and remembering that $f \geq 0$, and $J_4 = qN_4$, we find $q = 0$; that is, this spacetime cannot be generated by charged particles in the kinetic theory approach. The physical reason for this result

is clear. Since the spacetime (4.1) is homogeneous, every event has an x^1 axis of symmetry passing through it. By symmetry the current must be along the x^1 axis, J^1. This current would produce a field transverse to the x^1 axis, but by the homogeneity of the spacetime this field must vanish. One can have source-free magnetic fields in this model (Collins, (1977)); however, source-free fields are not possible in kinetic theory since if $N_i = 0$ this implies $f = 0$, which corresponds to no particles at all.

The results of the last paragraph lead to the question of these same results occurring in other spatially homogeneous spacetimes. If these spacetimes cannot allow sources that generate the magnetic fields they contain, then these magnetic fields become part of the initial conditions of the universe, or are produced in earlier epochs that are inhomogeneous. This is an important point considering, for example, the recent work of Fennelly (1980) on the condensation of protogalaxies using plasma instabilities in the early universe.

V.   CONCLUSIONS

We have extended the killing vector approach to kinetic theory to include charged particles. The rule for forming the distribution function in the charged case is to replace the momentum by the canonical momentum which has the stipulation that the Lie derivative of the vector potential be zero. This supports the claim by Israel and Stewart (1980) that the canonical momentum has a fundamental significance in kinetic theory. In charging Fackerell's static spherically symmetric models we obtain a static distribution of charged particles

where the electrical repulsion is balanced by the gravitational attraction.

In applying kinetic theory to the G_4 VIII spacetime we arrive at the conclusion that this spacetime cannot be generated by a perfect fluid in kinetic theory. This agrees with the conclusions reached by Kantowski (1966) earlier. These results call into question the physical meaning of perfect fluid solutions in such spacetimes. Presumably many of the Bianchi models would share this same difficulty. Since fluid models are often employed in these spacetimes to interpret cosmological data it is important to understand fully the implications of this apparent paradox. It is not possible for the G_4 VIII spacetime to be generated by charged matter. This implies that the source-free magnetic fields present in exact solutions for this model must arise from an earlier, more chaotic state of the universe or be given as initial conditions.

In future work one needs to return and treat the neutral plasma problem in detail using the methods outlined above. Also, it would be important to construct models which could evolve to homogeneous cosmologies and contain sources for the fields. This would presumably take one to models that are valid at earlier epochs. Lastly, one should study the kinetic theory of the other Biachi modes. Recent work by Treciokos and Ellis (1980) is in this direction. Earlier work by Matzner (1971) is also an important contribution. It is clear that more work needs to be done in exploring exact solutions in the Einstein-Boltzmann and Einstein-Maxwell-Boltzmann theories.

ACKNOWLEDGMENT

The author thanks Al Fennelly for interesting and helpful discussions concerning this work.

REFERENCES

Collins, C. B.: 1977, J. Math. Phys. 18, 104, and references therein.

Ehlers, J.: 1971, General Relativity and Cosmology, N.Y., Academic Press.

Fackerell, E.D.: 1968, Ap. J. 153, 643.

Fennelly, A.J.: 1980, Phys. Rev. Lett. 44, 955.

Hakim, R.: 1967, J. Math. Phys. 8, 1315.

Israel, W. and Stewart, J.M.: 1980, General Relativity and Gravitation Vol. 2, N.Y. Plenum Press.

Kantowski, R.: 1966, Ph.D. thesis, Univ. of Texas.

Kantowski, R. and Sachs, R.K.: 1966, J. Math. Phys. 7, 443.

Kompaneets, A.S. and Chernov, A.S.: 1965, Sov. Phys. JETP 20, 1303.

Matzner, R.A.: 1971, Ann. Phys. 65, 438.

Ray, J.R. and Zimmerman, J.C.: 1977, Il Nuovo Cimento, 42B, 183.

Stewart, J.M.: 1971, Nonequilibrium Relativistic Kinetic Theory, Berlin Springer-Verlag.

Treciokas, R. and Ellis, G.F.R.: 1971, Commun. Math. Phys., 23, 1.

Treciokas, R. and Ellis, G.F.R.: 1980, "Anisotropic Solutions of the Einstein Boltzmann Equations", preprints.

ACKNOWLEDGEMENT

1981

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

SURFACE CONTOUR OF A BIAXIALLY STRAINED,
DOUBLY CURVED ANTENNA REFLECTIVE MESH

Prepared By: Annie V. Saylor, Ph.D.

Academic Rank: Assistant Professor

University and Department: The University of Alabama in Huntsville
                                  Department of Mathematics

NASA/MSFC:
    (Laboratory)   Preliminary Design Office
    (Division)   Systems Engineering
    (Branch)   Systems Laboratory

MSFC Counterpart: Carl E. Colley

Date: August 7, 1981

Contract No.:   NGT 01-008-021
                                  (The University of Alabama in Huntsville)

XX

SURFACE CONTOUR OF A BIAXIALLY STRAINED, DOUBLY CURVED ANTENNA REFLECTIVE MESH

BY

Annie V. Saylor
Assistant Professor of Mathematics
The University of Alabama in Huntsville
Huntsville, Alabama

ABSTRACT

The purpose of this study is to develop and solve the elastic membrane equations which describe the deflection of a biaxially tensioned reflective mesh from an idealized surface. The conditions of equilibrium of the forces acting on a membrane element furnish equations which may be used to solve a variety of problems. Configurations considered include continuous, as well as incremental, mesh attachment to structural members to form shapes which are approximately parabolic or spherical.

INTRODUCTION

Large space antennas are desirable for use in communications, radio astronomy, radar surveillance and submillimeter astronomy. These antennas will have diameters of 50 meters or more and will require deployment or construction in space due to their size. Currently, emphasis is being placed upon deployable structures which will support both a reflector and a phased array lens. Several designs for such structures have been proposed [1]; a common feature is the formation of the surface of the antenna via membrane(s) attached to a framework.

Among the many challenging problems faced in the design of these antennas is that of controlling surface accuracy. The deployed antenna's shape will be an approximation to a desired form, and the resulting error must be within tolerances established by the particular use intended.

OBJECTIVES

Surface accuracy is affected by several sources - approximation errors, environmental effects, aging effects and other uncertainties. This paper concerns itself only with the error due to pillowing which occurs in the membrane as it is used to approximate the desired surface. Our purpose is to derive equations describing this error and to obtain either exact or approximate solutions to these equations for various antenna designs.

SURFACE ACCURACY REQUIREMENTS

Suppose the desired surface is given by $u = f(x,y)$ where $(x,y) \in \mathcal{A}$, some planar region. Further assume that the shape of the deployed antenna can be described by $z = g(x,y)$ for $(x,y) \in \mathcal{A}$, and let $w = u - z$. Then the root-mean-square error is given by

(1)
$$\mathcal{E} = \sqrt{\frac{\iint_{\mathcal{A}} w^2 dA}{\text{area of } \mathcal{A}}}.$$

It has been established [1] that the restrictions on $\mathcal{E}$ set by most of the intended uses range from $10^{-5}d$ to $10^{-3}d$, where d is the diameter of the antenna.

For all designs under consideration, $\mathcal{A}$ is divided into subregions, usually triangles. We note that if

$$\mathcal{A} = \bigcup_i \mathcal{A}_i$$

with

$$\iint_{\mathcal{A}_i} w^2 dA = \iint_{\mathcal{A}_j} w^2 dA$$

and

area of $\mathcal{A}_i$ = area of $\mathcal{A}_j$ for all i,j, then

$$\mathcal{E} = \sqrt{\frac{\iint_{\mathcal{A}_i} w^2 dA}{\text{area of } \mathcal{A}_i}}$$

for any i.

GENERAL SURFACE EQUATION

For the surface in question, we assume that the membrane theory for shallow shells of arbitrary double curvature applies. (See [2] for a complete discussion of the implications of these assumptions.) Consider an element of the shell, with the x and y axes placed in the local principal curvature directions (see Figure 1 below). Let R_1 and R_2 be the local principal radii and suppose that the tensions N_x and N_y are known constants.

The shell equilibrum equation [2] gives

$$(2) \qquad \frac{N_x}{R_1} + \frac{N_y}{R_2} = 0,$$

and by the shallow-shell assumption,

$$(3a) \qquad z_{xx} = \frac{1}{R_1},$$

$$(3b) \qquad z_{yy} = \frac{1}{R_2};$$

hence

$$(4) \qquad N_x z_{xx} + N_y z_{yy} = 0.$$

If we consider this equation on a subregion $\mathcal{M}_i$ of $\mathcal{M}$ and require that appropriate additional conditions on z be met, we arrive at a unique choice for z on $\mathcal{M}_i$.
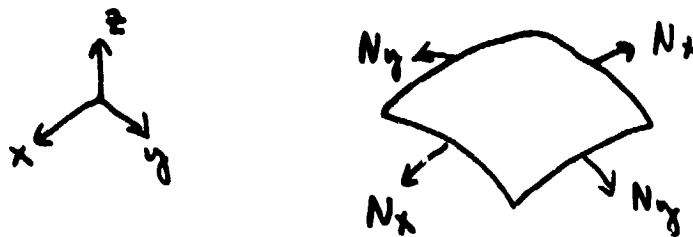
Figure 1

SURFACE BY CURVATURE ASSUMPTIONS

Suppose that the desired surface $u = f(x,y)$ has a maximum radius of curvature R which is independent of x and y. Further suppose that we require $R_1 = R$; then by (3), we have

(5) $\qquad z_{xx} = \dfrac{1}{R}$,

and by (4),

(6) $\qquad z_{yy} = \left(- \dfrac{N_x}{N_y} \right) \left(\dfrac{1}{R} \right)$.

This gives

(7) $\qquad z(x,y) = \dfrac{1}{2R} \left(x^2 - \dfrac{N_x}{N_y} y^2 \right) + ax + by + c$,

where a, b, c are constants which may be chosen by placing additional restrictions on z.

For example, let

(8) $\qquad u = \dfrac{x^2 + y^2}{2R}$,

$\mathcal{N}$ be a disc of diameter d and $\mathcal{N}_i$ as shown in Figure 2, where N is a

positive integer.
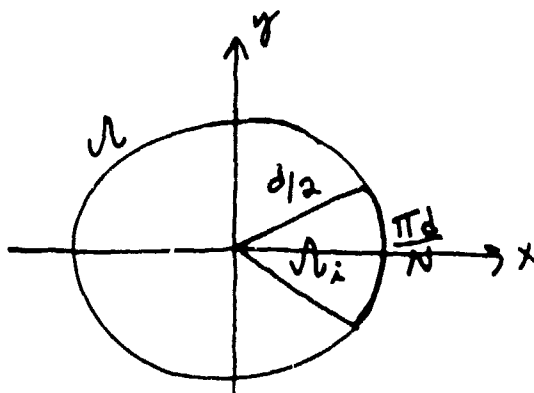
Figure 2

Consider z and u in polar coordinates and assume that z is symmetric about the x axis with

$$z(0,0) \quad = \quad u(0,0)$$

and

$$z(d/2, \pi/N) \quad = \quad u(d/2, \pi/N).$$

Then

$$z = \frac{r^2}{2R}\left(\cos^2\theta - \frac{N_x}{N_y}\sin^2\theta\right) + \frac{d}{4R}\left(1 + \frac{N_x}{N_y}\right)\frac{B^2}{A} r\cos\theta$$

and

$$u - z = \frac{1}{2R}\left(1 + \frac{N_x}{N_y}\right)\left[r^2\sin^2\theta - \frac{d}{2}\frac{B^2}{A}r\cos\theta\right]$$

where $A = \cos \pi/N$ and $B = \sin \pi/N$. Hence

$$\varepsilon^2 = \frac{\left[\frac{1}{2R}\left(1 + \frac{N_x}{N_y}\right)\right]^2}{\frac{\pi(d/2)^2}{N}} \int_0^{d/2}\int_{-\pi/N}^{\pi/N}\left[r^2\sin^2\theta - \frac{d}{2}\frac{B^2}{A}r\cos\theta\right]^2 r\,d\theta\,dr$$

which gives

$$\varepsilon^2 = \frac{N}{\pi}\left[\frac{1}{2R}\left(1 + \frac{N_x}{N_y}\right)\right]^2 \left(\frac{d}{2}\right)^4\left[-\frac{1}{12}B^3A - \frac{1}{8}AB + \frac{\pi}{8N} + \frac{\pi}{4N}\frac{B^4}{A} - \frac{1}{60}\frac{B^5}{A}\right]$$

and

(9) $$\frac{\varepsilon}{d} = \frac{1}{2Rd}\left(1 + \frac{N_x}{N_y}\right)\left(\frac{d}{2}\right)^2 \sqrt{\frac{N}{\pi}\left(-\frac{1}{12}B^3A - \frac{1}{8}AB + \frac{\pi}{8N} + \frac{\pi}{4N}\frac{B^4}{A} - \frac{1}{60}\frac{B^5}{A}\right)}.$$

As an example, suppose R = 2F, where F is the focal length, and that F/d = 2. Then for N=12 we have

$$(10) \qquad \frac{\varepsilon}{d} \quad = \quad .00116(1 + \frac{N_x}{N_y}) ;$$

which gives, for $N_x = N_y$,

$$(11) \qquad \frac{\varepsilon}{d} \quad = \quad .00232.$$

Similarly, for N = 60 we have

$$(12) \qquad \frac{\varepsilon}{d} \quad = \quad .93822 \times 10^{-4}.$$

Another feasible restriction to place on z is to require that 't agree with u on the boundary of $\mathcal{N}_i$. That is, z is the solution of

(13a) $\qquad N_x z_{xx} + N_y z_{yy} = 0$ on $\mathcal{N}_i$

(13b) $\qquad\qquad\qquad z = u$ on $\partial \mathcal{N}_i$.

If we let

$$\hat{y} = \sqrt{\frac{N_y}{N_x}} \; y,$$

$\hat{\mathcal{N}}_i$ be the image of $\mathcal{N}_i$ under this transformation,

$$\hat{z}(x,\hat{y}) = z(x,\hat{y})$$

and

$$\hat{u}(x,\hat{y}) = u(x,\hat{y}),$$

then (13) becomes

(14a) $\qquad \Delta \hat{z} = 0$ on $\hat{\mathcal{N}}_i$

(14b) $\qquad \hat{z} = \hat{u}$ on $\partial \hat{\mathcal{N}}_i$.

The solution to (14) may be approximated by a variety of numerical techniques. As a simple example, consider a finite difference approach for the case in which $\mathcal{N}_i$ is a rectangle. Suppose

$\qquad N_x = .004$ lb/in,

$\qquad N_y = .0015$ lb/in,

$\qquad d = 1968.5$ in (50m),

$$R = 2F,$$

$$F = d/2,$$

$$\mathcal{N}_i = [600, 700] \times [-25, 25],$$

$$u = \frac{x^2 + y^2}{2R},$$

$$\Delta x = .5$$

and

$$\Delta y = .5 \sqrt{\frac{N_y}{N_x}},$$

then by the Gauss-Seidel method one obtains a root-mean-square error

$$\varepsilon_i = .00211$$

for $\mathcal{N}_i$.

CONCLUSIONS AND RECOMMENDATIONS

The equilibrium equation (4) can be used to find the root-mean-square surface error for a variety of designs, provided that N_x and N_y are known. In general this may not be the case; a more useful model of the antenna surface would be an iterative approach:

i) Given the loads applied to the mesh and the present surface, find the resultant stresses,

ii) Using the stresses, find the deformed surface.

These steps could be repeated until the rms error reached the desired accuracy. In applying step ii) the nonlinearity of the stress-strain law for the mesh [3] should be taken into account.

REFERENCES

1. Large Space Systems Technology-1980, NASA CP-2168.

2. Flugge, W., Stresses in Shells, Springer-Verlag, 1960.

3. Tang,S.,Boyle, R., Whiteside, J. and Anderson, R.," A Nonlinear Stress-Strain Law For Metallic Meshes", Conference on Advanced Technology for Future Space Systems, Hampton,VA,May 8-10, 1979.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

RAY TRACING STUDIES OF JUPITER'S MAGNETOSPHERE

Prepared By: N. Frank Six

Academic Rank: Professor/Department Chairman

University and Department: Western Kentucky University
                                    Department of Physics & Astronomy

NASA/MSFC: Space Sciences Laboratory
  Division: Solar-Terrestrial
  Branch: Magnetospheric Physics

MSFC Counterpart: James L. Green, ES53

Date: October, 1981

Contract No.:   NGT 01-008-021
                                    The University of Alabama in Huntsville

RAY TRACING STUDIES OF JUPITER'S MAGNETOSPHERE

BY

N. Frank Six
Department of Physics & Astronomy
Western Kentucky University
Bowling Green, Kentucky

ABSTRACT

Recently the Planetary Radio Astronomy Experiment (PRA) on the Voyager 1 and 2 spacecrafts obtained dramatic new information both because of its proximity to Jupiter and because it was able to measure the radio emissions to much lower frequencies than has been possible from earth and near-earth observatories. The data reveal a nested family of arcs in the decametric spectral region and strong kilometric/hectometric emissions.

The objectives of this study are the following:

1) calculate raypaths for decametric wavelength radiation in Jupiter's magnetosphere;

2) compare the model-dependent raypaths with the Voyager observations; and

3) deduce characteristics of the source regions and the influence of propagation effects.

Since the Jovian magnetosphere with its complex magnetic field and plasma distribution (including Io's plasma torus) is highly anisotropic for the propagation of low frequency radio waves, a three-dimensional ray tracing program has been employed to calculate the raypaths. Families of rays have been launched at particular angles with respect to the magnetic field lines to generate conical sheets of radiation for various frequencies and various source locations. As the planet's magnetic field rotates, these warped sheets of radiation sweep past the observer, producing signatures in frequency versus time plots. These signatures match some of those found in the Voyager data. The greatest propagation effects occur in and around the source regions in the Io auroral oval.

ACKNOWLEDGEMENT

LIST OF FIGURES

INTRODUCTION

For nearly three decades, intense and variable decametric radio emissions from Jupiter have been observed from the earth over the frequency range 5 MHz to 40 MHz. Near-earth satellites have extended the low frequency limit set by the terrestrial ionosphere down to a frequency of 425 kHz. The radiation is characterized by a high degree of elliptical polarization, complex dynamic spectra, and an upper cutoff frequency of ∿39.5 MHz. It's probability of occurrence, dynamic spectral characteristics, polarization, and intensity are related to the central meridian longitude of Jupiter and to the orbital phase of Io. A review of these observations is given by Carr and Desch (1976) and references contained within.

While the nature and origin of these emissions have stimulated much theoretical work (e.g. Smith, 1976), the emission mechanisms, the location of the sources of emission, the Io-magnetosphere interaction, and many properties of the magnetospheric plasma are still unknown. Recently the Planetary Radio Astronomy Experiment (PRA) on the Voyager 1 and 2 spacecrafts obtained dramatic new information about Jupiter both because of its proximity to Jupiter and because it was able to measure the radio emissions to much lower frequencies than has been possible from earth and near-earth observatories. The data reveal a nested family of arcs in the decametric spectral region and strong kilometric/hectometric emissions. Encounter observations show unique spectral characteristics due to the spacecrafts' proximity to Jupiter.

This study involves implementing a versatile ray tracing program in conjunction with detailed models of Jupiter's magnetic field and magnetospheric plasma distribution, and the PRA observations, in order to understand the emission mechanisms, source locations, and propagation effects.

The Jovian magnetosphere with its complex magnetic field and plasma distribution (including the recently discovered plasma torus) is highly anisotropic for the propagation of low frequency radio waves. This medium not only influences the emission mechanisms, it also influences the propagation of the radiation from its source to the observer. Ray tracing studies carried out by Green and Gurnett (1980) and by several members of the Planetary Radio Astronomy Voyager team indicate that strong bending of kilometric emissions takes place in the plasma torus. In addition, from ray tracing calculations of whistlers in the Jovian magnetosphere, Menietti and Gurnett (1980) have been able to match the whistler dispersion characteristics and the torus temperature with the Voyager 1 observations. The whistler dispersion characteristics are heavily dependent on the density along the ray path and thereby produce an additional check on not only the densities within the torus (as determined by Warwick, et al., 1979) but also the density along the field lines. The conclusion

from these early studies is that much of the emission spectrum of Jovian kilometric radiation observed by both Voyagers 1 and 2 can be explained by propagation effects.

One result of the ray tracing studies of Green and Gurnett (1980) is to show that there exists shadow regions in which the source cannot transmit directly to the observer. A knowledge of these regions combined with the spacecraft data places strong constraints on both the source location and the emission mechanisms.

The results of previous studies highlight the importance of ray tracing. Propagation mode, source location, angular distribution, plasma distribution, and magnetic field configuration all play a role in the calculations. These characteristics for Jovian decametric radiation are virtually unknown and from ray tracing calculations much knowledge can be gained. A ray tracing study of Jovian decametric radiation which is in agreement with the observed emission patterns should give important clues as to the generation mechanism of this radiation.

OBJECTIVES

Using techniques which have proven successful in explaining the
interactions of particles and waves in the terrestrial magnetosphere,
models of Jupiter's magnetic environment have been formulated and tested
for validity using the data accumulated by the Pioneer and the Voyager
spacecraft. Three dimensional ray tracing studies of the Jovian magneto-
sphere have been performed in order to understand the origins, emission
mechanisms, and characteristic modes of propagation of the low frequency
(1 khz to 50 MHz) radiation from Jupiter. The objectives of the inves-
tigation are the following:

1. Using a three-dimensional ray tracing program which includes
   magnetic field and plasma models, calculate raypaths for dec-
   ametric wavelength radiation in Jupiter's magnetosphere.

2. Compare the model-dependent, theoretical raypaths with the
   observations made by the Voyager Planetary Radio Astronomy
   experiments.

3. Deduce, where possible, the location of the source regions,
   the influence of propagation effects, and characteristics
   of the generating mechanism.

VOYAGER DATA

The goal of this project, simply stated, is to extract as much information as possible from the comprehensive Voyager/Planetary Radio Astronomy (PRA) experiment data taken near the planet Jupiter. Conclusions regarding the source of the sporadic emissions - locations, mechanism, propagation effects - are the desired outcome.

The PRA experiment packages on the Voyager 1 and 2 spacecrafts each consist of an orthogonal pair of 10 meter monopoles, connected so that left hand and right hand polarizations can be distinguished. The radiometers measure the total power and the polarization sense of the incoming signal by sweeping through a low frequency band and a high frequency band. In the "stepping" mode, the receiver step tunes through 70 channels in the low band: 1.2 kHz - 1.3 MHz (each step spaced 19.2 kHz from the previous step; bandwidth: 1 kHz), and through 128 channels in the high band: 1.2 MHz - 40.5 MHz (steps separated by 307.2 kHz; bandwidth: 200 kHz). A complete sweep through both bands (198 frequency channels) is accomplished in 6 seconds. Each step alternates in polarization between right hand and left hand. In this study, the data used for comparison purposes had been averaged over 48 seconds (8 sweeps). The noise from the on-board high speed switching logic was so bad that the useable high band data is limited to the period $\pm$ 1½ months from encounter. Figure 1 shows observations of the Jovian decametric radiation made by Voyager 1. The data is presented in the frequency time domain, and much of the spacecraft interference has been removed. Note the nested "arcs" representative of the sporadic decametric emission. The horizontal bands are noise which was unsuccessfully repressed. A great deal of energy has been expended developing algorithms which would remove the spacecraft interference and reveal the less intense Jupiter signals. The nested arcs, both vertex-early and vertex-late, are the predominant characteristic observed in frequency-time plots such as shown in Figure 1.
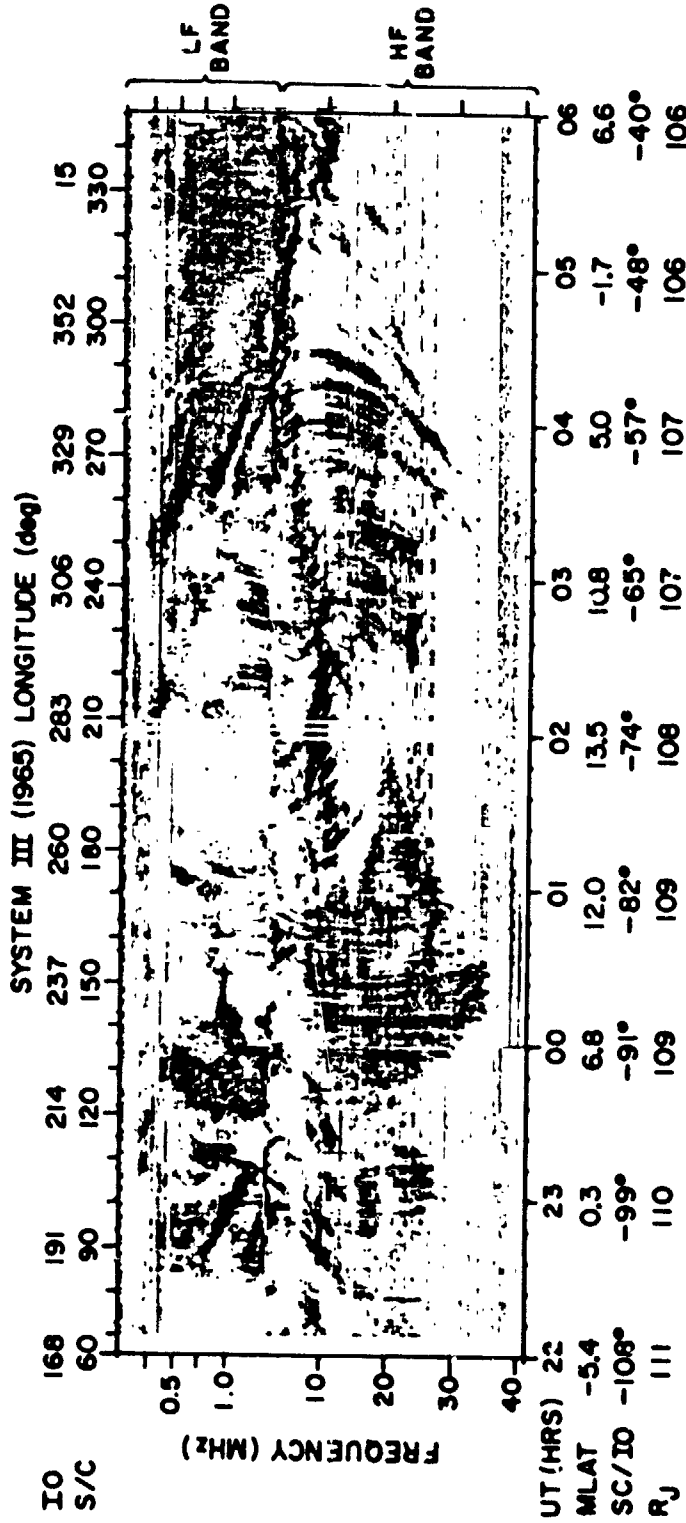
Figure 1: JOVIAN DECAMETRIC RADIATION SIGNATURES

(from Warwick et al., 1979)

RADIATION PATH MODELING

The major effort has been the development of a Jovian radiation path
model using techniques which have proven successful in explaining the
interactions of particles and waves in the terrestrial magnetosphere.
The description of Jupiter's magnetic field, adopted for this study, is
that published by Acuna and Ness (1976). A spherical harmonic analysis
was used to calculate the components of $\vec{B}$ and their derivatives. The
plasma distribution chosen was that of Sentman and Goertz (1977) augmented
by the Warwick torus (1979). Figure 2 illustrates the chosen plasma en-
vironment (plasma frequency contours) in a magnetic meridian plane.

Radiation paths are generated by a three dimensional ray tracing
program based on the Stix cold plasma formulation of the index of re-
fraction and Haselgrove's set of first order differential equations (1955).
A source position (radial distance from the center of Jupiter, System III
longitude, and System III latitude) from which waves of a specified fre-
quency are to be launched is input, along with the specified launch angle
with respect to the magnetic field. At the origin of the ray, the pro-
gram calculates an index of refraction surface based on the plasma and
magnetic field parameters. Next the program takes an incremental step
in the direction perpendicular to the index of refraction surface, i.e.,
in the direction of propagation, and determines the coordinates of this
new position on the raypath. Snell's law is used to conserve the compo-
nent of the refractive index parallel to planes of constant electron
number density. Another index of refraction surface is calculated at
this new point and the procedure is repeated. In this manner, a raypath
is generated. The behavior of a ray's path is determined by the follow-
ing initial conditions: the frequency of the wave, the wave normal angle,
the coordinates of the source (launch point), the propagation mode, and
the magnetospheric model. The flowchart of the ray tracing program is
shown in Figure 3. The nine subroutines are described below.

1. DAMJUP is the main program which controls logic flow, specifies
    parameters, specifies accuracy, controls action of all sub-
    routines.

2. MAGNET describes the magnetic field of Jupiter (Acuna and Ness)
    and utilizes a spherical harmonic analysis to calculate the
    magnitude of the field, B_0, the spherical components B_r, B_θ,
    B_ϕ, and the derivatives dB/dr, dB/dθ, dB/dϕ at a specified
    point.

3. DHPCG is a subroutine (IBM) which obtains an approximate solu-
    tion to a general system of first order ordinary differential
    equations (Haselgrove) with given initial values. A Runge-
    Kutta procedure is used to obtain the starting values, then
    Hamming's modified predictor-corrector method is initiated.

ELECTRON PLASMA FREQUENCY (kHz) CONTOURS
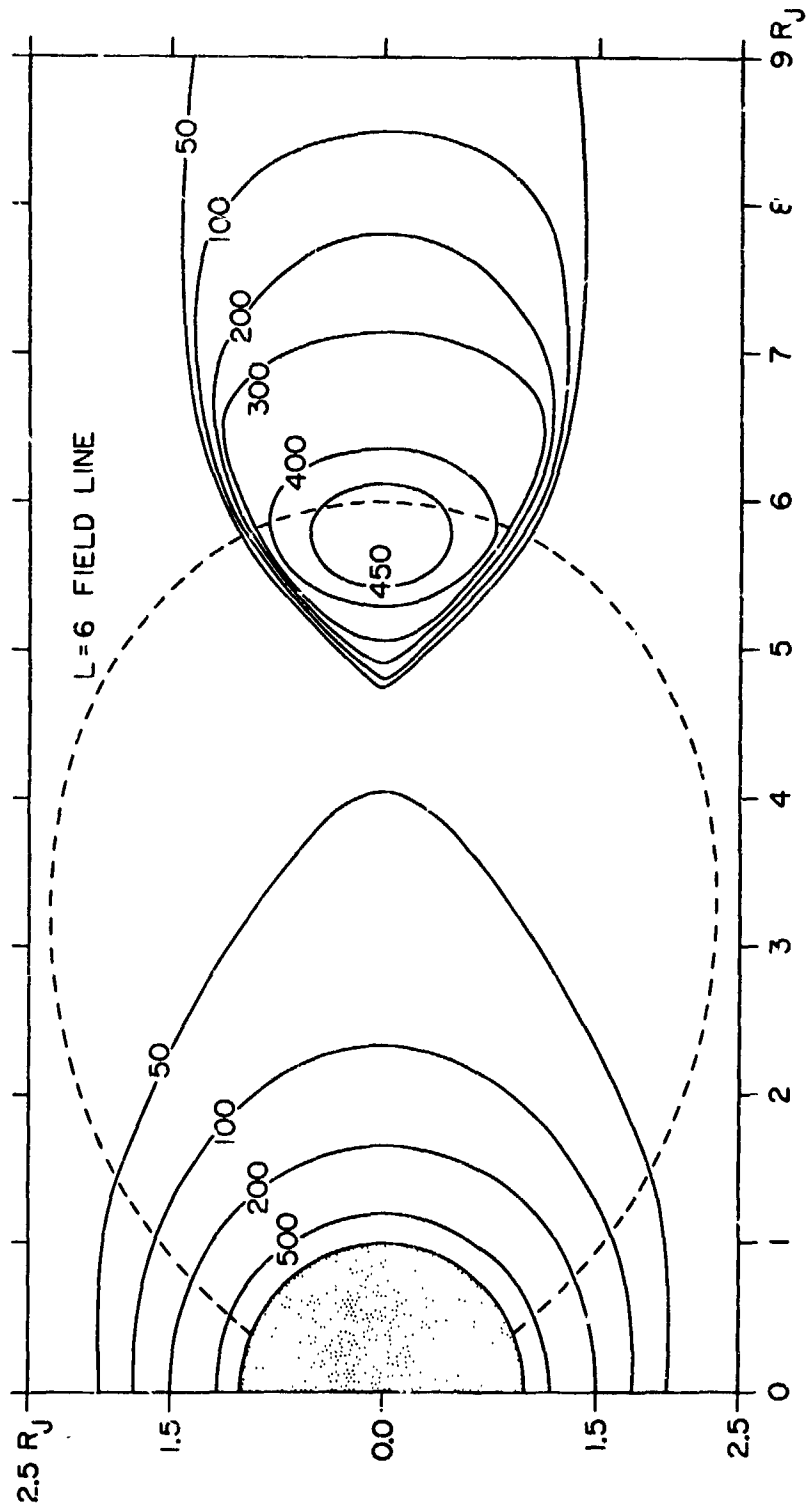
Figure 2: MODEL OF JUPITER'S MAGNETIC FIELD AND PLASMA
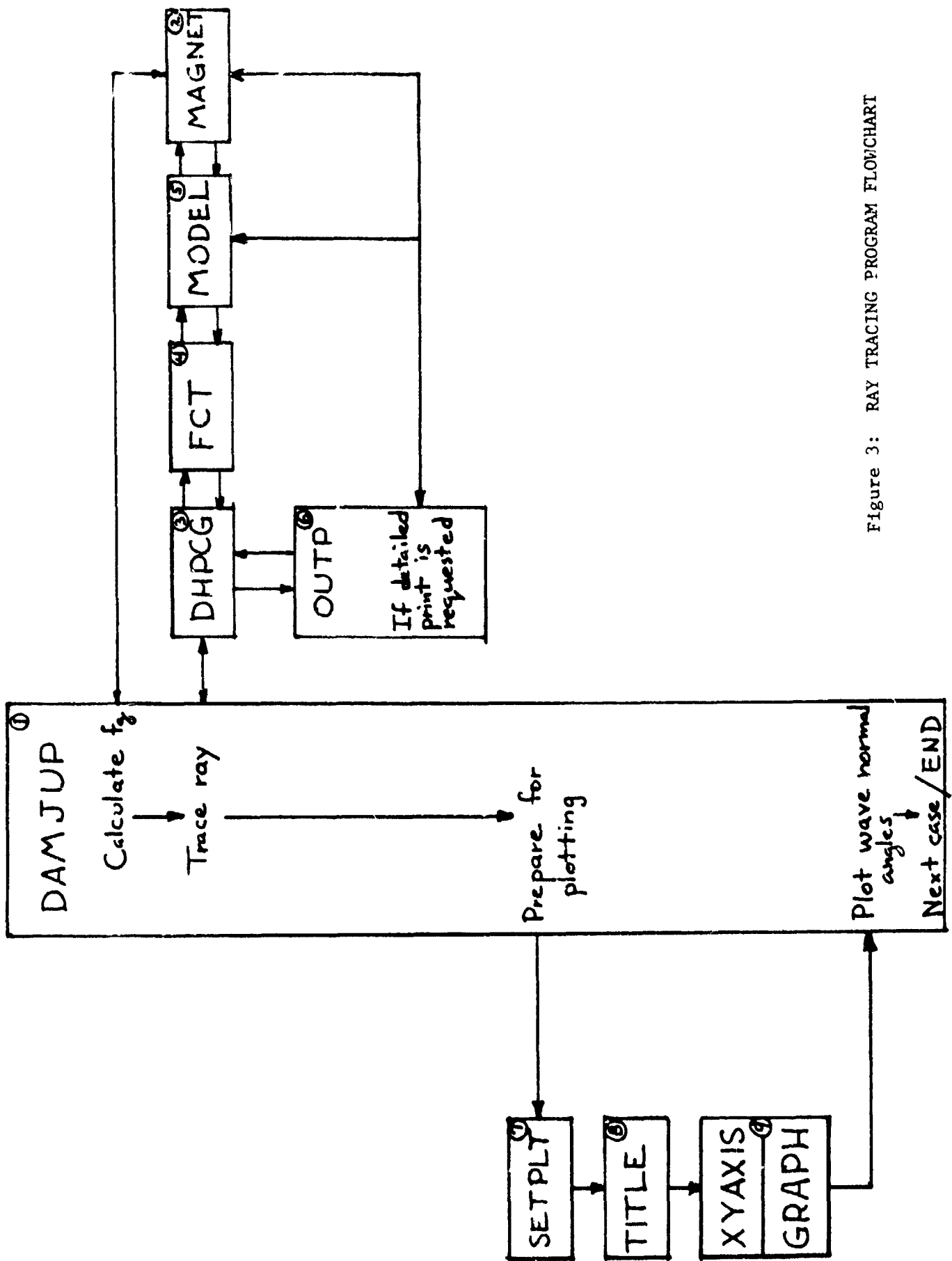
Figure 3: RAY TRACING PROGRAM FLOWCHART

This procedure controls the accuracy and adjusts the step size h. Knowing the results (index of refraction surface) at the equidistant points X_{j-3}, X_{j-2}, X_{j-1}, and X_j, the results at point (X_j + h) are calculated. If the error function value is greater than a given tolerance, then the increment h is halved, and the calculation repeated. If the error function is less than a given tolerance, the result is passed on to the FCT subroutine. When the error function is less than 1/50 the established tolerance, the next calculation is carried out with a step size of 2h.

4. FCT calculates the index of refraction surface $n(\psi)$ at a point on the propagation path using the Stix (cold plasma) formulation. Also, the tangent to this surface at a point on the propagation path is determined. These results are passed on to DHPCG, which takes a step of increment h in the direction perpendicular to the tangent to $n(\psi)$, that is, in the direction of the group velocity (the propagation path*). DHPCG determines the coordinates of the next point along the propagation path, and asks FCT to calculate another $n(\psi)$ surface. (*Direction of energy flow; direction of mean Poynting vector.)

5. MODEL describes the plasma environment (Sentman and Goertz) of Jupiter, including the Io torus (Warwick). It calculates the plasma frequency f_p, the gyrofrequency f_g (calls MAGNET), and the derivatives at a given point.

6. OUTP stores in the plotting arrays the coordinates of a point on the ray path which has been generated in DHPCG. These coordinates are r, θ, ϕ, ψ (wave normal angle), and two angles, delta and epsilon, which designate the direction of the phase velocity. If a detailed printout has been called for by selection of that PRINT mode, then OUTP delivers this information to the printer and calls on MAGNET and MODEL to provide field and plasma parameters at the updated ray path location.

7. SETPLT sets microfiche scales for plotter.

8. TITLE sets fiche title.

9. XY AXIS and GRAPH determines XY grid and labels the axes.

Let us assume that the source of the decametric emission is located on a field line through Io. This is a likely source location since the Io flux tube footprint on Jupiter's ionosphere is expected to be the site of field line currents which could be responsible for particle acceleration. This probable source location is illustrated in Figure 4. Further, let us assume the radiation is cyclotron emission, hence the origin of the rays is given by the magnetic field subroutine, MAGNET. The radiation would escape only in the R-X mode. Families of rays launched at the same wave normal angle produce conical sheets of radiation, warped by the plasma and by the asymmetries in the field.

JUPITER'S MAGNETOSPHERE
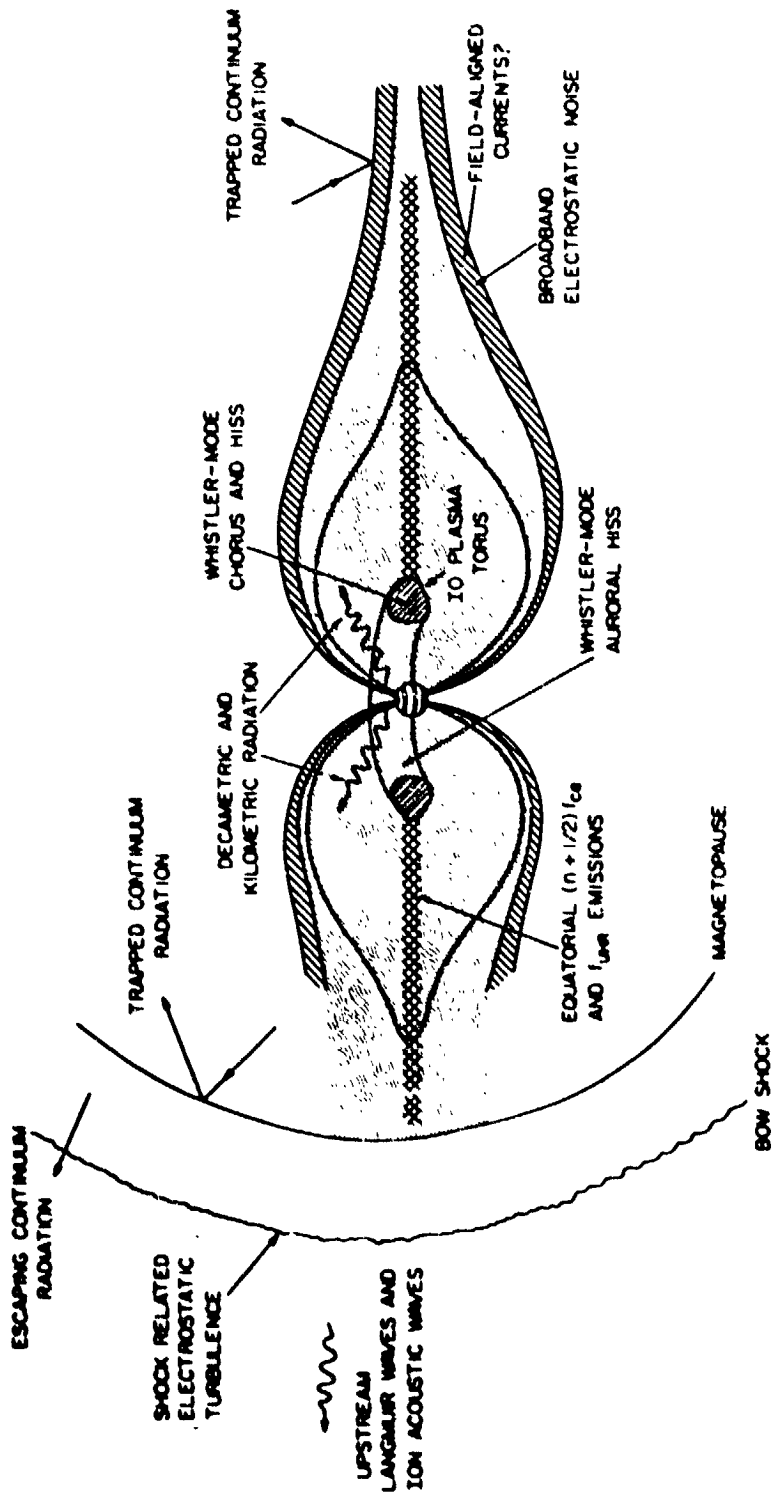
Figure 4: RADIATION SOURCES IN JUPITER'S MAGNETOSPHERE

(from Gurnett 80-39)

RESULTS: CALCULATED RADIATION SHEETS

It is believed that the decametric radiation is generated in the R-X
mode. Choosing the appropriate sign in the Stix formulation of the index
of refraction, raypaths have been calculated at seven frequencies (2, 5,
10, 15, 20, 25, and 30 MHz) originating from both northern and southern
hemisphere sources at 36 separate longitudes (spaced every 10° around the
planet) and at wave normal angles in 10° increments between 60° and 130°.
For each different combination of the four parameters: hemisphere,
longitude, frequency and wave normal angle, 36 rays are launched at 10°
increments around B̌. This family of raypaths can be imagined as the skel-
eton of a conical sheet of radiation, just as the ribs of an umbrella
relate to the fabric. Intersections of these conical sheets of radiation
with radial shells around Jupiter at distances of 2, 4, 10, 20, 40, 60,
80, 90, 100, and 150 Jupiter radii are presented as Mercator projections.
Figure 5 illustrates a set of four radiation sheets, one for each of the
wave normal angles 60°, 70°, 80° and 90°. These rays were launched with
a frequency of 2.0 MHz from a northern hemisphere source at a longitude
of approximately 160°. The Mercator projection is of the Jovicentric
radial shell at 150 Jupiter radii. Note that there are 36 rays in each
radiation sheet. This figure represents a vast catalog of raypaths cal-
culated by the three dimensional ray tracing program described earlier
in this report.
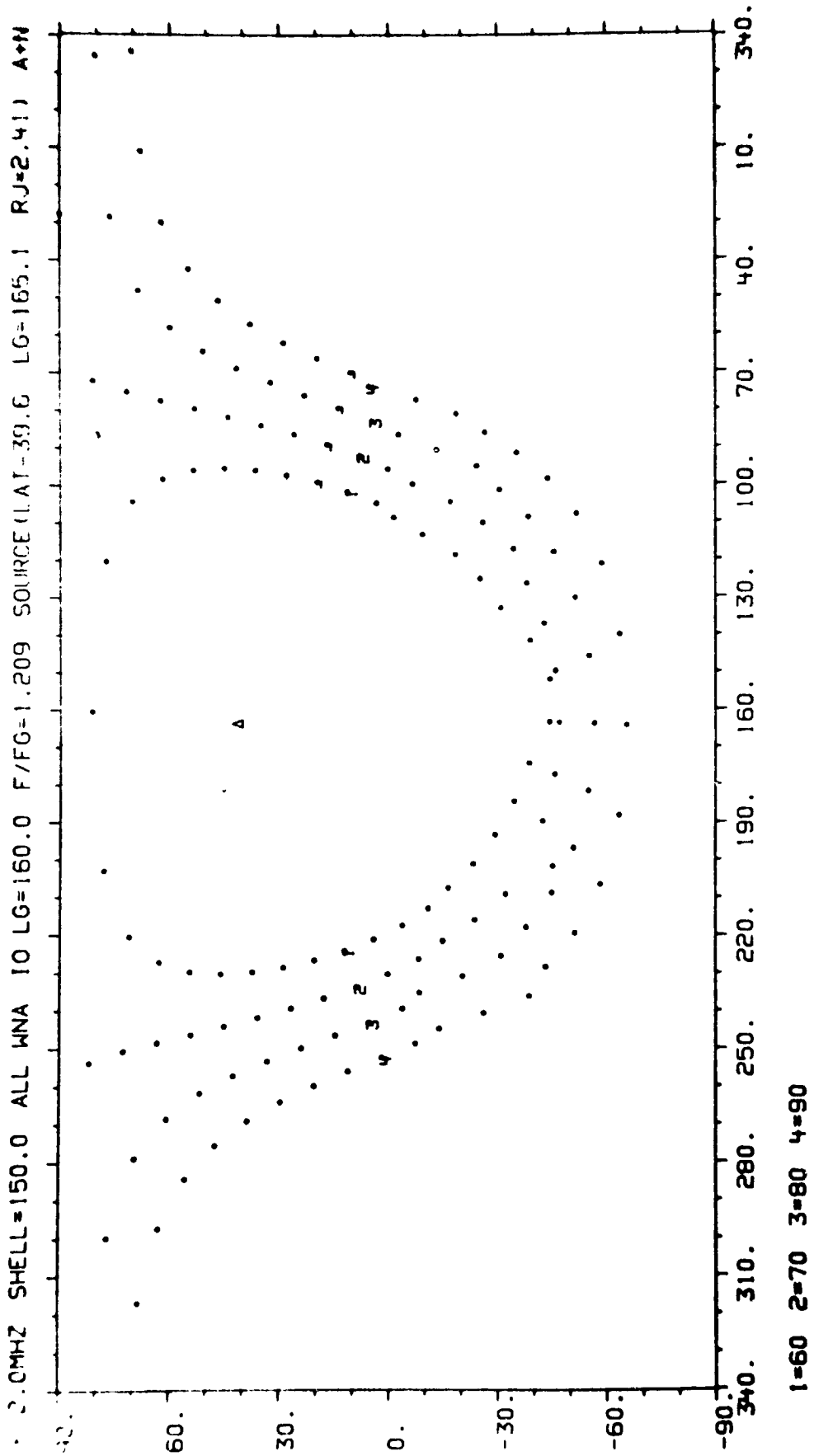
Figure 5: 2 MHz R-X MODE RADIATION SHEETS ORIGINATING
FROM A SOURCE IN JUPITER'S NORTHERN HEMISPHERE

CONCLUSIONS

The reason for calculating raypaths, such as those shown in Figure 5, is to examine their intersection with the Voyagers' trajectories and find out to what extent the model-dependent radiation sheets yield signatures (in frequency-time space) which match the actual observations. In Figure 6, the 20 MHz radiation sheet is shown superimposed with radiation sheets of 12.8, 8, 5 and 2 MHz. The spacecraft is slightly above the equatorial plane as illustrated by the dashed line. As Jupiter rotates, the radiation sheets sweep over the spacecraft (alternatively, one might imagine the spacecraft moving along the dashed line from right to left toward higher longitudes). Should Voyager be near Jovian longitude 90°, the PRA experiment would receive the various frequencies of radiation in descending order: first 20 MHz, then 12.8, 8, 5 and 2 MHz. Should Voyager be near Jovian longitude 270°, the radiation sheets would fall on the antennas in ascending order: first 2 MHz, then 5, 8, 12.8 and 20 MHz. Signatures of both types are actually observed. The model we have employed does indeed generate decametric arcs as seen by an observer riding with Voyager. In Figure 6, the location of the Northern hemisphere source of the Jupiter emission is designated by a small triangle. The Mercator projection is made at the 150 R_J shell. Io was at longitude 180° and all waves were launched at a wave normal angle of 90°.1. If the mechanism of radiation produces nested cones of emission originating on field lines at high northern and southern latitudes, then the signatures observed in the Voyager data can be explained as a geometric effect: the radiation sheets are warped by the plasma near the planet and by the twist (azimuthal asymmetry) in the magnetic field such that overlapping occurs. An observer riding with Voyager intermittantly crosses thin sheets of radiation as the source co-rotates with the planet. It is as though the observer were viewing the ruffles in a twirling ballerina's tutu.

An examination of the catalog of decametric radiation sheets generated by the ray tracing procedures described in this report, shows that the Io plasma torus has minimal effect at these wavelengths. The major distortion by plasma occurs at higher latitudes nearer the surface of the planet. This is also the region of greatest azimuthal asymmetry in the magnetic field.

A preliminary conclusion regarding the high frequency limit of the decametric emissions has been drawn. Examining the Voyager 2 data of July 16, 1979, one notices a well-defined boundary to the highest frequencies observed during the period 0200-0800 UT. The ray tracing program was used to calculate raypaths for frequencies which were observed and for frequencies which were just above the high frequency boundary. Those rays generated by the program for the frequencies actually observed by the spacecraft all propagated through the magnetized plasma environment. The rays for the frequencies which were not observed by

Figure 6: DECAMETRIC RADIATION SHEETS CALCULATED BY RAY TRACING

F=20.0MHZ SHELL=150.0 PSI=90.1 IO LG=180.0 F/FG=1.212 SOURCE(LAT=47.6 LG=175.6 RJ=1.20)
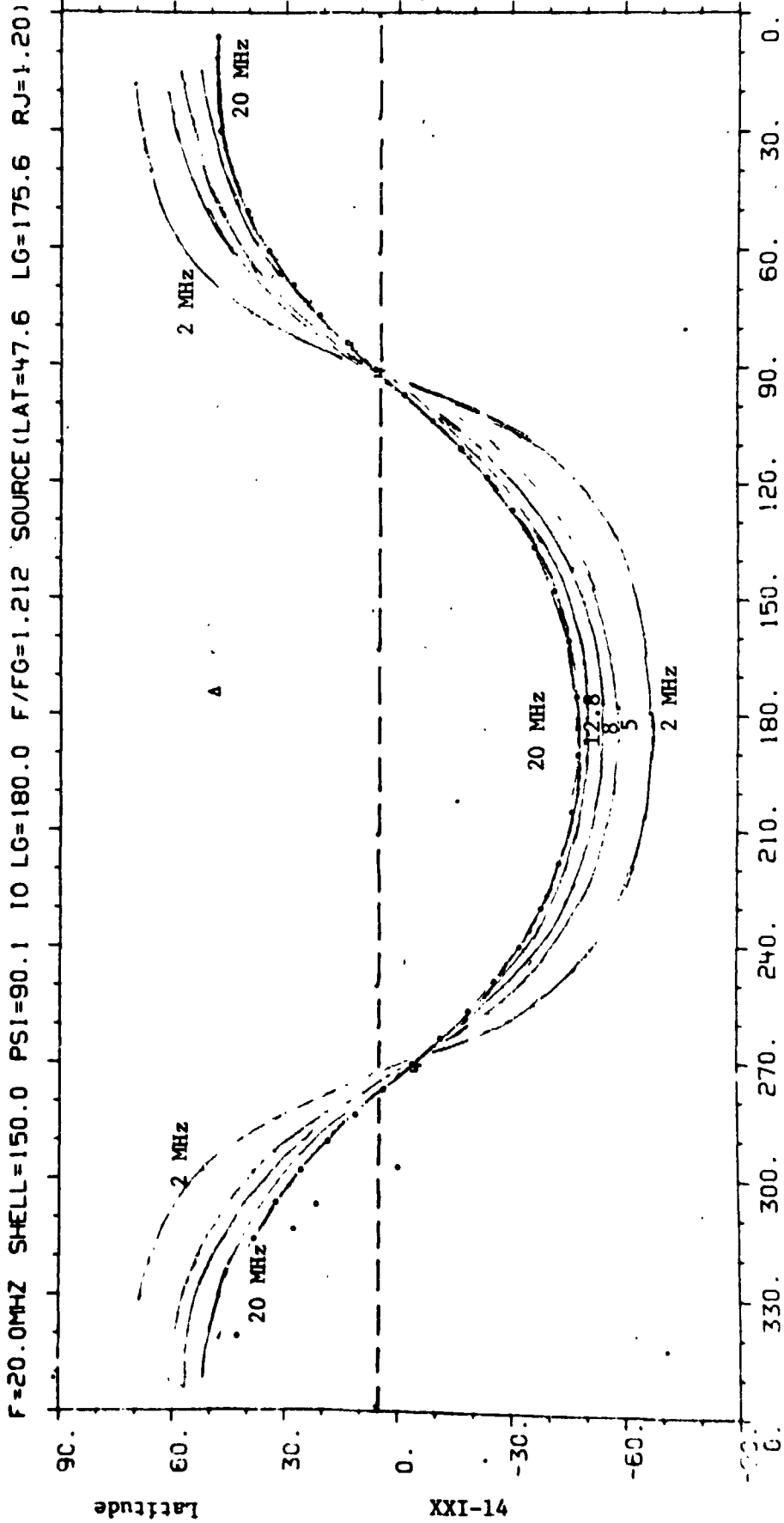
Jovian Longitude System III

Latitude

XXI-14

Voyager did not propagate through the magnetized plasma. Propagation, for these raypaths, was inhibited at the R=0 cutoff, i.e., where the extraordinary wave was reflected.

Additional comparisons of the model dependent raypaths with the PRA observations is necessary. To facilitate this cross-referencing, the observational data should be reformatted. Once the spacecraft generated interference is repressed, the data should be sorted according to fixed Io longitudes, as though Io were anchored in the Jovian magnetic field and the spacecraft circled the planet. If field aligned currents in the Io flux tube are triggering the emissions, then the patterns in the data, when formatted as suggested, should pinpoint the source locations. This software effort is now underway.

REFERENCES

1. Carr, T.D., and Desch, M.D., "Recent Decametric and Hectometric Observations of Jupiter," Jupiter, ed. T. Gehrels, 1976, pp. 693-737.

2. Smith, R.A., "Models of Jovian Decametric Radiation," Jupiter, ed. T. Gehrels, 1976, pp. 1146-1189.

3. Green, J.L., and Gurnett, D.A., "Ray Tracing of Jovian Kilometric Radiation," submitted to Geophysical Research Letters, 1980.

4. Menietti, J.D., and Gurnett, D.A., "Whistler Propagation in the Jovian Magnetosphere," Geophysical Research Letters, Vol. 7, No. 1, January 1980, pp. 49-52.

5. Warwick, J.W., Pearce, J.B., Riddle, A.C., Alexander, J.K., Desch, M.D., Kaiser, M.L., Thieman, J.R., Carr, T.D., Gulkis S., Boischot, A., Harvey, C.C., Pedersen, B.M., "Voyager 1 Planetary Radio Astronomy Observations Near Jupiter," Science, Vol. 204, 1 June 1979, pp. 995-998.

6. Acuna, M.H., and Ness, N.F., "Results from the GSFC Fluxgate Magnetometer on Pioneer 11", Jupiter, ed. T. Gehrels, 1976, pp 830-847.

7. Sentman, D.C., and Goertz, C.K., "Whistler Mode Noise in Jupiter's Inner Magnetosphere", Journal of Geophysical Research 83, 1977, 322.

8. Haselgrove, J., "Ray Theory and a New Method of Ray Tracking" London Physical Society, Report on Conference on the Physics of the Ionosphere, 1955, pp 355-364.

9. Gurnett, D.A., and Scarf, F.L., "Plasma Waves in the Jovian Magnetosphere", U. of Iowa, 80-39.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

Title: THUNDERSTORM ELECTRIC FIELD MODELLING

Prepared by: Arjun Tan
               Math-Physics Department
               Newberry College
               Newberry, SC 29108

NASA/MSFC: Atmospheric Sciences Division
               Space Sciences Laboratory

MSFC Counterpart: Dr. Hugh J. Christian, Jr.

THUNDERSTORM ELECTRIC FIELD MODELLING

by

Arjun Tan
Newberry College, Newberry, SC   29108

Principal: Dr. Hugh J. Christian, ES-83, NASA-MSFC, Huntsville, AL   35812

ABSTRACT

A handful of electric field models of thunderstorms are found in the literature. Holzer and Saxon (1952) calculated the electric field and conduction currents of a simple dipolar thundercloud, assuming exponentially increasing conductivity of air with altitude. Anderson and Freier (1969) studied the effects of a time-varying potential. Park and Dejnakarintra (1973) extended the mapping of electric fields to ionospheric heights for upper midlatitudes, where the conductivity is a tensor. Burke (1975) extended the modelling to disk-shaped charge distributions and compared their results with high altitude balloon flight data.

Since considerable amounts of data now exist of electric field measurements inside thunderclouds (both from horizontal and vertical flights) apart from ground-based observations, the need for a more realistic model of thunderstorm electricity is now called for (Christian, 1976). In this proposed model, calculation of electric field vectors within and near an isolated thundercloud (which has a given volume charge distribution) is envisaged. Of particular interest is the calculation of the maximum field strength within the thundercloud. Later, the effects of screening layers, both above and below the thundercloud may be investigated, as well as time-dependent potential problem. The study can lead to a better understanding of the charge distributions in a thundercloud, which in turn may shed some light on the actual mechanism of charging of a thundercloud.

INTRODUCTION

While considerable amounts of experimental data exist in thunderstorm field measurements, only a handful of electric field models of thunderclouds are found in the literature. It was well-known that the ionosphere is a good conductor of electricity, the ground is a fairly good conductor while the atmosphere is a poor conductor. For this reason, the earliest theoretical studies of atmospheric electricity were confined to the "spherical capacitor" model (for a review, see Chalmers, 1967). These models did not take into account the variation of electrical conductivity of air with altitude. Since conductivity is proportional to the mobility of ions present in the atmosphere and since the mobility increases roughly exponentially with altitude on account of the exponential decrease of neutral number density, the increase of conductivity of the atmosphere with altitude is also roughly exponential (Cf., Cole and Pierce, 1965).

Holzer and Saxon (1952) were the first to incorporate the exponential increase of conductivity of air to calculate the electric fields and conduction currents due to a simple dipolar thundercloud. Freier (1967) and Anderson and Freier (1969) studied the effects of a time-varying potential. Park and Dejnakarintra (1973) extended the mapping of electric fields for mid- and upper-mid-latitudes, taking into account the anisotropy of the electrical conductivity there. Dejnakarintra and Park (1974) further studied how the effects of a time-varying field are propagated into the ionosphere. Burke (1975)

extended the modelling to disk-shaped charge distributions and compared her results with high altitude ballon flight data.

While the above-mentioned models serve as useful tools for comparing with the ground-based and high altitude electrical field measurements, they cannot furnish information regarding the distribution of charges in the thundercloud. The actual distribution is, of course, a volume distribution, quite unlike the point-dipole or disk-shaped charges hitherto considered. For this reason, a more realistic thundercloud model with a volume distribution of charges is a logical next step to be undertaken.

OBJECTIVES

The objective of this study is two-fold. First, beginning with
a more realistic volume distribution of charges, we propose to compute
the electric fields above, below and around the charge distribution,
and investigate what kind of distribution best fits the experimental
data. This model would also permit comparisons with the other existing
models. Of particular interest would be the study of the effects of
screening layers of charges (see MODEL) on the resulting electric fields
above and below the thundercl ud.

Secondly, the model would permit calculation of the electric fields
within the thunderstorm itself. This will enable us to compare with
the various in situ measurements that have been carried out and are
currently being undertaken. These include the mountain-top mwasurements
of Rust and Moore (1974), rocket measurements of Winn et al.(1974),
balloon-borne measurements of Benbrrok et al.(1974) and Winn and Byerly
(1975) and parachute-deployed measurements of Bering et al.(1977).
Although the actual charge distribution and electric field are highly
complex, from the calculated maximum field and the variations of the
of the fields, some rough estimates of the charge densities could be
inferred.

MODEL

The size and shape of a thundercloud vary considerably from cloud to cloud, with latitude and geographical location. The height of the thundercloud may range from 4 km to 20 km or more (Vonnegut and Moore, 1959) but a typical tropical thundercloud rises to an altitude of 8 to 12 km. While the size of the thundercloud can be easily estimated, finding the electrical structure of the thundercloud is a much more difficult problem. This is complicated by several factors, such as, highly complex electrical structure, high temporal variability caused by charging, lightning discharge and subsequent charging, high intra-cloud variability etc. All these plus the general lack of agreement on the basic electrification process contribute to the difficulty in modelling the thundercloud electrical structure.

Nonetheless some broad general features of thundercloud electricity have been observed ever since Benjamin Franklin. Observations continue to affirm an overall dipole nature of the thundercloud with a net positive charge in the upper regions and a net negative charge in the lower regions. The positive charge appears to be more diffused than the negative charge (Reynolds, 1953) and sometimes the negative charge is in a vertical column (Malan and Schonland, 1951). Frequently the dipole can be tilted from the vertical (Reynolds and Neill, 1955). Further, lightning and strong winds do complicate the charge structure continuously (Moore, et al., 1964).

The amounts of positive and negative charges in a thundercloud are also quite uncertain. The earlier estimate of 20C and -20C by Malan (1963) is now believed to be too conservative while Kasemir's (1965) estimate of a negative charge of -340C may be overly high.Charges of 100C as estimated by MacGorman and Few (1974) may be a more acceptable figure.

In addition to the primary positive and negative charges, there is evidence of "screening layers" both above and below the main charges (Cf., Hoppel and Phillips, 1971). Further, since Schonland (1928), "corona discharge" or flow of positive charges from the surface of the earth has been observed.

Whereas any representative picture of a typical thundercloud charge may be a far cry from actual reality, for our model, the distribution shown in Fig. 1 is assumed.

EQUATIONS

The electric field is derived from the negative gradient of the potential

$$\underline{E} = -\nabla\phi .$$

(1)

The electric field is also related to the current density in air $\underline{J_a}$

$$\underline{J_a} = \sigma\underline{E} ,$$

(2)

where the conductivity σ increases roughly exponentially with altitude with a scale height $1/(2k)$

$$\sigma_z = \sigma_o e^{2kz} .$$

(3)

Under steady-state conditions,

$$\nabla\cdot\underline{J} = 0 ,$$

(4)

where the total current density $\underline{J}$ is the sum of $\underline{J_a}$ and the convection current density at the source region $\underline{J_s}$

$$\underline{J} = \underline{J_a} + \underline{J_s} .$$

(5)

From (4) and (5), we get

$$\nabla\underline{J_a} = -\nabla\cdot\underline{J_s} .$$

(6)

Poisson's equation for the problem is given by

$$\nabla^2\phi = -\nabla\cdot\underline{E} = -\frac{\rho}{\varepsilon_u} ,$$

(7)

where ρ is the total space charge density and ε_u the permittivity of air.

Combining (1), (2), (3) and (6) we get

$$\sigma_2 \left(\nabla^2 + 2k \frac{\partial}{\partial z} \right) \phi = \nabla \cdot \vec{J}_s .$$

(8)

Substituting into (7), we get

$$\rho = 2k\varepsilon_c \frac{\partial \phi}{\partial z} - \frac{\varepsilon_c}{\sigma_2} \nabla \cdot \vec{J}_s$$
$$= \rho_a + \rho_s ,$$

(9)

where ρ_a is the space charge density in air and ρ_s is the space charge density of the source.

Equations (8) and (9) now give

$$\left[\nabla^2 + 2k \frac{\partial}{\partial z} \right] \phi = - \frac{\rho_s}{\varepsilon_c} .$$

(10)

In cylindrical coordinates, Eq.(10) takes the form

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \phi}{\partial \varphi^2} + \left(2k \frac{\partial}{\partial z} \right) \phi = - \frac{\rho_s}{\varepsilon_c} .$$

(11)

Equation(11) is solved numerically with the boundary conditions

$$\phi = 0 \quad at \quad z = 0 ,$$
$$\phi = V \quad at \quad z = H ,$$

where H is the upper boundary of the atmosphere.

XXII-8

REFERENCES

Anderson,F.J. and G.D.Freier, J.Geophys.Res. 74, 5390,1969.

Benbrooke,J.R.,J.W.Kern and W.R.Sheldon, J.Geophys.Res.79, 5289,1974.

Bering,E.A.,J.R.Benbrook and W.R.Sheldon, J.Geophys.Res.82, 1925,1977.

Burke,H.K.,Ph.D.Thesis,Rice University, 1975.

Chalmers,J.A.,Atmospheric Electricity,Pergamon, New York, 1967.

Cole,R.K. and E.T.Pierce, J.Geophys.Res. 70, 2735, 1965.

Dejnakarintra,M. and C.G.Park, J.Geophys.Res. 79, 1903, 1974.

Freier,G.,J.Atmos.Sci.36,1967,1979.

Holzer,R.E. and D.S.Saxon, J.Geophys.Res.57, 207,1952

Hoppel,W.A. and B. B.Phillips, J.Atmos. Sci. 28, 1258,1971.

Kasemir, in Problems of Atmospheric and Space Electricity, Elsevier, 1965.

MacGorman,D.R. and A.Few,Trans.AGU 56, 1130, 1974.

Malan,D.J.,Physics of Lightning, The English Univ. Press, London, 1963.

Malan, D.J. and V.F.J.Schonland,Proc.Roy.Soc. A209, 158, 1951.

Moore,C.B., B.Vonnegut, B.A.Stern and H.J.Survilas, J.Geophys.Res.65,1907, 1960.

Park, C.G. and M.Dejnakarintra, J.Geophys.Res.78, 6623, 1973.

Reynolds,S.E.,US Signal Corps Rept., Soccorro, 1954.

Reynolds,S.E. and H.W.Neill,J.Meteorol.12, 1, 1955.

Rust, W.D. and C.B.Moore, J.Roy.Met.Soc. 100, 450, 1974.

Schonland,B.F.J.,Proc.Roy.Soc.A118, 252, 1928.

Vonnegut,B. and C.B.Moore in Recent Advances in Atmospheric Electricity, Pergamon, 1959.

Winn,W.P. and L.G.Byerly,J.Roy.Met.Soc.101, 979, 1975.

Winn,W.P.,G.W.Schwede and C.B.Moore,J.Geophys.Res.,79,1761, 1974.

Altitude(km)

— 10

— 9   -5 C

— 8

— 7   +100 C

— 6

— 5

— 4   -100 C

— 3   +5 C

— 2

— 1
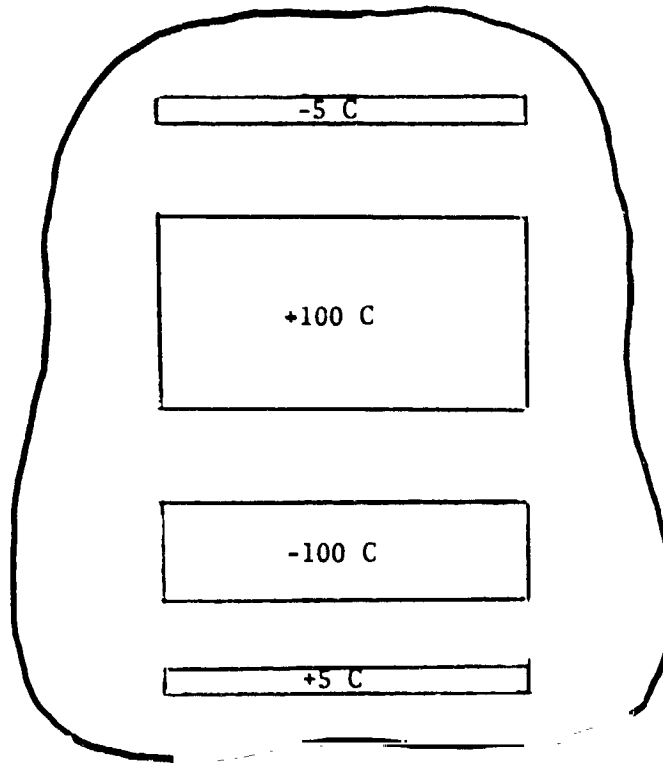
Fig.1. Stratified Thundercloud charge model.

NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

DETERMINATION OF THE EFFICIENCY OF A
PROPORTIONAL COUNTER FOR DETECTING X-RAYS

Prepared By: Dr. Alton C. Williams

Academic Rank: Assistant Professor

University and Department: Alabama A & M University
                                    Department of Physics

NASA/MSFC:
     Laboratory: Space Sciences
     Counterpart: Mr. M.C. Weisskopf

Date: August 14, 1981

Contract No.:   NGT 01-008-021
                                    The University of Alabama in Huntsville

I. INTRODUCTION

The Advanced X-ray Astrophysics Facility (AXAF) is to be a national x-ray observatory in space. At present, NASA is utilizing its x-ray test facilities to conduct studies to determine how to best construct the AXAF as well as to obtain empirical relations between x-ray scattering and the theories of visible light.

Basically, the testing at NASA consists of a parallel beam of x-rays (approximately monoenergetic) passing through a rectangular slit and scattering from two flat mirrors. The two mirrors are arranged in a periscope geometry so that the final reflected beam is parallel to the incident beam but displaced laterally by a given amount. One of the detectors used to intercept the reflected x-rays is a one dimensional gas-filled proportional counter which is sensitive to the position (in 1-dimension) at which the x-rays are incident within its detecting "window".

The total length of the anode wire of the proportional counter is 120 mm and this length can be divided electronically into a maximum of 1,024 parts. Hence, the output of an experimental run would be the number of incident x-rays that registered on each of the 1,024 channels. Each channel would represent an x-ray at a different spatial location and, hence, at a different scattering angle. In order to look at a wider range of scattering angles, the detector is placed on an optical table which can be rotated.

One of the problems that arises in the interpretation of the data has to do with the fact that the efficiency of the detector is a function of position. That is, the efficiency of the 1,024 channels are conceivably all different. What one would like in this case is to determine the relative efficiency of each channel with respect to some arbitrary reference channel. This would then allow the variation in the counts in each channel due to the variation in the efficiencies of the channels to be normalized out.

The following pages describe how the relative efficiencies can be determined from a particular set of experimental data.

PART II

I. ### The Desired Quantity

Each channel number, i, will have an absolute channel efficiency given by $E(i)$. It is desired to have the ratio

$$RCE(i) = \frac{E(i)}{E(i_0)} \tag{1}$$

for an arbitrary reference channel number i_0.

II. ### Relation Between Measured Quantities And The Relative Counter Efficiencies

What is measured is the quantity $C_k(i)$, the number of counts in channel number i when the table is at position P_k (arcseconds). The counts in channel i are related to the absolute efficiency via the relation.

$$C_k(i) = E(i) \, I_k(i) \tag{2}$$

where $I_k(i)$ is the flux at the position of channel number i when the table is at position P_k. The relative counter efficiencies are given by:

$$RCE(i) = \frac{C_k(i)}{C_{k_0}(i_0)} \frac{I_{k_0}(i_0)}{I_k(i)} \tag{3}$$

III. ### Relation Between Counts and Flux

Let $C_{k'}(i)$ = counts in channel i when the table is at position P_k'.
Since $RCE(i)$ is independent of table position, then from eq. (3)

$$\frac{C_k(i)}{C_{k_0}(i_0)} \frac{I_{k_0}(i_0)}{I_k(i)} = \frac{C_{k'}(i)}{C_{k_0}(i_0)} \frac{I_{k_0}(i_0)}{I_{k'}(i)}$$

or

$$\frac{C_{k'}(i)}{C_k(i)} = \frac{I_{k'}(i)}{I_k(i)} \tag{4}$$

IV. ### Calculation Of Channel Number For A Given Position

Let $\bar{D}$ = the number of channels/arcseconds of table movement   Then,
The number of arcseconds channel number i (at table position P_k)

is from the reference channel is given by

$$O_k(i) = \frac{(i - i_0)}{\bar{D}} - (P_k - P_{k_0})$$ (5)

Hence, if channel i (position P_k) and channel i' position $P_{k'}$ correspond to the same location in space, then

$$O_k(i) = O_{k'}(i')$$

or

$$i' = i + \bar{D}(P_{k'} - P_k)$$ (6)

V. <u>Why Calculation Is Not Straightforward</u>

To obtain RCE(i) from the measured counts, we need the ratio $\dfrac{I_{k_0}(i)}{I_k(i)}$ in eq. (3). This can be obtained using eq. (4) if

$$I_{k'}(i) = I_{k_0}(i_0) \quad .$$

That is, the spatial position of channel i (position k') must be same as the spatial position of channel i_0 (position k_0). From eq. (5) this means

$$i = i_0 + \bar{D}(P_{k'} - P_{k_0})$$

Since i_0 is fixed, i can vary only if $P_{k'}$ vary. That is, as many table positions as channels are needed to do the calculation straightforwardly.

VI. <u>The METHOD</u>

Condsider the quantity

$$Q^N(i) = \frac{C_{k_1}(i)}{C_{k_1}(i_0)} \frac{C_{k_1}(i_1)}{C_{k_2}(i_1)} \frac{C_{k_1}(i_2)}{C_{k_2}(i_2)} ---- \frac{C_{k_1}(i_N)}{C_{k_2}(i_N)}$$ (7)

where

$$i_k = i + k\bar{D}(P_{k_2} - P_{k_1})$$ (8)

and
N is chosen such that

$$i_0 = i_N = i + N\bar{D}(P_{k_2} - P_{k_1})$$

$$N = \frac{i_0 - i}{\bar{D}(P_{k_2} - P_{k_1})}$$ (9)

For definiteness

Choose $i_0 = 1024$, $P_{k_2} > P_{k_1}$

1. __Proof That $Q^N(i) = RCE(i)$__: (with $k_0 = k_1$)

Using eq. (4), $Q^N(i)$ can be written as

$$Q^N(i) = \frac{C_{k_1}(i)}{C_{k_1}(i_0)} \; \frac{I_{k_1}(i_1)}{I_{k_2}(i_1)} \; \frac{I_{k_1}(i_2)}{I_{k_2}(i_2)} \; --- \; \frac{I_{k_1}(i_N)}{I_{k_2}(i_N)} \tag{10}$$

But, since

$$i_k = i + k\overline{D}(P_{k_2} - P_{k_1})$$

$$i_{k+1} = i + k\overline{D}(P_{k_2} - P_{k_1}) + \overline{D}(P_{k_2} - P_{k_1}) = i_k \cdot \overline{D}(P_{k_2} - P_{k_1})$$

Then according to (5)

$$O_{k_1}(i_k) = \frac{(i_k - i_0)}{\overline{D}} - (P_{k_1} - P_{k_0})$$

and

$$O_{k_2}(i_{k+1}) = \frac{(i_k - i_0)}{\overline{D}} + (P_{k_2} - P_{k_1}) - (P_{k_2} - P_{k_0}) = \frac{i_k - i_0}{\overline{D}} - (P_{k_1} - P_{k_0})$$

Hence,

$$O_{k_1}(i_k) = O_{k_2}(i_{k+1})$$

That is, channel i_k (position k_1) is in the same spatial position as channel i_{k+1} (position k_2).

This means that

$$I_{k_1}(i_k) = I_{k_2}(i_{k+1})$$

and eq. (10) can be written as

$$Q^N(i) = \frac{C_{k_1}(i)}{C_{k_1}(i_0)} \; \frac{1}{I_{k_2}(i_1)} \; \frac{I_{k_1}(i_N)}{1} ---$$

But channel 1, (position k_2) is in the same spatial location as channel i (position k_1); also $i_N = i_0$ by definition of N.

Hence

$$Q^N(i) = \frac{C_{k_1}(i)}{C_{k_1}(i_0)} \frac{I_{k_1}(i_0)}{I_{k_1}(i)}$$

and since $k_1 = k_0$

$$Q^N(i) = \frac{C_{k_1}(i)}{C_{k_0}(i_0)} \frac{I_{k_0}(i_0)}{I_{k_1}(i)} = RCE(i) \qquad Q.E.D.$$

2. To Get Exact Values From Experiment

To get exact values for all i, N must be an integer for all i

$$N = \frac{i_0 - i}{\bar{D}(P_{k_2} - P_{k_1})}$$

This would be true only if $\bar{D}(P_{k_2} - P_{k_1}) = 1$.

That is, the two table positions are separated by 1 channel width (~ 8.5 arcseconds).

Then $N = i_0 - i = 1024 - i$.

(Note that RCE(1024) is not needed since it is alread known - RCE(1024) = 1)

3. To Get Approximate Values From Experiment

Let [N] represent the closest integer to the value N

hen

$$N = \left[\frac{i_0 - i}{\bar{D}(P_{k_2} - P_{k_1})} \right]$$

Since $i_N \leq 1024$, there is also a maximum value for N so that

$$i_N = i + N\bar{D}(P_{k_2} - P_{k_1}) \quad 1024.$$

It is easily seen that the approximate values obtained with this method re based on the assumption that the flux is constant over k channels where

$$k = \frac{1}{2} \bar{D}(P_{k_2} - P_{k_1}).$$

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

SOLIDIFICATION OF A BINARY MIXTURE

Prepared By: Basil N. Antar, Ph.D.

Academic Rank: Associate Professor

University and Department: University of Tennessee Space Institute
                                      Department of Engineering Science and
                                            Mechanics

NASA/MSFC:
    Division: Atmospheric Science
    Branch: Fluid Mechanics

MSFC Counterpart: George H. Fichtl

Date: July 20, 1981

Contract No.:   NGT 01-008-
                                      The University of Alabama in Huntsville

ABSTRACT

The time dependent concentration and temperature profiles of a finite layer of a binary mixture is investigated during solidification. The coupled time dependent Stefan problem is solved numerically using an implicit finite differencing algorithm with the method of lines. Specifically, the temporal operator is approximated via an implicit finite difference operator resulting in a coupled set of ordinary differential equations for the spatial distribution of the temperature and concentration for each time step. Since the result ng differential equations set form a boundary value problem with matching conditions at an unknown spatial point, the method of invariant imbedding is used for its solution.

NOMENCLATURE

Symbol	Definition

Latin Symbols

C	Concentration
d	Initial depth of melt
D	Mass diffusivity
l	Latent heat of fusion
h	Partition coefficient
Q	Heat transfer rate
S	Liquid - solid interface position
t	Time
T	Temperature
T_o	Solidification temperature at zero concentration
x	Spatial direction

Greek

ρ	Density
κ	Thermal diffusivity

Subscripts

l	Liquid state
s	Solid state

INTRODUCTION

A great deal of interest has been generated recently in the possibility of producing new materails in the reduced gravity environment provided during the forthcoming missions of spacelab. The range of possibilities extend from producing large crystals of uniform properties to manufacturing materials with unique properties. Most of these processes involve the solidification of materials from the liquid state. Convective motion within the liquid during solidification can influence the local material composition and the shape of the solid-liquid interface which may result in solids with non-uniform properties and crystal defects. The microgravity environment of Spacelab is being viewed as one in which the boyancy forces are eliminated so that convection driven by thermal gradients will not occur. It is hoped that this will lead to an improved solidification process. However, convection may occur in other reasons and whether convection is negligible or not during solidification constitutes a vital question bearing on the value of future material processing in low gravity environment. Little information exists presently on convections during solidification under such circumstances.

The work reported here constitutes the first step in a complex analytical investigation concerning the onset of convective motion in the liquid due to surface tension forces during the solidification of a finite thickness slab of a binary fluid. The onset of convection will be determined through a stability analysis. Stability analyses usually require the determination of a basic state whose stability is being studied. The rest of this report will deal with the definition of such a basic state.

STATEMENT OF THE PROBLEM

The stability analysis to be performed requires the complete description of the temperature and concentration fields together with the speed and position of the solid-liquid interface during the solidification process. Although the equations describing these fields are linear (the diffusion equations) the problem is greatly complicated by the fact that the position of the solid liquid interface is unknown, making the total problem much more difficult. The complexity is increased further here by the fact that the solidification temperature is dependent on the concentration which is also a variable in this problem. Thus it can be seen that the total problem is complex and highly transient making a closed form solution for the complete problem almost impossible. Boley [1] has outlined an algorithim for obtaining a closed form solution for the problem at hand but for short times only making his technique not very useful for the present application. Thus a numerical method of solution is chosen which is outlined below.

Before getting into the numerical details of the solution, the mathematical model for the problem is first outlined. Consider a layer of binary fluid which is infinite in the horizontal extent but of a finite depth d. Initially the liquid is at a uniform temperature and concentration with the appropriate boundary conditions on the top and bottom boundaries of the fluid. A sketch of the model is shown in Fig. 1.

At any time during the solidification process the temperature and concentration fields in both the solid and the liquid are generated by the following equations

$$\frac{\partial T_s^*}{\partial x^*} = \kappa_s \frac{\partial^2 T_s^*}{\partial x^{*2}}$$

(1a)

$$\frac{\partial C_s^*}{\partial x^*} = D_s \frac{\partial^2 C_s^*}{\partial x^{*2}}$$

(1b)

$$\frac{\partial T_\ell^*}{\partial x^*} = \kappa_\ell \frac{\partial^2 T_\ell^*}{\partial x^{*2}}$$

(1c)

$$\frac{\partial C_\ell^*}{\partial x^*} = D_\ell \frac{\partial^2 C_\ell^*}{\partial x^{*2}}$$

(1d)

In the above equations the densities of both the liquid and solid were taken to be the same which is a constant ρ.

To completely describe the field variables initial and boundary conditions must to be specified as well as the solid-liquid interface conditions. For initial conditions it will be assumed that the temp. of the lower liquid surface (at $x = 0$) has just reached the solidification temperature, $T_m^*(t_0)$ for the initial uniform concentrations of the liquid.

The temperature distribution in the liquid will be linear with a maximum at the upper surface and the initial concentration in the liquid, $C*_{1_0}$, will be taken to be uniform throughout the liquid.

The boundary conditions at all time t>to are given by:

$$\frac{\partial T_s^*}{\partial x^*}(o,t) = -\frac{Q_o(t)}{k_s} \qquad (2)$$

$$\frac{\partial C_s^*}{\partial x^*}(o,t) = O$$

at the lower surface and by

$$\frac{\partial T_\ell^*}{\partial x^*}(d,t) = -\frac{Q_\ell(t)}{k_\ell} \qquad (3)$$

$$\frac{\partial C_\ell^*}{\partial x^*}(d,t) = O$$

at the upper surface. The conditions at the solid-liquid interface, s(t) are given by:

$$T_\ell^*(s^*,t^*) = T_s^*(s^*,t^*) = T_m^* = F(C_\ell^*(s^*,t^*)) \qquad (4)$$

$$C_s^*(s^*,t^*) = k\, C_\ell^*(s^*,t^*) \qquad (5)$$

$$k_s \frac{\partial \bar{T}_s^*}{\partial x^*} - k_l \frac{\partial \bar{T}_l^*}{\partial x^*} = \rho L \frac{\partial S'}{\partial t^*} \qquad (6)$$

$$D_s \frac{\partial C_s^*}{\partial x^*} - D_l \frac{\partial C_l^*}{\partial x^*} = (1-k) C_l^* \frac{\partial S^*}{\partial t^*} \qquad (7)$$

Condition (4) shows that the solidification temperature T_m^* is a function of the concentration which may be obtained from the eutectic phase diagram of the specific materials in question. However, to simplify the analysis the functional dependence will be taken to be linear and of the following form

$$\bar{T}_m^*(S^*) = T_o^* - m\, C_l^*(S', t^*) \qquad (8)$$

where m is the slope in the phase diagram and is taken to be a constant.

At this point it is desirable to nondimensionalize the variables of the problem described above. The initial depth of the liquid, d, will be used for the length scale and since the problem is mainly diffusion dominated, the diffusion time d^2/κ will be used for the time scale. The original uniform liquid concentration, C_i^*, will be used for the concentration scale and the difference o between the solidification temperature at zero concentration T_o^*, and the solidification temp. of the initial concentration $T_m^*(t_o)$ will be used for the temperature scale. Thus, the folloiwng nondimensional variable will be used in the problem.

$$x = x'/d \qquad ; \qquad s = s''/d$$

$$t = \varkappa_s t''/d^2$$

(9)

$$c = c''/c'_{lo}$$

$$T = \frac{T'' - T_m'' (t_o)}{T_o'' - T_m'(t_o)}$$

Upon substituting the above definition into the problem and the boundary condition we obtain

$$\frac{\partial T_s}{\partial t} = \frac{\partial^2 T_s}{\partial x^2}$$

(10a)

$$\frac{\partial C_s}{\partial t} = P_s \frac{\partial^2 C_s}{\partial x^2}$$

(10b)

$$\frac{\partial T_l}{\partial t} = \varkappa \frac{\partial^2 T_l}{\partial x^2}$$

(10c)

$$\frac{\partial C_l}{\partial t} = P_l \varkappa \frac{\partial^2 C_l}{\partial x^2}$$

(10d)

with the boundary conditions given by:

$$\frac{\partial T_s}{\partial x} = - q_0$$

$$\frac{\partial C_s}{\partial x} = 0 \qquad \Bigg\} \quad \text{at} \quad x = 0 \qquad (11)$$

$$\frac{\partial T_l}{\partial x} = - q_1$$

$$\frac{\partial C_l}{\partial x} = 0 \qquad \Bigg\} \quad \text{at} \quad x = 1 \qquad (12)$$

The interface conditions are given by

$$T_l(s,t) = T_s(s,t) = 1 - M C_l(s,t) \qquad (13)$$

$$C_s(s,t) = k C_l(s,t) \qquad (14)$$

$$\frac{\partial T_s}{\partial x} - c \frac{\partial T_l}{\partial x} = L \frac{\partial s}{\partial t} \qquad (15)$$

$$\frac{\partial C_s}{\partial x} - D \frac{\partial C_l}{\partial x} = (1-k) \frac{C_l}{P_s} \frac{\partial s}{\partial t} \qquad (16)$$

The nondimensionalizations process adopted above results in several nondimensional numbers which are defined as follows:

$$P_s = D_s / \kappa_s \quad ; \quad P_\ell = D_\ell / \kappa_\ell \quad ; \quad K = K_\ell / K_s \quad ;$$

$$q_o = \frac{Q_o d [T_o^* - T_m^*(t_o)]}{\kappa_s} \quad ; \quad q_\ell = \frac{Q_\ell d}{k_\ell} [T_o^* - T_m^*(t_o)] \quad ;$$

$$M = \frac{\partial [T'/(T_o^* - T_m^*(t_o)]}{\partial (C'/C_{\ell_o}^*)} \quad ; \quad L = \frac{\ell}{C_f [T_o^* - T_m^*(t_o)]} \quad ; \quad D = D_\ell / D_s$$

METHOD OF SOLUTION

In order to obtain a full description of the basic state, the problem defined by equations (10) and conditions (11)-(16) need to be solved for some realistic initial conditions. Although the differential equations governing the problem are linear, the boundary conditions through the latent heat terms introduce strong nonlinearities and complicates the solution considerably. The problem is further complicated in that the position of the solid-liquid interface is not known a priori and need to be determined in the process of solving the problem at each instant in time. There exists in the literature a closed form solution when the domain in the x-direction is infinite (see Carslaw and Jaeger[2]). Also, there is a solution to the problem when the domain is finite, but the depth of the liquid is constant for all times (Antar et al [3]). Recently, Boley [1] has outlined a closed form solution to the problem which is valid only for short times. Thus it seems that if a uniformly valid solution for all times is needed and without any of the restrictions discussed above, then the only recourse is to obtain tne solution through numerical means.

It is clear, since this problem is important in many engineering applications, that a great deal of analytical and numerical techniques have been developed for its solution (see for instance the books of Ockendon and Hodgkins [4] and Wilson, Solomon and Boys [5] for an up to date review of these techniques). Thus there exists in the literature

a large number of numerical techniques for solving this problem which broadly fall into the two catagories finite difference and finite elements approximations. For an interesting review of the numerical techniques used in this area see the article by Fox [6]. The method which is used in the present application is that due to Meyer [7].

The method of Meyer [7] may be summarized as follows: first descretise the time operator through an implicit finite difference approximation. This will then yield a set of ordinary differential equations at any time level n, which form a boundary value problem Since there are interface conditions which have to be satisfied at the location of the interface Meyer used the invariant imbedding method to solve the problem. Note that Meyer applied this technique to solve only the temperature field, while in this report we will illustrate how his technique may be extended to solve for both the temperature and concentration fields. The details of the method are given below.

Let $N>0$ be an integer and define a time step $\Delta t = T/N$. If a_i^n etc. denotes $a_i(h\Delta t, x)$, etc. for $n = 0,\ldots, N$, then an implicit one level approximation of the equations given in (10) may be written as

$$\frac{d^2 T_s^n}{dx^2} - \frac{T_s^n}{\Delta t} = - \frac{T_s^{n-1}}{\Delta t} \tag{17a}$$

$$\frac{d^2 C_s^n}{dx^2} - \frac{C_s^n}{P_s \Delta t} = - \frac{C_s^{n-1}}{P_s \Delta t} \tag{17b}$$

$$\frac{d^2 T_\ell^n}{dx^2} - \frac{T_\ell^n}{\kappa \Delta t} = - \frac{T_\ell^{n-1}}{\kappa \Delta t} \tag{17c}$$

$$\frac{d^2 C_\ell^n}{dx^2} - \frac{C_\ell^n}{P_s \kappa \Delta t} = - \frac{C_\ell^n}{P_s \kappa \Delta t} \tag{17d}$$

The boundary conditions will become

$$\frac{dT_s^{\,n}}{dx} = -\,q_{,}^{\,n} \qquad ; \qquad \frac{dC_s^{\,n}}{dx} = 0 \qquad\qquad (18)$$

$$\frac{dT_l^{\,n}}{dx} = -\,q_{,}^{\,n} \qquad ; \qquad \frac{dC_l^{\,n}}{dx} = 0 \qquad\qquad (19)$$

while the interface conditions are expressed by

$$T_l^{\,n}(s^n) = T_s^{\,n}(s^n) = 1 - M\,C_l^{\,n}(s^n) \qquad\qquad (20a)$$

$$C_s^{\,n}(s^n) = k\,C_l^{\,n}(s^n) \qquad\qquad (20b)$$

$$\frac{dT_s^{\,n}}{dx} - c\,\frac{dT_l^{\,n}}{dx} - \frac{L}{\Delta t}\left(s^n - s^{n-1}\right) = 0 \; ; \qquad\qquad (20c)$$

$$\frac{dC_s^{\,n}}{dx} - D\,\frac{dC_l^{\,n}}{dx} - (1-k)\,\frac{C_l^{\,n}}{r_s^2\,\Delta t}\left(s^n - s^{n-1}\right) = 0 \qquad\qquad (20d)$$

First we split the second order operators in (17) into two first order differential equations in the following form

$$\frac{d\overline{T_s}^n}{dx} = v_s$$ (21)

$$\frac{dv_s^n}{dx} = \frac{T_s^n}{\Delta t} - \frac{T_s^{n-1}}{\Delta t}$$

$$\frac{dC_s^n}{dx} = w_s^n$$

(22)

$$\frac{dw_s^n}{dx} = \frac{C_s^n}{\rho_s \Delta t} - \frac{C_s^{n-1}}{\rho_s \Delta t}$$

and similar operator splitting for T_1^n and C_1^n. Then, since the boundary conditions (18) and (19) are given in a derivative form (Neumann conditions), we choose the solution v^n and w^n such that the equations

$$v_s^n(x) = Y_s^n(x)/\beta_s^n + Z_s^n(x)$$ (23)

$$w_s^n(x) = X_s^n(x)Y_s^n + \zeta_s^n(x)$$ (24)

and similar solution for v_1^n and w_1^n, hold. Y_s^n, X_s^n, Z_s^n and r_s^n are solutions to the following initial value problems

$$\frac{dY_s^n}{d\lambda} = \frac{1}{\Delta t} - \left(Y_s^n\right)^2 \tag{25}$$

$$Y_s^n(0) = 0$$

$$\frac{dX_s^n}{d\lambda} = \frac{1}{P_s \Delta t} - \left(X_s^n\right)^2 \tag{26}$$

$$X_s^n(0) = 0$$

$$\frac{dZ_s^n}{d\lambda} = - X_s^n Z_s^n - \frac{1}{\Delta t} T_s^{n-1} \tag{27}$$

$$Z_s^n(0) = - q_0^n$$

$$\frac{dr_s^n}{d\lambda} = - X_s^n r_s^n - \frac{1}{P_s \Delta t} C_s^{n-1} \tag{28}$$

Similar initial value problems for T_1^n and C_1^n are produced where the initial conditions are applied at $\lambda = 1$.

Assuming s^n is known we then find T_1^n, T_s^n, C_1^n and C_{ns}^n by integrating the following initial value problems for s^n:

$$\frac{d\bar{T_s}^n}{dx} = Y_s^n \bar{T_s}^n + Z_s^n \quad ; \quad \bar{T_s}^n(s^n) = \beta_s^n \tag{29}$$

$$\frac{d\bar{T_\ell}^n}{dx} = Y_\ell^n \bar{T_\ell}^n + Z_\ell^n \quad ; \quad \bar{T_\ell}^n(s^n) = \beta_\ell^n \tag{30}$$

$$\frac{dC_s^n}{dx} = X_s^n C_s^n + r_s^n \quad ; \quad C_s^n(s^n) = \gamma_s^n \tag{31}$$

$$\frac{dC_\ell^n}{dx} = X_\ell^n C_\ell^n + r_\ell^n \quad ; \quad C_\ell^n(s^n) = \gamma_\ell^n \tag{32}$$

The complete solutions $\{T_1^n, T_s^n, C_1^n, C_s^n, S^n, C_1(s^n)\}$ for $n = 1$, ..., N is advanced from time level to time level in the following manner. Assume $\{T_s^{n-1}, T_1^{n-1}, C_s^{n-1}, C_1^{n-1}, S^{n-1}, C_1(s^{n-1})\}$ are given, we then compute v_s^n, v_1^n, w_s^n and w_1^n according to (23) and (24). Substitution of v_s^n, v_1^n, w_s^n and w_1^n into the interface conditions (20) resulting

$$\varphi_1^n = v_s^n - c\, v_\ell^n - \frac{L}{\Delta t}(s^n - s^{n-1}) = 0$$

$$\varphi_2^n = w_s^n - \lambda w_\ell^n - (1-k)\frac{C_\ell^n}{\rho_s \Delta t}(s^n - s^{n-1}) = 0$$

If there exist two roots s^n and C_1^n for $\varphi_1^n = 0$ and $\varphi_2^n = 0$, then s^n is the desired interface position and C_1^n (s^n) is the desired interface concentration and T_s^n, T_1^n, C_s^n and C_1^n may be computed as described above.

To implement the algorithm outlined above, a numerical code was written and verified with the following specifications: All of the initial value problems were integrated using a 7-(8) Runge-Kutta Fehlberg [8] integration. Due to the massive bookkeeping problems only equal steps were used for the integrations. All of the interpolations were done using linear interpolation, although the program was written with a cubic spline interpolation capabilities. The roots for s^n and C_1^n (s^n) were found through a Newton-Raphson interative technique.
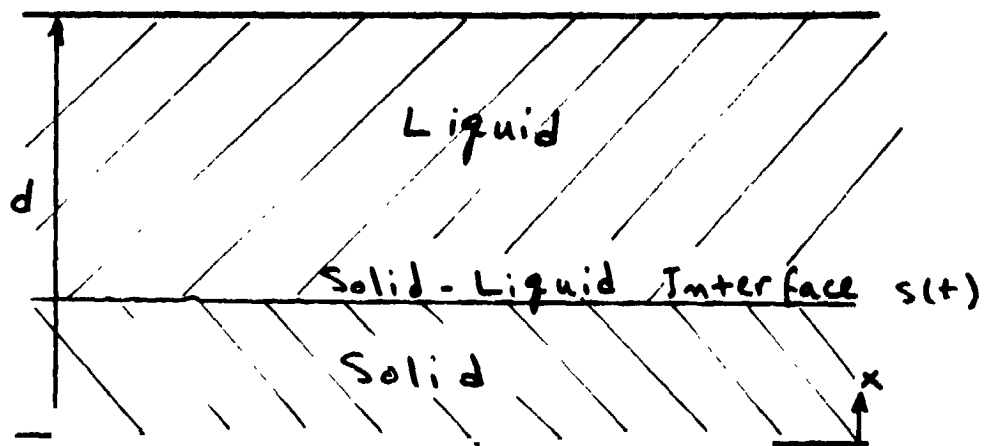
FIGURE 1

A sketch of the model used.

REFERENCES

1. Boley, B., "Time Dependent Solidification of Binary Mixtures," Int. J. Heat Mass Transfer, Vol. 21, 1978.

2. Carslaw, H. S. and Jaeger, J. C. Conduction of Heat in Solids, Oxford, (Larendon Press, 1959).

3. Antar, B. N.; Collins, F. G. and Fichtl, G. H., "Influence of Solidification on Surface Tension Driven Convection", Heat Mass Transfer. Vol. 23, 1980.

4. Ockendon, J. R. and Hodgkins, W. R., Moving Boundary Problems in Heat Flow and Diffusion. Oxford, Clarendon Press, 1975.

5. Wilson, D. G.; Solomon, A. D. and Bogs, P. T., Moving Boundary Problems. Academic Press, New York, 1978/

6. Fox, L., "What are the Best Numerical Methods, in Moving Boundary Problems in Heat Flow and Diffusion, J. R. Ockendon and W. R. Hodgkins edt., Oxford, Clarendon Press, 1975.

7. Meyer, G. H., Initial Value Methods for Boundary Value Problems. Academic Press, New York, 1973.

8. Fehlbery, K. "Classical Fifth-Sixth-Seventh and Eighth Order Runge-Kutta Formulas with Step Size Control", NASA TR R-287, Washington, D.C., 1968.

1981

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

ASSESSMENT OF MSFC'S SUPERVISORY TRAINING PROGRAMS
AND COURSES

Prepared By: Thomas A. Brindley, Ph.D.

Academic Rank: Associate Professor

University and Department: The University of Alabama
                                in Huntsville
                                Department of Education

NASA/MSFC:
     Office   Personnel Office
     Division   Personnel Development

MSFC Counterpart: C. D. Bean

Date: August 3, 1981

Contract No.:   NGT-01-008-021
                                The University of Alabama
                                in Huntsville

ASSESSMENT OF MSFC'S SUPERVISORY TRAINING PROGRAMS AND COURSES

BY

Thomas A. Brindley, Ph.D.
Associate Professor of Education
The University of Alabama in Huntsville
Huntsville, Alabama

ABSTRACT

This task is to review all the significant courses and special programs for the training of supervisors at MSFC to determine the adequacy of the present MSFC Supervisory Training Program and to recommend changes, if appropriate. This entails a review and analysis of the content, procedures, and student evaluations of the Basic Required Office of Personnel Management (OPM) 80 hours training for supervisors, the Optional 120 hours, the MSFC Management Development Program (MDP), NASA's Management Education Program (MEP), various OPM and special contract programs, pertinent procedural guidelines, regulations, and letters, as well as various MSFC computer reports which indicate who took what training. Various interviews with MSFC personnel involved in training supplement this review.

The research indicates the preferred types of training believed necessary to benefit the supervisor as student, as professional, and as manager, and to benefit the NASA/MSFC. Conclusions of this study are that changes should be made in the existing programs and courses offered for supervisors and that a more extensive concept of development be implemented. Recommendations are: (1) that the choice of courses selected for the Basic Required OPM 80 hours be improved; (2) that the optional 120 hours be discontinued and a shorter module be developed dealing with managerial decision making and human relations skills; (3) that the MDP and MEP be continued as at present; and (4) that a broad array of developmental strategies be incorporated to provide a variety of opportunities for supervisory improvement.

ACKNOWLEDGEMENTS

INTRODUCTION

The task assigned to me this summer was to conduct an in-depth study and assessment of the training programs and courses for supervisors at the Marshall Space Flight Center (MSFC). The purpose was to review all the significant courses and special programs designed for the training of supervisors to determine the adequacy of the present MSFC Supervisory Training Program and to recommend changes, if appropriate. The findings and recommendations of this study should indicate the preferred type of training and development believed necessary to benefit the supervisor as student, as professional, and as a manager, and to benefit the NASA-MSFC organization and mission.

Materials reviewed include the curricular content and procedures of courses offered at MSFC, student evaluations of courses and programs, training plans, course catalogs, in-house reports, directives, and computer printouts of who took what training. Specifically, an intensive review was made of the NASA Management Education Plan (MEP), the MSFC Management Development Program (MDP), the Basic OPM required 80 hours training for supervisors, the Optional 120 hours additional training, and special contract programs, such as Stirling Institute and Harbridge House. Various interviews with MSFC personnel involved in training supplement this review.

BACKGROUND

A Center directive was sent out by the Associate Director on February 17, 977, which (1) called for the identification of specific training needs of Center supervisors, and (2) structured a training plan for each supervis. This directive stated that all supervisors complete the 80 hours required by the Office of Personnel Management (OPM). It suggested that supervisors receive another 120 hours on an optional basis to be determined by the supervisor to receive instruction and his/her own supervisor, "at the discretion of the directorate head or program manager."

The content of the mandatory Civil Service Commission (now OPM) 80 hours was arranged in two "modules," dealing with required and suggested topics which are stated in the Federal Personnel Manual (410-A-1; Appendix A, Supervisory Training; September 6, 1974). The MSFC Training Center utilized basic courses which the Civil Service Commission (OPM) had designed specifically to meet the requirements of the general topics. Module One dealt with a general supervisory orientation, 16 hours, which described the overall NASA and MSFC missions, which explained the various support services and personnel practices common to all supervisors (such as performance appraisal and pay management), and which indicated the necessary safety and health requirements. Invited representatives from various support services and offices spoke to the group about their respective areas. Module Two dealt with General Management theory and practice, divided into two main categories: "Basic Supervision and

Management," 40 hours, including concep*s of organization, communication processes, human relations, and learning-motivation; and, "Problem Solving," 24 hours, including analysis of problems and decision i aking.

The 120 optional hours were arranged into four modules. The first, Module Three, called Resource Management, 24 hours, dealt with costs, manpower utilization, paperwork handling, and computer support. The next, Module Four, 40 hours, dealt with procurement, contracts, legal and patent matters, and technology utilizati.n in space sciences and engineering. Module Five, titled "Building Relationships," 40 hours, grouped the psychological concerns of human relations with some rather programmatic aspects of labor relations, grievance, and public affairs. Module Six, called "Plant Management," 16 hours, dealt with facilities, equipment, civil emergencies, and security. These optional courses -- Modules Three through Six -- were, for the most part, "in-house" MSFC courses, which were taught by local MSFC personnel. The idea was to use in-house expertise, except in certain instances, such as the use of Harbridge House or other contractors in special management and human relations courses.

A cursory glance at the content of these optional modules indicates that they deal with the nuts and bolts, the programs and procedures, of the ongoing daily affairs of the Marshall Center and with expected NASA operations. Even beginning supervisors who have been with NASA on-the-line for some time have a fairly thorough acquaintanceship with the various technologies utilized at NASA, with various facilities, with security, and with the public image of NASA. They may not know the details on how to handle personnel procedures and the decision-making involved in procurement, in contracts, and in the management of property and vehicles, but they certainly have been involved in these procedures and decisions as a regular employee. Thus, the obvious question emerges whether an old-line supervisor, or even a beginning one, needs to take courses which describe many procedures and programs that are common to Center operations. Are such information-giving courses necessary? Or do they supply information that is already known? Only the human relations aspects deal with the basic attitudinal problem of the engineer-scientist turned supervisor, and that is budgeted only 16 hours of the total 120 hours option.

The February 1977 directive provides a means to select courses to be taken and a means to record their completion. When an individual is appointed as a supervisor, the Personnel Office sends out a "Supervisory Training Plan" for completion by the individual's superior. The superior and the subordinate supervisor sit down together to discuss possible areas of need. Choices can be obtained by referring to the MSFC Training Course Catalog. Courses to be taken are suggested. These plans are submitted and categorized by the Personnel Development Division. Thus, a record of the training needs of each supervisor is developed, and a compilation of tne courses projected for the coming year is drawn up. The courses deemed appropria.e within budgetary and policy considerations are commited and scheduled. Supervisors are encouraged to participate.

Most supervisors understand the training needs of their subordinates in accord with the needs of their own organization. The Personnel Development

Division provides information beyond the catalog descriptions. What sometimes happens is that the higher level supervisors open the door to the supervisors to pick and choose the courses they would like to take without providing direction or consultation whether the courses selected were in the best interests of the Center, of importance to the job, and really helpful to the person requesting the course. Sometimes, it has been reported, counselors and Employee Development Specialists (EDS) get calls such as, "I've got to come up with some training, can you recommend some courses to take?" Most often these requests are for needed information, but occasionally it appears the supervisor is merely fishing to fill in a form.

Since the 200-hour MSFC Supervisory Training Program was started in 1977, 229 courses have been conducted through May 31, 1981, with 3313 instances of training. (An instance ' one person enrolled in a class. Thus, a class of 20 equals 20 instances, whereas one person who attends three different classes equals three instances of training.) There have been 181 supervisors who have received the 24-hour Basic Supervision Course, and 304 have taken the 40-hour Problem Solving Course. The remaining participation was in various courses which make up the optional portion of the program

There have been problems with scheduling courses, with attendance, and with compliance to the 1977 request for training. At first, certain directors of the laboratories supported the request by issuing firm orders to comply But as some complaints were heard from the labs, the directors began to back off and in effect allowed the lab supervisors to take a lackadaisical approach regarding attendance. Apparently there were some real complaints that demanded correction. First, some of the computer printouts, which were to identify background, previous training, and course needs, were in error. Courses were listed as required in some cases that were not necessary. A few old line supervisors were asked to attend Basic Supervision, even though OPM guidelines clearly state that if a supervisor has not taken the basic required 80 hours after having been a supervisor for two or more years, he is not required to do so. Some old line supervisors who had been managing for 10 to 20 years felt that the courses did not apply to them, that they were redundant, unnecessary, a waste of time, and even an insult to their expertise. Some new supervisors became immediately so deeply involved in their projects that they felt they could not break away from their work. In the two year period of CY-78 and CY-79, 67 employees became new supervisors. Due to various scheduling problems, 17 of these had not yet received the Supervisory Orientation, 16 had not received the Basic Supervision Course, and 20 had not received the course in Problem Solving, as of February 19, 1981.

Some remarks from various attendees of courses might indicate the nature of the content and how they were taught Some comments were: "too much overlap of one course in supervision with another," "too much repetition," "boring, elementary material," "three days is too long for what could be taught in a half day," "materials irrelevant to my job," "what I really need isn't offered," "a lousy instructor," "knows his field but didn't get it across," and "some of the people attending were just getting out of work."

Apart from the occasional displeasure expressed, the evaluations show that most all courses received high praise, were deemed important to the recipient, and were well worth the money that NASA spent to sponsor them. Consistently, the evaluations show 90% support or better, except in a few courses. The Training Branch employees have a keen understanding of which courses and which instructors are beneficial and worthwhile. They have been very careful to select the best courses and request the best instructors.

A basic problem in the training of professionals who later become supervisors and managers resides at the outset in the selection of supervisors. There is a fundamental difference in outlook of the usual engineer-scientist-turned manager. Oftentimes an engineer or scientist is oriented, probably in his/her initial education, probably by the nature of his profession, and perhaps by the very manner in which the person thinks, toward the resolution of problems dealing with abstractions or of things or of systems. Yet, being a manager or a supervisor requires a person skilled in the understanding of people and able to deal with them. When an engineer or scientist becomes a manager or supervisor there are sometimes inherent difficulties in adjustment.

Also, there is a basic question whether professional artisans can indeed be remade or educated into successful and respected managers and supervisors. Yet, the overall answer to this is that management as an art and science can be offered to managers in such a way that the intelligent student will grasp the meaning of the content and will be able to transpose this understanding into better practical supervisory approaches. There is also an underlying feeling, pervasive among modern employees, that the supervisor must learn to deal well and fairly with employees or the entire organization will suffer. The MSFC Center view supports the concept that training and development are vitally important to the organization. This view was expressed casually by the Center Director that supervisors, in crossing the great divide from being a technician and highly skilled professional to becoming a manager, must learn to succeed in their new roles, otherwise they will not be effective. Training and development can facilitate that crossing over the great divide.

Special supervisory training programs have been devised for selected managers. These high impact programs reach about sixty supervisors a year, only a fraction of the approximate 500 supervisors at MSFC. The nagging question is whether the Center can accept a policy that provides high level professional management training for about ten percent of the number of supervisors per year and that maintains an optional, open-ended training beyond the Basic OPM 80 hours.

C - 6

The Management Education Program (MEP) is designed for mid-level managers from field centers and Headquarters located in the training facilities at Wallops, Virginia. It is a two-week, live-in program composed of interaction sessions with specialists in administrative science, and with top NASA managers. The content focuses upon the examination and improvement of managerial skills and practices, an examination of the roles of a NASA manager in a changing organizational and national environment, and an understanding of NASA as a system. MSFC sends three participants a term, or twelve a year.

Evaluations and participant responses indicate that the MEP program is extraordinarily beneficial and well received. Respondents indicate that they gained in learning to improve their managerial skills, that they learned more about themselves and their role as a manager, that they acquired more understanding of the missions, policies, functions, and procedures of NASA as a whole and their own part in NASA, and that they were stimulated by new ideas. They were particularly pleased with the opportunity to meet top NASA leaders and to discuss major problems and philosophies with them. Nearly all participants welcomed the opportunity to examine and improve their own leadership styles and the impact they have in dealing with others. In conclusion, the MEP serves as a model for development courses and programs, as it is presently structured.

It was upon this Agency model that the MSFC Personnel Office designed a Center-wide program, called the Management Development Program (MDP). This is a one-week, live-in program for branch level supervisors and managers which is held away from the Center at a selected location and consists of twenty-four selected supervisors. The program is built around (1) the selected topics presented by three management consultants (contracted by the Alabama A&M University in cooperation with the Executive Development Program of Pennsylvania State University); and, (2) presentations and discussions of top MSFC management officials, including the Director and associate directors. There have been two MDP week long sessions, one held at Wheeler State Park and the other at Guntersville State Park.

Participants unanimously were extremely pleased with both workshop conferences and said that they were well worth their time away from their regular assignments. They reported that the individual presenters from the MSFC Directorate were well prepared and willing to share their own ideas with them. The participants welcomed the informality of the workshops and the cooperative openness and frankness of the top level managers. They reported that the three academicians presented a variety of relevant, practical, and informative materials which delved into aspects of self-inquiry, motivation, human relations, interaction techniques, and participative management. In summary, five main objectives were accomplished. One, the participants learned the views and philosophies of top level MSFC management. Two, they began to see their own individual styles, actions, and values in relation with others and with the NASA organization. Three, they learned ways which might improve their quality of leadership in self-perspective. Four, they learned ways in which the individual supervisor might better deal with subordinates in helpful,

supportive, and positive ways. And five, they were inspired to build better team spirit, morale, and pride in MSFC-NASA. The successes of both MDP conferences are similar to the success of the MEP programs at Wallops.

Apparently a new direction in the training and development of supervisors is being considered. The new approach seems to be a general, holistic, multi-faceted one that, while suggesting a well grounded initial preparation, provides a continuing developmental scheme. The growth of an individual's performance, efficiency, and productivity is perceived dynamically as an on-going affair, changing with situations, and priorities, and influenced by the needs of the organization. This dynamic approach is discussed in a recently published OPM letter (FPM Letter 411-1, dated June 4, 1981) titled Supervisory Development. The title alone indicates that the betterment of supervisors is conceived to be more extensive and inclusive than just formal training, that is, beyond the classroom. In fact, the letter indicates that the Basic OPM 80 hours will no longer be required.

Development is conceived to be based upon proper planning to meet individual, group, and organizational needs. The interaction and sharing between the supervisor and his/her immediate manager leads to a needs assessment based upon previous information about the supervisory position and about the selected individual, upon the competencies needed to perform the supervisory functions, and selected problems arising in day-to-day operations out of which selected needs are identified. The outcome of the developmental approach is that each supervisor's further growth can be directed and planned on an individual basis.

The assessed needs of a particular supervisor are met by a variety of strategies. Along with formal training through various courses, modules, workshops, conferences, and programs, there are other important methods. One is the mentor relationship whereby a new supervisor receives guidance and advice from a more experienced agency management team member who is not the supervisor's immediate supervisor. Another strategy is continued coaching in the development of skills by immediate supervisors. Rotational assignments, another stratagem, broaden the perspective and provide a wider knowledge base while indicating the interrelationships and function of other units. Still another strategy is to assign the supervisor a special project which provides opportunities to broaden skills and knowledge apart from normal duties.

FINDINGS

1. At the beginning of the supervisory training spectrum, 80 hours are required by OPM. These Basic 80 hours have been, for all practical purposes, completed by nearly all new and long-term supervisors at MSFC. The courses are standard, obligatory, and well taught. Student evaluations show they are worth NASA's money and are appreciated by the students. However, the new FPM Letter 411-1 does not state a requirement of 80 hours coursework for beginning supervisors.

2. The MDP, first conducted in November, 1980, and again in May, 1981, has been well received. It is an excellent program for selected mid-management supervisors. The cost of each session is about $17,000, well worth spending for the Center. Each session reaches 24 people.

3. NASA Headquarters is considering contracting with Stirling Institute for a special management training program which can be utilized by each center and can train large numbers at a time. It contains various modules which can be arranged according to need. Parts of it can be taught by a center's own personnel who would be trained specifically.

4. The Optional 120 hours are taught primarily by MSFC Center personnel and consist mainly of modules dealing with Center activities and Center support services (such as Procurement). They are not mandatory for supervisors. Attendance is optional and determined on a case-by-case basis. When taken in addition to the Basic OPM 80 hours, they provide a total training program of 200 hours.

    a. There are mixed feelings and reactions about the continuing need of these 120 Optional hours.

    b. A number of people whom I have talked with in this research has suggested that changes should be made in this area.

    c. In some areas, there has been less than enthusiastic compliance by supervisors and next level supervisors in the completion of these optional hours.

    d. The curriculum of these 120 hours is directed to inform supervisors about the details of activities and services at the Center, but the reaction by some supervisors, according to verbal reports and student evaluations, has been that the courses are repetitive, boring, unimportant, and not necessary. On the other hand, some comments and the official student evaluations reflect general high level course performance and praise.

    e. There are problems of non-compliance by supervisors in fulfilling the guidelines of the training program: a problem of some supervisors cancelling course registration at the last moment and failing to attend; a problem of getting an organization to meet its quota or a commitment for a particular class;

problems of partial attendance at class meetings; and, problems of determining which courses might or might not be appropriate. All these problems are fully known and understood by the educational specialists in the Training Branch and by the counselors in the Development Branch of the Personnel Office and by Center management.

f. Some dissatisfaction with the Supervisory Training Program in general, but particularly with the Optional 120 hours, seems to have arisen because of a few errors in an occasional record of the assignment of courses and in recording past education, training, and experience. Some of these errors in recording early training experiences and in assigning courses to be taken by a supervisor who already took a comparable course may have had to do with the changeover from manual to a computer system some years ago.

g. Ascertaining what an ideal training program ought to be is sometimes difficult for an organization as diverse and talented as NASA/MSFC. The state-of-the-art of management sciences, which changes with the demands of dynamic organizations and with the times, is difficult enough even for the academicians in the management profession. Certain questions arise:

1. Beyond the basics, should a wide variety of course choices be permitted, or should one learn more in depth the scientific and engineering specialities of one's own particular lab or basic work?

2. What kinds of training are absolutely essential for managers and supervisors at MSFC? What kinds are peripheral? Do they need work and new ideas in management/supervisory principles and practices, in problem solving, in human relations, in technical skills? Do they need interaction sessions?

3. Do top managers at MSFC have time to sit down informally with these supervisors to share experiences and to restate purposes as a part of their training?

h. The overall spectrum or continuum of supervisor training has room for alternative choices and can be revised without disrupting other Center functions.

RECOMMENDATIONS

My interpretations of the findings lead to the following recommendations:

1. Although the Basic OPM 80 hours are no longer required by OPM, according to FPM Letter 411-1, it is advised that the Center retain this requirement. The most streamlined, most efficient, most diverse, and most dynamic courses must be chosen for these Basic OPM 80 hours.

    a. The present Module 1 should be kept intact, because it introduces the total MSFC Center concept to the new supervisor.

    b. The general purpose of Module 2 (to cover General Management) should be kept, but the most beneficial and meaningful courses should be selected and they should be sequential and developmental in format. That is, the beginning managerial course should lead into the next, and follow upward thereon -- going from A to B to C to D, not A to Z to M to B.

2. The MSFC Management Development Program (MDP) must be continued and offered at least twice a year. It is nearly perfect. (Guntersville State Park is the preferred place.)

3. The NASA Management Education Program (MEP) should be utilized continuously as it has been in the past. It too is nearly perfect.

4. However, the Optional 120 hours of the 200 hour supervisor training program should be reworked entirely and changed to include the following points.

    a. The 120 hours should not be officially required.

    b. Supervisors and their superiors should continue to discuss together the training needs of each person. Course and program needs surveys should be utilized following similar procedures as presently used.

    c. A mid-level option or module should be created between the OPM Basic 80 hours and the MDP level.

        1. This option should imitate the content and format of the MDP or MEP.

        2. Because of cost, it should be a forty-hour "day-in" at the MSFC Training Center or at UAH or A&M, or at a hotel such as Huntsville Hilton. (A five day conference at the Hilton might cost minimally $625, while preferred arrangements might cost $2000).

        3. Various existing package programs might be the most suitable for this mid-level. Among these are:

a. Stirling Institute.

b. Middle Management Institute of OPM.

c. Alabama A&M Executive Development Program.

d. A MSFC-NASA built program consisting of selected courses hired in, using Training Branch facilities.

5. I have no specific recommendation for supervisory training, beyond the MEP, except for several considerations.

a. A type of training most beneficial to senior-level management would be block time discussion periods with MSFC directors -- "rap sessions" dealing with general wide-ranging NASA problems and national policy. Selected guest speakers, at the level of international and national concern, could be invited to serve as a catalyst for discussion.

b. Other senior level training would have to be coordinated, it seems to me, out of NASA Headquarters.

6. Evaluations and reporting.

a. A more subjective and extensive course evaluation form should be used for all MSFC courses, whether sponsored by OPM or not. They should be anonymous -- no name of student required or requested. They should adapt the forms used in the MDP program as appropriate. They should allow enough time at the end of each session to write a lengthy explanatory critique of instructor and content. These should be reviewed and tallied thoroughly by the instructor and by a training specialist and kept for several years on file. They would be used to document the adequacy of the course content, the receptivity of the students, and whether changes are needed, etc.

b. The need for training should include consideration of a supervisor's previous education and experience.

c. The records should clearly indicate courses at all levels so that there is never any omission or request to take something that has already been completed.

7. Subjective considerations in the building of courses are very important to the further development of supervisors. Among these considerations are:

a. Provide opportunities for supervisors to meet informally with top NASA directors and top MSFC management -- "rap sessions" -- as is being achieved currently in MDP and MEP.

b. Get "think tanks" going where supervisors of different areas and offices meet to share ideas. This will provide a "synergistic effect." (The outcome will be greater than the sum of its parts.) Base these groups on the

cadre, core, cell idea of six to ten supervisors who meet together infrequently or regularly.

        c. Build "networks" across disciplines and areas.

          1. This should build team spirit and morale -- an "esprit de corps."

          2. This provides a non-threatening environment or a sounding board for a manager who might discover problems are similarly shared by others in other areas and who has a chance to discuss matters away from direct responsibility.

          3. This exposes participants to ways and means of dealing effectively with mutual problems and to find ways to build better interpersonal relationships.

    8. Following the suggested guidelines of FPM Letter 411-1, the Personnel Development Division of the MSFC Personnel Office might devise and implement a comprehensive supervisory development plan. This plan might include a variety of developmental strategies (as discussed in this Final Report, page 6).

    9. A Center-wide Training Advisory Committee should be formed, meet quarterly, act in an advisory manner, and make recommendations about the entire Supervisory Training Program -- its purposes, its intent, its method of compliance, its curricula, and its relationships to MSFC Center and NASA activities. Membership on the committee might include directorate heads, program/project heads, and other key Center officials.

    Because supervisors are deeply involved in their own projects and organization, training is often perceived as a luxury. But it can become a necessity if it is truly educative and helpful to the person as an individual, to the person as supervisor/manager, and to the MSFC. Directors and supervisors and all employees should be involved in the development and constant renewal of curricula and instruction for supervisors of all levels.

REFERENCES

1. NASA Headquarters, Management Education Program-IX, Course Curriculum and Workbook, Wallops Island, Virginia, November 27 - December 8, 1978, NPD-31.

2. NASA Headquarters, Management Education Program: Follow-Up Study, Review of MEP Classes 8-11 (September 1978 - June 1979), NPD-31.

3. NASA-Marshall Space Flight Center, Management Development Program-I, Course Curriculum, Wheeler State Park, November 16-21, 1980.

4. NASA-Marshall Space Flight Center, Management Development Program I, Wheeler State Park, November 16-21, 1980, Final Review: Course Evaluation Summary, February 26, 1981, and course evaluation raw data.

5. NASA-Marshall Space Flight Center, Management Development Program-II, Course Curriculum, Guntersville State Park, May 17-22, 1981.

6. NASA-Marshall Space Flight Center, Management Development Program-II, Guntersville State Park, May 17-22, 1981, Final Review and Analysis Report, August 1981, and course evaluation raw data.

7. NASA-Marshall Space Flight Center, Honorary Awards/Personnel Development Programs/Full-Time Study Handbook, March 1980, pp. 211-265.

8. NASA-Marshall Space Flight Center, MSFC Training Course Catalog, USGPO Atlanta Region IV, January 1980.

9. NASA-Marshall Space Flight Center, Annual Training Plan FY-80 (October 1, 1980 - September 30, 1981), CM21.

10. NASA-Marshall Space Flight Center, Training Needs Survey: Courses FY-80, and Training Needs Survey: Courses FY-81, CM21.

11. NASA-Marshall Space Flight Center, Supervisory Training Program Status Report, Report No. 41, 1981.

12. NASA-Marshall Space Flight Center, Subject: Supervisory Training, Letter from John S. Potate/DM01, February 17, 1977.

13. NASA-Marshall Space Flight Center, MSFC Supervisory Training, Curricula and Modules, used to develop Supervisory Training Plan, February 17, 1977, CM21.

14. Office of Personnel Management, Interagency Training: 1979-1981 Catalog of Courses, Washinton, D.C., Pamphlet T-9, USGPO, February 1979.

15. Office of Personnel Management, Federal Personnel Manual, Appendix A. Supervisory Training, 410-A, Inst. 212, September 6, 1974, pp. 1-7.

16. Office of Personnel Management, Federal Personnel Manual System, Supervisory Development, FPM Letter 411-1, June 4, 1981, pp.1-12.

D26

1981
NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

DIGITAL COMPENSATION TECHNIQUES FOR THE EFFECTS OF
TIME LAG IN CLOSED-LOOP SIMULATION USING THE
6 DOF MOTION SYSTEM

Prepared By: Richard Brown, Ph.D.

Academic Rank: Associate Professor

University and Department   University of Alabama
                                University, AL
                                Department of Mathematics

NASA/MSFC: Data Systems Laboratory
                                Engineering Computers Division
                                Simulation Systems Branch

MSFC Counterpart: Frank L. Vinz

Date: August 25, 1981

Contract No: NGT-01-008-021
                                (University of Alabama in Huntsville)

DIGITAL COMPENSATION TECHNIQUES FOR THE EFFECTS OF THE LAG IN CLOSED LOOP SIMULATION USING THE 6 DOF MOTION SYSTEM

BY

Richard C. Brown
Associate Professor of Matehmatics
University of Alabama
Tuscaloosa, Alabama

ABSTRACT

In this study we continue efforts begun in [1] to develop digital filter compensation schemes for the correction of momentum gains observed in the closed loop simulation of the docking of two satellites using the 6 DOF motion system.

Several filters will be shown to work well for small delays (< .100ms) and a non-preloaded probe.

ACKNOWLEDGEMENT

I.   INTRODUCTION AND OBJECTIVES

In a previous report [1] we have investigated the employment of the Link Six-Degree-of-Freedom (6 DOF) motion system in contact dynamics simulation and proposed some compensation schemes to lessen instability in the system caused by time delay. The present paper is a continuation of this effort. Several compensation techniques (old and new) will be presented and intensively studied. First, however, we review the operation of the 6 DOF system and restate the underlying problem whose correction is the aim of this paper.

A docking probe assembly is placed on the platform of the Link 6 DOF Motion System. This platform having 6 degrees of freedom can perform any general motion (involving translations, pitches, yaws, or rotations) within certain limits of travel. Above the probe is a fixed target. The system software simulates initial conditions such as velocities, accelerations and rotations prior to docking, as well as effects due to orbital mechanics. When the probe is in contact with the target sensors placed in the target record the forces and torques encountered by the probe. In the closed-loop simulation mode, these forces and torques are used in a feed-back loop to control the motion of the probe itself. Essentially two integrations are performed to give velocity and position functions of the probe relative to the target. The position function is converted into commands to change the lengths of the six legs supporting the platform. These commands are then executed by hydraulic actuators in the legs causing the probe to move. In theory, therefore, the probe should behave as if it were freely responding to the forces generated by contact with the target. Unfortunately, the actual performance of the system does not quite conform to this model. The presence of a time lag between the production of a force signal and the execution of the leg length corresponding to it (currently about 200ms), causes an unstable response. To explain how this can occur, we introduce a simplified block diagram (fig. 1) appropriate to simple vertical translation:
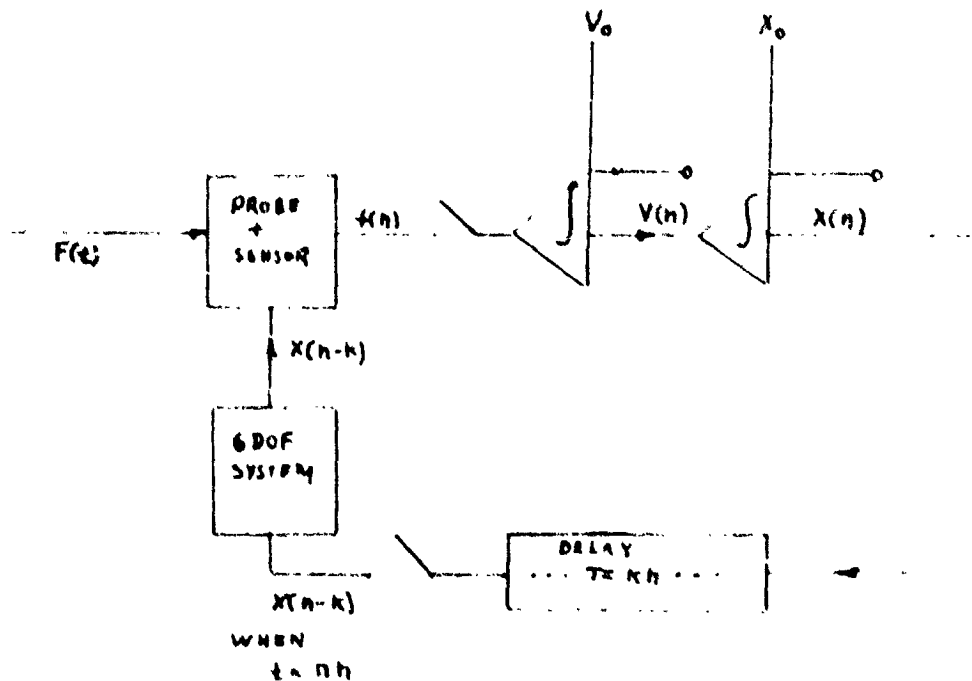
fig 1

In the above diagram h stands for the sampling interval. It is assumed that the delay is some integral multiple k of h . f(t) is some time dependent force, perhaps caused by thrusters. We adopt the abbreviations $x(n) := x(nh)$ $v(n): = v(nh)$, etc. For simplicity the delay is lumped in the feed back loop although each element in the system contributes to some part of it.

It is easy to see from fig. 1 how delay can cause an unstable response. Suppose at time $t=nh$ the probe has just reached its maximum compression. The compression produces a force signal, which is integrated twice becoming a displacement command $x(n)$. In the ideal situation $x(n)$ would be executed instantaneously. However, because of the delay it is not executed until kh seconds later (at $t= (n+k)h$). In this interval the probe executes the commands $x(n+1-k)$, $x(n+2-k)$, ---, $x(n+k-1-k)$. This, of course, causes compression beyond the theoretical maximum and in turn the communication of additional force to the system. The result is an apparent gain of momentum and energy. If the system did not cut off after the probe reaches its rest position, the process would be repeated causing a sequence of oscillations of increasing amplitude.

We can express the general situation in a mathematical way as follows:

The 6 DOF system is a device to solve the initial value problem

$$\ddot{X} = F(X, \dot{X}, t)$$

(1)
$$\begin{pmatrix} X(0) \\ \dot{X}(0) \end{pmatrix} = \begin{pmatrix} X_o \\ \dot{X}_o \end{pmatrix} \, .$$

Here X is a six dimensional state vector (one component for each degree of freedom) and F is a vector valued function of t, X and $\dot{X}$. Because of the time delay τ the system actually solves the retarded initial value problem.

(2)
$$\ddot{X} = F(X(t-\tau), \dot{X}(t-\tau), t)$$

$$\begin{pmatrix} X(t) \\ \dot{X}(t) \end{pmatrix} = \begin{pmatrix} X_o(t) \\ \dot{X}_o(t) \end{pmatrix} \, , \qquad 0 \le t \le \tau \, .$$

(Note that in this case the initial conditions have to be prescribed on an interval rather than at a point.)

It is easy to find examples where the solutions of (2) are unstable. If, for example, we restrict our attention to vertical translation and $F(X, \dot{X}, t) = \omega_o^2 x$ so that the probe is an ideal spring, then it is known the solutions are oscillatory (at first with a frequency close to ω_o/τ but of exponentially increasing amplitude.

If F is known and the initial conditions are not too complicated, it is not difficult to develop software simulating dynamic interaction of the probe with the target. This option bypasses the delay problem of closed-loop simulation but at the price of an over-simplification of the model. In actuality, F may be either unknown or too complicated to model accurately. (In some systems for docking mechanism can produce implulsive forces and torques at unpredictable times.)

Also, the initial conditions can be variable and complicated. Hence, from an engineering standpoint, accurate closed-loop simulation is desirable.

Since the delay can be reduced but not eliminated our task requires the synthesis of a control which when applied to (2) gives something nearly approximating the solution of (1).

This problem falls outside of classical control theory. If F is unknown, there is no apparent cost function to minimize or concrete performance criterion to meet. In fact, a little thought shows that if F is completely unknown then the problem is unsolvable. (More will be said on this issue later). We follow here a heuristic approach of extrapolation. At time t, $X(s)$ and $\dot{X}(s)$ are available $0 \leq s \leq t$. This suggests using this information to produce predicted vectors $X^*(t)$ $\dot{X}^*(t)$. If $X^*(t)$ and $\dot{X}^*(t)$ are sufficiently close to $X(t)$ and $\dot{X}(t)$, the solution of

(3)
$$\ddot{X} = F(X^*(t), \dot{X}^*(t), t)$$
$$\begin{pmatrix} X^*(0) \\ \dot{X}^*(0) \end{pmatrix} = \begin{pmatrix} X_0 \\ X_t \end{pmatrix}$$

should be close to that of (1).

We end this section with an outline of the remainder of the paper. Section 2 consists of a collection and analysis of several predictive algorithms. Section 3 and 4 consist of data and graphs relating to the performance of the algorithms. Section 5 contains the conclusions of the paper and discussion of some other approaches.

II. METHODS OF COMPENSATION BY EXTRAPOLATION

In this section we will consider explictly only the simplified situation of fig. 1, although our methods can be applied to the state vector case.

Since the 6 DOF system is a sampled data system, we begin with some comments and notation appropriate to a discrete setting. At time t=nh x(1),...x(n), v(1),... v(n), f(1),...f(n) are available. From this information, we will construct estimates x*(n), v*(n), f*(n) which approximates x(n+k), v(n+k) or f(n+k). These predictions will then replace x(n), v(n) or f(n). There is considerable freedom of choice since we can compensate position, velocity, or force.

Some possibilities are illustrated in fig. 2.
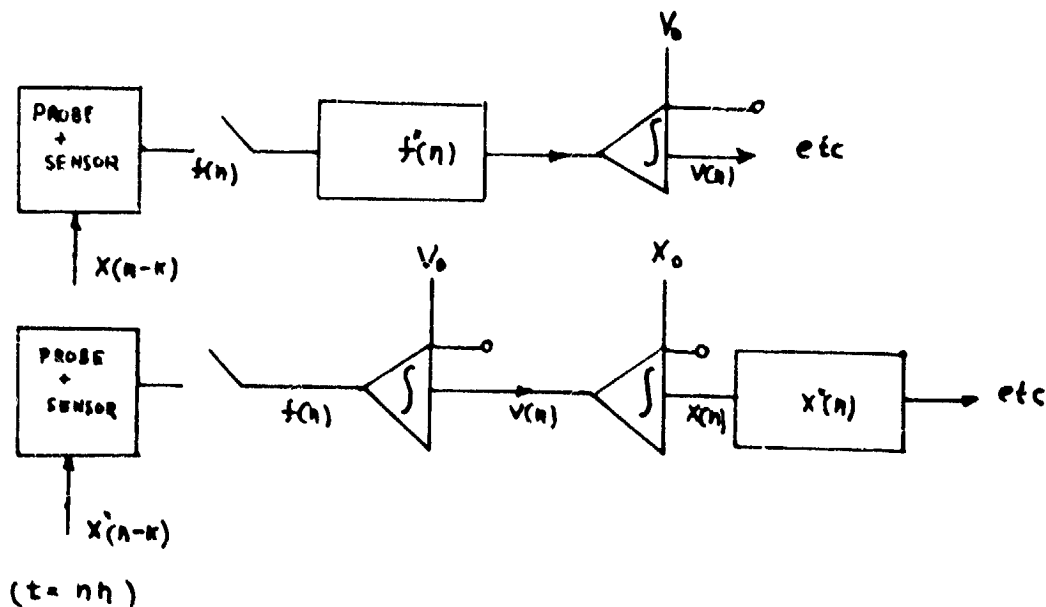
fig.2

To avoid repetitiveness we use the notation s(n) to refer to any one of the signals x(n), v(n) or f(n). It is then understood that s*(n) can mean x*(n), v*(n) or f*(n).

Our methods are all examples of digital filters. Let us call them "extrapolation filters". To motivate what follows we introduce some terminology from filter theory. Our filter will be said to be one step if s*(n) depends only on s(n); otherwise, the filter is multistep. The filter is recursive if the determination of s*(n) involves previous values of s*; otherwise, the filter will be called non-recursive. If there is a fixed upper bound on the number of past values of s or s* needed to calculate s* we say the filter has fixed memory; otherwise, it has expanding memory.

We test our extrapolations only in the simplest case of a pure spring. The results are presented in Section 3. The effects of noise are studied in Section 4.

2.1 One Step Taylor Filters

Since $x(t) = x(t-\tau) + \tau \dot{x}(t-\tau) + \tau^2/2) \ddot{x}(t-\tau) + \ldots + \tau^n/n! \; x^{(n)}(u)$, $t-\tau < u < t$. We can truncate the series and define

(T1) $\quad x^*(n) = x(n) + \tau v(n)$

or

(T2) $\quad x^*(n) = x(n) + \tau v(n) + \tau^2 f(n)/2M$

No higher degree of approximation is possible without differentating the force signal. However, it has been found that the formulas

(ET2) $\quad x^*(n) = x(n) + 9/8\tau v(n) + \tau^2 f(n)/2M$

or

(MM2) $\quad x^*(n) = .989039\, x(n) + 1.120184\tau v(n) + .554041\tau^2 f(n)/2M$

can give better results than a 4 term Taylor compensation (T3). The second formula (MM2) being about as good as the first.

2.2 Lagrange filters or polynomial extrapolation

To the points $s(n)$, $s(n-k)$, $s(n-ik)$, $i \leq m$ we fit an interpolating polynomial of degree m $P(t)$ and define $x^*(n)$ to be $P((n+k)h$ (Other choices for the spacing are possible but this seems to be the most stable.) The algorithm is initialized by applying a one step predictor such as T2, ET2 or MM2 to generate the points $s(k+1)$, ... $s(m+1)k$. The polynomial algorithm is applied to obtain $s^*((m+1)k+i)$ by extrapolation based on the points $s(mk+i)$, ... $s(k+i)$, $s(i)$, and so on.

The interpolation is done by the Newton forward difference interpolation formula:

$$s^*(n) = \sum_{i=0}^{m} \Delta^i s(n-mk) \binom{m}{i} .$$

For a derivation of the Newton forward difference formula and a FORTRAN subroutine see Conte and De Boor [2] p.45.ff. We give explicitly a few cases of this formula:

Linear (Ps1):

$$s^*(n) = 2s(n) - s(n-k)$$

Quadratic (Ps1):

$$s^*(n) = 3s(n) - 3s(n-k) + s(n-2k)$$

Cubic (Ps3):

$$s^*(n) = 4s(n) - 6s(n-k) + 4s(n-2k) - s(n-3k)$$

2.3 Osculatory Extrapolation

Since $x(n+1)$ is produced by a double integration of $f(t)$, $f(n+1)$ and $v(n+1)$ are available before $x(n+1)$. This suggests that it is wasteful to compute $x^*(n+1)$ only from knowledge of $x(n+1)$ and its past values. Instead, we might construct a polynomial such that

$$p(nh) = x(n) \qquad p((n-ik)h) = x(n-ik)$$

$$p'(nh) = x'(n) \qquad p'((n-ik)h) = x'(n-ik), \; i = 1, \ldots m.$$

This would give a polynomial of degree 2m-1. The additional boundary condition

$$p''(nh) = f(n)/M \qquad p''((n-ik)h) = f(n-ik)/M, \quad i = 1, \ldots m$$

gives a polynomial of degree 3m-1. Again $x^*(n) = p((n+k)h)$. Formulas in terms of divided differences and software can be found in [1], Ch. 1. In the low order cases it is possible to write down formulas explicitly. For example, the cubic polynomial interpolating $x(n)$, $v(n)$ and $x(n-k)$, $v(n-k)$ gives the estimate.

$$(\text{HP3}) \quad x^*(n) = 5\, x(n) - 4\, x(n-k) + 2\, (2v(n) + v(n-2k)).$$

2.3 Least Square PolynomialFilters

Instead of interpolating x, v or f, we can fit a low degree polynomial such that the least square deviation between it and x, v, or f is minimized at t=nh, ..., (n-mk)h. Such polynomials have long been used in smoothing and prediction problems.

2.3.1 Fixed Memory Filters

Here p is selected so that

$$(p(nh) - s(nh))^2 + (p((n-k)h - s((n-k)h))^2 + \ldots$$

$$(p((n-m)h) - s((n-m)h))^2$$

is minimized for all n and fixed m. Although the degree of p in theory can be up to m-2. In practice it is no higher than 3. (High degree least square polynomials give too much freedom to

simulate noise in the signal). We pass over the derivation of the form of the polynomial which is best done using discrete Legendre polynomials and give only the final formulas. For the details see Morrison [3] Ch. 7.

Linear   3 Points (LLS 3):

$$s^*(n) = 4/3\ s(n) + 1/3\ s(n-k) - 2/3\ s(n-2k)$$

4 points (LLS(4):

$$s^*(n) = s(n) + 1/2\ s(n-k) - 1/2\ s(n-k)$$

5 points (LLS 5):

$$s^*(n) = 4/5\ s(n) + 1/2\ s(n-k) + 1/5\ s(n-2k) - 1/10\ s(n-3k) -$$

$$2/5\ s(n-4k)$$

Quadratic   4 Points (QLS 4):

$$s^*(n) = 9/4\ s(n) - 3/4\ s(n-k) - 5/4\ s(n-2k) +$$

$$3/4\ s(n-3k)$$

Cubic   5 points (CLS 5):

$$s^*(n) = A\ B\ C\ (\ s(n),\ s(n-k),\ s(n-2k)\ s(n-3k),\ s(n-4k))^t$$

where A is the 1 x 4 matrix (1, -3/2, 7/2, -14), B ix the 4 x 4 diagonal matrix

$$
\begin{vmatrix}
1/5 & & \bigcirc & \\
& 3/10 & & \\
& & 4/14 & \\
& \bigcirc & & 1/10
\end{vmatrix}
$$

and C is the 4 x 5 matrix

$$
\begin{vmatrix}
1 & 1 & 1 & 1 & 1 \\
-1 & -1/2 & 0 & 1/2 & 1 \\
1 & -1/2 & -1 & -1/2 & 1 \\
-1 & 2 & 0 & -2 & 1
\end{vmatrix}
$$

2.3.2 Expanding Memory Filters

This filter fits a least square polynomial to the entire sequence s(n), s(n-k), s(n-2k), ... Thus as time elapses the polynomial is based on more and more points. It is possible to do the calculations recursively so that information from the previous polynomial is used for the new polynomial. Again we bypass the theory and give only the formulas. A complete development is given in [3] Ch. 9

Let e(n) denote s*(n) - s(n+k). Then

Linear:

$$s*(n) = 2s*(n-k) - s*(n-2k + 43e(n)/n+1) - 2(2n-1)$$

$$e(n-k)/(n+1)n.$$

Quadratic:

$$s^*(n) = 3s^*(n-k) - 3s^*(n-2k) + s^*(n-3k) + 9e(n)/(n+1) - 18(n-1)e(n-k)/(n+1)n + 3(3n-9n+8)\ e(n-2k)/(n+1)n(n-1)$$

Cubic:

$$s^*(n) = 4s^*(n-k) - 6s^*(n-2k) + 4\ s^*(n-3k) - s^*(n-4k) + 16\ e(n)/(n+1) - 24(2n-3)\ e(n-k)/n+1)\ n + 48(n-4n+5)\ e(n-2k)/(n+1)n(n-1) - 8\ (2n-15n+43n-45)\ e(n-3k)/(n+1)\ n\ (n-1)\ (n-2)$$

III. SIMULATION RESULTS

The tables and graphs in this section contain the results of some of the extrapolation schemes discussed in the previous section. We regard the probe as an ideal spring where M=311 slugs and the Hookes constant K is either 1200 lbs/ft or 60 lbs/ft. In the first case, a zero initial position (X_0) is taken. In the second a 69 lb preload is considered. The preload can be modeled by assuming an initial condition X =69/60=1.15 ft.

Several measures of error are employed to measure the effectiveness of competing compensation methods. In the first place if $x^*(t)$, $x(t)$ and $v^*(t)$, $v(f)$ denote the compensated and theoretical positions or velocities of the probe as functions of time t, we define the relative percentage error of position (RPEP) by

$$100|(x^*(t) - x(t))/x(t)|$$

and similarly the relative percentage error with respect to velocity (RPEV) by

$$100|(v^*(t) - v(t))/v(t)|.$$

Because no compensation method can work in the time interval [0,] there will be significant relative percentage errors in position or velocity during the first half of the trajectory. Velocity errors are especially prominent after reaching several thousand percent at the time of maximum compression. Relative position errors are less severe throughout most of the trajectory. However, they become large at the end (when $x(t) = x$) because of distortion in the period. Because of these sensitivities, we supplement RPEP and RPEV with the following measures based on the magnitude of the residuals in the least square sense:

(a) $E2P(t): = \int_0^t (x^*(s)-x(s))^2 ds$

(b) $E2V(t): = \int_0^t (v^*(s)-v(s))^2 ds$

(c) $RPE2P(t): = E2P(t)/\int_0^t x(s)^2 ds$

(d) $RP\ E2V(t): = E2v(t)/\int_0^t v(s)^2 ds$

(In the data below the above integrals are computed by a
rectangle integration rule with step size h.)

Fig 3

Behavior of Typical
Computation Algorithms

$k = 1200$ lbs/ft $M = 3.1$ slugs

$T = 3$ sec $V_0 = 1.44$ ft/sec

1 = Thrombical

2 = ET2

3 = $\rho \times 3$

4 = 915 9

5 = Uncompensated
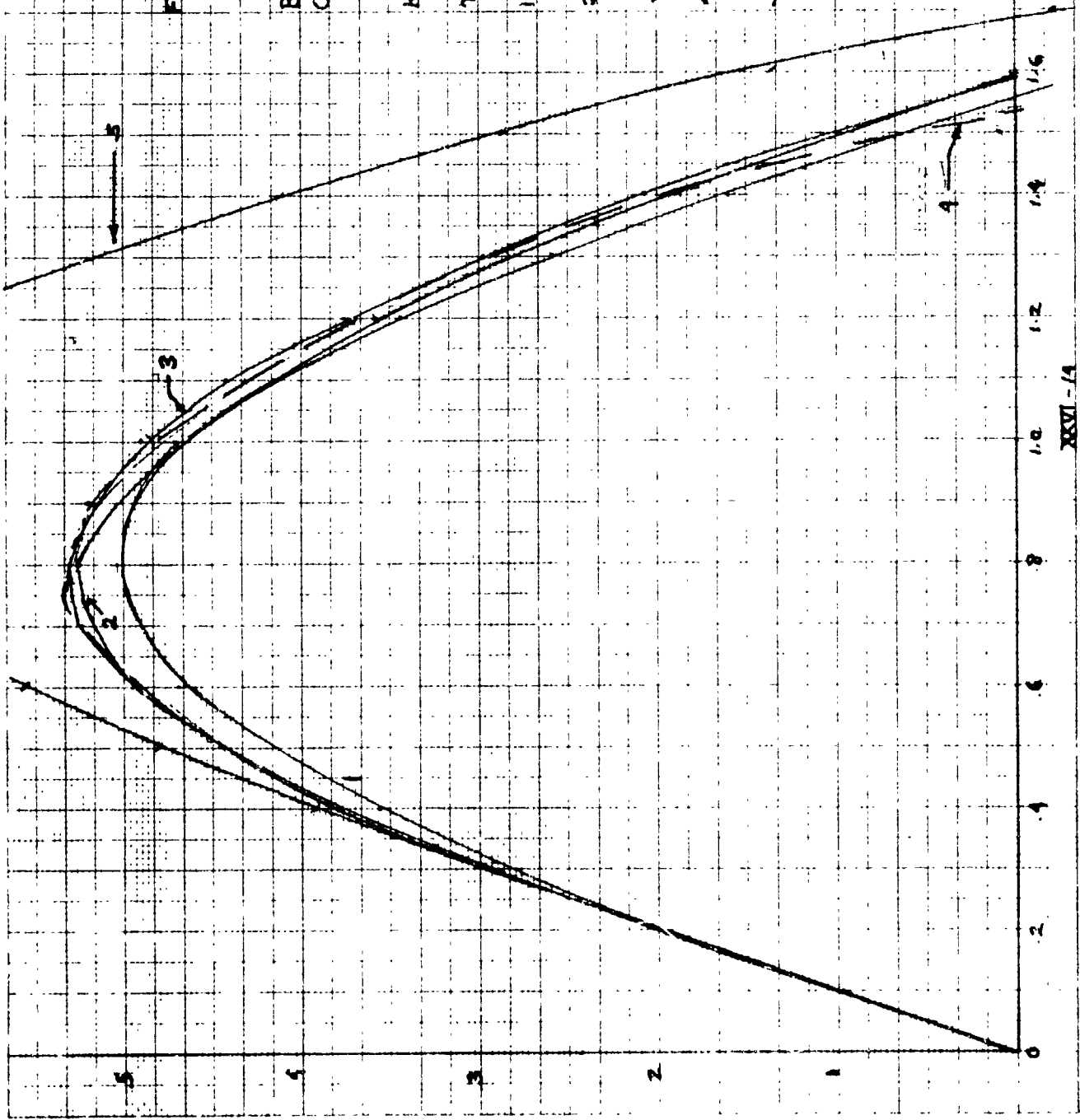
Residual Errors for Some Compensation Methods

K=1200 lbs. M=311 Slugs γ =.2 V_0 =.1

Algorithm	Position		Velocity		
	$E2P/10^{-6}$	$RPE2/10^{-6}$	$E2V/10^{-6}$	$RPE2V/10^{-5}$	Vf
UC[1]	403	19.6	1179	14.7	.156
T1	9.08	.443	103	1.29	.110
T2	8.28	.404	87.4	1.09	.110
T3	17.0	.829	43.9	.547	.108
ET2	4.4	.216	20.4	.254	.101
MM2	5.87	.287	54.6	.681	.103
Ps1[2]	34.8	1.70	256	3.19	.116
Ps2	3.3	.162	32.8	.41	.997
Ps3	5.21	.252	16.8	.209	.103
Ps4	6.77	.330	60.4	.753	.105
Ps5	8.52	.416	73.1	.912	.105
H3	15.8	.77	50.0	.94	.105
E5	7.3	.36	23.7	.47	.106
H7	9.3	.454	40.6	.763	.101
H9	6.8	.33	52.7	.97	.103
QLS4	4.18	.204	33.3	.416	.101
QLSM	6.49	.316	68	.85	.106
C80[3]	12.6	.613	26.2	.327	.106

See next page for notes   XXVI-15

1. Uncorrected

2. Polynomial extrapolation applied to position and force. Insignificant differences observed.

3. Best algorithm of 1960 "Fix & C" in [1].

XXVI-15a

Relative Percentage Errors in Velocity (RPEV)

K=1200 lbs   M=311 Slugs   T =.2   V₀ =.1

Algorithm	.4	.6	.8	1.0	1.2	1.4	1.6	$\bar{u}^{*}$
			Time					
UC	30.7	82.1	10233	61.1	5.4	26.3	56.4	43.7
TI	8.9	6.4	782.5	19.1	15.5	13.3	10.8	12.3
T2	9.6	9.4	304	14.8	13.3	12.0	10.2	11.6
T3	10.3	14.3	903.4	.2	4.3	6.1	7.5	7.1
ET2	7	3.7	453.1	8.2	.5	2.8	.81	3.8
MM2	7.1	1.2	1282.2	19.1	12.0	7.5	2.9	8.3
Ps1	7.1	.61	2115.7	35.9	26.7	27.7	16.2	18.0
Ps2	7.1	3.8	1536.3	15.5	7.3	3.2	.3	6.2
Ps3	7	3.9	900..	4.0	2.8	2.7	2.7	3.9
Ps4	7.1	.36	1593.5	23.1	11.7	7.2	4.6	9.0
Ps5	7.1	.36	1593.5	23.1	14.5	8.8	5.1	9.8
H3	6.8	8.7	1898.2	2.5	4.2	5.0	5.5	5.5
H5	6.9	5.8	2188.2	.64	2.9	4.4	5.9	4.4
H7	6.9	5.8	1660.1	3.4	3.8	2.8	1.35	4.0
H9	6.9	3.8	1869.2	6.7	3.5	3.6	3.1	4.5
QLS4	6.9	3.7	453.1	13.4	9.01	5.4	1.20	6.6
C80	8.9	11.4	3121.1	1.01	3.7	4.8	5.6	5.9

* $\bar{u}$ computed discounting error at .8.

$XXVI$-16
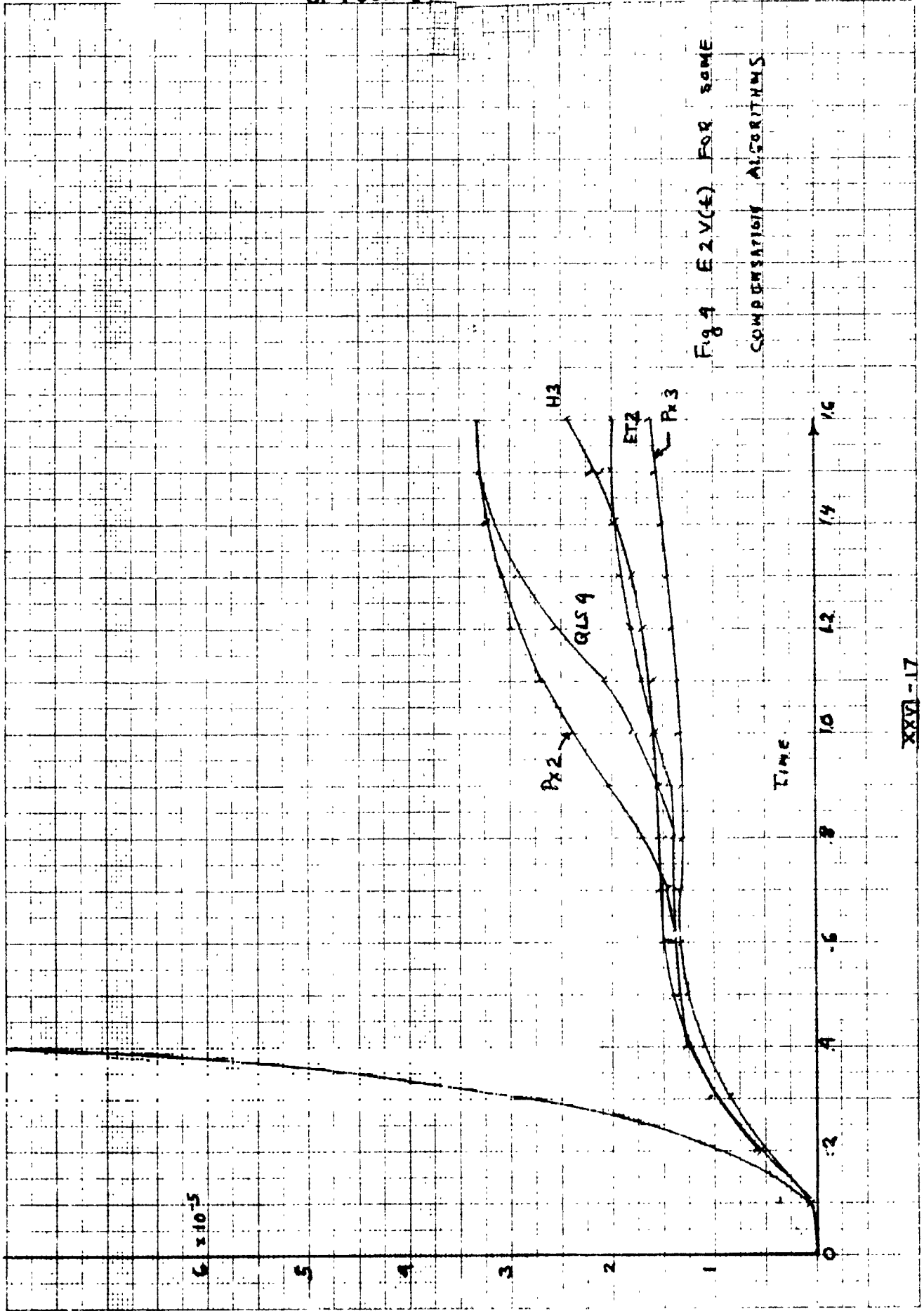
Fig 4. E2 V(t) For some compensation Algorithms

Relative Percentage Errors in Position (RPEP)
K=1200 lbs   M=311 Slugs $T=.2$ $V_o =.1$

Time

Algorithm	.4	.6	.8	1.0	1.2	1.4	1.6	$\bar{q}$ *
UC	9.9	19.0	30.4	45.3	67.5	115.4	4338.6	48
T1	5.6	6.2	5.7	4.1	.08	12.2	1453.4	5.6
T2	5.7	6.6	7.7	5.7	3.1	5.2	984.9	5.5
T3	5.8	7.2	8.3	9.6	11.7	17.2	601.9	10.0
ET2	5.2	5.3	4.9	4.2	3.6	3.5	100.3	4.5
MM2	5.2	5.2	4.0	1.7	2 3	12.6	1067	5.2
Ps1	5.2	5.2	3.5	.63	10.1	34.6	3485.1	10.7
Ps2	5.2	4.8	3.1	1.0	1.9	7.8	545.3	4.0
Ps3	5.0	5.2	4.8	4.7	5.1	7.1	650.9	5.3
Ps4	5.2	5.1	3.7	.99	3.6	14.6	1213.9	5.3
Ps5	5.2	5.1	3.7	.99	4.2	17.5	1448.66	6.1
H3	4.9	5.4	5.9	6.5	7.4	9.8	1344.6	6.7
H5	4.8	4.9	5.3	6.0	7.3	10.7	1721.3	6.5
H7	4.8	5.0	5.1	5.7	6.5	8.8	2058.8	5.9
H9	4.9	5.0	4.3	3.3	1.9	1.2	1428.2	3.4
QLS4	5.2	5.3	4.9	3.8	1.6	3.1	381.0	4
QLSM	5.1	5.3	4.6	3.2	.31	11.0	1184.0	5.4
C80	5.3*	.3	7.1	8.1	9.7	14.4	2700.7	8.5

* $\bar{q}$ computed discounting error at 1.6.   XXVI-18

Fig. 5 E2 P(t) for some

Compensatory Algorithms
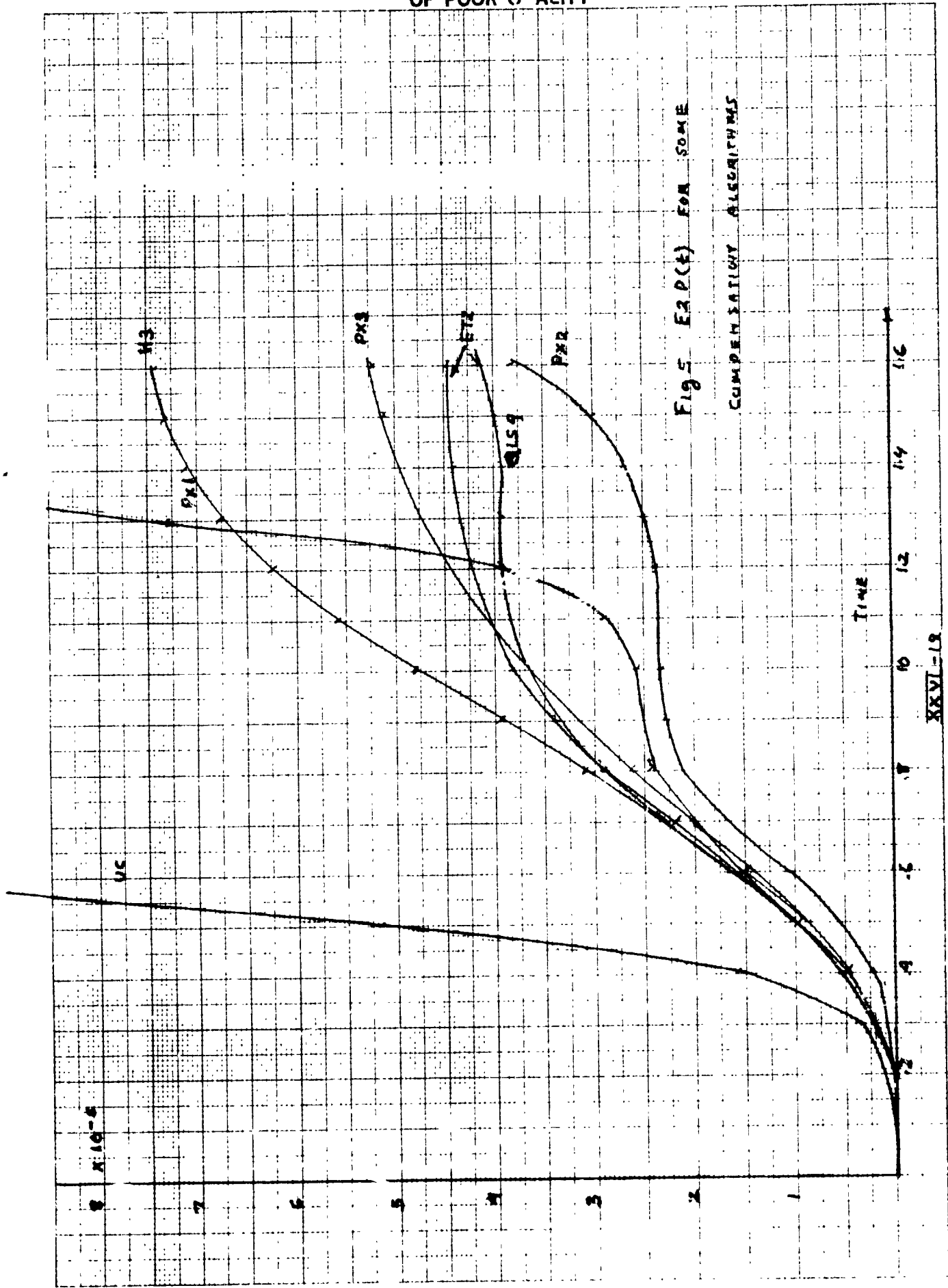
TABLE 4

Residual Errors in the Preloaded Case

K=60lbs/Ft   M=311 Slugs   Preload=69 lbs X_o =1.15 V_o =.1

Algorithm	V(.88)*	E2P(.88)/10^{-6}	RF2P(.88)	E2V(.88)/10^{-6}	RE2V/10^{-6}	V	t
		Position			Velocity		
UC	-.55	330	.028	1507	52.3	-.137	124
ET2	-.56	333	.028	1470	50.9	-.137	124
MM2	-.54	304	.027	1524	52.8	-.137	124
Px2	-.55	339	.028	1513	52.4	-.135	124
Px3	-.54	341	.029	1524	52.8	-.135	124
Px4	-.54	341	.029	1524	52.8	.137	124
QLS4	-.55	332	.028	1470	50.9	.137	124

* Approximate half period of theoretical motion

t = half period of compensated or uncorrected motion.

IV. EFFECTS OF NOISE

An important consideration in compensation design relates to sensitivity with respect to noise. Table 5 studies the performance of the probe in both uncorrected and corrected modes when the force signal is corrupted by a uniformly distributed noise whose maximum amplitude is 10% of the signal. Also, we illustrate phase angle and amplitude response functions for some of the algorithms in Table 6 and figs. 6-7.

TABLE 5

Some Data on Response to Noise

K=1200lbs/ft   M=311 Slugs V_θ =.1

	Approximate Standard Deviation of $E2P/10^{-6}$	Approximate Standard Deviation of $E2V/10^{-6}$
UC	104	268
ET2	20.8	97.2
Px2	15.9	59.2
Px3	16.7	39.6
Px4	25.4	79.5
Pf3	19.6	30.7
QLS4	10.7	39.3
QLSM	11.58	33.6

XXVI-22

DISCUSSION OF TABLE 5

Standard deviations were computed from four trials with a 10% uniformly distributed noise corrupting the force signal. We conclude that Px2, PF3, QLS4 and QLSM are probably less sensitive to noise than the others. The performance of all compensating algorithms except Px4 and ET2 is adequate. In no case did the exit velocity differ from the noise free by more than .03 ft/sec. PF2 is also good. However the tests on this were incomplete and are not shown.

TABLE 6

Some Transfer Functions

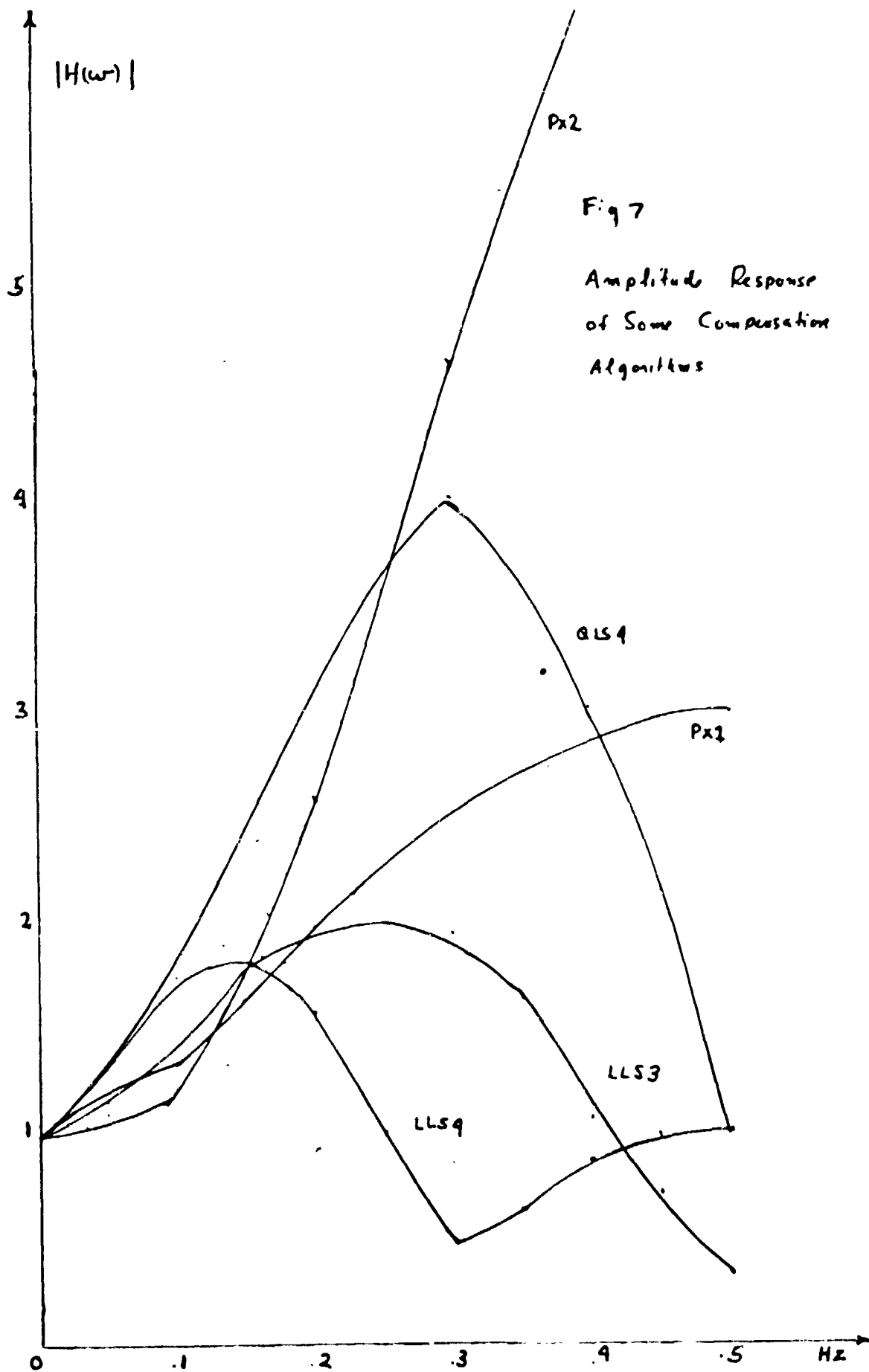| Filter	Re $H(\omega)$	Im $H(\omega)$
T1	1	
T2	$1-(\omega\tau)^2/2$	
T3	$1-(\omega\tau)^2/2$	$-(\omega\tau)^3/6$
ET2	$1-(\omega\tau)^2./2$	$9/8\ \omega\tau$
MM2	$.989039-.554041\,(\omega\tau)^2$	$1.130184\ \omega\tau$
Ps1	$2-\cos\omega\tau$	$\sin\omega\tau$
Ps2	$3-3\cos\omega\tau+\cos2\omega\tau$	$3\sin\omega\tau-\sin2\omega\tau$
Ps3	$4-6\cos\omega\tau+4\cos2\omega\tau-\cos3\omega\tau$	$6\sin\omega\tau-4\sin2\omega\tau+\sin3\omega\tau$
LLS3	$4/3+1/3\cos\omega\tau-2/3\cos2\omega\tau$	$-1/3\sin\omega\tau+2/3\sin2\omega\tau$
LLS4	$1+1/2\cos\omega\tau-1/2\cos3\omega\tau$	$-1/2\sin\omega\tau+1/2\sin3\omega\tau$
LLS5	$4/5+1/2\cos\omega\tau+1/5\cos2\omega\tau-1/10\cos3\omega\tau-2/5\cos4\omega\tau$	$-1/2\sin\omega\tau-1/5\sin2\omega\tau+1/10\sin3\omega\tau+2/5\sin4\omega\tau$
QLS4	$5/4-3/4\cos\omega\tau-5/4\cos2\omega\tau+3/4\cos3\omega\tau$	$3/4\sin\omega\tau+5/4\sin2\omega\tau-3/4\sin3\omega\tau$

Fig. 6 Phase Response of Some Compensation Algorithms

$|H(\omega)|$

Px2

Fig 7

Amplitude Response
of Some Compensation
Algorithms

OLS4

Px1

LLS3

LLS4

Hz

V. CONCLUSIONS

Tables 1-3 show that ET2, MM2, Ps2, Ps3, H5, H7 and QLS 4 give errors in the non-preloaded case which are noticeably smaller than the others. Of these ET2 is perhaps the best in all around accuracy. Truncated Taylor series, linear extrapolation, and linear least squares (not shown) do not work well. The use of high degree polynomials for data fitting is notoriously unreliable. Although we do not give the data here, our experience confirms this judgment. Polynomial or osculatory extrapolation for degrees greater than 5 gave progressively poorer results. Fitting a polynomial of degree 5 simultaneously to the position, velocity and force signals also was unstable. Similarly cubic least squares fixed or expanding memory filters gave poor results. Here the problem was not instability, but rather cubic least squares fitting allows so much freedom that $s^*(n)$ mimics the uncorrected signal.

Although we have managed to reduce considerably the position and velocity errors of the best 1980 filter ("C80" - Table 1), we hesitate to predict that significantly greater improvement is possible. Because of the delay the position and velocity residuals are 6.2×10^{-6} and 3.4×10^{-8} at 200ms. At $t = 2\tau = 400$ms these have increased to 1.1×10^{-6} and 4.9×10^{-7}. It follows that these values are probably a lower bound to the errors of any algorithm The least gain in velocity residual over the rest of the trajectory is given by Ps3-16.8×10^{-6} (about a 270% increase) while Ps2 gives slightly less than a 10 fold increase in position residual. Looking at it from a different point of view, these reduce the errors in the uncorrected trajectory in velocity by a factor of .014 and in position by a factor of .008. Likewise, relative percentage errors in the magnitude of velocity and position can be reduced by a factor of 10 to between 3 and 5%. Therefore, because of the "burden" of error early in the trajectory, it is unlikely that other methods can give better results unless performance at the beginning is improved. Likewise, we have found no way of eliminating the bad errors in velocity at the time of maximum compression or in position at return. Fortunately, these errors are only relative; the quantities involved are very small and so they have no qualitative significance or implications for the hardware (they are virtually unnoticeable on a graph).

The algorithms were also found to work well for damped springs, other initial velocities and delays. As delay is reduced the results improve dramatically.

In the case of a preloaded spring and low Hooke's constant as Table 6 shows, all our methods fail. None even significantly

perturb the uncorrected trajectory. This phenomenon was also found in [1]. We do not understand the mathematics but physically it seems to result from the potential energy locked up in the spring by the preload which is magnified by the delay early in the trajectory, "swamping" the compensation.

Any method of extrapolation depends on assumptions about the function being extrapolated. If nothing is known about the function except its past history, then prediction amounts to fortune telling. The methods in this paper all depend on the assumption that the function is locally "like" a polynomial. In particular, if n is the degree of the polynomial, the derivatives are assumed to exist up to n+1 and not to grow too fast. Error bounds then depend on the product of a power of the step and the highest derivative. This implies that if the function is not smooth then the results of extrapolation will not be good. We would not expect any compensation scheme to work therefore if as the probe was compressed impulsive forces were generated.

The behavior of an extrapolation algorithm, however, also depends on other factors than just accuracy of extrapolation. T3 is the most accurate extrapolation of the group for functions with two continuous derivatives. Yet its behavior is mediocre. Px2 or QLS2 are poor extrapolations but give good performance. The reason for this may lie in the frequency response characteristics of the algorithm.

In the frequency domain reversing a delay amounts to the construction of a transfer function $H(\omega)$ with linear phase advance $\phi(\omega) = $ arctan $($ Im $H(\omega)/$Re $H(\omega)) = $ such that $| H(\omega) | = $ 1. None methods look especially impressive in this regard (cf. fig.). A digital filter with exactly this transfer function would be noncausal. However, the requirement suggests the following approximation approach. If the filter is of the form

$$\delta'(n) = a_0 \delta(n) + a_1 \delta(n-k) + \cdots + a_m \delta(n-mk) +$$

$$b_1 \delta'(n-k) + b_2 \delta'(n-2k) + \cdots + b_m \delta'(n-mk) .$$

The filter has the transfer function

$$H(\omega) = \frac{\left(a_0 + a_1 e^{-j\omega\tau} + \cdots + a_m e^{-mj\omega\tau} \right)}{\left(1 + b_1 e^{-j\omega\tau} + \cdots b_m e^{-ja\omega\tau} \right)}$$

A possible strategy would be the determination of the parmeters a_i, b_i so as to minimize either

$$\int_0^\pi \left(\left(\omega\tau - \tan^{-1}\left(Im\ H(\omega) / Re\ H(\omega) \right) \right)^2 + \lambda \left| 1 - H(\omega) \right|^2 \right) d\omega$$

or

$$Max\ \left| \omega\tau - \tan^{-1}\left(Im\ H(\omega) / Re\ H(\omega) \right) \right| + \lambda \left| 1 - H(\omega) \right|$$

for ω in the Nyquist interval $[0,\pi]$

where λ is some parameter reflecting emphasis either on the phase gain or the amplitude.

A related strategy would be to approximate in the S plane. We approximate $e^{\tau s}$ by a rational transfer function having a stable denominator. This can be done for $e^{-\tau s}$ by Pade approximation. But the writer is unaware of any work for $e^{\tau s}$.

It would be of interest to pursue these avenues further and compare the results with the time domain methods emphasized here.

REFERENCES

Brown R. C., "The Effect of Time Delay in the 6 DOF Motion System and its Compensation," NASA Contractor Report NASA CR-161511, Marshall Space Flight Center, October 1980

Conte, S. D., and de Boor C., Elementary Numerical Analysis, McGraw-Hill, 1980

Morrison, N., Introduction to Sequential Smoothing and Prediction, McGraw-Hill 1969

1981

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

A NEW SPS ANTENNA DESIGN TO REDUCE GRATING LOBES

Prepared By:	C. H. Chan, Ph.D.
Academic Rank:	Professor of Physics
University and Department:	University of Alabama in Huntsville Department of Physics

NASA/MSFC
  (Laboratory)
  (Division)
  (Branch)
	
	Preliminary Design Office
	Subsystems Design Division
	Electrical System Branch

MSFC Counterpart:	Woolsey Finnell III
Date:	August 5, 1981
Contract No.:	NGT-01-008-021 (University of Alabama in Huntsville)

ACKNOWLEDGEMENT

A New SPS Antenna Design to Reduce Grating Lobes

By

C. H. Chan, Ph. D.
Professor of Physics
University of Alabama in Huntsville
Huntsville, Alabama 35899

ABSTRACT

In the SPS system, the antenna's misalignment will produce well-defined grating lobes. These grating lobe peaks occur every 440 km and are potentially hazardous to the environment. One way to suppress these peaks is to phase control every power module. The cost due to the increase in receiving electronics and processors, however, could prove to be prohibitive.

We propose here a new design of the antenna involving the addition of two broad gaps, one along the x axis and another along the y axis. The gap distance is exactly one half of the distance between the center of two neighboring subarrays. It is shown in this report, by detailed calculation of far field radiation patterns, that our design does reduce grating lobe peaks without sacrificing power in the main lobe.

I. Introduction

In the SPS system, the systematic tilt of the entire antenna structure due to attitude control system errors will produce well-defined grating lobes.[1] These grating lobe peaks occur every 440 km if the phase control is provided to the 10.4 x 10.4 meter subarray level. Based upon environmental considerations, the grating lobe peaks are constrained to be less than 0.01 mW/cm^2 which in turn sets a severe upper limit on the pointing error for the antenna to within an arc-minute. One way to lower this severe alignment requirement is to phase control every power module (approximately 98000 of them). In this case the upper limit on the pointing error can be lowered to within three arc-minutes. However the cost due to the increase in receiving electronics and processors could prove to be prohibitive. R. H. Durrett, in a recent NASA publication[2] suggested that one may design an antenna with varying subarray sizes and varying gap distances between subarrays to suppress the grating lobes. This report gives a specific design along the line of Durrett's suggestion and presents a detailed calculation of far field radiation patterns and its results.

The design we present is very simple. It involves the addition of two broad gaps, one along the x-axis and another along the y-axis. The gap distance is exactly one-half of the distance between the centers of two neighboring subarrays. To understand how the grating lobes will be suppressed in this design, we note that grating lobe peaks occur at places on the ground where the signals from each of the subarrays add in-phase. With the added gap, at the location where the first grating

lobe peak was, the signal from each of the subarrays from the left half of the antenna will be exactly 180° out of phase with those from the right half of the antenna, e.g. they cancel each other and produce a null there. In fact, all locations of odd numbered peaks (the 1st, 3rd, 5th, etc.) are nulls. With the first grating lobe suppressed and since the peak of the second grating lobe is lower, our simple design may be used as a means to relieve the severe constraint on the pointing accuracy.

We present our calculations and results in Section II and Section III. It is shown there that our design does reduce grating lobe peaks without sacrificing power in the main lobe.

II. Calculation

Figure 1 shows the geometry of the m,n^{th} subarray in relation to the ground receiving points. The center of the space antenna is located at (x_o, y_o, z) and the center of the rectenna at the origin. Based on Kirchhoff's scalar diffraction theory, the electric field at the receiving site (x, y) due to field at m,n^{th} subarray centered at (ξ_{mn}, η_{mn}, z) is

$$E_{mn}(x,y) = -\frac{i}{\lambda} \int \int E'_{mn}(x',y') \frac{e^{ikR}}{R} dx' \, dy' \qquad (1)$$

where $k = 2\pi/\lambda$ and $R = [(x-x')^2 + (y-y')^2 + z^2]^{\frac{1}{2}}$. $E'_{mn}(x',y')$ is the m,n^{th} subarray's field distribution and may be written as

$$E'_{mn}(x',y') = A_{mn} e^{-ikR_{mn}}$$

where A_{mn} is a constant over the entire m,n^{th} subarray and kR_{mn} is the phase determined from the phase control system

$$R_{mn} = [\xi^2_{mn} + \eta^2_{mn} + z^2]^{\frac{1}{2}}$$

XXVII-5

At far field, the integration over the entire m,n^{th} subarray in Eq. (1) can be carried out

$$E_{mn}(x,y) = \frac{-ia^2}{\lambda z} e^{ik(\frac{x^2+y^2}{2z})} A_{mn} e^{-ikx\xi_{mn}/z} e^{-iky\eta_{mn}/z}$$

(2)

$$\left\{ sinc \ [k(x-\xi_{mn})a/2z] \ sinc \ [k(y-\eta_{mn}) \ a/2z] \right\}$$

where sinc α = sin α/α is the sinc function and a is the size of the subarray. The total electric field due to the entire array is then the sum of the contributions from all the subarrays.

$$E(x,y) = \sum_{m,n} E_{mn} (x,y) = F_e F_a$$

(3)

where

$$F_e = \frac{a^2}{\lambda z} sinc \ [k(x-x_o)a/2z] \ sinc \ [k(y-y_o)a/2z]$$

(4)

is the element factor and

$$F_a = \sum_{m,n} A_{mn} e^{-ikx\xi_{mn}/z} e^{-iky\eta_{mn}/z}$$

(5)

is the array factor. In deriving Eq. (3), we have ignored the unimportant overall phase

$$-ie^{ik(\frac{x^2+y^2}{2z})}$$

and approximated x - ξ_{mn} by x-x$_o$, and y-η_{mn} by y-y$_o$ in the sinc functions. The element factor here is a slowly varying sinc function peaking at the boresight of the array x$_o$, y$_o$. The array factor is a more complicated function depending both on the field distribution and spatial arrangement of subarrays. Its peaks are well defined and their locations

do not move with misalignment changes of the subarray. This behavior is due to the operating characteristics of the retrodirective phase control system. When the subarrays are regularly spaced at 10.4 m, the array factor gives grating peaks every 440 km. If there is a pointing error or small gaps between subarrays, the grating peaks no longer fall on the zeroes of the element factor, creating undesirable grating lobes. To suppress these grating lobes, we propose here a new design involving the addition of two broad gaps, one along the x axis and another along the y axis; the gap distance is exactly one half of the distance between the center of two neighboring subarrays. Due to the cancellation between the contributions from the left half of the antenna and those form the right half of the antenna (also those from the upper half and those from the lower half), all the odd numbered grating lobe peaks are suppressed.

III: Results

To proceed to detailed numerical calculation of the far field radiation pattern, following values for various parameters are assumed.

Subarray size $a = 10.4$ meter

Far field distance $z = 37500$ km

Frequency $= 2.45$ GHz

Gap between subarrays $= 3.0$ cm

Pointing error $x_0 = -32.7$ and -10.9 km (or, misalignment 3 and 1 arc-minute)

$y_0 = 0$

We shall also assume a 10 db gaussian taper for the aperture illumination. The exact form of the distribution is shown in Figure 2 where

a ten step and a total of 7224 subarrays are assumed. We note here that
our distribution is slightly different from the one given in the refer-
ence system. In carrying on our numerical calculation, we noticed
that the distribution given in the reference system[1] is not symmetric
with respect to the interchange of the x and y axes and hence is not
correct. We modified it and made it symmetric with respect to the
interchange of the x and y axes.

With A_{mn} given, the $E(x,y)$ and hence the intensity at far field
can be calculated. Figures 3 and 4 show far field distributions on
ground receiving points along the x axis. Figure 3 is the one without
the extra broad gaps. The solid curve is that for a pointing error
of 3 arc-minutes and dashed curve is that for a pointing error of 1
arc-minute. The differences for the main lobe and side lobes between
these two misalignments are small and cannot be distinguished in our
plot. The first grating lobe peaks are located at 440 km on both sides
of the main lobe. These two grating lobes are not symmetric because of
the pointing error. The sinc function of the element factor shifts
toward one peak and away from the other, hence enhancing one peak and
reducing the other. Figure 4 is the same calculation with the addition
of two broad gaps. The gap distance is 5.245 meters between the two
halves of the antenna. We note that the main lobe remains unchanged,
the side lobes are slightly altered, but the grating lobes are greatly
distorted. Each grating lobe splits into two; this is because our new
design forces nulls exactly at locations where the peaks were, making
each grating lobe doubly peaked.

It is seen from Figures 3 and 4 that there is essentially no change
in the amount of energy contained in the main lobe. Without extra gaps,
the two nearby grating peaks are high, with 0.16 mW/cm² and 0.10 mW/cm²

for a 3 arc-minute pointing error, and 0.019 mW/cm^2 and 0.010 mW/cm^2 for a 1 arc-minute pointing error. With extra gaps, the grating peaks are lowered by a factor of about two and widths about doubled, making the total power contained in the grating lobes in the two designs about the same. Since it is the high power density which is potentially hazardous to the environment, evidently a lower peak, broad lobe is more tolerable.

IV. <u>Conclusion</u>

In this report we showed that by introducing two extra broad gaps crossed at the center of the antenna, we can reduce the grating lobe peaks without sacrificing power in the main lobe. Unfortunately, our design did not remove the grating lobes completely as one would like to. We believe that there are designs that will achieve this goal. Our calculation is a step in the right direction. Further investigation is required.

FIGURE CAPTIONS

Figure 1. Geometry of the antenna in relation to the ground receiving points.

Figure 2. Microwave array power distribution.

Figure 3. Power density on ground along the x axis, without the added broad gaps. Solid curve is the one for 3 arc-minute pointing error, dashed curve for 1 arc-minute.

Figure 4. Power density on ground along the x axis, with the added broad gaps. Solid curve is the one for 3 arc-minute pointing error, dashed curve for 1 arc-minute.

REFERENCES

1) Satellite Power System: Concept Development and Evaluation Program, Reference System Report, U.S. Department of Energy and the National Aeronautics and Space Administration. DOE/ER-0023, October 1978, pp 29-37. ·

2) Informal note "An Alternate Concept for SPS" by R. H. Durrett, MSFC/EC33, Feb. 12, 1981.

FIGURE -1 GEOMETRY OF THE ANTENNA IN RELATION TO THE GROUND RECEIVING POINTS

(X_0, Y_0, Z)

Y'

(ξ_{mn}, η_{mn}, Z)

mnTH SUBARRAY

X'

Z

R

R_{mn}

Y

O

(X_0, Y_0, O)

X

(X, Y, O)
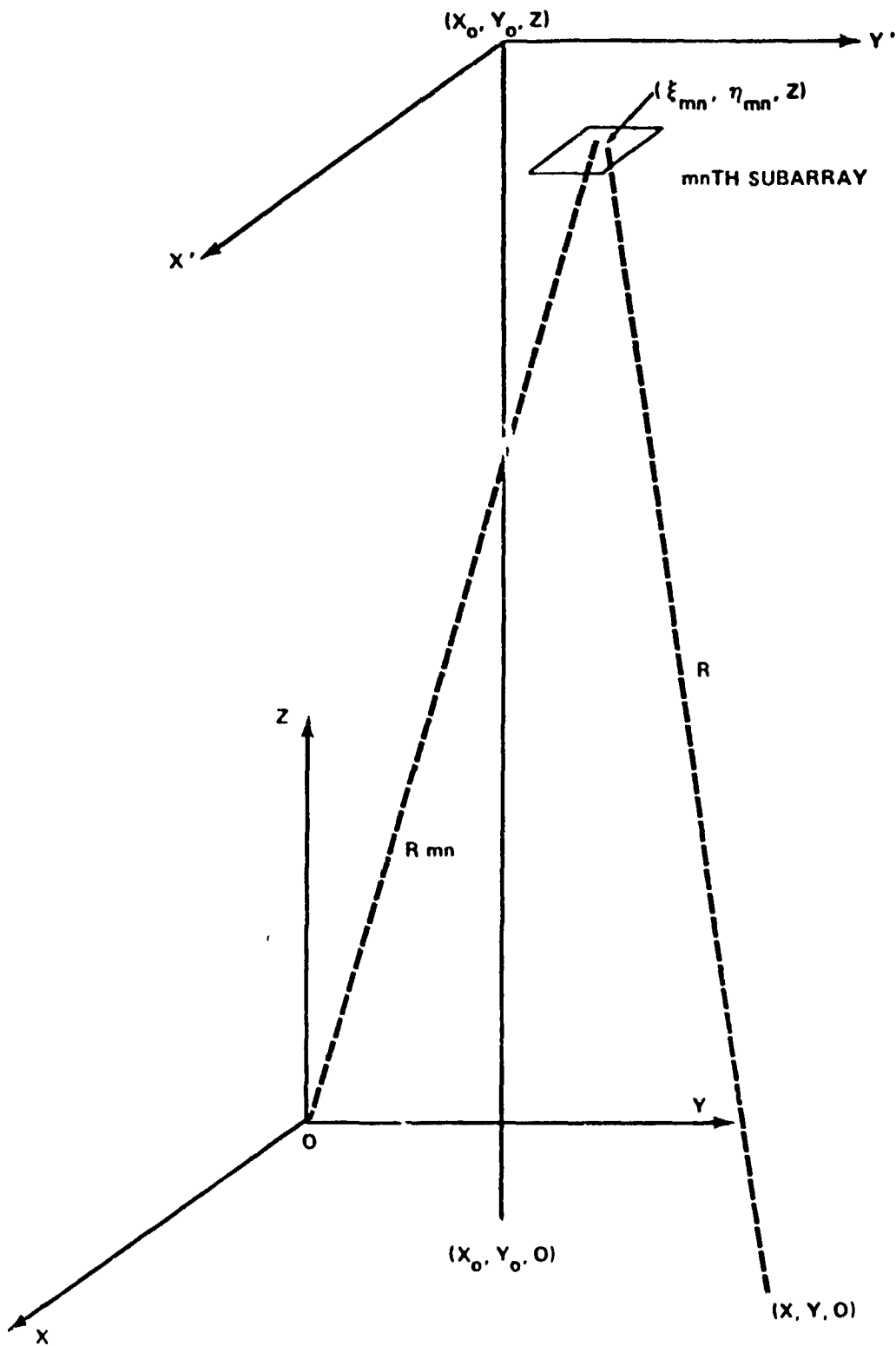
FIGURE 2 - MICROWAVE ARRAY POWER DISTRIBUTION

STEP	POWER DENSITY kW/M²	NUMBER SUBARRAYS
1	22.14	276
2	18.45	636
3	14.76	640
4	12.30	628
5	9.84	780
6	7.38	900
7	5.54	664
8	4.92	664
9	3.69	1008
10	2.46	1028
		7224
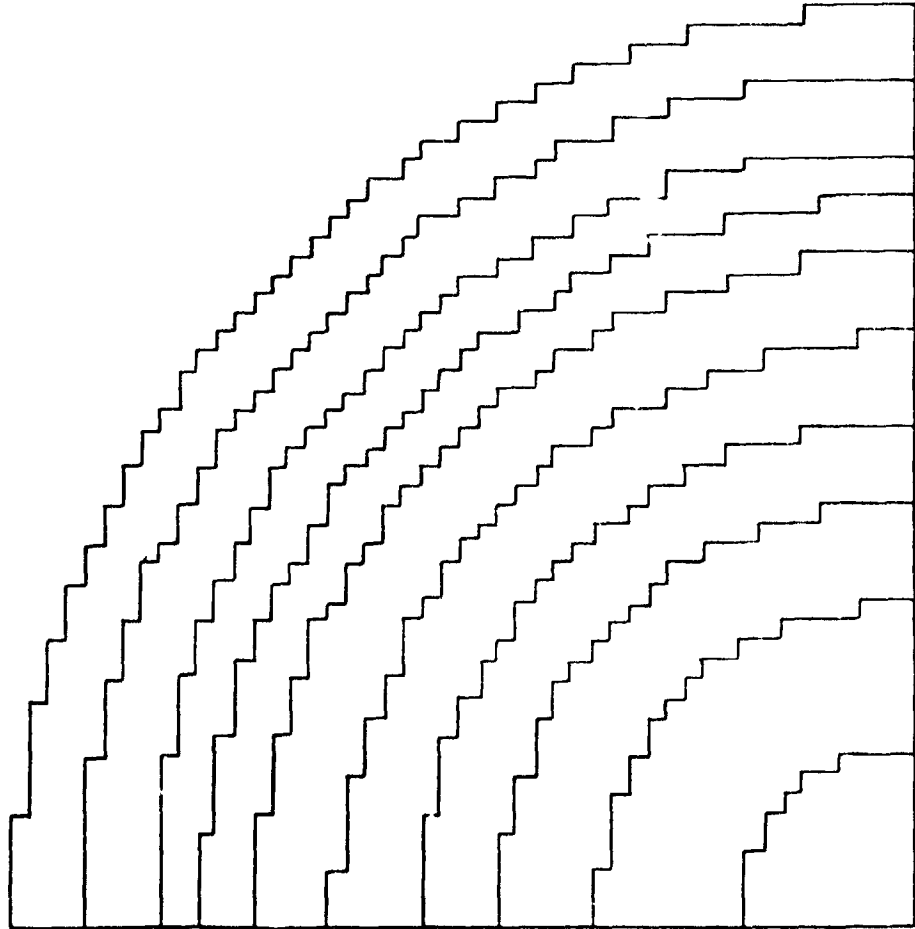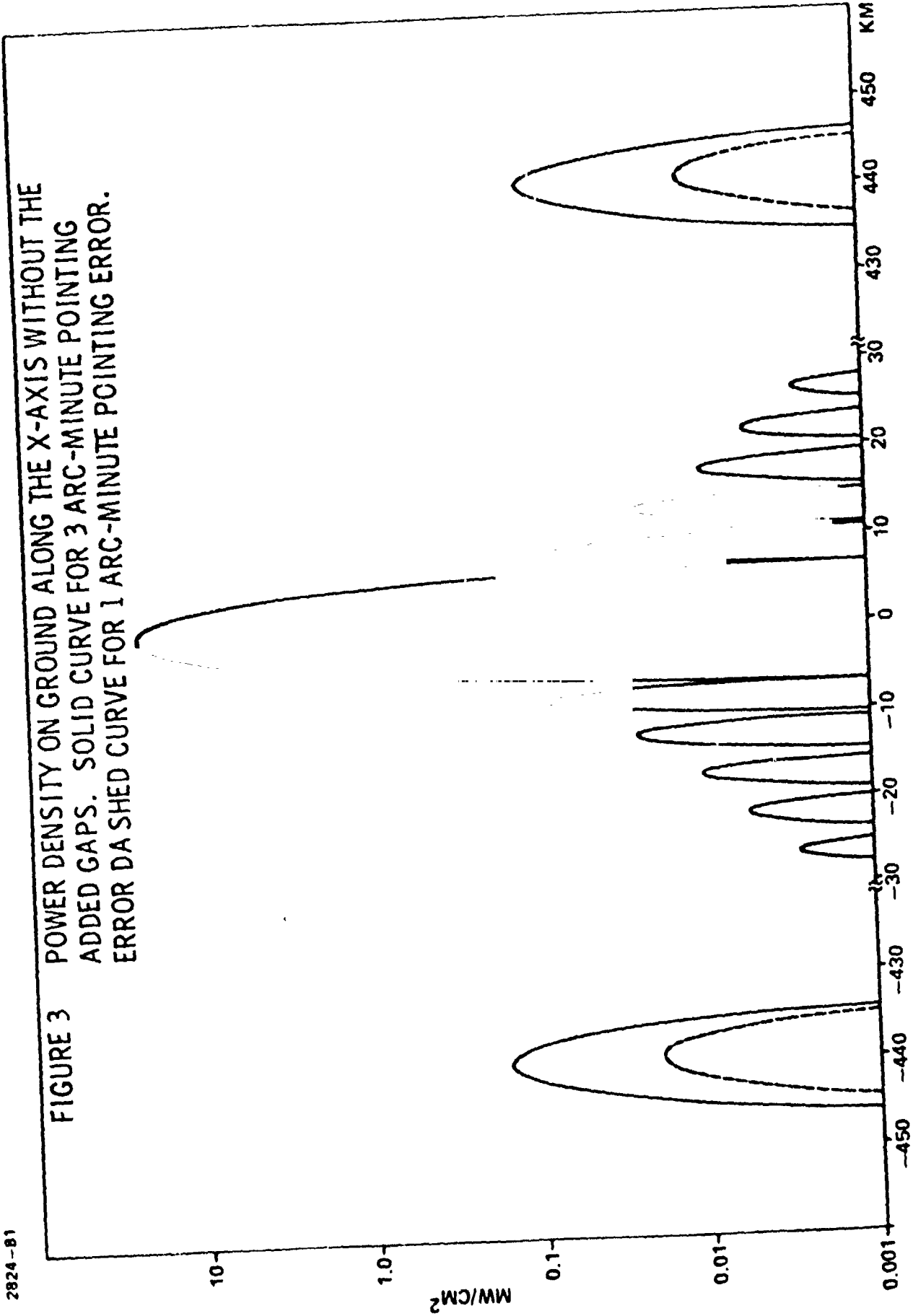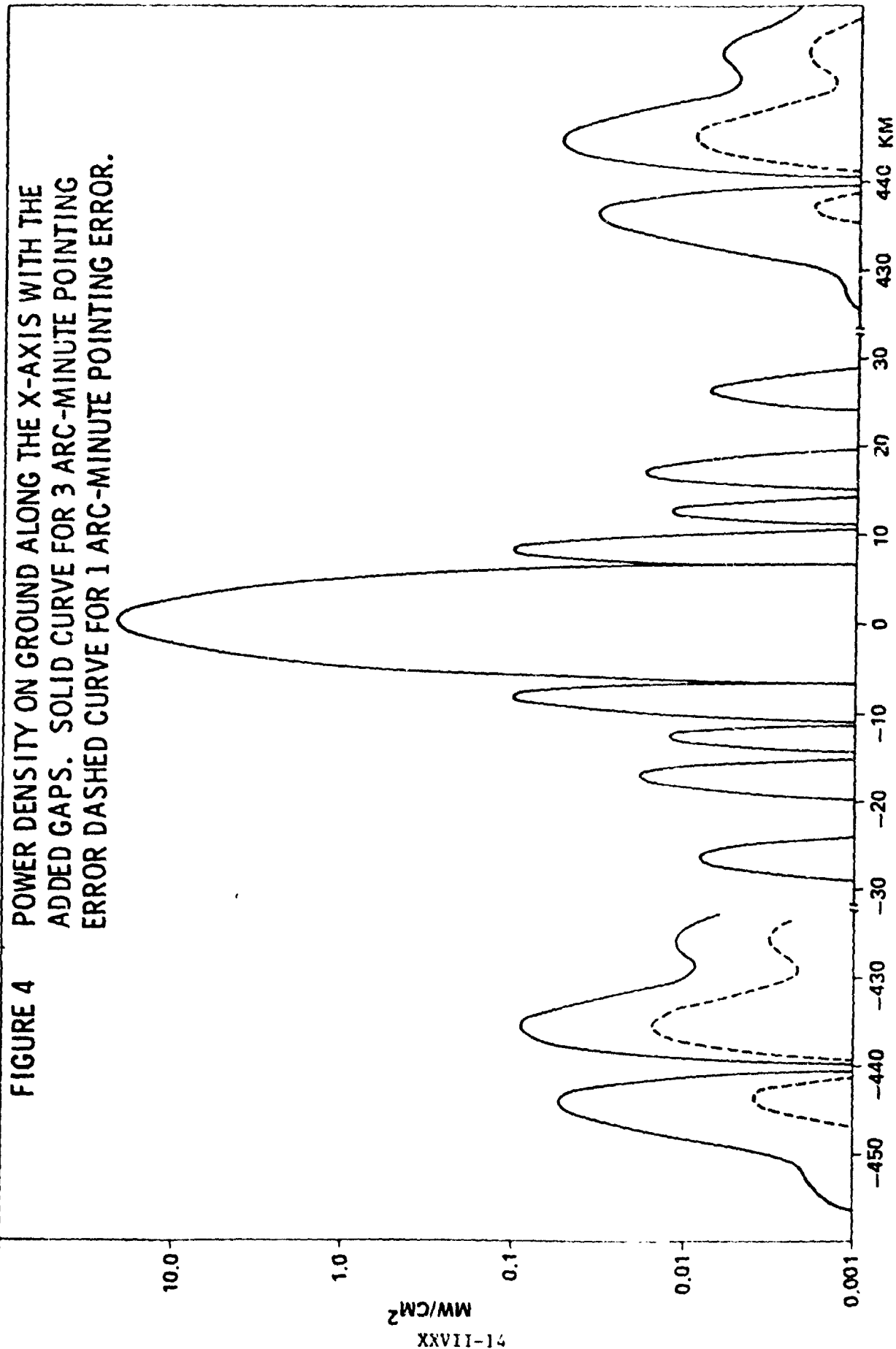
2822-81

FIGURE 3   POWER DENSITY ON GROUND ALONG THE X-AXIS WITHOUT THE ADDED GAPS. SOLID CURVE FOR 3 ARC-MINUTE POINTING ERROR DASHED CURVE FOR 1 ARC-MINUTE POINTING ERROR.

2824-81

XXVII-13

FIGURE 4 POWER DENSITY ON GROUND ALONG THE X-AXIS WITH THE ADDED GAPS. SOLID CURVE FOR 3 ARC-MINUTE POINTING ERROR DASHED CURVE FOR 1 ARC-MINUTE POINTING ERROR.

2825-81

1981

NASA/ASEE SUMMER FACULTY RESEARCH
FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

ALUMINUM AND CHROMIUM
ION PLATING STUDIES FOR
ENHANCEMENT OF SURFACE PROPERTIES

Prepared by: Jack H. Davis, Ph.D.

Academic Rank: Associate Professor

University and Department: The University of Alabama
                                      in Huntsville
                                      Department of Physics

NASA/MSFC:
      (Laboratory)   Materials and Processes
      (Division)   Engineering Physics
      (Branch)   Physical Sciences

MSFC Counterparts: Ann F. Whitaker and
                                      Dr. Alan P. Biddle

Date: August 11, 1981

Contract No.:   NGT-01-008-021
                                      University of Alabama
                                      in Huntsville

I.A. Abstract

An experimental project was undertaken which produced ion plated coatings on steel substrates. About 20 tensile samples of 4340 steel were ion plated in the Denton system with aluminum using resistance heating evaporation boats. In the V.T.A. 7375 system, ten samples were chromium ion plated; four on 4340 steel disks and the other six onto 440-C stainless steel rods for roller bearing wear improvement testing. Each of the samples was plated on a separate run to correlate the film parameters with the run parameters. Some of the chromium literature was reviewed, and improvements to the vacuum system were made.

I.B. Acknowledgements

I.C. Introduction

Holliday[12] gave the following concise introduction to ion plating.

Since first being reported in the literature in 1963 (ref. 1), ion plating has progressed to the point that today it is used in several commercial processes including the aluminum coating of fasteners used in the manufacture of aircraft (ref. 2). Ion plating has several positive qualities. Probably the most important of which are outstanding film adhesion and deposition on all sides of the substrate (including coverage into cavities) whereas normal vacuum deposition gives the usual line-of-sight coverage. Other qualities obtainable, depending on the material deposited, include exceptional corrosion resistance, high film purity, fine grain structure, very low-coefficient of friction, and improved mechanical properties of metals (ref. 3-5).

Several articles have been published describing the ion-plating process (e.g., ref. 6 and 7). Basically the process consists of two phases. In the initial cleaning phase the object to be plated (substrate) is made the cathode of a dc inert gas discharge or plasma. Some of the inert gas atoms are ionized and accelerated towards the substrate. The bombardment of the substrate by these ions having high kinetic energy produces a clean surface preparatory to the actual plating process. In the plating phase the material to be plated is evaporated from the anode while maintaining the inert gas discharge. Some of the coating material atoms are also ionized and accelerated towards the substrate. These ions follow the electric field lines which terminate on all sides of the biased substrate and thereby help coat not only the front but all sides of the substrate. However, since the degree of ionization in many plasmas is very low (0.1-2%), it has been suggested that this mechanism is probably secondary to gas scattering of neutral atoms in contributing to the high throwing power of ion plating (ref. 8). Thus the energies of the impinging particles range from that obtained from ions accelerated by the potential difference between the anode and cathode to the thermal energy of unionized atoms. These factors, along with the heating of the substrate surface due to its continued bombardment, produce the graded-fused interface, which provides the superior adherence and improved mechanical properties characteristic of ion plating.

I.D. Background

Last year's report[9] ended on a mixed note with the e-beam inoperative during the last five weeks due partly to a faulty transformer. A replacement was ordered. During the winter Dr. A. Biddle joined MSFC and replaced the above transformer and made numerous improvements both to the vacuum chamber and to the e-beam generator while coating several substrates. This year the V.T.A. system was nearly faultless except once; however, cleaning the high voltage feed thru contacts corrected the problem within an hour.

Literature surveys by the RSIC computer search have generated over 300 pages of abstracts in the following categories:

Ion Implantation

Ion Plating

Hydrogen Embrittlement

Chromium Vacuum Deposition

II. Aluminum Ion Plating Overview

The aluminum ion plating objective is twofold. The first is to develop and refine the ion plating technique in the Denton vacuum system. The second objective is to ion plate hardened tensile test specimens with aluminum so as to determine, via stress-strain curve analysis, the degree to which the aluminum film protects against 5000 psi hydrogen embrittlement.

II.A. Denton System Improvements

The Denton vacuum system consists of a diffusion pump coupled via a cold trap (LN_2) to 24" bell jar of stainless steel.

During the ion plating one desires to bleed about a 30µ Ar pressure into the jar. It is being pumped by a 10^{-6} torr diffusion pump vacuum. However the valves on the system are automatic pneumatic types being either fully open or fully closed. A new crank adjustable high vacuum valve would be expensive in our case. A micrometer adjustable stop could preserve the desired crack as the valve flips to the closed position. Again the reason for the concern for an adjustable high vacuum valve is that most of the Al ion plating was done, contrary to ideal procedure, by pumping the system un-trapped through the roughing line valve. On a few runs in order to pump through a cold trap the diffusion pump heater was turned off and the jar was pumped through the diffusion pump cold trap.

A less expensive, more reversible, solution would be the deployment of a 14" dia. disk with a center orifice hole about 1/2" in diameter across the pumping throat to achieve the desired differential pressure. The glow discharge current and voltage were supplied by a Denton Model 801 (e-beam supply).

The filament, or boat supply, a 9V D.C., 500A Molectric Summa Processing, sends current through Cu electrodes to the intermetallic ($BNTiB_2$) boat. Graphite sheets have been procured for interfacing the Cu to the boat, but not in time for these Al ion platings.

II.B. Tensile Specimen Testing

Ten tensile specimens were formed from a 1" dia. 4340 (cold rolled) steel rod at UAH. They were heat treated at NASA to ~ R.C. 50 and then fine ground at UAH.

Six of our ion plated tensile test samples were
Auger evaluated by Dr. J. McClure. The only rough samples
[T-?, T-7 (run 18), T-13, and T-19] showed evident of gross
deep carbon and oxygen contamination. After such refinements
as (a) pumping through the cold trap, (b) replacing a Teflon
plate with a glass plate in the ionization area as suggested
by Dr. Biddle, (c) improving the electrode to boat contact
to stop electrode melting at high current, (d) cleaning the
vacuum jar insides (especially the electrodes) after each
use, samples T-2 (run 48) and T-21 (run 53) showed contami-
nation reduction by 1 to 2 orders of magnitude both on the
surface and after 45 minutes of etch (~ 500 $\overset{o}{A}$).

Quench oil black due to the hardening in the threads
could contribute to the carbon, however T-21 (run 53) was
the only unhardened mild steel substrate used and was no
better than T-1 (run 48).

III. Chromium Ion Plating Overview

The first objective is to develop and refine the ion
plating technique in the V.T.A. system.

The second objective is to ion plate Cr films onto
hardened 440C stainless steel roller bearing surfaces so
that others may evaluate the degree to which the bearing
capacity or life is improved.

III.A. Literature Findings

Nestell and Christy report[10] chromium 200 to 2000 $\overset{o}{A}$
thick vacuum evaporated films have much reduced reflectivity
relative to the bulk, as long as the substrate temperature
is below 700 ^{o}C. Since ion plated surfaces are normally
duller[14] than vacuum evaporated it is not surprising that
our chromium films were not lusterous.

III.B. Sample Configuration and Treatment

The 440-C roller specimen geometry and heat treatment
procedures are shown in Fig. 1. The only variation in the
above treatment was the step B where liquid nitrogen was
used (77 ^{o}K) in place of the -100 ^{o}F. The righthand center
hold was "removed" by making the original sample longer and
then cutting short on one side.

Heat treat to RC 53 minimum per RA1111-002 except:

A. Harden at 1935 $\pm$ 25 $^\circ$C Oil Quench

B. Refrigerate A minus 320F $\pm$ 10F 30 minutes minimum.

C. Temper 325F $^{+25F}_{-\ 0F}$ one hour minimum.

D. Refrigerate in LN_2 (-320F)

E. Temper 325F $^{+25F}_{-\ 0F}$ one hour minimum.

(76.2 $\pm$ .15 mm)
3.00 $\pm$ .05

6-8

.00635 mm

.00025

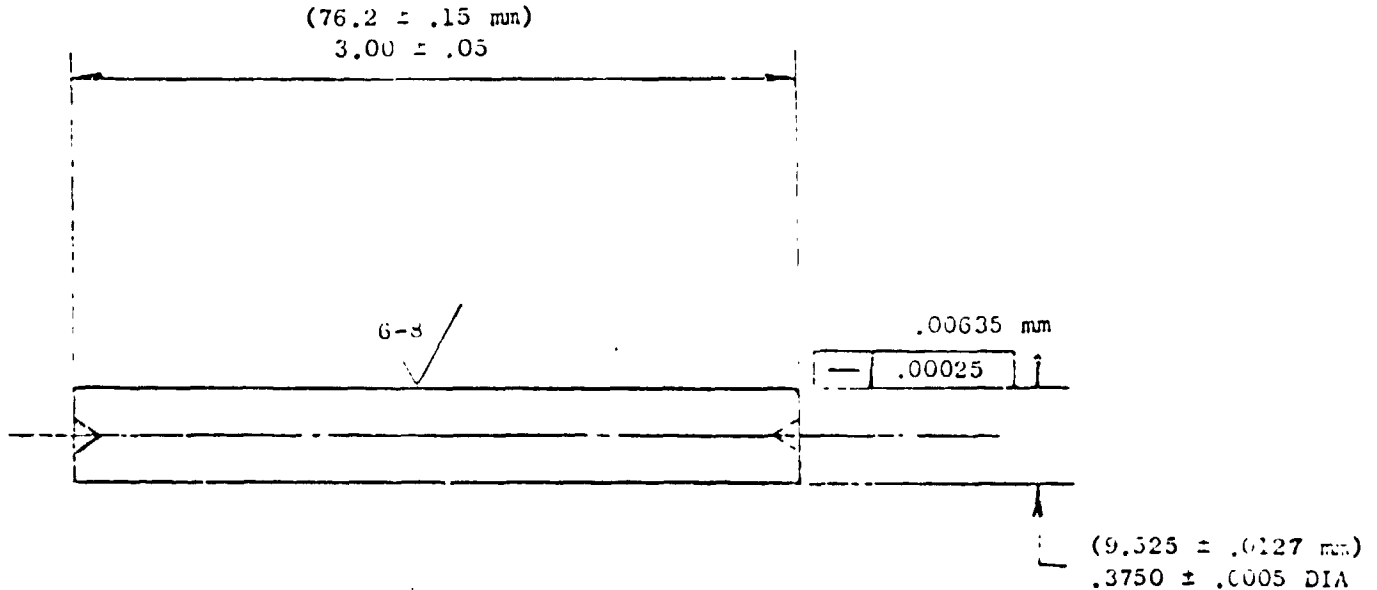(9.525 $\pm$ .0127 mm)
.3750 $\pm$ .0005 DIA

Figure I. RC Rig Specimen Geometry and Specification

Two rod sample holders were fabricated. The first being magnetic and thus allowing the complete cylindrical surface to be coated. The second was a sleeve or collar type providing better mechanical stability but shielding the top cm of the sample.

III.C. Substrate Temperature Measurements

In order to measure the 3 KV sample surface temperature without the problem of arcing thermocouple wires, an infrared pyrometer was used. Even the thick pyrex jar walls caused only about a $\frac{1}{1.6}$ x drop in temperature meter reading (^{O}C) over the tested range 80 ^{O}C to 400 ^{O}C. So when reading through the bell jar glass a 1.6X correction factor is used. Of course the disadvantage is that the window slowly plates over.

Fortunately (and surprisingly) the optical infrared thermometer (Omega Model 600) did not vary noticeably (i.e., 10 ^{O}C) upon extinguishing the glow discharge surrounding the sample. A flip-up cradle would be ideal for aiming the handheld I.R. gun at the thin sample.

III.D. Experimental Findings

Sample #25 polished 1.5" disk with polished Cr film generated the following peaks from X-ray diffractometer.

h l	2θ	I(CPS)
(110)	44.5	150
(200)	64.7	6
(211)	81.7	15
(310)	115	3

Note that the film grains have a highly nonrandom orientation with most of the cell cube faces lying in the surface plane of the film.

Higham's[13] ion plated chromium films initially have the (110) plane dominant. But as a stronger, denser discharge is generated by voltage applied to a 3rd electrode, his film orientations change to (200).

Our coating's preferred (110) orientation was typical of cooler substrates while Higham and Teer's[13] was approaching the character of films coated either at higher substrate temperatures or at greater voltages on their 3rd electrode.

Thorton[15] shows (see Figure II) graphically in 3-D
the effect of both temperature and pressure on the sputtered
thin film texture. The column structure existing from
0 < T/Tm < .3; however at higher substrate temperature the
atoms have enough lateral mobility to merge the grains
together smoothly. Thorton reports that increasing the
argon pressure enhances the column structure perhaps by
argon adhesion in the intercolumn voids. In fact he also
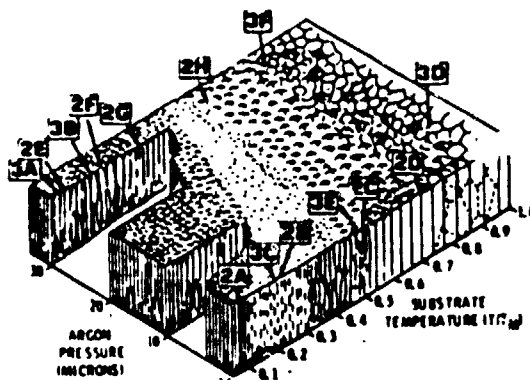gets dull dark gray films at 30μ Argon~room temperature
substrates.

Figure II. The effect of temperature and argon
            pressure on sputtered film texture. (Thorton[15])

The most dramatic observation from the 7 rod chromium
coatings is that the dark gray dull films were coated at
27μ and 28μ average coating pressure while the other 4 coatings
were brighter and shinier (but not lusterous) at pressures
of 14, 16, 16, 13 + 2u.   This is identical to the trend
noted by Thorton[15].

Average film thicknesses based on mass gained for
samples 1-7 were roughly 0, 3, 3, 6, 5, 0.2, and 6.6μ,
respectively. The Cr films were soft and could be marred with
a pocket knife blade. Dec-Tac scans across the knife marks
showed a film depth of ~.8 near the middle of sample #3.
The scan showed a bulge in film thickness where the film inter-
faces the shielded area.

Auger impurity analysis (highly qualitative) of Cr on
940C sample #5 (roller rod) shows the following estimated
concentrations relative to Cr. Before etch C (.1), O (1.2),
Fe (.01), Cr (1), after 135 min. etch C (.05), O (.3),
Fe (.01), Cr (1). Auger analysis of this Cr film showed the
expected large concentration of C and O on the surface. There
is some disappointment in the large O and C found inside. The

steel vertical sample holder (collar and rod) were not cleaned between runs and contained chromium to chromium oxide which could have sputtered onto the sample. Also, some gas (CO_2 and O_2) may be trapped in the column structure the film should have.

Carpenter[18] points out that when possible (i.e., if the extra throwing power is not needed) to stop ion plating after coating $\sim$.01μ and continue at lower pressures with regular vacuum deposition. He also points out the vacuum evaporation in general and ion plating in particular are wasteful of source material.

The strong surface brightness - Ar pressure dependence went unnoticed on the Al samples. Since other parameters were being varied the pressure was always held to 30μ.

The handbook etch for Cr, one part HCl, one part water and some zinc satisfactorily etched the Cr at about a mil per minute. Of course Cr plated tools and hardware must be protected.

IV. Black Powder

Previous investigators have expressed concern about a thick black deposit which covers the system inside walls after ion plating. Since some of the resulting samples have a darkish tinge the concern is justified. Suggestions that the glow discharge breaking down such hydrocarbons as pump oil or Teflon were investigated. One file letter suggests the Al black to be aluminum oxide with impurity color centers.

X-ray diffractometer analysis of black looking Al ion plated films on Cu substrates indicate only the presence of Al (small peaks) and Cu (large peaks) with no Al_2O_3 peaks detectable. So probably no more than 10% of the Al had oxidized into crystalline Al_2O_3.

An X-ray diffractometer scan of the black powder indicates mainly the presence of Cr, but also indicates Cr_3O_4 with no peaks found for the three of the other oxides (CrO, Cr_2O_3, Cr_3O_2).

In fact, Allen[16,17] patented a device for forming .1 to .01 micron sized metal powders. The device is almost identical to our ion plating unit except that it has a funnel shaped bottom to catch the particles and his Ar pressure (50 to 500μ) was slightly higher than ours ($\sim$30μ).

Carpenter[18] in mentioning the advantages and disad-
vantages of ion plating points out that it is even more
wasteful of source material than vacuum evaporation.

After ion plating chromium and allowing the bell jar
to cool overnight under vacuum, air was vented in the next
day which caused an amber-like fire glow around the hearth
and some flashes of sparks on the bell jar walls. Presumably
this is due to Cr oxidation. In fact, on one occasion the
acetone soaked cleaning napkin was turned into a flaming
torch. Auger analysis of the black powder indicates the
presence of chromium and oxygen. Such behavior would be
expected from fine chromium particles with large surface-to-
mass ratios.

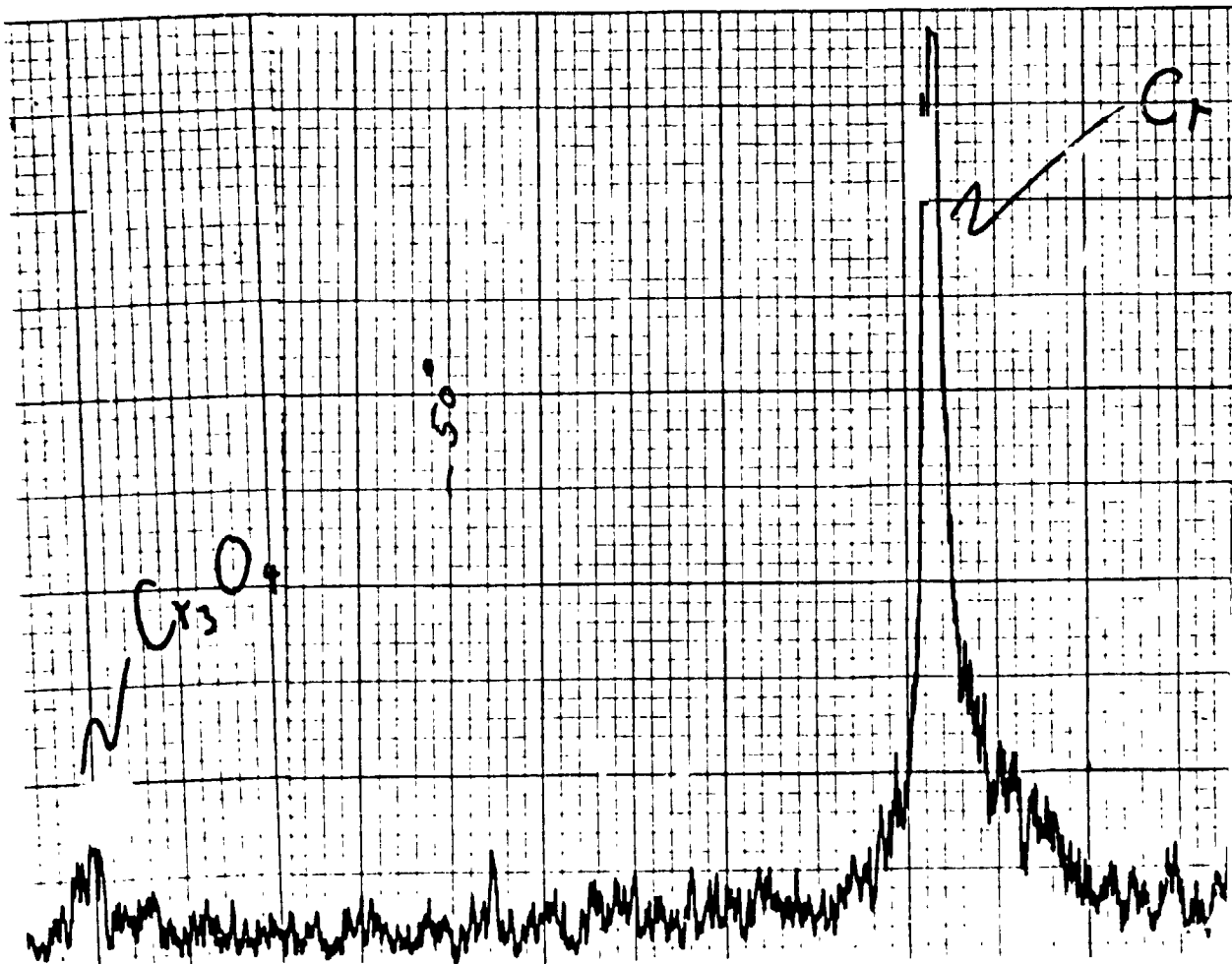Figure III. X-ray diffractometer scan of "black powder"
showing both chromium and chromium oxide.

V. Conclusions

Both vacuum systems have just proven themselves capable of producing good ion platings on a regular basis.

A host of two dozen samples are now available for further analysis. SEM-EDAX, Knoop hardness testing, and surface reflectivities could be easily done before destructive tensile and roller testing.

REFERENCES

1. Mattox, D. M., "Film Deposition Using Accelerated Ions," Rept. No. SC-DR-281-63, Sandia Corp., (Nov., 1963); Mattox, D. M., Electrochem. Technol., 2 295 (1964).

2. "Versatile Aluminum Coating Controls Corrosion," Iron Age, (February 9, 1976).

3. S alvins, T., "The New Applications of Sputtering and Ion Plating," NASA TMX-73551 (1977).

4. Chambers, D. L. and Carmichael, D. C., Research/Development, 22, 32 (1971).

5. Spalvins, T., J. Am. Soc. Lub. Eng., 27, 40 (1971).

6. Mattox, D. M., Proceedings of the Conference on Sputtering and Ion Plating, NASA SP-5111, NASA Lewis Research Center, Cleveland, Ohio, p. 41 (1972).

7. Mattox, D. M., J. Vac. Sci. Technol., 10, 47 (1973).

8. Aisenburg, S. and Chabot, R. W., J. Vac. Sci. Technol., 10, 194 (1973).

9. Davis, J. H., Ion Plating Studies for HIgh Temperature Application, p. V, Oct. 198C, NASA CR-161611.

10. Nestell, J. E. and Christy, R. W., J. Vac. Sci. Technol., 15(2) p. 366 (1978).

11. Holliday, R. J., "Ion Plating Investigation," BER Report No. 224-94, The University of Alabama, p. Xlll (Sept. 1978).

12. Holliday, R. J., "Ion Plating Studies for h Temperature Applications," University of Alabama i. Huntsville, Report No. 227, p. Xlll (Oct. 1979).

13. Higham, P. A. and Teer, D. G., Thin Solid Films 58, p. 121 (1979).

14. Jones, K., Griffith, A. J., and Williams, E. W., Ion Plating with Electron Beam Evaporation Conference on Ion Plating and Allied Techniques, p. 115, Edenburg 1977.

15. Thorton, J. A., J. Vac. Sci. Technol., Vol. II, No. 4, p. 666 1974.

16. Allen, L. R., et al, U. S. Patent #3,049,421, 14 Aug. '62.

17. Allen, L. R., Chem. Week 91, 47 sec. 1 '62.

18. Carpenter, R., "The Basic Principles of Ion Plating," Science and Technology of Surface Coatings, p. 393, ed. by Chapman, B. N. and Anderson, J. C. Academic Press (1974).

1981

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

STATISTICAL ANALYSIS OF DAILY MEAN TEMPERATURES

Prepared by:	Deva C. Doss, Ph.D.
Academic Rank:	Professor
University and Department:	The University of Alabama in Huntsville Department of Mathematics
NASA/MSFC (Laboratory) (Division)	Space Sciences Atmospheric Sciences
MSFC Counterpart:	O. E. Smith
DATE:	August 26, 1981
Contract No.:	NGT-01-008-021 (University of Alabama)

ACKNOWLEDGEMENTS

I would like to express my gratitude to the NASA/ASEE Summer
Faculty Fellowship Program for providing me an invaluable oppor-
tunity to do research this summer and Dr. G. Karr, Director of
the Program for his able administration and cooperation. My very
special thanks are due to my NASA counterpart, Mr. O. E. Smith,
for suggesting the problem and for his useful discussions. My
profound thanks are due to Dr. F. Lee Cook, Chairman of the Depart-
ment of Mathematics, UAH, for extending me the computer facilities
of the University.

STATISTICAL ANALYSIS OF DAILY MEAN TEMPERATURES

BY

Deva C. Doss, Ph.D.
Professor of Mathematics
University of Alabama in Huntsville
Huntsville, Alabama

ABSTRACT

The present study deals with the data of daily mean temperatures recorded at the Kennedy Center during the period of 1957-1977. The primary objective of the study is to forecast daily mean temperatures and their thirty-day moving averages for a period of ten to fifteen days in a given month. This investigation is of exploratory nature and employs fitting a cubic spline with four knots to the trend.

Since it is found that the standard deviation is linear in the mean, a logarithmic transformation of the data is used for finding an integrated moving average process IMA by the Box-Jenkins approach. The first differences of the transformed data seem to fit a moving average model with parameter value 2, MA(2). The consideration of seasonality factor makes the fit worse.

INTRODUCTION

Daily mean temperatures for the period 1957-1977 collected at Cape Canaveral form the data base for the present study. The daily mean temperature is obtained by averaging hourly temperatures of a day.

1.1 Purpose of the Study The primary objective of this study is to build a forecasting model for daily mean temperatures at Cape Canaveral. This will be used in forecasting daily mean temperatures and thirty-day-running mean temperatures at Kennedy Space Flight Center where the Space Shuttle is scheduled to be launched by the end of September, 1981.

1.2 Methodology A descriptive study of the daily mean temperature is made in order to have a general understanding of the temperature in the time domain by computing descriptive statistics such as mean, standard deviation, maximum, minimum, quartiles and autocorrelations for each day of the year based on 21 observations corresponding to 21 years. The second approach is to reduce the data to a stationary time series and apply to them the recent Box-Jenkins ARIMA techniques explained in Box and Jenkins (1970) and Anderson (1976).

DESCRIPTIVE APPROACH

Since there are 21 observations for each day of the year (except for 4 days for which only 20 observations are available) descriptive statistics such as $\bar{x}$, s.d., quartiles are computed for each day. Their plots are done by using a UNIVAC computer program.

2.1 Mean and Standard Deviation. A plot of mean and s.d. is presented in Figure 1. For the sake of clarity, mean $\bar{x}$ and $\bar{x} \pm 2$ s.d. are plotted every seventh day. One can immediately note two facts.

1. $\bar{x}$ is almost linear for a period that includes four months, i.e., June through September, descends rapidly on either end of this period, and settles in fifties during the period of three months, December through February.

2. Variation of daily mean temperatures as suggested by $\bar{x} \pm 2$ s.d. bands increases as the mean temperature drops. In other words, the Summer range 1.6-3.0 of s.d. doubles during the transitional periods from and to the summer and almost triples in the most of the winter.
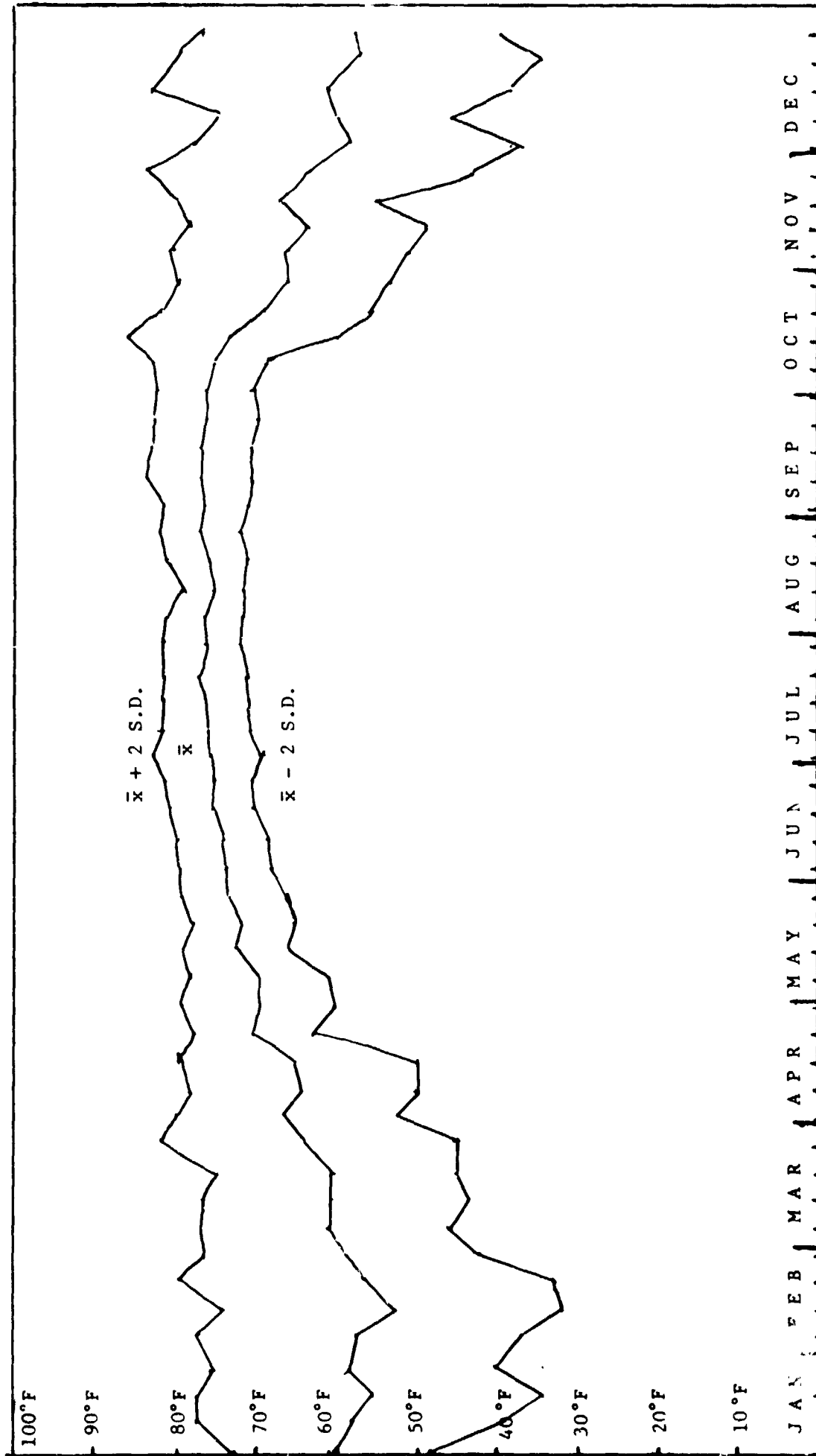
Figure 1. $\bar{x} - 2$ S.D., $\bar{x}$, $\bar{x} + 2$ S.D. of Daily Mean Temperatures

XXIX-5

It follows from (1) that it is futile to fit a curve by harmonic analysis since it will involve at least more than one hundred terms. However, since the plot of $\bar{x}$ in Figure 1 suggests one maximum, one minimum and two points of inflection, a cubic spline with 4 or 5 knots may adequately smoothen the plot of $\bar{x}$.

2.2 **Cubic Spline Fit**   Using an IMSL computer program a cubic spline by the least squares method is fit to the 365 $\bar{x}$'s of daily mean temperatures. The first knot is chosen to be August 8 since a maximum plateau seems to occur on this day. Then the last knot corresponds to August 7. Three knots between the first and the last are selected so that the mean square error for the fit is minimized. This means the cubic spline curve is made of four cubic equations which have continuous second order derivatives at the knots. On examining the Table 1, one may as well extend the domain of the second cubic equation from [114, 219] to [114, 228] without increasing substantially the least square error. In fact, it is worth trying to fit a cubic spline with only three knots, i.e., one cubic equation may essentially cover the winter while the other covers the rest of the year.

Table 1.   Cubic Spline Fit for $\bar{x}$

Knot	c_0	c_1	c_2	c_3
1 (Aug. 8)	74.0036039	.2519037	-.0056720	.0000219
114 (Nov. 30)	61.6039710	-.1919427	.0017446	.00000005
219 (Mar. 15)	60.7353034	.1759009	.0017591	-.0000733
228 (Mar. 2.)	62.4311957	.1897002	-.0002469	-.0000033

Least Square Error = 1.344099

2.3 **Logarithmic Transformation**   The second fact that the s.d. depends upon the mean $\bar{x}$ suggests investigation of regression of s.d. on $\bar{x}$. It is found that s.d. is negatively correlated with $\bar{x}$, i.e. with a high correlation of -0.93. Moreover, the regression line of s.d. s on $\bar{x}$ is given by

$$s = .3704(83.533 - \bar{x}) . \tag{1}$$

It is a well-known fact that if s.d. is linear in $\bar{x}$, a logarithmic transformation of data points is considered. Incidentally, it is interesting to note that the maximum daily mean temperature for 21 years is 83°F. Throughout the study only the logarithmic transformation

$$y = \log(84-x) \tag{2}$$

is considered

The quartiles of the transformed variable are plotted in Figure 2. It appears that the range (maximum-minimum) is nearly uniform indicating the logarithmic transformation has a constant variance in the time domain. Once again, cubic splines with 3 or 4 knots can be fit for the quartiles

? 4 <u>Autocorrelations</u> The computed autocorrelations for the raw and transformed data for each day indicate that the autocorrelations at lag k, $k \geq 5$ are not significant. The autocorrelations at lag 3 and 4 are not substantially different from zero. From practical point of view, it appears the temperature of a day is dependent upon temperatures of the preceding two or three days.

THE BOX-JENKINS METHOD

Box and Jenkins (1970) have developed in analyzing a stationary time series Auto-Regressive Integrated Moving Average (ARIMA) models. Indentification, estimation. verification and forecasting are the procedures to be followed in selecting a model (see Anderson (1976)).

3.1 <u>Identification</u> For identification purposes, autocorrelations (strictly speaking, serial correlations) and partial autocorrelations at lag k, $k \leq 30$ of daily mean temperatures are calculated. Given a stationary series $x_1, x_2, \cdots, x_N$ one can estimate mean, variance and autocorrelations respectively by

$$\bar{x} = \sum_{i=1}^{N} x_i/N, \quad c_0 = \sum_{i=1}^{N} (x_i - \bar{x})^2/N$$

$$r_k = \sum_{i=k+1}^{N} (x_i - \bar{x})(x_{i-k} - \bar{x})/Nc_0 . \tag{3.1}$$

The estimated partial autocorrelation at log k is

$$\hat{\phi}_{kk} = |P_k^*|/|P_k| \tag{3.2}$$

where P_k is the estimated $k \times k$ autocorrelation matrix $[a_{ij}]$, $a_{ij} = r_{i-j}$ and P_k^* is the matrix obtained from P_k by replacing the last column with $[r_1, r_2, \cdots, r_k]'$. Observe $r_{-k} = r_k$.

Since the number of observations is very large, in fact, $N = 7699$, one can use Bartlett's formula (see Bartlett (1946)) for testing whether autocorrelations are significantly different from zero for a stationary process,

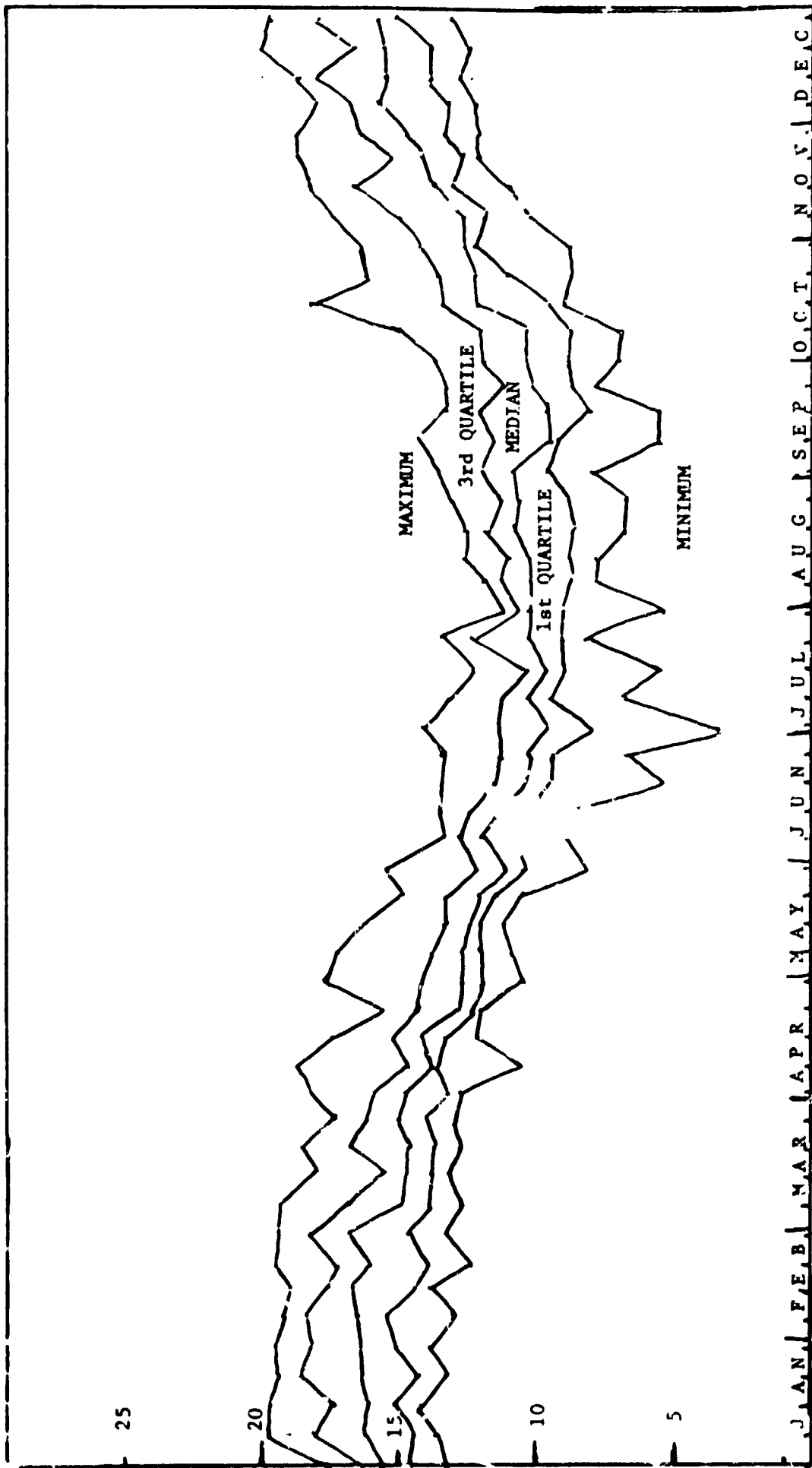$$Var[r_k] = \sum_{-K}^{K} \rho_i^2/N = (1 + 2 \sum_{1}^{K} r_i^2)/N \tag{3.3}$$

Figure 2. Quartiles of Transformed Variable

XXIX-8

and Quenouille's formula (see Quenouille (1949)) for testing significant partial autocorrelations if any,

$$\text{Var}(\hat{\phi}_{kk}) \simeq 1/N . \tag{3.4}$$

Moreover, autocorrelations and partial autocorrelations of a stationary processes are asymptotically distributed normal.

In Table 2 all autocorrelations and partial autocorrelations (under zero order of differencing) of transformed series are all significant and taper very slowly. This suggests that the data need to be differenced at least once. That is, instead of considering a series $\{x_i, i=1, \cdots, N\}$, one considers a differenced series $\{y_i, i=2, \cdots, N\}$ where

$$y_i = x_i - x_{i-1} \tag{3.5}$$

The transformed series is differenced once, twice, and three times and its autocorrelations and serial autocorrelations are presented in Table 2. The computed autocorrelations at lag k, $k \leq 4$ are only significant whereas the partial autocorrelations, although gradually decreasing, are significant. Therefore the plausible model is a moving average model MA(q) with order $q, 1 \leq q \leq 4$. The differencing has cut down the variance by two-thirds. Further differencing has increased the variances and partial autocorrelations. By the principle parsimony, a MA(q) model should be fit to the first differences instead of second or third differences.

3.2 **Estimation**   The moving average model of q, the MA(q) process is defined as

$$z_i = a_i + \theta_1 a_{i-1} + \cdots + \theta_q a_{i-q} \tag{3.6}$$

where a_i are normally, independently and identically distributed with zero mean and constant variance σ_a^2. For the MA(q) process.

$$\sigma_a^2 = \text{Var}(z)/(1 + \theta_1^2 + \cdots + \theta_q^2) \tag{3.7}$$

$$\rho_k = 0, \; k > q . \tag{3.8}$$

The last equation is used for identifying a MA process.

The moving average models MA(q), $q = 1,2,3$ are fit to the once differenced transformed series using a maximum likelihood estimation algorithm. The maximum likelihood estimators of the parameters are presented in Table 3. The selection of order q is based on successive significant reduction in variance of the series. Then the choice seems to be MA(2).

Table 2. $\bar{x}, s^2, r_k, \phi_{kk}, k \leq 12$, for Undifferenced and Differenced Data

		ORDER OF DIFFERENCING			
		0	1	2	3
	Mean	2.6215	-.0001	-.0001	-.0001
	Variance	.3598	.1237	.3119	.9651
Autocorrelations at Lag	1	.8280	-.2610*	-.5471*	-.6795*
	2	.7457	-.1419*	.0085	.1717*
	3	.7121	-.0441*	.0328*	.0185
	4	.6937	-.0286*	.0001	-.0091
	5	.6851	-.0137	-.0045	-.0096
	6	.6814	.0124	.0205	.0189
	7	.6733	-.0131	-.0132	-.0141
	8	.6697	-.0047	-.0032	-.0007
	9	.6676	.0116	.0097	.0048
	10	.6618	.0025	.0069	.0117
	11	.6551	-.0231	-.0319*	-.0361*
	12	.6563	.0320	.0412*	.0462*
Partial Autocorrelations at Lag	1	.8280	-.2610	-.5471	-.6795
	2	.1910	-.2254	-.4151	-.5390
	3	.1797	-.1486	-.3239	-.4405
	4	.1351	-.1296	-.?730	-.3674
	5	.1244	-.0927	-.2523	-.3401
	6	.1107	-.0968	-.2049	-.2956
	7	.0770	-.0865	-.1846	-.2561
	8	.0833	-.0630	-.1786	-.2376
	9	.0747	-.0541	-.1608	-.2343
	10	.0534	-.0748	-.1225	-.1748
	11	.0452	-.0250	-.1527	-.1925
	12	.0669	-.0489	-.1098	-.1590
Significant at .05		ALL	Starred Autocorrelations And All Partial Autocorrelations		

Table 3. M.L.E. Estimation of MA(q), q = 1,2,3

q	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\sigma}_a^2$	% Reduction in Var(z)
1	-.5622			.1074	13.2
2	-.4791	-.2898		.0991	19.9
3	-.4564	-.2371	-.0987	.0981	20.7

3.3 **Verification** A verification of a model fit is based on the analysis of the residuals. For a moving average process one has to analyze the estimated series $\{\hat{a}_i\}$ and see if it has not satisfied the independence and normality conditions. The independence condition can be checked by considering autocorrelations of $\{\hat{a}_i\}$ and the normality by a low-powered portmanteau lack-of-fit test given by Box and Pierce (1970). The statistic

$$R = N \sum_1^K r_k^2(\hat{a}) \qquad (3.9)$$

is calculated, where K is 30. Then should the fitted model be appropriate,

$$R \simeq \chi_{K-q}^2 \qquad (3.10)$$

and significant R indicates model inadequacy.

For N = 1000, a fit of MA(2) model is made and the largest auto-correclation is found to be .0605 when signs are ignored. Moreover, the value of R = 27.8888 which lies below the mean of χ_{28}^2. Therefore, the second order MA model seems to be an adequate model.

If the analysis of $\{\hat{a}_i\}$ is based on most of the data N = 7630, only one of the autocorrelations equals 0.078 and four of them exceed numerically .03 but are less than .042. All others lie in the interval (-.03,.03). The R value for K = 30 is 137.8956 and for K = 1000 is 1560.6238. The increase in R value is contributed by the largeness of N. Although R is significant in both cases, it should be noted that in reality one cannot realize a perfect process with zero autocorrelations. For all practical purposes with N as large as 7630 and autocorrelations numerically less than .1, the process should be considered stationary and adequate.

3.4 **Forecasting** Using an IMSL routine and the MA(2) model on the first differences of the transformed data one can statistically forecast the daily mean temperatures and provide 95% confidence intervals. For instance, Table 4 presents forecasts for 10 days, November 23-December 2; ther

forecasts are based on preceding 1000 observations. How good are they?

Table 4. 95% Confidence Interval for Forecasts with Lead of
10 Days, Based on Preceding 1000 Observations.

Lower Bound	Forecast	Upper Bound	Actual Value
58	70	76	71
53	69	76	66
53	69	76	62
52	69	77	64
52	69	77	40
51	69	77	55
51	69	77	73
51	69	77	72
50	69	77	73
50	69	77	69

Except one all observations are found in the confidence belt. However,
one may argue that the width of the confidence belt is so large that it
cannot fail, but contain the observations. The maxima for these ten days
from the 21 year observations are 75-77 which coincide with the upper
bounds of the forecasts and the minima lie in forties whereas the lower
bounds are in fifties. The forecasting model does reduce the length of
confidence interval.

3.6 Seasonality Factor The above sequence of procedures of identifica-
tion, estimation and verification is done with seasonality factor. It is
surprising to discover that no seasonal model fits better than the MA(2).
In fact, a seasonal model in the ARIMA class does not satisfy the criteria
for fitness of the data.

Finally, the running mean follows the trend of the daily mean tem-
perature, but possesses a smaller variation.

CONCLUSIONS

The daily mean temperature follows an integrated moving average model
IMA(1,2) after logarithmic transformation. The trend of the series is
best described by a cubic spline with, at most, four knots. It should be
borne in mind that a forecast, however sophisticated it may be, is itself
an educated guess and may not prove to be better than a forecast arrived
at by the shadow of a ground hog.

REFERENCES

1. Anderson, O. D., *Time Series Analysis and Forecasting, The Box-Jenkins Approach*, Butterworths, London, 1976.

2. Bartlett, M. S., "On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series.", *J. R. Statist. Soc.*, B 8, 27-41, 1946.

3. Box, G.E.P. and Jenkins, G. M., *Time Series Analysis Forecasting and Control*, Holden-Day, San Francisco, 1970.

4. Box, G.E.P. and Pierce, D. A., "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," *J. Am. Statist. Ass.*, 65, 1509-1526, 1970.

5. Quenuille, M. H., "Approximate Tests of Correlation in Time Series", *J. R. Statist. Soc.*, B 11, 68-84, 1949.

1981

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

SURFACE CONTROL TEMPERATURES FOR THE BRIDGMAN-STOCKBARGER
AND FLOAT ZONE TECHNIQUES

Prepared By: Larry M. Foster, Ph.D.

Academic Rank: Assistant Professor

University and Department   University of Alabama
                                       in Huntsville
                                      Mathematics Department

NASA/MSFC:
  (Laboratory)   Space Sciences
  (Division)   Space Processing
  (Branch)   Semiconductors and Devices

MSFC Counterpart   R. J. Naumann

Date: August 21, 1981

XXX

SURFACE CONTROL TEMPERATURES FOR THE
BRIDGMAN-STOCKBARGER AND FLOAT-ZONE TECHNIQUES

By

Larry M. Foster, rn.D.
Assistant Professor of Mathematics
University of Alabama in Huntsville
Huntsville, Alabama

ABSTRACT

The surface control temperature required to properly pose the equations of state for the Bridgman-Stockbarger and float-zone techniques will be investigated. In particular, let

(1) $\quad \Delta T = P \frac{\partial T}{\partial x}$, $\quad x < 0$ and $0 < r < 1$

(2) $\quad T(0,r) = A(r), \quad 0 < r < 1$

(3) $\quad \frac{\partial T}{\partial x}(0,r) = B(r), \quad 0 < r < 1$

and

(4) $\quad T(x,1) = f(x), \quad x < 0$

where $A(r)$ and $B(r)$ are given and T is the material temperature. If $f(x)$ is arbitrary, then (1)-(4) are over-posed and usually have no solution. We will ᴉ vestigate the properties the surface control temperature $f(x)$ must possess to properly pose (1)-(4).

NASA/MSFC
   (Laboratory)   Space Sciences
   (Division)   Space Processing
   (Branch)   Semiconductors and Devices

MSFC Counterpart: R. J. Naumann

ACKNOWLEDGEMENTS

1. INTRODUCTION

The scaled partial differential equations defining crystal growth by the Bridgman-Stockbarger technique [1] are

$$(1) \quad \nabla^2 T = P_\ell \frac{\partial T}{\partial x} \quad , \quad x < 0 \, , \, 0 < r < 1$$

and

$$(2) \quad \nabla^2 T = P_s \frac{\partial T}{\partial x} \quad , \quad x > 0 \, , \, 0 < r < 1$$

where P_ℓ and P_s are the liquid and solid Péclèt numbers respectively, T is the working material temperature and x is the axial distance from an assumed flat solid-melt interface. To conserve energy at the interface,

$$(3) \quad -K_\ell \frac{\partial T}{\partial x} \Big)_{x=0^-} + K_s \frac{\partial T}{\partial x} \Big)_{x=0^-} + \mathcal{L} = 0$$

where K_ℓ and K_s are the respective liquid and solid conductivities and $\mathcal{L}$ is the product of the crystal growth rate, the melt density, and the latent heat of solidification. The boundary heating and cooling mechanisms in the idealized case are

$$(4) \quad -K_\ell \frac{\partial T}{\partial r} = h_\ell \left(T - T_\ell \right) \quad , \quad r = 1, \quad x < 0$$

and

$$(5) \quad -K_s \frac{\partial T}{\partial r} = h_s T \quad , \quad r = 1, \quad x > 0$$

where the ambient temperature for the solid region, x > 0, has been scaled to zero and h_ℓ and h_s are the liquid and solid heat transfer coefficients at crucible surface, r = 1. Equations (1)-(5) are mathematically properly posed and may be solved by the separation of variables method. However, if M is the working material's melting temperature and we adjoin to equation (1)-(5)

$$(6) \quad T \Big)_{x=0^-} = M = T \Big)_{x=0^+} \quad , \quad 0 < r < 1$$

the system (1)-(6) is overposed. If the system (1)-(6) has a solution,

then the process parameters P_ℓ, P_s, K_ℓ, K_s, χ, h_ℓ, h_s, T and M are necessarily functionally dependent [2]. However, in general, this necessary functional dependence is not sufficient to guarantee (1)-(6) has a solution. The problem is further complicated by the addition of a more realistic adiabatic zone from x = -Q to x = L containing the melt (x = 0) interface. The resulting system is then

(7) $\quad \nabla^2 T_1 = P_s \dfrac{\partial T_1}{\partial x}$, $x > L$, $0 < r < 1$

(8) $\quad \nabla^2 T_2 = P_s \dfrac{\partial T_2}{\partial x}$, $0 < x < L$, $0 < r < 1$

(9) $\quad \nabla^2 T_3 = P_\ell \dfrac{\partial T_3}{\partial x}$, $-Q < x < 0$, $0 < r < 1$

(10) $\quad \nabla^2 T = P_\ell \dfrac{\partial T}{\partial x}$, $x < -Q$, $0 < r < 1$

(11) $\quad -K_s \dfrac{\partial T_1}{\partial r} = h_s T_1$, $x > L$, $r = 1$

(12) $\quad \dfrac{\partial T_2}{\partial r} = 0$, $0 < x < L$, $r = 1$

(13) $\quad \dfrac{\partial T_3}{\partial r} = 0$, $-Q < x < 0$, $r = 1$

(14) $\quad -K_\ell \dfrac{\partial T}{\partial r} = h_\ell (T - T_\ell)$, $x < -Q$, $r = 1$

(15) $\quad T_1 = T_2$, $x = L$, $0 < r < 1$

(16) $\quad T_2 = M = T_3$, $x = 0$, $0 < r < 1$

(17) $\quad T_3 = T$, $x = -Q$, $0 < r < 1$

(18) $\quad \dfrac{\partial T_1}{\partial x} = \dfrac{\partial T_2}{\partial x}$, $x = L$, $0 < r < 1$

(19) $\quad \dfrac{\partial T}{\partial x} = \dfrac{\partial T_3}{\partial x}$, $x = -Q$, $0 < r < 1$

and

(20) $\quad -K_\ell \dfrac{\partial T_3}{\partial x} + K_s \dfrac{\partial T_2}{\partial x} + \mathscr{L} = 0$, $x = 0$, $0 < r < 1$

Let System I include equations (7)-(9), (11)-(13), (15), (16), (18) and (20). Then System I is properly posed if we include the realistic conditions

(21) $\quad \dfrac{\partial T_i}{\partial r}\bigg)_{r=0^+} = 0 \quad , \quad i = 1, 2 \text{ and } 3$

and

(22) $\quad \sup\limits_{0 < r < 1} \lim\limits_{x \to \infty} |T_1(x,r)| < \infty$

However, if (14) and (19) were adjoined to System I, the resulting system would be overposed and generally insolvable. To circumvent this dilemma, we propose the following. First solve System I and let

(23) $\quad A(r) = T_3(x,r) \, , \quad x = -Q \, , \, 0 < r < 1$

and

(24) $\quad B(r) = \dfrac{\partial T_3}{\partial x} \, , \quad x = -Q \, , \, 0 < r < 1$

Then replace condition (14) by

(25) $\quad T(x,r) = f(x) \, , \quad x < -Q \, , \, r = 1$

where

$\quad f(-Q) = A(1) \text{ and } f'(-Q) = B(1) \, ,$

but f(x) is otherwise unknown. The control f(x) is constructed such that the solution of the well posed System II:

(26) $\quad \nabla^2 T = P_1 \dfrac{\partial T}{\partial x} \, , \quad x < -Q \, , \, 0 < r < 1$

(27) $\quad T(x,r) = A(r) \, , \quad x = -Q \, , \, 0 < r < 1$

(28) $\quad T(x,r) = f(x) \, , \quad x < -Q \, , \, r = 1$

also satisfies

(29) $\quad \dfrac{\partial T}{\partial x}(x,r) = \beta(r) \; , \quad x = -Q, \; 0 < r < 1$

In other words, we wish to construct a heating control $f(x)$ such that if (14) is replaced by (28), the system (7)-(13), (15)-(20) and (28) has a solution. If, in addition, a Newton heating control $g(x)$ is desired, i.e.,

(30) $\quad -K_{\ell}\dfrac{\partial T}{\partial r} = h_{\ell}(T - g(x)) \; , \quad x < -Q \; , \quad r = 1$

instead of an infinite sink heating control $f(x)$, then we only need solve

$$-K_{\ell}\dfrac{\partial T}{\partial r} = h_{\ell}(f(x) - g(x)) \; , \quad x < -Q \; , \quad r = 1$$

for $g(x)$ since $f(x)$ and T (and hence $\dfrac{\partial T}{\partial r}$) have already been found. In Section 2 of this paper, the control $f(x)$ is constructed such that (26), (27), (28) and (29) has a solution. The solution of System I is constructed in Section 3 and coupled with the results of Section 2. A numerical case study is given in Section 4 followed by various remarks and conclusions in Section 5.

2. CONSTRUCTION OF THE CONTROL SURFACE

If $f(x)$ is the desired control such that (26)-(29) has a solution, then define $\Theta(x,r)$ by

$$T(x,r) = \Theta(x,r) + f(x), \; x < -Q \text{ and } 0 < r < 1$$

If we define $G(x) = Pf' - f''$, $A(r) - f(-Q) = \alpha(r)$ and $B(r) - f'(-Q) = \beta(r)$, then

(31) $\quad \Theta_{xx} + \Theta_{rr} + \dfrac{1}{r}\Theta_r = P_{\ell}\Theta_x + G(x) \; , \quad x < -Q, \; 0 < r < 1$

(32) $\quad \Theta(x,r) = \alpha(r) \; , \quad x = -Q, \; 0 < r < 1$

(33) $\quad \Theta_x(x,r) = \beta(r) \; , \quad x = -Q, \; 0 < r < 1$

and

(34) $\quad \theta(x,r) = 0 \quad , \quad x < -Q, \quad 0 < r < 1$

Although the internal heat generation term $G(x)$ in (31) complicates the partial differential equation, the reduction of the nonhomogeneous boundary condition (28) to that of a homogenous one (34) greatly simplifies the solution process. To solve (31), we first solve the Helmhotz eigenvalue problem:

$$(35) \quad \begin{cases} \nabla^2 \psi(r) + \lambda^2 \psi(r) = 0 \quad , \quad 0 < r < 1 \\ \psi(1) = 0 \\ \psi'(0) = 0 \end{cases}$$

The solutions of (35) are the Bessel functions

$$(36) \quad \psi_m(r) = J_0(\lambda_m r) \, , \quad m = 1, 2, \cdots$$

where the eigenvalues of (35) are the real positive roots of

$$(37) \quad J_0(\lambda_m) = 0$$

Moreover,

$$(38) \quad \int_0^1 \psi_N(r) \psi_M(r) r \, dr = \begin{cases} 0 \quad \text{if } N \neq M \\ \frac{1}{2} J_1^2(\lambda_M) \quad \text{if } M = N \end{cases}$$

if we let

$$(39) \quad \theta(r,x) = \sum_{m=1}^{\infty} C_m(x) \psi_m(r)$$

then (38) implies

$$(40) \quad \int_0^1 \theta(x,r) \psi_M(r) r \, dr = C_M(x) \frac{J_1^2(\lambda_M)}{2}$$

Equations (39) and (40 form the dual integral transform pair of $\theta(x,r)$:

$$(41) \quad \begin{cases} \Theta(x, r) = \sum_{M=1}^{\infty} \frac{2 \psi_M(r)}{J_1^2(\lambda_M)} \bar{\Theta}_M(x) \\ \\ \bar{\Theta}_M(x) - \int_0^1 \Theta(r, x) \psi_M(r) r \, dr \end{cases}$$

If we define

$$(42) \quad \bar{G}_M(x) = \int_0^1 G(x) \psi_M(r) r \, dr$$

multiply (31) by $\psi_m(r) r \, dr$, integrate each side of the resulting equation from $r = 0$ to $r = 1$ and apply Green's Theorem, then

$$(43) \quad -\lambda_M^2 \bar{\Theta}_M + \bar{\Theta}_M'' = P_\ell \bar{\Theta}_M' + \bar{G}_M$$

Since

$$(44) \quad \bar{\Theta}_M(-Q) = \int_0^1 Q(r) \psi_M(r) r \, dr$$

and

$$(45) \quad \bar{\Theta}_M'(-Q) = \int_0^1 B(r) \psi_M(r) r \, dr$$

we need only solve the initial value problem (43)-(45) to compute $\bar{\Theta}_m(x)$. Because $Q(1) = B(1) = 0$, we may expand $Q(r)$ and $B(r)$ as

$$(46) \quad Q(r) = \sum_{m=1}^{\infty} Q_m J_0(\lambda_m r)$$

and

$$(47) \quad B(r) = \sum_{m=1}^{\infty} B_m J_0(\lambda_m r)$$

where

$$(48) \quad Q_m = \frac{2}{J_1^2(\lambda_m)} \int_0^1 J_0(\lambda_m r) Q(r) r \, dr$$

C - 7

and

$$(49) \quad \mathcal{B}_m = \frac{2}{J_1^2(\lambda_m)} \int_0^1 J_0(\lambda_m r) \mathcal{B}(r) \, r \, dr$$

Combining (38), (44)-(47),

$$(50) \quad \bar{\Theta}_m(-Q) = \alpha_m \frac{J_1^2(\lambda_m)}{2}$$

and

$$(51) \quad \bar{\Theta}_m'(-Q) = \mathcal{B}_m \frac{J_1^2(\lambda_m)}{2}$$

For notational convenience, let $S_m = \sqrt{P_\lambda^2 + 4\lambda_m^2}$. Then the solution of (43), (50) and (51) is

$$(52) \quad \bar{\Theta}_m(x) = \frac{J_1^2(\lambda_m)}{2 S_m} \left(\mathcal{B}_m - \frac{P_\lambda - S_m}{2} \alpha_m \right) \text{Exp}\left(\frac{P_\lambda + S_m}{2}(x+Q) \right)$$

$$+ e^{\frac{P_\lambda + S_m}{2}x} \int_{-Q}^x \frac{\bar{G}_m(t)}{S_m} e^{-\frac{P_\lambda + S_m}{2}t} dt$$

$$+ \frac{J_1^2(\lambda_m)}{2 S_m} \left(\frac{P_\lambda + S_m}{2} \alpha_m - \mathcal{B}_m \right) e^{\frac{P_\lambda - S_m}{2}(x+Q)}$$

$$- \text{Exp}\left(\frac{P_\lambda - S_m}{2}x \right) \int_{-Q}^x \frac{\bar{G}_m(t)}{S_m} e^{-\frac{P_\lambda - S_m}{2}t} dt$$

Since we expect $\lim_{x \to -\infty} T(x,r) = f(x)$, we require $\lim_{x \to -\infty} \bar{\Theta}_m(x) = 0$. Hence, if we assume $G(x) = Pf' - f'' \to 0$ as $x \to -\infty$ (a realistic assumption), then (52) implies we must also require

$$(53) \quad \lim_{x \to -\infty} \left[\frac{J_1^2(\lambda_m)}{2 S_m} \left(\frac{P_\lambda + S_m}{2} \alpha_m - \mathcal{B}_m \right) \text{Exp}\left(\frac{P_\lambda - S_m}{2} Q \right) \right.$$

$$\left. - \int_{-Q}^x \frac{\bar{G}_m(t)}{S_m} e^{-\frac{P_\lambda - S_m}{2}t} dt \right] e^{\frac{P_\lambda - S_m}{2}x} = 0$$

For notational convenience, let

(54) $\quad G(x - Q) = E(-x), \quad x < 0$

Then (53) implies

(55)
$$\frac{\lambda_m J_1(\lambda_m)}{2}\left(\beta_m - \frac{P_1 + S_m}{2} a_m\right)$$
$$= \int_0^\infty E(t) e^{-\frac{S_m - P_1}{2} t} dt$$

Because we assume $G(x) \to 0$ as $x \to -\infty$, we expand $E(t) = \sum_{k=1}^\infty a_k e^{-kt}$.
However, for computational purposes, we approximate $E(t)$ by $\sum_{k=1}^N a_k e^{-kt}$
and then solve the approximate identity

(56)
$$\frac{\lambda_m J_1(\lambda_m)}{2}\left(\beta_m - \frac{P_1 + S_m}{2} a_m\right)$$
$$= \int_0^\infty \sum_{k=1}^N a_k e^{-\left(\frac{S_m - P_1}{2} + k\right)t} dt$$

for $a_1, a_2, \cdots, a_N$. To approximate $f(x)$, we need only then solve

(57)
$$\begin{cases} \sum_{k=1}^N a_k e^{-kt} = -P g'(t) - g''(t), \quad t > 0 \\ g(0) = A(1) \\ g'(0) = -B(1) \end{cases}$$

and set

(58) $\quad f(t) = g(-Q - t), \quad t < 0$

(57) is easily solved by

(59)
$$g(t) = A + B e^{-P_1 t} + \sum_{k=1}^N \left[\frac{1}{k} - \frac{1}{k - P_1}\right] \frac{a_k}{P_1} e^{-kt}$$

where

$$\begin{pmatrix} 1 & 1 \\ 0 & P_1 \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} A(1) \\ B(1) \end{pmatrix} - \sum_{k=1}^N \left[\frac{1}{k} - \frac{1}{k - P_1}\right]\frac{a_k}{P_1}\begin{pmatrix} 1 \\ k \end{pmatrix}$$

3. SOLUTION OF SYSTEM I

To complete the method described in Section 2, the functions $a(r)$ and $\beta(r)$ must be known. Because $T_3(x,r) = a(r)$ and $\frac{\partial T_3}{\partial x}(x,r) = \beta(r)$ at $x = -Q$, we need only solve System I (equations (7)-(9), (11)-(13), (15), (16), (18), and (20)-(22)). Assume $h_s \neq 0$ and let β_n and γ_n be the real roots of

$$(60) \quad \begin{cases} K_s \beta_m J_1(\beta_m) = h_s J_0(\beta_m) \\ \text{and} \\ J_0(\gamma_m) = 0 \quad \text{with} \quad \gamma_i > 0 \end{cases}$$

Separating variables, the solutions of (7)-(9) are

$$(61) \quad T_1(x,r) = \sum_{m=1}^{\infty} A_m J_0(\beta_m r) \, Exp\left(\frac{P_s - \sqrt{P_s^2 + 4\beta_m^2}}{2} x\right)$$

$$(62) \quad T_2(x,r) = \sum_{m=1}^{\infty} J_0(\lambda_m r)\left[B_m e^{\frac{P_s - \sqrt{P_s^2 + 4\gamma_m^2}}{2} x} - C_m e^{\frac{P_s + \sqrt{P_s^2 + 4\gamma_m^2}}{2} x}\right]$$

and

$$(63) \quad T_3(x,r) = \sum_{m=1}^{\infty} J_0(\lambda_m r)\left[D_m e^{\frac{P_\ell - \sqrt{P_\ell^2 + 4\gamma_m^2}}{2} x} - E_m e^{\frac{P_\ell + \sqrt{P_\ell^2 + 4\gamma_m^2}}{2} x}\right]$$

Combining (62) and (63) with (16),

$$(64) \quad \begin{cases} B_1 + C_1 = M = D_1 + E_1 \\ \text{and} \\ B_m + C_m = 0 = D_m + E_m, \quad m = 2, 3 \cdots \end{cases}$$

and hence (20) coupled with the partial derivatives of (62) and (63) with respect to x imply (note $\lambda_1 = 0$)

$$(65) \quad \begin{cases} K_\ell E_1 P_\ell - K_s C_1 P_s = \mathcal{L} \\ -K_\ell D_m \sqrt{P_\ell^2 + 4\gamma_m^2} + K_s B_m \sqrt{P_s^2 + 4\gamma_m^2} = 0, \quad m = 2, 3, \cdots \end{cases}$$

Since

$$\int_0^1 J_0^2(\gamma_N r) \, r \, dr = \frac{1}{2} J_1^2(\gamma_N), \quad N = 1, 2, \cdots$$

and

$$\int_0^1 J_0(\beta_m r) J_0(\gamma_N r)\, r\, dr = \frac{\beta_m J_0(\gamma_N) J_1(\beta_m)}{\beta_m^2 - \gamma_N^2}$$

if (61) and (62) are substituted 'n (15) and (18), then, after multiplying each of the resulting equations by $J_0(\gamma_N r) r\, dr$ and integrating from r = 0 to r = i,

(66)
$$\sum_{m=1}^{\infty} A_m \frac{\beta_m J_1(\beta_m)}{\beta_m^2 - \gamma_N^2} \operatorname{Exp}\left(\frac{P_s - \sqrt{P_s^2 + 4\beta_m^2}}{2} L \right)$$

$$= \begin{cases} (B_1 + C_1 e^{P_s L})/2 & , \ N=1 \\ B_N \left(e^{\frac{P_s - \sqrt{P_s^2 + 4\gamma_N^2}}{2} L} - e^{\frac{P_s + \sqrt{P_s^2 + 4\gamma_N^2}}{2} L} \right) \frac{J_0(\gamma_N)}{2} & , \ N = 2, 3, \cdots \end{cases}$$

and

(67)
$$\sum_{m=1}^{\infty} A_m \frac{P_s - \sqrt{P_s^2 + 4\beta_m^2}}{2} \cdot \frac{\beta_m J_1(\beta_m)}{\beta_m^2 - \gamma_N^2} e^{\frac{P_s - \sqrt{P_s^2 + 4\beta_m^2}}{2} L}$$

$$= \begin{cases} C_1 P_s e^{P_s L}/2 & , \ N=1 \\ B_N \left[\frac{P_s - \sqrt{P_s^2 + 4\gamma_N^2}}{2} \operatorname{Exp}\left(\frac{P_s - \sqrt{P_s^2 + 4\gamma_N^2}}{2} L \right) - \frac{P_s + \sqrt{P_s^2 + 4\gamma_N^2}}{2} \operatorname{Exp}\left(\frac{P_s + \sqrt{P_s^2 + 4\gamma_N^2}}{2} L \right) \right] \frac{J_0(\gamma_N)}{2} & , \\ & N = 2, 3, \cdots \end{cases}$$

clearly, if we knew $A_1, A_2, \cdots$, then we could compute all the necessary coefficients for (62) and (63) by merely invoking (64)-(66). However, for computational purposes, we will only approximate $A_1, A_2, \cdots, A_K$.

If, for N = 1 (note (64)), we combine (66) and (67), eleminate B_1 and C_1 and compute only the first K terms of the resulting infinite series, we produce the approximate identity

(68)
$$\sum_{m=1}^{\infty} A_m \left[\frac{2}{1 - e^{P_s L}} - \frac{2}{P_s e^{P_s L}} \frac{P_s - \sqrt{P_s^2 + 4\beta_m^2}}{2} \right] \frac{J_1(\beta_m)}{\beta_m} e^{\frac{P_s - \sqrt{P_s^2 + 4\beta_m^2}}{2} L} = \frac{M}{1 - e^{P_s L}}$$

Similarly, for N = 2, $\cdots$, K, we construct the approximate identity

$$(69) \quad \sum_{m=1}^{K} A_m \frac{\beta_m J_1(\beta_m)}{\beta_m^2 - \gamma_N^2} \left[\frac{1}{e^{\frac{P_s - \sqrt{P_s^2 + 4\gamma_N^2}}{2}L} - e^{\frac{P_s + \sqrt{P_s^2 + 4\gamma_N^2}}{2}L}} \right.$$

$$\left. - \frac{P_s - \sqrt{P_s^2 + 4\beta_m^2}}{2\left(\frac{P_s - \sqrt{P_s^2 + 4\gamma_N^2}}{2} Exp\left(\frac{P_s - \sqrt{P_s^2 + 4\gamma_N^2}}{2}L \right) - \frac{P_s + \sqrt{P_s^2 + 4\gamma_N^2}}{2} Exp\left(\frac{P_s + \sqrt{P_s^2 + 4\gamma_N^2}}{2}L \right) \right)} e^{\frac{P_s - \sqrt{P_s^2 + 4\beta_m^2}}{2}L} \right]$$

$$= 0, \quad N = 2, 3, \cdots, K$$

Solving (68) and (69) for $A_1, \cdots, A_K$, we may then approximate $B_1, \cdots, B_K$, $C_1, \cdots, C_K, D_1, \cdots, D_K, E_1, \cdots$, and E_K as outlined above. The two functions, $\alpha(r)$ and $\beta(r)$, required in Section 2 are then computationally approximated by

$$(70) \quad \begin{cases} \tilde{\alpha}(r) = \sum_{m=1}^{K} J_0(\gamma_m r) \left[D_m e^{-\frac{P_\ell - \sqrt{P_\ell^2 + 4\gamma_m^2}}{2}Q} + E_m e^{-\frac{P_\ell + \sqrt{P_\ell^2 + 4\gamma_m^2}}{2}Q} \right] \\[2ex] \tilde{\beta}(r) = \sum_{m=1}^{K} J_0(\gamma_m r) \left[D_m \frac{P_\ell - \sqrt{P_\ell^2 + 4\gamma_m^2}}{2} Exp\left(-\frac{P_\ell - \sqrt{P_\ell^2 + 4\gamma_m^2}}{2}Q \right) \right. \\[2ex] \qquad\qquad \left. + E_m \frac{P_\ell + \sqrt{P_\ell^2 + 4\gamma_m^2}}{2} Exp\left(-\frac{P_\ell + \sqrt{P_\ell^2 + 4\gamma_m^2}}{2}Q \right) \right] \end{cases}$$

4. NUMERICAL CASE STUDY

In this section we present a numerical case study of the methods described in Sections 2 and 3. Let P_Q = .1, P_s = .2, h_s = K_s = K_ℓ = L = Q = $\mathcal{L}$ = 1.0, and M = 10.0 . First, the set of coefficients $A_1, \cdots, A_K$ for K = 2, $\cdots$, 7 are computed as outlined in Section 3 and are partially displayed in Table 1

	K = 2	K = 4	K = 6	K = 7
A_1	.1869+02	.1863+02	.1862+02	.1862+02
A_2	-.4982+02	-.49829+02	-.4814+02	-.4811+02
A_3		.4151+03	.4115+03	.4106+03
A_4		-.5288+04	-.5162+04	-.5147+04
A_5			.7723+05	.7681+05
A_6			-.1284+07	-.1267+07
A_7				.2241+08

Table 1

The size of K appears to have little affect on $A_1, \cdots, A_K$. Also, since the terms in (61) have rapidly decaying exponential factors, the increasing magnitudes of the A_n's is to be expected. For K = 7, the various B_n, C_n, D_n and E_n are given in Table 2.

n	B_n	C_n	D_n	E_n
1	.6146+02	-.5146+02	.3073+02	-.2073+02
2	-.1017-01	.1017-01	-.1016-01	.1016-01
3	.1928-03	-.1928-03	.1927-03	-.1927-03
4	-.4959-05	.4959-05	-.4960-05	.4959-05
5	.1463-06	-.1463-06	.1463-06	-.1463-06
6	-.4670-08	.4670-08	-.4700-08	.4670-08
7	.1570-09	-1570-09	.1570-09	-.1570-09

Table 2

To avoid numerical instabilities, we approximate the coefficients $\alpha_1, \cdots, \alpha_5, \beta_1, \cdots$ and β_5 in (46) and (47) by solving the overposed systems of linear equations

$$\sum_{m=1}^{5} \alpha_m J_o(\lambda_m r_k) = \tilde{\alpha}(r_k) \ , \ k = 1, \cdots, 10$$

and

$$\sum_{m=1}^{5} \beta_m J_o(\lambda_m r_k) = \tilde{\beta}(r_k) \ , \ k = 1, \cdots, 10$$

for $\alpha_1, \cdots, \alpha_5, \beta_1, \cdots, \beta_5$ by least squares methods where $\tilde{\alpha}(r)$ and $\tilde{\beta}(r)$ are defined in (70) and r_k = (k-1)/10 . The resulting approximations for $\alpha_1, \cdots, \alpha_5, \beta_1, \cdots$ and β_5 are given in Table 3.

	n = 1	2	3	4	5
α_n	-.1349+03	.8784+02	- .658+02	.5472+02	-.2974+02
β_n	.5475+02	-.3494+02	.2577+02	-.2123+02	.1122+02

Table 3

For N = 3, the control f(x) generated by (56)-(59) is tabulated in Table 4

x	f(x)	x	f(x)
-1	14.055	- 6.5	107.70
-1.5	14.967	- 7.	118.57
-2.	13.405	- 7.5	128.55
-2.5	14.701	- 8.	137.65
-3.	20.701	- 8.5	145.94
-3.5	30.600	- 9.	153.47
-4.	42.918	- 9.5	160.31
-4.5	56.359	-10.	166.5
-5.	70.012	-10.5	172.11
-5.5	83.307	-11	177.19
-6.	95.922		

Table 4

5. CONCLUSIONS.

In this paper we have presented a method for constructing a control surface temperature such that the partial differential equations and boundary conditions describing Bridgman-Stockbarger crystal growth are not over-posed if we require the solid-melt interface to be at $x = 0$. This method can also be applied to symmetric float zone methods described by:

$$(71) \quad \nabla^2 T_s = P_s \frac{\partial T_s}{\partial x} \; , \quad x < -Q \; , \; 0 < r < 1$$

$$(72) \quad \nabla^2 T_\ell = P_\ell \frac{\partial T_\ell}{\partial x} \; , \quad -Q < x < 0, \; 0 < r < 1$$

$$(73) \quad -K_\ell \frac{\partial T_\ell}{\partial r} = h_\ell (T_\ell - C(x)) \; , \quad -Q < x < 0$$

$$(74) \quad \frac{\partial T_\ell}{\partial x} = 0 \; , \quad x = 0 \; , \; 0 < r < 1$$

$$(75) \quad -K_\ell \frac{\partial T_\ell}{\partial x} + K_s \frac{\partial T_s}{\partial x} + \mathcal{L} = 0 \; , \quad x = -Q, \; 0 < r < 1$$

and

$$(76) \quad T_\ell = M = T_s \quad at \quad x = -Q$$

where $C(x)$ is a priori known. Using dual integral transform methods, equations (72)-(74) and (76) can be solved. Using (75), if we then define $B(r) = \frac{\partial T_s}{\partial x}$, $x = Q$, and $A(r) = M$, we may use the methods of Section 2 to construct the cooler control $f(x)$, i.e.,

$$(77) \quad T_s = f(x), \quad r = 1, \quad x < -Q$$

As a numerical aside, the reader is warned not to use large N(equation (56)) because system (56) is similar to the Hilbert matrix system for such N. This, of course, is due to our method of expanding $E(t)$ in terms of e^{-kt}, $k = 1,2,\cdots$. Undoubtedly, a better basis exists but its form is still an open question to the author.

6. REFERENCES

1. Chang, C.E., and Wilcox, W.R., "Control of Interface Shape in the Vertical Bridgman-Stockbarger Technique, "Journal of Cyrstal Growth", Vol. 21, pp. 135-140, 1974.

2. Foster, L. M., "Material and Process Constraints for a Flat Interface in the Bridgman-Stockbarger Technique", N.A.S.A. CR-161511, Sec. VIII, Oct. 1980.

NASA/ASEE Summer Faculty Research Fellowship Program

Marshall Space Flight Center
The University of Alabama

Design Considerations of a Thermally Stabilized
Continuous Flow Electrophoresis Chamber II

Prepared by: Thomas S. Jandebeur, PhD

Academic Rank: Assistant Professor

University and Department: Athens State College
                                Department of Biology

NASA/MSFC:
    Division: Space Processing
    Branch: Separation Processes

MSFC Counterpart: Robert S. Snyder, PhD

Date: August 21, 1981

Contract No :   NGT 01-008-021
                                The University of Alabama
                                    in Huntsville

Design Considerations of a Thermally Stabilized Continuous Flow Electrophoresis Chamber II

by

Thomas S. Jandebeur
Assistant Professor of Biology
Athens State College
Athens, Alabama

Abstract

The overall objective of this continuing research is to investigate the basic adjustable parameters of a Beckman Continuous Particle Electrophoresis (CPE) Apparatus to determine the optimum conditions for ground-based operation for comparison with space experiments.

Previous work has determined the optimum electrical field strength and curtain flow rate for standard cell populations in phosphate buffer, (3), the optimum concentration and flow rate of the sample (2), and has investigated the postulation by Ostrach (4) that an upflow electrophoresis chamber would be more efficient than a downflow chamber (1,2).

The purpose here is to evaluate the possible application of electrically insulated copper/aluminum chamber walls as a means to thermally stabilize or equilibrate lateral temperature gradients which exist on the walls of conventional plastic chambers and which distort the rectilinear base flow of buffer through the chamber, significantly affecting sample resolution.

Introduction

The rectilinear base flow through a cooled electrophoresis chamber is extremely sensitive to any lateral temperature gradients which exist on the chamber inner walls. The uniform velocity front has been seen to be affected by lateral temperature gradients as small as $0.04^{\circ}C/cm$ in a 0.55 cm thick test chamber (5). The plastic walls of conventional continuous flow electrophoresis chambers support significantly higher gradients. Recognizing that the thermal conductivity of all transparent materials is poor, the use of an efficient heat conductor, such as copper or aluminum, in construction of an electrophoresis chamber to allow equilibration of lateral temperature gradients should result in significantly better sample resolution.

Objectives

A major problem encountered in the development of the metal walled electrophoresis chamber is the application of a thin, buffer resistant coating which will provide electrical insulation without significantly compromising the metal's thermal conductivity (5). The purpose of this research, then, is to evaluate the performance of coatings recommended for copper and/or aluminum as applied to copper test strips, to copper inserts of a modified Beckman Continuous Particle Electrophoresis (CPE) chamber, and to the aluminum walls of a prototype electrophoresis chamber, in each case with the coating being evaluated under conditions similar to those encountered during continuous flow electrophoretic separation of a mixture of fixed red blood cells (1,2,3).

Materials and Methods

The following coatings recommended for copper, applied individually to clean copper test strips by dipping, were exposed to a 0.15 M NaCl solution for varying intervals not exceeding 30 days and examined periodically for evidence of corrosion (5):

> Vinylidene flouride powder resin/
>   "Kynar" (Penwalt Corporation)
>   in dimethyl formamide

> "Formvar" (Monsanto)

> "Butvar" (Monsanto) coating formulation
>   for metal cans made up from "Epon 1007"
>   (Shell Oil Company) and "Methylon 75103"
>   (General Electric)

71000 epoxy resin coating (Armstrong)

R-4-3117 silicone conformal coating
(Dow Corning)

"Temp-R-Tape C" (CHR Industries)

Where appropriate, the coatings were preceded by an "Iridite # 7P" precoat (Allied Research Products). The Kynar film was crosslinked using a 10 mega rad dose of gamma radiation from a Co^{60} source. A fluidized bed technique was employed in the application of the 71000 epoxy resin coating.

Several coatings recommended for copper, applied individually to clean copper test strips by dipping (1202 Glyptal – General Electric; PDS Air Dry Plastic Coating – Industrial Products) or by spraying (Krylon No. 1303 Crystal Clear acrylic spray coating), were evaluated while immersed in a 0.15 M NaCl solution supporting an electric field (395 volts, 120 milliamps), with the coated copper serving as one electrode, and for varying intervals not exceeding 6 hours. Failing to conduct the current, the coating might then be judged to be an acceptable electrical insulator.

The more resistant coatings, including R-4-3117 silicone conformal coating, 71000 epoxy resin coating, Temp-R-Tape C and PDS Air Dry Plastic Coating, were subsequently evaluated in a Beckman CPE chamber modified as described in the following paragraph.

The frontplate glass insert was replaced by a copper insert of the same dimension. A plexiglass backplate fitted with a copper insert replaced the standard all plexiglass CPE backplate. Two vertical channels in the plexiglass portion of the backplate each containing 2.5 mm diameter spectra/pore semi-micro cellulose membranes enclosing platinum electrodes communicate with the separation chamber through two 28 cm by 0.2 cm slots in the copper insert. A copper "boat" fixed to the back of the copper insert comprised the backplate cooling chamber. The copper insert with the copper "boat" attached was placed, as a unit, into a recession milled into the front of the plexiglass backplate.

Results

The R-4-3117 silicone conformal coating and the 71000 epoxy resin coating, applied at thicknesses of approximately .001 inch, failed to completely insulate the copper wall inserts

of the modified Beckman CPE chamber (5). Adequate insulation was provided by Temp-R-Tape C (5), and the PDS Air Dry Plastic Coating, but the configuration of the backplate copper insert electrode slots did not prove compatible with the application of either coating.

To eliminate the electrode slot coating problem, a prototype aluminum chamber was designed such that electrode channels communicated with the separation chamber from the sides of, rather than from the rear of, the separation chamber. Platinum electrodes enclosed by 2.5 mm diameter spectra/pore semimicro cellulose membranes were employed to establish an electrical field across the separation chamber. The electrodes and membranes were secured to plexiglass top and bottom plates designed to allow for electrode cooling and curtain buffer flow.

Applied to the aluminum chamber walls, Temp-R-Tape C provided adequate electrical insulation when tested for approximately 5 hours at an electrical field strength of 600 volts and 18 milliamps.

Conclusions and Recommendations

Temp-R-Tape C (CHR Industries, New Haven, Connecticut), a thermal curing pressure-sensitive extruded fluorinated ethylene propylene (FEP) copolymer provides an acceptable combination of mechanical, thermal, electrical, and chemical properties for electrophoretic application. The tape has negligible moisture absorption and exhibits freedom from pinholes and voids as evidenced by its high dielectric strength (2800 volts/mil). Further, it is extremely thin (3 mils), is relatively easy to apply, and adheres well to a copper or aluminum substrate.

Dupont Company manufactures several Teflon finishes which may be applied with electrostatic or conventional compressed air paint spraying equipment. Teflon FEP finishes, designed to give a non-porous surface with high release characteristics, a low coefficient of friction and excellent chemical resistance, may be applied to aluminum with a total dry film thickness of from 0.9 to 1.3 mils, including a primer thickness of approximately 0.3 mils. Teflon-S non-stick finishes are similarly characterized.

The following recommendations are made, to be carried out in the indicated sequence:

(1) that Teflon FEP and/or Teflon-S finishes be applied
to the prototype aluminum wall chamber by a licensed industrial
applicator and evaluated under test conditions similar to those
described in this paper;

(2) that Type C Teflon (FEP) film, 0.5 mils thick (50 gauge),
be applied to the prototype aluminum wall chamber using 46960
polyester adhesive and RC803 curing agent (both available
from Dupont Company), and evaluated under test conditions
similar to those described in this paper;

(3) that, in order to be compatible with the application
of Teflon finishes or films, the Beckman CPE chamber be modi-
fied such that the electrode placement is at the side of the
separation chamber, as opposed to from the rear of the
separation chamber, and to accept aluminum wall inserts;

(4) that, in addition to Temp-R-Tape C, should the Teflon
FEP and/or S finishes and/or the Type C Teflon (FEP) film
prove to be effective electrical insulators, these coatings
be applied to the aluminum wall inserts of a Beckman CPE
chamber modified as described in (3) above; and

(5) that the Beckman CPE chamber modified as described
in (3) and (4) above be utilized to generate data for compari-
son with results achieved in the separation of a mixture of
fixed red blood cells in the conventional Beckman CPE chamber (2).

References

1. Jandebeur, Thomas S., Continuous Flow Electrophoresis: The effect of Sample Concentration on Throughput and Resolution in an Upward Flowing System.   NASA CR+161511

2. McGuire, J. K. and R. S. Snyder, Operational Parameters for Continuous Flow Electrophoresis of Cells, in press.

3. McGuire, J. K. and R. S. Snyder, Characterization of Continuous Flow Electrophoresis for Improvement of Resolution and Throughput, in press.

4. Ostrach, S., Convection in Continuous Flow Electrophoresis, J. Chromat. 140, 187 (1977).

5. Rhodes, Percy H., Teresa Y. Miller and Robert S. Snyder, Design Considerations of a Thermally Stabilized Continuous Flow Electrophoresis Chamber, in press.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

MATHEMATICAL PROGRAMMING TECHNIQUES FOR SCHEDULING
SPACELAB CREW ACTIVITIES AND EXPERIMENT OPERATIONS

Prepared By: Frank H. Mathis

Academic Rank: Assistant Professor

University and Department: Baylor University
                                     Department of Mathematics

NASA/MSFC:
      Division: Mission Analysis
      Branch   Mission Integration

MSFC Counterpart: William C. Askew

Date: August 14, 1981

Contract No.:   NGT 01-006-021
                                     The University of Alabama in Huntsville

MATHEMATICAL PROGRAMMING TECHNIQUES FOR SCHEDULING SPACELAB CREW ACTIVITIES AND EXPERIMENT OPERATIONS

BY

Frank H. Mathis
Assistant Professor of Mathematics
Baylor University

ABSTRACT

The objective of this report is to investigate several mathematical programming techniques which may be applied to the scheduling of experiments and crew activities for Spacelab missions. We will discuss the use of currently known methods, in particular zero-one programming and heuristic dispatching. In addition we propose a new type of scheduling algorithm and present examples to illustrate and test its use.

ACKNOWLEDGEMENT

INTRODUCTION

In order to obtain a schedule of activities for a Space-lab mission , a number of experiments must be assigned to specific times in the mission so that a variety of constraints are satisfied. Since this problem may be formulated as a limited resource - project scheduling task, several well-tested algorithms exist for obtaining a good solution.

The use of zero-one programming as developed by Pritsker, Watters and Wolf [4] and improved by several authors (see, for example, [5]) has proved effective on small scheduling problems. However, run time and computer storage requirements increase rapidly with the number of variables present and may prove prohibitive for use on a problem the size of a Spacelab mission.

On the otherhand, dispatching techniques, which take the experiments to be scheduled in a specific order and assign each to the first valid time in the mission, produce results with relatively fast run-time and small storage requirements. However, these methods may fail to yield a good schedule. The result is that a great deal of time must be spent investigating and editing the timeline obtained from a dispatching scheduler to obtain a better schedule.

Clearly there is a need for algorithms which can obtain a good quality timeline and not exceed unreasonable time and storage requirements.

OBJECTIVES

The objective of this report is to investigate current techniques for scheduling with resource-constraints. In particular, we discuss the use of zero-one programming and dispatching methods. We will then introduce a new algorithm which combines characteristics of both existing methods to generate a valid schedule of experiments. Examples and test results are presented to illustrate the feasibility of the use of this algorithm on data typical to Spacelab missions.

DEFINITION OF PROBLEM

In order to formulate the problem of scheduling a Space-
lab mission, we assume that each experiment has been divided
into a number of individual tasks which we denote as models.
A model in turn is divided into separate steps each of which
demand a specific set of requirements. In addition, each
model may have to be scheduled several times. We denote this
by the number of performances of a model.

The constraints imposed on the schedule are of the fol-
lowing types. A step may have several resource requirements.
These include the use of a crewman or a certain piece of
equipment as well as electrical power, computer memory, data
transmission and so forth. Two steps may have a sequencing
constraint. That is, step A must be scheduled before step B
with specified minimum and maximum delays between the two
steps. Each model will have an early and a late start time be-
tween which the first step of each performance must be sched-
uled. Finally, many steps must be scheduled during one of
several specific time intervals of the mission. For example,
a step may require the ability to view a certain star. Thus
that step can only be scheduled during those times in the
Spacelab orbit that the star is visible. We refer to this
type of requirement as a target. We assume that we know the
times at which any target is available during the mission.

Thus the problem is to schedule as many as possible of
the specified number of performances of each model such that
all constraints are satisfied. We refer to the resulting
schedule as a timeline.

As an example of the type problem encountered in a Space-
lab mission we will use a data file which was constructed in
June 1980 for Spacelab I. We point out that this file is not
currently in use; however, we feel it can serve to typify the
size and structure of the problem. Table I summerizes this
data.

TABLE I

DATA for SPACELAB I (June 1980 version)

Length of timeline   165.61   Hours

Number of models   218

Number of performances   935

Total number of steps   3497

Number of crew available   6

Types of equipment   38

Other types of resource constraints 4

Number to targets   72

TYPES OF SCHEDULERS

I. Zero-one Programming.

The use of zero-one programming was first applied to scheduling problems in a series of papers by Pritsker, Watters and Wolf [4]. In this approach the timeline is divided into discrete time intervals. Corresponding to each interval a variable must be defined for every step in the problem. A variable is assigned the value of "one" if the corresponding step starts during that time interval and a value of "zero" otherwise. One advantage of this formulation is that the problem is easily represented by a set of linear equalities and inequalities. We illustrate this below. Refer to Table II for a listing of the required variables.

TABLE II
Variables for zero-one programming

i	-	Subscript referring to a model.
j	-	Subscript referring to a performance.
k	-	Subscript referring to a step.
t	-	Subscript referring to a time interval.
x (ijkt)	-	A zero-one variable assigned the value of 1 if step k of performance j of model i starts in time period t, and 0 otherwise.
d (ik)	-	Duration of the kth step of model i.
min (ik)	-	Minimum delay after step k of model i.
max (ik)	-	Maximum delay after step k of model i.
E (i)	-	Early start time for model i.
L (i)	-	Late start time for model i.
N (i)	-	Number of steps for model i.
P (i)	-	Number of performances for model i.

TABLE II

Variables for zero-one programming (cont.)

m	–	Subscript referring to a resource.
r (ikm)	–	Amount of resource m required by step k of model i.
R (mt)	–	Amount of resource m available at time t.

The various constraints may be defined as follows:

Sequencing examples for steps k and k+1 of model i, performance j:

$$\sum_{t=E(i)}^{L(i)} \left\{ t \cdot X(ijkt) \right\} + d(ik) + \min(ik) \leq \sum_{t=E(i)}^{L(i)} \left\{ t \cdot X(ijk+1t) \right\}$$

and

$$\sum_{t=E(i)}^{L(i)} \left\{ t \cdot X(ijk+1t) \right\} \leq \sum_{t=E(i)}^{L(i)} \left\{ t \cdot X(ijkt) \right\} + d(ik) + \max(ik).$$

Resource example (including targets) for resource m at time $\hat{t}$:

$$\sum_{i} \sum_{j=1}^{P(i)} \sum_{k=1}^{N(i)} \left\{ r(ikm) \sum_{t=\hat{t}-d(ik)}^{\hat{t}} [X(ijkt)] \right\} \leq R(m\hat{t}).$$

In addition we require a step definition constraint for each step. This will assure that the step is assigned to at most one start time.

$$\sum_{t=E(i)}^{L(i)} X(ijkt) \leq 1, \quad \text{for all i,j and k.}$$

Also a performance completion constraint is needed for each performance. This will allow no step in a performance to be scheduled unless the last step is also scheduled.

$$\sum_{t=E(i)}^{L(i)} \sum_{k-1}^{N(i)} X(ijkt) = \sum_{t=E(i)}^{L(i)} \left\{ N(i) \cdot X(ij\ N(i),t) \right\}, \text{for all } i \text{ and } j.$$

Using the above constraints one may program the problem to optimize several objective functions. We list some examples.

To obtain the most steps possible we select the X(ijkt) to maximize.

$$\sum_{i} \sum_{j=1}^{P(i)} \sum_{k=1}^{N(i)} \sum_{t=E(i)}^{L(i)} X(ijkt).$$

Similarly maximizing the following functions produces the most performances, models and experiment time respectively.

$$\sum_{i} \sum_{j=1}^{P(i)} \sum_{t=E(i)}^{L(i)} X(ij\ N(i)\ t),$$

$$\sum_{i} \sum_{t=E(i)}^{L(i)} X(i\ P(i)\ N(i)\ t),$$

$$\sum_{i} \sum_{j=1}^{P(i)} \sum_{k=1}^{N(i)} \sum_{t=E(i)}^{L(i)} \left\{ d(ik) \circ X(ijkt) \right\}.$$

Since good algorithms exist to handle integer programming with zero-one variables (see, for example, chapter XI of [3]), the above approach appears as an attractive means to solve the scheduling problem. However, if one applies this technique to Spacelab data, obvious problems arise.

First the largest time interval which would reasonably
define the timeline is one minute. This means that a formu-
lation of 3497 steps over 165 hours would require well over
3 million variables. Although techniques such as the method
of Talbot and Patterson [5] can greatly decrease this storage
requirement, the time required to solve the zero-one problem
remains substantial. This is because the most efficient zero-
one algorithms employ an implicit enumeration technique in
which after a feasible solution is obtained, the method back
tracks until all possible solutions are directly or indirectly
observed. The larger the problem the longer it takes to ac-
complish the back tracking. In [5] results for test problems
with up to 800 zero-one variables and 300 constraints are
reported. Using an IBM 370/168 the authors obtain run-times
of 6 to 30 seconds CPU. If this were extrapolated to a prob-
lem the size of a Spacelab mission, run times of 5 to 30 hours
might be expected. These numbers were obtained assuming that
the time increases proportionally with size. A more reason-
able assumption would dictate that run-time increases with
the cube of the size. Indeed this is usually the case for
linear programming applications. Extrapolating under this
assumption would result in a run-time of over 9000 years!

II. Dispatching Techniques.

Because of the large storage and long run-time require-
ments of the zero-one formulation, there is a need for alter-
native methods if one is faced with a problem of any size.
The most common approach for scheduling large problems is a
dispatching technique. In these methods one places an order
on the list of steps to be scheduled, then following this
order the steps are assigned to the first time at which all
constraints are satisfied. This allows for a schedule to be
obtained very rapidly with minimal storage requirements.
Clearly, the order on the steps is crucial to the structure
of the resulting timeline. In some cases, one may obtain a
good timeline by choosing the order based on a heuristic rule
or "rule of thumb". For example, one may wish to schedule
those steps which require the most resource first. Davis
and Patterson [1] have tested several heuristic dispatching
rules on various scheduling problems. Also see Grone and
Mathis [2] for a discussion of dispatching methods applied
to Spacelab missions.

We note that the current scheduler in use at MSFC is a
dispatching type program which may use either a random ordering
or an order fixed by the operator. We will refer to this pro-
gram as TLP for Timeline Program. In a test envolving 30
runs of TLP using the random ordering on our sample data, we
found that a valid timeline is obtained in less than 2 minutes
of elapsed time.

None of the runs produced a timeline with all of the desired models scheduled; the average run scheduled approximately 80% of the models and 87% of the performances.

There are several problems with a dispatching scheduler. There is no guarantee that the resulting schedule is near optional. Also, heuristics seem to be very problem dependent. That is, a rule which works well on a particular problem may prove disastrous for the next problem. Thus, a dispatching scheduler may generate a valid timeline very quickly, but then a considerable amount of time must be spent editing the schedule in order to improve its quality. This is indeed the case with TLP. The timeline produced by TLP can be converted into one in which 93% of the models and 96% of the performances schedule, but only after much investigation and alteration requiring several man-hours of effort.

A NEW ALGORITHM

We now propose a new algorithm which combines some elements of both zero-one and dispatching techniques. We wish to construct a method which has the flexibility of zero-one programming so that after a step is scheduled we may later move the step to another place in the timeline. However, in order to speed up the algorithm, we will not attempt a full back tracking procedure but rather depend on a heuristic rule to dictate which steps should move.

To define our method, we divide constraints into two groups. A constraint is called step dependent if the amount available to a step at a particular time will change as other steps are scheduled at the same time. Crew usage, equipment and other resources are examples of step dependent constraints. Step independent constraints are those which do not depend on how many steps are running at the same time. Examples are early and late start time, sequencing, and targets.

We say that a step is assigned to a feasible time if no step independent constraint is violated, and the timeline is valid if all constraints are satisfied.

The general idea of the algorithm is to start with a feasible timeline and then proceed from the front to the back of the timeline checking each time for validity. If at any time it is found that a step dependent constraint is violated, a particular step scheduled at that time is selected and moved to a future time in the timeline.

Whenever a step is moved, all other steps in the same model must be checked and possibly moved to assure that maximum and minimum delay requirements are satisfied. We will not got into the details of this process but rather denote it by the phrase "move other steps as necessary". If, during an attempt to move step j to a new time, any step in the same model can not be reassigned to a feasible time, we will say that step j cannot be reassigned.

If at any time, a step dependent constraint is violated but no step can be reassigned, then a step is selected and removed from the timeline. In this case all other steps in the same performance must be removed. Of course, the method of choosing a step to be moved or removed is vital to the quality of the resulting timeline. As with a dispatching method, one may use any of several heuristic rules to accomplish this. In the algorithm to follow, we will move the step which uses the most of the violated resource.

This strategy is used in an attempt to cause as many steps as possible to schedule at the present time and to minimize the number of moves required. We do not claim that this rule is necessarily the best.

We refer the reader to table III for a list of variables used in the algorithm. A formal statement of the algorithm follows:

TABLE III

A Variables List for the Scheduling Algorithm

N	-	Number of steps.
st(j)	-	Starting time of step j.
d(j)	-	Duration of step j.
T	-	Current time under investigation for validity.
T (old)	-	Previous time investigated.
T (new)	-	Time to which a stp may be re-assigned.
J (T)	-	Set of indices j such that step j is scheduled during time T. That is, $st(j) \le T \le st(j) + d(j)$
L	-	Set of indices corresponding to steps for which time reassignments were not possible.
M	-	Number of step dependent constraints.
R (k,T)	-	Amount of resource k available at time T.
r (jk)	-	Amount of resource k required by step j.
S (k,T)	-	Scaled requirement for resource k at time T; $$S(k,T) = \left\{ \sum_{j \in J(T)} r(jk) \right\} / R(k,T).$$
ϕ	-	The empty set.

ALGORITHM

1. Assign st (j), j=1,2,··· N, to its earliest feasible start time. Set T=-1.

2. Set T (old) = T ; Set L = ϕ.

3. If $\left\{ j \text{ such that } st (j) > T (old) \right\} = \phi$, stop ; otherwise set $T = \min\limits_{j} \left\{ st(j) \text{ such that } st(j) > T (old) \right\}$.

4. Find all j in J (T).

5. Find k which corresponds to the maximum S (k,T), k=1,2, ···M.

6. If S (k,T) $\leq$ 1, go to 2 ; otherwise continue.

7. If J (T) - L = ϕ, let j be the first index in L, remove step j from the timeline and go to 4; otherwise continue.

8. Find i in J(T)-L to maximize r(ik).

9. Set $T (new) = \min\limits_{j} \left\{ st(j) + d(j) \text{ so that } j \text{ is in } J(T) \text{ and } e(jk) \neq 0 \right\}$.

10. Attempt to reassign st(i) to the first feasible time $>$ T (new); more other steps as necessary.

11. If st(i) can be reassigned go to 4; otherwise, add i to L and go to 7.

We point out that because reassignment of a step may cause other steps to move to a time that has already been checked for validity, a single application of the algorithm may not produce a valid timeline. In practice the algorithm must be repeated until no moves are necessary.

EXAMPLES

In order to illustrate the operation of the algorithm, we present several examples. The first is a test problem used in [4]. Also see chapter XII of [3]. The problem contains 6 models, 8 individual steps and 3 types of resources. The data is summerized in Table IV.

TABLE IV
DATA FOR EXAMPLE 1.

Index	Model	Step	Earliest Start	Duration	Latest Start	Resource Requirements		
						1	2	3
1	1	1	0	4	1	5	3	2
2	1	2	4	3	5	0	1	1
3	2	1	0	3	5	2	0	2
4	3	1	1	3	4	1	1	1
5	3	2	4	2	7	2	2	0
6	4	1	1	2	7	2	0	0
7	5	1	2	5	4	2	1	1
8	6	1	2	1	8	1	3	0
Amount of Resource Available at all times						8	5	4

The initial feasible timeline is shown in figure 1. At T=1 the timeline is valid. Moving to T=2 we find resource 1 has the maximum scaled requirement and step 1 is the greatest user. We attempt to reassign step 1 to T(new) = 3; however, this would cause step 2 to be scheduled after its latest start. Hence, step 1 cannot be reassigned. We next attempt to reassign step 3 to T(new). This results in the timeline shown in figure 2.

The timeline is now valid at T=2. We proceed with the algorithm and after six more reassignments we obtain the timeline shown in figure 3. Note that all times are valid.

It is interesting to note that McMillian [3] obtained the schedule shown in figure 4 using zero-one programming. Here the problem is formulated with 33 variables, 37 constraints, and the objective to minimize schedule length. McMillian reports a run-time of 2.3 seconds CPU on an IBM 7044. McMillian also applied a dispatching rule at "minimum slack-time first" to the above example. Slack time is the amount of time between the earliest start time and the latest start time. The result is shown in figure 5. Note that model 3 did not schedule.

As a second example, a computer program was designed to impliment our algorithm on the DIGITAL VAX-11/780. The first 100 models of the Spacelab data file were selected as an initial test case. This represents a problem with 1911 steps. The algorithm successfully ran in 11 minutes 23 seconds of CPU time. A valid timeline was produced with 1624 steps scheduled. This represents about 85% of the steps requested.

Finally the same program was applied to the full 218 models in the Spacelab data file. Although not all constraints were used - some of the maximum and minimum delays were not checked - the algorithm produced a timeline in 48 minutes of CPU time. Of the 3497 steps requested, 3001 or about 86% were scheduled.

RESOURCE USAGE										
k=1	7	10	13	8	4	4	2	0	0
k=2	3	4	8	5	4	4	2	0	0
k=3	4	5	6	4	2	2	2	0	0
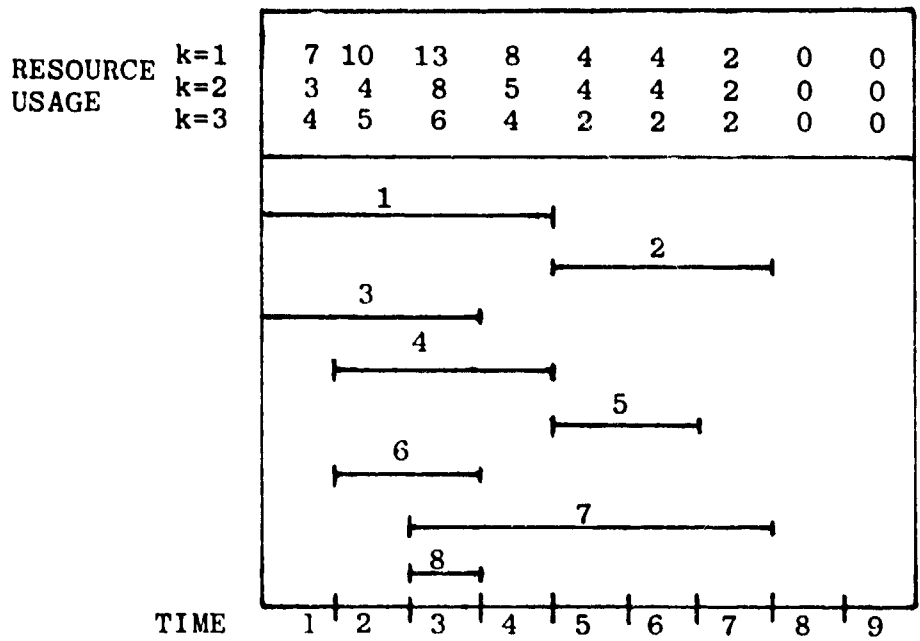
Figure 1. Initial Feasible Timeline.

RESOURCE USAGE										
k=1	5	8	11	11	6	6	2	0	0
k=2	3	4	8	8	4	4	2	0	0
k=3	2	3	4	6	4	4	2	0	0
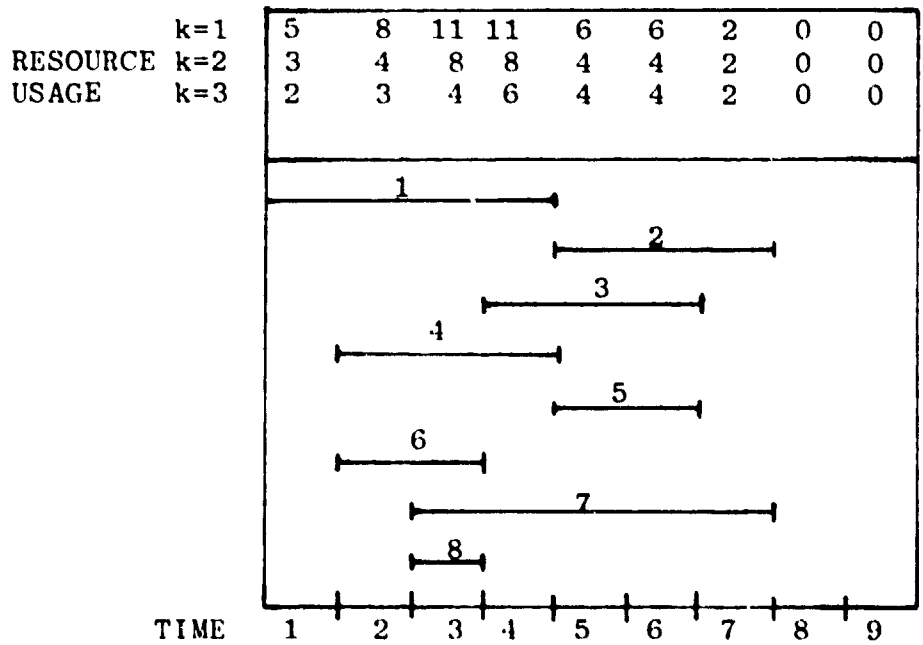
Figure 2. Timeline after one move.

XXXII-15
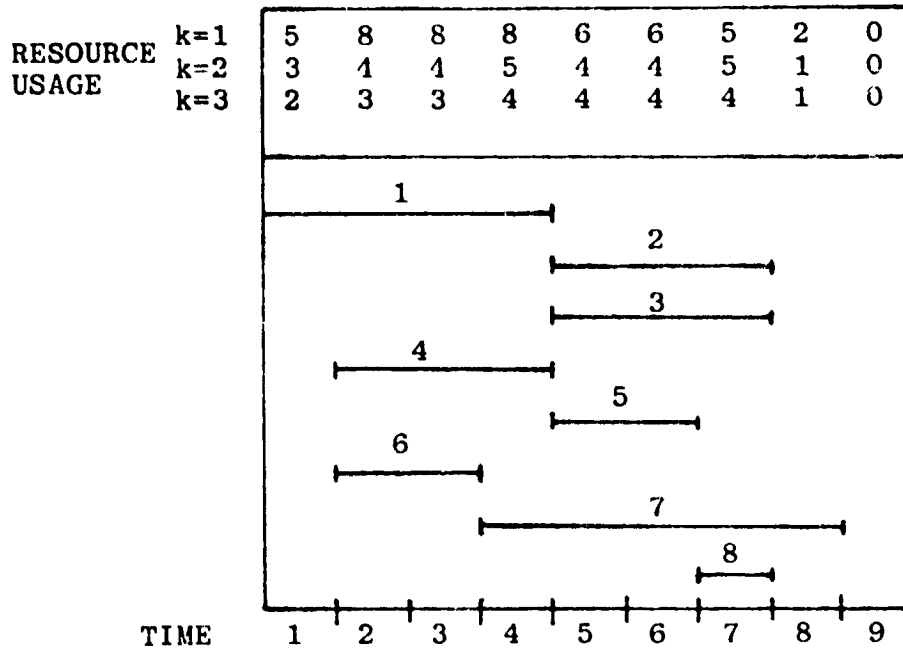
Figure 3.   Final Valid Timeline.
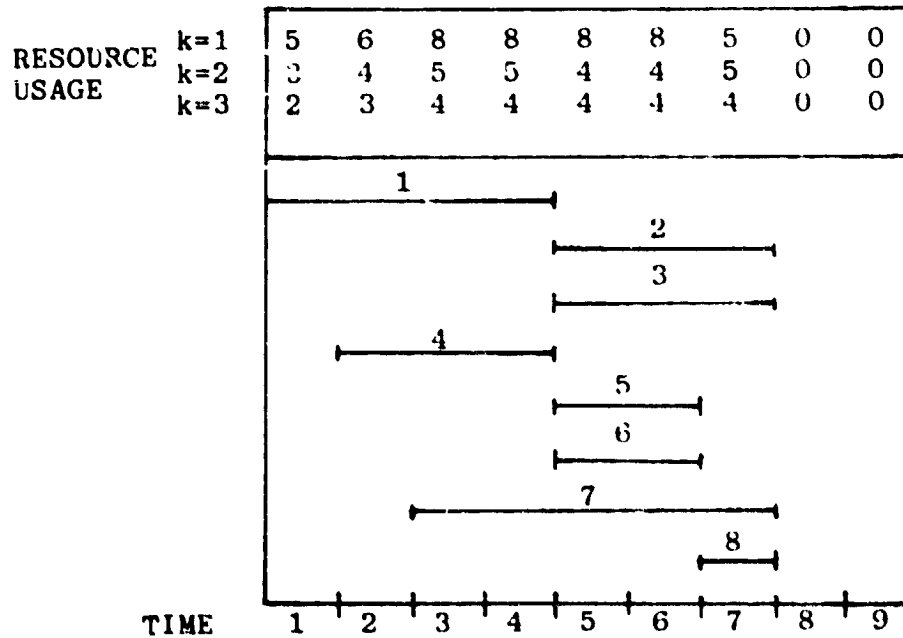
Figure 4.   Timeline from zero-one method.
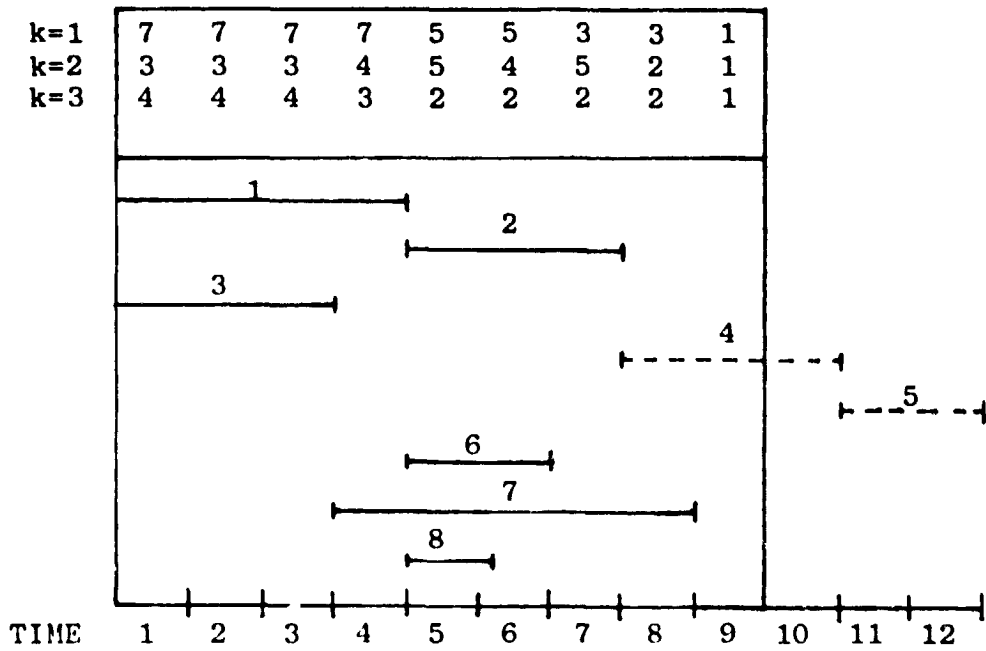
Figure 5. Timeline from dispatching method.

CONCLUSIONS AND RECOMMENDATIONS

Based on our preliminary experimental results, the algorithm presented is feasible for use on Spacelab scheduling problems, although the resulting timeline seems of comparable quality with those obtained by TLP while the expense of run-time is substantially greater than that of TLP. Nevertheless, we feel that the flexibility furnished by the algorithm warrents further investigation at least with respect to two points. First more efficient searching and sorting techniques may be employed in the coding of the algorithm to decrease run-time. More importantly, the use of alternative heuristic rules for the choice of steps to move or remove should be examined. We feel that this investigation could lead to great improvement in the performance of the algorithm as well as the quality of the resulting timeline.

REFERENCES

1. Davis, E. W. and J. H. Patterson, " A Comparison of Heuristic and Optimum Solutions in Resource - Constrained, Project Scheduling", Management Science, Vol. 21 (1975), pp. 944-955.

2. Grone, R. D. and F. H. Mathis, "A Ranking Algorithm for Spacelab Crew and Experiment Scheduling", NASA CR-161511, Marshall Space Flight Center, October 1980.

3. McMillan, C., Mathematical Programming, John Wiley and Sons, Inc., New York, 1970.

4. Pritsker, A., L. Watters and M. Wolf, "Multiproject Scheduling with Limited Resources: A Zero-one Programming Approach", Management Science, Vol. 16 (1969), pp. 93-108.

5. Talbot, F. B. and J. H. Patterson, "An Efficient Integer Programming Algorithm with Network Cuts for Solving Resource - Constrained Scheduling Problems", Management Science, Vol. 24 (1978), pp. 1163-1174.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

A STUDY OF THE RELATIVISTIC RISE IN IONIZATION AS
MEASURED BY ION CHAMBERS OF DIFFERENT DESIGN

Prepared By: Julian B. Shand, Jr., Ph.D.

Academic Rank: Professor

University and Department: Berry College
                                    Department of Physics

NASA/MSFC: Space Sciences Laboratory
    Division: Space Physics
    Branch   Astrophysics

MSFC Counterpart: Thomas A. Parnell

Date: August 14, 1981

Contract No.:   NGT 01-008-021
                                    The University of Alabama
                                     in Huntsville

A STUDY OF THE RELATIVISTIC RISE IN IONIZATION AS MEASURED BY ION CHAMBERS OF DIFFERENT DESIGN

BY

Julian B. Shand, Jr.
Charles A. Dana Professor of Physics
Professor of Computer Science
Berry College
Mount Berry, Georgia

ABSTRACT

This investigation has reconciled ion chamber data obtained by MSFC on ultrarelativistic heavy cosmic rays with recent theoretical approaches which successfully describe ion chamber data at powerful particle accelerators. The data taken by MSFC do not show a relativistic rise as great as the rise for particles of charge one at accelerators. The difference is explained by the greater thickness of the ion chambers used for cosmic rays, the greater energy transfer to the gas of the ion chamber for cosmic rays of charge 8 to 26, and those delta-ray electrons which escape from the ion chamber and are thus not recorded.

INTRODUCTION

Cosmic rays have been investigated since their discovery in 1911. Their flux in space is nearly independent of time and direction. They consist of nuclei, with small amounts of electrons and electromagnetic radiation. The nuclei are mostly protons (around 90 percent) and helium (around 9 percent), but heavy nuclei are present also. The energy spectrum extends to the greatest energies known, reaching as much as 10^{20} eV.

The origin of cosmic rays is a mystery, as is the mechanism by which they gain extremely high energy. Exploding galaxies, supernovae, pulsars, etc., have been proposed as sources, but none have been clearly identified. Measurement of the energy and identification of the various nuclear species are important for an understanding of the phenomena. For energies from 1 GeV/nucleon to several hundred GeV/nucleon, conventional time-of-flight measurements and Cerenkov counters are inadequate. Energy loss measurements become important here. Ionization measurements give an indirect estimate of the energy loss of a charged particle by measuring the number of electrons ejected from their atoms as the particle passes. This information can be used for particle identification and estimation of its energy.

OBJECTIVES

It was desired to find out why the relativistic rise in ionization detected in cosmic-ray experiments by MSFC does not agree with experiments at accelerators on singly charged particles. A related objective was to study recent theoretical treatments in the literature to see if they could be applied or needed to be applied to the cosmic-ray setup.

BODY OF REPORT

Until recently, theoretical treatments of the ionization loss of charged particles in thin gas-filled proportional counters predicted considerably greater losses than experiments showed (Refs. 1, 2). On the other hand, the experiments actually measure energy <u>deposit</u> in the chamber, not energy loss, and this may be an important distinction.

The relativistic rise in ionization has been well understood qualitatively for some time. Relativistic charged particles passing through a gas-filled chamber lose energy mainly by excitation and ionization of the gas (Ref. 3). Fano (Ref. 4) gives one way to look at the interaction between the incident particle and the atoms of the absorber. The electromagnetic field of the particle is represented by two terms. One, called the longitudinal term, is merely a static Coulomb interaction. The second term, called the transverse term, involves emission and absorption of virtual photons. The longitudinal interaction is effectively constant at high velocities. The transverse term, however, increases in strength as particle speed v increases. This is due to the increase in the transverse field strength, which is proportional to γ, where

$$\gamma = (1 - \beta^2)^{-1/2} \quad \text{and} \quad \beta = v/c$$

This relativistic rise would continue indefinitely if it were not for the polarization of the absorber atoms. The energy loss levels off in what is called the Fermi plateau.

In what we will call the Bethe-Born equation for energy loss S (Ref. 5), obtained by integrating over possible collisions,

$$S \propto \ln 2mv^2 - \ln I - \ln(1 - \beta^2) - \beta^2 - \delta/2$$

Here m = electron mass, v = particle velocity, and the absorber atoms are described only crudely by the mean ionization potential I and a density correction δ.

Sternheimer (Ref. 6) improved the theory by express-
ing I and δ in terms of the oscillator strengths and
energy levels of isolated atoms. However, the oscillator
strengths were poorly known. This Bethe-Born-Sternheimer
interaction model predicted losses which were greater than
the energy deposits measured experimentally.

Ispirian, et al. (Ref. 7) realized that for a singly
charged particle and a thin ($\sim$ 1.5 cm) gas absorber, the
number of ionizing collisions is small ($\sim$ 30) and, therefore,
an integration over collisions is inappropriate. They used
instead a Monte Carlo method with energy transfer proba-
bilities based on rather simple assumptions. Results were
somewhat closer to experiment.

Cobb, et al. (Ref. 1) used Monte Carlo methods and a
more realistic model of the absorber atom. The most pro-
bable energy loss in a thin chamber and for singly charged
particles was less than the binding energy of K-shell
electrons, so these electrons did not contribute appreciably
to the most probable energy loss. Therefore, use of a mean
ionization potential was not appropriate for singly charged
particles in a thin absorber. The following procedure was
then used. The energy absorption of longitudinal photons
was described by delta functions at the ionization energies
followed by the standard Rutherford cross section. The
absorption related to the transverse field was described
by an imaginary term in the dielectric constant of the gas.
The dielectric constant was calculated from energy levels
and plasma frequency for the gas. Good agreement with
experiment was achieved.

Allison, et al. (Ref. 8) made a further improvement by
calculating the dielectric constant of the gas directly
from the recent measurements of photoionization cross
section for the gas. A numerical convolution method was
used to calculate energy loss for a singly charged particle
in a thin absorber. Because the probability of a large
energy loss in a single collision is small, such events
were not considered; also the energy could be carried away
from the chamber in such a case. Excellent agreement with
accelerator experiments was achieved.

Applicability to Cosmic-Ray Measurements

Cosmic-ray experiments conducted by MSFC in coopera-
tion with other institutions have involved thick (8.4 cm)
ion chambers with xenon gas at atmospheric pressure. The
cosmic-ray particles of interest had charges from $Z = 6$ to
30. Of course, the strength of interactions is proportional

to Z^2. With the thick absorber, the high density of electrons, and large Z, the number and violence of collisions is large. The most probable energy loss is much greater than the binding energy of K-shell electrons. Therefore, a mean ionization potential has some validity, as Sternheimer used it. Besides this, the distribution of energy losses becomes less sensitive to the details of the single-collision cross section for thick absorbers (Ref. 9). Since Cobb and Allison neglect energy carried away by energetic delta-ray electrons, their good results may be due partly to having as much energy brought into the ion chamber as lost from the ion chamber by this mechanism. The cosmic-ray chambers are significantly different and cannot automatically be expected to have such luck. Parnell, Watts, and Gregory (Ref. 10) have done a calculation of energy deposit which keeps track of energetic delta-ray electrons entering and leaving the ion chamber. They used the Bethe-Borne-Sternheimer treatment of the absorber atom instead of the more rigorous approach of Allison and Cobb. Their results agree with their cosmic-ray experiments.

CONCLUSIONS AND RECOMMENDATIONS

The recent successful theoretical treatments of Cobb and Allison add understanding of the ion chamber measurements and the interaction of relativistic particles with matter. It seems that cosmic-ray experiments do not need as much rigor in the details of the absorber atoms but require more attention to energy carried out of the chamber by energetic delta-ray electrons. A further Monte Carlo calculation using the best available atomic model is recommended, because this should remove any suspicion of errors that cancel. A better understanding of the rigorous approach of Allison and Cobb applied to high Z particles and matching it to sophisticated transport calculations of electrons, developed at MSFC, is required. This author plans to aid the MSFC group in such an endeavor.

REFERENCES

1. Cobb, J. H., W.W.M. Allison, and J. N. Bunch, "The Ionization Loss of Relativistic Charged Particles in Thin Gas Samples and Its Use for Particle Identification," Nuclear Instruments and Methods 133, 1976.

2. Ermilova, V. C., L. P. Kotenko, and G. I. Merzon, "Fluctuations and the Most Probable Values of Relativistic Charged Particle Energy Loss in Thin Gas Layers," Nuclear Instruments and Methods 145, 1977.

3. Jackson, J. D., Classical Electrodynamics, Wiley, New York, NY, 1962.

4. Fano, V., "Penetration of Protons, α Particles, Mesons," Annual Review of Nuclear Science 13, 1963.

5. Ahlen, S. P., "Theoretical and Experimental Aspects of the Energy Loss of Relativistic Heavily Ionizing Particles," Reviews of Modern Physics 52, 1980.

6. Sternheimer, R. M., in Methods of Experimental Physics, 5A, edited by L.C.L. Yuan and C. S. Wu, Academic Press, New York, NY, 1961.

7. Ispirian, K. A., A. T. Margarian, and A. M. Zverev, "A Monte-Carlo Method for Calculation of the Distribution of Ionization Losses," Nuclear Instruments and Methods 117, 1974.

8. Allison, W.W.M. and J. H. Cobb, "Relativistic Charged Particle Identification by Energy Loss," Annual Review of Nuclear and Particle Science 30, 1980.

9. Bichsel, H. and R. P. Saxon, "Comparison of Calculational Methods for Straggling in Thin Absorbers," Physicsl Review A, Vol. 11, #4, 1975.

10. Gregory, J. C., T. A. Parnell, and J. Watts, "Relativistic Rise Measurement for Heavy Cosmic Rays in Xenon," Papers of the 17th International Cosmic Ray Conference, Paris, 1981 (in press).

AN INVESTIGATION OF THE UNIFORM
RANDOM NUMBER GENERATOR

Prepared by: Enoch C. Temple, Ph.D

Academic Rank: Assistant Professor

University and   Alabama A&M University
Department: Department of Mathematics

NASA/MSFC
Division: Systems Dynamics
Laboratory: Space Sciences

MSFC Counterpart: Mario Rheinfurth

Date: August, 1981

Contract No.:   NGT 01-003-021
                                The University of Alabama in Huntsville

Acknowledgements

INTRODUCTION

Since 1946 statisticians have been interested in the development and evaluation of arithmetic random number generators that may be easily programmed into a digital computer. One well known and once widely used arithmetic generator is the mid-square method. This method consists of squaring a K-digit number and taking the K middle digits of the squared number to form a second K-digit number. To get a third K-digit number, the process is repeated on the second number. This continues until n random numbers are generated.

One major disadvantage of the mid-square method is that its generated sequence often degenerates to a cycle which is very small, consisting in the worst case of one single number. Another short-coming is that it is difficult to analytically evaluate certain properties of this generator. Hence, the mid-square method has been superseded by other arithmetic methods whose mathematical structures are easier to analyze. For a more indepth treatment of random number generators that have received wide application, the reader should consult, Jansson (1966). Also, a complete list of recent references may be found in Sowey (1978).

Most random number generators that are in use today are of the congruential form

$$X_{i+1} \equiv AX_i + C \bmod M \tag{1}$$

where A, C, and M are nonnegative integers. If C=0, the gene ator is called the multiplicative type and those for which C≠0 are called mixed congruential generators. It is easy to see that congruential generators will repeat a sequence of numbers after a maximum of M values have been generated. The number of numbers that a procedure generates before restarting the sequence is called the length or the period of the generator. Generall , it is desirable to make the period as long as possible. The next section gives a more detailed discussion of congruential generators.

The Linear Congruential Generator

Let $\{X_i\}$ denote a sequence of numbers generated by the linear congruential generated by the linear congruential generator module M. Then $X_{i+1} \equiv AX_i + C \bmod M$. Let the initial value of the sequence be X_0. From equation (1) we have

$$X_1 \equiv AX_0 + C \bmod M \;\text{--}> X_1 = AX_0 + C - MK_1$$

$$X_2 \equiv AX_1 + C \bmod M \;\text{--}> X_2 = AX_1 + C - MK_2 = A^2X_0 + C\left(\frac{A^2 - 1}{A - 1}\right) - Md_2$$

$$X_n \equiv AX_{n-1} + C \bmod M \;\text{--}> X_n = A^nX_0 + C\left(\frac{A^n - 1}{A - 1}\right) - Mdn \tag{2}$$

where $d_n = \sum_{i=1}^{n} K_i A^{n-i}$ and $K_1, K_2, \ldots, K_n$ are integers.

Since the sequence is periodic, there exists a period P such that $X_p = X_0$ which implies that

$$X_p - X_0 = (A^P - 1)(X_0 + \frac{C}{A-1}) - MK \tag{3}$$

where K is an integer. From (3), it can be seen that the period P of a given congruential generator is the minimal P that satisfies $(A^P - 1)(X_0 + \frac{C}{A-1}) \equiv 0 \bmod M$. Ideally, A, C, M, and X_0 are selected so that the minimal P is maximized. It is worth while to remark that a maximal P is not sufficient to guarantee that the generator will produce numbers whose behavior resemble a set of random numbers.

For example, if $A = 1 = C = X_o$, the generator is (1) has period M and the generated sequence is $1,2,\ldots, m-1,0$ which is not a random sequence. Thus a large portion of the research on linear congruential generators has been devoted to the task of finding values for A, C, M and X_o so that the period P is maximized and so that the numbers generated behave as though they are random. In general, for a fixed value of M, A and C are chosen to provide speed in computation, good statistical behavior as well as a long period.

By applying the tools of number theory to equation (3), the following values of P can be determined. (See Knuth (1969) for details.)

(i) $P = M$ if $M = 2^\beta$, $A \equiv 1 \bmod 4$ and C is relatively prime to M.

(ii) $P = \frac{M}{4}$ if $M = 2^\beta$, $A \equiv 3 \bmod 8$ or $1 \bmod 4$, X_o is odd and $C = 0$.

(iii) $P = M - 1$ if M is prime, $C = 0$, $X_o \neq 0$.

In the sequel that follows, items (i), (ii) and (iii) above will be called generator (i), (ii) or (iii).

Generator (i) is a full period mixed congruential generator. Since the period of this generator is M, the choice of the starting value X_o is not important because all possible values $0,1,\ldots, M-1$, will be generated. Generator (ii) is a maximal period multiplicative generator with period $M/4 = 2^{\beta-2}$. Since A and X_o are odd and M is even, it follows that all X_i values are odd. That is, one period of this generator contains exactly one-half of the odd numbers between 0 and M. Generator (iii) removes the restriction against even numbers imposed by generator (ii). Its initial value X_o may be any number between 0 and M.

Except for the restriction that C be relatively prime to M, the choice of C is usually discussed only in terms of whether or not it should be zero. That is, whether or not the generator should be multiplicative or mixed. Greenberger (1961) used a function of C to place bounds on the correlation between X_{i+1} and X_i. However, his result does not appear to be very useful in selecting a good value for C. The most widely used generators have been those where $C=0$, $M=2^\beta$ and A takes on a value so that the generator has maximal period.

In closing the discussion on selecting a value for C, it is worthwhile to remark that a multiplicative generator module 2^β can be represented equivalently as a mixed full period generator module $2^{\beta-2}$. To see this, consider the multiplicative generator $Y_i = AY_{i-1} \bmod 2^\beta$ with $A \equiv 5 \bmod 8$. Let $Z_i = (Y_i - 1)/4$. Then $Z_i = AZ_{i-1} + C \bmod 2^{\beta-2}$ where $C = 2d + 1$, d is an integer. This is an indication that as long as $M = 2^\beta$, one need not place emphasis on the selection of C to improve the period or the statistical behavior of a congruential generator. Also note that Allard, Dotell and Hull (1963) ran extensive frequency and serial tests on congruential generators. They concluded that X_o and C are far less important than the multiplier A in establishing statistical behavior of a generator.

Although (i), (ii) and (iii) have identified values for A that will provide a full or maximal period, many studies have focused on the choice of A that will produce a sequence of numbers with properties that one would expect to find in a sequence of random numbers. A necessary condition for the approval of a sequence of numbers is that the numbers pass a collection of standard statistical tests on equi-distribution, runs, serial correlation, etc. The book by Jansson (1966) describes nine difference emperical tests of randomness.

Knuth (1969) describes two theoretical tests that may be used to assess the randomness of a congruential generator. They are the lattice test, introduced by Marsaglia (1972) and the spectral test developed by Coveyou and MacPherson (1967). These two tests are limited to full-period linear congruential generators and to certain multiplicative generators.

The development of the lattice and spectral tests followed an observation by Lach (see Greenberger (1965)) of straight line patterns in the two-dimensional plots of X_i against X_{i+1}, where the sequence X_i is produced by a congruential generator. These straight line patterns are referred to by Marsagl a (1972) as the lattice structure of the generator. The lattice structure of congruential generators were once thought to be caused by the size of the modulus, the initial value X_0 or even the characteristic of the multiplier. Chambers (1967) concludes that the straight line patterns would disappear when the period is sufficiently large. His conclusion was perhaps based on a small sample taken from a relatively long period where, for a small sample, the lattice structure was not immediately apparent. Marsaglia (1968) showed that random numbers produced by congruential generatores lie in a rela-tively small number of hyperplanes. Thus, the existence of the lattice structure is independent of the parameters. As an example, figure 1 shows a plot of successive pairs (X_i, X_{i+1}) from the simple modulus generator $X_{i+1} = 7 X_i \mod 61$.

Many authors have recommended various values for the multiplier A. Some recommended values for A as a result of empirical studies and at least two papers, Marsaglia (1972) and Atkinson (1980), make an excellent comparison of theoretical and empirical tests for a family of multiplica-tive and mixed generators. Marsaglia compared the lattice test to em-pirical tests while Atkinson made comparisons on both the lattice and spectral tests. Some recommendations of good generators were made by Marsaglia and by Atkinson. Both authors claimed that their research is a clear indication that the lattice and spectral tests measure a quantity which is related to bad performance of generators. In particular, results from Marsaglia (1972, page 280) indicate that the widely used multiplier $A = 2^{\alpha} + \gamma$, α and γ are integers, are likely to produce nonrandom genera-tors. Empirical studies by Hull and Dobell (1964) support this claim.

Other Methods

Many other methods for generating random numbers have been proposed. Many of these proposals consist of slight modifications to procedures discussed above.
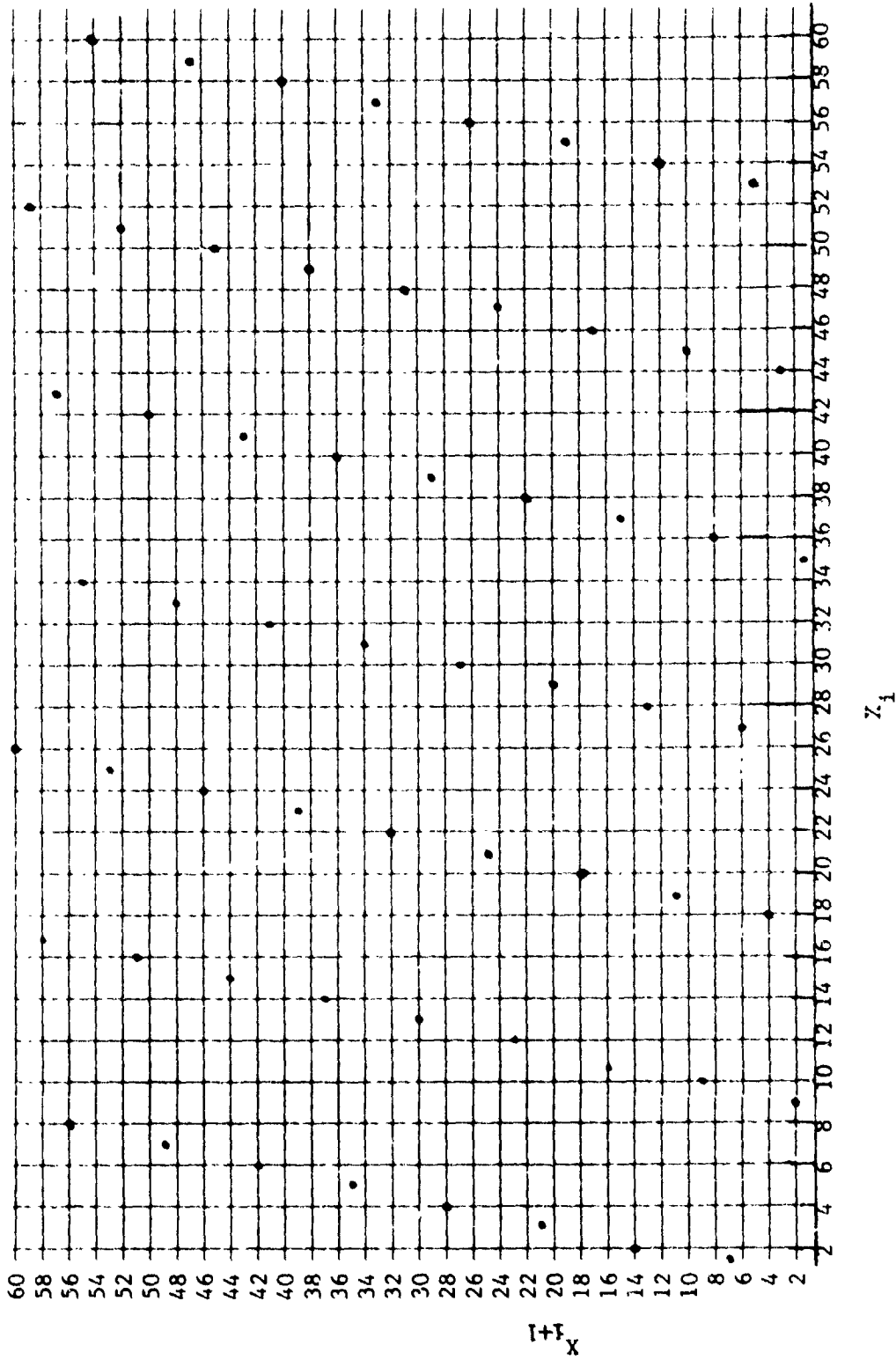
FIGURE 1

In this section several promising procedures that differ from the multiplicative and mixed procedure will be briefly discussed. The first method is a strictly additive congruential generator introduced by Green et al. (1959). The ith element of the sequence is generated by the relation

$$X_i = X_{i-1} + X_{i-2} + \ldots + X_{i-K} \bmod M \qquad (4)$$

where M is prime and X_1, X_2, $\ldots$, X_K are K initial values selected by some process. As far as this author has determined, very little mathematical theory has been applied to the study of the behavior of additive generators. However, several empirical studies have been made with at least two studies reporting conflicting results. See, for example, Burford (1973) and Perkins and Menzefricke (1975).

The second method is a more complex additive method where the ith term is computed by

$$X_i = a_1 X_{i-1} + a_2 X_{i-2} + \ldots + a_K X_{i-K} \bmod M \qquad (5)$$

M is a prime, $a_i (i = 1,2,\ldots,K)$ are nonnegative integers and X_1, X_2, $\ldots$, X_K are the K initial values. Through the theory of finite fields it is possible to show that integers a_1, a_2, $\ldots$, a_K exist so that the sequence defined by (5) has period M^K-1. However, it is difficult to compute their values. Knuth (1969) gives an approach for selecting the multipliers a_1, a_2, $\ldots$, a_K. This author has not been able to locate a reference which statistically evaluated a generator of the type given in (5) for values other than $a_i = 1$ (i = 1, 2, $\ldots$, K). If $a_i = 1$ for i = 1, 2,$\ldots$,K, (5) becomes the generator given (4).

The third method was introduced by Maclaren and Marsaglia (1965). They conducted a series of tests on the multiplicative and mixed generators and on two improved generators which they developed. Both of their improved methods essentially consist of coupling two generators where one generator shuffles the sequence of the other. They concluded that their procedures are superior to both the multiplicative and mixed generators.

Brown and Solomon (1976) showed that if two vectors $\underline{X}_i$ and $\underline{Y}_i$ contain independent components, then the sequence defined by $\underline{Z}_i = \underline{X}_i + \underline{Y}_i$ converges to the vector $\underline{\gamma}$ where $\underline{\gamma}$ is a vector of independent uniformly distributed random variables. They also indicated that

$$\underline{W}_n = \sum_{i=1}^{n} X_i \qquad (6)$$

converges to $\underline{\gamma}$.

This author made an effort to use (6) to define a generator that exhibits a better uniform random behavior than now exists in the multiplicative or the mixed generator. Here it should be recalled that the period of multiplicative or mixed generators is some number P. If the pairs (X_i, X_{i+1}) are truly random, one would expect P^2 different points with each point occurring with probability P^{-2}. However, the plot of the ordered pairs (X_i, X_{i+1}) gives only P points in the two-dimensional space. Thus, it seems that P(P-1) points are missing.

If a new generator could be constructed, from a multiplicative or mixed generator having period P, so that a total of P^2 different ordered pairs (X_i, X_{i+1}) could possibly occur, then our intuitive expectation would be satisfied. Furthermore, a generator of this type would have a period of P^2 and the best possible two-dimensional lattice structure. In general, a generator with a period of P^K would have the best possible K-dimensional lattice structure.

To evaluate the period of the sequence of partial sums $W_n = \sum_{i=1}^{n} X_i$, we let

$$X_{i+1} = AX_i + C \mod M \tag{7}$$

where P is the period of sequence X_i. Then if (7) is a full period (P=M) generator with M even, $\sum_{i=1}^{M} X_i = \frac{(M-1)M}{2}$

(The sum of M-1 consecutive positive integers) implies that the maximum period of W_n is 2M. If (7) is a maximal period generator with period M-1, C=0 and M odd, $\sum_{i=1}^{M-1} X_i = \frac{(M-1)M}{2}$

implies that the maximal period of W_n is M. Through empirical evaluations, it appears that if $C \neq 0$ and if $X_{i+1} = AX_i + C \mod M$ has period P<M-1, then the maximal period of $W_n = \sum_{i=1}^{n} X_i$

is PM. Hence it is safe to conclude that W_n does not have period M^2 or M^2-1. Additionally, the 2-dimensional lattice structure of generator W_n is worse than that of (7) where A in (7) is selected to give a good lattice.

A concluding remark should be made about a generator of the form

$$X_i = B \sum_{j=1}^{K} X_{i-j} \mod M \tag{8}$$

where M is prime and $X_1, X_2, \ldots, X_K$ are selected by some random process and B>2. A check through examples have shown that there is a B, for K=2, such that this generator has period M^2-1. It could be the case that the examples investigated are special cases of the generator mentioned in (5). Therefore, this author recommends a more thorough investigation of the theoretical properties of generator (8).

Conclusion

According to the lattice structure test and the fact that every multiplicative generator can be transformed to a mixed generator, there is no point in trying to show that the multiplicative generator is better or worse than the mixed generator. Basically, the generators' performance depend on the selected value of the multiplier. By Theorem 2 page 253 of Marsaglia, the inference is that any good (bad) multiplicative generator module 2^β can be transformed to a good (bad) mixed generator. Beyond the point of selecting a good multiplier, the subject of mixed and multiplicative generators appears to be exhausted.

REFERENCES

Allard, J., Dobell, A., and Hull, T., (1963),"Mixed Congruential Random Number Generators for Decimal Machines," J. Assoc. Conput. Mach., vol. 10, pp. 131-141.

Atkinson, A.C. (1980), "Test of Pseudo-random Numbers," Appl. Statist, vol. 29, No. 2, pp. 164-171.

Brown, M. and Solomon, H. (1976), "On Combining Pseudorandom Number Generators," U.S. Govt. Report AD-A030693/6CA.

Burford, R. (1973), "A Better Additive Congruential Random Number Generator," Decision Sciences, vol. 4, No. 2, pp. 190-193.

Chambers, R. (1967), "Random Number Generation," IEEE Spectrum, vol. 4, part 1, No. 2, pp. 48-56.

Coveyou, R. and MacPherson, R. (1967), "Fourier Analysis of Uniform Random Number Generators," J. Assoc. Comp. Mach., vol. 14, pp. 100-119.

Green, B., Smith, J. and Klem, L. (1959), "Empirical Tests of an Additive Random Number Generator," J. Assoc. Comp. Mach. vol. 6, pp. 527-537.

Greenberger, M. (1961), "An a Priori Determination of Serial Correlation in Computer Generated Random Numbers," Math. Comput., vol. 15, pp. 383-389.

Greenberger, M. (1965), "Method in Randomness," Commun. Assoc. Comput. Mach., vol. 8, pp. 177-179.

Hull, T. and Dobell, A. (1964), "Mixed Congruential Random Number Generators for Binary Machines," J. Assoc. Comput. Mach., vol. 11, pp. 31-40.

Jansson, B. (196), Random Number Generators, Almqvist and Wiksell, Stockholm.

Knuth, D.E., (1969), The Art of Computer Programming: Seminumerical Algorithms, vol. 2, Addison-Wesley.

MacLaren, M.D. and Marsaglia, G.(1965), "Uniform Random Number Generators," J. Assoc. Comput. Mach., vol. 12, pp. 83-89.

Marsaglia, G. (1972), "The Structure of Linear Congruential Sequences," In Applications of Number Theory to Numerical Analysis, S.K. Zaremba (ed.), Academic Press.

Perkins, W. and Menzefricke, U. (1975), "A Better Additive Congruential Random Number Generator," Decision Sciences, vol. 6, pp. 194-198.

Sowey, E.R. (1978), "A Second Classified Bibliography on Random Number Generation and Testing," Internat. Stat. Rev., vol. 46, pp. 89-102.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

ON THE RELATIONSHIP BETWEEN MICROFISSURING AND MICROSTRUCTURE IN THE HAZ OF INCONEL 718

Prepared By: Raymond G. Thompson

Academic Rank: Assistant Professor

University and Department: Clemson University
                          Materials Engineering Program

NASA/MSFC:
    Laboratory: Materials and Processes
    Division: Process Engineering
    Branch: Metals Processes

NASA Counterpart: A. C. Nunes

Date: August 27, 1981

Contract No.:   NGT 01-008-021
                          The University of Alabama in Huntsville

ACKNOWLEDGEMENTS

ON THE RELATIONSHIP BETWEEN MICROFISSURING AND MICROSTRUCTURE IN THE HAZ OF INCONEL 718

BY

Raymond G. Thompson
Assistant Professor of Materials Engineering
Clemson University
Clemson, South Carolina

ABSTRACT

Inconel 718, as well as many other metals, is susceptible to inter-granular hot cracking in the weld heat-affected-zone (HAZ). These cracks form near the solidus temperature of the metal during the welding process. This problem is of particular concern to NASA/MSFC because the SSME is primarily constructed of welded Inconel 718.

The present program studied microfissuring in the weld HAZ of Inconel 718 by simulating HAZ thermal cycles with the Gleeble machine. Previous researchers have studied the instantaneous mechanical properties of the HAZ using the Gleeble. The present study examines the instantaneous microstructure of the HAZ. This approach showed that second phase structures, high in niobium, melt and wet the grain boundaries in the HAZ during welding. It is postulated that the resultant HAZ grain boundaries, enriched in niobium, act as preferred sites for microfissure nucleation as the weld zone cools.

TABLE OF CONTENTS

Page

INTRODUCTION

The intimate relationship between the mechanical properties and the microstructure of a metal has been appreciated for many years. It has also been realized that the microstructure of a metal continuously evolves during the processing and use of the metal. Examples of the evolutionary nature of microstructure and its relationship to mechanical properties are; natural age hardening aluminum alloys, sigma phase embrittlement of certain stainless steels during long time elevated temperature use, hot-work cracking of alloy steel due to the development of low melting temperature sulfur eutectics, and the post-weld strain age cracking of certain nickel base alloys due to the rapid precipitation of the σ phase during cooling. These are only a few examples of how the evolution of microstructure is revealed through its sometimes detrimental effect on mechanical properties. It is this evolution of microstructure during the welding process which the present study relates to microfissuring.

Microfissuring is the occurrence of intergranular cracks near the solidus temperature of the metal immediately after tne heat source (electron beam, arc, etc.) has passed and the weld begins to cool. These cracks are believed to initiate in either the solid state or when incipient intergranular liquid still wets the HAZ grain boundary during cooling. The plastic strain necessary to cause both types of cracking is on the order of 0.1% to 0.5%.[1] Greater strains than these will cause multiple or large cracks. Thus it is imperative to study intergranular microstructures for both the incipient melting condition and the solid state condition in order to understand the relationship between grain boundary microstructural phases and the microfissuring phenomena.

The evolution of mechanical properties in the weld HAZ during the welding process has heen studied using simulated welding thermal cycles.[2] A Gleeble machine is often used to simulate the welding thermal cycle. Most studies have investigated the instantaneous mechanical properties of the HAZ by conducting fracture tests imposed during the simulated thermal cycle. Useful data on tensile strength and fracture ductility have been obtained in this way. The instantaneous microstructures which give rise to the observed mechanical properties during the welding cycle have not yet been studied. The present study is an initial investigation into the evolution of microstructure in the HAZ during the welding process.

EXPERIMENTAL PROCEDURE

A model 510 Gleeble machine was used to simulate a welding thermal cycle such that the specimen experienced a similar microstructural evolution as the HAZ. The method used was similar to that used by Owczarski, Duvall and Sullivan[3] on Inconel 718. This method was used in hopes of making a first order comparison between our microstructural study and their mechanical study.[2] The thermal cycle used in both studies is characteristic of arc welded, 1½ inch thick, stainless steel with an energy input of 70,000 joules per inch.

Small diameter, 0.043 inches, specimens of Inconel 718 with the composition given in Table 1 were held in the water cooled jaws of the Gleeble and subjected to the synthetic HAZ thermal cycle. The small diameter wire was chosen instead of the typical 0.25 inch diameter specimens because the small wire can be quenched more reapidly. The typical HAZ thermal cycle is given in Figure 1. The microstructural evolution in the HAZ was studied by interrupting the thermal cycle at predetermined temperature with a rapid quench technique. This was accomplished by stopping the current flow to the resistance heated specimen and simultaneously quenching with tap water. This procedure resulted in surface quenching rates shown in Figure 1. The thermal cycles in all cases were recorded with a 0.001 inch diameter Chromel-Alumel thermocouple percussively welded to the specimen surface. The evolution of microstructure in the HAZ was studied for both the solution annealed and fully heat treated conditions. These heat treatments are given in Table 2. The solution annealed material was quenched on heating from 2181F, 2275F and 2300F. It was quenced on cooling from a peak temperature of 2300F at 2216F and 2150F. The fully heat treated material was quenched on heating at 2150F, 2235F and 2300F. It was quenched on cooling from the peak temperature of 2300F at 2200F, 2161F and 2050F.

Longitudinal and transverse sections were taken from the specimen midspan for metallographic analysis as shown in Figure 2. These specimens were polished through 0.05 micron alumina and electrolytically etched in ASTM etch 111 at 15 volts for 15 seconds.

RESULTS

Metallography

Figure 4 shows the evolution of microstructure during a simulated HAZ thermal cycle for Inconel 718 in the fully heat treated condition. It can be readily seen on the original photographs that secondary phases melt during the thermal cycle and wet the grain boundaries. Figure 5 shows the evolution of microstructure during a simulated HAZ thermal cycle for Inconel 718 in the solution annealed condition. The high magnification inserts in this figure show the melting of the second phase stringers and their absorption into the grain boundaries.

Chemical Analysis of Microstructure

A Cambridge M4 stereomicroscope equipped with an EDAX was used for the chemical analysis of microstructural phases. All microstructures shown in Figures 4 and 5 had the typical matrix composition given in Table 3. This can be compared with the wet chemical analysis given in Table 1. A high niobium (Nb) concentration was found in the phases observed in all microstructures. Table 3 gives the composition of the stringers found in Figure 5A which is essentially the as received structure. This composition was found to modify as the second phase melts and penetrates the grain boundaries as given in Table 3. A bulky, circular phase was also found which had the compositional range given in Table 3. This was apparently an MC carbide with Nb and Ti freely substituting for each other. An appreciation can be gained for the composition of the various microstructures developed in the HAZ during welding by comparing the numbered phases in Figure 5 with the corresponding compositions given in Table 3.

DISCUSSION

Mechanical Testing

The strength and ductility of Inconel 718 in the HAZ during welding has been studied using the hot ductility test.[3,4] Typical results of these tests (Figure 6) show that Inconel 718 has two distinct stages of low ductility in the HAZ during welding. These are from 2000F to melting during the heating cycle and from melting to 1800F during the cooling cycle. Microfissuring is most likely to occur during the cooling cycle when thermal contraction strain is imposed on the HAZ. Intergranular microfissuring which may occur on cooling has been attributed to low melting eutectic, liquid films on the grain boundaries.[4,5,6] However, recent investigation has shown that a liquid film is not a prerequisite for microfissuring. Instead, the grain boundary microstructure becomes inherently crack sensitive near the solidus temperature. The fact that Inconel 718 shows a reduced ductility at lower temperatures on cooling than on heating (Figure 6) points out that microstructural changes occur in the HAZ during welding. It is this microstructural evolution which is responsible for microfissuring sensitivity in Inconel 718.

Microstructural Evolution

Figures 4 and 5 clearly show that an intergranular microstructural change occurs in the HAZ during a simulated welding thermal cycle. The EDAX chemical analysis and phase morphology seen in the above figures suggest that Laves phases melt in the HAZ during welding and are redistributed into the grain boundaries. It is not known why the molten Laves phase flows into the grain boundaries, but it is certain that this phase forms an intergranular network during the simulated welding cycle. The Laves phase is characterized by a (Nb + Mo) content of 47% before melting. As the molten phase spreads on the grain boundary it is diluted to a (Nb + Mo) content of 28%. This represents a significant increase over the normal (NB + Mo) content of 9%.

The question immediately arises as to whether or not this redistribution of (Nb + Mo), in the form of an intergranular network microstructure, could lead to reduced ductility. The answer is probably yes. Several studies have determined that Laves phases, concentrated in Nb, lead to reduced ductility. Gordine[8] found that if intergranular Laves phases form in weld fusion metal the ductility is reduced. As the Laves is gradually dissolved through solution heat treatment the ductility of the metal is restored. Kaufman and Palty[9] distinguish the Laves phase of Inconel 718 as embrittling and Eiselstein[10] notes a reduction in ductility when Laves phase is present in the microstructure. The above investigations suggest that Laves phases in the HAZ of Inconel 718 will lead to intergranular brittleness and possible microfissuring.

XXXV-3

Figure 4E shows a specimen quenched from 2050F on cooling which fractured due to thermal contraction on cooling. Intergranular micro-cracking can clearly be seen to be associated with the intergranular Nb-rich phase. It is not known at what temperature these cracks formed or if the intergranular phase was molten or solid at the time of rupture. It is obvious that the intergranular brittleness is related to the distribution of the Nb-rich phase in the grain boundaries.

Origin of Laves Phase

In order for the Nb-rich phase to redistribute into the grain boundaries, it must first be present in the forged, cast or wrought product which is to be welded. The processing steps which precede the as-welded condition are typically air arc melting (AM), air induction melting (IM), or vacuum induction melting (VIM) followed by vacuum arc or electroslag remelting. These processing steps are then followed by casting, hot and/or cold working. Figure 7 shows a TTT diagram for phase formation in Inconel 718.[10] The Laves phase will form in about 60 minutes at 1800F. This condition could easily be met in cooling ingots after the primary processing, during cooling after casting and during hot working. The cold working of microstructures containing Laves phases tends to form stringers and distribute the phase randomly throughout the structure. When the phase is thus distributed (such as the as-received material of the present study) it has a minimal effect on mechanical properties. However, after an intergranular network is formed as in Figures 4 and 5 the phase should be considered very harmful.

The detrimental effect of Laves phase has been recognized and attempts are usually made to minimize its formation during all processing steps.[11] Niobium segregation occurs during primary processing and leads to Laves formation in secondary processing steps. The degree of segregation is controlled by solidification rate which controls dendrite formation and by homogenization heat treatment.[12] During casting Laves formation would be controlled by cooling rate and homogenization. The Laves formation during hot work would be controlled by hot working above 1900F and cooling between 1900F and 1600F in less than one hour.[10] The above schedule is approximate and would not work if significant Nb segregation existed before hot working.

SUMMARY

The present study is the first to conclusively show Nb-rich phase redistribution in the HAZ during simulated welds. Thompson[4] observed similar phase redistribution in Inconel 718 but was undecided as to its composition and sequence of development. Bearing in mind the precautions taken to eliminate Laves phases during other processing steps, it would seem prudent to show similar precautions before, during and after welding.

The tendency for microfissuring is greatly enhanced when Laves phases are redistributed onto the grain boundaries of the HAZ during welding. The detrimental effect of intergranular Laves phase is two fold. First, the Laves phase remains molten to temperatures 100F below the normal grain boundary melting temperature (Figure 3). Second, the Laves phase forms a brittle intergranular phase in place of the normally ductile grain boundaries.

CONCLUSIONS

1. An apparently harmless distribution of Laves phase in the form of stringers and randomly distributed particles can be transformed into a dangerous, low ductility intergranular network in the HAZ of Inconel 718 welds.

2. Inconel 718 should be more prone to HAZ microfissuring when Laves phases are redistributed into the grain boundaries. This would be due both to the Laves' low melting point and to its inherent brittleness.

3. The formation of brittle intergranular networks in the HAZ of Inconel 718 welds can probably be suppressed by the appropriate control of alloy microstructure through suitable primary and secondary process techniques.

RECOMMENDATIONS FOR FUTURE RESEARCH

1. It is recommended that all Inconel 718 parts of the SSME which are subject to welding be evaluated to determine if their pre-weld processing promotes or discourages Laves phase formation.

2. It is recommended that pre-weld heat treatments, aimed at dissolving Laves phases, be studied as to their potential effect on the critical service properties of Inconel 718. This could be contrasted to the possible service properties of the HAZ if pre-weld heat treatments are not used.

3. A correlation should be made between the percentage of Laves phase in the as welded metal and the magnitude of HAZ microfissuring in real welds. A similar correlation could be made with service properties.

1. The grain size effect on microfissuring (Appendix C) could be studied in two experiments. Hot ductility tests could be run on four groups of the same metal which differs only by grain size. The grain sizes would be 10^{-3}mm, 10^{-2}mm, 10^{-1}mm and 1mm in diameter. This would evaluate the critical grain boundary displacement theory. A second experiment would evaluate the effect of grain size on the redistribution of Laves phase in the HAZ. Four groups of Inconel 718 would be processed so that they varied only by grain size (grain boundary area). Each group would have the same volume percent of Laves phase. The group would then be exposed to the simulated HAZ thermal cycle to determine the critical grain size for complete grain boundary wetting.

BIBLIOGRAPHY

1. R. G. Thompson, Final Progress Report, NASA Contract # NAS8-33524, 1980.

2. W. A. Owczarski, D. S. Duvall and C. P. Sullivan, Weld. J. 45, 1966, 145-s.

3. D. S. Duvall and W. A. Owczarski, Weld. J. 46, 1967, 423-s.

4. E. G. Thompson, Weld. J. 48, 1969, 70-s.

5. B. Weiss, G. E. Grotke and R. Strickler, Weld. J 49, 1970, 471-s.

6. W. Yeniscavich, Weld. J 45, 1966, 344-s.

7. R. G. Thompson, unpublished research, NASA Contract # NAS8-33805, 1981.

8. J. Gordine, Methods of High Alloy Weldability Evaluation, Proceeding of Welding Research Council Symposium, 1970, pg. 28.

9. M. Kaufman and A. E. Palty, Trans. TMS-AIME 221, 1961, 1253.

10. H. L. Eiselstein, ASTM-STP 369, 1965, 62.

11. R. P. Daykin and A. J. DeRidder, Superalloys - Proceeding of the Second International Conference, 1972, F-1.

12. R. P. DeVries and G. R. Mumau, J. of Metals, 1968, pg. 33.

APPENDIX A

TABLE 1: WET CHEMICAL ANALYSIS OF INCONEL 718

Al	Fe	Cr	Mn	Mo	Ni	V	Ti	Nb	Co	C	S
0.31	13.01	18.47	0.23	3.29	57.8	0.27	0.91	5.12	1.51	0.06	0.003

TABLE 2: HEAT TREATMENT (VACUUM)

Solution Anneal Heat Treatment: 1750F for 1 hour and furnace cool

Full Heat Treatment: 1750F for 1 hour, furnace cool to 1400F

1400F for 8 hours, furnace cool to 1200F

1200F for 10 hours, furnace cool to room temperature

TABLE 3: EDAX ANALYSIS OF MICROSTRUCTURE

SPOT # from Figure 5

A1, B1, C1, D1, B2, C2	4.8	4.3	1.1	53.0	17.5	17.9	0.8
A2, A3	36.7	10.7	3.5	26.1	10.1	9.7	2.8
B3, C3, D2	21.2	7.3	2.2	42.0	12.9	11.6	2.3
B4	64.6	12.2	8.9	5.0	2.2	1.9	4.8
C4	15.3	6.9	57.6	8.9	4.8	4.0	3.5

APPENDIX B

FIGURES
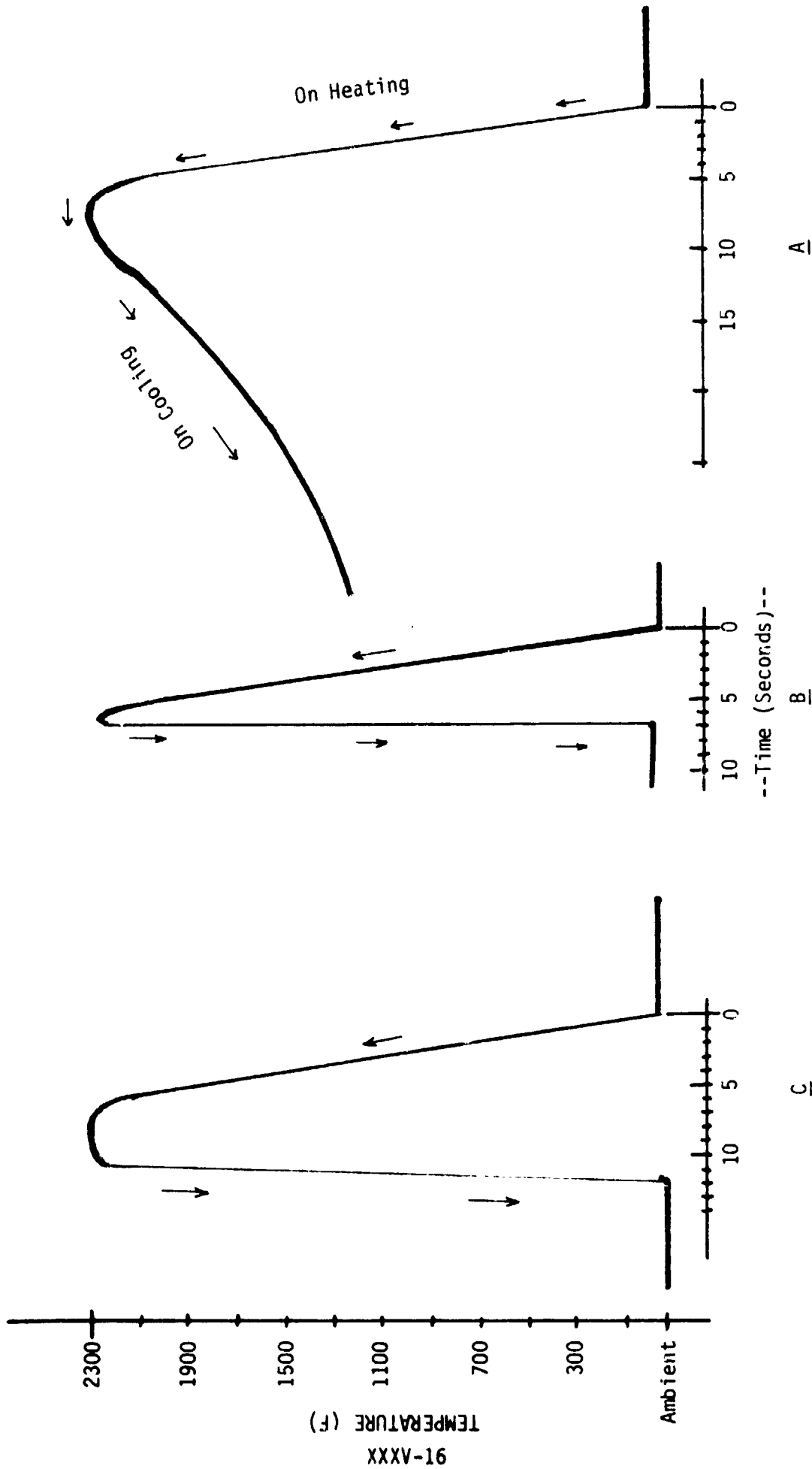
FIGURE 1 - SIMULATED HAZ THERMAL CYCLE AND RAPID QUENCH RESULTS.

A. SIMULATED HAZ THERMAL CYCLE.
B. QUENCH FROM 2250 ON HEATING.
C. QUENCH FROM 2250 ON COOLING.

Thermocouple leads

Gleeble specimen

Transverse section

Longitudinal section
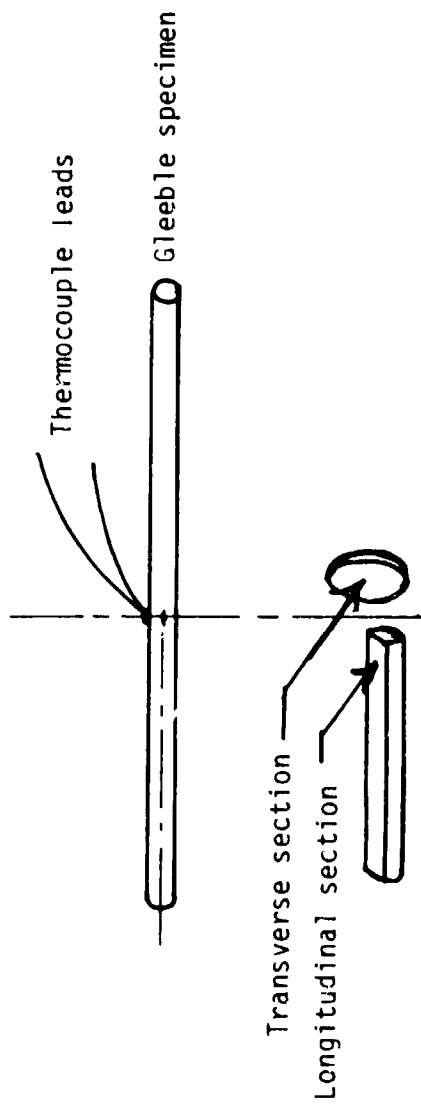
FIGURE 2 - SECTION OF GLEEBLE SPECIMEN FOR METALLOGRAPHIC OBSERVATION

Laves solidus on heating (2250F)
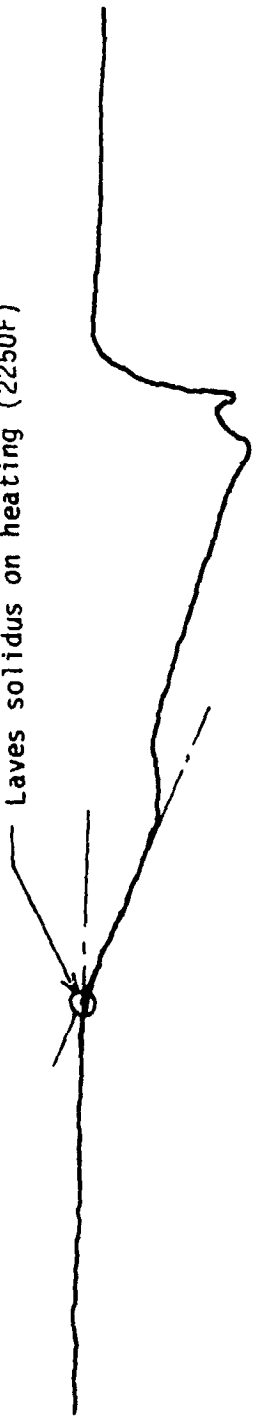
Bulk solidus after remelting (2335F)

1400

00

1200

TEMPERATURE (C)
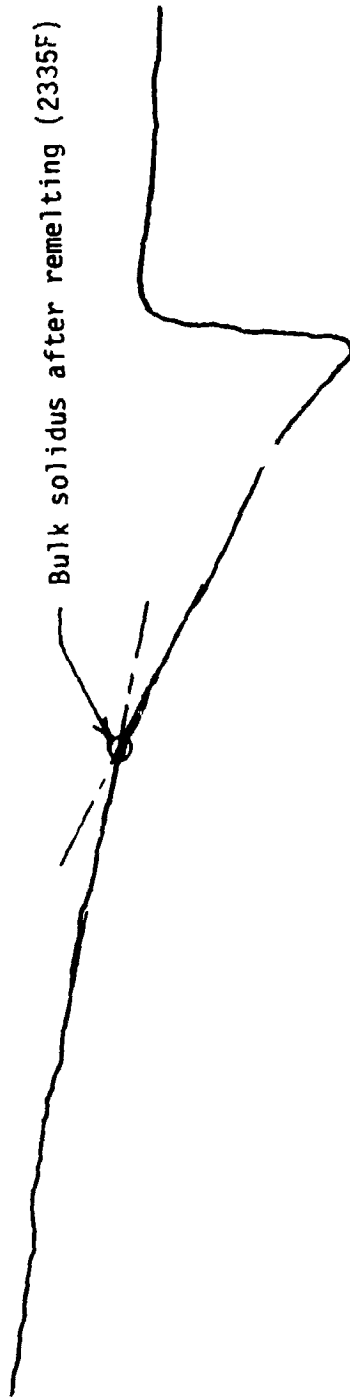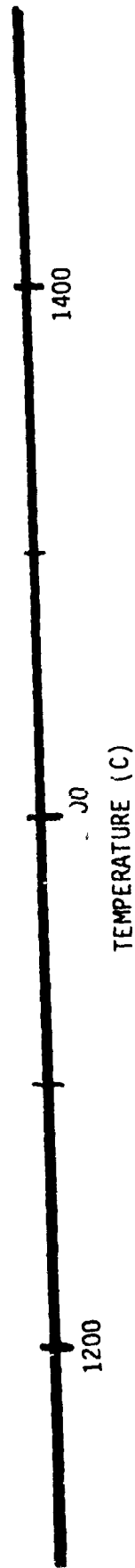
FIGURE 3 - DTA OF INCONEL 718 SHOWING SOLIDUS WITH AND ITHOUT LAVES PHASE PRESENT.

XXXV-18

A. QUENCHED ON HEATING FROM 2150F
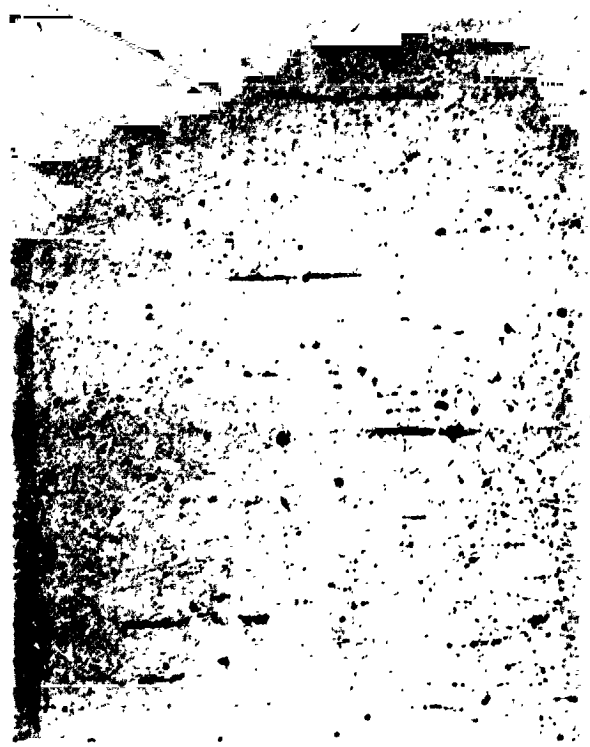
B. QUENCHED ON HEATING FROM 2236F
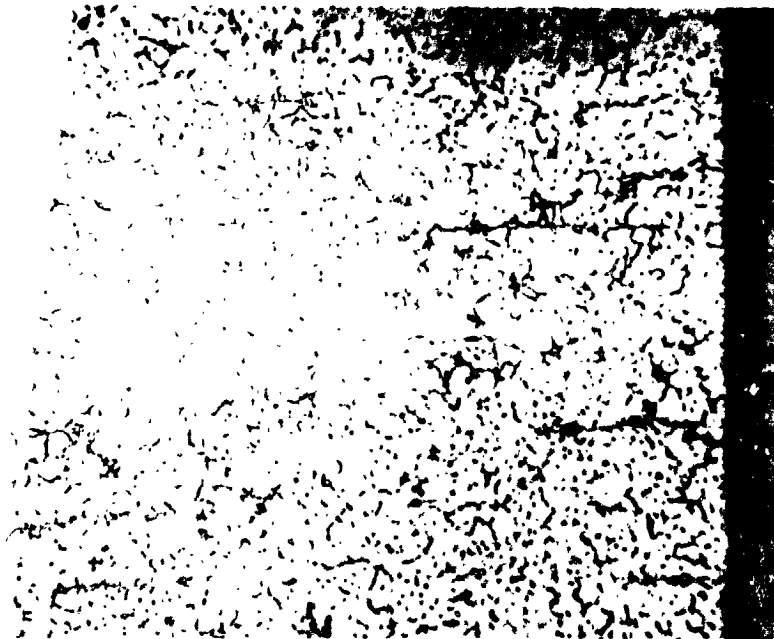
C. QUENCHED ON HEATING FROM 2300F

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

Figure 4: EVOLUTION OF HAZ MICROSTRUCTURE DURING WELDING FOR
THE FULLY HEAT TREATED CONDITION (200 Mag.)
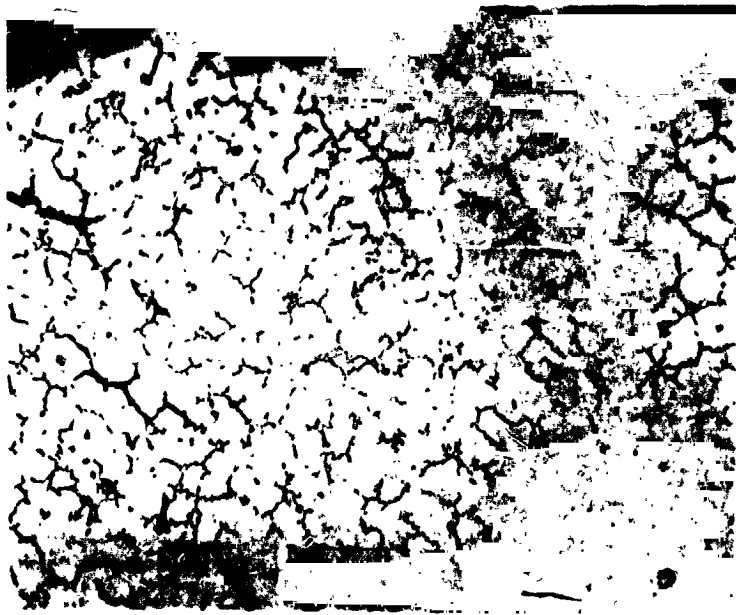
E. QUENCHED ON COOLING FROM 1050F

D. QUENCHED ON COOLING FROM 2200F

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

Figure 4:   EVOLUTION OF HAZ MICROSTRUCTURE DURING WELDING FOR
(cont.)   THE FULLY HEAT TREATED CONDITION (200 Mag.)
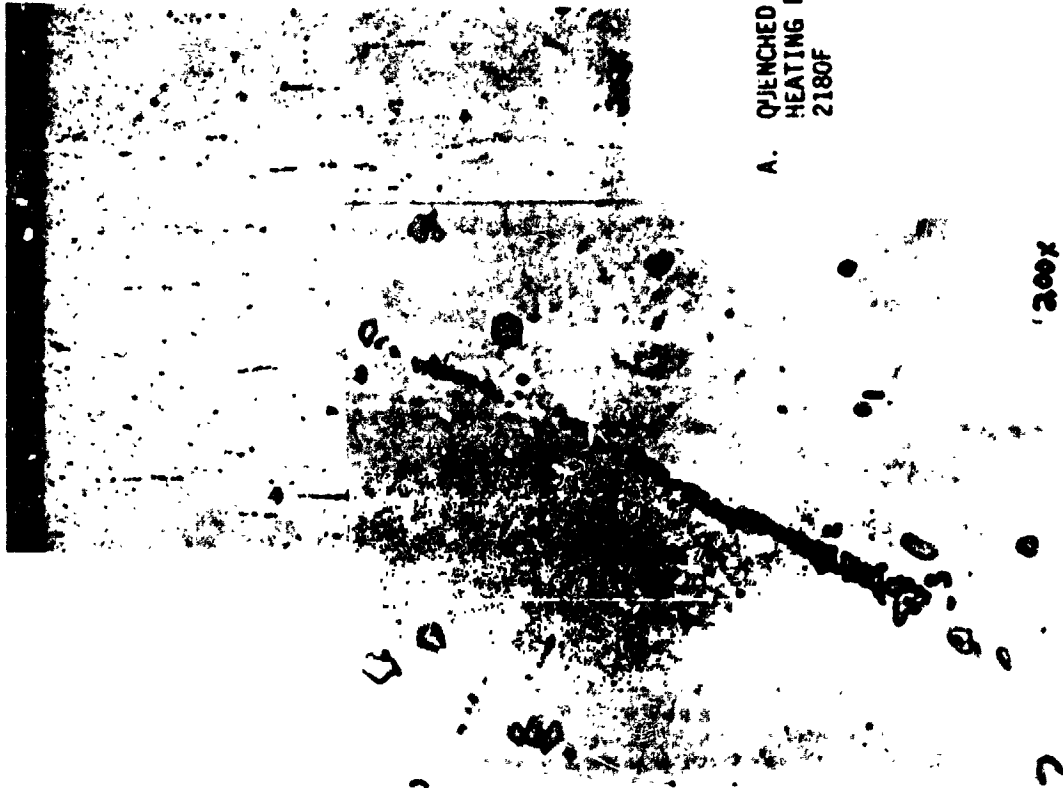
A. QUENCHED ON HEATING FROM 2180F

200X
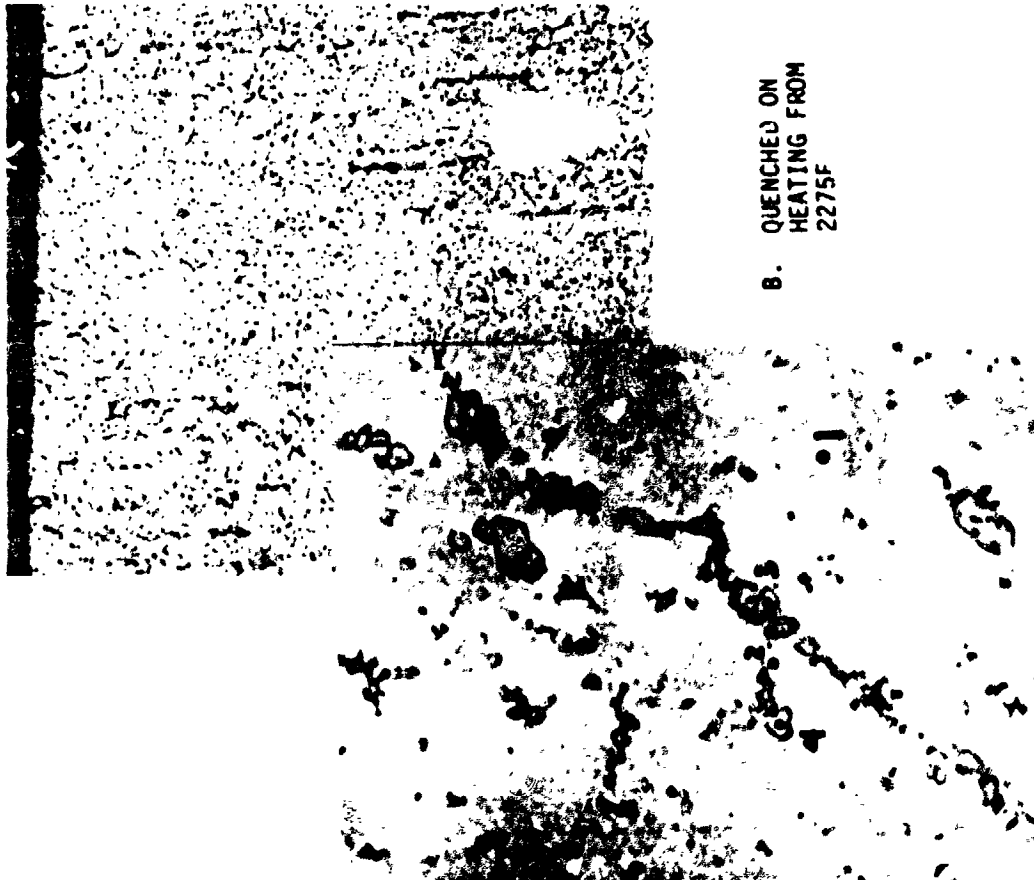
B. QUENCHED ON HEATING FROM 2275F

Figure 5: Evolution of HAZ Microstructure During Welding For The Solution Annealed Condition (200 Mag.); (200 Mag. on Insert)

C. QUENCHED ON
HEATING FROM
2300F
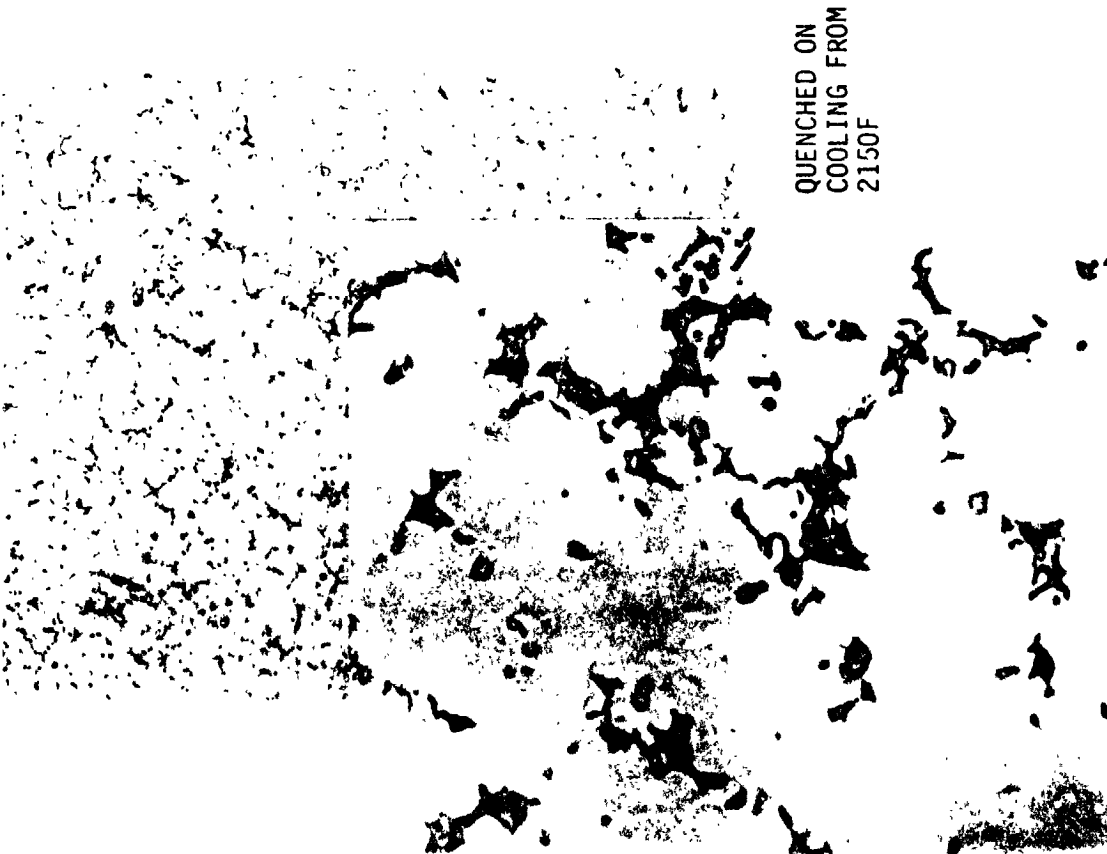
QUENCHED ON
COOLING FROM
2150F

Figure 5: Evolution of HAZ Microstructure During Welding For
The Solution Annealed Condition (200 Mag.) (1200 Mag. on Insert)
(cont.)
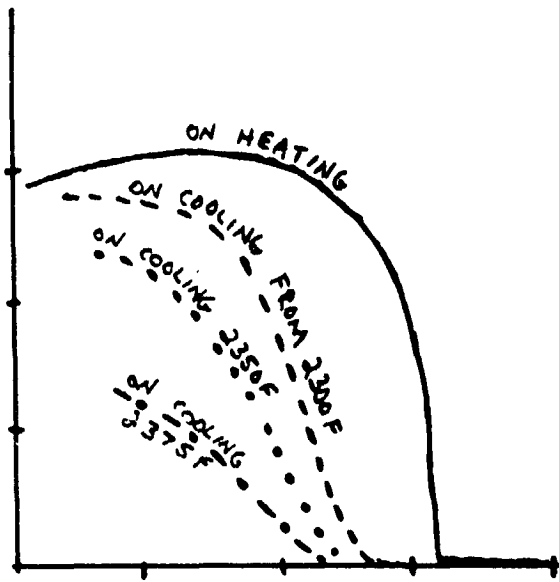
FIGURE 6 - TYPICAL HOT DUCTILITY FOR INCONEL 718.[3]
EXACT CURVE WILL VARY FROM HEAT TO HEAT AND
WITH PROCESSING TECHNIQUES.
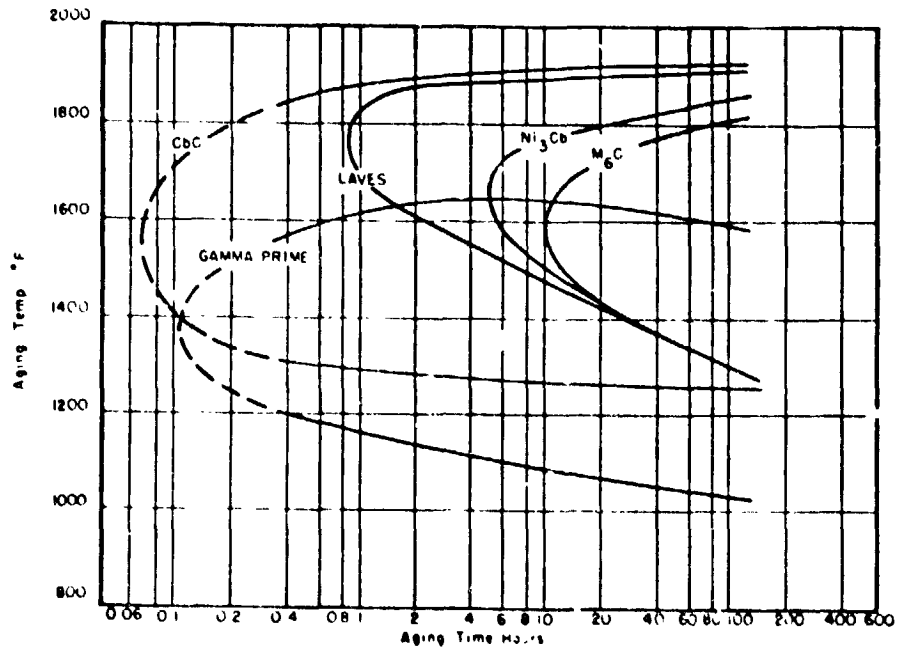
FIGURE 7 - TTT DIAGRAM FOR INCONEL 718 AFTER ANNEALING AT 2100F
FOR ONE HOUR AND WATER QUENCHED.[10]

XXXV-23

APPENDIX C

The Grain Size Effect in Microfissuring

It is widely known that large grain size in the as-welded metal promotes HAZ microfissuring while small grain size prevents HAZ microfissuring. Many hypotheses have been suggested to explain this phenomenon. The present author finds the following explanations most satisfying based on his personal research.

Grain boundaries are known to suffer high temperature creep due to accelerated grain boundary diffusion and dislocation flow. Large grain sizes are more creep resistant because they have less grain boundary area to aid deformation. Small grain sizes are super plastic because of their large grain boundary area contributing to plastic flow. By analogy, large grain size metals will be inherently more prone to micro-fissuring because they do not have enough area to distribute the imposed plastic strain. At temperatures near the melting point the grain boundary becomes increasingly more ductile compared to the grain so that shear is confined to the very small g in boundary volume which gives the appearance of "brittle" failure. This is in contrast to creep tests run at lower temperatures, relative to the melting point, where the grain is able to plastically shear and follow the flow of the grain boundary. Under microfissuring conditions the grain boundary area of the small grain size is able to plastically deform and absorb the thermal contraction strain responsible for microfissuring.

This suggests that for a given microstructure and temperature of deformation there may be a critical grain boundary displacement ΔX_c which would result in microfissuring. Given a grain diameter d the critical strain ϵ_c would be roughly

$$\epsilon_c = \frac{\Delta X_c}{d}$$

Further, for grain size variations without changes in grain boundary structure the critical strain resulting in microfissuring would be expected to be inversely proportional to grain size.

A second grain size effect is seen in the present investigation. As the HAZ reaches the Laves melting temperature, the liquid wets the grain boundary. A large grain size (small grain boundary area) allows for more complete intergranular wetting by the liquid. A small grain size requires more liquid and discourages complete wetting. Thus, for a given volume of Laves' phase, the microfissuring tendency will increase with an increase in percent of intergranular wetting which is promoted by large grain size.

1981

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

HEAT LOSS AND DRAG OF SPHERICAL DROP TUBE SAMPLES

Prepared By:	Donald B. Wallace, Ph.D., P.E.
Academic Rank:	Associate Professor
University and Department:	The University of Alabama in Huntsville
	Department of Mechanical Engineering
NASA/MSFC:	
(Laboratory)	Space Science
(Division)	Space Processing
(Branch)	Optical Physics
MSFC Counterpart:	Robert J. Naumann, Ph.D.
	Michael B. Robinson
Date:	August 7, 1981
Contract Number:	NGT 01-008-021 The University of Alabama in Huntsville

XXXVI

HEAT LOSS AND DRAG OF SPHERICAL DROP TUBE SAMPLES

By

Donald B. Wallace, Ph.D., P.E.
Associate Professor of Mechanical Engineering
The University of Alabama in Huntsville
Huntsville, Alabama 35899

ABSTRACT

This study investigates and proposes analysis techniques for three aspects of the performance of the NASA/MSFC 32 meter drop tube:

    (a) Heat loss through the support wire in a pendant drop sample.

    (b) Temperature history of a drop falling through the drop tube when the tube is filled with helium gas at various pressures.

    (c) Drag and resulting g-levels experienced by a drop falling through the tube when the tube is filled with helium gas at various pressures.

The developed methods for (b) and (c) apply to systems with sufficiently small Knudsen numbers for which continuum theory may be applied.

Sample results are presented, using niobium drops, to indicate the magnitudes of the effects. Helium gas at one atmosphere pressure can approximately double the amount of possible undercooling but it results in an apparent gravity levels of up to 0.1 g.

ACKNOWLEDGEMENTS

NOMENCLATURE

a — acceleration (cm/s^2)

A — sphere surface area $= \pi D^2$ (cm^2)

A_p — projected sphere surface area $= \pi D^2/4$ (cm^2)

$B_1(T)$ — function of temperature, see equation (9)

$B_2(T)$ — function of temperature, see equation (11)

c — heat capacity $(kJ/kg\ K)$

C — constant

C_D — drag coefficient

d — diameter of wire (cm)

D — diameter of sample drop (cm)

F — drag force on sphere $(g\ cm/s^2)$

g — acceleration of gravity $(980\ cm/s^2)$

h — convective heat transfer coefficient $(W/cm^2\ K)$

k — thermal conductivity $(W/cm\ K)$

K_n — Knudsen number $= \lambda/D$

m — mass (g)

Nu — Nusselt number $= hD/k$

p — gas pressure $(torr)$

p_c — critical pressure of a gas $(torr)$

Pr — Prandtl number $= c\mu/k$

q — heat flow rate in the wire (W)

Q — heat loss reat from sample drop (W)

Re — Reynolds number $= \rho VD/\mu$

t — time (s)

T — temperature (K)

T_c –   critical gas temperature   (K)

T_o –   ambient (tube) temperature   (K)

V –   velocity   (cm/s)

X –   position   (cm)

ΔX –   incremental length   (cm)

Z –   dummy variable $Z = \dfrac{\partial T}{\partial x}$

β –   coefficient of sliding friction

ϵ –   total hemispherical emissivity

ϵ_w –   total hemispherical emissivity of wire material

λ –   mean free path of gas molecules   (cm)

μ –   absolute viscosity   (g/cm s)

π –   pi = 3.14159...

ρ –   density   (g/cm^3)

σ –   Boltzmann's constant = 5.67×10^{-12} W/cm^2 K^4

TABLE OF CONTENTS

INTRODUCTION

With the advent of space travel in the past decade, the interest in low gravity/containerless processing of materials has dramatically increased. Experiments have been performed on Skylab [1,2]* and on the joint space mission Apollo-Soyuz Test Project [2,3]. A recent book by Naumann [4] reviews work that has been recently done in these areas.

The use of the space environment (low gravity) has three major consequences on the solidification process:

(1) Containerless processing - this allows the molten material to be free of contamination from the container and to be free from external nucleation sites for crystal growth (this permits a molten material to significantly undercool before solidifying).

(2) Reduction of natural convective flows - this permits a sample to solidify without a forced mixing due to density variations with temperature changes.

(3) Elimination of sedimentation and Stokes flow - a more dense phase will not have a tendency to drop to the bottom of the melt.

The disadvantages of space experimentation and space processing include the expense and long lead times required for scheduling on a space flight. Several alternatives are availabel for low-g experiments which require only short periods of low gravity time. The alternatives include: airplanes flying parabolic trajectories, drop towers, and drop tubes.

A NASA KC-135 airplane can simulate low gravity environments for periods of approximately 30 seconds by essentially free falling in a parabolic trajectory at altitudes of approximately ten kilometers. Typical missions of two to three hours will consist of flying thirty to fourty low gravity parabolas in roller coaster fashion. The flights are manned so that experiments can be modified between runs.

Drop towers and drop tubes also simulate low gravity with free falls but they are limited in the vertical distance available. A 30 meter free fall takes about 2.5 seconds and a 100 meter free fall takes about 4.5 seconds. A drop tower drops an apparatus which contains the experiment as well as any necessary supporting equipment (such as furnaces, quenching media, and temperature monitoring devices). The drop tower experiments are low gravity but not necessarily containerless. NASA operates a 100 meter drop tower at the Marshall Space Flight Center (Huntsville, AL).

A drop tube allows the material under investigation to free fall down a tube while monitoring it from fixed stations. This paper deals with the analysis of drop tube samples, specifically the samples used in the 32 meter drop tube located at NASA/MSFC in Huntsville, Alabama.

---

* Numbers in brackets refer to references at the end of the paper

NASA / MSFC 32 Meter Drop Tube Facility

A schematic of a 32 meter drop tube facility is shown in Figure 1. The containerless melting apparatus is enclosed in a stainless steel bell jar which is located directly over a 10 centimeter inside diameter stainless steel tube. The bell jar and tube assembly can be evacuated by two roughing pumps (RP_1 and RP_2) and two turbomolecular pumps to a typical pressure of 2 mPa (1×10^{-5} torr). Various instrumentation ports (I_1, I_2, ... , I_6) and view ports (VP_1, ... , VP_6) are loacted on each floor level. Electro-pnemumatic (EV_1 and EV_2) and hand operated vacuum valves (V_1, ... , V_3) maintain the vacuum conditions within the tube during retrieval of a sample or insertion of a new sample. Helium gas can be backfilled into the tube to provide for increased cooling of low temperature metals or alloys.

The containerless melting apparatus [5] is based upon the pendant drop technique [6] using omnidirectional electron bombardment. A schematic of the vacuum melting apparatus is shown in Figure 2. The bombarding electrons are supplied by a hot circular cathode which is maintained at a high negative potential. The focusing grid, which is torroidal in shape, is maintained at the same negative potential as the cathode and helps concentrate the electrons onto the sample surface. The cylindrical equi-potential grid helps prevent electron loss to the bell jar assembly and prevents heating of the support wire. The sample is maintained at the ground potential and is suspended from a support wire preferably made out of the major constituent element (i.e. niobium for niobium alloys). The sample can be raised or lowered into the heating apparatus by means of a gear box which is controlled from outside the bell jar using a magnetic-coupled feed-thru. The emission current striking the sample is accurately measured by electrically isolating the support assembly and letting the current flow to the ground potential through an isolated ammeter circuit. The current striking the sample, and consequently the heating power, can be accurately controlled by using a ten-turn potentiometer located on the power supply or by the position of the sample in the melting apparatus. The brightness temperature of the sample can be continuously monitored and digitally recorded by an automatic fast-response (0.1 second) pyrometer. The true thermal history of the sample, during the heating and melting operation, is calculated once the spectral emissivity of the sample is determined.

To study undercooling and low-gravity containerless solidification of various high-temperature metals and alloys, a drop tube apparatus can complement orbital or sub-orbital space experiments, and in two areas drop tubes offer advantages over space experiments:

(a) Economy of construction and operation, allowing numerous samples to be prepared with various parameters and with a low cost per sample.

(b) Flexibility to change parameters during the experiment along with a short turn-around time to capitalize on unexpected results.

The drop tube has several limitations when compared to the other methods of attaining low gravity:
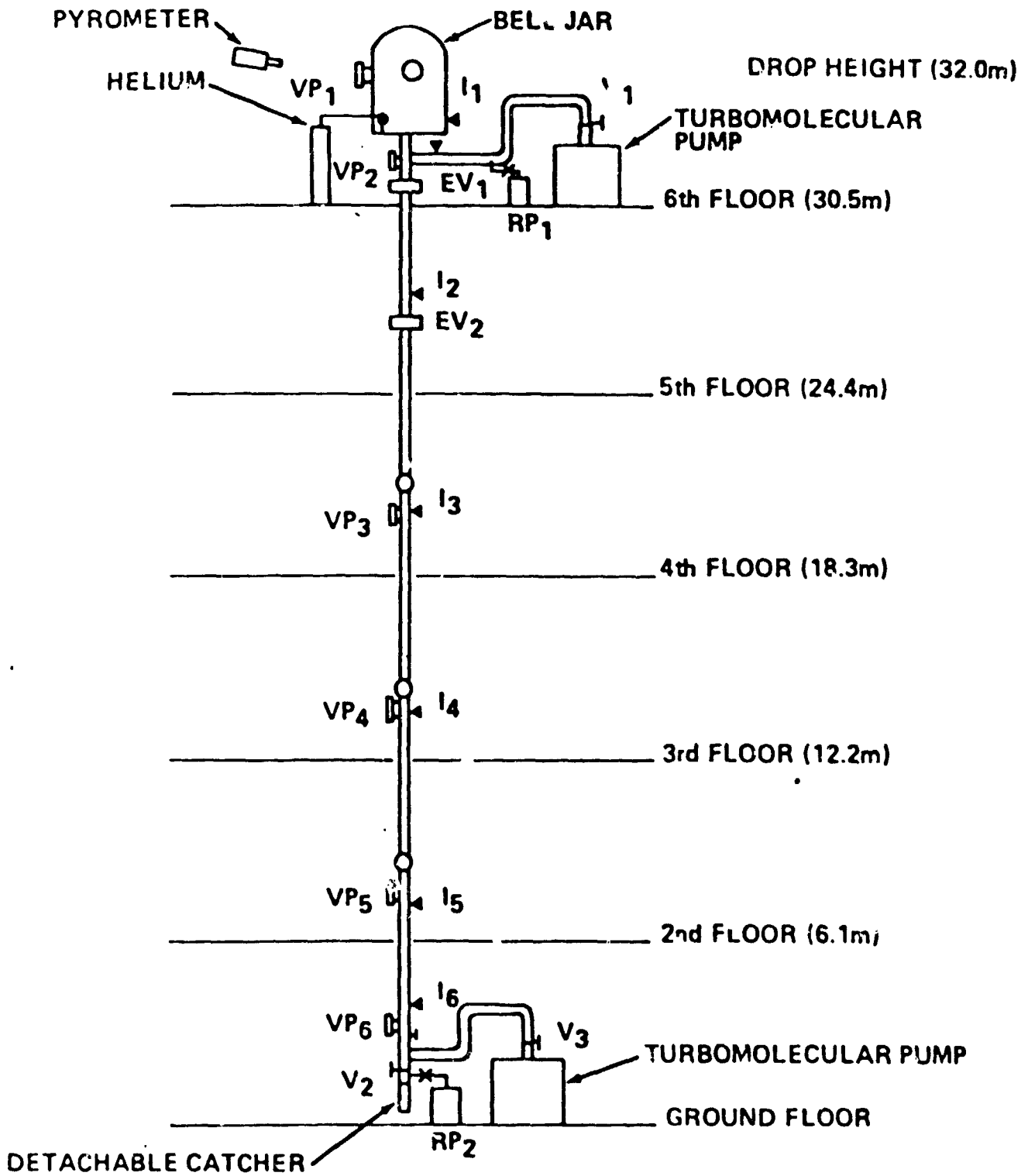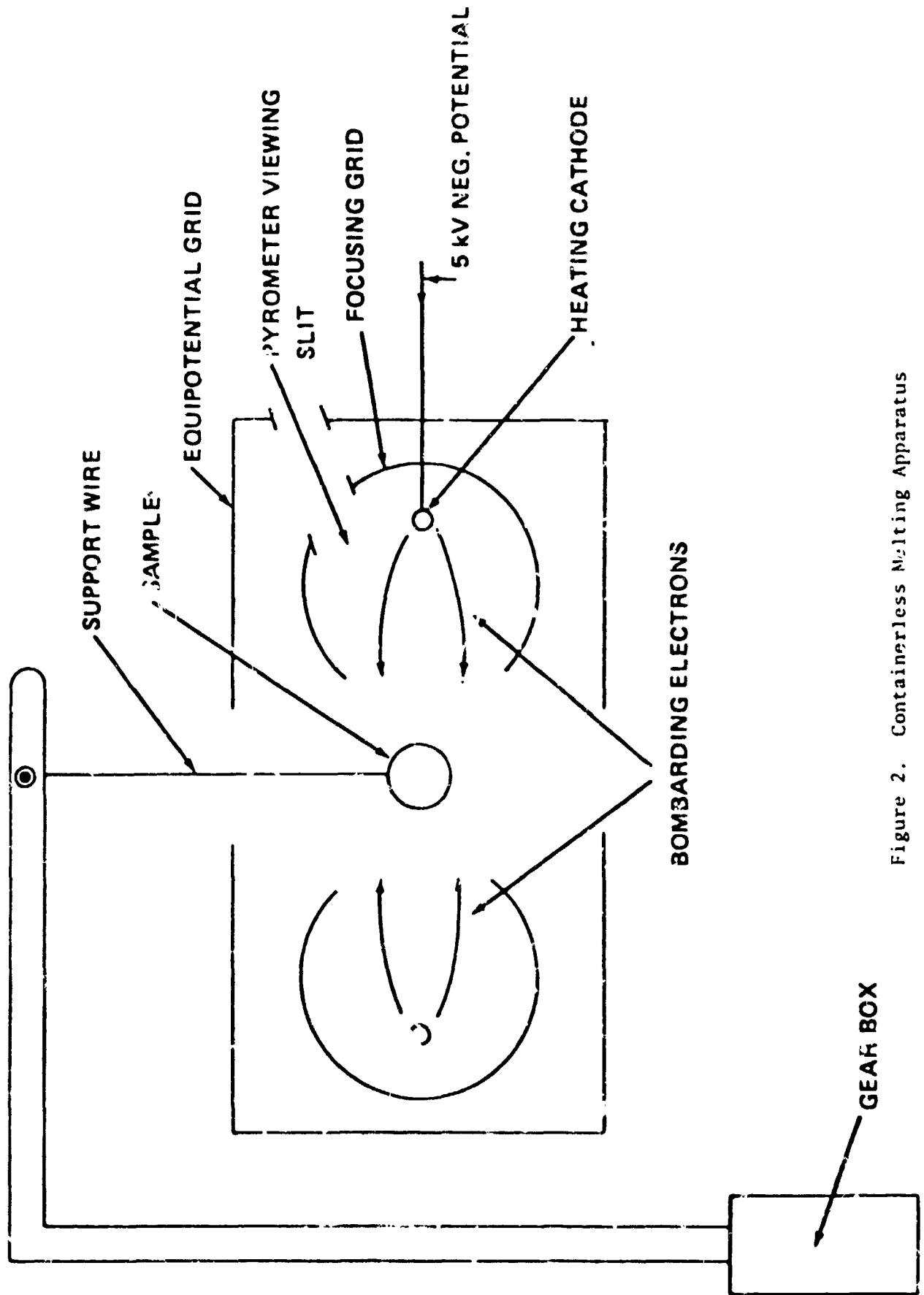
Figure 1. NASA/MSFC 32 m Drop Tube Facility

Figure 2. Containerless Melting Apparatus

Labels in figure:
SUPPORT WIRE
SAMPLE
EQUIPOTENTIAL GRID
PYROMETER VIEWING SLIT
FOCUSING GRID
5 kV NEG. POTENTIAL
HEATING CATHODE
BOMBARDING ELECTRONS
GEAR BOX

(a) Due to limited vertical distances, times during which low gravity exists are limited to several seconds.

(b) Resulting from the first limitation, relatively high melting point materials must be used so that sufficient heat transfer may take palce during the fall to solidify the sample or attain sufficient undercooling. Helium gas may be placed in the tube to increase the heat loss rate [7].

(c) The sample is small (typically 3 to 5 mm in diameter) and moving at high velocities (about 24 meters per second after 30 meters of free fall) making accurate temperature histories difficult to measure.

(d) Accurate high temperature properties (e.g. heat capacity and spectral missivity) for foth solids and liquids are not generally available for use in supporting analytical models.

The containerless melting apparatus at the top of the drop tube can be used to measure the high temperature properties by performing a heat balance on the pendant drop [8]. The accuracy of the properties depend on the determination of the heat loss to the supporting wire.

OBJECTIVES

The overall objective of this study was to review and refine the thermal analysis models used to simulate the drop tube performance. Specific objectives were:

(a) Refine the estimate of heat loss up the pendant drop wire.

(b) Refine the heat transfer model for molten metal drops falling in helium gas at various pressures.

(c) Determine the effective g-level experienced by a falling drop due to the drag forces of the gas.

C - 8

THERMAL LOSS IN A PENDANT DROP SUPPORT WIRE

While the evaluation of high temperature properties of a material is being conducted in the containerless melting apparatus (Figure 2), the sample remains suspended from the support wire. This wire acts like a cooling sting in that heat can be conducted up this support wire and then radiated away. Accurate material characteristics can be determined for the sample only if this heat loss can be closely predicted.

A model for the support wire is shown in Figure 3 along with a standard heat balance for a typical element [9,10]. The heat balance equation is (assuming that no radiant energy is received directly from the pendant drop sample):

Heat Stored = Heat Conduction In - Heat Conduction Out
  - Heat Radiated Out

$$m c \frac{\partial T}{\partial t} = - k \frac{\pi d^2}{4} \frac{\partial T}{\partial x} - \left[-k \frac{\pi d^2}{4} \left(\frac{\partial T}{\partial x} + \frac{\partial^2 T}{\partial x^2} \Delta x \right) \right]$$

$$- \epsilon_w (\pi d \Delta x) \sigma (T^4 - T_0^4)$$

Substituting the fact that:

$$m = \rho \frac{\pi d^2}{4} \Delta x$$

and simplyfing, the heat balance equation becomes:

$$\rho c \frac{\pi d^2}{4} \Delta x \frac{\partial T}{\partial t} = k \frac{\pi d^2}{4} \frac{\partial^2 T}{\partial x^2} \Delta x - \epsilon_w \sigma \pi d \Delta x (T^4 - T_0^4)$$

or

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c} \frac{\partial^2 T}{\partial x^2} - \frac{4 \epsilon_w \sigma}{\rho c d} (T^4 - T_0^4) \qquad ---(1)$$

Equation (1) is the governing equation provided that the wire is in a vacuum.

By letting the time variation in temperature (the left side of Equation (1)) go to zero, the steady state equation is developed as:

$$\frac{\partial^2 T}{\partial x^2} = \frac{4 \epsilon_w \sigma}{k d} (T^4 - T_0^4) \qquad --- (2)$$

Solution - Temperature Invariant Material Properties

A solution to Equation (2) may be found by making the following assumptions:

(1) The wire is semi-infinite.

(2) Thermal expansion of the wire is negligible.

(3) The material properties do not vary with temperature (i.e. emissivity and thermal conductivity are independent of temperature).
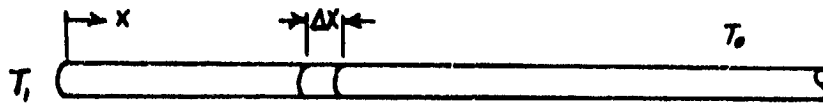
(a) Model of entire support wire
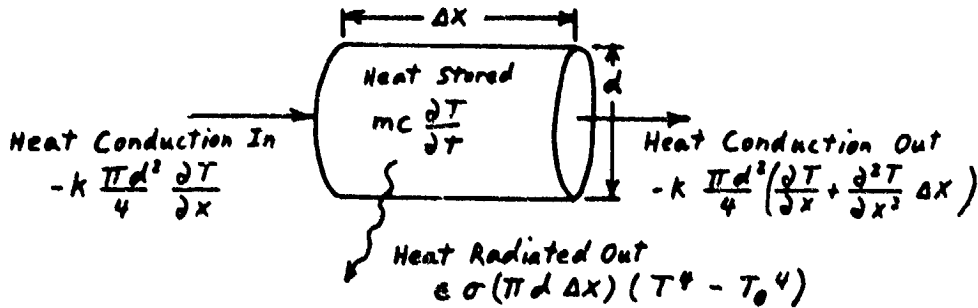
(b) Model of a segment of wire

Figure 3. Model of Support Wire

Let $\qquad \mathbf{z} \equiv \dfrac{\partial T}{\partial x}$

Therefore

$$\frac{\partial^2 T}{\partial t^2} = \frac{\partial z}{\partial x} = \frac{\partial z}{\partial T}\frac{\partial T}{\partial x} = z\frac{\partial z}{\partial T}$$

Substituting in equation (2):

$$z\frac{\partial z}{\partial T} = \frac{4\epsilon_w \sigma}{k d}\left(T^4 - T_J^4\right)$$

$$z\,\partial z = \frac{4\epsilon_w \sigma}{k d}\left(T^4 - T_o^4\right)dT$$

Integrating both sides (since emissivity and thermal conductivity are not functions of temperature):

$$\frac{z^2}{2} = \frac{4\epsilon_w \sigma}{k d}\left(\frac{T^5}{5} - T T_o^4\right) + C$$

where C is the constant of integration. Using the boundary condition that the wire asymptotically approaches the ambient temperature:

when $\qquad T = T_o \qquad \dfrac{\partial T}{\partial x} = 0 = z$

The constant of integration is evaluated and the resulting equation is:

$$\frac{z^2}{2} = \frac{4\epsilon_w \sigma}{k d}\left(\frac{T^5}{5} - T T_o^4 + \frac{4 T_o^5}{5}\right)$$

Solving for z defines the temperature gradient as:

$$\frac{\partial T}{\partial x} = \left[\frac{8\epsilon_w \sigma}{k d}\left(\frac{T^5}{5} - T T_o^4 + \frac{4 T_o^5}{5}\right)\right]^{\frac{1}{2}} \qquad -(3)$$

Heat conducted across any cross section of the wire is given by:

$$q = - k \frac{\pi d^2}{4} \frac{\partial T}{\partial x} \qquad ---(4)$$

Combining equations (3) and (4) and simplyfing the results yields:

$$q = \pi \left[.1 \epsilon_w \sigma k d^3 \left(T^5 - 5 T T_0^4 + 4 T_0^5 \right) \right]^{1/2} \qquad ---(5)$$

Equation (5) can be used to predict the heat loss up the wire provided that T is taken as the drop temperature and that the previously defined assumptions are valid.

Solution - Temperature Variant Material Properties

During the analysis of the molten drop it is difficult to prevent the drop from dissolving enough of the wire that is supporting it to reach a critical mass and fall. For this reason tungsten is often used as the support wire due to its high melting point.

Material characteristics of tungsten [10] are shown in Figure 4. These curves show that the assumption of temperature invariant material properties is not valid. The ratio of emissivity to thermal conductivity is not constant either (as shown in Figure 4(b)). A fourth order polynomial curve fit to the data in Figure 4(b) yields the following expression:

$$\frac{\epsilon_w}{k} = a_4 T^4 + a_3 T^3 + a_2 T^2 + a_1 T + a_0 \qquad ---(6)$$

$$\text{where}: a_4 = 4.96 \times 10^{-15}$$
$$a_3 = -5.16 \times 10^{-11}$$
$$a_2 = 1.66 \times 10^{-7}$$
$$a_1 = -4.55 \times 10^{-5}$$
$$a_0 = 2.14 \times 10^{-2}$$

Using the polynomial approximation of the ϵ_w/k ratio, equation (6), in the steady state equation, equation (2), results in:

$$\frac{\partial^2 T}{\partial x^2} = \frac{4 \sigma}{d} \left(a_4 T^4 + a_3 T^3 + a_2 T^2 + a_1 T + a_0 \right)\left(T^4 - T_0^4 \right)$$

Solving this equation in a manner similar to the previous section yields a temperature gradient equation:

$$\frac{\partial T}{\partial x} = \pm \left\{ \frac{8\sigma}{d} \left[a_4 \left(\frac{T^9}{9} - \frac{T^5 T_0^4}{5} + \frac{4 T_0^9}{45} \right) + a_3 \left(\frac{T^8}{8} - \frac{T^4 T_0^4}{4} + \frac{T_0^8}{8} \right) + a_2 \left(\frac{T^7}{7} - \frac{T^3 T_0^4}{3} + \frac{4 T_0^7}{21} \right) \right. \right.$$
$$\left. \left. + a_1 \left(\frac{T^6}{6} - \frac{T^2 T_0^4}{2} + \frac{T_0^6}{3} \right) + a_0 \left(\frac{T^5}{5} - T T_0^4 + \frac{4 T_0^4}{5} \right) \right] \right\}^{1/2} \qquad ---(7)$$

where the coefficients $a_0, a_1, \ldots, a_4$ are defined in equation (6) for tungsten properties. The heat conduction equation is still defined by equation (4) where the thermal conductivity, k, is evaluated at the
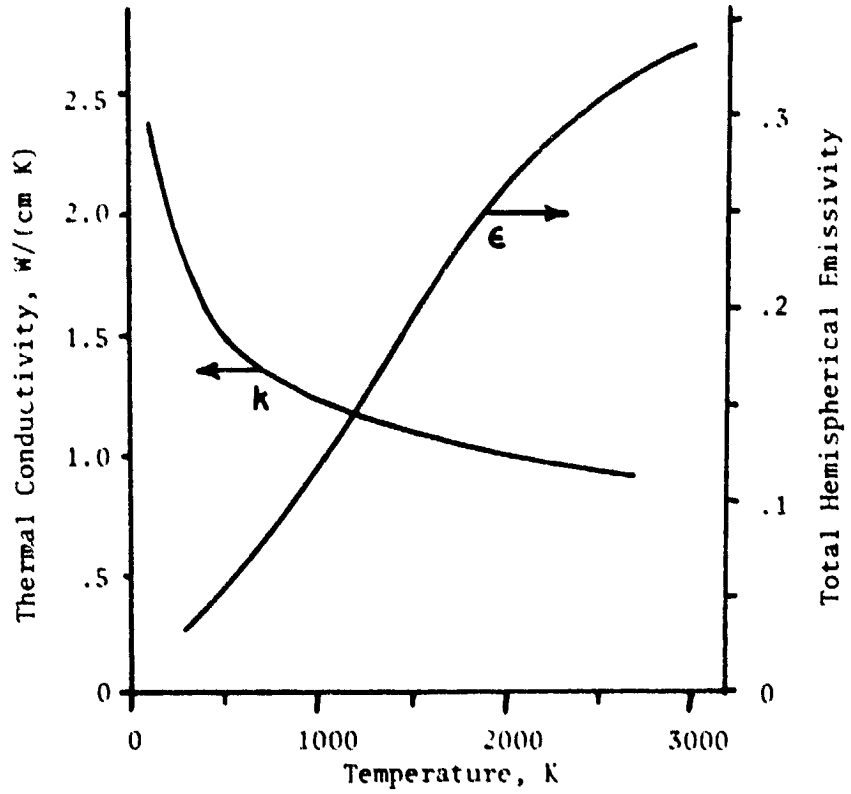
(a) Thermal Conductivity **(k)** and Total Emissivity **(ε_w)**
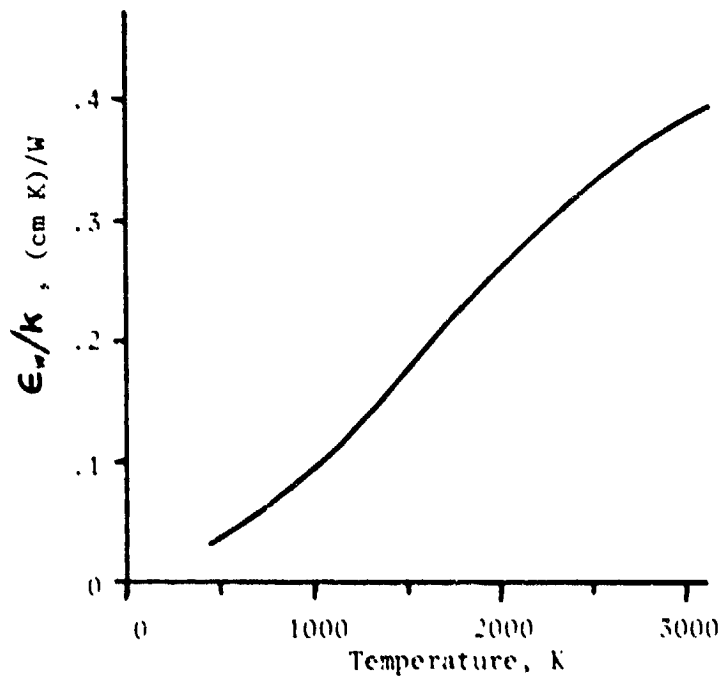
(b) **ε_w/k**

Figure 4. Tungsten Properties

temperature, T, of the appropriate cross section. To evaluate the heat loss om a pendant drop sample, the cross section adjacent to the drop is cl use. so that the temperature used in evaluating equations (4) and (7) is the sample temperature.

Equations (4) and (7) are combined and rewritten as:

$$\frac{q}{d^{1.5}} = \pi k \left\{ \frac{\sigma}{2} \left[a_4 \left(\frac{T^9}{9} - \frac{T^5 T_o^4}{5} + \frac{4 T_o^9}{45} \right) + a_3 \left(\frac{T^8}{8} - \frac{T^4 T_o^4}{4} + \frac{T_o^8}{8} \right) + a_2 \left(\frac{T^7}{7} - \frac{T^3 T_o^4}{3} + \frac{4 T_o^7}{21} \right) \right. \right.$$
$$\left. \left. + a_1 \left(\frac{T^6}{6} - \frac{T^2 T_o^4}{2} + \frac{T_o^6}{3} \right) + a_o \left(\frac{T^5}{5} - T T_o^4 + \frac{4 T_o^4}{5} \right) \right] \right\}^{1/2} \quad —(8)$$

or

$$\frac{q}{d^{1.5}} = B_1(T) \quad\quad —(9)$$

The temperature dependent function $B_1(T)$ in equation (9) represents the right hand side of equation (8) since it is a function of temperature alone. If the wire is tungsten then the coefficients $a_4, \ldots, a_o$ are defined in equation (6). The temperature variation in the thermal conductivity of tungsten is obtained by curve fitting a fourth order polynomial to the data in Figure 4(a) [11] and the resulting equation is:

$$k = 7.92 \times 10^{-14} T^4 - 5.83 \times 10^{-10} T^3 + 1.59 \times 10^{-6} T^2 - 2.10 T + 2.22$$

The heat loss rate from the pendant drop sample itself is given by the equation:

$$Q = (\pi D^2) \epsilon \sigma (T^4 - T_o^4) \quad\quad (10)$$

This equation assumes that radiation occurs equally from the entire surface area of the drop to a gray body enclosure at temperature T_o (whose emissivity is unimportant [12]). This equation may be written as:

$$\frac{Q}{\epsilon D^2} = \pi \sigma (T^4 - T_o^4)$$

Dividing equation (9) by this equation results in:

$$\left(\frac{q}{Q} \right) \left(\frac{\epsilon D^2}{d^{1.5}} \right) = \frac{B_1(T)}{\pi \sigma (T^4 - T_o^4)} = B_2(T) \quad\quad (11)$$

The right hand side of equation (11) is also just a function of temperature so it is represented by the function $B_2(T)$. Note that these are functions of temperature alone because the material properties of thermal conductivity (k) and total hemispherical emissivity (ϵ_w) of the wire were defined as functions of temperature.

The relationships represented by equations (9) and (11) are plotted in Figure 5. The abscissa is absolute temperature. The functions $B_1(T)$ and $B_2(T)$ are plotted so that the left hand sides of equations (9) and (11) may be determined graphically if the support wire is tungsten.
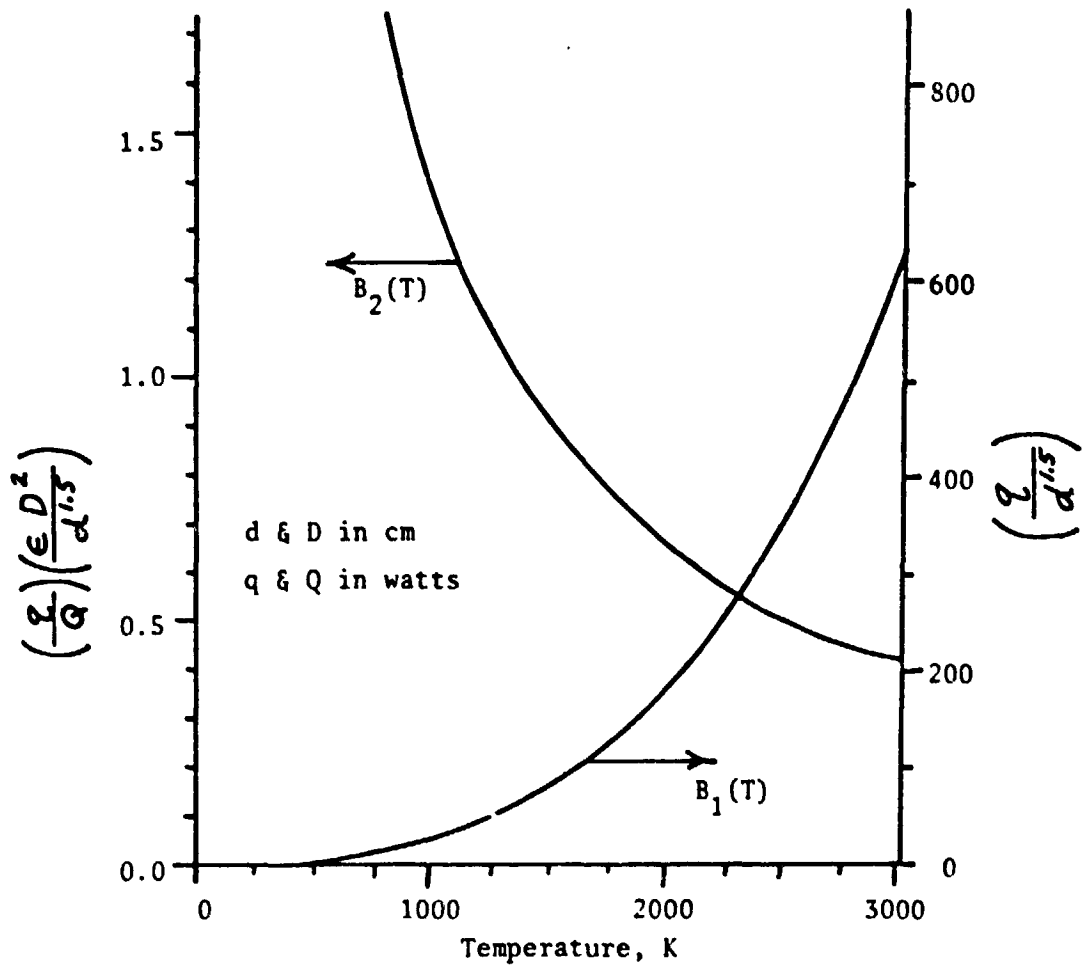
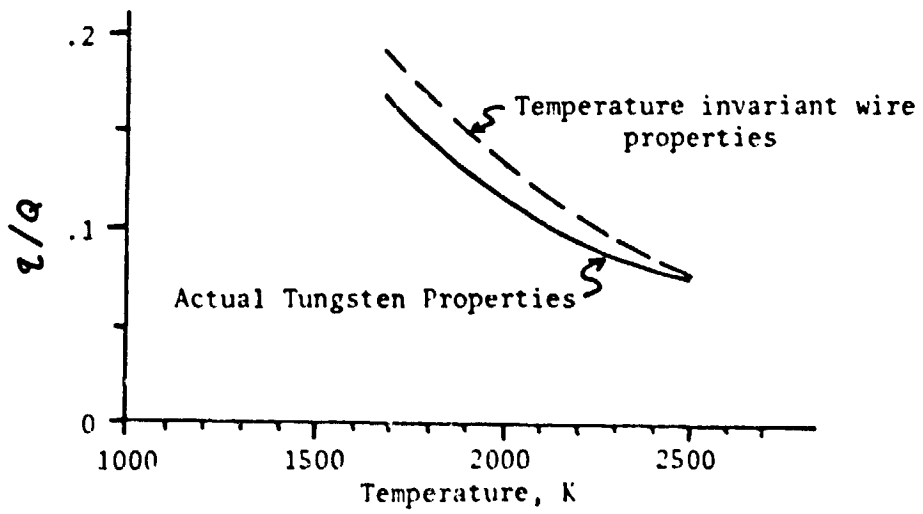Figure 5.   Heat Loss in a Tungsten Wire

Figure 6.   Heat Loss Example

Example

This example will determine the amount of heat being lost in the tungsten support wire using Figure 5. Suppose a .3 cm diameter niobium sample drop is suspended from a .025 cm (.010 inch) tungsten wire and is maintained a an absolute temperature of 2100 K. The function $B_1(T)$ is determined from Figure 5 to be 220. Therefore:

$$q/d^{1.5} = B_1(T) = 220$$

$$q = 220 \, (.025)^{1.5}$$

$$q = .89 \text{ watts}$$

The ratio of heat conducted up the wire to the heat radiated from the sphere is determined from function $B_2(T)$. Assume the emissivity of niobium is known to be 0.265.

$$\frac{q}{Q}\left(\frac{\epsilon D^2}{d^{1.5}}\right) = B_2(2100 \, K) = .63$$

$$\frac{q}{Q} = .63 \, \frac{d^{1.5}}{\epsilon D^2}$$

$$\frac{q}{Q} = .63 \, \frac{(.025)^{1.5}}{.265 \, (.3)^2} = .107$$

Therefore the heat loss up the wire is 10.7% of the heat loss from the sample drop itself.

Using reported values of emissivity for niobium [8], the heat loss ratio for the previously defined sample (.3 cm niobium drop on an .025 cm tungsten wire) was determined over a range of temperatures. The results are shown in Figure 6 as the solid line.

Also plotted in Figure 6 are the results obtained using the temperature invariant property solution of equation (5). The tungsten properties were evaluated at the drop temperature for this solution. It can be noted that this approach overestimates the heat loss in the wire in this example by up to 2 percent.

Typical temperature profiles along the wire are shown for both cases in Figure 7. These were obtained by numerically integrating the temperature gradient equations (equations (3) and (7)).

Conclusions

Temperature variations along the pendant drop support wire cannot be ignored. If the temperature variations of the support wire material properties are known, then the heat loss in the wire can be determined by knowing the drop temperature alone.
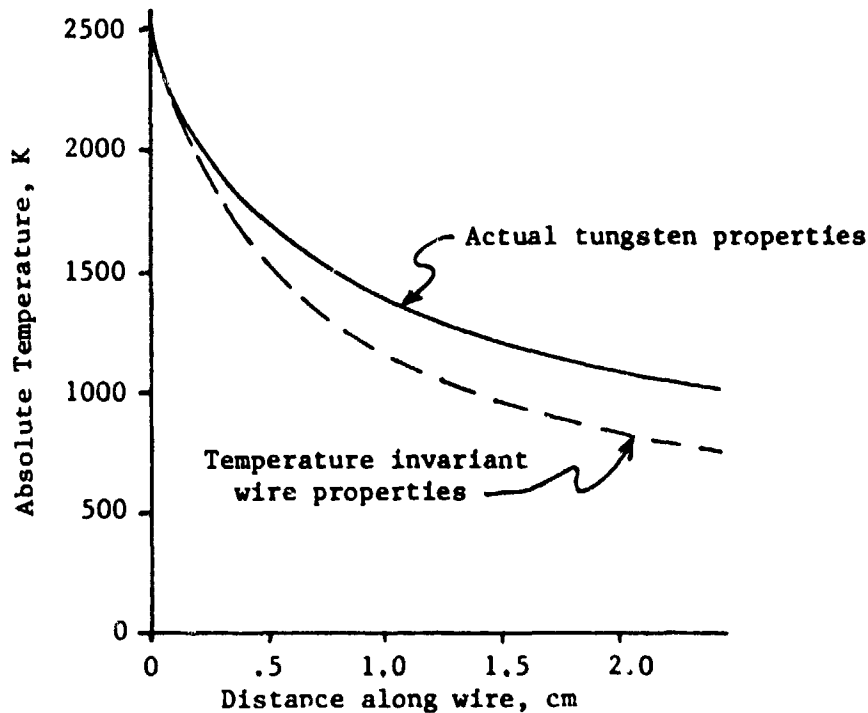
Figure 7. Temperature variation along support wire

TEMPERATURE OF A FALLING MOLTEN DROP

Heat transfer to and from spheres has been a topic of interest for many years. Many experiments were performed using different fluids (air, water, and oils) at various flow rates. From the results many empirical relationships have been developed to describe the heat transfer process [9, 10, 13 - 18]. Several of the investigators developed relationships to bit a number of experimental data sets using the dimensionless constants of Reynolds number (Re), Nusselt number (Nu) and Prandtl number (Pr). These are:

McAdams [14] - $Nu = .37 \ Re^{.6}$ ————(12)
$$17 < Re < 17,000$$

White [18] - $Nu = 2 + .3 \ Pr^{1/3} \ Re^{.6}$ ————(13)
$$Re < 100,000$$

Whitaker [17] - $Nu = 2 + (.4 \ Re^{1/2} + .06 \ Re^{2/3}) Pr^{.4} \left(\frac{\mu_0}{\mu_T}\right)^{.25}$ ——(14)
$$3.5 < Re < 80,000$$
$$.7 < Pr < 380$$
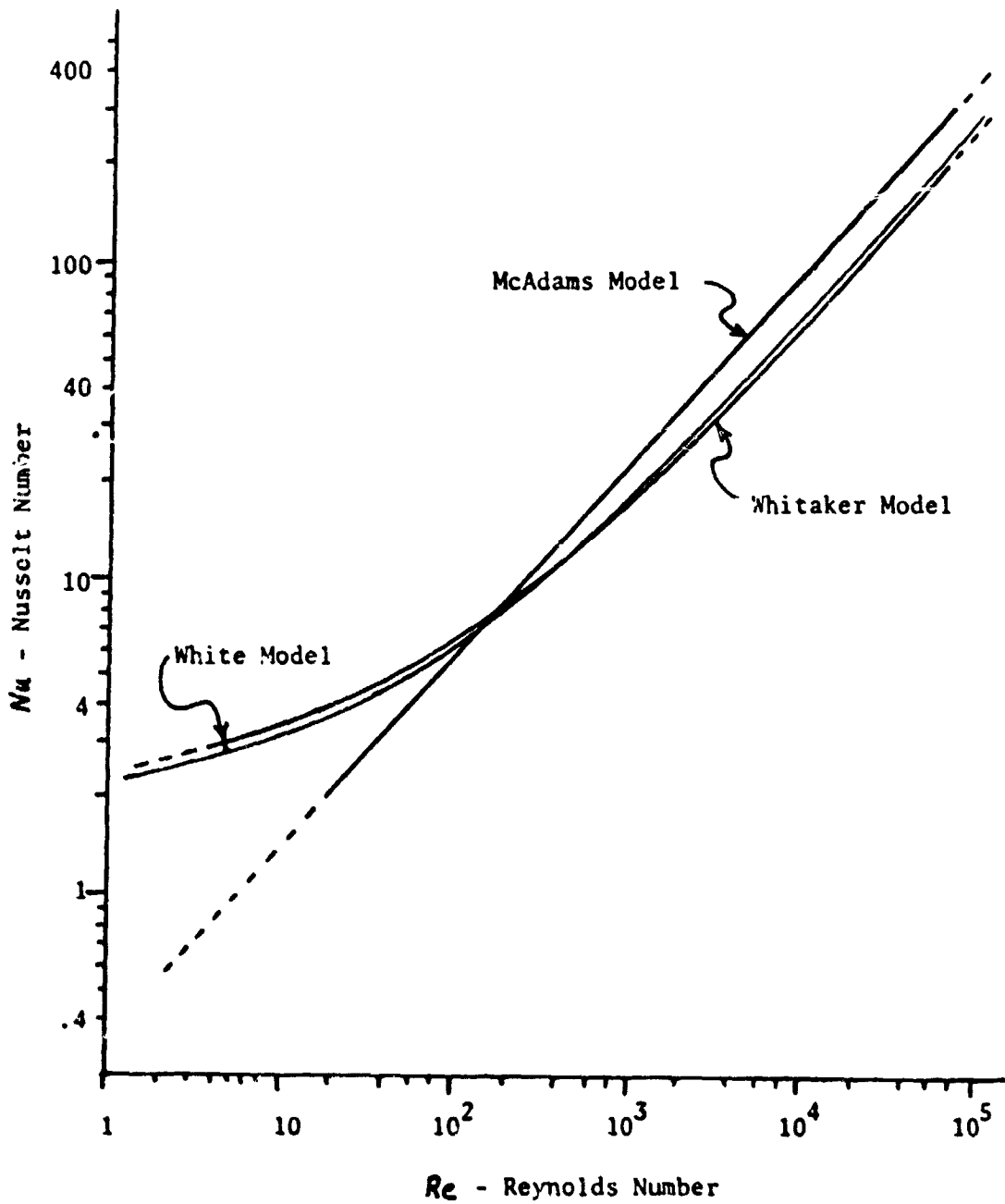
Figure 8. Comparison of Heat Transfer Models when T_{film} approximately equals T_{sphere}

Equations (12) and (13) have all the fluid properties evaluated at the film temperature as defined by:

$$T_f = \frac{T + T_o}{2} \qquad\qquad (15)$$

Equation (14) on the other hand evaluates all properties at the bulk temperature (T_o) except for μ_T, the absolute viscosity of the gas which is evaluated at the drop temperature (T). Figure 8 compares these functions under the conditions that the Prandtl number is 0.7 and the bulk temperature (T_o) is nearly the same as the sphere temperature (T). The lower limit of the Nusselt number being 2 for the last two equations represents steady state heat transfer by conduction from a stationary sphere in an infinite, non-moving fluid media.

Drop tube conditions which limit the applicability of the previous equations to determination of heat transfer from a falling molten drop include:

(a) Conditions are not steady state, velocity continually increases.

(b) Large temperature differences exist between the specimen and ambient conditions (~2000 K) as compared with the cited experiments (200 K maximum).

(c) The gas in the tube (if any) is below atmospheric pressure and most of the experimenters used atmospheric pressure and above.

(d) As a result of low gas densities (and pressures), Reynolds number for the drop tube are typically very low (on the order of 0.1 to 100).

(e) At very low drop tube gas pressures the mean free path of the gas molecules will become so large that the gas will no longer behave like a continuum (a basic assumption for the three models presented).

A solution to this problem has been formulated by Robinson [7] using an average film temperature to evaluate the gas properties and McAdams model of heat transfer (equation (12)). Since properties are assumed constant a closed form solution is obtained. Helium gas was used as the backfill gas due to its high thermal conductivity and inert behavior.

The present study determines the lower pressure applicability of the continuum assumption in this problem, develops an iterative scheme to include helium property variations during the fall, and investigates the effect of different heat transfer models.

Helium Properties

A gas can be considered a dilute gas if it exists at a temperature significantly above its critical temperature and at a pressure below its critical pressure. For helium these properties are [18] $T_c = 5$ K and $P_c = 2.26$ atmospheres. Thus for the drop tube conditions helium may be accurately modeled as a dilute gas.

Viscosity of a dilute gas may be expressed (with an accuracy of plus or minus 4%) by a power law equation [18]:

$$\mu_{He} = 4.46 \times 10^{-6} \; T^{.666} \qquad\qquad —(16)$$

Using the ideal gas law the density may be expressed by the equation (note that the pressure is in units of torr):

$$\rho_{He} = 6.41 \times 10^{-5} \; p/T \qquad\qquad —(17)$$

White [18] states that thermal conductivity of gasses may also be expressed as power law functions. A curve fit of published helium thermal conductivities [11] yields the relationship (which fits the data to plus or minus 4% over the range of temperatures of interest in the drop tube):

$$k_{He} = 2.20 \times 10^{-5} \; T^{.74} \qquad\qquad —(18)$$

Helium Mean Free Path

The Knudsen nember is defined as the ratio of mean free path to some characteristic dimension. For the drop tube the characteristic dimension is the drop diameter, D. Therefore,

$$K_n = \lambda/D$$

When the Knudsen number is very small the gas may be assumed to be a continuum. For Knudsen numbers greater than one the gas must be treated as free molecules with intermolecular collisions becoming rare. Gas heat transport will depend to an increasing extent on collisions of molecules with the bounding surfaces.

Mean free path is the average distance traveled by an atom or molecule of a gas between collisions. For a gas at rest in local thermo-dynamic equilibrium is defined by [19]:

$$\lambda = \frac{\mu}{\rho} \left(\frac{\pi}{2RT}\right)^{1/2}$$

But since

$$p\mathcal{V} = p/\rho = RT$$

Therefore

$$\lambda = \mu \left(\frac{\pi}{2\rho p}\right)^{1/2} \qquad\qquad —(19)$$

Combining equations (16), (17), and (19) result in:

$$\lambda_{He} = 1.91 \times 10^{-5} \; T^{1.166}/p \qquad\qquad —(20)$$

Figure 9 is a graphic representation of equation (20) and shows helium mean free path as a function of temperature and pressure.
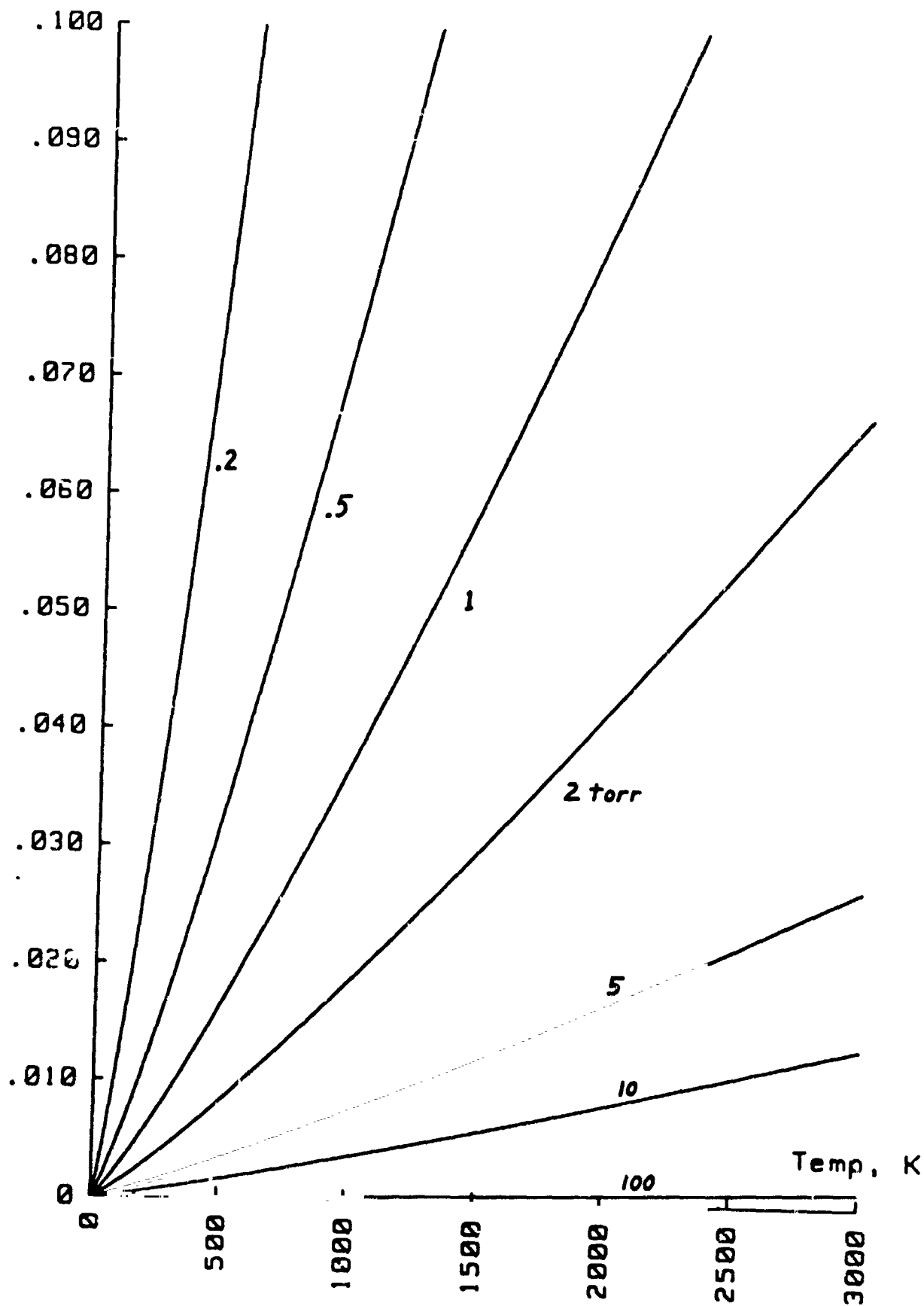
Figure 9. Mean Free Path of Helium Atoms

For film temperatures of around 1500 K and a drop diameter of .3 cm, the applicability of the continuum assumption at pressures below 2 torr would be suspect. At a pressure of 10 torr the gas should behave as a continuum.

This analysis will include pressure computations down to 1 torr although the accuracy of the results at this pressure should be suspect for small diameter samples.

Computation Method

The rate at which a heated sphere will lose energy during a free fall in the drop tube filled with a gas is given by Robinson [7] and Zemansky [20] as:

$$\frac{dQ}{dt} = -\epsilon A \sigma (T^4 - T_0^4) - hA(T - T_0)$$

The heat lost in a small time increment may be estimated from:

$$\Delta Q = \left[-\epsilon A \sigma (T^4 - T_0^4) - hA(T - T_0) \right] \Delta t \qquad \text{(21)}$$

The temperature drop of the sphere can be estimated from:

$$\Delta T = \Delta Q / cm \qquad \text{(22)}$$

The time increment can be determined for a given displacement of the drop from the general dynamics equations:

acceleration is the acceleration of gravity minus the effect of drag –

$$a = g - F/m \qquad \text{(23)}$$

velocity change in a distance of Δx –

$$V_{new} = \left(V_{old}^2 + 2a \Delta x \right)^{1/2} \qquad \text{(24)}$$

$$\Delta V = V_{new} - V_{old}$$

time to travel a distance Δx –

$$\Delta t = \Delta V / a \qquad \text{(25)}$$

The iteration scheme developed determines the temperature, acceleration, and heat transfer histories of a drop by examining the conditions at incremental positions along the drop tube. Conditions are assumed to remain constant during the small travel distance and new conditions are determined at the end of the interval to be used in the next interval. The iteration scheme is:

(a) Define the system conditions (gas pressure, drop material properties, etc.) and initial conditions (initial drop temperature).

(b) Determine the film temperature -- equation (15)

(c) Determine the helium properties at the film temperature -- equations (16), (17), and (18).

(d) Evaluate Reynolds number using current helium properties (at the film temperature).

(e) Evaluate drag force (discussed in a later section).

(f) Determine acceleration, velocity change and time increment -- equations (23), (24), and (25).

(g) Determine the current value of Nusselt number, Nu, using one of the models given in equations (12), (13), and (14). With the Whitaker model (equation (14)) the bulk Reynolds and Prandtl numbers are used.

(h) h is determined from the definition of the Nusselt number:

$$h = \frac{Nu \; k}{D}$$

(i) Heat loss from the sphere is determined from equation (21).

(j) The change in temperature is determined from equation (22) and thus a new drop temperature is found.

(k) Repeat steps (b) through (j) until the drop reaches the bottom of the tube.

It was determined that steps of 10 cm provided accurate results. The results, after forcing all material characteristics to remain constant throughout the fall, were compared to the closed form solution of Robinson [7]. No appreciable differences were found in the results. Reducing the integration step size also resulted in no appreciable change.

Results

The iteration scheme is applied to the example of a niobium drop in order to compare the various heat transfer models. Table 1 shows the high temberature niobium properties which were assumed [8, 11].

A typical simulation result for a 3 mm niobium sphere dropped in a 100 torr (13.4 kPa) helium atmosphere with an initial temperature at its melting point is shown in Figure 10. The White model of convective cooling (equation (13)) was assumed in this example. Note that the radiation provides less than 40% of the total cooling for the sphere.

Table 1. Niobium Properties at High Temperature

Melting Temperature	2741 K
Specific Gravity	8.57
Heat Capacity	.435 kJ/kg K
Thermal Conductivity	.548 W/m K
Total Hemispherical Emissivity	.317

Diameter   = .30 cm   Init. Temp. = 2741 K   Emissivity   = .317
Specific Gr. = 8.57 g/cm^3   Therm. Cond. = .548 W/cm K) Heat Capacity = .435 kJ/kg)
Tube Temp.   = 300 K   He Pressure = 100 torr

Sphere Mass = .1212 g   Surface Area = .28274 cm^2

White model: $Nu=2+.3*(Pr^.333)*(Re^.6)$

Dist	Time	Temp	g-level	Qrad	Qconv	H	Biot	Nu	Re	Pr
10	.143	2597	.002059	4.11	3.48	.0352	.0096	2.12	.3	.613
100	.452	2357	.006577	10.18	10.08	.0344	.0094	2.28	1.2	.619
200	.640	2236	.009476	12.86	13.73	.0344	.0094	2.37	1.8	.621
300	.784	2153	.011777	14.58	16.39	.0344	.0094	2.43	2.3	.623
400	.906	2088	.013770	15.84	18.55	.0345	.0094	2.48	2.8	.624
500	1.013	2035	.015567	16.83	20.40	.0345	.0094	2.52	3.3	.625
600	1.110	1988	.017225	17.64	22.02	.0345	.0095	2.56	3.7	.626
700	1.199	1948	.018779	18.33	23.47	.0346	.0095	2.60	4.1	.627
800	1.282	1912	.020250	18.92	24.80	.0346	.0095	2.64	4.5	.628
900	1.360	1879	.021654	19.43	26.02	.0347	.0095	2.67	4.9	.628
1000	1.434	1848	.023003	19.89	27.15	.0347	.0095	2.70	5.3	.629
1100	1.504	1821	.024304	20.29	28.21	.0347	.0095	2.73	5.7	.629
1200	1.571	1795	.025565	20.66	29.21	.0348	.0095	2.76	6.1	.630
1300	1.636	1771	.026789	20.99	30.15	.0348	.0095	2.79	6.5	.631
1400	1.698	1748	.027983	21.30	31.04	.0349	.0095	2.81	6.8	.631
1500	1.758	1726	.029148	21.58	31.89	.0349	.0096	2.84	7.2	.632
1600	1.816	1706	.030287	21.83	32.70	.0350	.0096	2.87	7.5	.632
1700	1.872	1687	.031404	22.07	33.48	.0350	.0096	2.89	7.9	.633
1800	1.926	1669	.032501	22.29	34.23	.0351	.0096	2.91	8.3	.633
1900	1.979	1651	.033578	22.50	34.94	.0351	.0096	2.94	8.6	.633
2000	2.031	1634	.034638	22.69	35.63	.0352	.0096	2.96	9.0	.634
2100	2.082	1618	.035682	22.87	36.30	.0352	.0096	2.98	9.3	.634
2200	2.131	1603	.036711	23.04	36.94	.0353	.0097	3.00	9.6	.635
2300	2.179	1588	.037727	23.20	37.57	.0353	.0097	3.03	10.0	.635
2400	2.226	1574	.038729	23.35	38.17	.0354	.0097	3.05	10.3	.635
2500	2.273	1560	.039720	23.49	38.76	.0354	.0097	3.07	10.7	.636
2600	2.318	1547	.040699	23.62	39.33	.0355	.0097	3.09	11.0	.636
2700	2.362	1534	.041668	23.75	39.88	.0355	.0097	3.11	11.4	.636
2800	2.406	1521	.042627	23.87	40.42	.0355	.0097	3.13	11.7	.637
2900	2.449	1509	.043577	23.99	40.94	.0356	.0097	3.15	12.0	.637
3000	2.491	1497	.044518	24.10	41.45	.0356	.0098	3.17	12.4	.637
3100	2.533	1486	.045450	24.20	41.95	.0357	.0098	3.19	12.7	.638
3200	2.573	1475	.046375	24.30	42.44	.0357	.0098	3.21	13.0	.638

Figure 10.   Typical Simulation Results

Figures 11, 12 and 13 graphically present the cooling curves for the heat transfer models of McAdams, Whitaker, and White, respectively, applied to the example of a 3 mm niobium sphere. Tube atmosphere varies from a vacuum to one atmospheric pressure (760 torr) of helium gas.

Figures 14 and 15 show the cooling curves for a 5 mm and a 7 mm niobium sphere, respectively, using White's heat transfer model.

Figures 16, 17 and 18 compare the three different methods of heat transfer prediction. Figure 16 is for a 3 mm drop with a helium atmosphere of 10 torr. Figure 17 is for a 3 mm drop at 760 torr helium pressure. Finally, Figure 18 is for a 5 mm drop at 760 torr.

In the beginning of the fall down the drop tube, the McAdams model predicts the slowest cooling rate. This is due to the fact that it does not have a lower bound of 2 on the Nusselt number as the other two models do (see Figure 8). This lower bound was included by White and Whitaker to account for conductive heat transfer into an infinite media. Equation (13) indicates, however, that the McAdams model applies only to situations where Reynolds number exceeds 17. This is true, in the examples presented, only after falling about 2 meters in a helium atmosphere of 760 torr. The Reynolds number is below 17 in all other cases, therefore the results for 1, 10 and 100 torr of helium pressure should not be trusted.

In all cases the White model (equation (13)) predicts the most convective heat transfer. Figure 8 shows that the White model and the Whitaker model are very similar -- provided that the bulk temperature and the film temperature are nearly identical. Since this is not the case for a niobium sphere in a drop tube, the models differ widely in the predicted temperature profiles (see Figures 16-18). The variation in the predicted undercooling varies between the two models by up to 160 K at the bottom of a 32 meter drop tube -- 1390 K versus 1230 K (see Figure 16).

The White model predicts undercoolings of between 1200 K and 1500 K are possible for a 3 mm niobium sphere dropped at its melting point into atmospheres of helium varying in pressure from 10 torr to 760 torr. About 700 K undercooling is possible in a vacuum under the same conditions.
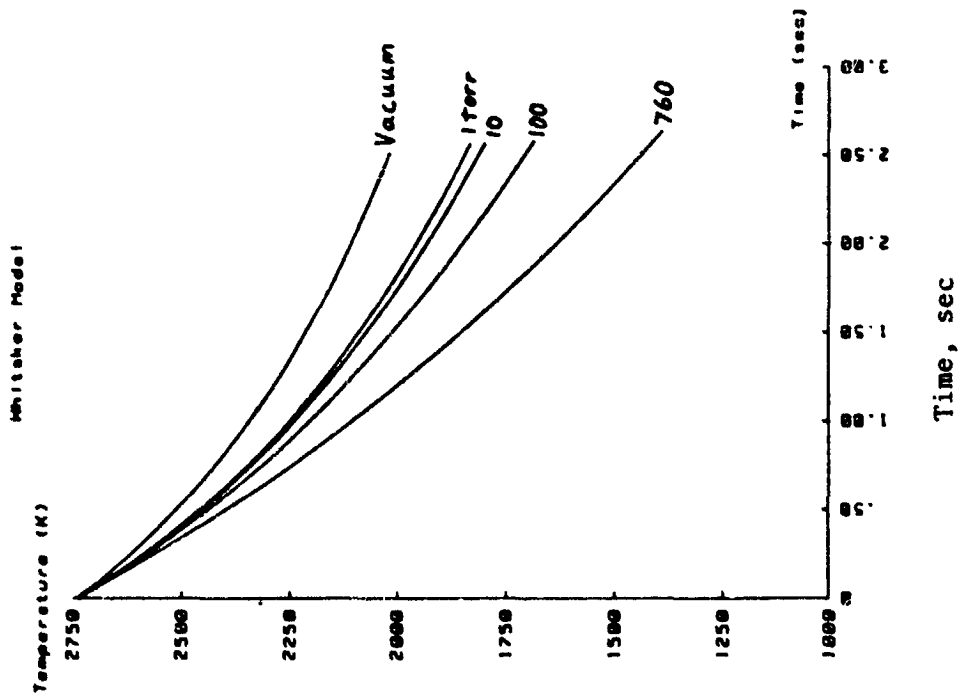
Figure 12. Whitaker Model for 3mm Niobium
            Sphere
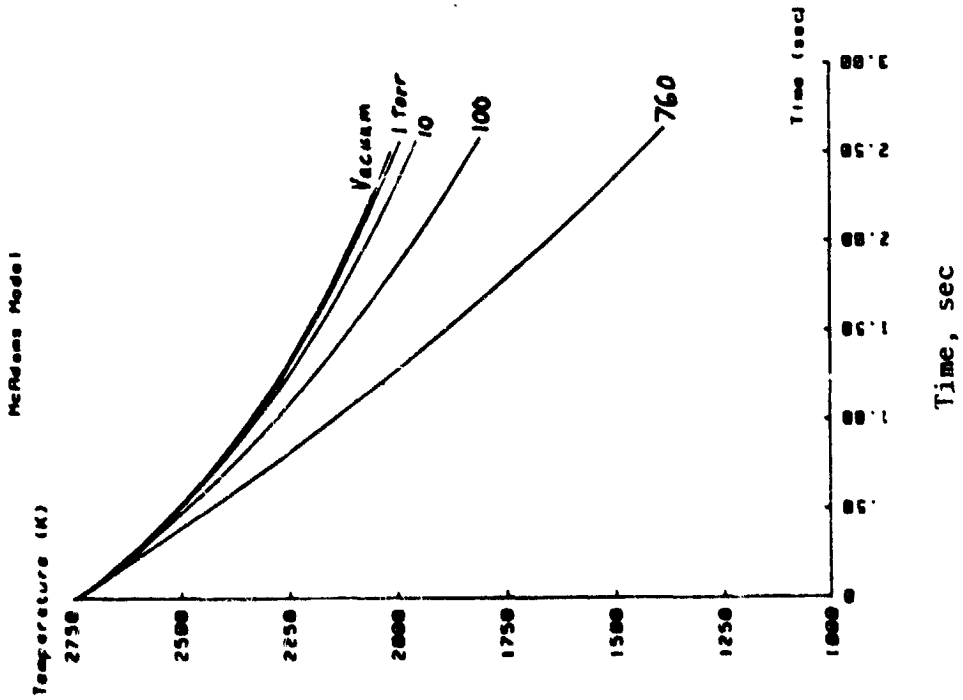
Figure 11. McAdams Model for 3mm Niobium
            Sphere
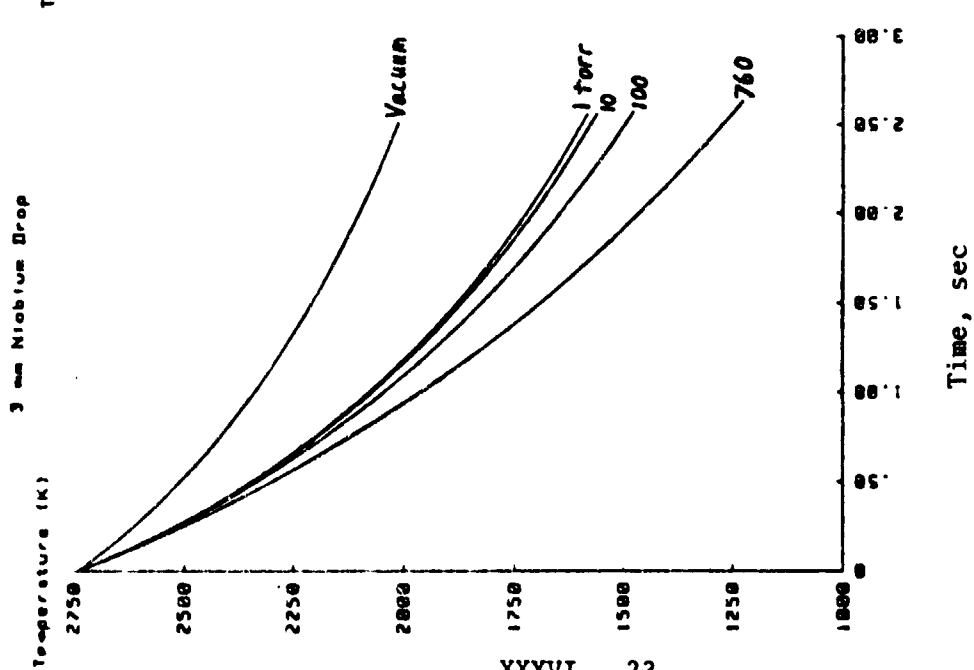
Figure 13. White Model for 3mm Niobium Sphere
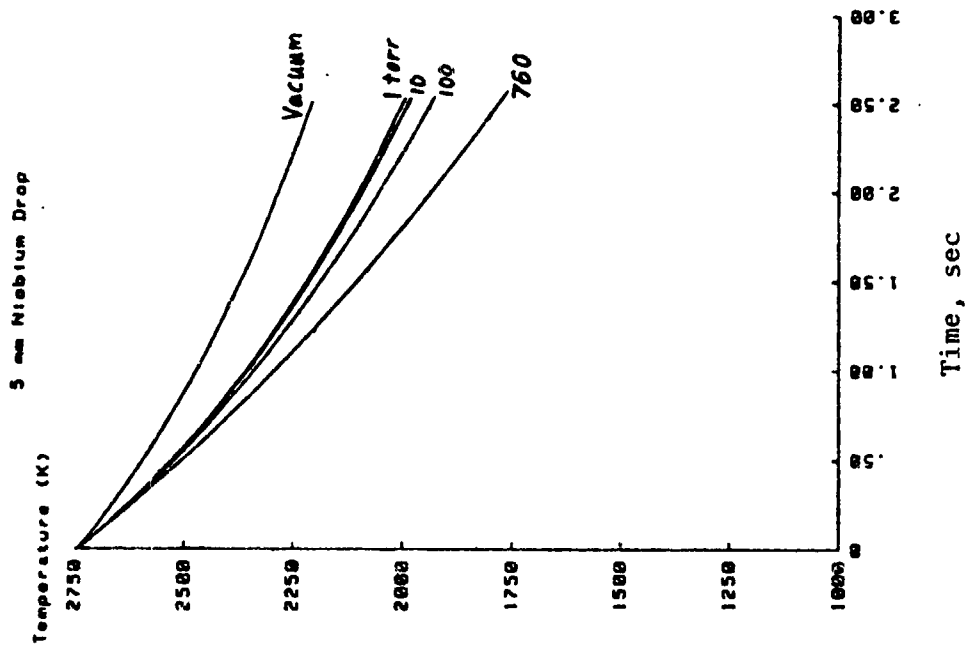
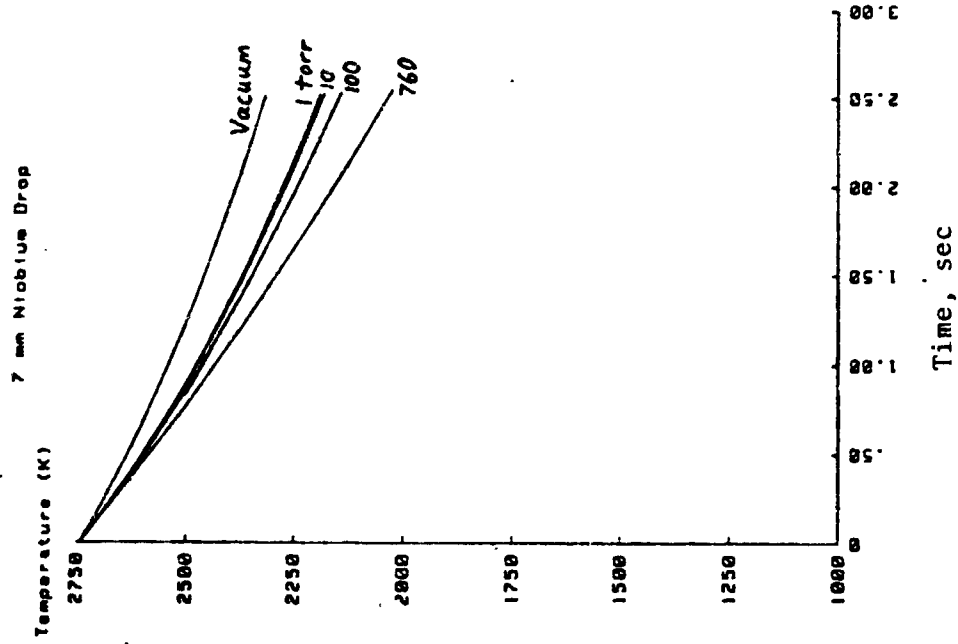Figure 14. White Model for 5mm Niobium Sphere

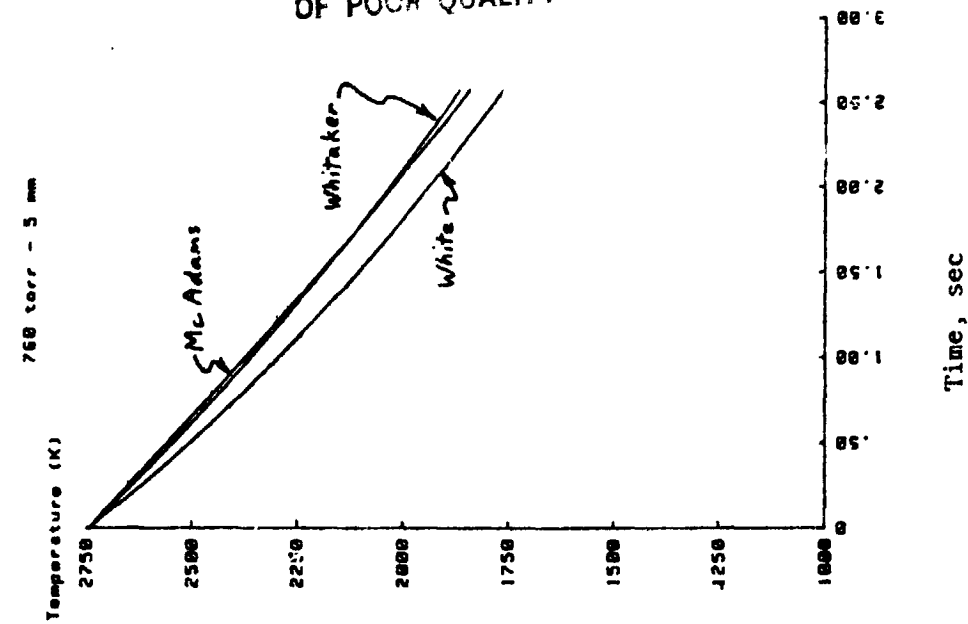Figure 15. White Model for 7mm Niobium Sphere

Figure 16. Model Comparison for
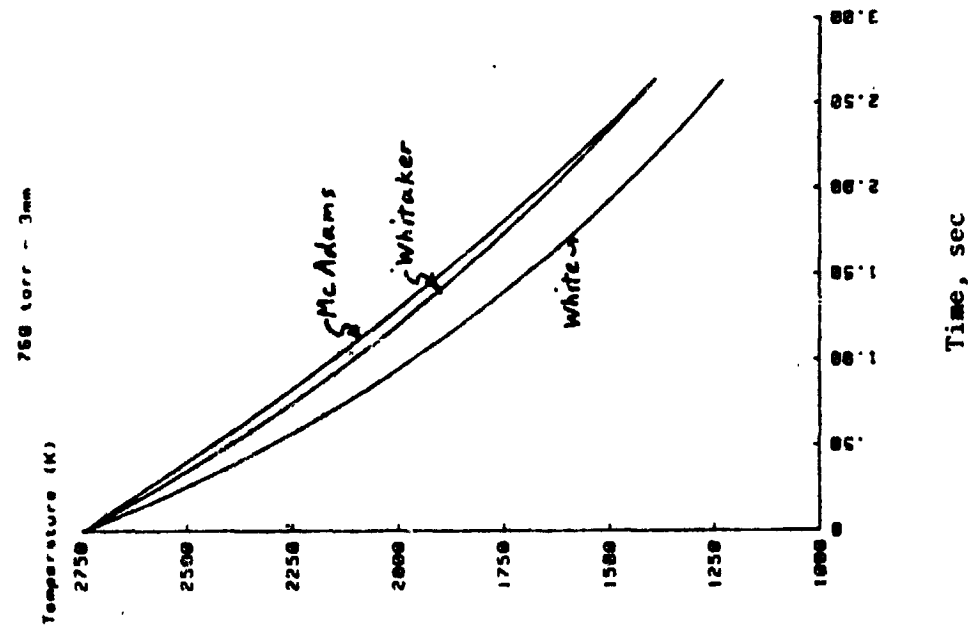3mm Niobium Sphere at
10 torr Helium Pressure

Figure 17. Model Comparison for
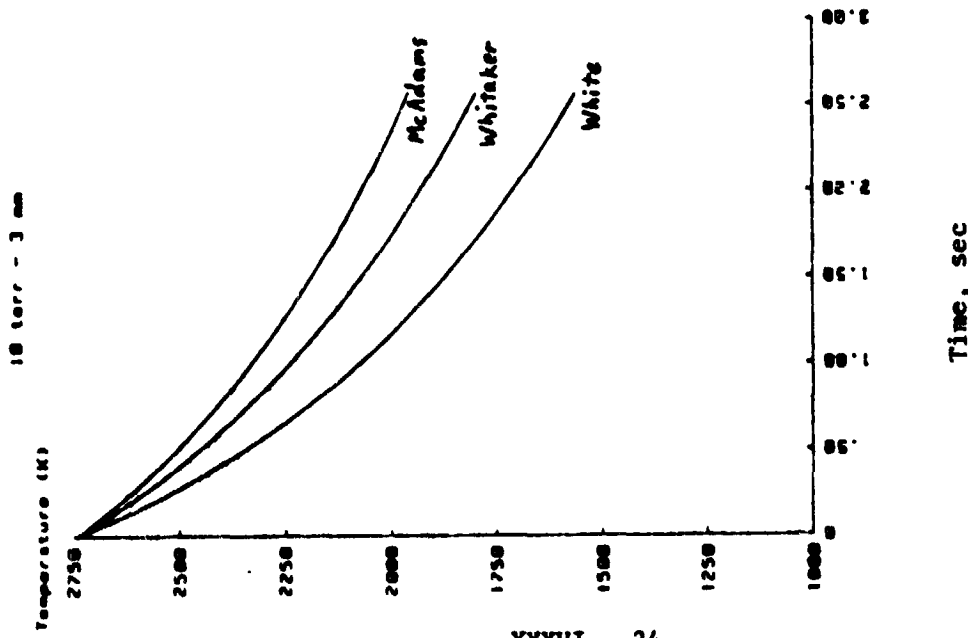3mm Niobium Sphere at
760 torr Helium Pressure

Figure 18. Model Comparison for
5mm Niobium Sphere at
760 torr Helium Pressure

DRAG ON A FALLING MOLTEN DROP

Drag on spheres has been extensively studied, both analytically and experimentally, and summaries of the results can be found in most fluid flow books [18, 21, 22]. Flow around a sphere can be catagorized into one of four types:

(a) Free molecular flow

(b) Creeping flow (laminar flow)

(c) Turbulent flow

(d) Supersonic flow

Only the first three types of flow can exist in the typical drop tube so only they will be discussed. Turbulent flow results in a drag force, F, which can be predicted by the equation [21, 22]:

$$F = C_D A_p \frac{\rho V^2}{2} \qquad\qquad \text{---(26)}$$

where C_D is the drag coefficient defined by Figure 19 in terms of the Reynolds number, Re, and A_p is the projected area of a sphere:

$$A_p = \frac{\pi}{4} D^2$$

The drag coefficient curve (Figure 19) is a result of experimental studies.

The creeping flow around a sphere can be solved analytically yielding the famous formula determined by Stokes in 1851 (see [18, 21]):

$$F = 3 \pi D \mu V \qquad\qquad \text{---(27)}$$
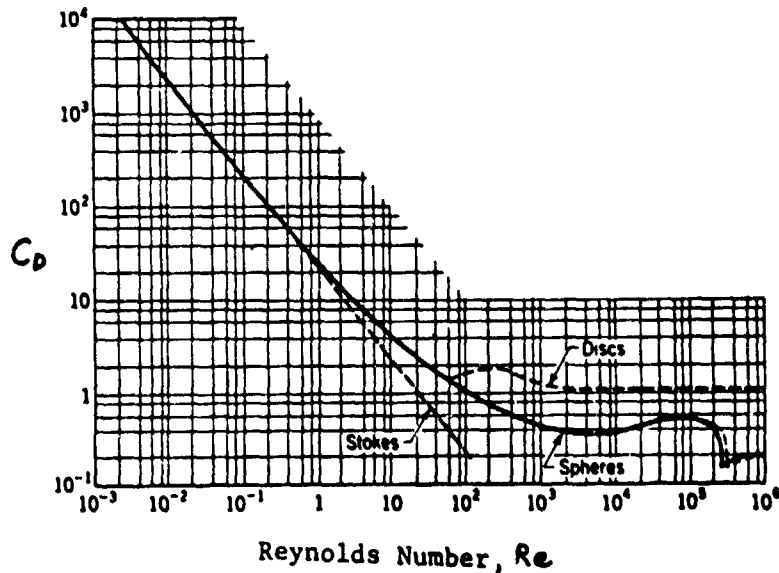
Figure 19. Drag Coefficients for Spheres

This equation does not functionally depend on Reynolds number (or fluid density). However this equation is valid only for low values of Reynolds numbers ($Re < 1$). If the drag coefficient of equation (26) is chosen as

$$C_D = {}^{24}/_{Re}$$

the resulting equation is Stokes equation (equation (27)). This lower asymptotic limit for C_D is shown in Figure 19. White [18] suggests a formula for the drag coefficient, C_D, which closely matches the data of Figure 19 over a Reynolds number range of 0 to 200,000:

$$C_D = \frac{24}{Re} + \frac{6}{1 + \sqrt{Re}} + .4 \qquad \text{———(28)}$$

Using this formula both creeping flow and turbulent flow drag can be accurately predicted.

As pointed out by Happel and Brenner [23], the previous equations are based on the assumption that the fluid behaves like a continuum and that no slippage occurs at solid surfaces. For very dilute gasses these assumptions are not valid.

Epstein [24] developed a detailed theory of motion of small spheres relative to gasses from a kinetic theory viewpoint. The result is a first order perturbation scheme that is applicable when deviations from creeping motion are not large [23]:

$$F = 3 \pi D \mu V \left(1 - \frac{2 \mu}{D \beta} \right) \qquad \text{———(29)}$$

where β is a coefficient of proportionality, designated as the "coefficient of sliding friction." The "slip coefficient," μ/β, has been determined experimentally and found to agree with the general kinetic theory formulation based on the assumption that there is only a small proportion of specular reflection. Happel and Brenner [23] suggest that the slip coefficient can be evaluated from:

$$\frac{\mu}{\beta} = C \lambda$$

where the constant C has a value between 1.3 and 1.4 and λ is the mean free path of the gas molecules. Equation (29) then becomes (using C with a value of 1.35):

$$F = 3 \pi D \mu V \left(1 - 2.7 \, Kn \right) \qquad \text{———(30)}$$

where Kn is the Knudsen number defined previously in the discussion of the helium gas. Happel and Brenner [23] suggest that equation (30) is valid when the value of the Knudsen number is less than 0.1 (i.e. when the square of the Knudsen number is negligible). For a 3 mm diameter niobium sphere, Figure 9 indicates that this corresponds to a pressure greater than 2 torr.

For very small Knudsen numbers, say $K_n < .01$, the correction factor of equation (30) becomes negligible and Stokes equation (equation (29)) is sufficiently accurate. For a 3 mm niobium sphere, Figure 9 indicates that a helium pressure greater than 10 torr is required.

Computation Method

For this study, the gas was assumed to be sufficiently dense so that the resulting flow could be considered either creeping or turbulent. Therefore equations (26) and (28) were used to compute the drag force on a falling drop.

The apparent g-level experienced by the drop was determined by dividing the frag force, F , by the weight of the drop:

$$g - level = \frac{F}{mg} \qquad\qquad (31)$$

As pointed out in the previous section, this computation procedure is valid only for Knudsen numbers less than .01.

The computation scheme is incorporated in the previously described iteration method. The film temperature is used to evaluate the helium properties of density and viscosity.

Results

Figures 20, 21 and 22 show the g-levels which result from drops of niobium 3 mm, 5 mm, and 7 mm in diameter respectively. The drops were started at the melting point of niobium into various pressures of helium gas.

The accelerations experienced by the spheres due to the drag force under these conditions ranged from approximately .005 times the acceleration of gravity to over .1 times the acceleration of gravity at the bottom of the 32 meter drop tube. These levels are sufficiently high to make a claim of "low - g" questionable.

Note that the g-levels predicted for the helium atmosphere of 1 torr are probably high. The methods presented in the previous section could be used to refine the drag predictions.
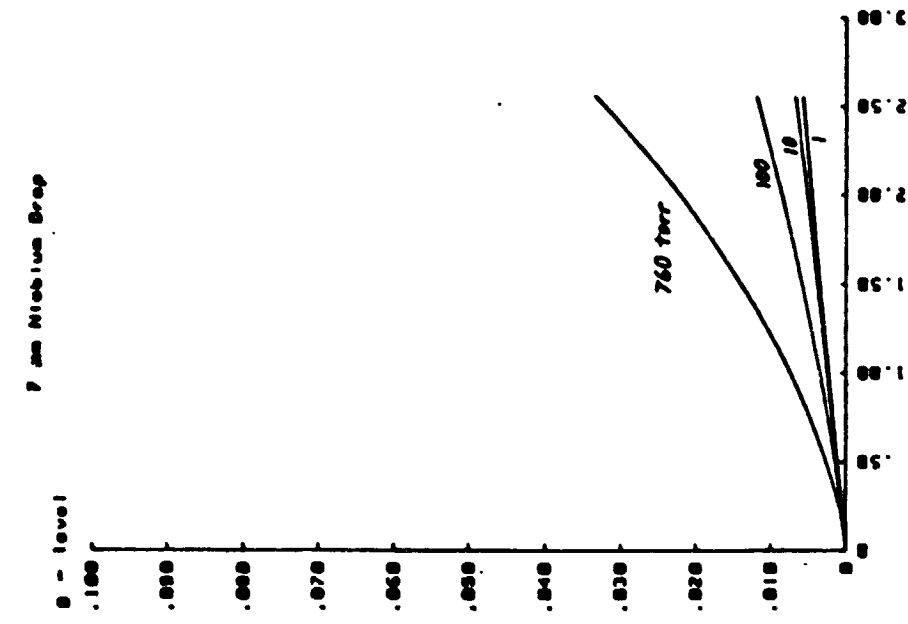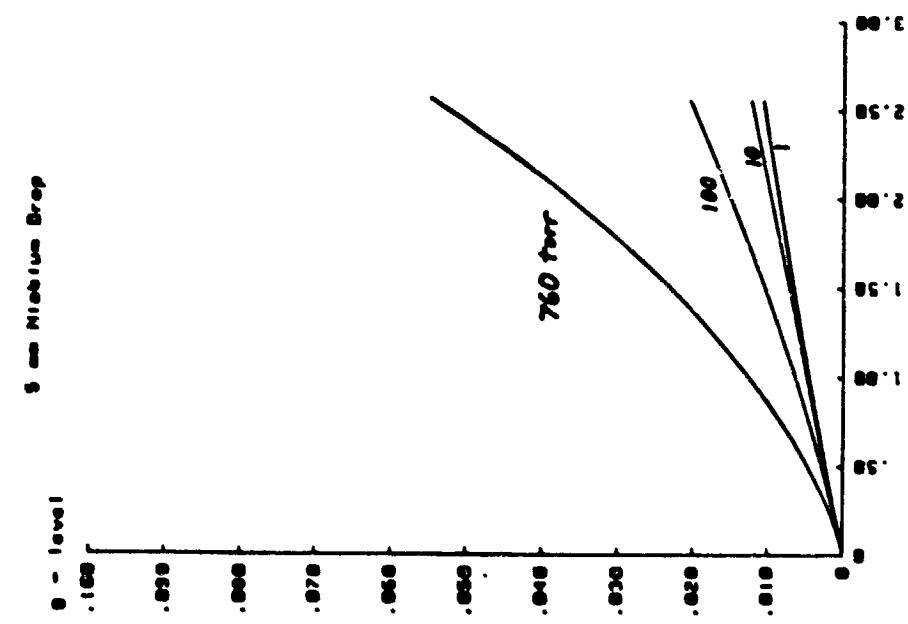
Figure 20. 3 mm Niobium Drop
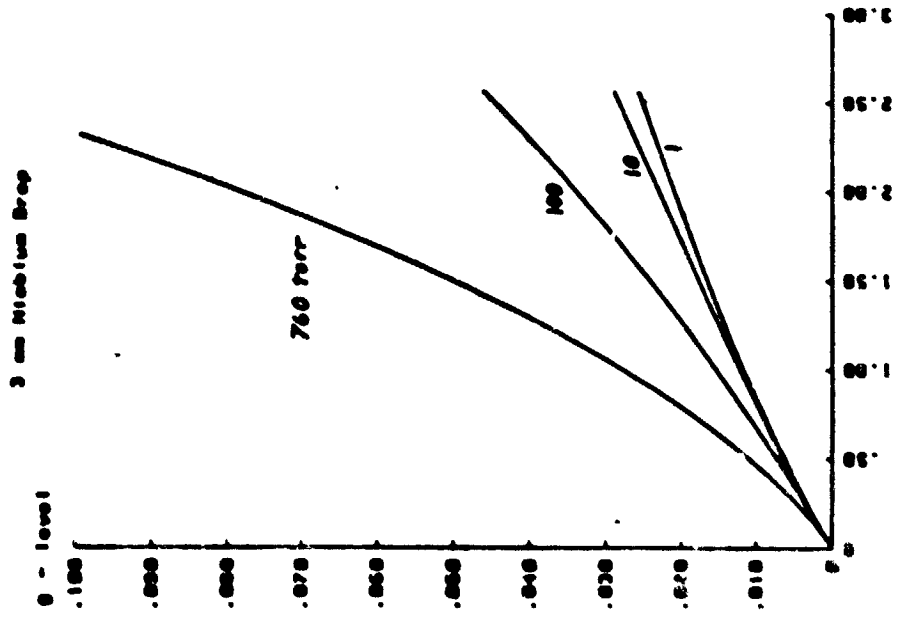

Figure 21. 5 mm Niobium Drop

Figure 22. 7 mm Niobium Drop

CONCLUSIONS AND RECOMMENDATIONS

This study has resulted in the development of methods to determine:

    (a) The heat loss from a pendant drop sample to its support wire.

    (b) The temperature history of a molten metal drop falling in a helium atmosphere,

    (c) The drag force and resultant g-level of a molten metal drop falling in a helium atmosphere.

Examples were given using niobium as the drop material but the method applies to any metal. The analyses schemes developed assume continuum behavior of the gas and are therefore limited to helium pressures greater than 5 torr for most drop tube specimens (the actual limit depends upon the Knudsen number -- the ratio of the mean free path to the helium to the drop diameter).

The results indicate that while cooling rates (and thus possible undercooling) increases as a result of introducing a gas (helium) into the drop tube, the resulting g-levels experienced by the drop are not negligible. The result is containerless but not necessarily low-g material processing.

This study confirms the results of previous studies and refines the accuracy of the predictions. It allows the incorporation of temperature dependent material characteristics not feasible with closed form solutions. However, additional work is needed in the following areas:

    (1) Experimental temperature profiles of the pendant drop support wire should be compared with the theoretical predictions to test the validity of the assumptions.

    (2) More research is needed to determine the most appropriate heat transfer model for a hot sphere moving through a dilute helium gas. The available models show too much variation.

    (3) Experimental data points are needed to confirm the results of analytical predictions of the temperature of a drop falling in a tube. This might be accomplished by catching the drop at various elevations in the drop tube and then determining what its temperature was at that point by calorimetry (accurate high temperature properties must be known).

    (4) Experimental data points are needed to evaluate the drag predictions. This can be done by accurately timing a drop over a certain known distance.

    (5) The theory on the falling drop needs to be extended into regions of high Knudsen numbers. The application of the kinetic theory of gasses is needed.

REFERENCES

1. "Proceedings of the Third Space Processing Symposium: Skylab Results," Vol I and II, NASA Report M-74-5, June, 1974.

2. Steg, L. (ed.), "Material Sciences in Space with Applications to Space Processing," Progress in Astronautics and Aeronautics, 52, AIAA, New York, 1977.

3. "Apollo - Soyuz Test Project Summary Science Report," Vol. I, NASA Report SP-412, 1977, pp. 353-563.

4. Naumann, R.J., "Early Space Experiments in Materials Processing," NASA Report SP- 443, 1980.

5. Lacy, L.L., Robinson, M.B., and Nisen, D.B., "Containerless High Temperature Calorimeter Apparatus," U.S. Patent 4,278,083, 1981.

6. Murr, L.E., Interfacial Phenomena in Metals and Alloys, Addison - Wesley, Reading, Mass., 1975, pp 95-98.

7. Robinson, M.B., "Radiative and Gas Cooling of Falling Molten Drops," NASA Report TM-78189, 1978.

8. Robinson, M.B., "Undercooling Measurement in a Low-Gravity Containerless Environment," M.S. Thesis submitted to The University of Alabama in Huntsville, 1981.

9. Kreith, F., Principles of Heat Transfer, 3rd Edition, Intext Publishing Co., New York, N.Y., 1973.

10. Holman, J.P., Heat Transfer, 5th Ed., McGraw-Hill, New York, N.Y., 1981.

11. CRC Handbook of Chemistry and Physics, 54th Ed., The Chemical Rubber Company, Cleveland, OH, 1974.

12. Seigel, R. and Howell, J.R., Thermal Radiation Heat Transfer, McGraw-Hill Company, New York, NY, 1972.

13. Kramers, H., "Heat Transfer From Spheres to Flowing Media," Physica, XII, No 2-3, June, 1946, pp. 61-80.

14. McAdams, W.H., Heat Transmission, McGraw Hill Co., New York, NY, 1954.

15. Yuge, T., "Experiments on Heat Transfer From Spheres Including Combined Natural and Forced Convection," Trans. ASME J. of Heat Transfer, C82, 1960, pp. 214-220.

16. Vliet, G.C., and Leppert G., "Forced Convection Heat Transfer from an Isothermal Sphere to Water," Trans. ASME J. of Heat Transfer, C83, 1961, pp. 163-175.

17. Whitaker, S., "Forced Convection Heat Transfer Correlations for Flow in Pipes, Past Flat Plates, Single Cylinders, Single Spheres, and for Flow in Packed Beds and Tube Bundles," AIChE Journal, Vol. 18, No. 2, March, 1972, pp. 361-371.

18. White, F.M., Viscous Fluid Flow, McGraw-Hill Company, New York, NY, 1974.

19. Rohsenow, W.M., and Hartnett, J.P., (eds.), Handbook of Heat Transfer, McGraw-Hill Co., New York, NY, 1973.

20. Zemansky, M.W., _Heat and Thermodynamics, 5th Edition_, McGraw-Hill Co., New York, NY.

21. Streeter, V.L., _Fluid Mechanics_, McGraw-Hill Company, New York, NY, 1962.

22. Shames, I.H., _Mechanics of Fluids_, McGraw-Hill Company, New York, NY, 1962.

23. Happel, J., and Brenner, H., _Low Reynolds Number Hydrodynamics_ Noordhoff Int. Publishers, Leyden the Netherlands, 1973.

24. Epstein, P.S., "On the Resistance Experienced by Spheres in Their Motion Throuth Gasses," _Phys. Rev._ 23, 1924.

1981 NASA/ASEE SUMMER FACULTY
RESEARCH FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA-HUNTSVILLE

INCIPIENT FAILURE DETECTION (IFD)
OF SSME BALL BEARINGS

Prepared By: David Y. Wang

Academic Rank: Professor

University and Department: University of Wisconsin-Platteville
                                      Department of Civil Engineering

NASA/MSFC:
    Division: Structural Dynamics
    Branch: Environmental Analysis

MSFC Counterpart: J. H. Jones

Date: August 7, 1981

Contract No.:   NGT 01-008-021
                                      The University of Alabama in H'ville

INCIPIENT FAILURE DETECTION (IFD)
OF SSME BALL BEARINGS

by

David Y. L. Wang, Ph.D.
Professor of Civil Engineering
University of Wisconsin - Platteville
Platteville, Wisconsin

ABSTRACT

Because of the immense noise background during the operation of a large engine such as the SSME, the relatively low level unique ball bearing signatures were often buried by the overall machine signal. As a result, the most commonly used bearing failure detection technique, pattern recognition using power spectral density (PSD) constructed from the extracted bearing signals, is rendered useless.

In order to retrieve the deeply buried bearing defect signals from the background noise, a survey of the possible usage of a noise-cancelling filter [1]* in conjunction with two data enhancement techniques: time domain averaging (TDA)[2] and random decrement analysis (Randomdec)[3], was conducted.

TDA consists of averaging consecutive segments of the signal, one period apart. This is equivalent to cross-correlating the signal with a train of unit impulses having the same repetition rate. If this impulse train has the same repetition rate as the desired signal, the signal is, thus, extracted. Obviously, a prior knowledge of the desired signal period is required.

Randomdec averages segments of time history which start at a given amplitude. This averaging process yields a curve representing the response of a free, damped vibration system with an initial displacement. Advantages of using Randomdec to extract bearing defect signals are as follows:

1. Results do not vary with intensity of the input.

2. Randomdec eliminates the effect due to nonlinearity of the system.

The above mentioned data enhancement techniques were carried out by using a HP5451C Fourier Analyzer. The signal was preprocessed by a Digital Audio Corp. DAC-1024I noise cancelling filter in order to estimate the desired signal corrupted by the background noise. Reference levels of good bearings were established. Any deviation of bearing signals from these reference levels would indicate the incipient bearing failures.

---

*Numbers within the parentheses refer to the references at the back of the report.

ACKNOWLEDGEMENT

NOMENCLATURE

Symbols	Definition
$x(t)$	Complex signal.
$s(t)$	Periodic signal.
$n(t)$	Noise.
m, n	Number of samples or rolling elements of bearing.
T_o	Period of desired signal.
k, n	Positive integers.
f_t	Synchronizing or triggering frequency.
Δt	Sampling rate.
y_s	Selected amplitude of randomdec.
$\delta(\cdot)$	Randomdec function.
IFD	Incipient failure detection.
HPOTP	High pressure oxygen turbopump.
f_o, f_i, f_s	Ball pass frequency over outer, inner race, and ball spin frequency, respectively.
d	Rolling element diameter.
D	Bearing pitch diameter.
R	Shaft rotative speed in Hz.
α	Bearing contact angle

INTRODUCTION

Instrumentation for early-warning fault detection is needed in many areas such as air transportation and space vehicles. High reliability and safety standards are required in these areas in order to prevent catastrophic failures that may result in tragedy. Vibration and sound measurements have been used for the detection of faults in rotating machinery for many years. These set-ups, with their moderate amount of data reduction, have been proven beneficial for small scale machines.[4] As machines become larger and more complex, the conventional method of pattern recognition became inadequate because the desired component signals are often buried by the immense background noises. This report describes two methods to extract the desired bearing defect signals from a complex signal with a high level background noise.

The first method is a time domain averaging (TDA) technique with "order normalization".[5] This technique will increase the signal to noise ratio. It, therefore, can extract the periodic signals due to the bearing defects from the complex waveforms. This process is coherent, requiring that the period of the periodic bearing signals be known or assumed.

The second method (Randomdec)[6], averages segments of time history of the complex wave with a constant amplitude. This method assumes that the signal consists of a deterministic part (a free damped response due to an initial desplacement) and a random part which we can assume to have a zero mean. The averaging process eliminates the random part of the signal yielding a unique signature corresponding to the deterministic part of the signal. This signature is, then, compared to a standard signature and changes of the signature are interpreted in terms of possible bearing failure.

The above mentioned two methods were carried out using facilities of the environmental analysis branch of the System Dynamics Laboratory. The possible usage of either method as an on-line IFD of SSME ball bearings was examined accordingly.

OBJECTIVES

The objective of this project is to evaluate two data enhancement techniques (TDA and Randomdec) on the extraction of a desired periodic component from a complex waveform and large background noises. This evaluation will yield valuable information in the determination of adopting a possible IFD system for SSME ball bearings.

TIME DOMAIN AVERAGING (TDA)

Consider a signal, $x(t)$, where:

$$x(t) = S(t) + N(t) \qquad (1)$$

$S(t)$ is the desired signal with a known period T_O, and $N(t)$ is the background noise. When summing up subsequent $x(t)$, each one a period, T_O, apart, the periodic signal $S(t)$ will add coherently while noise $N(t)$ incoherently. After M summations, we get:

$$x(t_i) = m S(t_i) + \sqrt{m} \, N(t_i)^{(7)} \qquad (2)$$

therefore, the signal to noise ratio is enhanced by a factor of $\sqrt{m}$.

Summing m repetitions of the signal $S(t)$ is equivalent, mathematically, to convolve the signal, $S(t)$, with a train of m unit impulses at a time delay of the known period T_O. After averaging we get:

$$g(t) = \frac{1}{m} \int_{-\infty}^{\infty} S(t-\epsilon) \sum_{k=1}^{m} \delta(\epsilon - kT_O) d\epsilon$$

$$= \frac{1}{m} \sum_{k=1}^{M} S(t-kT_O) \qquad (3)$$

This linear averaging operation is equivalent to a comb filter with center frequencies of nft ($f_t = \frac{W_t}{2\pi}$), and is described by: (7, 8)

$$|H(jw)| = \frac{1}{m} \frac{\sin \pi \, m \, \frac{w}{W_t}}{\sin \pi \, \frac{w}{W_t}} \qquad (4)$$

$$\phi(jw) = -\pi(m-1) \frac{w}{W_t} \qquad (5)$$

Two techniques have been used in performing the above theory. The first of these is to determine the appropriate time shift according to the period, T_O, of the desired signal and sampling rate, Δt. Then we stack consecutive segments of signal, without any gap, m times successively. The second technique is the so-called "trigger and averaging". It is accomplished by generating a trigger signal that time locks m consecutive segments in the analog-to-digital data acquisition process. Block diagram of this process is shown in the following figure:
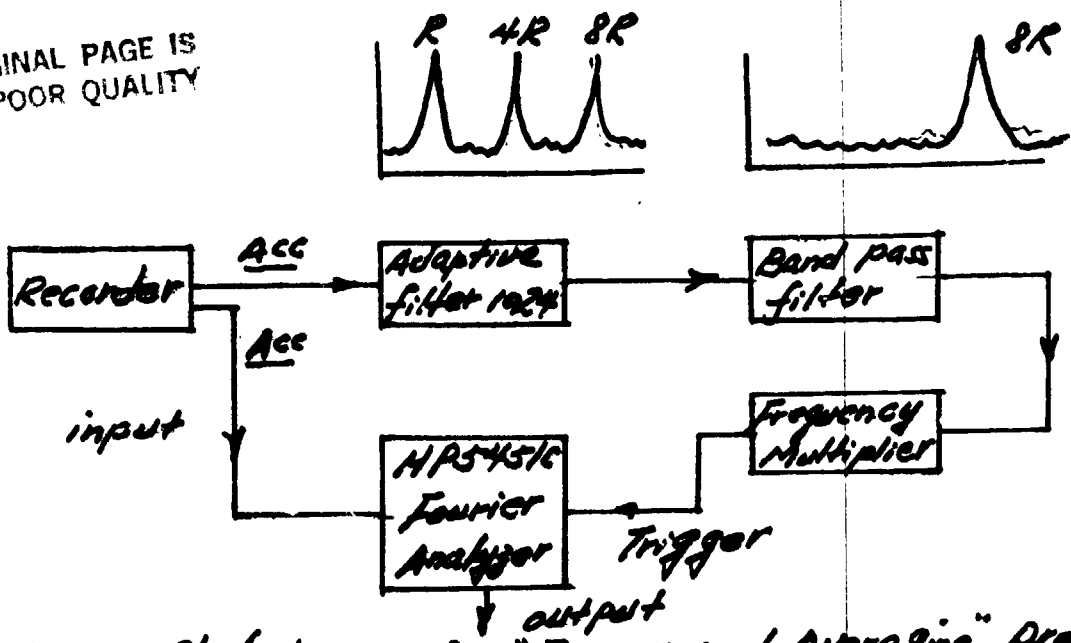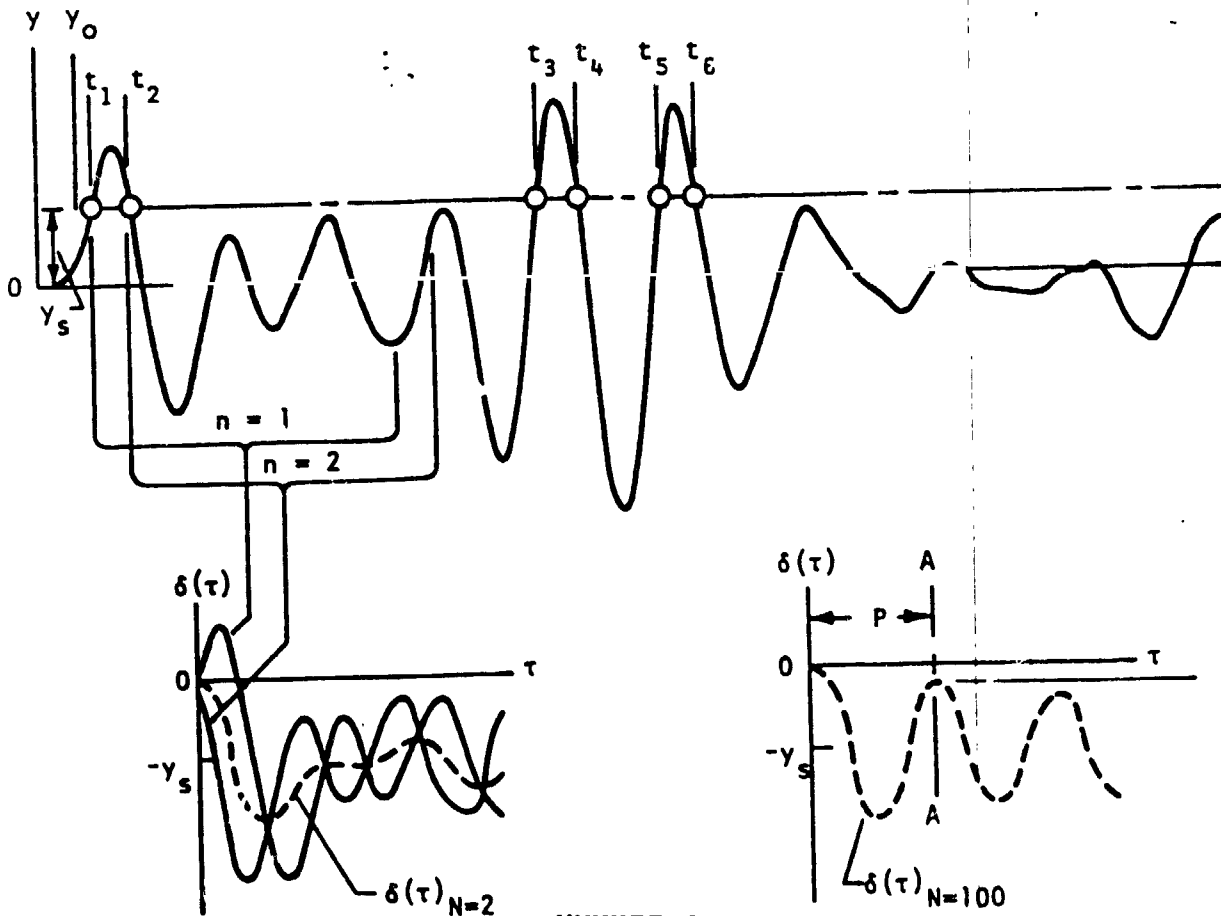
Fig. 1. Block diagram for "Trigger and Averaging" Process

RANDOM DECREMENT AVERAGING (RANDOMDEC)

Random signature is obtained by averaging segments of a random time history which start at a constant amplitude, y_s, as shown in Figure 2.

It can be defined as:

$$\delta(T) = \frac{1}{N} \sum_{n=1}^{n} y_o(t_n + T) \qquad (6)$$

Where: $\qquad y_o = y - y_s$ and

$\qquad t_n = t$ when $y_o = o$

As more and more samples are taken, the signature converges to a curve with a form as shown for N = 100. This curve can be considered as the free, damped, response of the system with an initial displacement, y_s. Even though, it has an appearance similar to the autocorrelation, but they are different in that the randomdec signature:

a). is independent of the input intensity.

b). has the same dimension as that of the time history.

Another interesting property of the randomdec is that if the desired signal is periodic and selected constant amplitude, y_s, equals to rms of the complex signal, then, the samples would be taken periodically.

To implement randomdec in failure detection, a possible setup is shown in Figure 3.
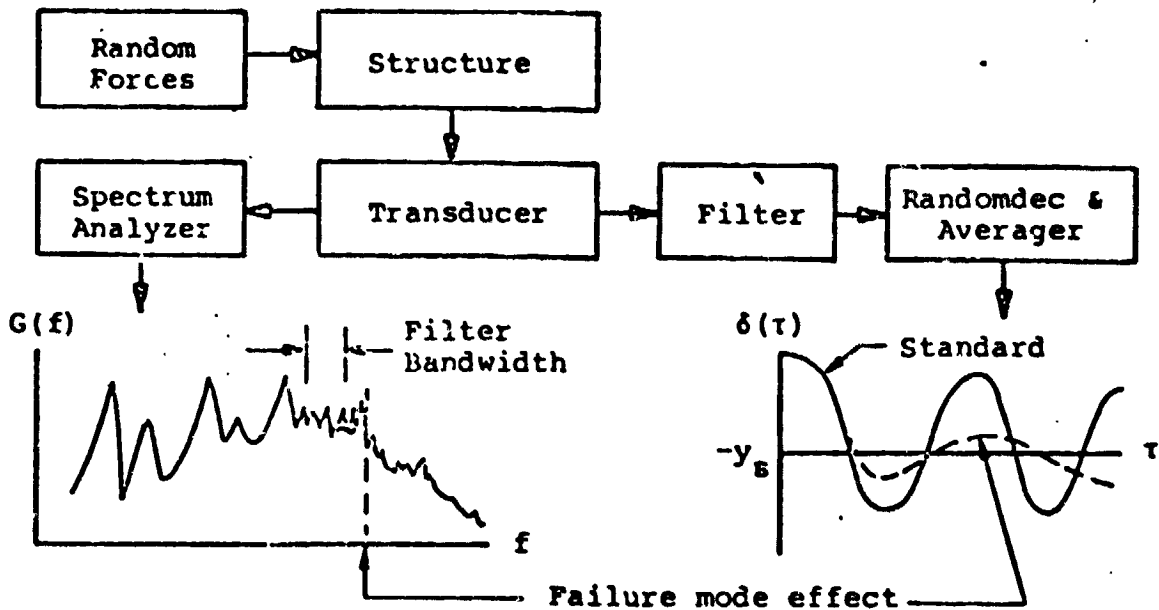
Fig. 3 Possible Set-up of Randomdec

The spectrum analyzer in the figure provides a broad view of the system response which may be used as an aid to specifying filtering requirements. When a failure or failures develops in the system, the number of degrees of freedom of the system will be increased so that the frequency of the failure mode is expected to decrease. Therefore, to detect ball bearing defect at an early stage, it is advantageous to apply the randomdec check in a high enough frequency such that corrective action can be taken and complete failure avoided.

DATA COLLECTION AND PROCESS

Failures of ball bearings within the high pressure oxy-
gen turbopump were experienced during tests of the SSME.
Spallings, damage cuts, pittings, and discoloration appear on
both inner and outer races, rolling elements and cages. The most
serious failure seems to occur at bearings closest to the turbine
of HPOTP. Data collected were usually low pass filtered at 10k
Hz or under, therefore, possibility of adopting various IFD
techniques in the higher frequency range (30kHz and higher) of
the SSMD bearings is precluded.

To evaluate the two data enhancement techniques, signals
were first processed by a adaptive noise-cancelling filter (DAC
1024) before feed into the HP5451C Fourier analyzer system to
obtain the final results. Programs used in this final data
reduction process are listed in the appendix. To implement the
"trigger and averaging" of TDA, the signal from a nearby accele-
rometer was conditioned and is used as the trigger source, while
a frequency multiplier is used to do the "frequency order norma-
lization" as shown in Figure 1.

In order to analyze the obtained results, ball pass frequen-
cies over defect, for the given bearing configuration at 100%
SSME power level, were computed according to the following equa-
tions:

$$f_o = \frac{z}{2} R (1 + \frac{d}{D} \cos\alpha) \tag{9}$$

$$f_i = \frac{z}{2} R (1 - \frac{d}{D} \cos\alpha)$$

$$f_s = \frac{D}{d} R (1 - (\frac{d}{D})^2 \cos^2\alpha) \tag{7}$$

DATA ANALYSIS

Three data reduction techniques were used in analyzing data obtained from known good bearings that were damaged in various degrees. They are: spectrum analysis (with adaptive filter), TDA, and randomdec. The general procedure in identifying bearing defects is to establish signatures of good bearings in either time or frequency domain as a reference of comparison. When bearing signatures start to deviate from the reference signature, it usually indicates that the failure mechanism of the bearing is in its developing stage. The extent of the bearing damages is suggested by the degree of deviation of bearing signatures from the reference signature.

Figure 4 shows a power spectral density (PSD) of a typical good bearing. The fundamental frequency and its higher harmonics are distinctly shown. It can be used as the reference signal for comparison. Figures 5 and 6 are PSD's for damaged bearings. Two observations can be made in comparing these two figures with that of Figure 4. The first is the disappearance of higher harmonics (4th harmonic and up) and the growth of the first two fundamental modes. These changes may be explained by the following scenario. As bearing defects start to grow, the failure modes gradually decrease because of the additional degrees of freedom due to the added constraints by the bearing defects. The higher harmonics, therefore, will vanish with the formation of the defects. Evergy associated with these higher harmonics will shift to the lower modes and, therefore, cause their growth. The second observation is the mushrooming of narrow band peaks in the range of bearing defect frequencies (5.5N to 8N approximately where N is the shaft speed of HPOTP). This is probably caused by the couplings of bearing defect and other periodic events or components during the operation of the SSME.

Figures 7 and 8 are the PSD's obtained by time domain averaging with a frequency sweep in the interested interval (2500 Hz to 4000 Hz). The increment of narrow band peaks in the ball pass frequencies over inner and outer races (approximately 3500 Hz, 2500 Hz) and ball spin frequency (3100 Hz) is obvious. Figure 9 simulates the application of randomdec in bearing defects detection. The decrease of failure mode as bearings start to fail is apparent. Reason for this phenomena is the same as explained previously.

CONCLUSIONS

Three data reduction and enhancement techniques for detecting bearing failures have been described in this report. Advantages and disadvantages of adopting each method as part of an on-line IFD system are discussed as follows.

1. Spectrum Analysis with an Adaptive Filter:

    a. It is easiest of the three to implement.
    b. It is incapable to predict the initial **stage** of bearing failures.

2. TDA:

    a. It is hardest of the three to implement.
    b. It is imperative to know the period of the extracted signal.
    c. It does give source(s) of bearing defects.

3. Randomdec:

    a. It is the most sensitive of the three in predicting the initial stage of bearing defects.
    b. Sources of system defects can not be pin-pointed.

One alternative method in ball bearing defects detection is to utilize the bearing high frequencies, especially for large machines such as SSME. Natural frequencies of ball bearings are usually far beyond that of the noice, therefore, less noice interference to the bearing signal.

REFERENCES

1. Widrow, B., Glover, J. R. et. al. "Adaptive Noise Can-
   celling: Principles and Applications, "Proceedings of
   the IEEE, Vol. 63, No. 12, Dec. 1975.

2. Braun, S., "The Extrachon of Periodic Waveforms by Time
   Domain Averaging." Acoustica, Vol. 32, No. 2, 1975.

3. Cole, H. A., Jr., "Failure Detection of A Space Shuttle
   Wing Flutter Model by Random Decrement," NASATM x-62041,
   May 1971.

4. Houser, D. R., "Signal Analysis Techniques for Vibration
   Diagonostics. "Proceedings of 22nd Meeting of the Mechanical
   Failures Prevention Group, Dec. 1975. PB248254.

5. Babkin, A. S., Anderson, J. J., "Mechanical Signature
   Analysis, "Sound and Vibration 7(4), April 1973.

6. Cole, H. A., Jr., "On-Line Failure Detection and Damping
   Measurement of Aerospace Structures By Random Decrement
   Signatures," NASA CR 2205, Mar. 1973.

7. Trimble, C. R., "What is Signal Averaging?" Hewlett-Packard
   Journal, April 1968,

8. Gold, B., Rader, C. M., "Digital Processings of Signals,"
   McGraw-Hill, New York.

9. Harris, T. A., "Rolling Bearing Analysis," Wiley, New York,
   1966.

Figure 4

6-9

Figure 5

Figure 6

Figure 7

Figure 8

Figure 9