NASA CR-61396

GEORGE C. MARSHALL SPACE FLIGHT CENTER

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

1972 NASA-ASEE
SUMMER FACULTY FELLOWSHIP
RESEARCH PROGRAM

RESEARCH REPORTS

Skylab

CONDUCTED BY

AUBURN UNIVERSITY
UNIVERSITY OF ALABAMA

SEPTEMBER 1972

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
NASA CR-61396		

4. TITLE AND SUBTITLE	5. REPORT DATE
1972 NASA-ASEE SUMMER FACULTY FELLOWSHIP PROGRAM — RESEARCH REPORTS	September 1972
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Submitted by Mr. J. Fred O'Brien, Jr., Dr. Donald C. Raney and Mr. Marion I. Kent	8. PERFORMING ORGANIZATION REPORT #
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. WORK UNIT NO.
Auburn University, Auburn, Alabama University of Alabama, Tuscaloosa, Alabama	11. CONTRACT OR GRANT NO. NGT-01-003-045
	13. TYPE OF REPORT & PERIOD COVERED
12. SPONSORING AGENCY NAME AND ADDRESS National Aeronautics and Space Administration Washington, D.C. 20546	NASA Contractor Report
	14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The 1972 NASA-ASEE Summer Faculty Fellowship Program (Research) conducted at George C. Marshall Space Flight Center was administered by Auburn University and the University of Alabama. This program was one of six such programs co-sponsored by the Office of University Affairs, NASA Headquarters, and the American Society of Engineering Education. Twenty-five professors participated in the 1972 program in which more than eighty percent of their time was spent in research activity within the laboratory to which they were assigned. This report is a compilation of the documented research of the individual professors.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
	Marion I. Kent Unclassified — Unlimited

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
Unclassified	Unclassified	676	NTIS

MSFC - Form 3292 (May 1969)

FACULTY FELLOWS

Seated, L - R: Dr. Judy M. Olson, Dr. Gary A. Sawyer, Dr. Gerald R. Karr, Dr. John W. Sheldon, Dr. Paul D. Smith, Dr. Amnon Sitchin, Dr. Duen-Ren Jeng, Dr. Murl W. Parker
Standing, L -R, 1st. row: Dr. Joel M. Elliott, Dr. Jerome C. Glaser, Dr. Donald J. De Smet, Dr. John H. Kusmiss, Dr. Frank E. Martin, Dr. Paul H. DeHoff, Dr. James E. Dudgeon, Dr. Edwin I. Griggs, Dr. Jerry D. Moore
Standing, L -R, 2nd. row: Mr. J. Fred O'Brien,* Dr. Keith H. Hawks, Dr. Richard A. Morrison, Dr. Fred E. Williams, Dr. Jerry W. Rogers, Dr. Jimmy L. Dodd

* Staff

Page Intentionally Left Blank

PREFACE

The 1972 ASEE-NASA Summer Faculty Fellowship Program (Research) conducted at George C. Marshall Space Flight Center was administered by Auburn University and the University of Alabama. This program was one of six such programs co-sponsored by the Office of University Affairs, NASA Headquarters, and the American Society of Engineering Education. The fundamental objectives of this program were to further the professional knowledge of qualified engineering and science faculty members, to simulate an exchange of ideas between participants and NASA, and to enrich and refresh the research and teaching activities of participants' institutions.

Twenty-five professors participated in the 1972 program in which more than eighty percent of their time was spent in research activity within the Laboratory to which they were assigned. This report is a compilation of the documented research of the individual professors. Because of the nature of the work performed by some of the Fellows, their report may not adequately describe the summer's activities. However, identification of their program is given in the title of their paper.

If further information is desired, contact the authors directly. Authors' addresses are at the end of this report.

J. Fred O'Brien, Jr., Director
Auburn University

Donald C. Raney, Assoc. Dir.
University of Alabama

Marion I. Kent
Assistant Director For
University Affairs
MSFC Representative

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

TIME-DEPENDENT STRUCTURAL ANALYSIS OF METAL-EPOXY COMPOSITES

Prepared by:	Paul H. DeHoff
Academic Rank:	Associate Professor
University:	Bucknell University
Laboratory: (Division) (Branch)	Astronautics Laboratory Analytical Mechanics Division Advanced Technology Branch
Research Counterpart:	John E. Key
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author considers himself fortunate to have had the opportunity to participate in the ASEE/NASA Summer Faculty Fellowship Program and wishes to express his appreciation to all the people associated with the program. He also wishes to thank his counterpart, Mr. John E. Key, as well as all members of the Advanced Technology Branch for their cooperation and friendship throughout the program.

TIME-DEPENDENT STRUCTURAL ANALYSIS OF METAL-EPOXY COMPOSITES

By

Paul H. DeHoff

ABSTRACT

In an effort to reduce structural weight while maintaining structural integrity on the Space Shuttle and other proposed space vehicles, various composite materials are being evaluated to determine their applicability and reliability in load carrying situations. While most of the composites being considered for structural components are of the metal-metal type, there are some applications for which metal-epoxy composites are proposed. Since the metal-epoxy composites demonstrate time-dependent behavior, it will be necessary to update the existing NASA structural analysis programs to include viscoelastic effects.

In this study an approximate method is presented to analyze plane stress orthotropic linear viscoelastic problems under isothermal conditions. A computer program which is based on the method was used to solve a simple uniaxial creep problem for a Maxwell type material. This solution is presented and compared with the analytical solution.

INTRODUCTION

It seems apparent that for the foreseeable future one of the major design concerns for space vehicles will be a continuing search for new and improved lightweight materials which can be used in structural applications. One of the relatively recent advances has been the development of advanced filementary laminated composites. These composites are currently being evaluated to determine their applicability and reliability in load carrying situations.

Most of the composite materials being considered for structural components are of the metal-metal type and are, therefore, amenable to structural analysis using available elastic-plastic finite element structural analysis programs such as NASA's NASTRAN general purpose program. However, there have been some applications for which metal-epoxy composites have been proposed and these composites are known to demonstrate some time-dependent material behavior. Thus, if these components are to be properly analyzed, it will be necessary to update the existing structural analysis programs to account for viscoelastic effects.

The present study is intended to be an initial step toward the eventual inclusion of a complete thermal viscoelastic option into the existing finite element programs. Although the program developed here is limited to plane stress orthotropically linear viscoelastic materials under isothermal conditions, it is believed that thermal effects and additional cases can be added with some additional effort.

PRELIMINARY THEORY

General Viscoelastic Equations

We assume that the time-dependent behavior of each layer of the metal-epoxy composite can be characterized as a linear homogeneous anisotropic material having orthotropic symmetry. In that case, the isothermal stress-strain behavior can be described by a constitutive equation of the type

$$\sigma_{ij}(t) = 2 \int_{-\infty}^{t} \left[\phi_1(t-\tau) \frac{\partial \varepsilon_{11}(x_k,\tau)}{\partial \tau} + \phi_2(t-\tau) \frac{\partial \varepsilon_{22}(x_k,\tau)}{\partial \tau} \right.$$

$$\left. + \phi_3(t-\tau) \frac{\partial \varepsilon_{12}(x_k,\tau)}{\partial \tau} \right] d\tau \, \delta_{ij} + 2 \int_{-\infty}^{t} \phi_4(t-\tau) \frac{\partial \varepsilon_{ij}(x_k,\tau)}{\partial \tau} +$$

$$+ \phi_4(t-\tau)\frac{\partial \varepsilon_{22}(x_k,\tau)}{\partial \tau} + \phi_5(t-\tau)\frac{\partial \varepsilon_{33}(x_k,\tau)}{\partial \tau}\Big] d\tau\, m_{ij}$$

$$+ 2\int_{-\infty}^{t}\Big[\phi_3(t-\tau)\frac{\partial \varepsilon_{11}(x_k,\tau)}{\partial \tau} + \phi_5(t-\tau)\frac{\partial \varepsilon_{33}(x_k,\tau)}{\partial \tau}$$

$$+ \phi_6(t-\tau)\frac{\partial \varepsilon_{33}(x_k,\tau)}{\partial \tau}\Big] d\tau\, n_{ij} + 2\int_{-\infty}^{t}\phi_1(t-\tau)\Big[\delta_{i1}\frac{\partial \varepsilon_{12}(x_k,\tau)}{\partial \tau} m_{2j}$$

$$+ m_{i2}\frac{\partial \varepsilon_{21}(x_k,\tau)}{\partial \tau}\delta_{1j}\Big] d\tau + 2\int_{-\infty}^{t}\phi_3(t-\tau)\Big[\delta_{i1}\frac{\partial \varepsilon_{13}(x_k,\tau)}{\partial \tau} n_{3j}$$

$$+ n_{i3}\frac{\partial \varepsilon_{31}(x_k,\tau)}{\partial \tau}\delta_{1j}\Big] d\tau + 2\int_{-\infty}^{t}\phi_4(t-\tau)\Big[m_{i2}\frac{\partial \varepsilon_{23}(x_k,\tau)}{\partial \tau} n_{3j}$$

$$+ n_{i3}\frac{\partial \varepsilon_{32}(x_k,\tau)}{\partial \tau} m_{2j}\Big] d\tau \qquad (1)$$

where $$\varepsilon_{ij}(x_k,\tau) = \frac{1}{2}\Big[\frac{\partial u_i(x_k,\tau)}{\partial x_j} + \frac{\partial u_j(x_k,\tau)}{\partial x_i}\Big] \qquad (2a)$$

$$\delta_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};\quad m_{ij} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},\quad n_{ij} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (2b)$$

and the ϕ_i's are the six independent relaxation functions describing the material behavior.

If we consider only deformations that start at time equal to zero and if we further restrict the study to plane stress situations, then it can be shown that equations (1) reduce to the following form upon integration by parts.

$$\sigma_{11}(t) = \ldots \qquad (3a)$$

$$\sigma_{22}(t) = \ldots \qquad (3b)$$

$$\sigma_{12}(t) = \ldots \qquad (3c)$$

With these basic relations we now proceed to develop the equations necessary to apply the finite element method to time dependent deformations.

Finite Element Development

While several approaches to the development of the necessary finite element equations are possible, an approximate technique similar to one reported by White (1) is used because of its adaptability to thermoviscoelasticity and acceptance of relaxation information in numerical form.

In order to simplify the development, we consider only isothermal deformations and consider the body forces to be zero. The principle of virtual work can then be written over a time interval Δt in the matrix form:

$$\int_t^{t+\Delta t} \left\{ \int_{VOL} \delta\underline{\varepsilon}^T(x,y,z,t)\,\underline{\sigma}(x,y,z,t)\,dV - \int_{S_2} \delta\underline{u}^T(s,t)\,\underline{F}_s(s)\,ds \right\} dt = 0 \qquad (4)$$

where $\underline{F}_s$ represents the surface forces and $\delta\underline{u}$ the virtual displacements. For a plane stress constant strain triangular element having the geometry shown in figure 1, the strains can be written in terms of the nodal displacements $\underline{\delta}$ by

$$\underset{\sim}{\varepsilon}(t) = \underset{\sim}{B}\,\underset{\sim}{\delta}(t) \tag{5}$$

where

$$\underset{\sim}{B} = \frac{1}{2A}\begin{bmatrix} Y_{bc} & 0 & Y_{ca} & 0 & Y_{bc} & 0 \\ 0 & X_{cb} & 0 & X_{ac} & 0 & X_{ba} \\ X_{ca} & Y_{bc} & X_{ac} & Y_{ca} & X_{ba} & Y_{ab} \end{bmatrix} \tag{6a}$$

$$Y_{bc} = Y_b - Y_c \;,\; X_{bc} = X_b - X_c \quad \text{, etc.} \tag{6b}$$

and A = area of element triangle.

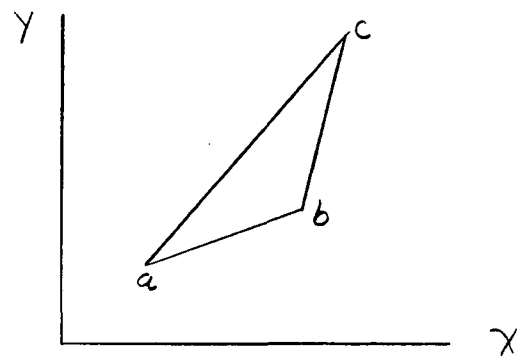


Fig. 1 - Element Geometry

For ease of notation we write equation (3) in the nonseparable operator form

$$\underset{\sim}{\sigma}(t) = \underset{\sim}{\mathcal{L}}\,\underset{\sim}{\varepsilon}(t) = \underset{\sim}{C}\,\underset{\sim}{\xi}(x_A, t) + \underset{\sim}{\mathcal{L}}\,\varepsilon(x_A, \tau) \tag{7}$$

where $\underline{\varepsilon}(x_i, t_0)$ represents the initial elastic response and $\underline{\varepsilon}(x_i, \tau)$ represents the time contribution. Using this notation and introducing the nodal displacements, equation (4) becomes

$$\int_{t_1}^{t_2} \{\delta\underline{\delta}^T [\int_{vol} \underline{B}^T \underline{L} \underline{\varepsilon}(x_i, t) \, dv] - \delta\underline{\delta}^T \underline{F}(t)\} \, dt = 0 \qquad (8)$$

where we have replaced the surface forces with statically equivalent nodal forces. For arbitrary nodal displacements we then obtain

$$\int_{vol} \underline{B}^T \underline{L} \underline{\varepsilon}(t) \, dv = \underline{F}(t) \qquad (9)$$

This is in a form similar to the usual elastic analysis except that the operator form must be integrated over time. This integration would normally require that the relaxation functions, as well as the strains, be expressible in analytical form. However, since we are interested in having numerical data as input, it is convenient to use a numerical integration technique. Toward that end we consider only the x or r component of the "pseudo" operator. For example, utilizing equations (3) and taking the x component we have

$$[\underline{L}\underline{\varepsilon}(t)]_x = 2[\theta(0)\varepsilon_x(t) + \theta_1(0)\varepsilon_x(t)$$

$$- \int_0^t [\frac{\partial \theta_1(t-\tau)}{\partial \tau} \varepsilon_x(x_i, \tau) + \frac{\partial \theta_2(t-\tau)}{\partial \tau} \varepsilon_x(x_i, \tau)] \, d\tau] \qquad (10)$$

Now the derivative of the relaxation function can be approximated by a straight line between any two discrete times t_i and t_{i+1} as shown in figure 2.

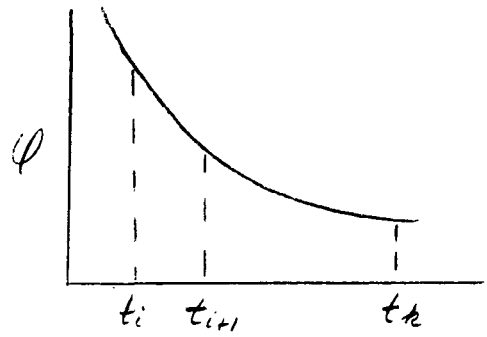


Fig. 2 - Typical Relaxation Function

It can be shown that this approximation leads to

$$\left.\frac{\partial \varphi}{\partial \tau}\right|_{t_i}^{t_{i+1}} \simeq \frac{\varphi(t_k - t_{i+1}) - \varphi(t_k - t_i)}{t_{i+1} - t_i} \tag{11}$$

where the present time has been taken equal to t_k .

The integral term can now be evaluated by successive application of the trapezoidal rule. This yields

$$\int_0^{t_k} \left[\underset{\sim}{\mathcal{L}}\, \underset{\sim}{\varepsilon}(\tau)\right]_{II} d\tau \simeq \sum_{i=1}^{k-2} \left\{ \sum_{r=1}^{2} \left[\frac{\varphi_r(t_k - t_{i+1}) - \varphi_r(t_k - t_i)}{2}\right]\left[\varepsilon_{rr}(t_{i+1}) + \varepsilon_{rr}(t_i)\right]\right\} + \sum_{r=1}^{2} \left[\varphi_r(0) - \varphi_r(t_k - t_{k-1})\right]\left[\frac{\varepsilon_{rr}(t_k) + \varepsilon_{rr}(t_{k-1})}{2}\right] \tag{12}$$

Incorporating the initial response from equation (10) and rearranging terms yields finally

$$[L'\varepsilon(t_k-)] = \sum_{r=1}^{2}\left\{\left[\frac{\phi_r(0)+\phi_r(t_k-t_{k-1})}{2}\right]\varepsilon_{rr}(t_k) - \left[\frac{\phi_r(0)-\phi_r(t_k-t_{k-1})}{2}\right]\varepsilon_{rr}(t_{k-1})\right\} - \sum_{i=1}^{k-2}\left[\sum_{r=1}^{2}\left\{\left[\phi_r(t_k-t_{i+1}) - \phi_r(t_k-t_i)\right]\varepsilon^*_{rr}(t_i)\right\}\right] \quad (13)$$

where $$\varepsilon^*_{rr}(t_i) = \frac{\varepsilon_{rr}(t_{i-1}) + \varepsilon_{rr}(t_i)}{2} \quad (14)$$

We define the relaxation function differences by

$$\Phi_m(t_k) = \frac{1}{2}\left[\phi_m(0) - \phi_m(t_k-t_{k-1})\right] \quad k = 2,3,\ldots \quad (15a)$$

$$\Psi_m(t_k) = \frac{1}{2}\left[\phi_m(0) + \phi_m(t_k-t_{k-1})\right] \quad k = 2,3,\ldots \quad (15b)$$

$$\Phi_m(t_i) = \left[\phi_m(t_k-t_{i+1}) - \phi_m(t_k-t_i)\right] \quad k = 3,4\ldots \quad (15c)$$

By a similar procedure the remaining terms can be found so that we finally obtain.

$$\underline{\sigma}(t_k) = \underline{B}^T\underline{\Psi}_m(t_k)\underline{B}\,\underline{\varepsilon}(t_k) - \underline{B}^T\underline{\Phi}_m(t_k)\underline{B}\,\underline{\varepsilon}(t_{k-1}) - \sum_{i=1}^{k-2}\underline{B}^T\underline{\Phi}_m(t_i)\underline{B}\,\underline{\varepsilon}^*(t_i) \quad (16)$$

where the relaxation function difference matrices are given by

$$\underset{\sim}{\Phi}_m = \begin{bmatrix} \Phi_1 & \Phi_2 & 0 \\ \Phi_2 & \Phi_4 & 0 \\ 0 & 0 & \Phi_7 \end{bmatrix} \tag{17a}$$

$$\underset{\sim}{\Psi}_m = \begin{bmatrix} \Psi_1 & \Psi_2 & 0 \\ \Psi_2 & \Psi_4 & 0 \\ 0 & 0 & \Psi_7 \end{bmatrix} \tag{17b}$$

In equation (16) we can identify time dependent equivalent stiffness matrices. These are

$$\underset{\sim}{K}_{\Psi}(t_k) = \underset{\sim}{B}^T \underset{\sim}{\Psi}_m \underset{\sim}{B} \tag{18a}$$

$$\underset{\sim}{K}_{\Phi}(t_k) = \underset{\sim}{B}^T \underset{\sim}{\Phi}_m \underset{\sim}{B} \tag{18b}$$

$$\underset{\sim}{K}_{\Phi}(t_i) = \underset{\sim}{B}^T \underset{\sim}{\Phi}_m(t_i) \underset{\sim}{B} \tag{18c}$$

Upon substitution of equations (18) into equation (15), the force-displacement relation takes a form similar to the usual finite element relation. This becomes

$$\underset{\sim}{F}(t_k) = \underset{\sim}{K}_{\Psi}(t_k)\,\underset{\sim}{\varepsilon}(t_k) - \underset{\sim}{K}_{\Phi}(t_k)\,\underset{\sim}{\varepsilon}(t_{k-1}) - \sum_{i=1}^{k-2} \underset{\sim}{K}_{\Phi}(t_i)\,\underset{\sim}{\varepsilon}^*(t_i) \tag{19}$$

The solution for the forces now requires an iterative procedure to evaluate the stiffness matrices and the nodal displacements at each time increment. In the next two sections we discuss briefly the programming technique and a simple problem which was used to test the method.

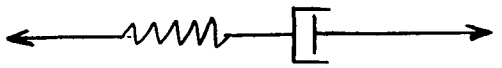
COMPUTER PROGRAM

The basic program which was used in this study is a modification of a modular program reported in Zienkiewicz (2) for plane strain linear elastic analysis. The program was modified to accept orthotropic relaxation data and to perform the time iterations necessary to determine the displacements. The program in its present form requires that the time increments be equally spaced real time intervals. This, of course, is a serious limitation since the viscoelastic response is generally logarithmic; however, it is anticipated that a modification of the program to accept log time increments can be made. This would require that an interpolation procedure be introduced to evaluate the relaxation functions between input data points.

The actual program is not included in this report, however, a complete listing will be available through the author or through Mr. John Key of MSFC.

SAMPLE PROBLEM

In order to test the procedure we now look at a simple uniaxial tension creep test of an orthotropic linear viscoelastic material with relaxation functions which are describable by a simple Maxwell spring-dashpot model. (Figure 3) The properties of the model are chosen arbitrarily and are not intended to represent any real material. The simple model is chosen so that an analytical solution is readily available for comparative purposes.



$$\phi_1(t) = \phi_4(t) = 10^{-5} e^{-t/10} \; ; \quad \phi_2(t) = \tfrac{1}{2}\phi_1(t), \; \phi_3(t) = \tfrac{1}{4}\phi_1(t)$$

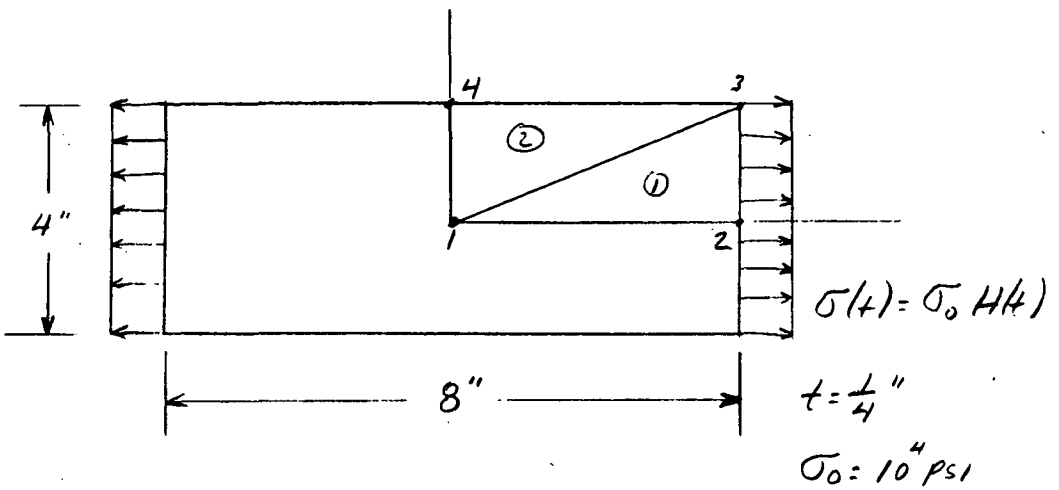


Fig. 3 - Creep Tension Plate and Maxwell Model

It can be shown that the solution for the uniaxial creep strain is given by

$$\varepsilon_{11}(t) = \frac{4}{30}\left[1 + \frac{t}{10}\right] (\text{in/in}) \tag{20}$$

where the units have been selected so that t is in seconds. Due to symmetry, only 1/4 of tension plate was modeled, and two simple triangles were used as shown in figure 3. The comparison with the analytical solution is shown in figure 4 for a period of 5 seconds.

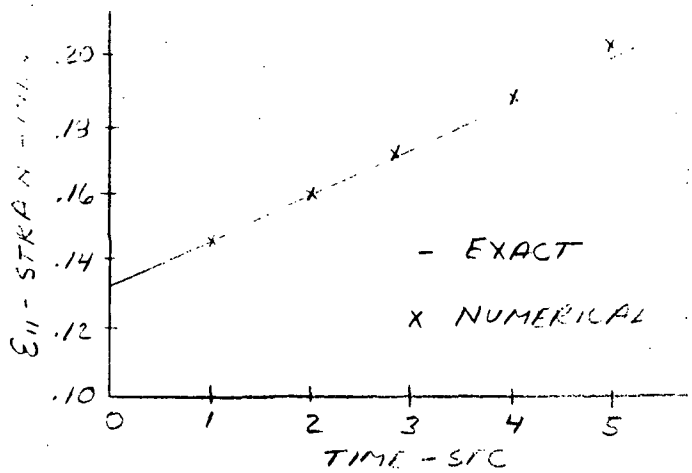


Fig. 4 - Comparison Between the Finite Element Method and Analytical Solution for Simple Creep

As can be seen from figure 4, the finite element procedure deveioped in this study yields strains for this simple creep problem which are slightly higher than the exact values at the longer times. This discrepancy is introduced because the slopes of relaxation curves have been approximated by assuming the functions to be represented by linear functions whereas the actual functions were taken to be exponential. It is believed that the accuracy can be improved by taking smaller time intervals or by using a higher order fit as an approximation to the relaxation functions. Unfortunately, time limitations precluded any additional effort along these lines.

CONCLUSIONS AND RECOMMENDATIONS

This study has shown that the finite element method can be successfully applied for the determination of stresses and strains in anisotropic viscoelastic materials. The program which was developed is presently limited to plane stress isothermal deformations and, in fact, has only been used to solve simple uniaxial creep and relaxation type problems. Thus, before any real analysis of the usefulness of the program can be made, it would be necessary to test it with more complex geometries using a large number of elements. Moreover, if only isothermal deformations were of interest, then approaches other than the one pursued here would be available. As was previously stated, this method was chosen because of its usefulness for thermal viscoelastic strains, and it is anticipated that this area will be further pursued.

REFERENCES

1. White, J. L., "Finite Elements in Linear Viscoelasticity," AFFDL-TR-68-150, Boeing Company, Kent, Washington.

2. Zienkiewicz, O. C., "The Finite Element Method in Engineering Science," McGraw-Hill, 2nd Edition, 1971.

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

A STUDY OF THE GROWTH OF SINGLE CRYSTALS

Prepared by:	Donald J. DeSmet
Academic Rank:	Assistant Professor
University:	University of Alabama
Laboratory: (Division) (Branch)	Space Sciences Laboratory Space Thermophysics Division Thermal Enviornment Physice Branch
Research Counterpart:	Mirt C. Davidson
Date:	August 11, 1972
Contract No:	NGT-01-003-045

A STUDY OF THE GROWTH OF SINGLE CRYSTALS

By

Donald J. DeSmet

ABSTRACT

Research on the growth of single crystals of the III-V compound indium antimonide has been pursued in order to gain a firsthand knowledge of some of the factors which govern the growth of single crystals of this compound. These studies were conducted using a standard Czochralski crystal pulling furnace. This type of furnace allows crystals to be formed without being constrained to a crucible. Several factors were fornd to have a large detrimental effect on the growth of single crystals of indium antimonide, the most important of which seems to be the degree to which the enviornment surrounding the crystal durning growth can be controlled. Ways in which to overcome these effects have been determined and are outlined.

INTRODUCTION

The phenomina of crystalization and the study of single crystals has become increasingly important with the advancement of technology. Single crystals of many materials have been admired and considered valuable for as long as man has appreciated beauty and symmetry. As technology has advanced naturally occuring single crystals have found many technological and scientific applications. Indeed, in many cases scientific advancement would have been greatly hindered if natural single crystals had not been available. One typical example of this is the study of the polarization of light and the interaction of polarized light with matter. If natural crystals of calcite, quartz, and other anisotropic transparent materials had not been avialable many important scientific discoveries would have been delayed. Even today the quality of natural crystals of optical materials is such that most precision polarization instruments use natural rather than artificial crystals for optical components.

In recent years a need for single crystals of materials not commonly found in nature has developed. This need has come about in part because of technological advances(for example many solid-state devices are manufactured from single crystals of semi-conductors), and in part because scientific investigations of the solid state are at present baised for the most part on the behavior of single crystals. In addition the actual process of single crystal formation is of scientific interest.

PRESENT WORK

The present work was undertaken in order ot obtain firsthand knowledge of some of the factors which influence the growth of crystals of the III-V compound indium antimonide. Indium antimonide was chosen it is in many ways typical of III-V compounds(which are interesting as a group because of their wide range of physical properties), and because neither component of this compound is particularly volitile.

The materials used to initially compound the indium antimonide used in this work were 99.999% pure indium powder and 99.999% pure antimony powder. Equal atomic parts of these two powders were mixed and subsequently melted under an argon atmosphere. Since the constituent materals used were powders the major impurity in the initial compound was oxide. Because this oxide tends to gather in small pockets

and move toward the surface of the compound when it is molten, the following procedure was used to remove this impurity:

The initially formed compound was immersed briefly in a mixture of 4 parts nitric acid, 1 part glacial acetic acid, and 1 part nitric acid. This removed the surface oxide and polished the indium antimonide so that any small pockets of oxide which had drifted to the surface appeared as dull grey patches. After rinsing, these patches were removed with a small wire brush and the material was given a second chemical polish in order to remove any surface impurities introduced by the brush. The material was then melted again, causing some of the oxide remaining in the material to move to the surface. This entire process was then repeated until there was no evidence of further oxide in the material.

The furnace used for pulling crystals, commonly known as a Czochralski crystal puller or Czochralski furnace, is shown schematically in Figure 1. The large chamber, A, is water cooled in order to prevent external parts from overheating. For optimum operation it should be airtight and capable of holding a vacuum of about 10^{-5} torr. The heating element, B, was a large tungstern filiment with a current carrying capability of 2000 Amperes. A graphite crucible, C, was used to hold the melted indium antimonide, D. When growing a crystal from the melt a seed crystal which has been suitabley oriented, E, is attached to a shafe, F, which can be rotated at a controlled rate in order to stir the melt and in addition can be raised or lowered at a controlled rate. A temperature sensor, T, is needed to measure the temperature of the melt, and a small water-cooled window is included so that the growth process can be observed, Not shown in the figure but necessary for operation are a quartz support stand for the crucible and a temperature regulator.

In theory the growth of a single crystal using this technique is relatively simple. The material to be grown is melted and maintained at a suitable temperature, usually a few degrees above the melting point, by means of the external temperature regulator. The seed is lowered until it touches the melt at the center of the crucible, and a small portion of the seed is melted. The seed rotation is then started and the seed is slowly withdrawn from the melt. The temperature, pull rate, and rotation are adjusted so that a crystal of the desired size is formed on the end of the see. Since, ideally, the crucible and the melt are the hottest parts in the furnace, the latent heat of fusion is conducted away through the crystal and the support shaft. Since the crystal can only grow as rapidly as the heat of fusion is removed, the thermal properties of the crystal and the temperature and temperature gradients are important factors in determining the rate of growth of the crystal. For this reason it is essential the

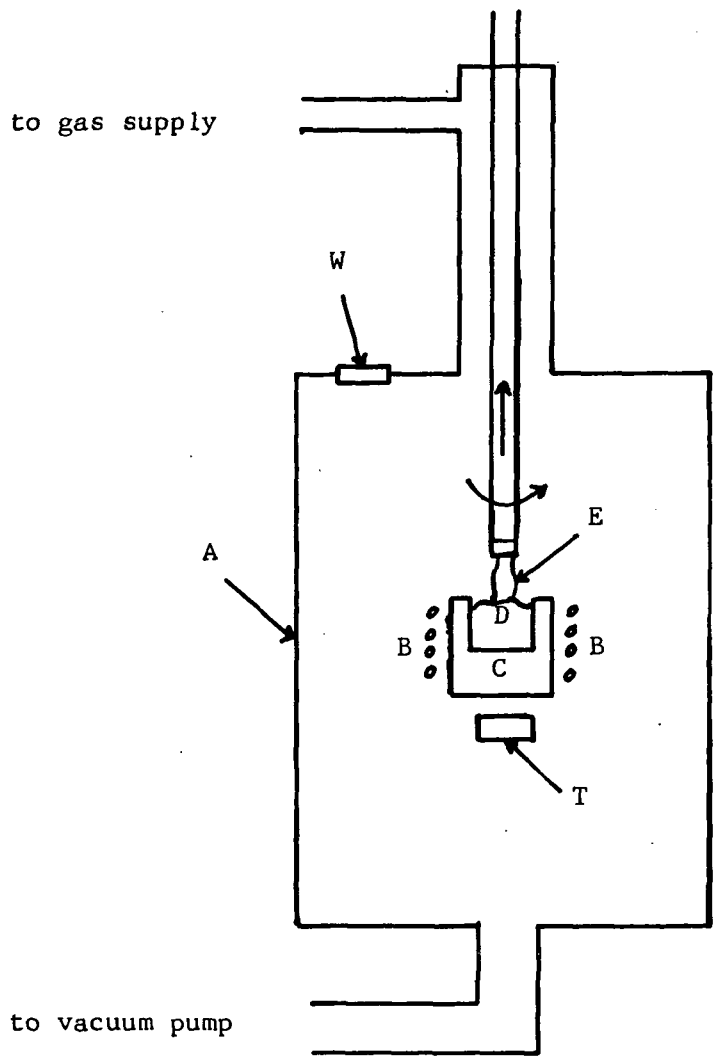


Figure 1. Crystal Growing Furnace (schematic).

temperature within the furnace be maintained at an accurate and constant value. If the temperature becomes too high, heat will not be conducted away rapidly enough and the crystal will seperate from the melt or, worse, the seed may melt. If the temperature becomes too low the melt itself will solidify.

In practice the growth of a single crystal of indium antimonide is not quite this simple. One complicating feature is that any oxide which forms on the surface of the melt does not evaporate at the melting point. This means that any oxide formed on the surface of the melt during the melting procesess will remain on the surface of the melt and act as a contaminant. In addition, as noted by Haggerty and Wenckus(reference 1), any oxide present on the surface of the melt may cause a polycrystaline solid to be formed rather than a single crystal. Therefore the atmosphere which surrounds the crystal during growth must be accurately controlled, especially with regard to oxygen and water vapor content.

Another complicating feature which is typical of III-V compounds is that they are polar. Indium antimonide has a ZnS crystal structure, and an analysis of this structure shows that planes such as the (III) and the ($\bar{1}\bar{1}\bar{1}$) are not equivalent. For indium antimonide the surfaces terminating with indium atoms are designated an ($\bar{1}\bar{1}\bar{1}$), whereas those terminating with antimony atoms are designated as (III). These crystal faces differ both chemically and physically, although with commonly used crystalographic techniques such as x-ray diffraction one cannot readily distinguish these two faces. For the growth of single crystals in th (III) direction the end of the seed terminating in antimony atoms must be used. If the end terminating in indium atoms is used the solid formed will be polycrystaline. The correct orientation of the seed can only be accomplished by the combined use of x-ray differaction (to determine the III orientation)and a chemical etch developed by Lavine, Gatos, and Finn(reference 2)to seperate the (III) and ($\bar{1}\bar{1}\bar{1}$) directions.

CONCLUSIONS AND RECOMMENDATIONS

Several attemps were made to grow single crystals of indium antimonide using both oriented single crystal seeds and polycrystaline seeds. These attempts were partially successful in that several pieces of polycrystaline material with large single crystal grains were produced, although no single crystals have yet been grown. The reasons for this lack of complete success have been determined. They are the lack of a completely controlled enviornment in the crtstal growing furnace and inadequate control of the temperature. The first of these

can be corrected by producing a better seal between the interior and exterior of the furnace and the second can be corrected by using a more sensitive temperature sensor and a more accurate temperature control. Steps have been taken to correct both of these problems and work now in progress should prove fruitful in producing good single crystals of indium antimonide and other materials which can be grown by the Czochralski technique.

REFERENCES

1. J. S. Haggerty and J. F. Wenckus, Development of Techniques for Growth of Bulk Single Crystals of Several III-V Compound Semiconductors, Final Technical Report prepared under Contract No. NAS 12-2020, Arthur D. Little, Inc., Cambridge, Mass. (June 1970).

2. M. C. Lavine, H. C. Gatos, and M. C. Finn, J. Electrochem. Soc., 108, 974 (1961).

3. J.J. Gilman, Editor, The Art and Science of Growing Crystals, John Wiley & Sons, London and New York (1963).

4. R. F. Strickland-Constable, Kinetics and Mechanism of Crystallization, Academic Press, London and New York (1968).

1972
ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

SHAPED BEAM KU-BAND ANTENNA DESIGN

Prepared by:	James E. Dudgeon
Academic Rank:	Assistant Professor
University:	University of Alabama (Tuscaloosa)
Laboratory: (Division) (Office)	Astrionics Laboratory Instrumentation and Communication Advanced Studies and Analysis
NASA Counterparts:	Grady H. Saunders Harlan Burke
Date:	August 11, 1972
Grant No.:	NGT-01-003-045

TABLE OF CONTENTS

SHAPED BEAM KU-BAND ANTENNA DESIGN

By James E. Dudgeon

ABSTRACT

Antenna systems for a Ku-band synchronous orbit communications satellite were considered. Desired was a wideband antenna system which is capable of simultaneously, or individually, illuminating the four Continental U.S. (CONUS) time zones and possibly Alaska and Hawaii. Possible applications would be for national or regional educational TV, medical information centers and consultation networks, disaster warning, law enforcement networks, and remote area broadcast coverage.

A multiple beam offset feed cluster used in conjunction with an offset parabolic reflector was chosen for the antenna realization. For radiating elements in the feed array, polyrod or dielectric rod radiators fed by waveguide are investigated. Polyrod antennas shrink the cross-section of the waveguide required, are fairly directive, and can be tapered to reduce sidelobes. As a consequence, it is thought that mutual coupling effects between adjacent elements will be minimal and the resultant physically small elements will allow greater packing densities for the drive elements in the feed array than would otherwise be possible.

Aperture distributions and corresponding ground patterns (footprints) produced by polyrod elements in an offset feed structure are found. Also the propagation from a ground area back onto the feed aperture is formulated and related to the use of physically realizable feed elements. The problems of high density feed elements, maintaining low sidelobe levels to meet international standards, aperture blocking, and feed placement (F/D) tradeoffs are considered.

CHAPTER 1

INTRODUCTION

Antenna systems for a Ku-band synchronous orbit communications satellite are considered. Desired is an antenna system which is capable of simultaneously, or individually, illuminating the four continental U.S. (CONUS) time zones and possibly Alaska, Hawaii, and Puerto Rico. Possible applications would be for national or regional educational TV, medical or other service information centers and consultation networks, disaster warning, law enforcement networks, and remote area broadcast coverage. Design objectives are wide bandwidths (≥500 MHz), at least one kilowatt of transmitted power per time zone, subregional coverage options, low sidelobe levels to meet international standards, and compatibility as a payload of an Atlas/Centaur or like launch vehicle (1360kg) [1].

A multiple beam offset feed cluster used in conjunction with a parabolic reflector was chosen for the antenna realization. For radiating elements in the feed array dielectrically loaded circular horns or dielectric rod radiators fed by waveguide are used, polyrod antennas shrink the cross-section of the waveguide required, are fairly directive, and can be tapered to obtain desired beamwidths and reduce sidelobes. It is thought that the serious problem of mutual coupling effects between adjacent elements can be minimized. The resultant physically small elements will allow greater packing densities for the radiators in the feed array than would otherwise be possible.

Among the antenna system that can be used for this application are (1) lenses [2], (2) phased arrays, [3], (3) parabolic dishes, and (4) special aperture configurations [4]. To obtain the beamwidths and pattern shapes desired at Ku-band, the lens appears to be plagued with design, construction, and weight problems. A Butler type phased array with multiple orthogonal beams would work, provided highly directive individual elements could be realized and provided the expense and loss factors at Ku-band could be overcome. The parabolic reflector with a multiple beam offset feed array is thought to be advantageous because of the structural simplicity, the considerable experience associted with its widespread usage, and the fact that it is amenable to low loss, high power waveguide feed elements. However, the needs of the shaped-beam application introduce several new problems, and their solution will be the undertaking of this report. As will be seen, the concept for illuminating odd-shaped

regions on the earth is that of superposition of multiple independent beams, individually pointed at different coordinates. The separate beams might have a 3 dB beamwidth on the order of 0.5 degree. For an equatorial orbit satellite at 129^{o} W. and boresighted at 17^{o} N, 118^{o} W, the required beam pointing deviations are $\pm 6.2^{o}$. To obtain good focusing and pointing accuracy for off boresight beams it is found that F/D's greater than 0.8 are preferred, and, consequently, to achieve $\pm 6.2^{o}$ squints, an aperture blockage of about 4% would result for symmetrical paraboloids with the beams aimed at the vertex. To avoid aperture blockage, offset reflector geometries are recommended. Beams are all aimed above the vertex and only an upper portion of the paraboloid is used as the reflector. Off focus placement of the feed elements is used to point the beams, and for large F/D the squint angle of the beam is approximately equal to the angle between the focal axis and the ray from the feed element to the vertex of the reflector. Even though previous work exists on the separate problems of offset reflectors [5] and offset feeds [6], the results in the literature tend to be sketchy and none deal with the complexities that arise when both problems are combined.

Chapter II contains a review of design fundamentals for parabolic reflectors antennas. Included are discussions of geometry and reflection results, offset feeds, offset reflectors, focusing and pointing errors, beam deviation factor, aperture blocking, choice of F/D, and need for an offset reflector configuration.

In Chapter III offset reflector and offset feed geometries are analyzed. Among the results obtained are aperture distributions for conical beams aimed off the paraboloid vertex, far field patterns including diffraction effects produced by such aperture distributions, and the use of a Petzval feed surface for minimizing coma effects for an offset feed cluster.

CHAPTER II

PARABOLIC REFLECTORS

A parabolic surface has the property of reflecting all rays issuing from its focal point back parallel to its axis and such that path lengths from focal point back to the focal plane are equal (See Fig. 2-1). Since all rays are consequently in phase over the focal plane aperture, the reflector is well suited from a small feed standpoint for constructing fairly high gain, large aperture antennas. Good treatments of the basics of parabolic reflector antennas are given by Sletten [5] and Wolff [9].

Throughout this report, the terms "offset feed" and "offset reflector" will occur repeatedly. By an offset feed is meant a reflector structure where the feed point or points are located away from the focal point of the paraboloid. An offset reflector geometry is one where the feed beam maximums are aimed at some point on the reflector other than the paraboloid vertex. As will be seen, this is done when bad aperture blockage cannot be avoided, and it generally results in using a portion of the parabolic surface which does not have the circular symmetry that is characteristic of vertex aiming.

Much of the work to follow depends on geometric relations for rays reflected from a paraboloid; so some of the results will be presented for later reference.

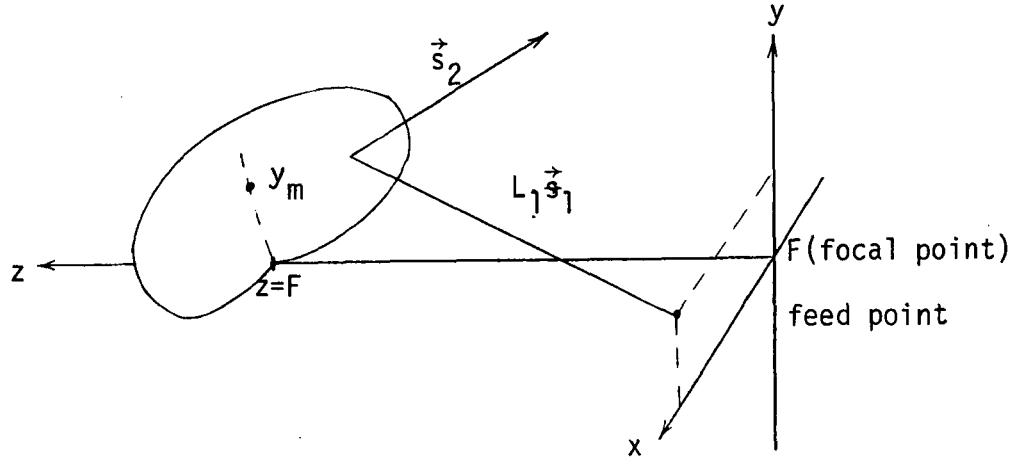


Figure 2-1 Paraboloid

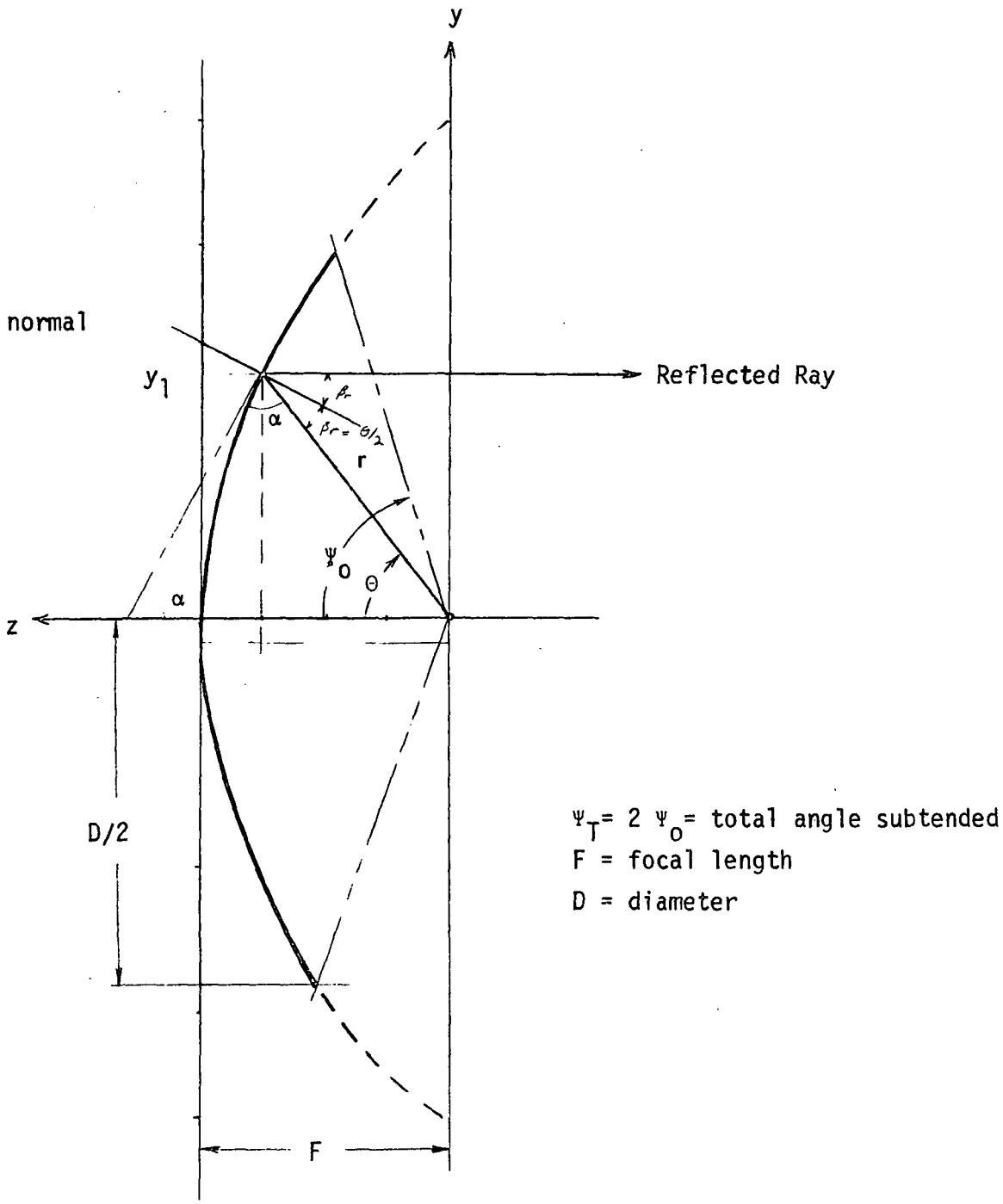


Figure 2-2 Paraboloid Cross Section

In Figure 2-1 is shown a portion of a paraboloid with vertex (V) and focal point (F). The aperture plane is located at z=0 and passes through the focal point for the surface. For a parabola with focal length, F, all ray paths from the focal point back to the focal plane equal 2F. The z axis passes through the vertex and focal point, and for the focal point at the origin, the xy plane is the focal plane. Referring to Figure 2-2 where a parabola cross section is plotted, some of the more important geometric relations are summarized as follows:

1. The equation for a paraboloid in rectangular coordinates is

$$y^2 + x^2 = 4F\ (F-z). \tag{2-1}$$

2. In spherical coordinates the radial distance is

$$r = \frac{2F}{1 + \cos\theta} = F \sec^2\left(\frac{\theta}{2}\right) = p \tag{2-2}$$

where (r, θ, ϕ) are the spherical coordinates measured from the focal point.

3. The slope of the reflector surface at a height y is

$$\frac{dy}{dz} = \frac{2F}{y} = \tan(\alpha) = \frac{1 + \cos\theta}{\sin\theta} \tag{2-3}$$

Note that ARF is an isosoles triangle.

4. $a = d_1$ (2-4)

5. The ray length $r = F + d_1 = \frac{2F}{1 + \cos\theta}$ (2-5)

6. At the point of reflection for a ray coming from the focal point, the angles of incidence and reflection are

$$\beta_i = \beta_r = \frac{\theta}{2} \tag{2-6}$$

7. For a symmetric paraboloid of diameter (D) the total angle subtended can be shown to be (2-7)

$$\Psi_T = 2\Psi_o = 2 \tan^{-1}\left[\frac{1/2\ F/D}{(F/D)^2 - 1/16}\right] = 4 \tan^{-1}\left[\frac{.25}{F/D}\right]$$

8. The focal length to diameter ratio (F/D) for a symmetric paraboloid of resolution can be shown to be

$$F/D = \frac{1 + \cos \Psi_o}{4 \sin \Psi_o} = \frac{0.25}{\tan (0.25\ \Psi_T)} \tag{2-8}$$

where Ψ_T is the total angle subtended and Ψ_o is the half cone angle subtended by the paraboloid.

The relationship between F/D and the angle subtended is plotted in Figure 2-3. In Table 2-1 are listed some of the commonly used F/D's along with the corresponding cone angles. Note that the plot of F/D found in many handbooks [7] is in error and Figure 2-3 is a corrected result.

A. Offset Feeds

Referring back to Figure 2-1, consider an incident ray of length L_1 from the source point (0, y_o, z_o) to the point (x, y, z) on the paraboloid. For simplicity it is assumed that the coordinate axes have been rotated so that the source point lies in the x = 0 plane. The equation for the parabola

$$r = \frac{2F}{1 + \cos \theta} = \frac{F}{\cos^2 \frac{\theta}{2}} \tag{2-9}$$

can be manipulated to give the surface

$$L = F - p \cos^2 \frac{\theta}{2} .$$

From L the unit normal to the reflector is just (2-10)

$$\vec{n} = \frac{\nabla L}{|\nabla L|} = \frac{\nabla (F - p \cos^2 \theta/2)}{|\nabla ( )|} = - \vec{a}_R \cos \frac{\theta}{2} + \vec{a}_\theta \sin \frac{\theta}{2}$$

The incident ray ($L_1 \vec{s}_1$) is of length

$$L_1 = \sqrt{x^2 + (y-y_o)^2 + (z-z_o)^2} \tag{2-11}$$

and has direction cosines

$$\left(\frac{x}{L_1}, \frac{y-y_o}{L_1}, \frac{z-z_o}{L_1} \right) = (\ell_1, m_1, n_1) \tag{2-12}$$

Ψ_T	Ψ_o	$F/D = \frac{1}{4 * \tan(\frac{\Psi_T}{4})}$
40°	20°	1.4178
56°	28°	1.000
60°	30°	.933
71°	35.66°	.800
80°	40°	.6868
90°	45°	.6036
100°	50°	.5362
106°		.5000
120°	60°	.4330
140°	70°	.3570
160°	80°	.2979

$$\tan\left(\frac{\Psi_T}{4}\right) = \frac{1}{(4)\ (F/D)}$$

$$\Psi_T = 4\ \tan^{-1}\left[\ \frac{.25}{F/D}\ \right]$$

TABLE 2-1 Commonly used F/D's with Corresponding Cone Angles

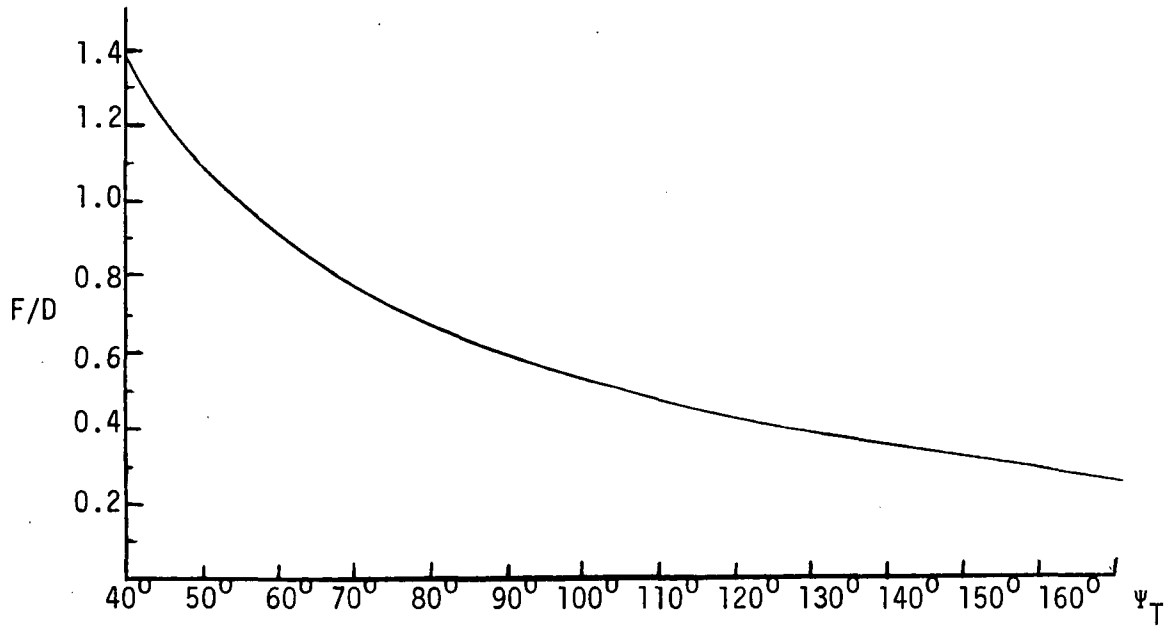


Figure 2-3 F/D Versus Angle Subtended

δ_y	$\Psi_o = \Psi_T/2$	F/D	θ_s	γ(min.)	γave.	Deviation
0.2F	28^o	1.0	11.31^o	8.9^o	10.1^o	$\pm 1.2^o$
0.1F	28^o	1.0	5.73^o	4.5^o	5.12^o	$\pm .61^o$
0.2F	50.92^o	.525	11.31^o	5.21^o	8.26^o	$\pm 3.05^o$ (37%error)
0.1F	50.92^o	.525	5.73^o	2.767^o	4.238^o	$\pm 1.471^o$(35%error)
0.1F	34.708^o	0.8	5.73^o	4.0725^o	4.89^o	$\pm .819^o$

Table 2-2 Focusing Errors

The angle of reflection referenced to the surface normal equals the angle of reflection and a reflected ray $\vec{s}_2$ is obtained with direction cosines

$$\cos(\alpha_x) = \ell_2 = \frac{2Fz_o + yy_o}{4F^2 - 2Fz} \quad \frac{x}{L_1} \tag{2-13}$$

$$\cos(\alpha_y) = m_2 = \frac{2Fz_o + yy_o}{4F^2 - 2Fz} \quad \frac{y}{L_1} \tag{2-14}$$

$$\cos(\alpha_z) = m_2 = \frac{2Fz_o + yy_o}{2F - z} \quad \frac{1}{L_1} + \frac{z - z_o - 2F}{L_1} \tag{2-15}$$

[Snell's law in vector form gives $\vec{s}_r = \vec{s}_{inc} - 2(\vec{n} \cdot \vec{s}_{inc})\vec{n}$]

B. Focusing Errors

In this section some of the aberrations produced by offset feed elements with a parabolic reflector are treated [8]. The nonparallel alignment of rays emanating from the same off-axis source point on a planar feed surface can lead to pattern degradations. As an illustration, angles of rays from several points on a focal plane (z=0) surface are sketched in Figure 2-4, and ranges of ray angles produced by feed points typical of those needed for a synchronous orbit antenna are calculated in Table 1-2. Roughly for a ± six degree ground range coverage, a focal deviation of $\delta_y \simeq 0.2F$ is required. For an F/D of 1.0, rays from $\delta_y = 0.2F$ were reflected ± 1.2 degrees from their average over the upper half of the paraboloid. In practice such variations could cause the beam to diverge, resulting in lower gain and possibly, what might be worse, higher sidelobe levels. For multiple feed points one might consider the possibility of obtaining parallel aperture illumination in a single direction by summing ray contributions in that direction from all the multiple feed points. However, switchable and optional regional coverage necessitates using multiple independent beams, thus, source elements would be required to simultaneously have many different values. Such an approach would only be worthwhile for a single feed aperture distribution which was to produce a single, unchangeable shaped-beam pattern.

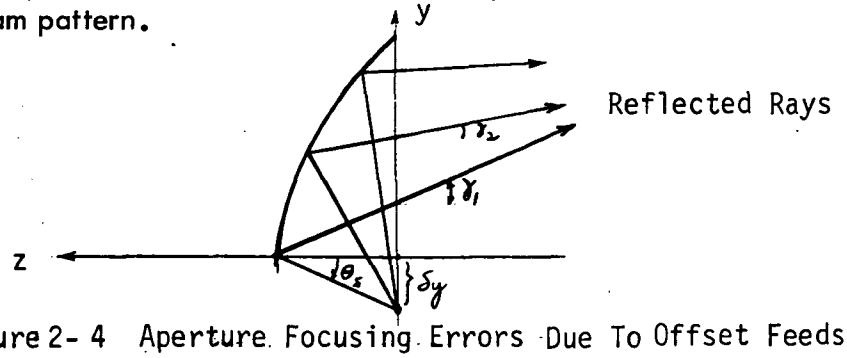


Figure 2-4 Aperture Focusing Errors Due To Offset Feeds

C Offaxis Feeds

Before the pattern aberrations produced by offset feeds can be examined, it is necessary to formulate the far field patterns for parabolic reflectors with such feed offsets. Consider the geometry shown in Figure 2-5.

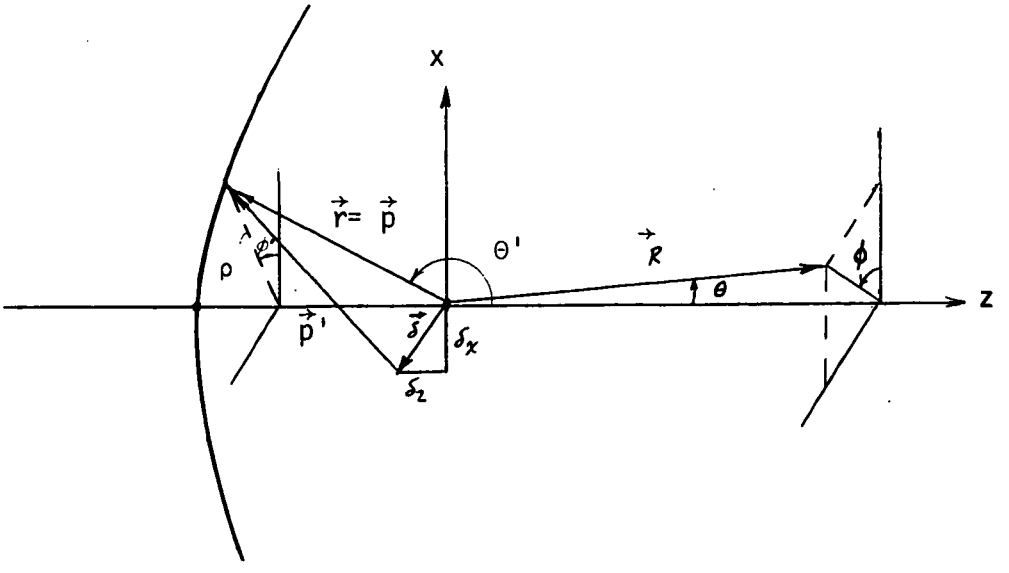


Figure 2-5 Offset Feed

1. Focal Feed

For a feed element located at the focal point, the far field pattern takes the form

$$E(\theta,\phi)=\int_0^{2\pi}\int_0^a f(\rho,\phi')e^{ik(\bar{p}-\bar{p}\cdot\bar{a}_R)}\,\rho\,d\rho\,d\phi' \qquad (2\text{-}16)$$

which can be obtained from Eqs. 2-10 to 2-15 using slightly altered coordinates to fit the geometry used by Ruze [6]. In the above

$$p=\frac{2f}{1-\cos\theta'} \qquad (2\text{-}17)$$

$$\vec{p}=p[\bar{a}_x\sin\theta'\cos\phi'+\vec{a}_y\sin\theta'\sin\phi+\vec{a}_z\cos\theta'] = \text{incident ray from the focal point.} \qquad (2\text{-}18)$$

$$\vec{a}_R=\vec{a}_x\sin\theta\cos\phi+\vec{a}_y\sin\theta\sin\phi+\vec{a}_z\cos\theta = \text{unit vector in } (\theta,\phi) \text{ direction.}$$

2. Offset Feed

When the feed is displaced by the vector, $\delta = \delta_x \bar{a}_x + \delta \bar{a}_z$, the p equation is modified to

$$\vec{p}' = \vec{p} + \vec{\delta} = \vec{p} + \delta_x \vec{a}_x + \delta_y \vec{a}_y \tag{2-20}$$

$$p' = p\{1+\frac{2\delta_x}{p}\cos\phi'\sin\theta' + \frac{2\delta_z\cos\theta'}{p} + \frac{\delta^2}{p^2}\}^{1/2} . \tag{2-21}$$

The far field pattern becomes

$$E = \iint f(\rho,\phi') e^{jk(p'-\vec{p}'\cdot\bar{a}_R)} \rho d\rho\, d\phi' \tag{2-22}$$

For feed displacements that are small compared to the focal length

$$\delta << f < p ,$$

it can be shown by using the first order terms in the binomial expansion for p' that

(2-23)

$$p' - \vec{p}'\cdot\bar{a}_R \simeq 2F - \delta_x\cos\phi\sin\theta - \delta_z\cos\theta - p\sin\theta'\sin\theta\cos(\phi-\phi') + \delta_x\cos\phi'\sin\theta' + \delta_z\cos\theta' + \frac{\delta_x^2}{2p} + p\cos\theta'(1-\cos\theta) - \frac{\delta_x^2}{2p}\cos^2\phi'\sin^2\theta'$$

$\rho = p\sin\theta'$

These phase contributions have the following effects:

1. The first three terms are independent of the integration.
2. The fourth term is the usual factor associated with a center fed aperture.
3. The fifth term corresponds to beam shift.
4. The next three terms produce field curvature which can cause aberration.
5. Last term is associated with astigmatism.

3. Petzval Surface

Ideally it would be desirable to eliminate all aberration producing terms, and part of the last term. The field curvature terms in Eq.2-23 can be eliminated by setting

$$\delta_z = \frac{\delta_x^2}{2F} \tag{2-24}$$

This feed focus is in a sense a "best" location for offset feed elements in that many of the aberration producing terms in (2-23) are cancelled. The generated feed surface is a paraboloid with a focal length of 0.5F which is located tangent to the focal plane [6, 10]. Thus, assuming astigmatism is negligible and lumping the constant phase shift terms into the constants associated with $f(\rho, \phi')$, the far field pattern can be written

$$F(\theta,\phi) = \int_0^{2\pi}\int_0^{a} f(\rho,\phi')\, e^{-jk[\rho \sin\theta \cos(\phi'-\phi) - \delta_x \sin\theta' \cos\phi']}\, \rho\, d\rho\, d\phi' \tag{2-25}$$

As simple as (2-25) may seem, it masks the most difficult problem of evaluating the field pattern for offset feed elements, namely, that of determing the aperture distribution, $f(\rho,\phi')$. This topic is covered in Chapter III.

The first term in the exponent of the above integrand is the phase delay normally obtained for the far field of a focus fed circular aperture. The second

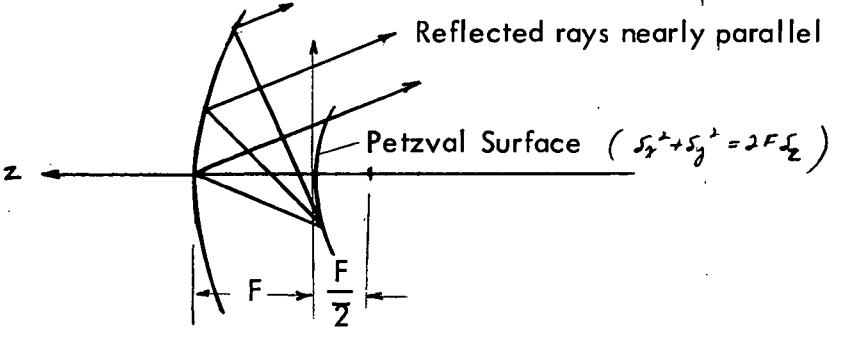


Fig.2 -6 Petzval Feed Surface

term in the exponential gives rise to the beam degradation caused by the offset feed. Using the geometric series expansion

$$\sin\theta' = \frac{r/F}{1+(r/2F)^2} = \frac{r}{F}\left[1-\left(\frac{r}{2F}\right)^2 + \left(\frac{r}{2F}\right)^4 - - -\right], \tag{2-26}$$

the offset feed phase departure term can be expanded as

$$k\delta_x \sin\theta'\cos\phi' = \frac{2\pi}{\lambda}\delta_x \cos\phi' \frac{r}{F}\left[1-\left(\frac{r}{2F}\right)^2 + \left(\frac{r}{2F}\right)^4 \ldots\right] \tag{2-27}$$

Ruze [6] defines the variable

$$U_S = \frac{\delta_x}{F} = \tan\theta_S \tag{2-28}$$

θ_S = feed squint angle

as a measure of feed squint. If only the first term in(2-27) is significant, the resulting beam would be scanned by the feed squint or offset angle, θ_S. The second term when significant produces "primary coma" which tends to shift the beam in the opposite direction (away from the geometric squint angle, θ_S) and distort its pattern shape. The remaining terms constitute higher order coma. Obviously the relative importance of the coma terms is dependent on the F/D chosen and are negligible for high F/D as characteristic of many optical systems.

As a measure of beam pointing effects the quantity, beam deviation factor (BDF) is defined as

$$\text{BDF} = \text{beam deviation factor} = \frac{\text{actual beam offset angle}}{\text{feed offset angle}} = \frac{\theta_m}{\theta_S} \tag{2-29}$$

The variations of BDF as a function of paraboloid F/D have been calculated by Ruze [7] and are plotted in Figure (2-7). The aperture distributions considered all assume circular symmetry and range from uniform illumination to the commonly used tapered illumination, $f(p) = 0.3 + 0.7\ (1-p^2)$, with a -10db edge pedestal and resulting far field pattern having -29 db sidelobes. From Figure (2-7) it is seen that for F/D's greater than 0.8 that the beam deviation factor is between 0.95 and 1.0 indicating that the feed offset angle and the beam angle are approximately the same. For this reason $F/D \geq 0.8$ is recommended for shaped-beam designs.

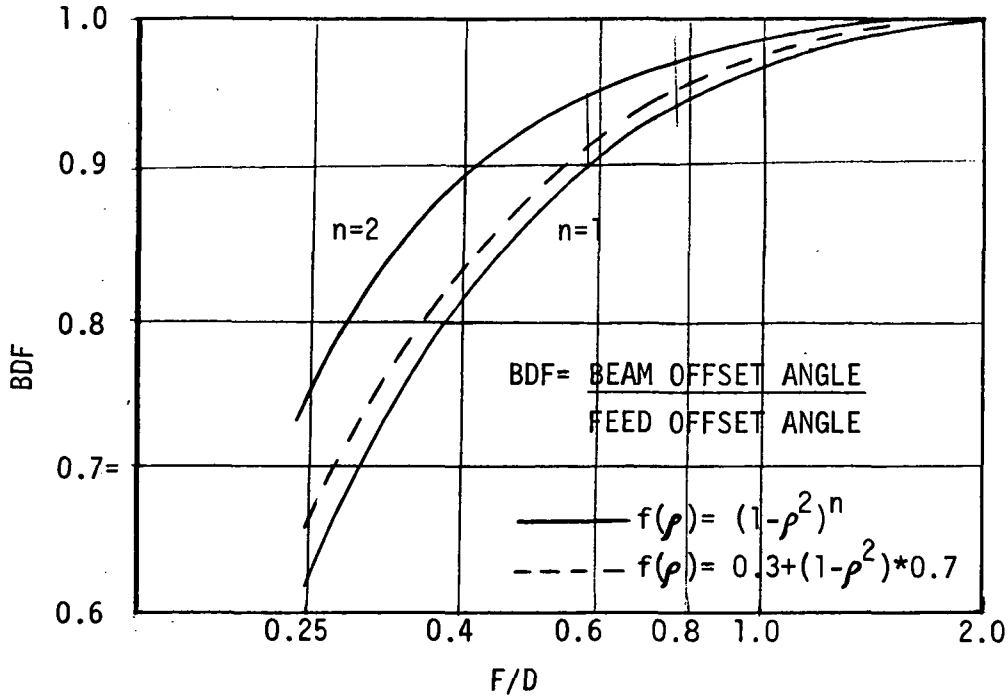


Figure 2-7 Beam Deviation Factor [6]

D. Choice of Reflector

For a shaped-beam antenna designed for fifty state coverage and employing a single reflector, a centered reflector dish with an F/D of at least 0.8 is recommended. The choice of F/D is fairly important because of its influence on beam quality and pointing. To better appreciate the tradeoffs involved in reflector selection, several examples will be given. Taking $\pm 6.2^{o}$ from boresight as an estimate of the beam squints required, we see from Table 2-2 that for an F/ of 0.525 and an offset position of $\delta_y = 0.1F$, the aperture illuminating rays vary from 5.73^{o} to 2.76^{o} over the aperture plane [feed squint $=\theta_s = 5.73^{o} = \tan^{-1}\left(\frac{\delta_y}{F}\right)$]

which is a range of $\pm 1.47^{o}$ from the average. Now if we are dealing with half-degree beams, this variation can seriously affect the pattern. As a general rule, the effective beam pointing direction for offset feeds is less than the feed squint angle (θ_s), and hence δ_y would have to be larger than the 0.1F originally allowed. The beam deviation factor (BDF) is useful for estimating feed location corrections, however, it does not show the magnitude of the beam degradation. For this example with F/D = 0.525, the BDF from Fig. 2-7 is about 0.88 for the -10 dB edge tapered illumination, and the

cone angle subtended by the dish is 101.8^0. The aperture blockage is thus on the order of

$$\left[\frac{2\theta_m/BDF}{\Psi_T}\right]^2 \times 100\% = \left[\frac{12^0/.88}{101.8^0}\right]^2 \times 100\% = 1.8\%$$

which is not terribly significant. The beam degradation for limiting squints, on the other hand, could be very unsatisfactory for half-degree beams. Resorting to larger F/D's remedies the beam deviation problem, but worsens the blockage for reflectors which are not offset. For instance, an F/D of 0.8 gives a BDF of 0.95 and the aperture blockage increases to

$$100 \left[\frac{12^0/BDF}{\Psi_T}\right]^2 = \left[\frac{12^0/.95}{71^0}\right]^2 \times 100\% = 3.2\%$$

which is still reasonably small. Probably a good compromise lies somewhere in the F/D range from 0.7 to 1.0. For centered reflectors the power loss due to blockage is usually greater than the above percentages since the feed beam maximums are blocked, however, it is still very tolerable for this design. The relative sidelobe level increase due to blockage will normally be of more consequence. If desired, aperture blockage can be entirely eliminated by using an offset reflector. On balance, the drawbacks to an offset reflector are that of designing a deployable structure for satellite application and the polarization characteristics.

At this point it might be appropriate to summarize some of the characteristics of the recommended system:

1. Off focus feed elements located on a Petzval surface would be used for pointing.

2. The F/D should be greater than, or equal to, 0.8 to give a beam deviation factor approximately equal to one.

3. A centered reflector geometry is suggested.

4. Adjacent feed beams would have -6dB points in common in first order or trial tests. Depending on results this could be altered.

5. A compromise choice of effective beamwidths for feed elements has to be made. Low sidelobe levels (less than 25 dB) and narrow beamwidths would enable sharp coverage up to the edges of complicated borders (Eastern Time Zone is the most difficult) with minimum

spillover into foreign neighbors (70% of the population of Canada is within 100 miles of the U. S. Border). Narrow beams require more elements for area coverage, however, feed complexity and mutual coupling problems are best satisfied by minimizing the number of elements. Fairly good coverage could probably be achieved with forty or fifty elements with about half-degree beamwidths. If time zone borders could be more loosely defined (smoothed) and if foreign neighbors desired inclusion in the coverage services, a simple and cheap design could be realized with about six broadbeam feeds (one for each time zone plus one for Hawaii and one for Alaska). This study is oriented toward the requirements of the more sophisticated systems.

CHAPTER III

MULTIBEAM APERTURE DISTRIBUTIONS AND PATTERNS

This chapter is concerned with the determination of aperture distributions and far field patterns for offset reflector and offset feed combinations using a parabolic reflector. For analysis purposes, the feed elements are assumed to have conically symmetric patterns such as might be realized using waveguide-fed polyrod antennas [9,11] or dielectrically loaded circular horns [13]. These are representative of the elements that can be used to obtain wide bandwidths, circular polarization, and small sizes. One pattern that is used in many of the sample calculations in this chapter is that of a polyrod [11] having a -10dB beamwidth of 64^{o}. In all the discussions that follow only a single feed element is considered at a time with the understanding that the net pattern for multiple feeds can be found by superposition of single feed patterns. Qualitatively the superposition will be marked by (1) vector tilts to the effective aperture planes associated with different feed positions as used to point the beams (this phenomenon might also be mathematically described by linear phase tapers across a reference aperture which produce the same beam steering), (2) shifts of phase center locations for the different apertures, and (3) skewing and edge truncation of amplitude distributions dependent on feed position.

A. Conical Beam Feed to an Offset Reflector

In this analysis an offset reflector geometry is used for the purpose of eliminating aperture blockage. A feed element with a conical pattern is located at the focal point of the paraboloid. The geometry of the offset reflector is detailed in Figure 3-1a with the plane PP' which is normal to the conical beam shown in Figure 3-1b.

The focal point is at the origin and the feed beam is centered at the point $(0, y_m, z_m)$ on the reflector. The length of the ray to the beam center is

$$P_m = \rho_m = \frac{2F}{1 + \cos \Psi_.} = r_m \qquad (3\text{-}1)$$

Given

γ = half cone angle

β = azimuth angle of the ray within the beam (measured from vertical)

L = ray length from feed point to the reflector,

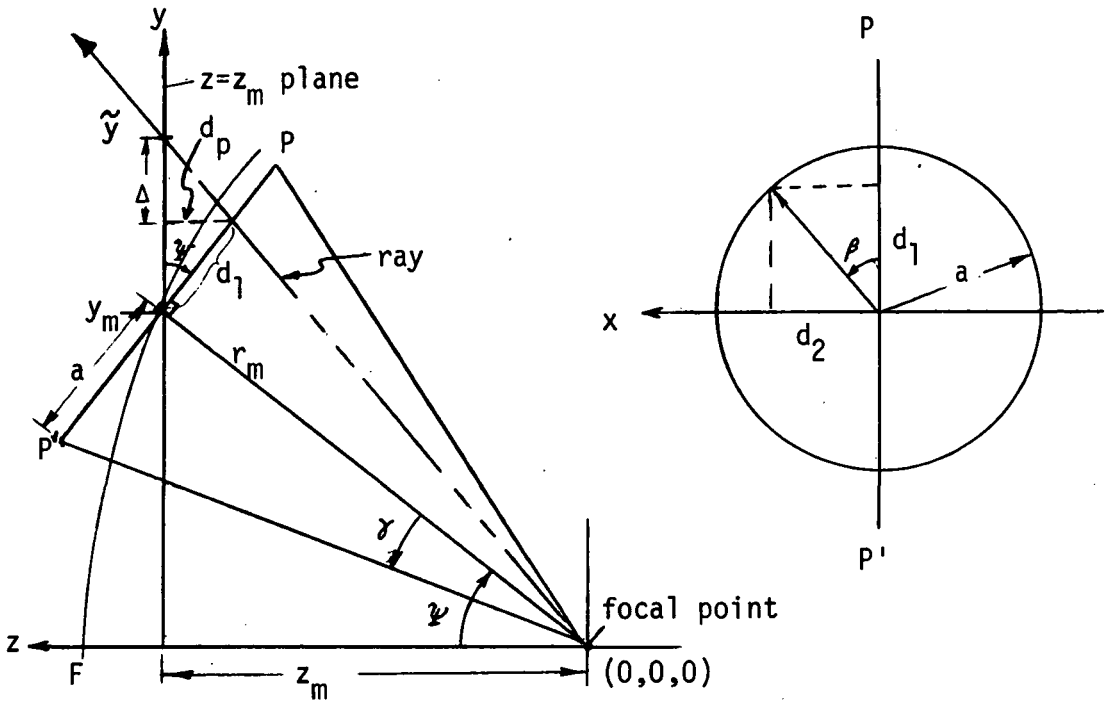


Figure 3-1 Offset Reflector with Conical Feed Pattern

the general footprint of a conical beam projected back to the focal plane is obtained. The analysis depends on finding the direction cosines for rays on the surface of the cone shown in Figure 3-1 and using them to determine the points of intersection of the ray with the paraboloid. The ray direction cosines are found by first locating the points of intersection, $(\tilde{x}, \tilde{y}, \tilde{z})$, of the rays with the $z = z_m$ plane. From Figure 3-1

$$a = P_m \tan\gamma \tag{3-2}$$

$$d_1 = a \cos(\beta) = P_m \cos(\beta) \tan(\gamma) \tag{3-3}$$

$$d_2 = a \sin(\beta) = P_m \sin(\beta) \tan(\gamma) \tag{3-4}$$

$$d_p = d_1 \sin\Psi = P_m \cos(\beta) \tan(\gamma) \sin(\Psi) \tag{3-5}$$

also

$$\tilde{z} = z_m = P_m \cos\Psi \tag{3-6}$$

$$\tilde{y} = y_m + d_1 \cos\Psi + \Delta \tag{3-7}$$

The effects of the ray flaring are illustrated in Figure 3-2

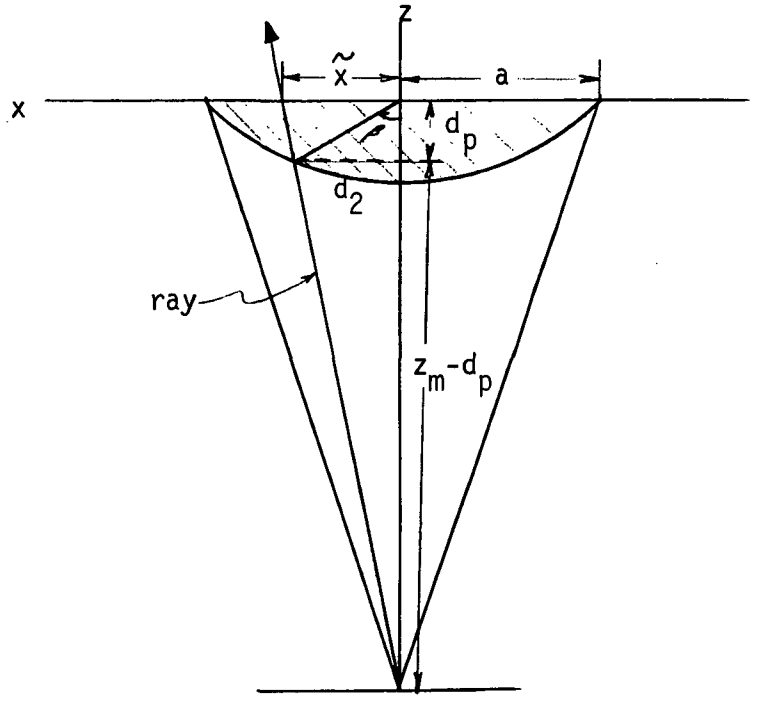


Figure 3-2 Ray Flaring in a Conical Beam

For similar triangles

$$\frac{\Delta}{d_p} = \frac{y_m + d_1 \cos\Psi + \Delta}{z_m} \tag{3-8}$$

yielding

$$\Delta = \frac{y_m + d_1 \cos\Psi}{-1 + z_m/d_p} \tag{3-9}$$

From the geometry shown in Figure 3-2

$$\frac{\tilde{x}}{z_m} = \frac{d_2}{z_m - d_p} \tag{3-10}$$

giving $$\tilde{x} = P_m \frac{\tan(\gamma)\ \sin(\beta)}{1 - \tan(\gamma)\ \tan(\Psi)\ \cos(\beta)} \tag{3-11}$$

Substituting (3-3) and (3-9) in (3-7) gives

$$\tilde{y} = \frac{\rho_m[\sin(\Psi) + \cos(\beta)\ \tan\gamma\ \cos\Psi\ ]}{1-\tan\gamma\ \tan\Psi\ \cos\beta} \tag{3-12}$$

Defining

$$LP = \sqrt{\tilde{x}^2 + \tilde{y}^2 + \tilde{z}^2} \tag{3-13}$$

$$= P_m\sqrt{\cos^2\Psi + (\frac{\sin\Psi+\tan\gamma\ \cos\Psi\ \cos\beta}{1-\tan\gamma\ \tan\Psi\cos\beta})^2 + (\frac{\tan\gamma\ \sin\beta}{1-\tan\gamma\ \tan\Psi\cos\beta})^2}$$

the direction cosines for the ray can be expressed

$$\cos(\alpha_z) = \cos\Theta = \tilde{z}/LP \tag{3-14}$$

$$\cos(\alpha_x) = \sin\Theta\ \cos\phi = \tilde{x}/LP \tag{3-15}$$

$$\cos(\alpha_y) = \sin\Theta\ \sin\phi = \tilde{y}/LP \tag{3-16}$$

From the direction cosines

$$\Theta = \cos^{-1}(\tilde{z}/LP) \tag{3-17}$$

$$\phi = \cos^{-1}(\tilde{x}/LP*\sin\Theta) = \sin^{-1}(\tilde{y}/LP*\sin\Theta\ ) \tag{3-18}$$

$$= \tan^{-1}(\tilde{y}/\tilde{x}).$$

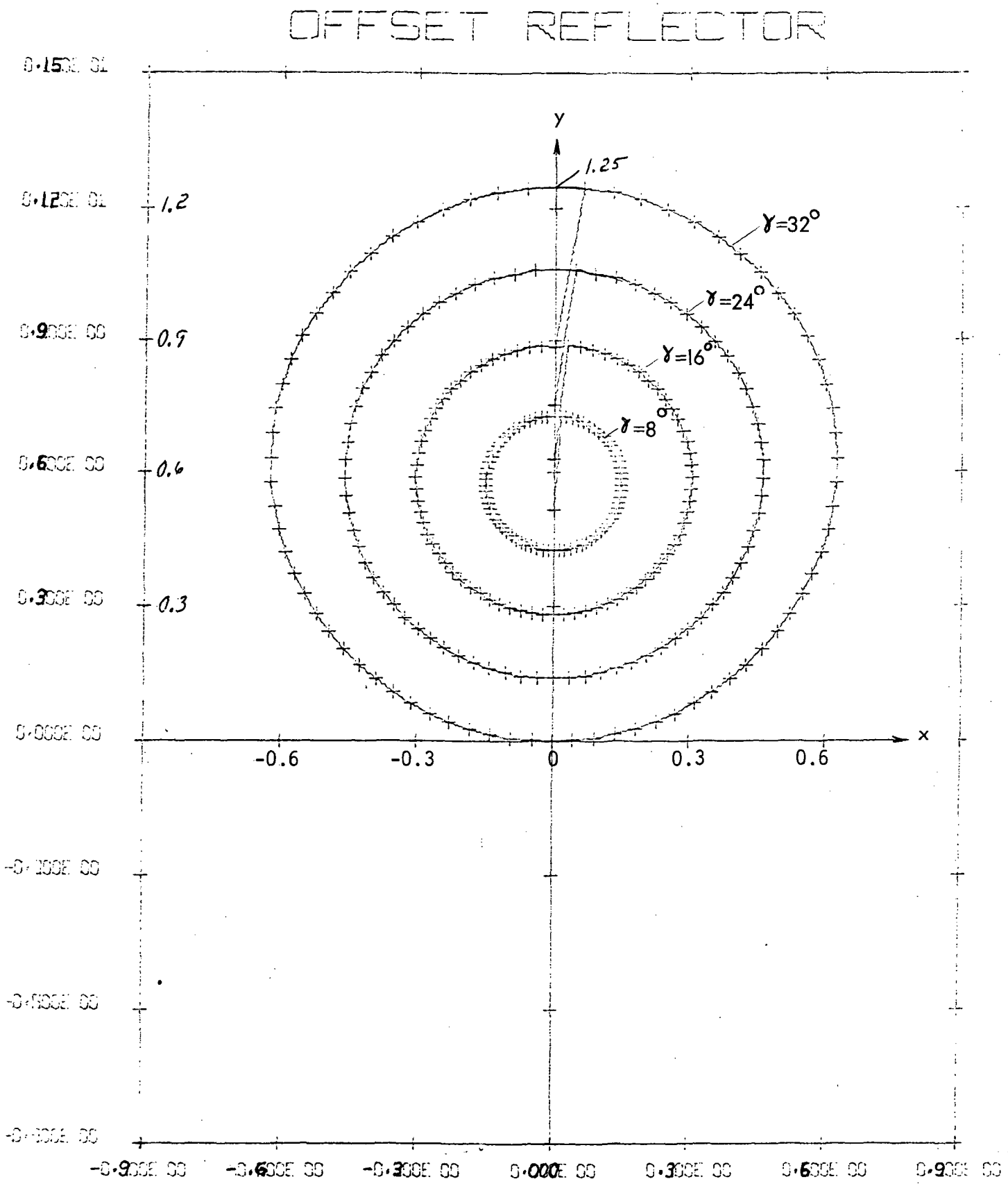


Fig. 3-3 Amplitude Contours in the Aperture Plane of an Offset Reflector

Once the ray polar angles (Θ, ϕ) are known for a beam position (Ψ, γ, β), the point of intersection of the ray with the paraboloid can be found using

$$x = \rho \sin\Theta \cos\phi = 2F \tan\frac{\Theta}{2} \cos\phi \qquad (3\text{-}19)$$

$$y = \rho \sin\Theta \sin\phi = 2F \tan\frac{\Theta}{2} \sin\phi \quad . \qquad (3\text{-}20)$$

The preceding equations have been computerized to produce the constant amplitude contours for half-cone angles of $\gamma = 8^0, 16^0, .., 32^0$, and these are plotted in Figure 3-3. As can be seen, these contours, which show how the aperture distribution is affected by the offset reflector, are nearly a family of circles with skewed center locations whose relative shift depends on position in the conical beam. The problem of determining far field patterns then becomes that of finding a mathematical model which best describes the aperture illumination. For a centered reflector(no offset)the aperture distribution has circular symmetry and is independent of ϕ'.

So that the ramifications of the integral used to determine far-field patterns from a knowledge of fields over an aperture surface of an antenna can be fully understood, a formulation of the pattern integral will be given. Let a planar aperture be located in the xy-plane such as shown in Figure 3-4.

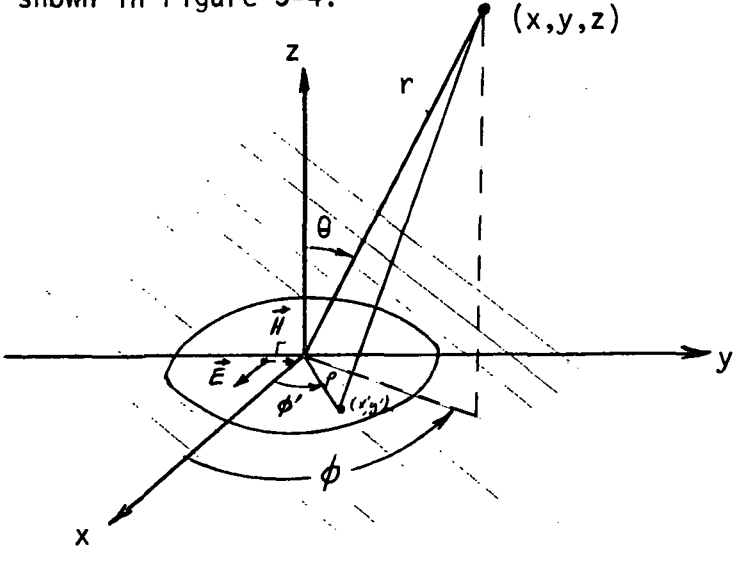


Fig. 3-4 Planar Radiating Aperture

For aperture fields which vanish outside the aperture and equal

$$\vec{E}_{ap} = E_x \vec{a}_x \tag{3-21}$$

$$\vec{H}_{ap} = H_y \vec{a}_y = \frac{1}{\eta} E_x \vec{a}_y \tag{3-22}$$

within the aperture, it can be shown that the far-field patterns are [9, 12]

(3-23)

$$E_\Theta = \frac{jk\eta}{4\pi} \cos\phi \, (1+\cos\Theta) \frac{e^{-jkr}}{r} \underbrace{\iint_{A_p} H_y(r',y') e^{+jk(xx'+yy')/r} dx'dy'}_{\triangleq F_H}$$

$$E_\phi(\Theta,\phi) = -\frac{jk\eta}{4\pi} \frac{e^{-jkr}}{r} \sin\phi \, (1+\cos\Theta) \, F_H(\Theta,\phi) \tag{3-24}$$

The above equations introduce the diffraction effects which are not included in a ray optics approach to the solution. In many applications, including this one, the aperture field integral

$$F_H(\theta,\phi) = \iint_{A_p} H_y(x',y') e^{+jk(xx'+yy')/r} dx'dy' \tag{3-25}$$

is used as the approximate pattern characteristic. For pattern behavior near the z axis (broadside) the approximation is excellent, while for θ large, the additional attenuation caused by the $\begin{pmatrix}\sin\phi \\ \cos\phi\end{pmatrix}$ $(1 + \cos\theta)$ multipliers should be included.

Circular Aperture Patterns

The following are some of the important illuminations and their derived far-field patterns for circular apertures:

(1) Uniform (f = constant)

$$F_{H_1} = \int_0^a \int_0^{2\pi} e^{+jkp \sin\theta \cos(\phi-\phi')} \, p\,dp\,d\phi' \tag{3-26}$$

$$= 2\pi \int_0^a J_0(kp \sin\theta) \, p \, dp \tag{3-27}$$

$$F_{H_1} = 2\pi a^2 \frac{J_1(k_a \sin\theta)}{ka \sin\theta} \tag{3-28}$$

$$P = \sqrt{x^2 + y^2} \tag{3-29}$$

which can be expressed in terms of the variable

$$X = ka \sin\theta \tag{3-30}$$

as

$$F_u = \pi a^2 \frac{J_1(x)}{(\frac{x}{2})} \quad .$$

(2) Gaussian $(e^{-\alpha p^2})$

$$f(p, \phi') = e^{-\alpha p^2} \tag{3-31}$$

$$F_{H_2}(\theta,\phi) = \int_0^a e^{-\alpha p^2} \, 2\pi J_0(kp \sin\theta) \, p \, dp \; . \tag{3-32}$$

For an infinite aperture the above can be evaluated in closed form to give

$$F_\infty = \frac{\pi}{\alpha} \, e^{-k^2 \sin^2\theta/4\alpha} \; . \tag{3-33}$$

(3) Tapered $[\,1-(\frac{p}{a})^2]^{\,n}$

This aperture distribution is typical of that obtained from a cosine^n feed.

$$f(p,\phi') = [1-(\tfrac{p}{a})^2]^{\,n} \tag{3-34}$$

Integrating the above yields

$$F_{H_3}(\theta,\phi) = \pi a^2 \; \frac{2^{m+1} m!\, J_{m+1}(ka\sin\theta)}{(ka\sin\theta)^{m+1}} \tag{3-35}$$

$$= \pi a^2 m! \; \frac{J_{m+1}(X)}{(\frac{X}{2})^{m+1}} \tag{3-36}$$

(4) Tapered on a pedestal $\{\, A + B[1-(\frac{p}{a})^2] \,\}$

This distribution is widely used in many practical designs. It is basically a compromise choice to keep aperture efficiency as high as possible and maintain sufficiently low sidelobe levels. For a -10 dB aperture edge level, the aperture distribution is

$$f(p,\phi') \cong 0.31 + 0.69\,[1-(\tfrac{p}{a})^2]\,. \tag{3-37}$$

The resultant field pattern is a linear combination of uniform and tapered (m=1) giving

$$F_{H_4} = \pi a^2 [0.31\,\frac{J_1(x)}{x} + 0.69\,\frac{J_2(x)}{(\frac{x}{2})^2}]\,. \tag{3-38}$$

The pattern in Eq. 3-38 is used as a standard for comparison of subsequent offset patterns. With the aid of a computer, the pattern for this illumination is plotted in Figure 3-5. From Figure 3-5 it is seen that sidelobes are down 29dB and that the half-power beamwidth is 64.6 λ/D degrees. Choosing a feed F/D = 0.8, a dish diameter of 2.58 meters (8.48 feet) is needed for a -3dB beamwidth of a half degree, while a diameter of 3.6 meters (11.8 feet) produces a -6dB beamwidth of a half degree.

(5) Offset Reflector

When a conically symmetric beam is aimed off axis($y_m \neq 0$), the aperture illumination derived in previous sections and plotted in Figure 3-3 is obtained. Because the constant amplitude contours were approximately circular, the outermost circle ($\gamma=32^o$) was chosen to define the aperture boundary (radius = 0.625F and F/D=0.8) and the y axis of the coordinates was shifted to make the new origin coincide with the center of the bounding circle. The location of the feed beam center was shifted from y=0.5735F to $y'=-y_m=-0.0515F$.

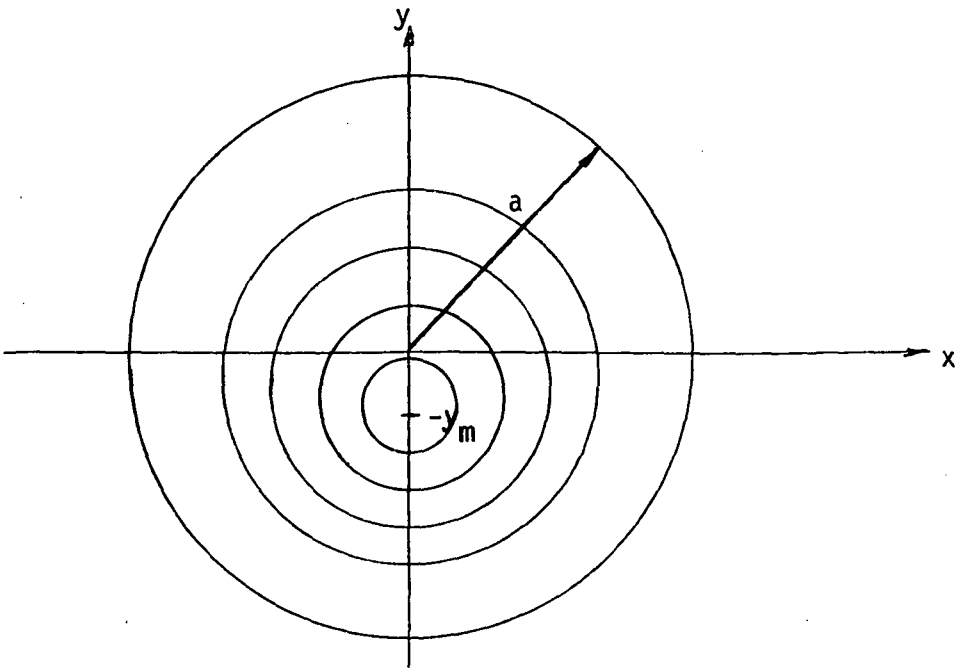


Figure 3-6 Shifted Coordinate Contours

Since the $\gamma=32^o$ contour defined the -10dB locus of the polyrod, a skewed tapered aperture distribution on a pedestal resulted. Several different tries were made to find a good mathematical fit for this aperture distribution with the result being

$$f(p,\phi) = f(x,y) = A+B\left[1-\frac{x^2+y^2}{a^2}\right]\left[1-\left(\frac{y+y_m}{y_{m+a}}\right)^2\right] \qquad (3\text{-}39)$$

Note that the shifted coordinate system shown in Figure 3-6 has been used. Since the azimuthal symmetry previously encountered has been lost, the far field pattern can be best expressed as

$$F_H=\int_{-a}^{a}\int_{-\sqrt{a^2-y^2}}^{\sqrt{a^2-y^2}}\left[A+B\left(1-\frac{p^2}{a^2}\right)\left(1-\left(\left(\frac{y+y_m}{y_m+a}\right)\right)^2\right)\right]e^{+jk(xx'+yy')/r}\,dx\,dy \qquad (3\text{-}40)$$

where the integration is now in terms of x and y. Regrouping (Eq. 3-40) gives

$$F_H(\theta,\phi)=\int_0^a\int_0^{2\pi} A + B[1-\frac{\alpha^2}{a^2}]e^{+jk\ \sin\theta\cos\ (\phi-\phi')}\ pdp\ d\phi' \qquad (3\text{-}41)$$

$$-B\int_{-a}^{a}\int_{-\sqrt{\ }}^{\sqrt{\ }} (\frac{y+y_m}{y_m+a})^2\ (1-\frac{\alpha^2}{a^2})e^{+jk(xx'+yy')/r}\ dx\ dy$$

which is recognized as the pattern of the preceding case minus a term produced by the offset reflector. Thus with an offset reflector

$$F_H = F_{H_4} - OFFSET(y_m,\ a,\ \theta,\phi)\ *B \qquad (3\text{-}42)$$

where

$$OFFSET = \int_{-a}^{a}\int_{-\sqrt{a^2-y^2}}^{\sqrt{a^2-y^2}} (\frac{y+y_m}{y_m+a})^2(1-\frac{x^2+y^2}{a^2})e^{+jk(xx_o+yy_o)/r}\ dx\ dy \qquad (3\text{-}43)$$

Integrating with respect to x gives (after a considerable effort)

(3-44)

$$OFFSET=\frac{4}{a2}\int_{-a}^{a} e^{+j\ w_y\ y}\ (\frac{y+y_m}{y_m+a})^2\{\frac{\sin[w_x\sqrt{a^2-y^2}]}{w_x^3} - \frac{\sqrt{a^2-y^2}\ \cos(w_x\sqrt{a^2-y^2})}{w_x^2}\}dy$$

where $w_x = k\ \sin\theta\ \cos\phi$ (3-45)

$w_y = k\ \sin\theta\ \sin\phi$

$x = ka\ \sin\theta$.

Two special cases for the offset correction term (OFFSET) are of interest, that for the E-plane pattern ($w_y=0$) and that for the H-plane pattern ($w_x=0$). Equation (3-44) simplifies to

E Plane ($w_y=0$):

$$\text{OFFSET}(w_x,0)= \frac{8}{a^2(y_m+a)^2} \int_0^a (y^2+y_m^2)\left[\frac{\sin(w_x\sqrt{a^2-y^2})}{w_x^3} - \frac{\sqrt{a^2-y^2}\cos(w_x\sqrt{a^2-y^2})}{w_x^2}\right]dy \tag{3-46}$$

H plane ($w_x=0$):

$$\text{OFFSET}(w_y)= \frac{8}{3a^2(y_m+a)^2}\left\{\int_0^a (y^2+y_m^2)\cos(w_y y)(a^2-y^2)^{1.5}dy+2j\int_0^a yy_m\sin(w_y y)(a^2-y^2)^{1.5}dy\right\}$$

The last two integrals were numerically integrated, and the E-plane and H-plane patterns are plotted in Figure 3-7 for the offset aperture distribution

$$f_{ap}=0.3 + 0.69\left(1-\frac{p^2}{a^2}\right) - 0.69\left(\frac{y+y_m}{y_m+a}\right)^2 \left(1-\frac{p^2}{a^2}\right) . \tag{3-48}$$

The parameters chosen are those of an offset reflector with an F/D of 0.8 and a half-degree beamwidth(-6dB) at the Ku-band frequency of 15 GHz. The required dish diameter was D=3.6 meters and the offset beam maximum position (y_m) was 14.82 cm. The feed antenna had a -10 dB beamwidth of 64^o with the beam centered at $\Psi=32^o$. Comparing Figures (3-5) and (3-7) it is observed that using the offset reflector has not changed the patterns drastically. The H-plane sidelobe level was increased about 1 dB above the centered reflector while the H-plane sidelobe level was decreased to -32dB.

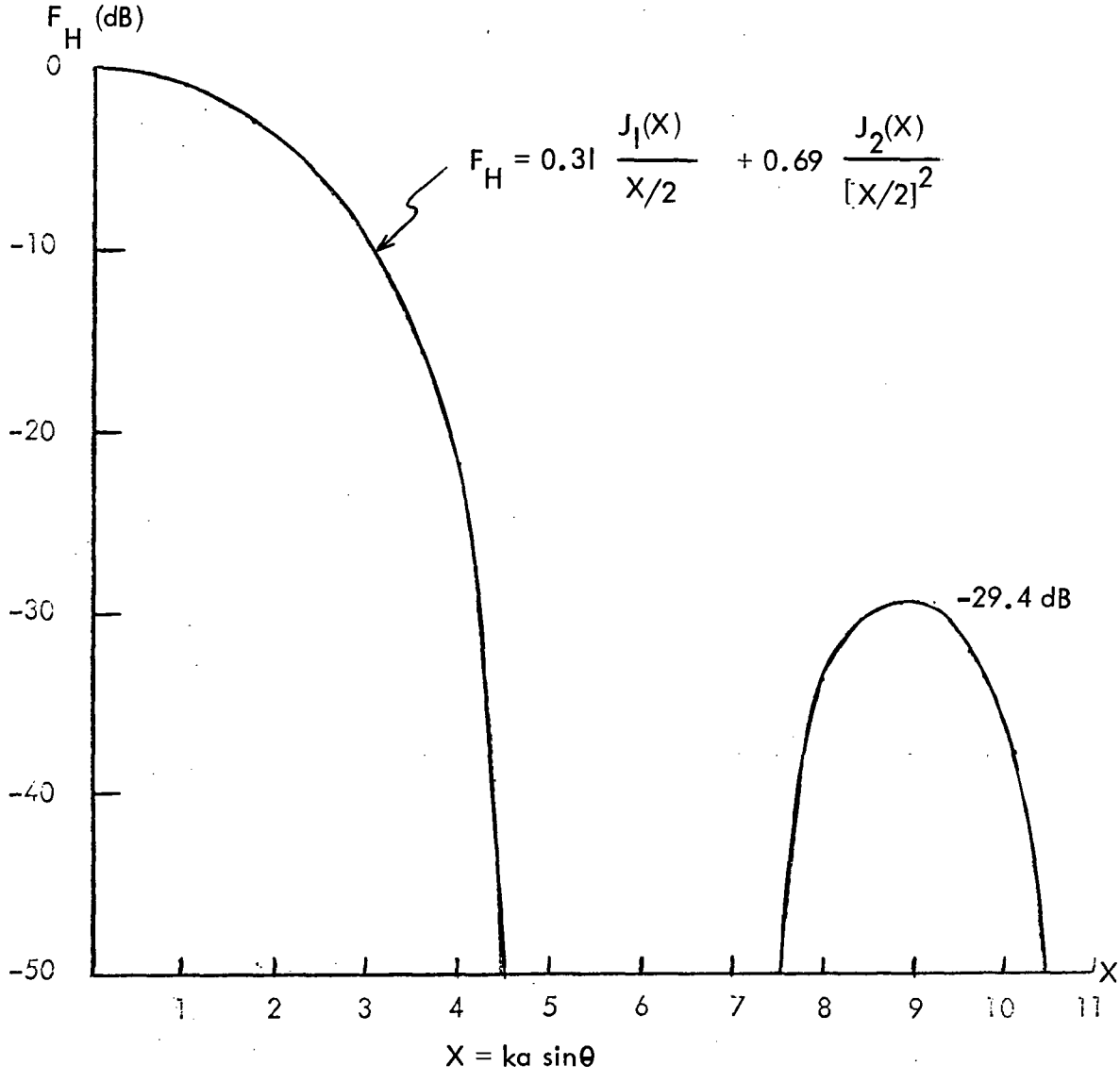


Fig. 3-5 Far-Field Pattern for a Paraboloid with a -10 dB Tapered Aperture Illumination

$$f_{aperture} = 0.31 + 0.69[1-\rho^2]$$

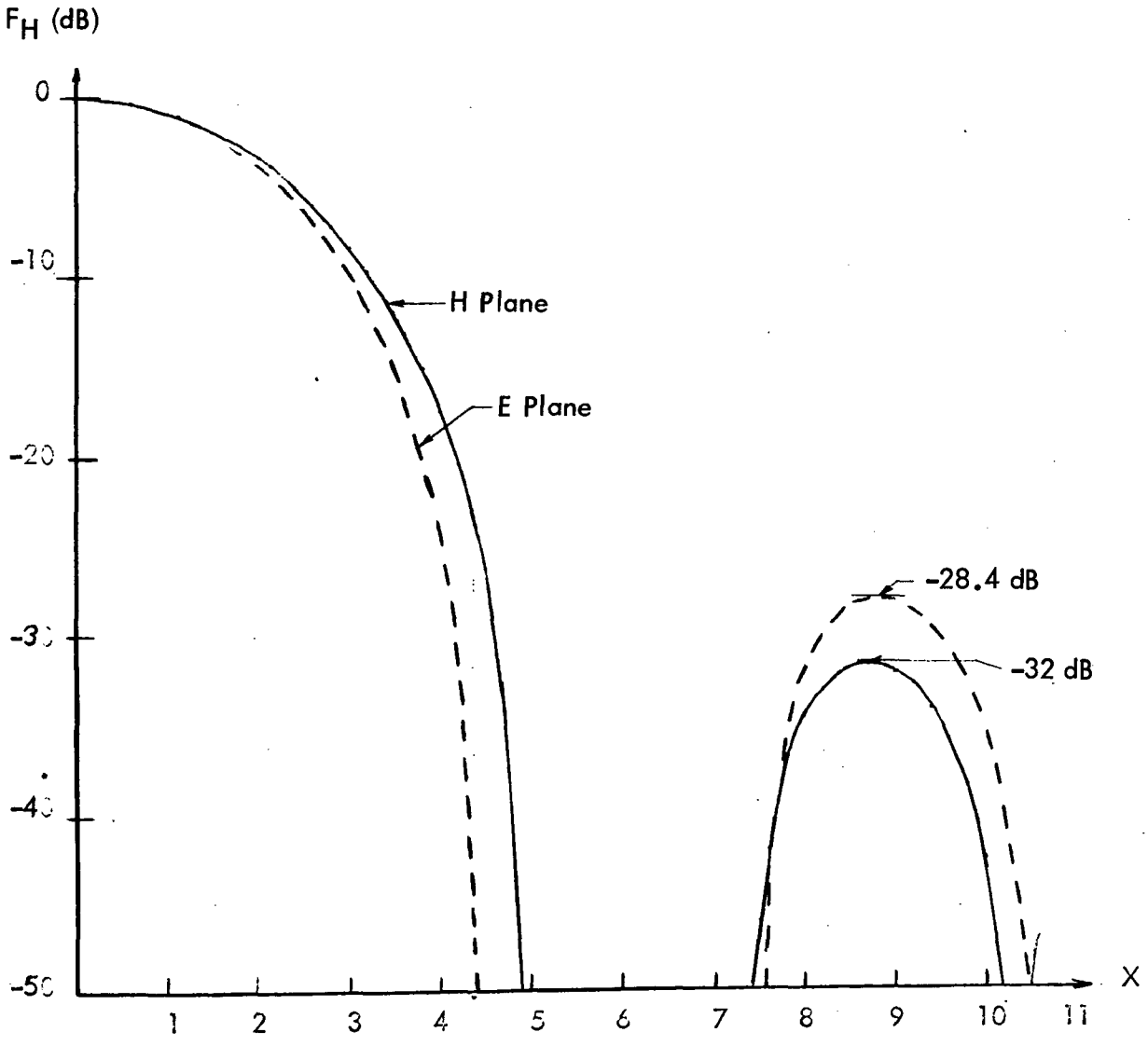


Fig. 3-7 Patterns for Paraboloid with Offset Reflector

B. Conical Beam Offset Feeds

Since offset feeds are used to point the beams, it is desirable to have design guidelines for dealing with such structures. Because some of Ruze's work [6], which was introduced in Chapter II, serves this purpose, it is worthwhile to highlight some of his more important results. The geometry used is drawn in Figure 2-5 (page), and the results to follow apply to circular apertures with feed elements offset by $\vec{\delta} = \delta_x \vec{a}_x + \delta_z \vec{a}_z$. The feed elements are placed on a Petzval surface to remove some of the aberration forming terms in the far field pattern that could otherwise result. For $2F\delta_z = \delta_x^2$ the far field magnitude is

$$|E(\theta,\phi)| = \int_0^a \int_0^2 f(p,\phi') e^{+jk[p \sin\theta \cos(\phi'-\phi) - \delta_x \sin\theta' \cos\phi']} p \, dp \, d\phi' \, . \tag{3-49}$$

The aperture illumination function, $f(p,\phi')$, is assumed to be circularly symmetric. This is done to simplify the analysis, and it must be realized that this is only approximately so. Thus the results are only approximate, however, they are close enough in most instances to be fairly good indicators. The analysis for nonsymmetric illuminations is a good topic for future study.

The exponent in (3-49) may be put in a more compact form by noting

$$\sin\theta' = \frac{p/F}{1+(\frac{p}{2F})^2} = \frac{p}{F}[1-(\frac{p}{2F})^2 = (\frac{p}{2F})^4 - \ldots] \tag{3-50}$$

and defining

$$M(p) = 1 + (\frac{p}{2F})^2 \tag{3-51}$$

$$U_S = \frac{\delta_x}{F} = \tan\theta_S \, .$$

Then using several trigonometric identities one can write

$$kp[\sin\theta \cos(\phi-\phi') - \frac{U_S \cos\phi'}{M(p)}] = Akp \cos(\phi' - \alpha) \tag{3-52}$$

where

$$A = \sqrt{\sin^2\theta - \frac{2\sin\theta\, U_S}{M}\cos\phi + \frac{U_S^2}{M^2}} \tag{3-53}$$

$$\alpha = \tan^{-1}\left[\frac{\sin\theta\ \sin\phi}{\sin\theta\ \cos\phi - U_S/M}\right] \tag{3-54}$$

Substituting (3-52) in (3-49) the far field magnitude becomes

$$|E(\theta,\phi)| = \int_0^a \int_0^{2\pi} f(p,\phi')e^{+jkp\ A\ \cos(\phi'-\alpha)}\, p\ dp\ d\phi' \tag{3-55}$$

Since the aperture illumination has been assumed to be circularly symmetric [f = f(p)], the integration with respect to ϕ' can be performed to give

$$|E| = 2\pi \int_0^a f(p)\ J_o\ (kp\ A)\ p\ dp \tag{3-56}$$

Ruze used the computer to evaluate (3-56) for various combinations of aperture illumination and F/D. Some of his results are plotted in Figures (3-8) to (3-11), and these can be useful in formulating an offset feed design. Several of the plots are expressed as a function of the quantity

$$X = \frac{\frac{w_m}{2w_o}\ (D/F)^2}{1 + 0.02\ (D/F)^2} \tag{3-57}$$

where

$$\frac{w_m}{2w_o} = \text{number of half-power beamwidths scanned.} \tag{3-58}$$

From Figure 3-9 it is seen that for large X, the sidelobe levels caused by coma(coma lobes) are quite high (order of -10dB). For instance, with an F/D of 0.8 and a half-power beamwidth of 0.5^o,

Figure 3-11 indicates that $w_m/2w_o$ is about 20. Hence X by (3-57) is about 30 giving a gain loss over 2 dB and a -10dB coma lobe.

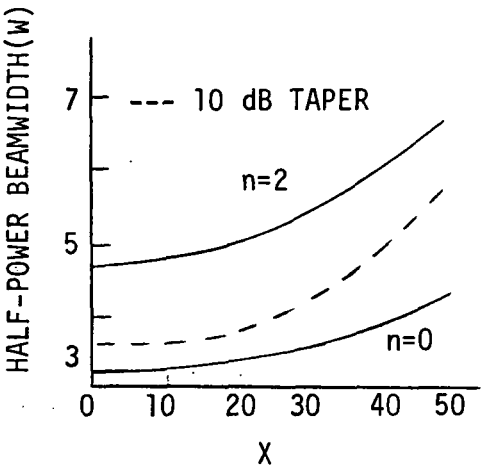


Fig. 3-8 Half-power Beamwidth

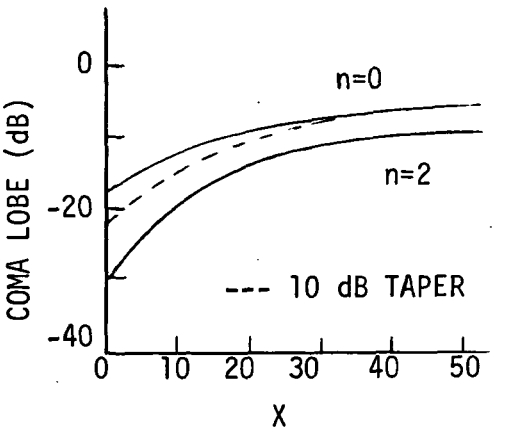


Fig. 3-9 Coma Lobe

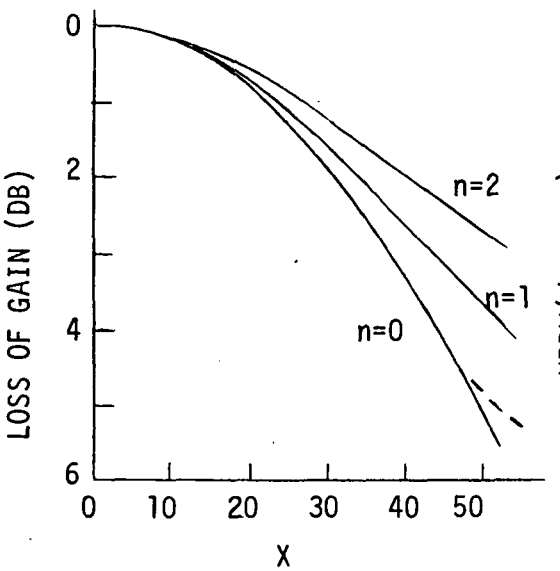


Fig. 3-10 Loss of Gain

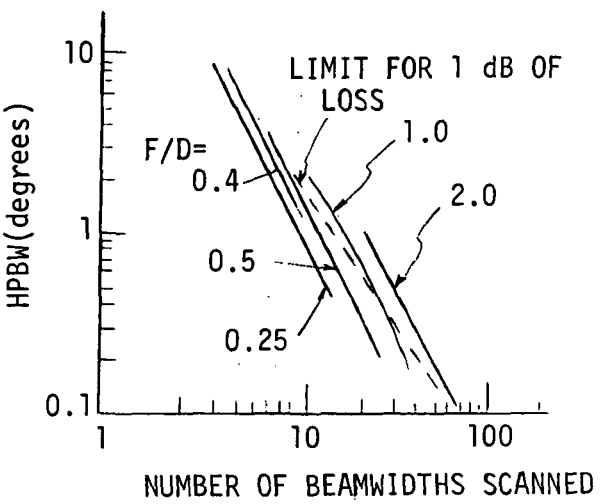


Fig. 3-11 Region of Validity

CHAPTER IV

CONCLUSIONS

At the end of Chapter II are recommendations for a practical shaped-beam antenna system using a parabolic reflector. Some analytic results are developed in Chapter III which include (1) aperture distributions and patterns for offset reflectors and, (2) simplified patterns for offset feed elements and their associated beam deviation and pointing effects. After solving the offset reflector problem, it was concluded that the offset reflector geometry, although interesting, is not necessary for shaped-beam coverage of the U. S. from a synchronous orbit satellite.

Some potential study areas for future effort would be:

1. A multielement feed array should be built so that mutual coupling can be measured and estimates obtained as to its effect on multibeam radiators.

2. The antenna pattern for single beams obtained in this study should be superimposed to synthesize an actual multibeam, shaped pattern. This investigation should result in recommendations regarding the number of beams needed to "best" illuminate a complicated time zone, the feed amplitudes and phases needed to realize patterns, and crossover levels between the patterns of adjacent beams (initially would use -6dB points).

3. The effects of element failures on the ground area coverage should be considered. It would certainly be desireable that an element failure not be catastrophic to an area. This goal would probably relate to the selection of the number of beams used per region.

4. Sidelobe level increases produced by aperture blocking should be calculated. The pattern modifications caused by this effect should be included in the multibeam pattern synthesis results.

5. The significance of polarization deterioration caused by a large F/D reflector should be analyzed or measured for offaxis feeds.

6. Recommendations which are complete down to specifications and dimensions of an actual shaped-beam antennas should be made.

REFERENCES

1. R. D. Briskman, "Domestic satellite design," Proceedings of 1970 ICC, pp. 38-16 to 38-23

2. M. I. Skolnik, Introduction to Radar Systems, McGraw-Hill, pp. 286-294.

3. R. C. Hansen (ed.), J. L. Bulter, Microwave Scanning Antennas, Vol. III, Academic Press, 1966, Chapter 3.

4. A. C. Ludwig, "Near field, far field expansions using spherical wave expansions," IEEE Trans. AP, March 1971, pp. 214-220.

5. Collin and Zucker (eds.), C. J. Sletten, Antenna Theory, Pt. 2, McGraw-Hill, 1969, Chapter 17.

6. J. Ruze, "Lateral feed displacement in a paraboloid," IEEE Trans. AP, September 1965, pp 660-665.

7. Microwave Journal, February 1970 (technical and buyers guide edition), pp. 47, 53.

8. J. W. Blaker, Geometric Optics - The Matrix Theory, Marcel Dekker, Inc., New York, 1971, Chapter 7.

9. E. A. Wolff, Antenna Analysis, John Wiley and Son, 1966, Chapter 7.

10. L. C. Martin, Technical Optics, Vol. 2, London, Sir Isaac Pitman and Sons, pp. 74.

11. L. Rainwater, 1972 Sperry Rand report to be published plus many helpful discussions, Huntsville, Alabama, August, 1972.

12. J. Dudgeon, internal memo, July 1972.

13. M. Hamid and A. Mohsen "Diffraction by dielectric-loaded horns and corner reflectors, " IEEE Trans. AP, September 1969, pp.660-662.

1972

NASA - ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

SMOKE PLUME ANALYSIS USING FAN-BEAM AND SINGLE-BEAM RADIOMETERS

PREPARED BY:	JIMMY L. DODD
ACADEMIC RANK:	ASSOCIATE PROFESSOR
UNIVERSITY:	MISSISSIPPI STATE
LABORATORY: (Division)	AERO-ASTRODYNAMICS FLIGHT DATA STATISTICS OFFICE
RESEARCH COUNTERPART:	JAMES W. BILBRO
DATE:	AUGUST 11, 1972
CONTRACT NUMBER:	NGT-01-003-045

SMOKE PLUME ANALYSIS USING FAN-BEAM AND SINGLE-BEAM RADIOMETERS

by

Jimmy L. Dodd

ABSTRACT

A fan-beam and single-beam radiometer system has been designed and constructed for NASA by the IIT Research Institute, Chicago, Illinois. This system has been previously used to study cloud movement and smoke movement.

Smoke plume data were taken using the stacks of the Mills Road Steam Plant on Redstone Arsenal, Huntsville, Alabama. The fan-beam unit was placed approximately 235 meters east of the Mills Road Steam Plant and the single-beam unit was placed approximately 80 meters north of the fan-beam unit. Results of the smoke plume measurements indicate that the smoke plume is a periodic process near the mouth of the smokestack.

The usefulness of the system is limited because of inaccurate and cumbersome positioning methods and base line constraints due to cable length. It is recommended that the radiometers be mounted on pedestals which have both azimuth and elevation positions determined by digitally controlled stepper drive motors. It is also recommended that a time multiplex system be used to transmit data from the radiometers to the electronics van.

INTRODUCTION

The fan-beam and single-beam radiometer system has been previously used by other investigators including Morrison (1) and Latimer (2). This system is described in detail in the IIT Research Institute's final report (3). A van is used to house most of the electronics of the system in addition to housing several other items of instrumentation. The fan-beam and single-beam units are mounted in small trailers to facilitate field measurements.

The usefulness of the system is limited to a great extent by the length of the cables connecting the fan-beam and single-beam units to the electronics van. Another limitation of the system is the imprecise and cumbersome method that must be used to point the telescopes of the fan-beam and single-beam units.

An open area north of Building 4311 was used as the location for setting up the system. The Mills Road Power Plant, Redstone Arsenal, Alabama, was used as the smoke plume source. A base line from fan-beam unit to single-beam unit of about 80 meters was used.

OBJECTIVES

The objective of the summer research program was to investigate the usefulness and limitations of the fan-beam unit and single-beam unit radiometer system. In addition, meaningful experiments using this system were to be recommended. Recommendations concerning modifications and improvements to the system were to be made.

It was hoped in the early part of the summer that a measurement of an inversion layer would be made with this system. Due to cable difficulties during the time of a lengthy inversion layer cover, no inversion layer measurements were made. After the cables had been repaired, not enough time remained to make inversion layer measurements and analyze the data.

TEST RESULTS

Figure 1 shows the physical arrangement used for the experiments that were made. One beam of the fan-beam unit and the single-beam

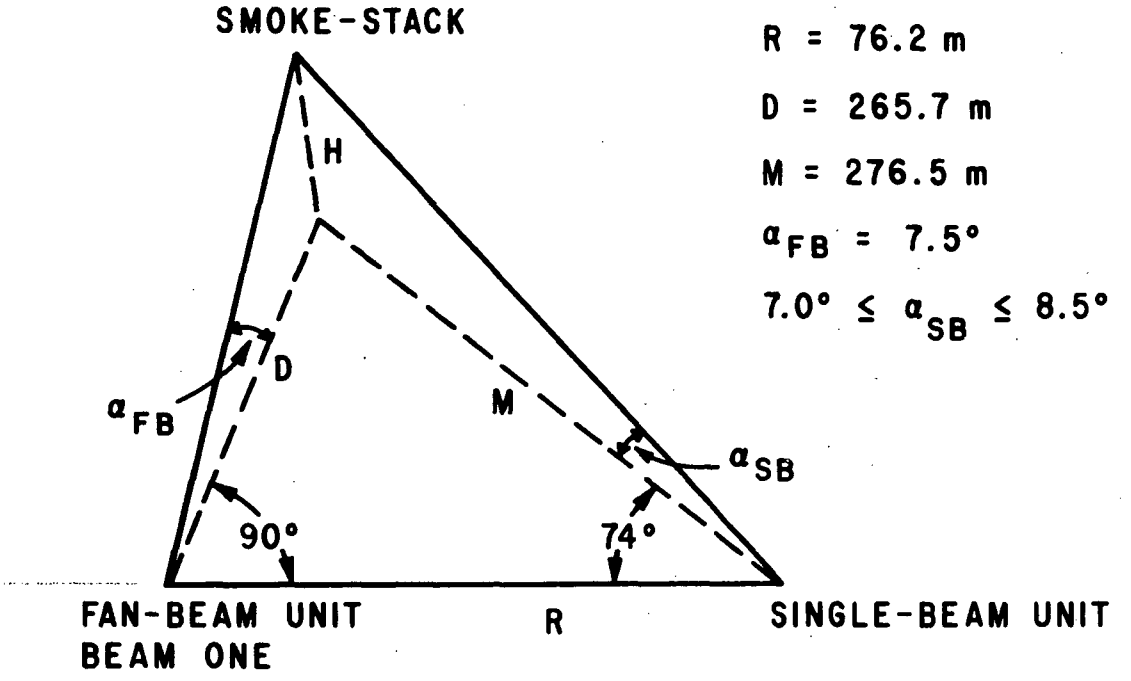


FIG. 1. GEOMETRY FOR FIELD TESTS

unit were both pointed just above the mouth of one of the smokestacks. The elevation of the fan-beam unit was maintained constant while the elevation of the single-beam unit was varied from 1/2 degree below the fan-beam elevation to 1 degree above.

Figures 2 through 5 show the movement of the cross-correlation peak from about 1 second to the left of zero time lag to about 2 seconds to the right of zero time lag corresponding to the single-beam angle of elevation being varied from 7 degrees to 8.5 degrees.

By using the change in the peak of the cross-correlation curve from -1 to +1 second corresponding to a change in single-beam angle of elevation of 1^{o}, the wind speed was calculated to be 2.45 m/s. The calculations are as follows:

$$H = D \tan \alpha_{SB}$$

Where H is the smoke plume height in meters, D is the horizontal distance from the single-beam unit to the smokestack in meters and α_{SB} is the angle of elevation of the single-beam unit in degrees. The wind speed equation is as follows:

$$S = \frac{H_8 - H_7}{T_8 - T_7}$$

Where S is the wind speed in m/s, H_8 is the smoke height in meters for an elevation angle of 8^{o}, H_7 is the smoke height in meters for an elevation angle of 7^{o}, T_8 is the cross-correlation peak time for an elevation angle of 8^{o}, and T_7 is the cross-correlation time for an elevation angle of 7^{o}.

RECOMMENDATIONS

The fan-beam and single-beam radiometer system has been shown to be useful in measuring smoke plume and cloud movement.

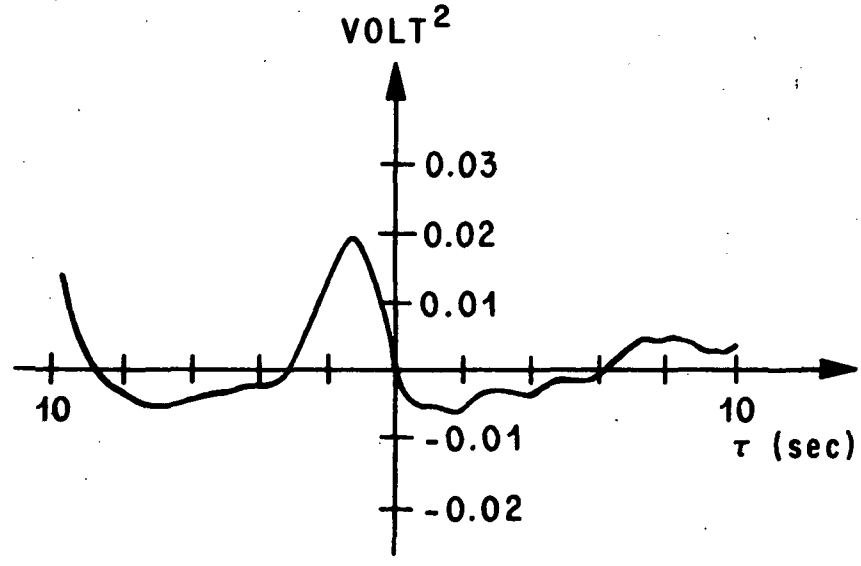


FIG. 2. CROSS-CORRELATION BETWEEN SINGLE-BEAM AND ONE BEAM OF FAN-BEAM, $\alpha_{SB} = 7.0°$, $\alpha_{FB} = 7.5°$.

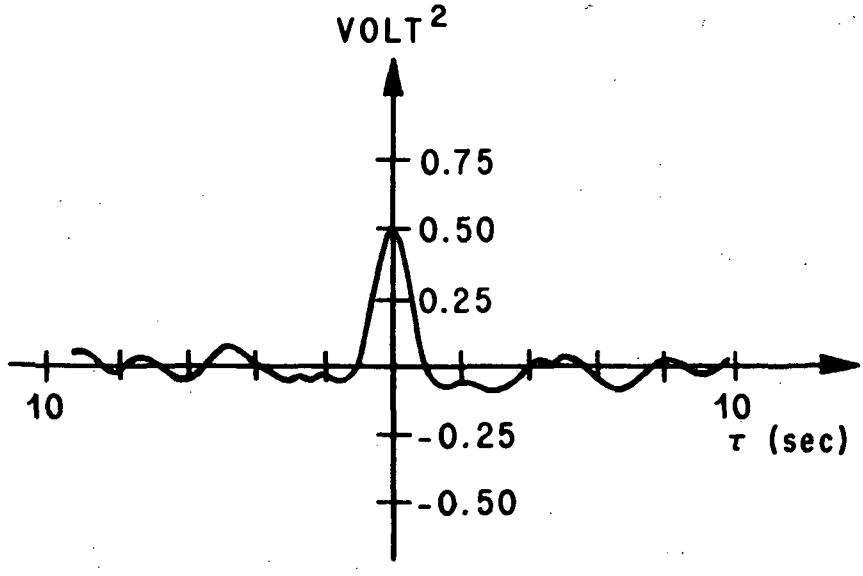


FIG. 3. CROSS-CORRELATION BETWEEN SINGLE-BEAM AND ONE BEAM OF FAN-BEAM, $\alpha_{SB} = 7.5°$, $\alpha_{FB} = 7.5°$.

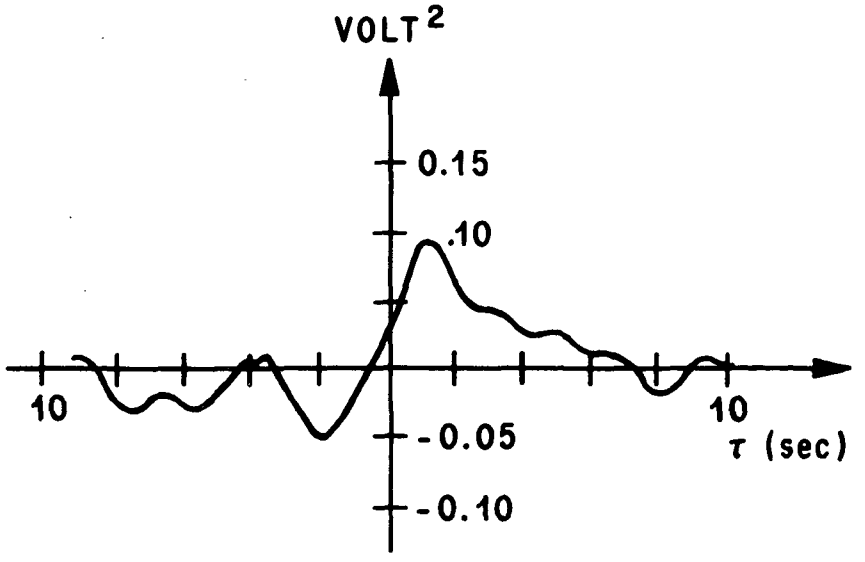


FIG. 4. CROSS-CORRELATION BETWEEN SINGLE-BEAM AND ONE BEAM OF FAN-BEAM, $\alpha_{SB} = 8.0°$, $\alpha_{FB} = 7.5°$.

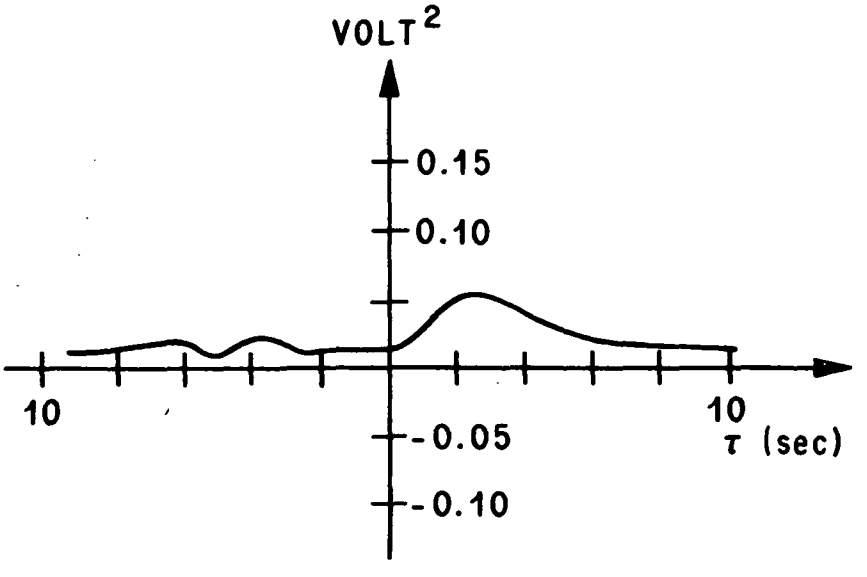


FIG. 5. CROSS-CORRELATION BETWEEN SINGLE-BEAM AND ONE BEAM OF FAN-BEAM, $\alpha_{SB} = 8.5°$, $\alpha_{FB} = 7.5°$.

Improvements need to be made in the positioning of the telescopes and in handling the data from the system. It is recommended that both fan-beam and single-beam unit be mounted on pedestals, the azimuth and elevation of which can be remotely controlled. It is recommended that digitally controlled stepper drive motors be used. A time shared multiplex data system should be used to transmit the data from the units to the electronic van. The data should be stored on magnetic tape and analyzed by a digital computer.

REFERENCES

1. Morrison, R. A, "Experience with the Crossed-Beam Photometer System," 1971 NASA-ASEE Summer Faculty Research Program Report, Addendum, Marshall Space Flight Center, September 1971.

2. Latimer, Paul, "Crossed-Beam Surveillance of High Stack Emissions," 1971 NASA-ASEE Summer Faculty Research Program Report, Addendum, Marshall Space Flight Center, September 1971.

3. Betz, H. T., "Modification of Fan-Beam and Single-Beam Radiometer," IIT Research Institute, Chicago, Illinois, Project V6124, Contract NAS8-28145, June 1972.

Page Intentionally Left Blank

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

A PRELIMINARY STUDY OF A BUOY SYSTEM FOR ACQUISITION, TRANSMISSION, AND MANAGEMENT OF HYDROLOGICAL DATA OBTAINED FROM IN-SITU MEASUREMENTS

Prepared by:	J. Mark Elliott
Academic Rank:	Assistant Professor
University:	University of South Alabama
Office:	Environmental Applications
Research Counterpart:	George F. McDonough
Date:	August 25, 1972
Contract No:	NGT-01-003-045

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to Dr. George McDonough, Mr. Rex Morton, and Mr. Jim Daniels of the MSFC - Environmental Applications Office for their support of this project. Also, the author would like to thank Dr. C. Everett Brett, Coordinator, Alabama Marine Environmental Sciences Consortium for his advice in this study.

PREFACE

The Environmental Applications Office, to which the author was assigned, hosted the Auburn University Engineering Systems Design summer program in 1972 and the subject of their study was an "Earth Resources Information Management System." Since the author's 1971 research report (1) considered the desirability of a data storage and management system for land use information, the author was invited to participate in the design group's activity. The objective was to gain insight into the management of earth resources information through interaction with a group in an intense study of the problem area. The concept derived by the group could then be extended by the author into a particular situation, e.g. an information system for the newly formed Alabama Marine Environmental Sciences Consortium. Although this was the desired objective, a final decision for a particular application could not be made until the design group finalized their concept since it had complete freedom to choose an area for application.

Two things occurred during the course of the summer that effectively precluded achievement of the author's objective. First, the design group did not clearly define their concept early enough for the author to use it for the purposes intended. Secondly, a requirement arose within the Environmental Applications Office concerning the Earth Resources Technology Satellite program which began its operational phase during the summer. Therefore it was necessary to essentially abandon the initial objective and proceed as indicated in this report.

The author does feel, however, that the subjects considered by the design group are excellent choices for research, further amplification, analysis, and application. It is recommended that this not be attempted in the same summer. Proper analysis and study of concepts proposed by the design group can only be accomplished after the work is completed.

It should also be evident that topics considered by the Research Fellows may precede consideration of related topics by the design group. The author's 1971 work bears this out.

Participation with the design group was a valuable experience and indeed it did impact on portions of the following report.

A PRELIMINARY STUDY OF A BUOY SYSTEM FOR ACQUISITION, TRANSMISSION, AND MANAGEMENT OF HYDROLOGICAL DATA OBTAINED FROM IN-SITU MEASUREMENTS

By

J. Mark Elliott

ABSTRACT

The requirements for a system of remotely located, data collection buoys are considered first for a prototype system to be used in conjunction with the Earth Resources Technology Satellite (ERTS-A), and then for a more advanced system. The necessary sensor characteristics for compatibility with the ERTS-A Data Collection Platforms are considered as well as possible sites for location of the prototype buoys. The advanced system is considered from the standpoint of continuous data collection both through satellite data relay and ground telemetry systems.

Management of the data from a buoy system is analyzed, especially with regard to the advanced system.

INTRODUCTION

With the successful launch of the first Earth Resources Technology Satellite (ERTS-A) on 23 July 1972, the earth observations program is proceeding in three dimensions: (1) a sensor carrying satellite is aloft in a sun-synchronous orbit; (2) coordinated aircraft underflights are in progress; (3) ground truth data, coordinated with satellite and aircraft passes, are being collected. While the ERTS program is essentially a research and development program, it is anticipated that it will provide much needed data for immediate requirements in such areas as oceanography, land use management, pollution monitoring, etc.

The aspect of the ERTS program most closely related to this study is the Data Collection System (DCS). The DCS consists of three subsystems: the Data Collection Platform (DCP); the receiving and transmitting equipment in the ERTS satellite; and the receiving site equipment. With the DCS, environmental data may be measured at remote locations, fed to the DCP at the remote site, transmitted through the ERTS to the ground receiving station. The data are then available to investigators for analysis and correlation with the satellite or aircraft gathered imagery if desired.

While many different types of data could be obtained through the ERTS/DCS for many different applications, this report will consider a data collection buoy system, equipped with various kinds of sensors, with proper equipment to interface with the DCP. These "data buoys" will be located in rivers, lakes, bays, etc., and will measure the desired parameters, transmitting the information through the DCP, then through the ERTS on to the ground. Depending on the kinds of parameters measured, and the frequency of measurement, and the distribution of the buoys, the data may be used to (1) develop analytical models of rivers, bays, etc., (2) monitor water quality, (3) obtain base-line data for use in follow-on modeling or monitoring (obtaining historical records in the process) or (4) provide inputs to specific research projects.

The requirements for each of these objectives will be discussed in the report, but basically a phased approach to a data buoy system will be considered. First of all, the requirements for a prototype system will be analyzed. This prototype system will consist of a small number of buoys and the uses of the data will be somewhat limited. However, the prototype system will be relatively inexpensive, and it will transmit data through ERTS; in short it will be a research and development system.

Requirements will then be considered for a truly operational system with a sufficient number of buoys and a sufficiently sophisticated data gathering system to provide near real time capability if desired. While such a system will be quite expensive, it will provide data with the potential for much wider applications.

PROTOTYPE DATA BUOY SYSTEM

University of Alabama/MSFC ERTS-A Project

The University of Alabama, Tuscaloosa, in coordination with the Marshall Space Flight Center is participating in an ERTS-A project designed to apply remotely sensed data in the management of natural resources and improvement of environmental quality in Alabama (2). One phase of this project is the placement of a data collection buoy system in the Black Warrior - Tombigbee - Mobile Bay River system shown in Figure 1 . Ideally, a sufficient number of buoys, collecting a variety of data at a high rate should be employed so that the entire river system could be characterized with respect to such items as temperature, current, water quality (dissolved oxygen, Ph, turbidity, salinity, etc.), and energy budget.

Such a system would provide input data for physical and analytical models, for monitoring the quality of the water throughout the system, and for special research projects. However, two constraints have been imposed on the buoy system: (1) the data collected by the buoys will be relayed through the ERTS/DCS; and (2) funds and manpower are available to construct only four operating buoys.

ERTS/DCS Constraints (3)

The use of the ERTS/DCS for data acquisition imposes several additional constraints on the data buoy system. First of all, the DCP can accept only a maximum of 8 analog inputs, or up to a maximum of 64 bits of serial digital or parallel digital data, or any combination of the above. In addition, the data will be relayed through ERTS at least once every 12 hours. The DCP itself transmits the data it "sees" at its terminals every 180 seconds in a 38-millisecond burst, but due to the orbital parameters and location of ground receiving sites, only one message relay every 12 hours is certain.

Required Measurements and Alternatives

Among the alternatives for data collection are:

1. Continuously measure up to 8 different parameters and relay this data whenever the ERTS is mutually available to the DCP and a ground station (approximately 9:20 AM and 9:20 PM CST).

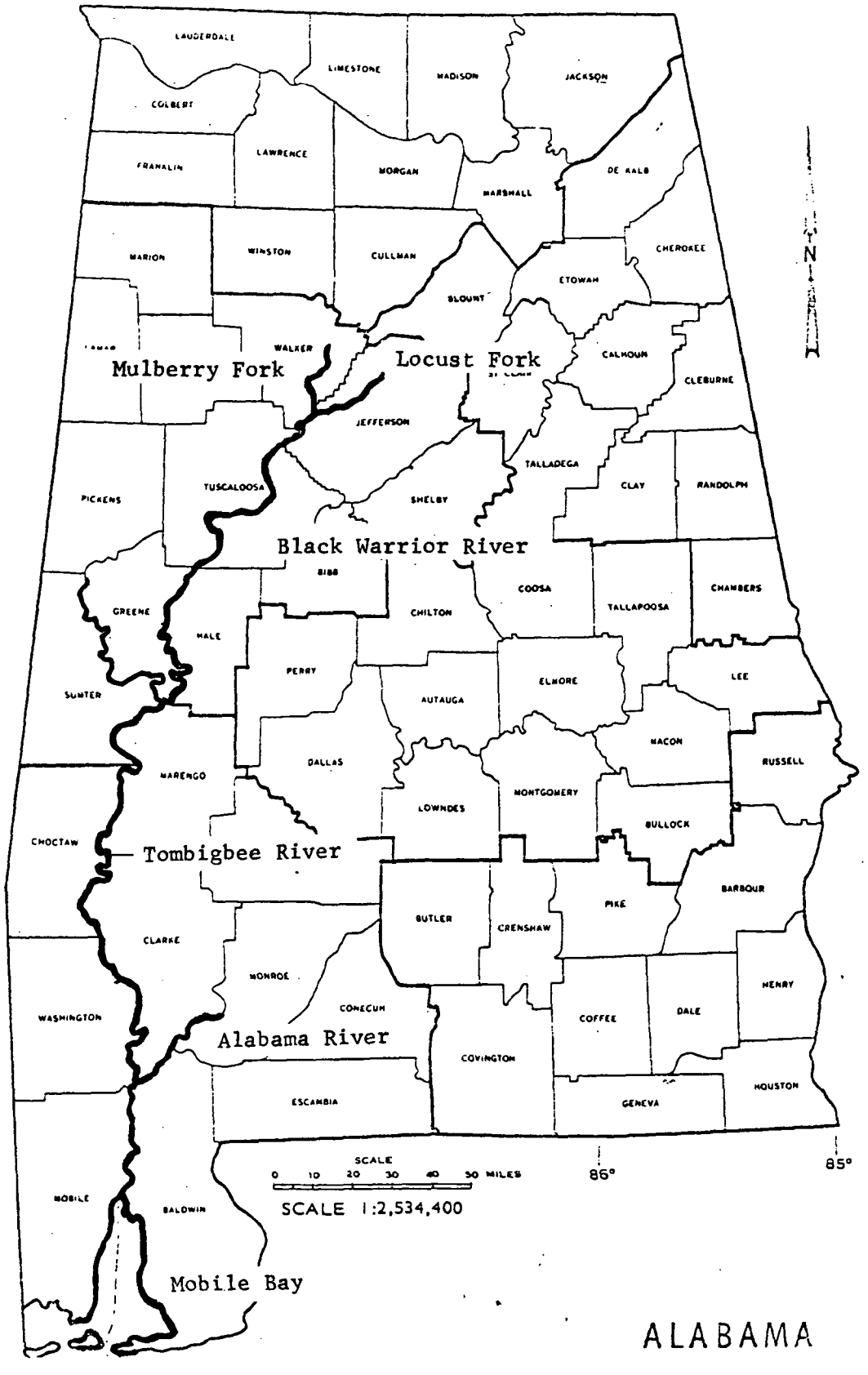


Figure 1: Black Warrior - Tombigbee - Mobile Bay River System

2. Measure up to 8 parameters twice a day at a specific time other than when ERTS is overhead, e.g. 12:00 AM and 12:00 PM, store this data on tape or buffer memory and make it available to the DCP during the ERTS pass.

3. Measure 4 parameters twice every 12 hours, store this information and relay through ERTS as in alternative No. 2. (This alternative can be carried to the ultimate alternative of measuring one parameter 8 times every 12 hours, etc.).

The first alternative would be the simplest in that no data storage equipment would be required. (A timer might be used to switch the sensors on and off in order to conserve power.) However, investigators on the project have selected only 4 parameters to be measured by buoys located in the river (the requirements for the Mobile Bay buoys will be discussed later)(4):

1. Temperature;

2. Specific conductivity;

3. Dissolved oxygen;

4. Ph.

With 4 measurements to be made, it would be possible to adopt alternative No. 3 if the necessary timer and auxiliary data storage equipment were used in the buoy. In the fresh water at least, data would be available for 6 hour intervals.

For the buoys to be located in Mobile Bay, 5 parameters are desired every hour (5):

1. Temperature;

2. Specific conductivity;

3. Dissolved oxygen;

4. Current direction;

5. Current velocity.

The requirement for 5 parameters will be difficult to attain if ERTS relay is the only method of data collection. Three alternatives for data collection through ERTS relay are:

1. Measure only 4 parameters and adopt alternative No. 3 for the fresh water buoys, i.e. measure each parameter twice every 12 hours;

2. Use two DCP's for each buoy. This would provide a 4 hour data interval measuring 5 parameters or a 3 hour data interval for 4 parameters;

3. Use one DCP and measure 5 parameters with a 12 hour data interval.

If the one hour data interval requirement must be met, a fourth alternative is to:

4. Incorporate a separate telemetry system in each buoy with a transmitter with the buoy and receiver at a ground station or a receiver only for receiving the signal from the DCP.

Alternative No. 4 appears to be the only feasible approach if a one hour data interval is required. If the DCP transmitter is used, the signal must be reformatted after it is received because the DCP encodes the data in a special format prior to transmission.

Distribution of Buoys

Ideally, a system of buoys should be placed throughout the river system with some minimum distance between each buoy. Also, several buoys should be distributed throughout Mobile Bay. However, as indicated previously, it appears that only four operational buoys will be constructed. Logical arguments could be made for each of the following alternatives:

1. Locate all buoys in Mobile Bay;

2. Locate all buoys in the river in a particular area;

3. Distribute the buoys evenly along the length of the river.

However, since this to be a prototype system, the buoys should be located to meet the following requirements:

1. At least one buoy be placed in salt water (Mobile Bay) and one in fresh water so that buoy performance can be observed in the two environments;

2. All buoys be located so that periodic maintenance as well as emergency repairs can be performed quickly and easily;

3. As many of the buoys as possible be located in such a manner as to indicate the impact of man's activities on water quality.

All three of the above requirements will be satisfied if the buoys are located in two distinct groups: one set in Mobile Bay and the other in the Black Warrior River near Birmingham or Tuscaloosa as shown in Figure 2 . This would allow maintenance of the bay buoys to be performed out of the Dauphin Island Sea Lab and maintenance for the river buoys out of the University of Alabama, Tuscaloosa. Both the Mobile and Birmingham/Tuscaloosa locations relate to requirements one and three.

One possible scheme for locating the buoys is to place one buoy in Mobile Bay, one buoy in the Locust Fork of the Black Warrior River, one buoy in the Mulberry Fork of the river, and one buoy in the Black Warrior downstream of the confluence of the Mulberry and Locust forks but upstream of the Bankhead lock and dam. This distribution would:

1. Provide a comparison of water quality between the Locust Fork, which flows past Birmingham and carries wastes from that city, and the Mulberry Fork which flows through a more sparsely developed region to the Northwest (although there are strip mines in this area);

2. Indicate water quality after the two forks of the river are mixed;

3. Meet all three of the stated requirements.

The obvious disadvantage to this scheme is that only one buoy will be located in Mobile Bay. However, the one buoy will provide the opportunity to observe buoy performance in a harsher environment with salt water, tidal fluctuations, and possible severe weather. Also, with only one buoy in the bay, two DCP's might be committed to a single buoy as discussed in the section on Required Measurements and Alternatives, or a separate telemetry capability might be used.

Application of Data

Data relayed through the ERTS will be available to the investigators within 24 hours after the measurements are relayed by the spacecraft so that information on the water quality will be available within a relatively short period of time. Further application of the data, especially for modeling purposes, and more especially with respect to Mobile Bay, will be somewhat limited. Specialized studies, such as

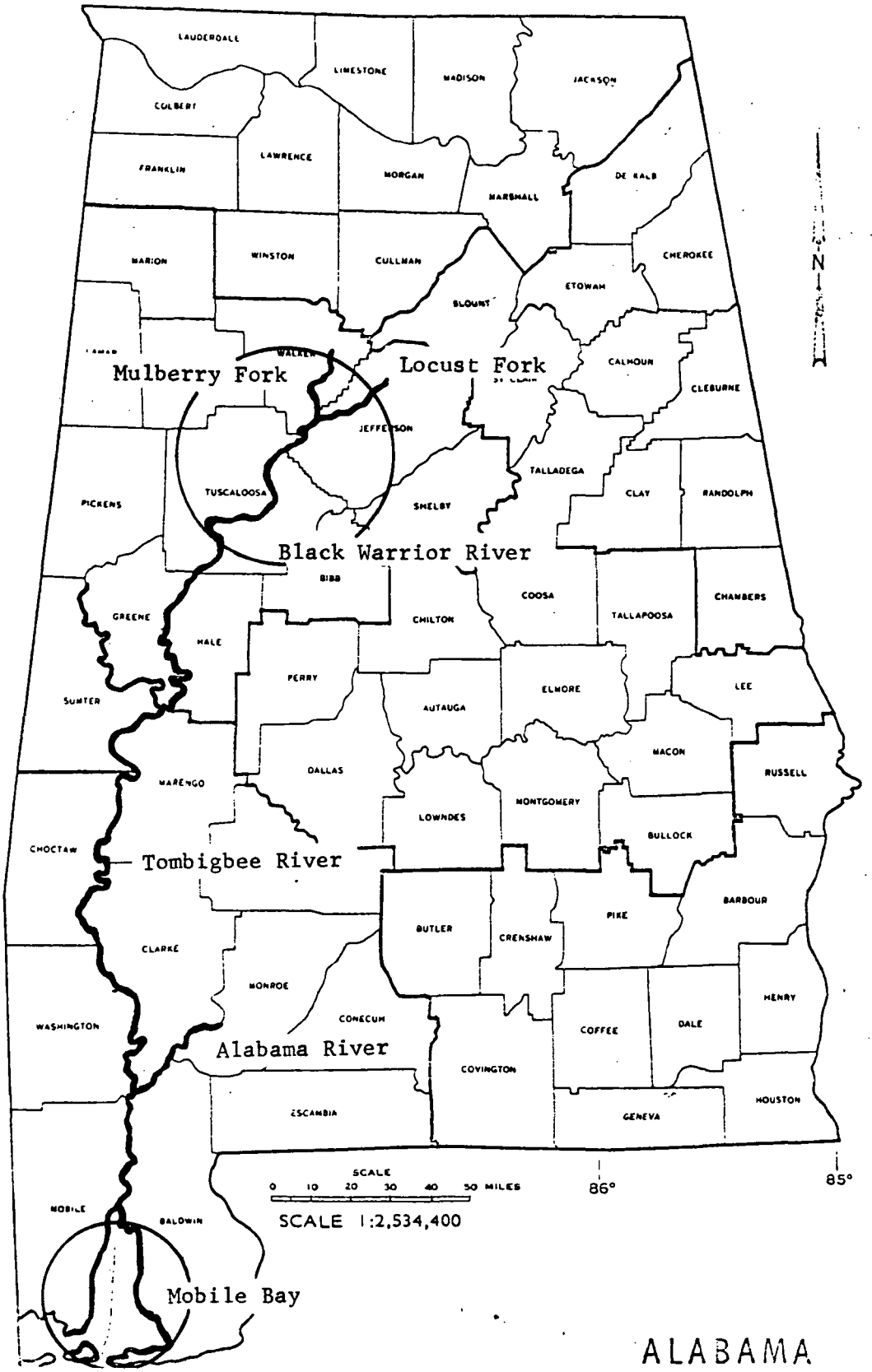


Figure 2: Areas for Prototype Buoy Location

those concerned with the effects of wastes from Birmingham or strip mining on water quality might benefit from the prototype system.

It should be re-emphasized that with the small number of buoys and the limited number of measurements being made, the prototype system is a research and development system which hopefully will lead to a more sophisticated operational system.

ADVANCED DATA BUOY SYSTEM

An advanced data buoy system should be designed to provide data for two primary objectives: 1) physical and analytical modeling of the river system (including the bay) and 2) water quality monitoring. By meeting these primary objectives, the secondary objective of providing data for special research projects would also be met for a broad range of problems. This section of the report will consider the requirements and various alternatives for an advanced data buoy system which is not subject to the same constraints imposed on the prototype system.

Data Requirements

Typical measurements used to monitor water quality are (6)

1. Temperature;
2. Dissolved oxygen;
3. Specific conductivity;
4. Ph;
5. Turbidity.

Additional parameters might be required at specific locations or for special studies (e.g. radioactivity near nuclear power plants) so that provisions should be made for extra data channels to be added as required. It should be noted that changes in the above parameters might be the result of changes in some non-measured parameter. Therefore these measurements indicate when it would be necessary to obtain water samples for more complete analysis.

For modeling, typical measurements required are (7)

1. Temperature (three-dimensional distribution);
2. Atmospheric pressure;
3. Incident solar radiation;
4. Outgoing radiation;
5. Relative humidity;
6. Cloud cover;
7. Wind velocity;

8. Wind direction;

9. Current velocity;

10. Current direction.

Combining both lists produces 14 separate parameters which need to be measured. Cloud cover and wind conditions might be obtained from Weather Bureau data, which reduces the number of required parameters to 11. It would be advantageous to have the capability for additional sensors which could be used to obtain additional parameters as required. Therefore, 15 to 20 separate channels of data might be measured by each buoy.

Table I lists each of the 11 parameters to be measured, indicates why each is important, and lists a representative sensor for obtaining this parameter. It should be noted that a three dimensional distribution of the various parameters would be desirable for modeling purposes. However, the number of separate sensors to accomplish this would be staggering. Therefore, three dimensional distributions will have to be considered relative to obvious constraints in the next section.

Buoy Distribution

Taking the Black Warrior - Tombigbee - Mobile Bay river system as an example, the distribution of buoys will be determined based on the following requirements:

1. Adjacent buoys should be a maximum of 50 miles apart;

2. A buoy should be placed in each main tributary 5 to 10 miles upstream from the confluence of the tributary and the main river;

3. A buoy should be located immediately downstream of industries and municipalities which discharge effluents into the river.

Figure 3 shows the minimum distribution that meets requirements one and two, and Table II gives the location of each site. A detailed inspection of the river using aerial photography as well as a first-hand inspection would be necessary to adequately meet requirement No. 3, but such an inspection should be made prior to final site selection for the individual buoys. Existing aerial surveys, e.g. (8) could be used to select tentative locations to meet requirement No. 3. Specific locations would depend on where effluent discharge points are actually located along the river.

In addition to existing panchromatic aerial photographs, a survey of the river should be made with 4-band multi-spectral photography (3 visible and one near IR) and a thermal scanner flight should be

Table I: Measurements and Typical Sensors for Water Quality Monitoring and Analytical Modeling

Parameter	Sensors	Requirement
Temperature	Thermocouple, thermistor	Indicates "thermal pollution"; affects other water parameters; measure of internal energy of water for thermal models.
Dissolved Oxygen	Electrochemical cell	Indicates capability of water to sustain life.
Specific Conductivity	Conductivity cell	Variations can indicate change in concentration of various ionized organic or inorganic compounds in water; together with temperature indicate salinity.
Ph	Hydrogen-ion conductivity cell	Provides indirect evidence of various compounds in water, especially those of limited ionizations.
Turbidity	Light source and photoelectric detectors for measuring reflectance or transmittance.	Influences photosynthetic activity in water (oxygen production); input to diffusion models.
Atmospheric pressure	Barometer	Influences amount of evaporation at water surface (heat and mass transfer).

Table I: (Continued)

Parameter	Sensors	Requirement
Incident solar radiation	Phyrheliometer, radiometer.	Influences photosynthetic activity; boundary condition for energy input at surface.
Outgoing radiation	Radiometer	Together with incident radiation provides net radiant energy input at water surface; indicates surface temperature.
Relative humidity	Dewcell (with temperature and pressure)	Influences evaporation at surface (heat and mass transfer).
Current direction and Current velocity	N. A.	Influences convection heat transfer and mass transfer.

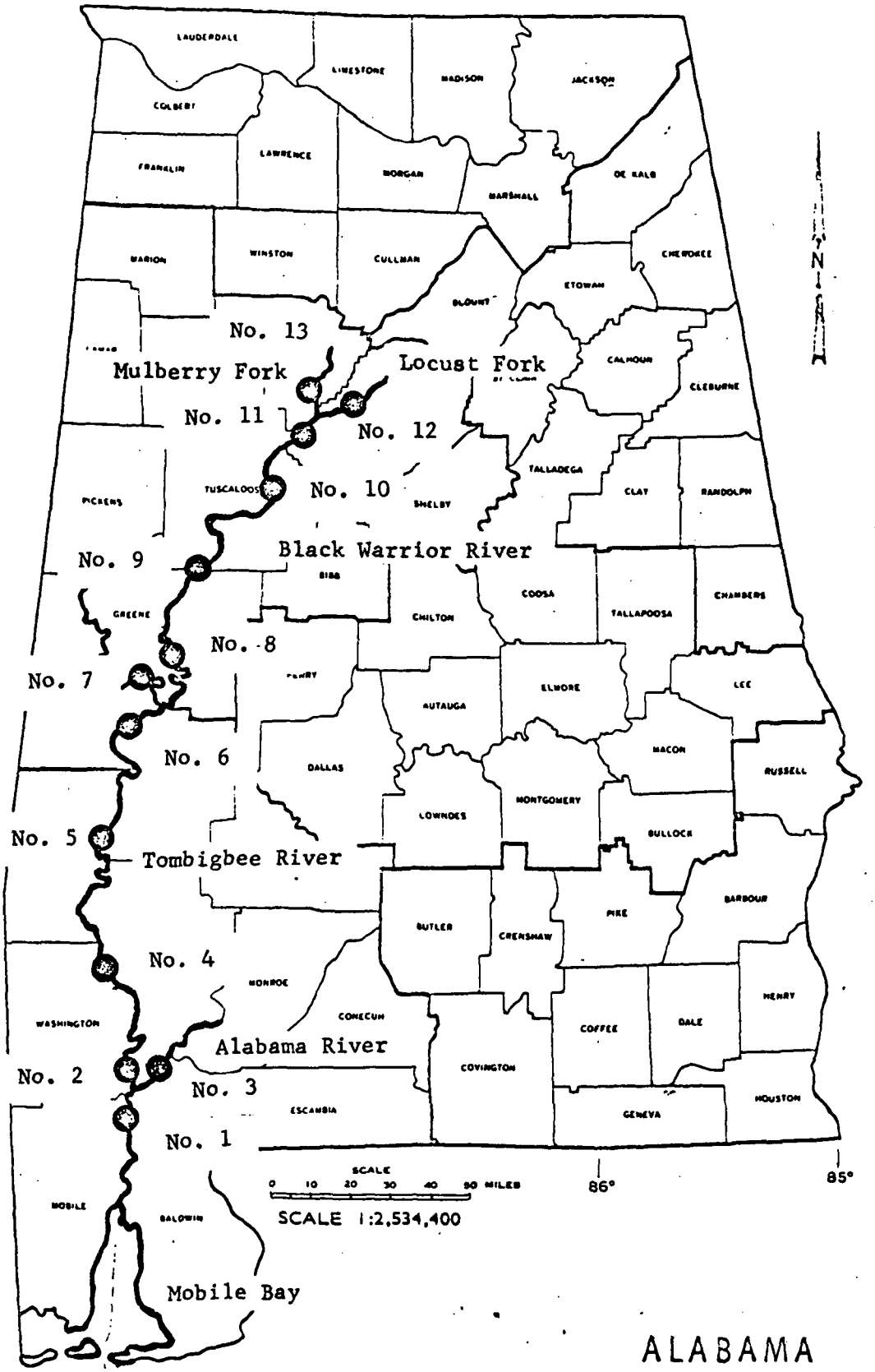


Figure 3: Minimum Distribution for Advanced Buoy System

Table II: Description of Buoy Locations in Figure 3 (8)

Buoy No.	Mile*	Location
1	40	.75 mi. above Tensaw River (Mobile R.)
2	55	Vicinity of Calverts Bar (Tombigbee R.)
3	55	Approximately 10 mi. above confluence of Alabama and Tombigbee (Alabama R.)
4	105	Vicinity of Little McGrews Shoals (Tombigbee R.)
5	155	1.25 mi. above Magnolia Landing (Tombigbee R.)
6	205	.5 mi. below Gulf States Paper Corp. (Tombigbee R.)
7	255	Approximately 38 mi. above confluence of Black Warrior and Tombigbee (Tombigbee R.)
8	255	.5 mi. below Dollarhide Creek (Black Warrior R.)
9	305	Vicinity of Von Cleave Bar (Black Warrior R.)

* Miles above mouth of Mobi River

Table II: (Continued)

Buoy No.	Mile	Location
10	355	.5 mi. above Bone Creek (Black Warrior R.)
11	375	.25 mi. below Little Shoal Creek (Black Warrior R.)
12	390	Vicinity of Day Beacon No. 29 (Mulberry Fork)
13	390	1 mi. below Alabama By-Products Co. Maxine Mine (Locust Fork)

made over the entire river system for each season prior to establishment of an operational data buoy system. Imagery from these surveys should be carefully analyzed and correlated with ground truth if necessary and used to determine if anomalous conditions might exist at specific locations which would warrant buoy placement other than that specified by requirements one through three.

The buoys should be anchored to dedicated pilings or platforms in the river sufficiently stable to hold buoy, power supplies, transmitters, antenna, and warning lights as required. Because the Black Warrior is a navigatable waterway, extensive, continuous three dimensional distributions of the various parameters can not be obtained. Vertical profiles of the various parameters could be obtained but continous horizontal profiling across the river would certainly interfere with river traffic. However, even vertical profiling would require many additional sensors per buoy. Therefore, extensive profiling at a particular point in the river should be accomplished by on-site measurements as required.

Distribution of buoys in Mobile Bay should be determined after a thorough study of the area. The bay area presents a very complex situation due to the tides, size, etc. and also due to the complex river system just north of the bay proper. Therefore, the recommendations in this section of the report apply to the Black Warrior - Tombigbee river system north of a point approximately 5 miles below the confluence of the Tombigbee and Alabama rivers.

Data Collection

With a primary objective of the data buoy being to provide data for water quality monitoring, a near real time (if not real time) capability is required. Also, data should be collected eventually at a central location, therefore the data should be "handled" as little as possible between actual measurements and deposition in the central data facility.

Three alternative means of data collection appear reasonable·

1. Strip chart or tape recorders in the buoy to be read at specific intervals;

2. Hard line data link from buoy to central station;

3. Dedicated telemetry system;

4. Satellite relay, e.g. ERTS/DCS.

The first alternative, although apparently the least expensive, would

be too inefficient and slow with over 13 buoys spread over approximately 350 miles of river (not counting Mobile Bay buoys) and so will receive no further consideration.

The second alternative could be accomplished in two ways. First, a dedicated data link network could be established between each buoy and the central data collection station. This approach would require an extensive network due to the large distances involved, etc. The second approach would be to use telephone line data links to the data collection station. This approach, i.e. data transmission over telephone is quite slow compared to a dedicated hard line data link, so that depending on how many variables are to be measured, frequency of measurement, etc., this approach is likely unfeasible.

Alternative No. 3 should provide for the most rapid access to the data. However, because of relief along the river, numerous ground relay stations might be required in order to get the data into the central station. Of course some combination of alternatives two and three may be the most effective.

The fourth alternative would obviously require a satellite. Relay through the ERTS/DCS system is, as indicated previously, limited to eight separate channels, at a sure data interval of 12 hours. It is anticipated that since a near real time capability will be required, another series spacecraft other than ERTS would be required. Such a spacecraft should be able to relay, during each pass, all data recorded since the previous pass. This would require that a tape recorder or memory be designed as part of a second generation DCP so that several hours of data may be stored and then relayed through the spacecraft. Also, the DCP and spacecraft should be able to handle more than eight channels of data.

Since a second generation spacecraft is not currently available, the approach at this time should be to have a dedicated telemetry system (with hard line data links as necessary) as an integral part of the data buoy system and that consideration be given to satellite relay if and when an appropriate spacecraft becomes available.

Finally, it should be noted that in a detailed analysis of the data buoy itself, the power supply will be of considerable importance. It appears that batteries will be sufficient for power since maintenence intervals will likely be set by the need to service the sensors due to contamination by substances in the water rather than by the batteries. However, with an operational system, it might be more economical and reliable in the long run to use commercial power. This decision of course will be influenced by the location of the buoys.

DATA MANAGEMENT

Efficient utilization of the data acquired by a data buoy system can be realized only if the data is collected and managed properly. This section of the report will consider in general terms a data management system as a part of the total data buoy system.

Central Data Collection Facility

The data measured by each buoy should be relayed, as discussed in the previous section, to a Central Data Collection Facility (CDCF) which will be the central point for storage and retrieval of the data. The CDCF would logically be a part of (or under the control of) the organization(s) responsible for the entire system, e.g. the University of Alabama, the Marine Environmental Sciences Consortium, the Water Resources Research Institute, etc. The CDCF might also be the central facility for other kinds of data, e.g. air pollution data and so forth. However, when the data from the data buoy is received at the CDCF through the data collection network, it would immediately be analyzed automatically by the Water Quality Monitoring System (WQMS). This system would scan the incoming data for anomalous behavior and would announce warnings of these anomalies at the appropriate location. Simultaneous with or immediately after analysis by the WQMS, data will be recorded on high density storage devices which serve as the data base. The WQMS might also store information that it processes in the data base. Data may be retrieved from this data base as required for modeling, special studies, etc. so proper software will be necessary to insure flexibility.

Consideration will need to be given on whether to keep all data as baseline data or whether to periodically purge the data base, retaining only limiting values of the data over periodic intervals. Certainly all data should remain intact for a predetermined length of time before they are purged. Decisions involving elimination of data should be coordinated with all organizations likely to use the data.

Location of the CDCF

No final recommendation can be made at this time for location of the CDCF. From a technical standpoint, there is no real requirement for a particular locale. Possible organizations to house and maintain the CDCF are:

1. Marine Environmental Sciences Consortium;

2. Marshall Space Flight Center;

3. Water Resources Research Institute;

4. State Soil and Water Conservation Committee.

No matter where the CDCF is located, each of the above organizations would need access to the data management system through remote terminals at least in order to obtain status reports on water quality and to retrieve data from the data base. In some organizations, it might be necessary to have an active interaction with the WQMS in order to know in real time water quality at various locations. Each of these organizations, plus others, should have a voice in the management of the data base, especially who may access the data base and, when and if and how data will be purged from the data base. An existing organization which may establish policy in this area is the Alabama Earth Resources Data Committee (9)

In addition, data stored in the data base would be available to proposed earth resources information management systems, e.g. ERISTAR (10), either directly or through existing information analysis centers, e.g. the Marine Environmental Sciences consortium. Figure 4 shows how the various components of the data and information management system fit together, and how it would interface with an ERISTAR system.

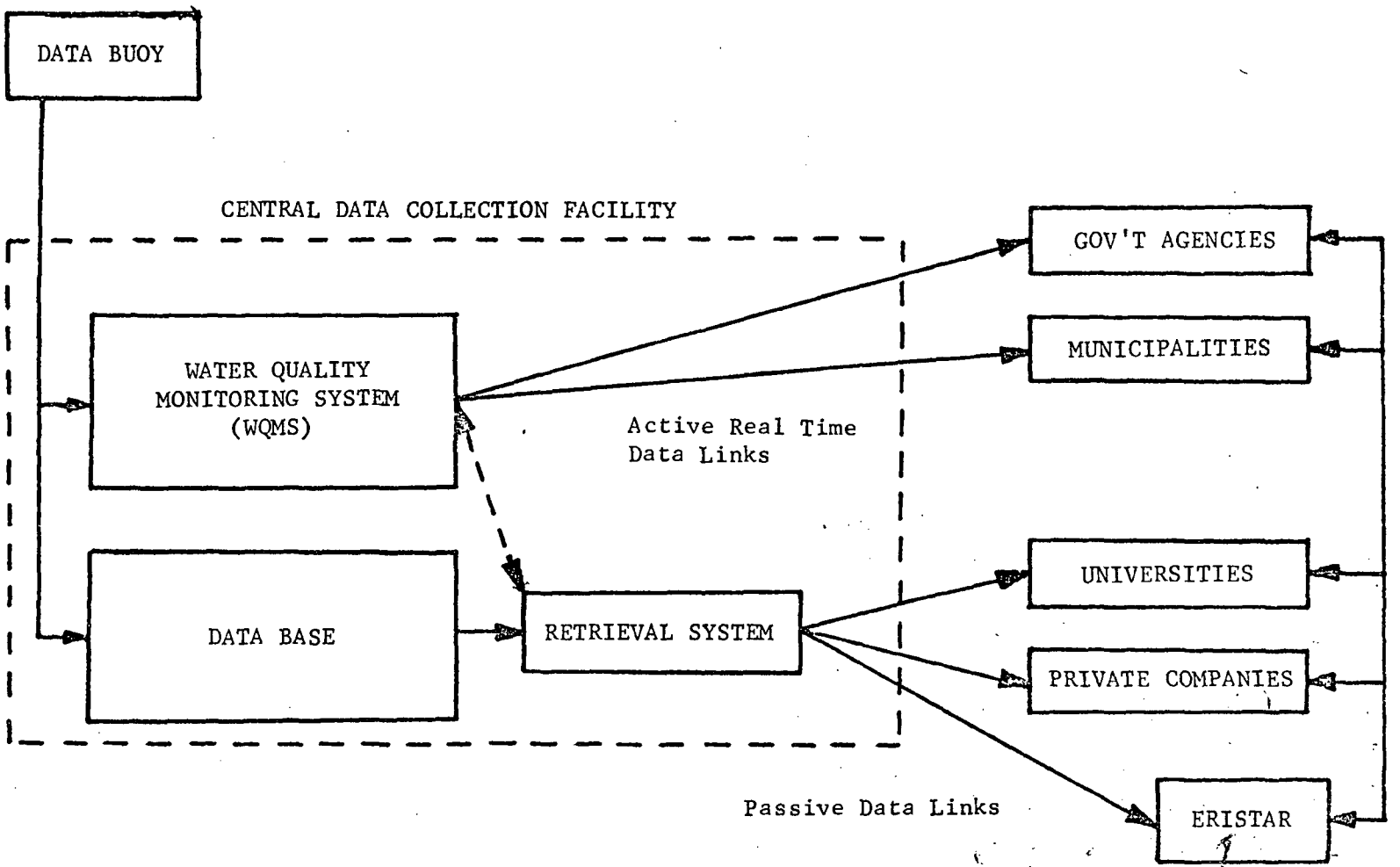


Figure Data Management System for Advanced Buoy System

RECOMMENDATIONS

1. A prototype data buoy system should be established as outlined in the initial sections of this report. This system will rely primarily on ERTS/DCS for data collection, with special provisions made for the Mobile Bay area buoy. Experience with the prototype system will be used in the development of an advanced system.

2. A detailed systems study and design should be completed for an advanced data buoy system with its own data collection system (adaptable for appropriate satellite relay) and data management system as discussed in this report.

3. The capabilities of the Marshall Space Flight Center in instrumentation, communications, and data management should be utilized to the fullest possible extent in the design, implementation, operation, and data management phases of the data buoy program.

REFERENCES

1. Elliott, J. M., "Determination of Land Use by Multi-Spectral Photointerpretation," 1971 NASA-ASEE Summer Faculty Fellowship Research Program Research Reports, MSFC, September 1971.

2. Henry, H. R., "Investigation Using Data in Alabama from Earth Resources Technology Satellite (ERTS-A)," Technical Proposal No. 321, The University of Alabama, Tuscaloosa, April 1971 (Revised: November 1971).

3. Earth Resources Technology Satellite Data Collection Platform Field Installation, Operation and Maintenance Manual, General Electric Co., Space Division, Valley Forge Space Center, March 1972.

4. Private Communication, Rex R. Morton, MSFC, Environmental Applications Office, July 1972.

5. Private Communication, C. Everett Brett, Alabama Marine Environmental Sciences Consortium, July 1972.

6. Parker, B. W., et al., "Automatic System for Monitoring Water Quality," J 1. Sanitary Engineering Division, ASCE Paper 2554, SA 4, Vol. 86, p. 25, July 1960.

7. Tokar, J. V., "Thermal Plumes in Lakes - Compilation of Field Experiments," Argonne National Lab, Center for Environmental Studies, ANL/ES-3, August 1971.

8. Navigation Charts, Black Warrior - Tombigbee Rivers, U. S. Army Engineer District, Mobile, March 1971.

9. Executive Order Number 43, State of Alabama, July 1972.

10. ERISTAR, Earth Resources Information Storage, Transfer, and Retrieval, 1972 NASA-Auburn-ASEE Summer Faculty Fellowship Program in Engineering Systems Design, MSFC, 1972.

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

A LIMITED SURVEY OF GENERAL PURPOSE FINITE ELEMENT COMPUTER PROGRAMS

Prepared by:	Jerome Charles Glaser
Academic Rank:	Assistant Professor
University:	Iowa State University
Laboratory: (Division) (Branch)	Astronautics Laboratory Analytical Mechanics Division Advanced Technology Branch
Research Counterpart:	John E. Key
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

ACKNOWLEDGEMENTS

The writer wishes to express his appreciation to his NASA counterpart, John E. Key, for a very enjoyable summer, stimulating discussions, and helpful advice. Thanks also to all the Analytical Mechanics Division, Astronautics Laboratory personnel for providing a pleasant working atmosphere. Particular thanks to all those of the Advanced Technology Branch who gave freely of their time and experience to provide assistance.

The author would also like to thank J. Fred O'Brien for his efforts as program director and MSFC for its support of the summer program.

A LIMITED SURVEY OF GENERAL PURPOSE FINITE ELEMENT COMPUTER PROGRAMS

By

Jerome Charles Glaser

ABSTRACT

General purpose programs based on the finite element method are considered a basic tool in structural analysis. To be effective, these programs should have "state-of-the-art" capabilities. Thus, there is a need for periodic review of these programs to assess their effectiveness.

Past surveys of finite element computer programs exist which point out general characteristics but lack details on available analysis methods and types of modeling elements. Also, more recent program development has been largely concerned with nonlinear structural analysis.

This report contains a comparison of ten representative programs. A listing of additional programs encountered during the course of this effort is also included. Tables are presented to show the structural analysis, material, load, and modeling element capability for the ten selected programs. These tabular comparisons provide a reasonable picture of the analyses and elements one would expect to find in a current general purpose program.

INTRODUCTION

The last decade has seen the development and implementation of a number of general purpose computer programs for structural analysis (Refs. 1-6). A typical example of these is the NASTRAN program. It was developed on the recommendation of an ad hoc group which had surveyed structural analysis computer capabilities throughout the country and found a wide range of programs but little interchange. This group also specified objectives for the NASTRAN program as follows: 1) be general purpose, 2) be based on the finite element method, 3) be "state-of-the-art" in analytical mechanics, numerical methods, and computer programming, 4) be capable of large 3-dimensional structural analysis, 5) be computer independent, 6) be easy to modify and update, 7) be user convenient, and 8) be well documented. Other general purpose programs have been developed with similar objectives. In general, one could say that the basic objective of these programs is to provide a broad base of computer solution capability in structural analysis for a wide range of users.

As indicated by the above specifications, these programs are to be maintained, as nearly as possible and practical, in a general "state-of-the-art" form. Thus, they are in need of periodic review and comparison to existing programs and current literature. The general objectives of this report are to provide a survey of existing finite element programs, a comparison of their characteristics with current literature, and to outline what features should appear in a "current" general purpose program.

These objectives represent a formidable task and can be considered analogous to a "population explosion." In this report, these objectives are limited to:

1) information from readily available sources, subject to the author's interpretation,

2) a list of finite element programs encountered,

3) a comparison of ten representative programs in terms of analysis capability, material characteristics, structural loading, and modeling elements, and

4) current literature comparisons as time permits.

FINITE ELEMENT PROGRAMS

Gallagher (1) provided an interesting overview of large-scale, general purpose programs in terms of program identification, characteristics, and future trends. He characterized two basic types;

those with many elements, but limited in analyses, and those with a large analysis range, but limited modeling capability. One could add a third type; those that are in wide usage. Hibbitt, et al, (3), and Marcal (4) have presented discussions concerning general purpose program requirements for solution of nonlinear problems. The major addition proposed here is the requirement of a capability for dealing with different material constitutive equations and numerical analysis schemes to handle both nonlinear material and geometric problems. The survey by Hartung (5), although limited to shell structures and not to finite element programs, supplies a valuable source of information about available computer programs. More recently, Perrone (6) compiled a description of approximately 400 computer programs used in structural mechanics; however, most can be classed as special purpose programs.

Marcal (4) points out an interesting distinction between programs which has a bearing on the question of machine independence. He separates programs in terms of program language used, either some form of matrix interpretive language or the familar FORTRAN language. He concludes the latter is better from a user standpoint. One can also draw a distinction based on whether a program is intended to be a basic research tool or a practical analysis tool. FORTRAN allows immediate use of the program and easy revision as needed in basic research. On the other hand, it would seem to allow generation of a multitude of special versions, contrary to other objectives. These distinctions in program form are related to the question of allowing user coded subroutines or having only rigid format capability in a program. A sample of each program type is included in the tabular comparisons which follow in this report.

Only programs based on the finite element method are considered in this report. This method has been the subject of numerous publications. The author has encountered several thousand references via various bibliographies, e.g., Akin, et al (83). The results of earlier published work in this area have been solidified in recent books. One can obtain a good feel for the applications capability of this method by scanning several of these books, e.g., Zienkiewicz (7), Przemieniecki (8), Oden (9), and Desai and Abel (10). Zienkiewicz is strongly oriented toward presentation of element formulation methods, particularly isoparametric elements but includes general considerations of dynamic and nonlinear problems. Przemieniecke provides the most detail on dynamic applications while Oden covers nonlinear applications. Desai and Abel present a very readable account of the method as well as a good description of its many uses and is recommended as a starting point. A cursory review of these books provides a reasonable idea of what one might expect to find in a general purpose program based on the finite element method.

A large number of finite element computer programs were noted during the course of this survey. Table 1 provides a list of these codes. It should not be considered a complete listing of programs

in current use. In this table, references for each code are shown and, when known, its origin, source of development funds, and whether the program is available (A) or proprietary (P). If no information was readily available concerning a code, reference is made to the document which referred to it. Also, if unnamed, the program is denoted by the developer's last name. Not all references indicated were reviewed but were listed to provide as complete a documentation list as possible for further study.

STRUCTURAL ANALYSIS CAPABILITY

Tabular comparisons of ten programs are presented in this section to provide an overview of their capability for structural analysis. These programs (NASTRAN, MAGIC, MARC2, ASKA, FORMAT, ASTRA, STARDYNE, ICES/STRUDL, ELAS, and SABOR/DRASTIC) are a representative sample of general purpose programs. Some of these programs are surely more versatile than this report will show. The author apologizes for this; however, it is felt that the information shown will be helpful in the assessment of what should constitute the overall capability of a general purpose program.

Table 2* presents a breakdown of the general analysis capability, program language, and number of modeling elements available in these programs. It can be seen that the capability for heat transfer analysis is only beginning to be exploited. The structural optimization category shown is interpreted here to mean self-contained routines as opposed to interface with other programs. Optimization routines have not received much attention with regard to inclusion into these general purpose programs. The utility of finite element methods in this area is well known and its development has been recently promoted, e.g., by Hartung (5) and Zienkiewicz (84). The additional categories of analysis shown are expanded upon in Tables 3 through 7.

Linear static analysis, Table 3, is expanded in terms of structural geometry. These general shapes appear to be adequately covered in most cases although coverage of thick-walled bodies of revolution and general 3-dimensional shapes is lacking in several cases. The capability of each program to solve a problem in any one of these areas is undoubtedly enhanced by the structural modeling experience of the user.

*The comparison tables in this report should be interpreted as follows:

1) the reference which demonstrated or strongly implied a particular capability, etc., is indicated by number as listed at the end of this report,
2) a blank space indicates unknown capability, etc.,
3) NONE indicates a capability, etc., known not to exist,
4) DEV indicates a capability, etc., proposed or under development,
5) lower case letters refer to footnotes.

In general, the terminology used corresponds to NASTRAN documentation.

Table 4 shows a breakdown of stability and eigenvalue analyses. Most programs show a capability for primary buckling analysis. It appears that local instability, as opposed to stability of a complete structure, is not considered. This is probably accomplished by re-analysis of components in most cases. Local instability solutions have recently been discussed by Przemieniecki (89). Higher-order buckling methods have not been greatly used. Pre- and post-buckling as shown here is a nonlinear problem and has been included only in the MARC2 and ASKA programs. Some recent reference works of interest are indicated for this area (88, 91, and 105).

Most of these programs also have the capability to solve the eigenvalue problem for the natural modes (free vibration deformation patterns and frequencies) of a structure. This is an important requirement for dynamic response analysis. Most use modal analysis where eigenvalues are extracted from matrices using the vibration modes as degrees of freedom rather than direct methods which use grid point displacements as degrees of freedom. The eigenvalues are extracted using either transformation or tracking methods (16). The former is an algebraic method generally used when a large number of roots are required, e.g., in a vibration problem. The latter is an iterative method generally used in buckling problems to obtain a small number of roots. Several programs include both types for computational efficiency in particular solutions. Only limited information was readily available about these program characteristics.

Linear dynamic analysis, Table 5, is outlined as in NASTRAN documents (11 and 16). Free vibration refers to the normal mode eigenvalue problem of Table 4. It represents the solution for natural frequencies or steady-state response to selected initial conditions of an ideal structure, having no damping or external forces. Most programs have the capability to determine the free vibration characteristics of a structure. McNeal (16) outlines the two solution methods, modal integration and direct integration, used in dynamic response analysis. He indicates that it is important to have both methods of formulation to allow for computational efficiency in different problems. He further indicates that the modal method is efficient when a small number of modes is sufficient to characterize a solution while the direct method is better when a large number of modes is desired. Hartung (5) provides further discussion about the advantages and disadvantages of these two methods; an important point made is that modal methods are not applicable to nonlinear problems. In both methods, finite difference numerical integration of the differential equations is used. In the modal method, the equations are integrated in terms of modal coordinates. This makes use of the fact that the equations are uncoupled when the displacements are formulated in terms of natural modes. The equation integration is done in terms of grid point displacements in the direct method. The Newmark Beta method is the integration algorithm most frequently mentioned. Because of integration simplicity, the modal method is the predominant one appearing in the selected programs.

Using either method, one can determine frequency response - the response of a structure to a known frequency, random response - structural response to a random forcing function, and transient response - response to a time-dependent forcing function. Strongest analysis capability is evident for frequency response using the modal method. Few of these programs allow analyses using the direct integration method; even the NASTRAN users' experiences document (15) provides little to show usage of the existing NASTRAN capabilities.

Table 6, Failure Analysis, has received limited attention although many published works on these topics have appeared recently. Only MARC2 and ASKA appear to provide much capability in these areas. These topics can also be interpreted in terms of program ability to analyze problems with material nonlinearity. Some of the more recent works in these areas, noted during this survey, are shown.

Nonlinear analysis capability is shown in Table 7. The NASTRAN capabilities provide, as noted, only approximate static solution methods for material and geometric nonlinear problems and allow for dynamic analysis of response to nonlinear transient loads. ASTRA and ICES/STRUDL provide for static solutions with geometric nonlinearity. Only MARC2 and ASKA appear here to be capable of analyzing static nonlinear problems in terms of geometry, materials, and buckling. The retention of nonlinear terms in the governing equations for these problems forces the use of iteration procedures. This may be a prohibitive cost factor and could prevent incorporation of nonlinear capability in some cases. It was not determined if ASTRA or ICES use iterative techniques. MARC2 and ASKA apply iteration methods and allow for the inclusion of various material constitutive relations. Both of these programs are research programs and were designed for the analysis of nonlinear problems while the other programs were not. Each can be packaged via subroutines to solve particular nonlinear problems. The computer costs would thus be less than if these programs were in the rigid format form of most general purpose programs.

Nonlinear geometric and material effects in dynamic analysis are not strongly indicated capabilities of any of these programs. Also, it is interesting to note that it has been proposed (15, 16) that aeroelastic solution capability be considered for NASTRAN. These analyses (flutter, etc.) are usually accomplished by use of auxiliary programs. In general, it is apparent that recently developed nonlinear analysis capability has not been included in the readily available general purpose programs.

Nonlinear analysis has been the subject of many recent articles; several of these works are shown in Table 7. Two of these appear to

be of immediate interest. Stricklen (90) provides an assessment of the solution procedures available for inelastic and/or large deflection structural behavior. This work includes a literature survey of material nonlinear and combined geometric-material nonlinear problems and a comparison of Newton-Raphson, incremental stiffness and self-correcting solution methods. Haisler (93) presents a comparative study of the available solution techniques which are applicable to the solution of the nonlinear algebraic or differential equations characterizing geometrically nonlinear behavior. He promotes a self-correcting initial-value form of solution as a possibility for inclusion into a general purpose program.

Table 8, Nonstructural Analysis, has been added to show other areas of application of the finite element method noted during the course of this survey. Table 9 shows the material effects that have been considered while Table 10 shows the structural loads mentioned in program literature. Both of these are directly related to a particular type of element and vary from element to element. No distinction has been made here as to element type and these tables should be interpreted only in general terms. It is apparent again, from Table 9, that nonlinear material analyses are not possible with most of these programs.

MODELING ELEMENT CAPABILITY

A large number of finite element models have been presented and surveyed in the published literature, and it is, perhaps, wise to consider some general background initially. Variational principles in energy form are generally considered to be the basis for finite element formulations. The minimum potential energy principle is the basis for the displacement method and is the most commonly used approach. In this case, generalized displacements are taken to be unknowns. In addition, there is the equilibrium method based on minimum complementary energy. Here, stress parameters are used as general unknowns. Also, there exist hybrid displacement, hybrid equilibrium, and mixed element formulation methods which use modifications and combinations of the two basic methods. More detailed information on these topics can be found, for example, in Desai and Abel (10).

There have been many new elements and applications presented in recent literature for each type of formulation. From these, several interesting general comments about future trends in element forms were noted during this survey. There has been a trend toward increasingly complex element forms. Zienkiewicz (84) has cited two economic reasons for this. Higher-order elements allow use of fewer elements to model a structure which lowers input data cost. Also, the required solution accuracy is maintained even though a decrease

in the total number of degrees of freedom occurs; this also allows a reduction in cost.

Gallagher (1) notes that equilibrium formulated elements have advantages in certain problems. He indicates that the minimum complementary energy method appears to be the most efficient procedure for cases where stresses are the important parameters and where the constitutive equations are awkward. Mallett (22) also indicates a trend toward more hybrid and mixed element formulations.

Mallett (22) also provides some thoughts on the problems of solid elements. He states that 3-d analysis should be restricted to rather modest problems or those of great import because of the large number of degrees of freedom required to describe 3-d elements. In his view, the most promising approach to the analysis of solids will be that of using sophisticated isoparametric element forms.

In terms of general purpose programs, Mallett (22) says that progress toward improved stress analysis capability will be realized simply by expansion of element libraries so that elements need not be used under circumstances where they are not really applicable. Thus, as Gallagher (1) states, it is essential that any general purpose program be able to accommodate, with ease, new elements as these appear in the literature.

Element modeling capability is obviously a very important part of a general purpose program. In this section, tabular listings of the finite element models available in the ten selected programs are presented. These tables follow the same format as in the previous section. In all these tables, the elements indicated are direct stiffness displacement models based on the minimum potential energy formulation.

Comparison tables are provided for scalar and 1-d, plate, shell, solid of revolution, and solid elements. In preparing these tables it was quite difficult, from available literature, to determine the actual element capability. Specific information was not available in many cases. A specific problem occurs when one tries to determine the special forms that general elements can assume. Footnotes are provided where this is thought to be a likely possibility. An additional problem is the fact that MARC2 and ASKA make extensive use of isoparametric elements whose specific form is not generally indicated. It was decided to use the actual and proposed element capability of NASTRAN as a comparison base, and details on these elements may be found in that documentation. Exceptions are the isoparametric and solid element capabilities.

It is seen immediately that the SABOR/DRASTIC program is limited to a doubly curved shell element and is not general purpose in this sense. Table 11 shows that beam elements are standard; however, curved and nonprismatic beam elements are not available. There are numerous forms of plate elements available as shown in Table 12. Most programs include triangular and quadrilateral element capability. Layered or composite elements have not received much attention. Only MARC2, ASKA, and ICES appear to contain capability for isoparametric plate elements. Table 13 shows that most of these programs have some provision for the analysis of shells.

Table 14 shows that a number of programs contain a triangular ring solid of revolution element, and a few include a trapezoidal or quadrilateral ring element. Most of these elements are limited to axisymmetric load and deformation criteria. Little provision for nonaxisymmetric conditions is available. Several programs have solid elements, as seen in Table 15; NASTRAN and MAGIC are notable exceptions, having shown this capability only in developmental versions.

Table 16 has been included to show special element capabilities noted during this survey. It mainly shows developmental elements for NASTRAN and indicates similar capability existing in the other programs. Of note here is the dummy element which will allow element research capability to be added to NASTRAN such as is available in MARC2. Also of interest is the fact that only a few of these programs use numerical integration techniques to form the element stiffness matrix. It was noted in several articles (1, 22) that this will be a major consideration in future work.

GENERAL OBSERVATIONS AND RECOMMENDATIONS

The basic objective of this effort has been to determine what capabilities should be available in a current general purpose structural analysis program. This survey has been limited to a study of available structural analyses, material considerations, and types of modeling elements. The tables presented in this report provide a reasonable overall picture of current capability in these areas and some general conclusions can be drawn from these tables.

Linear static and dynamic analysis capabilities as currently available in NASTRAN are definite requirements. It also appears that a current program should be able to handle material, geometric, and combined material-geometric nonlinear problems in the static case. The capabilities in these areas appear to be developed to such an extent that they should be available to a user. Nonlinear material and geometric problems in the dynamics area are not well-developed and cannot be considered at this point.

Eigenvalue and instability analysis capability should include the nonlinear capability of pre- and post- buckling analysis. This also appears to be sufficiently developed to be included in a general purpose program. A current program should also be able to handle various material constitutive equations and allow for consideration of plastic, nonlinear elastic, and creep behavior. The analysis capability which would perhaps best describe the state-of-the-art required for a current general purpose program would be that of a combined NASTRAN-MARC2 program. Whether all this capability can be practically incorporated into a single program such as NASTRAN is an interesting question; one which this author certainly cannot answer.

It is this author's opinion that the element capability available in and proposed for NASTRAN provides the most complete individual element library of any existing program. However, many of these elements seem to provide overlapping or redundant capability which leads to confusion. In this regard, it appears to this writer that isoparametric element formulations should be provided in any general purpose program because of their generality and versatility in describing families of elements.

As noted in an earlier section, there are an increasing number of elements appearing in the literature which are not based on the displacement method. A capability to handle and test these elements should be considered. Several papers have promoted this idea; for example, a paper by Robinson and Haggenmacher (98). They introduce a concept of characteristic matrices and an analogy which permits complete interchangeability of different types of elements between independently developed programs irrespective of the analysis method or basic element assumption. This is an interesting idea which may be useful in further generalizing the element capability of any program. It should be pursued to determine its usefulness for general purpose programs.

A general purpose program must be able to solve practical problems and allow for research and development work. It appears that both rigid format and user coded subroutine capability should be available. The addition of the dummy element to NASTRAN is an important step in this regard. This should allow for more research and testing of new elements. The requirement to learn and use new matrix languages (DMAP) rather than FØRTRAN is probably a bad feature to most working engineers, and this element will help in this respect. This will also increase the ability of the program to be a useful tool in nonstructural areas.

Any current program should also provide capability for the numerical integration of the element stiffness matrix since this seems, from several sources, to be of increasing importance.

Additionally, new structural optimization techniques and applications may provide methods which can be incorporated directly in a program and should be considered.

It is evident that a large problem exists in the area of program checkout. Someone should be given or take the responsibility of defining a set of practical problems which are well-documented in terms of experimental and analytical results. This problem set could serve as a minimum standard capability test that any potential user would have available for use with any program. These problems should be published in a form which is readily available. Along these lines, as new elements and applications are published, the specific problems in these articles should also be solved, if possible, using existing general purpose programs. This would help to determine if the work in recent publications represents a significant improvement over capability which already exists.

Many of the objectives of a general purpose program which were indicated earlier in this report have not been considered here. It is felt, however, that the basic objective has been satisfied within the stated limits.

Table 1 - Finite Element Computer Codes

Code Name	Developer/Funder	Status	References
1. NASTRAN	Mac-Neal-Schwendler/NASA-Langley	A	11-17
2. MAGIC	Bell Aerospace/AFFDL	A	18-22,80,82
3. MARC2	P. V. Marcal, Brown Univ./NSRDC	P	3,4,6,23,24
4. ASKA	H. Argyris, Univ. of Stuttgart	P	25-36,87,15
5. FØRMAT III	McDonnell-Douglas/AFFDL	A	37-46,15
6. ASTRA	Boeing Company/Boeing	P	6,47,48
7. STARDYNE	Mechanics Research, Inc./MRI	A	49
8. ICES/ STRUDL	MIT/IBM	A	50-53,79,84
9. ELAS	Jet Propulsion Co./JPL	A	54,55
10. SABOR/ DRASTIC	MIT/SAMSO	A	56-63
11. SAMIS	PhilcoFord/JPL	A	64,65
12. FESS	O. Zienkiewicz, Univ. of Wales, Swansea	P	7,66,67
13. EASE	Engineering Analysis/Engineering Analysis	A	68
14. MINI-ASKA	H. Kamel, Univ. of Arizona/LMSC	P	69
15. REXBAT 5	LMSC/LMSC	P	6,70
16. SNASOR	Texas A&M/NASA-MSC & Sandia	A	6,71
17. SNAP	Lockheed-Huntsville/MSFC	P	72,15
18. SAP	E.L. Wilson, Univ. of Calif. Berkeley	A	73
19. SLADE	Sandia Corp./Sandia Corp.	P	6,74,85,15
20. WASP	LMSC/LMSC	P	75
21. DAISY	H. Kamel, Univ. of Arizona/Lockheed Aircraft Co.	P	2,92
22. DYNA	H. Kamel, Univ. of Arizona	P	92
23. SAFE	Gulf Gen. Atomic, Inc.	A	6,10
24. FINESSE	O. Zienkiewicz, Univ. of Wales, Swansea	P	7
25. CØSMIC	Boeing Co./Boeing Co.	P	2,47
26. SAMECS	Boeing Co./Boeing Co.	P	2,78,86
27. DeVuebecke	F. DeVuebecke, Univ. of Liege	P	2
28. DYNES	Convair/San Diego	?	2
29. Gallagher	R.H. Gallagher, Cornell Univ.	?	2,88
30. DEMON	Douglas Aircraft Co.	?	2
31. MASS	General Electric Co.	?	2
32. ASTRAL	Grumman	?	2
33. FAMAS	Lockheed Aircraft Co.	?	2
34. SB039	Martin Co.	?	2,81

Table 1 - Finite Element Computer Codes (Continued)

Code Name	Developer/Funder	Status	References
35. Pian	T. H. H. Pian, MIT	?	2
36. STACUSS 1	MIT/SAMSO	A	5,77
37. Stephens/ Fulton	R. E. Fulton/NASA Langley	A	5
38. CRAB	TRW Systems, Redondo Beach, Calif.	?	6
39. Lindberg	G. M. Lindberg, NRC of Canada	A	6, 76
40. AS145A	TRW Systems, Redondo Beach, Calif.	?	6
41. FINPLA	A. Scordelis, Univ. of Calif., Berkeley	?	6
42. E-11101	L. Herrmann, Aerojet-Gen. Corp. Sacramento, Calif.	?	6
43. FEATS	Westinghouse Electric Corp. Pittsburgh, Pa.	?	6
44. SA 005	G.C. Pfaff, Martin-Marietta Corp., Denver, Colo.	?	6
45. Bader	R.M. Bader, AFFDL	A	6
46. EPAD	Universal Analytics, Inc., Los Angeles, Calif.	?	6
47. BATMAN	LMSC	?	6
48. GENESYS	British	?	84

Table 2 - Overall Program Characteristics

Analysis \ PROGRAM	NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Linear Static	11, 17, 16	1, 18, 21	3	25, 33	1, 46, 15	6, 47	1, 49	1, 10, 79	1, 10, 5	1, 5, 6	
Eigenvalue, Stability	11, 17, 16	1, 18, 21	3, 6	1, 5	1, 5, 46	1, 6, 47	5, 49	84	None	5	
Linear Dynamic	11, 17, 16	21	5	1, 33, 5	46, 15	6, 47, 5	1, 5, 49	1, 79, 84	None	5	
Failure	11, 14	None	3, 4	25, 87	None		None		None	None	
Nonlinear Static	11, 17, 16	None	1, 3, 4, 6	1, 5, 25	None	1, 6, 47	None	1, 79, 84	None	None	
Nonlinear Dynamic [a]	None	None	5		None	6, 47	None	None	None	None	
Heat Transfer	15 DEV	18, 22		87	None		None		None	None	
Structural Optimization	None	None	None	None	None	[c]	None	84 [d]	None	None	
Number of Modeling Elements	Large	10	User Supply	Large [b]	4	4	8	12	18	2	
Program Language	*Matrix* Inter.	*Matrix* Inter.	*Fortran* IV	*Matrix* Inter.	*Matrix* Inter.	*Fortran* IV		*Matrix* Inter.	*Fortran* II	*Fortran* IV	

a) does not refer to nonlinear (transient) loads
b) includes equilibrium and hybrid elements
c) capability by coupling to other programs
d) frames and bridges

Table 3 - Linear Static Analysis

Analysis / Linear Static	PROGRAM: NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Beam	11, 17	18, 21 22	5, 6	15	46, 41, 15	6, 47	49	84	10, 5, 54	None	
Plate	11, 17	18, 21 22	4	25, 15	46, 41, 15	6	49	10, 79 84	10, 5, 54	None	
Thin Shell	11, 17	18, 21 22	3, 4	25	5	6		10, 79 84	10, 5, 54	1, 5, 6	
Thick Shell	11, 17	18, 21 22			None	6	None		54	None	
3-d Solid	DEV	22	3, 4	25	None	47	49	10, 84	10, 54	None	
Inertial Relief a	17, 11, 14										

a) Converts steady-state dynamics to an instantaneous static solution, other programs may allow this thru user input.

Table 4 - Eigenvalue and Stability Analysis

Analysis / Eigenvalue, Stability	NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Primary Buckling	11, 14, 17	18, 21		25, 35	5, 46	5, 6, 47	5	84	None	None	
Local Buckling	None									None	89
Pre- and Post-Buckling	None	None	6, 4	5, 25	None		None			None	88, 91, 105
Natural or Normal Mode	11, 17, 15	18, 21		5	5, 46, 15	5, 6, 47	5, 49	84		5	
Direct Eigen. Analysis	11, 17, 16 a	None		None	None	None	None			None	
Modal Eigen. Analysis	11, 17, 16 a	21		5	5, 15	6, 47	5, 49			5	
Transformation Extraction	11, 17, 16					47	5				
Tracking Extraction	11, 17, 16		6			5, 47	5		None		

a) includes complex eigenvalues

Table 5 - Linear Dynamic Analysis

Analysis / Linear Dynamic	PROGRAM: NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Free Vibration (Normal Mode)	11, 17, 15	21, 22		5	46, 5	6, 47	5, 49	84	None	5	95, 96, 97
Direct Frequency Response	11, 17, 16	None			None	None	None			None	
Direct Random Response	11, 17, 16	None			None	None	None			None	
Direct Transient Response	11, 17, 16	None	5		None	None	None			None	
Modal Frequency Response	11, 17, 15	21		5	5, 46	5, 6, 47	5, 49, b			5	97
Modal Random Response	11, 17, 15	None				6, 47 a	None				
Modal Transient Response	11, 17, 16	None				47 a	49 b		None	5	85

a) if coupled to other programs
b) uses auxiliary program, DYNRE

Table 6 - Failure Analysis

Analysis / Failure	NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Yield [a]	11		3, 4	25, 87							87, 90, 91
Fracture	None	None			None	None	None		None	None	103, 104
Crack	14		3	25							104
Creep	None		3, 4	25, 87							87
Fatigue	None	None			None	None	None		None	None	

a) Most programs have a margin of safety feature, while several have definite yield criteria.

Table 7 - Nonlinear Analysis

Analysis / Nonlinear	PROGRAM: NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Static	11, 15, 17 b,c	None	3, 4, 6	1, 87, 25	None	1, 6, 47	None	1, 79, 84	None	None	
Dynamic	11, 15, 16 a		5			47 a		None			99, 100, 106
Large Displacement	11, 17, 15 b		1, 3, 4, 6	1, 5, 25, 32		1, 47		1, 79, 84			88, 90, 91, 93
Material	11, 14, 17 c		3, 4, 6	25, 87		None					90, 91, 101, 102
Buckling	None		4, 6	5, 25		None					88, 89, 91, 105
Flutter	15, 16 DEV		None	None		6, 47 d		None			
Acoustic Response	15	None			None		None	None	None	None	94

a) non-linear transient loads only

b) static analysis, differential stiffness, first order approximation

c) piecewise linear approximate static solution

d) coupled with other programs

Table 8 - Nonstructural Analysis

Non-structural \ Analysis / PROGRAM	NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Heat Transfer	15 DEV		87								
Fluid Flow						48					
Soils								84			
Control System Simulation	14, 17, 16					6					
Electrical Circuit	14, 16										
Acoustics	15										
Hydraulics	15							84			

Table 9 - Material Conditions

Materials [a] \ PROGRAM	NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Elastic Isotropic	11, 14, 17	18, 21	4	5	41	1	49	5	5, 54	5	
Elastic Anisotropic	11, 14, 17	18, 21	3, 4	25, 35							
Elastic Orthotropic	11, 14, 17	18, 21		15		6, 47			54	6	
Temperature Effects	11, 14, 17	18, 21	3, 4	25, 87			49		54		
Experimental Data	11	18, 21							54		
Plastic	11	None	3, 4	1, 2, 87	None	None	None		None	None	87, 90, 91 101, 102
Nonlinear Elastic	None			25, 87		6					
Inelastic	None		3, 1, 5			None					90
Creep	None		3, 4	25, 87		None					87
Viscoelastic	None	None	None		None	None	None	None	None	None	

a) no attempt was made to correlate to elements

Table 10 - Structural Loads

Loads [a] \ PROGRAM	NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Static	11, 16	18, 21	3	25	41, 15	6, 47	49		54	5	
Dynamic	11, 16			25		5, 47	49				
Concentrated	11, 14 17	18, 21	3		41, 15	47	49		54		
Distributed						47	49		54		
Temperature	11, 14 17	18, 21		25, 87	41, 15	6, 47	49		54	5	
Misfit	11, 14 17	18, 21			41		49		54		
Pressure	11, 14 17	18, 21	3			47			54		
Inertial	11, 14 17	21					49				
Initial					41					5	
Impact											

a) no attempt was made to correlate to elements

Table 10 - Structural Loads (Continued)

Loads	PROGRAM NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Surface	11, 14 17	21	3				49		54		
Body	11, 14 17	21					49		54		
Steady	17										
Unsteady											
Axisymmetric	17	18, 21	3			47			54	5	
Asymmetric	17 DEV					47				5	
Aero Loads						47					

Table 11 - Scalar and One-Dimensional Elements

Elements / Scalar & 1-d	PROGRAM NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Viscous Damper	11, 17	None			None				None	None	
Spring	11, 17	None			None		49		None		
Mass Damper	11, 17	None			None				None		
Extension Beam	11, 17	18, 21	5, 6		15		a	a	54		
Bending Beam	11, 17	a	5		15		a	a	54		
Torsion Tube	11, 17	a?	a?		a?		a?	a?	54		
Non-Prismatic Beam	15 DEV	None			None	47	None		None		
Curved Beam	15 DEV	None			None	47	None		None		95, 96
Extension, Bending Torsion Beam	11, 17	18, 21, 22	5, 6, 4		46, 41	6, 47	49	84	10, 54		
Isoparametric	None	None			None		None	84	None	None	

a) specialized case of other element

Table 12 - Two-Dimensional Plate Elements

Elements / 2-d, Plate	PROGRAM: NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Membrane Triangle	11, 17	18, 21		a	41, 46 15	a		10	54	None	
Bending Triangle	11, 17	18, 21		15	46	a			54		
Membrane-Bending Triangle	11, 17	18, 21		15		6, 47			10, 54		
Mid-Side Node Triangle	DEV	18, 21		a	None						
Higher-Order Triangle	15 DEV	None		33, 25 32	None						
Shear Panel	11, 17	18, 21, 22		a?	40, 46 15		49				
Twist Panel	11, 17	None		a?	None						
Membrane Rectangle	a	a		a?	41, 46			a	a		
Bending Rectangle	a	a		a?	a			10	a		
Membrane-bending Rectangle	a	a		a?				10	a	None	

a) specialized case of other element

Table 12 - Two-Dimensional Plate Elements (Continued)

Elements / 2-d, Plate	PROGRAM: NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Membrane Quadrilateral	11, 17	18, 21		15	15				54	None	
Bending Quadrilateral	11, 17	18, 21		a?	46				54		
Membrane-Bending Quadrilateral	11, 17	18, 21		a?		6, 47			10,54		
Mid-Side Node Quadrilateral	DEV	18, 21		a?	None						
Higher-Order Quadrilateral	15 DEV	None		87	None						
Sandwich, Composite	11, 17	22		15	None		49				101, 102
Multi-Layered Triangle	11, 15, DEV	a		a?	None						
Multi-Layered Rectangle	11, 15 DEV	a		a?	None						
Plate	11, 17	18, 21	3, 6, 4	25, 33, 34, 35	41, 46	6, 47	49	84	54		
Isoparametric	None	None		25, 33, 34, 35	None		None	84	None	None	84, 97

a) specialized case of other element

Table 13 - Shell Elements

Elements / Shell	PROGRAM: NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Conical	11, 14, 17	18, 21				47	None		10, 54	None	
Triangle	15 DEV	18, 21							54	None	
Quadrilateral	15 DEV	18, 21							54		85
Axisymmetric Load, Shell of Revolution	11, 14, 17	18, 21	5, 3, 6				None		54	1, 5, 6	91, 93
Asymmetric Load Shell of Revolution	11, 15						None		None	None	93
Doubly Curved	11, 14, 17 a	18, 21 a	5, 4			47	None		None	6	85, 91
Higher-Order Shell of Rev.	15 DEV						None		None	None	
Isoparametric	None	None		25, 36			None	84	None	None	84
Shell	11,17	18, 21	3,4	25, 36	5	47		10, 84	54	5	

a) toroidal, specializes to cone, cylinder, and shell cap element
b) specialized case of other element

Table 14 - Solid of Revolution Elements

Elements / Solid of Revolution	PROGRAM: NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Axisym. Triangular Ring	11, 17 a	18, 21 22	3, 4		None	6, 47	None		10, 54	None	
Axisym. Trapezoidal Ring	11, 17 a	21, 22 a	None						b		
Axisym. Quadrilateral Ring	b	b	None						10, 54		
Axisym. Rectangular Ring	b	22	None						b		
Axisym. Ring Stiffener	11, 17										
Asym. Load Triangular Ring	15 DEV	None							None		
Asym. Load Quadrilateral Ring	15 DEV	None	None						None		
Asym. Load Ring Stiffener	15 DEV	None	None		None		None		None	None	

a) specializes to core elements
b) specialized case of other element

Table 15 - Three-Dimensional Solid Elements

Elements / 3-d Solid	PROGRAM: NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Isoparametric	None		3, 4	25	None		None		None	None	84
Tetrahedron	16 DEV	22 DEV	6	25, 31, 32		47	49		10, 54		
Hexahedron	16 DEV			25, 29, 30			None		10, 54		
Cube	None		3, 4	a			49		a		
3-d Solid	16 DEV	22 DEV	4	25		47	49	10, 84	54		
Rectangular Prism	None	22 DEV		a			None		54 a		
Triangular Prism	None	22 DEV		a	None		None		a	None	

a) specialized case of other element

Table 16 - Special Elements

Special Elements / PROGRAM	NASTRAN	MAGIC	MARC2	ASKA	FORMAT	ASTRA	STARDYNE	ICES/STRUDL	ELAS	SABOR/DRASTIC	RECENT REFERENCES
Dummy [a]	15 DEV		3, 4								
General [b]	15, 14, 17 DEV										
Rigid Body	15 DEV						49				
Singularity [c]	None										
Hydroelastic [d]	15										
Heat Transfer	15 DEV	18, 22									
Multi-Layered	15 DEV									1, 5, 6	96
Num. Integ. of Elem. Stiffness	None	18, 21	3, 4. 6		None	None	None	None	None	5	

a) user supplied mass, damping, and stiffness matrices
b) define stiffness matrix by deflection influence coefficients
c) crack analysis element
d) fluid element

REFERENCES

1. Gallagher, R. H., "Large-Scale Computer Programs for Structural Analysis," On General Purpose Finite Element Computer Programs, P. V. Marcal (ed.), ASME, pp. 1-14., 1970.

2. Butler, T. G., "On the Reduction of Proliferation of Finite Element Programs," NASTRAN: User's Experience, NASA TM X-2378, pp. 813-825, Sept. 1971.

3. Hibbitt, H. D., et al., "On General Purpose Finite Element Computer Programs, P. V. Marcal (ed.), ASME, pp. 198-222, 1970.

4. Marcal, P. V., "On General Purpose Programs for Finite Element Analysis, with Special Reference to Geometric and Material Nonlinearities," Proc. Symp. on Numerical Solution of Partial Differential Equations, University of Maryland, May 1970.

5. Hartung, R. F., "Assessment of Current Capability of Computer Analysis of Shell Structures," AFFDL-TR-67-194, Wright-Patterson Air Force Base, Ohio, Feb. 1970.

6. Perrone, N., "Compendium of Structural Mechanics Computer Programs," Computers and Structures, Vol. 2, No. 3, pp. 305-437.

7. Zienkiewicz, O. C., "The Finite Element Method in Engineering Science," McGraw-Hill, Book Company, New York, NY, 1971.

8. Przemieniecki, J. S., "Theory of Matrix Structural Analysis," McGraw-Hill Book Company, New York, NY, 1968.

9. Oden, J. T., "Finite Elements of Nonlinear Continua," McGraw-Hill Book Company, New York, NY, 1972.

10. Desai, C. S. and Abel, J. F., "Introduction to the Finite Element Method," Van Nostrand-Reinhold Company, New York, NY, 1972.

11. _____, "The NASTRAN Theoretical Manual," NASA SP-221, Sept. 1970.

12. _____, "The NASTRAN User's Manual," NASA SP-222, Sept. 1970.

13. _____, "The NASTRAN Programmers' Manual," NASA SP-223, Sept 1970.

14. _____, "NASTRAN Demonstration Problem Manual," NASA SP-224, Sept. 1970.

15. _____, "NASTRAN: Users' Experiences," (report of Colloquium held at Langley Research Center, Hampton, VA, Sept. 13-15, 1971), NASA TM X-2378, Sept. 1971.

16. MacNeal, R. H., "Dynamic Structural Analysis with the NASTRAN Computer Program," On General Purpose Finite Element Computer Programs, P. V. Marcal (ed.), pp. 78-97, ASME, 1970.

17. Butler, T. G. and Michel, D., "NASTRAN - A Summary of the Functions and Capabilities of the NASA Structural Analysis Computer System," NASA SP-260, 1971.

18. Mallett, R. H. and Jordan, S., "MAGIC - An Automated General Purpose System for Structural Analysis: Engineer's Manual," Vol. I, AFFDL TR-68-56, Air Force Flight Dynamics Laboratory, WPAFB, Ohio, Jan. 1969.

19. Jordan, S., et al, "MAGIC - An Automated General Purpose System for Structural Analysis: User's Manual," Vol. II, AFFDL TR-68-56, Air Force Flight Dynamics Laboratory, WPAFB, OH, July 1969.

20. DeSantis, D., "MAGIC - An Automated General Purpose System for Structural Analysis: Programmer's Manual," Vol. III, AFFDL TR-68-56, Air Force Flight Dynamics Laboratory, WPAFB, OH, Jan. 1969.

21. Jordan, S, "MAGIC II: An Automated General Purpose System for Structural Analysis: Engineer's Manual (Addendum)," Vol. I, AFFDL TR-71-1, Air Force Flight Dynamics Laboratory, WPAFB, OH, May 1971.

22. Mallett, R. H. and Jordan, S., "Stress Analysis Using Finite Element Methods," On General Purpose Finite Element Computer Programs, P. V. Marcal (ed.), ASME, pp. 30-67, 1970.

23. Hibbitt, H. D. and Marcal, P. V., "Hybrid Finite Element Analysis with Particular Reference to Axisymmetric Structures," AIAA 8th Aerospace Sciences Meeting, New York, NY, Jan. 19-21, 1970.

24. Marcal, P. V., "Large Deflection Analysis of Elastic Plastic Shells of Revolution," Proc. of the AIAA/ASME 10th Structures, Structural Dynamics and Material Conference, pp. 369-379, April 1969.

25. Argyris, J. H., et al, "Some New Elements for the Matrix Displacement Method" Second Wright-Patterson Conference, AFFDL-TR-68-150, pp. 333-397, Dec. 1969.

26. Bergmann, H. W., "ASKA - A Large Scale Software System for Finite Element Analysis," U. S.-Japan Seminar on Matrix Methods of Structural Analysis and Design, Paper US-5-1, Tokyo, Japan (published only as a preprint) Aug. 1969.

27. Von Fuchs, G. and Schrem, E., "ASKA - A Computer System for Structural Engineers," Proc. ISD/ISSC, Symp. on Finite Element Techniques, Univ. of Stuttgart, June 10-12, 1969.

28. Argyris, J. H., et al, "Automatic System for Kinematic Analysis," ASKA, ISC Report No. 59, Nov. 1968.

29. Argyris, J. H. and Fried, I., "The LUMINA Element for the Matrix Displacement Method," The Aeronautical Jour., Vol. 72, No. 690, pp. 514-517, June 1968.

30. Argyris, J. H., et al, "The HERMES Element for the Matrix Displacement Method," The Aeronautical Jour., Vol. 72, No. 691, pp. 613-617, July 1968.

31. Argyris, J. H., et al, "The TET 20 and TEA Elements for the Matrix Displacement Method," The Aeronautical Jour., Vol. 72, No. 691, pp. 618-623, July 1968.

32. Argyris, J. H. and Scharpf, D. W., "The Curved Tetrahedron and Triangular Elements TEC and TRIC for the Matrix Displacement Method," The Aeronautical Jour., Vol. 73, pp. 55-65, Jan. 1969.

33. Argyris, J. H., et al, "The TUBA Family of Elements for the Matrix Displacement Method," The Aeronautical Jour., Vol. 72, No. 692, pp. 701-709, Aug. 1968.

34. Argyris, J. H. and Buck, K. E., "A Sequel to Technical Note 14 on the TUBA Family of Plate Elements," The Aeronautical Jour., Vol. 72, pp. 977-983, Nov. 1968.

35. Argyris, J. H. and Fried, I., "The PUBA Family of Plate Elements for the Matrix Displacement Method," ISD Report No. 56, Aug. 1968.

36. Argyris, J. H. and Scharpf, D. W., "The SHEBA Family of Shell Elements for the Matrix Displacement Method," The Aeronautical Journal, Vol. 72, No. 694, Oct. 1968.

37. Warren, D. S., "First Version of Fortran Matrix Abstraction Technique: Volume I, Analysis and Application of Abstraction Technique," AFFDL-TR-65-47, Vol. I, Sept. 1965.

38. Pickard, J., "FØRMAT II - Second Version of Fortran Matrix Abstraction Technique; Volume I. Engineering User Report," AFFDL-TR-66-207, Vol. I, Jan. 1967.

39. Cogan, J. P., "FØRMATII - Second Version of Fortran Matrix Abstraction Technique; Volume II. Description of Digital Computer Program," AFFDL-TR-66-207, Vol. II, March 1967.

40. Morris, R. C., "FØRMAT II - Second Version of Fortran Matrix Abstraction Technique; Vol. III. A User-Coded Matrix Generator for the Force Method," AFFDL-TR-66-207, Vol. III, March 1967.

41. Serpanos, J. E., "FØRMAT II - Second Version of Fortran Matrix Abstraction Technique; Volume IV. A User-Coded Matrix Generator for the Displacement Method," AFFDL-TR-66-207, Vol. IV, March 1967.

42. Pickard, J., "FØRMAT - Fortran Matrix Abstraction Technique; Volume V. Engineering User and Technical Report," AFFDL-TR-66-207, Vol. V, Sept. 1968.

43. Lackey, W. J. and Wild, R. E., "FØRMAT II - Second Version of Fortran Matrix Abstraction Technique; Volume II, Supplement I. Description of Digital Computer Program," AFFDL-TR-66-207, Vol. II, Supplement I, Sept. 1968.

44. Cogan, J. P., et al, "FØRMAT - Fortran Matrix Abstraction Technique; Volume VI. Description of Digital Computer Program - Phase I," AFFDL-TR-66-207, Vol. VI, Sept. 1968.

45. Morris, R. C., et al, "FØRMAT - Fortran Matrix Abstration Technique; Volume VII. Description of Digital Computer Program - Phase III," AFFDL-TR-66-207, Vol. VII, Sept. 1968.

46. Warren, D. S., "Application Experience with the FØRMAT Computer Program," Second Wright-Patterson Conference, AFFDL-TR-68-150, pp. 839-867, Dec. 1969.

47. Greene, et al, "ASTRA - Boeing's Advanced Structural Analyser," Proc. Int. Symp. on Structures Tech. for Large Radio and Radar Telescope Systems, MIT, Oct. 1967.

48. Martin, H. C., "Finite Element Analysis of Fluid Flows," Second Wright-Patterson Conf., AFFDL-TR-68-150, Dec. 1969.

49. _____, "STARDYNE User's Manual," Computer Systems Div., Mechanics Research, Inc., El Sequindo, Calif., June 1968.

50. Chu. S. L., "Analysis and Design Capability of the STRUDL Program," Inter. Symp. Numerical and Computer Methods in Structural Mechanics, U. S. Office of Naval Research, Urbana, Ill., Sept. 1971.

51. Logcher, R. D., et al., "ICES-STRUDL II, The Structural Design Language User's Manual,"
Vol. I: Frame Analysis, MIT, Research Report R68-91, Nov. 1968.
Vol. II: Additional Design and Analysis Facilities, MIT, R68-92, June 1969.

52. Shumacher, B., "An Introduction to ICES," MIT, Dept. of Civil Engr., MIT-DCIE-R67-47, Sept. 1967.

53. Ferrente, A. J., et al, "The solution of Finite Element Problems Using the STRUDL Language," U.S.-Japan Symp. on Matrix Methods of Structural Analysis and Design, Tokyo, Japan, Sept. 1-3, 1969.

54. Utku, S. and Akyuz, F. A., "ELAS - A General Purpose Computer Program for the Equilibrium Problems of Linear Structures, User's Manual," JPL TR-32-1240, Vol. I, Pasadena, Calif., Feb. 1968.

55. Utku, S., "ELAS - A General Purpose Computer Program for the Equilibrium Problems of Linear Structures, Documentation of the Program," JPL TR-32-1240, Vol. II, Pasadena, Calif., Sept. 1969.

56. Percy, J. H., et al, "SABOR I: A Fortran Program for the Linear Elastic Analysis of Thin Shells of Revolution Under Axisymmetric Loading Using the Matrix Displacement Method," ASRL TR 121-5, MIT Aeroelastic and Structures Research Laboratory, May 1965.

57. Percy, J. H., et al., "SABOR III: A Fortran Program for the Linear Elastic Analysis of Thin Shells of Revolution Under Asymmetric Loading Using the Matrix Displacement Method," ASRL TR 121-6, MIT, May 1965.

58. Mack, E. W., et al, "An Improved Discrete-Element Analysis and Program for the Linear-Elastic Static Analysis of Meridionally-Curved, Variable-Thickness, Branched Thin Shells of Revolution Subjected to General External Mechanical and Thermal Loads, Part II; The SABOR 4 Program," ASRL TR 146-4, Part 2, MIT, March 1968.

59. Witmer, E. A., et al, "An Improved Discrete-Element Analysis and Program for the Linear Elastic Static Analysis of Meridionally-Curved Variable Thickness Branched Thin Shells of Revolution Subjected to General External Mechanical and Thermal Loads, Part 1: Analysis and Evaluation," ASRL TR 146-4, MIT, March 1968.

60. Kotanchik, J. J., "Discrete Element Static Analysis of Bonded Double Layer, Branched, Thin Shells of Revolution, Part 1: Analysis and Evaluation," ASRL TR-139-6, Part 1, MIT, May 1968.

61. Kotanchik, J. J., "Discrete-Element Static Analysis of Bonded Double Layer, Branched Thin Shells of Revolution, Part 2: The SABOR 5 Program," ASRL TR-139-6, Part 2, MIT, May 1969.

62. Klein, S. and Sylvester, R. J., "The Linear Elastic Dynamic Analysis of Shells of Revolution by the Matrix Displacement Method," First Wright-Patterson Conf., AFFDL-TR-66-80, pp. 299-328, Nov. 1966.

63. _____, "DRASTIC 5," Computer Code Developed at MIT Aeroelastic and Structures Research Laboratory.

64. Melosh, R. J. and Christianson, "Structural Analysis and Matrix Interpretive System (SAMIS) Program: Technical Report," JPL TM-33-311, Pasadena, Calif., Nov. 1966.

65. Melosh, R. J., et al., "Structural Analysis and Matrix Interpretive System (SAMIS) Program Report," JPL TM-33-307, Pasadena, Calif., Dec. 1966.

66. _____, "Course on Recent Development of Finite Element Methods in Engineering and Physical Sciences," Univ. of Wales, Swansea, Jan. 5-9, 1970.

67. King, I. P., "Finite Element Solution Systems (FESS)," Internal Report Centre for Numerical Methods in Engineering, University of Wales, Swansea, 1967.

68. _____, "EASE - Elastic Analysis for Structural Eng., User's Manual and Application Guide," Prepared by Eng. Analysis Corpora tion, Distributed by CDC, Data Centers Div., Aug. 1969.

69. Kamel, H. A., "MINI-ASKA User's Manual," research performed for Lockheed Missiles and Space Co., under Research Contract No. PØAPz8J9280A, May 1968.

70. Loden, W. A., "User's Manual for REXBAT 5," Lockheed Missiles and Space Co., Palo Alto, Calif., (Unpublished report), July 1970.

71. Haisler, W. E. and Stricklen, J. A., "SNASOR II - A Finite Element Program for the Static Nonlinear Analysis of Shells of Revolution," Space Technology Report 70-72, Aerospace Engr. Dept., Texas A&M Univ., College Station, Texas, Oct. 1970.

72. Whetstone, W. D., "User's Manual for SNAP," Lockheed Huntsville Research and Engineering Center, Huntsville, Ala.

73. Wilson, E. L., "SAP - A General Structural Analysis Program," Rept. No. UCSESM 70-20, Dept. of Civil Engr., Univ. of Calif., Berkeley, Calif., Sept. 1970.

74. Key, S. W. and Beisinger, Z. E., "SLADE: A Computer Program for the Static Analysis of Thin Shells," SC-RR-69-369, Sandia Corporation, Albuquerque, New Mexico, Nov. 1970.

75. Strickland, et al, "WASP - A Digital Computer Program for the Linear Elastic Analysis of Hybrid, Symmetrically Loaded Bodies of Revolution," Report 9-87-68-2, Lockheed Missiles and Space Company, Palo Alto, Calif., June 1968.

76. Lindberg, G. M., et al, "Finite Element Dynamic Analysis of Shallow Shell Structures," National Research Council of Canada, National Aeronautical Establishment, NRC, NAE LR-540, July 1970.

77. Kotanchik, J. J. and Berg, B. A., "STACUSS 1: A Discrete-Element Program for the Static Analysis of Single-Layer Curved Stiffened Shells Subjected to Mechanical and Thermal Loads," MIT, ASRL TR 146-9, Dec. 1969.

78. _____"SAMECS User's Manual (TS-172), Structural Analysis Method for Evaluation of Complex Structures," Boeing Document No. D6-29059, May 1967.

79. Logcher, R. D., "ICES-STRUDL - An Integrated Approach to a Structural Computer System," Proc. Int. Sym. on Structures Tech. for Large Radio and Radar Telescope Systems, MIT, Oct. 1967.

80. Gallo, A. M., "MAGIC II: An Automated General Purpose System for Structural Analysis; Volume III, Programmer's Manual," AFFDL TR-71-1, Vol. 3, WPAFB, Jan. 1971.

81. Butler, T. G., "Rosman I Dynamic Analysis," Proc. Int. Sym. on Structures Tech. for Large Radio and Radar Telescope Systems, MIT, Oct. 1967.

82. Jordan, S., "MAGIC II: An Automated General Purpose System for Structural Analysis; Volume II, User's Manual," AFFDL-TR-71-1, WPAFB, Jan. 1971.

83. Akin, J. E., et al., "The Finite Element Method, a bibliography of its theory and applications," Dept. of Engineering Mechanics, Report EM 72-1, Univ. of Tenn, Knoxville, Tenn 37916, Feb. 1972.

84. Tottenham, H. and Brebbia, C. (eds.), "Finite Element Techniques in Structural Mechanics," Proc. of a Seminar at the Univ. of Southhampton, April 1970.

85. Key, S. W. and Beisinger, Z. E., "The Transient Dynamic Analysis of Thin Shells by the Finite Element Method," Third Wright-Patterson Conference, Oct. 1971.

86. Grisham, A. F., "The Boeing SST Prototype Internal Loads Analysis System and Procedures," Third Wright-Patterson Conference, Oct. 1971.

87. Argyris, J. H., "Finite Element Analyses of Thermomechanical Problems," Third Wright-Patterson Conference, Oct. 1971.

88. Gallagher, R. H., "A Procedure for Finite Element Plate and Shell Pre- and Post-Buckling Analyses," Third Wright-Patterson Conference, Oct. 1971.

89. Przemieniecki, J. S., "Finite Element Structural Analysis of Local Instability," AIAA Paper No. 72-354, AIAA/ASME/SAE 13th Structures, Structural Dynamics, and Materials Conference, San Antonio, Texas, April 1972.

90. Stricklin, J. A., "Evaluation of Solution Procedures for Material and/or Geometrically Nonlinear Structural Analysis by the Direct Stiffness Method," AIAA Paper No. 72-353, AIAA/ASME/SAE 13th Structures, Structural Dynamics, and Materials Conference, San Antonio, Texas, April 1972.

91. Sharafi, P. and Popov, E. P., "Nonlinear Finite Element Analysis of Sandwich Shells of Revolution," AIAA Paper No. 72-356, AIAA/ASME/SAE 13th Structures, Structural Dynamics, and Materials Conference, San Antonio, Texas, April 1972.

92. Kamel, H. A., "The Computer in Finite Element Analysis of Ship Structures," A Symposium and Short Course, Univ. of Arizona, Tucson, Arizona, March 1971.

93. Haisler, W. E., et al., "Development and Evaluation of Solution Procedures for Geometrically Nonlinear Structural Analysis," AIAA J., Vol. 10, No. 3, March 1972.

94. Hwang, C. and Pi, W. S., "Nonlinear Acoustic Response Analysis of Plates Using the Finite Element Method," AIAA J., Vol. 10, No. 3, March 1972.

95. Sabir, A. B. and Ashwell, D. G., "A Comparison of Curved Beam Finite Elements When Used in Vibration Problems," Jour. of Sound and Vibration, Vol. 18, pp. 555-563, 1971.

96. Akmed, K. M., "Dynamic Analysis of Sandwich Beams," Jour. of Sound and Vibration, Vol. 21, pp. 263-276, 1972.

97. Henshell, R. D., et al, "A New Family of Curvilinear Plate Bending Elements for Vibration and Stability," Jour of Sound and Vibration, Vol. 20, pp. 381-397, 1972.

98. Robinson, J. and Haggenmacker, H. W., "Basis for Element Interchangeability in Finite Element Programs," Third Wright-Patterson Conference, Oct. 1971.

99. Mei, C., "Nonlinear Vibration of Beams by Matrix Displacement Method," AIAA Jour., Vol. 10, No. 3, March 1972.

100. Gupta, K. K., "Dynamic Response Analysis of Geometrically Non-linear Structures Subjected to High Impact," Inter. Jour. of Numerical Methods in Engineering, Vol. 4, March 1972.

101. Lui, T. H. and Ito, Y. M., "Elastic-Plastic Analysis of Unidirectional Composites," Jour. of Composite Materials, Vol. 6, Jan. 1972.

102. Belytschko, T., "Finite Element Method for Elastic Plastic Sandwich Plates," ASCE, Jour. Engr. Mech. Div., Vol. 98, Feb. 1972.

103. Miyamoto, H., "Application of the Finite Element Method to the Fracture Mechanics," Tokyo Univ., Faculty of Engineering, Journal, Series B, Vol. 31, Sept. 1971.

104. Broberg, K. B., "Crack-growth Criteria and Nonlinear Fracture Mechanics," Jour. of the Mechanics and Physics of Solids," Vol. 19, Nov. 1971.

105. Mau, S., "A Finite Element for Nonlinear Prebuckling and Postbuckling," NASA CR-1936.

106. Hartzman, M., "Nonlinear Dynamics of Solids by the Finite Element Method," Computers and Structures, Vol. 2, Jan. 1972.

1972

NASA-ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

AERONAUTICS AND SPACE RESEARCH PROGRAM

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

MARSHALL SPACE FLIGHT CENTER

INVESTIGATION OF SOME CHARACTERISTICS RELATED TO PCM THERMAL CAPACITORS

Prepared by:	Edwin I. Griggs
Academic Rank:	Assistant Professor
University:	Tennessee Technological Univ.
Laboratory: (Division) (Branch)	Astronautics Laboratory Propulsion & Thermodynamics Life Support & Environmental
Research Counterpart:	W. R. Humphries
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

LIST OF FIGURES

LIST OF FIGURES
(Continued)

LIST OF FIGURES
(Continued)

LIST OF TABLES

NOMENCLATURE

Symbol	Definition	Units
ADI	Alternating-Direction-Implicit Numerical Method	
C	Thermal Capacitance of a Node	Btu/^{o}F
C_p	Constant Pressure Specific Heat	Btu/Lbm-F^{o}
GH	Horizontal Thermal Conductance	Btu/Hr-^{o}F
GV	Vertical Thermal Conductance	Btu/Hr-^{o}F

NOMENCLATURE
(Continued)

H	Height of PCM in Cell(Fig. 38)	inches
I	Nodal Designation Integer (Fig. 38)	
J	Nodal Designation Integer (Fig. 38)	
k	Thermal Conductivity	Btu/Hr-Ft-^{o}F
M	Number of Horizontal Nodes in PCM (Fig. 38)	
N	Number of Vertical Nodes in PCM (Fig. 38)	
PCM	Phase Change Material	
Q	Heat Transfer	Btu/Hr
S	Nodal Spacing in PCM	inches
S_1	Thickness of Bottom Plate	inches
S_2	Thickness of Fin	inches
T	Temperature	^{o}F
t	Time	Hr.
W	Width of PCM Cell	inches
x	Spatial Coordinate	Ft.
y	Spatial Coordinate	Ft.
α	Thermal Diffusivity	$Ft^2/Hr.$
Δ	Designates Finite Increment	
ρ	Density	Lbm/Ft^3

ACKNOWLEDGEMENTS

The author would like to express sincere appreciation to the entire staff of the 1972 NASA/ASEE Summer Faculty Fellowship Program for their role in providing a well organized and rewarding summer experience. Particular thanks are due Mr. J. Fred O'Brien, Director, Dr. Donald C. Raney, Associate Director, and Mr. Marion I. Kent, NASA Representative, for their respective contributions.

In addition, the author expresses his gratitude to each member of the Life Support and Environmental Branch of the Propulsion and Thermodynamics Division of the Astronautics Laboratory for their provision of a cooperative, challenging, and interesting environment in which to work during the summer.

Words of special thanks and appreciation are extended to Mr. W. R. Humphries for serving as my research counterpart on the program. His experience and interest in the area of study served to initiate the effort as well as to promote a continuing educational experience. His cooperation, technical discussion, and assistance were most helpful.

Finally, I want to thank those in NASA who have worked to promote and support this program whereby research and continuing educational opportunities have been made available to a significant number of university faculty across our country.

INVESTIGATION OF SOME CHARACTERISTICS RELATED TO PCM THERMAL CAPACITORS

By

Edwin I. Griggs

ABSTRACT

This report presents some results of a continuing experimental and analytical investigation of the thermal behavior of a PCM thermal capacitor. Fundamentally, a thermal capacitor is a device which provides attractive energy storage capability through utilization of a phase change material (PCM). Functionally, these units may play the role of a thermal flywheel by depressing temperature fluctuations caused by a changing thermal environment or they may serve as a heat sink to delay transient temperature rises. Since their operation is passive, these devices are particularly suitable for applications in the space program, a fact exemplified by the variety of current as well as proposed future applications.

In order for capacitor designs to fully and successfully exploit the advantages of the phase change concept, the PCM should have a high value of thermal diffusivity; this, however, creates a problem since the prime PCM candidates (n-paraffins), based on other considerations, have very low values of this property. To overcome this problem and achieve improved heat transfer in practice, metallic fins or some other form of metallic filler is embedded within the PCM. Such an arrangement, however, complicates an exact thermal analysis because of the introduced inhomogeneities and multidimensional effects.

The objectives of this work were (1) to continue to investigate and evaluate numerical techniques for predicting thermal performance within a thermal capacitor and (2)

to monitor a series of tests on an invertable thermal capacitor. The numerical study is a continuation of an effort initiated during participation in the 1971 NASA/ASEE Faculty Fellowship Program. During this former period, in addition to developing some background in the area, a numerical study utilizing an explicit formulation was directed to the prediction of the heat transfer within a single rectangular cell formed by metallic fins enclosing the PCM. Hopefully, in order to provide greater flexibility, shorter computer time, and provide potential for a parametric study, attention during this summer has been devoted to an implicit numerical scheme. A discussion of this effort is presented. The experimental work involved outlining and monitoring approximately fourteen tests of an invertable capacitor. Surface and PCM temperatures were measured at several locations and photographs were made of the PCM during melting and solidification. Some typical experimental results are presented and discussed.

INTRODUCTION

As a quantity of material undergoes a phase change, it absorbs or liberates an amount of energy equal to its heat of transition while remaining at constant temperature. During this transition, the effective thermal capacitance, ideally, is infinite. Practically, internal temperature gradients prevent this ideality from being completely realized. Utilization of a phase change to provide thermal control for transient systems is expanding. A thermal capacitor is a device which contains a phase change material(PCM) and is used for thermal control and/or energy storage purposes. While a liquid-vapor transition provides much larger energy storage capability per unit mass than does a solid-liquid transition, the latter has some practical advantages since the former presents containment problems due to the large volume change occurring in the liquid-vapor transition.

PCM thermal control is particularly attractive for applications in the space program since the technique is passive, and, therefore, should have good reliability. For short-term or one-time operated components where heat rejection, under normal conditions, may be insufficient to maintain a desired temperature level, PCM packages can serve as a heat sink to absorb the energy at near constant temperature. PCM devices can be used in conjunction with space radiators to store energy during high heat rejection periods and subsequently allow the radiator to radiate continuously at a lower temperature. This balancing scheme eliminates design of an extra large radiator to accomodate peak heating loads. Other applications include those of providing isothermal environments for experimental packages and biological specimen storage units. Thermal capacitors were employed on the Lunar Roving Vehicles and are to be used on Skylab(Ref. 1). The proposed thermal control subsystem for planetary descent probes will also involve use of a PCM (Ref. 2).

Effort is currently in progress toward application

of PCM technology together with other space-age technology to the problem of effectively utilizing solar energy. A present concept under study is that of incorporating a solar energy collector, a PCM storage unit, and an absorption refrigeration cycle to meet home and office cooling requirements. The need for a PCM which is compatible with the remainder of the system and its requirements is an indication of the expanding interest in PCM thermal control.

For a particular application, such as the one mentioned above, selection of an appropriate PCM is a prime design consideration. An ideal PCM should have the following characteristics (Ref. 3).

- High Heat of Fusion
- High Thermal Conductivity
- High Specific Heat
- High Density
- Low Volume Change During Phase Change
- Reversible Phase Change
- Long Life Reliability
- Dependable Freezing Behavior
- Low Vapor Pressure
- Low Toxicity
- Chemically Inert and Stable

For applications involving melt temperatures between 40 ^{o}F and 150 ^{o}F, the materials which emerged as the best PCM candidates in a study reported in Reference 4 were the normal (n-) paraffins containing from 14 to 30 carbon atoms. Potentially new PCMs including metallic PCMs, patented PCMs, freons, waxes and oils, low molecular weight compounds, organics, and composite PCMs have been discussed in Reference 5. Also Reference 3 has summarized a number

of PCMs and has tabulated many of their properties.

While many factors need to be evaluated in the choice of a PCM, the melt temperature is one of the major constraining considerations. For an application where the PCM is to protect a thermally connected component, the melt temperature must lie in the safe operating range of the component. Consider the schematic shown in Figure 1.

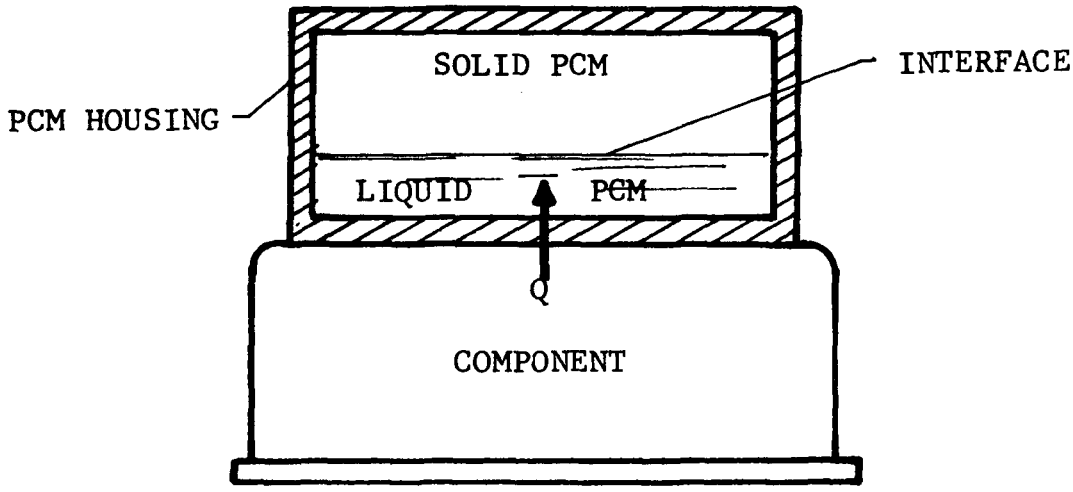


Figure 1 Schematic of PCM Used to Protect a Compone:

As energy is transferred from the component to the PCM, an ideal design would provide sufficient thermal communication to assure that the temperature of the component remains near the melt temperature of the PCM. Use of paraffins in this connection presents a practical problem. As the solid-liquid interface recedes from the surface adjacent to the component, a high resistance melt layer develops with significant internal temperature gradients. Consequently, the component temperature must rise appreciably above the PCM melt temperature. The high resistance is a consequence of the very low thermal conductivity values of normal paraffins. To partially alleviate this problem via provision of an improved conductivity between the interface

and the heated surface, metallic fillers have been used within the PCM. These fillers have included straight fins, honeycomb, foam, powder, and wool. Use and performance considerations of fillers have been discussed in References 3,4,and 5. The case of straight fins is schematically illustrated in Figure 2.

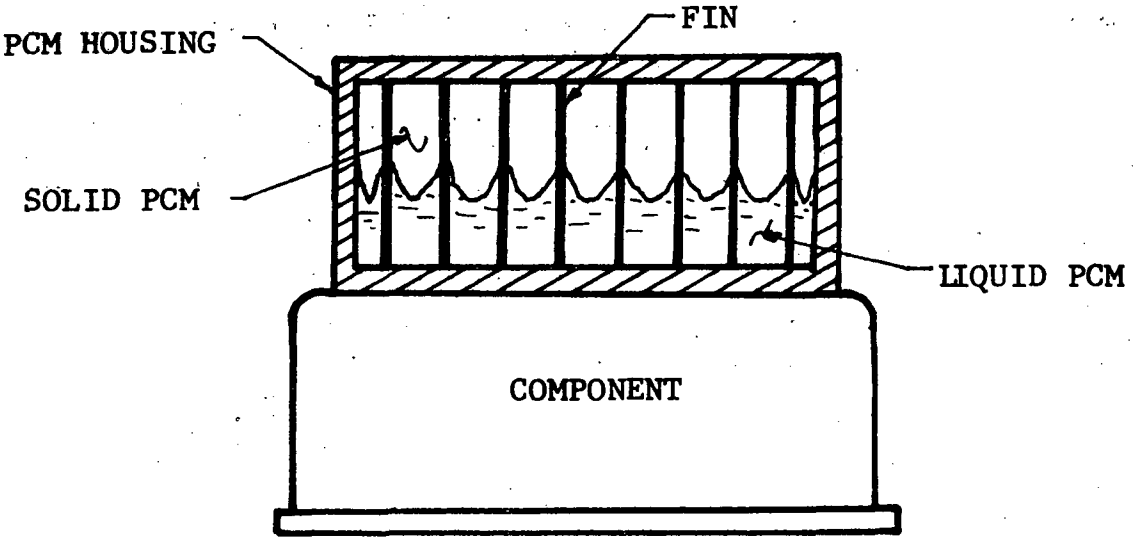


Figure 2 Illustration of Finned Thermal Capacitor

While the case illustrated in Figure 2 probably represents the simplest geometrical arrangement involving fillers, optimization of even this case for design purposes is still needed. The addition of a filler reduces the volume that is available to the PCM. Consequently, efficient designs should be based on an optimum arrangement that utilizes as much as possible the advantages of the phase change yet provides a sufficiently high conducting path to prevent large temperature gradients. In order to determine these optimum conditions, it is necessary to be able to understand the physics and to have an analytical model which adequately describes the thermal behavior.

As a first step in analyzing a finned thermal capacitor, attention has been focused on a typical cell within the

unit. A typical cell is schematically illustrated in Figure 3a.

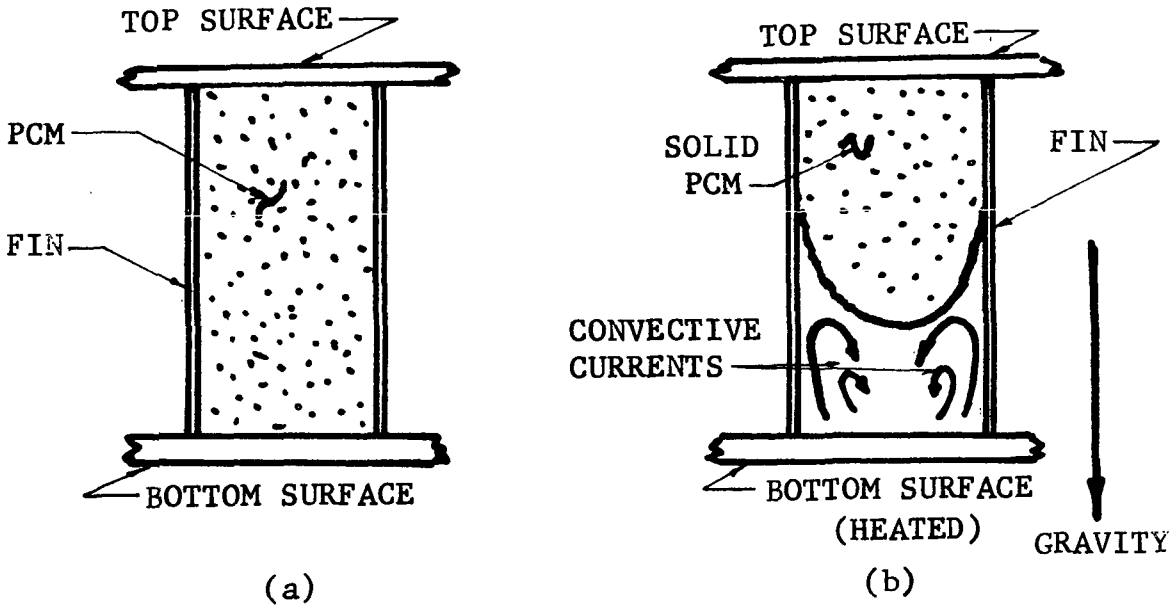


Figure 3 Typical Finned Capacitor Cell

If effects perpendicular to the plane of the sketch are neglected, the two-dimensional problem is still complicated because of the presence of two different materials demonstrating very different thermal properties. Also, attempts to correlate predictions with normal gravity test data are involved since convection exists when the heated surface is at the bottom such as that illustrated in Figure 3b. Prediction of the convective effects is a most formidable problem. There appears to be no information available in the literature which treats the problem of convection within a celluar enclosure where one boundary is moving. Even for steady conditions in stationary cells, the possible convective patterns are complex as illustrated by the study reported in Reference 6. Correlations of convective contributions to heat transfer in one-dimensional steady state experiments have been presented by Reference 7. Uses of these correlations in some phase change problems have been reported by References 8 and 9. More closely related but still unapplicable directly to phase

change problems is the work on liquid convection in the presence of honeycomb cells reported in Reference 10. Detailed mathematical treatment of natural convection in enclosures wherein phase change is occurring is a real need for accurate analysis of the problem under consideration. For space applications, however, gravity driven convection is not present. In such cases, a pure conduction model may suffice. An argument against this conjecture, however, is the possibility of surface tension driven convection as suggested in Reference 11.

The goal of the work during this summer was to continue an experimental and analytical investigation into the thermal characteristics of a PCM thermal capacitor. Specifically, the objectives were (1) to conduct a series of tests on an invertable, inhouse, PCM test capacitor and (2) to continue to examine numerical techniques for analyzing the phase change problem within a capacitor cell. The experimental work was pursued to provide additional data on the temperature responses as well as to provide some insight into the influence of gravity by testing the capacitor in a normal and an inverted configuration. Some details and results pertinent to this experimental effort are given in the next section. The continuation of the numerical work was pursued in an effort to gain a scheme that would eventually provide the capability for a parametric study and lead to treatment of a total unit. Further details on this aspect are given in the section on the numerical study.

EXPERIMENTAL STUDY

A portion of the effort during this summer was devoted to outlining and monitoring a series of tests on a previously built PCM thermal capacitor model. The model had been constructed to facilitate testing in both a normal and an inverted configuration. Motivation for pursuing similar tests in both orientations was the need to further consider and examine the influence of natural

convection. In addition the model was equipped with a series of fins arranged to provide three cells each of 3/4, 1/2, 1/4, and 1/8 inch spacing. This provision was included to facilitate further examination of the influence of fins. The side walls of the unit were made of plexiglass to allow visual observation. Further details are given below.

Test Apparatus

The primary housing for the PCM was constructed from 6061-T6 aluminum alloy. The PCM region was 1 1/2 inches deep by 3 1/2 inches wide by 5 1/8 inches long. The bottom of the housing was 1 1/4 inches thick by 3 1/2 inches wide by 7 inches long. The ends and top of the housing were approximately 5/16 inch thick. On the sides of the housing was a rectangular band made from 1/2 inch plate. The entire housing was welded and designed to withstand a working pressure of 20 psig. Five 1/2 inch dia. holes were drilled lengthwise through the bottom block to provide passages for the heating or cooling fluid used to thermally communicate with the PCM. A manifold and plenum arrangement was mounted on each end of the bottom block to distribute flow over the passages. The side faces of the band were milled flat and each side contained sixteen threaded holes to accomodate 1/4-20 bolts which held the plexiglass side faces against the housing. Sealing was provided with a rubber gasket. Thirteen fins were installed in the PCM region providing three each 3/4, 1/2, 1/4, and 1/8 inch cells. The fins were made from 5052-H38 aluminum alloy. A lengthwise cross section of the test unit is schematically shown in Figure 4. Complete details are shown on MSFC drawing, E90M04393. The entire apparatus was well insulated with the exception of the plexiglass side walls; the thermal conductivity of plexiglass is very low.

In order to provide controlled heating and cooling to the PCM, the capacitor was positioned in a flow loop which existed as part of a test setup previously used in building 4653. The flow diagram is schematically shown in

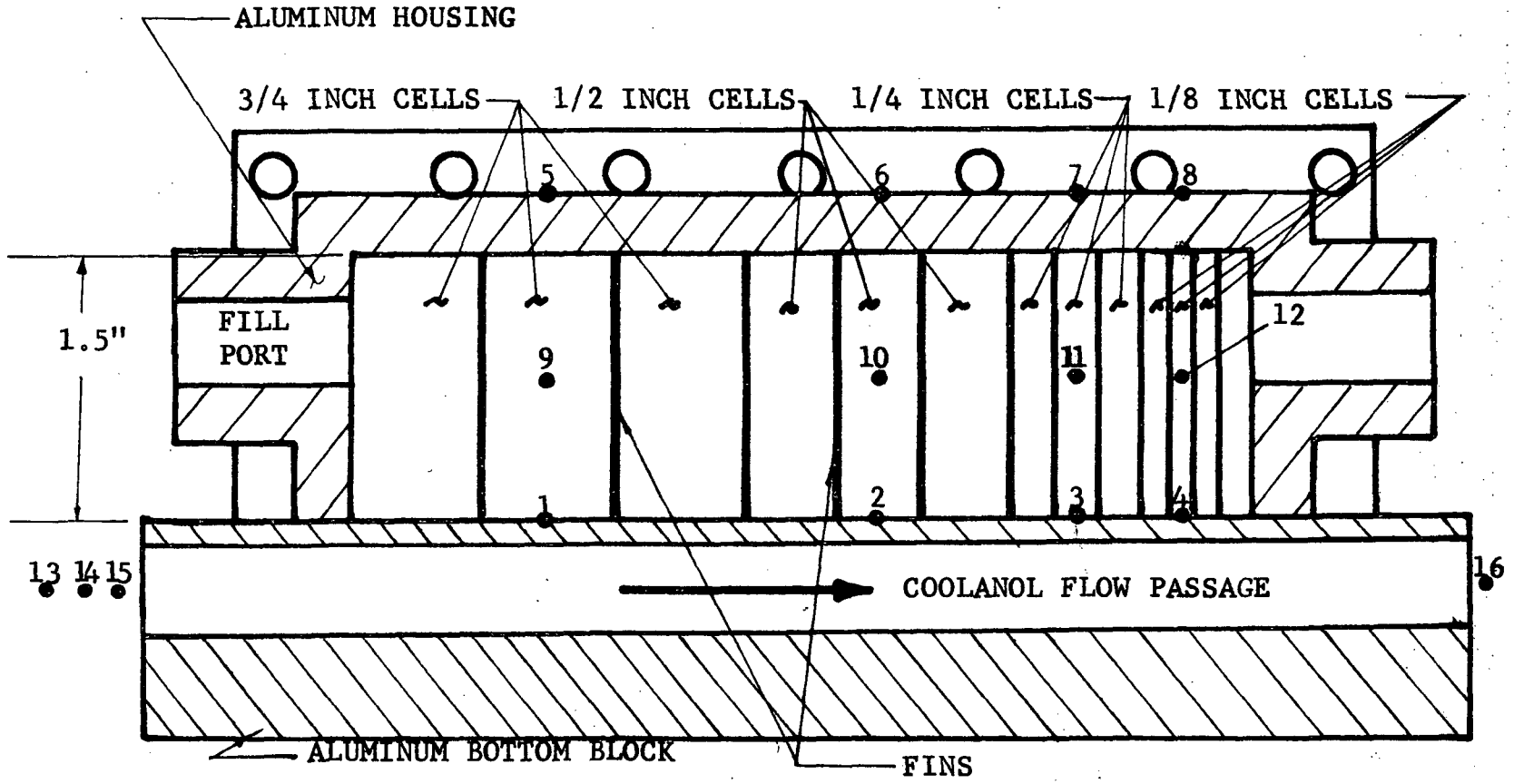


Figure 4 Schematic of Lengthwise Cross Section of Test Capacitor (● With Numbers Designate Thermocouple Locations)

Figure 5. The setup included a fluid reservoir, a variable speed pump, a turbine type flowmeter, a cold heat exchanger, a hot heat exchanger, mixing chambers, filters, and insulated lines with appropriate valves. The hot heat exchanger consisted of passing a coiled loop of the flow line through a bath of ethylene glycol which was maintained at a desired temperature by means of a thermostatically controlled electrical heater. Similarly, the cold heat exchanger involved passing a loop through a bath of ethylene glycol which was maintained at a low temperature with dry ice. A desired temperature at the inlet of the PCM capacitor was obtained by mixing the flow from the two heat exchangers by means of valve adjustments. The mixing chambers were incorporated to assure good thermal mixing and to avoid the possibility of stratified layers reaching the inlet to the test unit.

The PCM used for all tests was nonadecane ($C_{19}H_{40}$) which has a reported melt temperature of 89.8°F. The fluid used in the supporting flow loop was MonSanto Coolanol 15.

The system was instrumented with sixteen copper constantan thermocouples. Four were used to measure coolanol temperatures and the other twelve were used to measure temperatures in the PCM and on its housing. Three were mounted inside the inlet line upstream of the capacitor. Three were used to ascertain good mixing of the fluid. The remaining coolanol measurement was made in the exit line downstream of the capacitor. Four thermocouples were mounted on the bottom plate. These were located to provide a measurement in the middle of the center cell for each spacing group. Similarly, four were located along the centerline of the top of the housing and directly above the center cell of each of the spacing groups. The other four were supported by sheath covered probes and located within the PCM in the center cell of each group and midway between the top and bottom surfaces. The locations and the corresponding number designations are given in Figure 4. Those thermocouples located in the housing were inserted into small holes. The metal was peened around the bead and the location was covered with a layer of epoxy. The four PCM probes were inserted horizontally through one of the plexi-

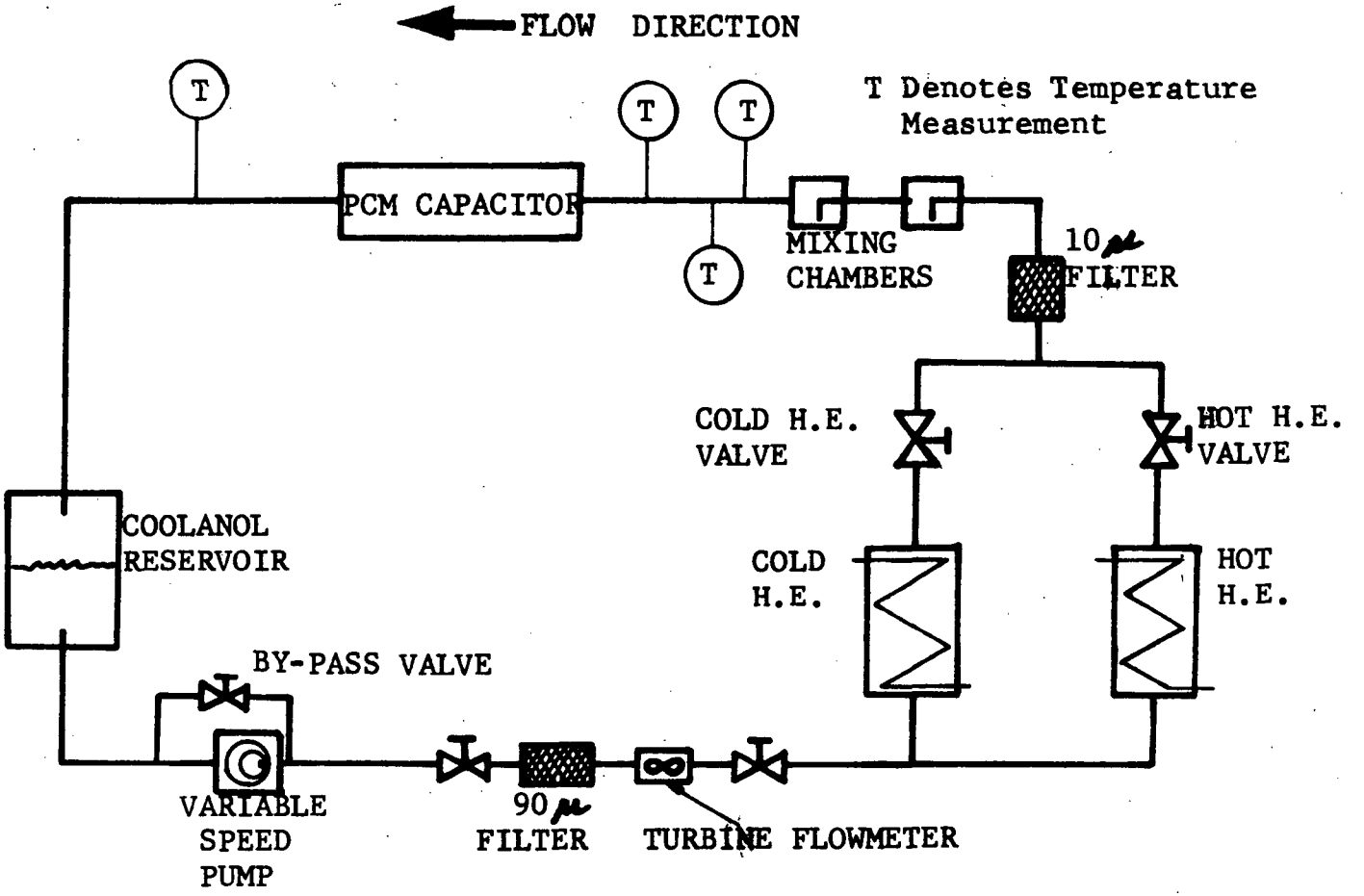


Figure 5 Flow Schematic of Test Setup

glass side windows.

The thermocouple emf outputs for all tests were recorded on a bank of Bristol strip chart recorders. The recording system included a 150° F reference junction for all readings. A common time base was provided by electrically controlled markings on each respective chart. The turbine flowmeter output was monitored with the aid of an electronic counter.

Photographs were made of the PCM during tests by means of a 35 mm camera focused through the plexiglass side walls.

A partial view of the test capacitor is given by the photograph in Figure 6. The capped port on the end plate was used in filling the unit with PCM. Figure 7 is a photograph of the test setup. The hot and cold heat exchangers are not visible in this view. The capacitor is located near the lower left hand corner of the photograph.

The capacitor was completely filled with nonadecane by supplying liquid to the port in one end and allowing air to bleed out of the port in the other end. The housing was heated during filling to prevent solidification. After filling, both ports were tightly sealed.

Experimental Tests and Procedure

It was desired to conduct several tests including melting and solidification with the capacitor in both a normal and an inverted orientation. The intent was to maintain identical conditions for comparable inverted and normal orientations.

Several preliminary steps were performed prior to the initiation of a run. The ethylene glycol bath in the hot heat exchanger was heated to the desired temperature level. Dry ice was placed in the bath for the cold heat exchanger until a low temperature solid-liquid mixture existed. The capacitor and PCM temperatures were allowed to reach an essentially uniform level.

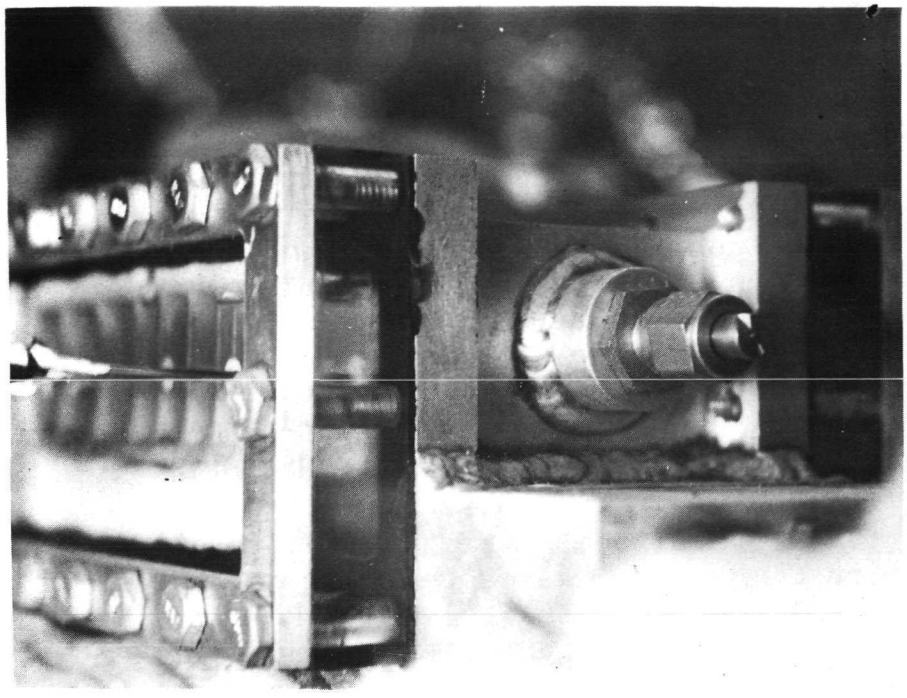


Figure 6 Photograph of Test Capacitor

Figure 7 Photograph of Test Setup

A test was started by initiating the coolanol flow. The pump speed was quickly set to produce the desired flow rate and this was periodically checked during a test. The inlet coolanol temperature was brought to the nominal desired value as rapidly as possible by adjusting the valves in the hot and cold lines upstream of the mixing chamber. This temperature was monitored visually by means of a recorder. Simultaneously with the initiation of the coolanol flow, the strip chart recorders were started and a common zero time base on each recording was provided by a signal controlled with a common switch. Throughout the duration of a test, the valves were adjusted as necessary in order to maintain the inlet temperature as near the desired level as possible. Also a number of photographs and visual observations were made throughout the test. A melt test was terminated when all PCM solid had melted. A solidification test was terminated when all liquid appeared to have solidified.

A summary of the tests is shown in Table 1.

Results

A portion of the data has been reduced from the strip charts. Typical temperatures for comparable melt runs 4 and 10 and comparable freeze runs 5 and 11 are shown in Figure 8 through 21. For each run, three sets of curves are shown. First, the four PCM temperatures and the coolanol inlet and outlet temperatures are given. Second, temperatures for the top and bottom surfaces are presented. Third, comparisons between the top, bottom, and PCM temperatures for the 3/4 inch cell locations and the 1/4 inch cell locations are shown. These three sets of comparisons are respectively presented in Figures 8,9, and 10 for melt run 4, Figures 11, 12, and 13 for melt run 10, Figures 14, 15, and 16 for freeze run 5, and Figures 17,18,19 for freeze run 11. In order to provide some direct comparison between the two orientations, Figure 20 shows temperatures of the bottom (heated) surface and the PCM at the 1/2 inch cell locations for melt runs 4 and 10. Figure 21 presents this same comparison for freeze runs 5 and 11. In the

Table 1 Summary of Tests on Invertable PCM Test Capacitor

TEST NO.	CAPACITOR ORIENTATION	TYPE RUN	NOMINAL INITIAL PCM TEMPERATURE	NOMINAL COOLANOL INLET TEMPERATURE	COOLANOL FLOW RATE
1	NORMAL	MELT	70 - 75 °F	110 °F	500 Lbm/Hr
2	NORMAL	MELT	70 - 75	135	500
2A	NORMAL	MELT	70 - 75	135	500
3	NORMAL	MELT	70 - 75	160	500
4	NORMAL	MELT	70 - 75	135	120
5	NORMAL	FREEZE	110	60	500
6	NORMAL	FREEZE	110	20	500
7	INVERTED	MELT	70 - 75	110	500
8	INVERTED	MELT	70 - 75	135	500
9	INVERTED	MELT	70 - 75	160	500
10	INVERTED	MELT	70 - 75	135	120
11	INVERTED	FREEZE	110	60	500
12	INVERTED	FREEZE	110	20	500
14	NORMAL	FREEZE	110	20	500

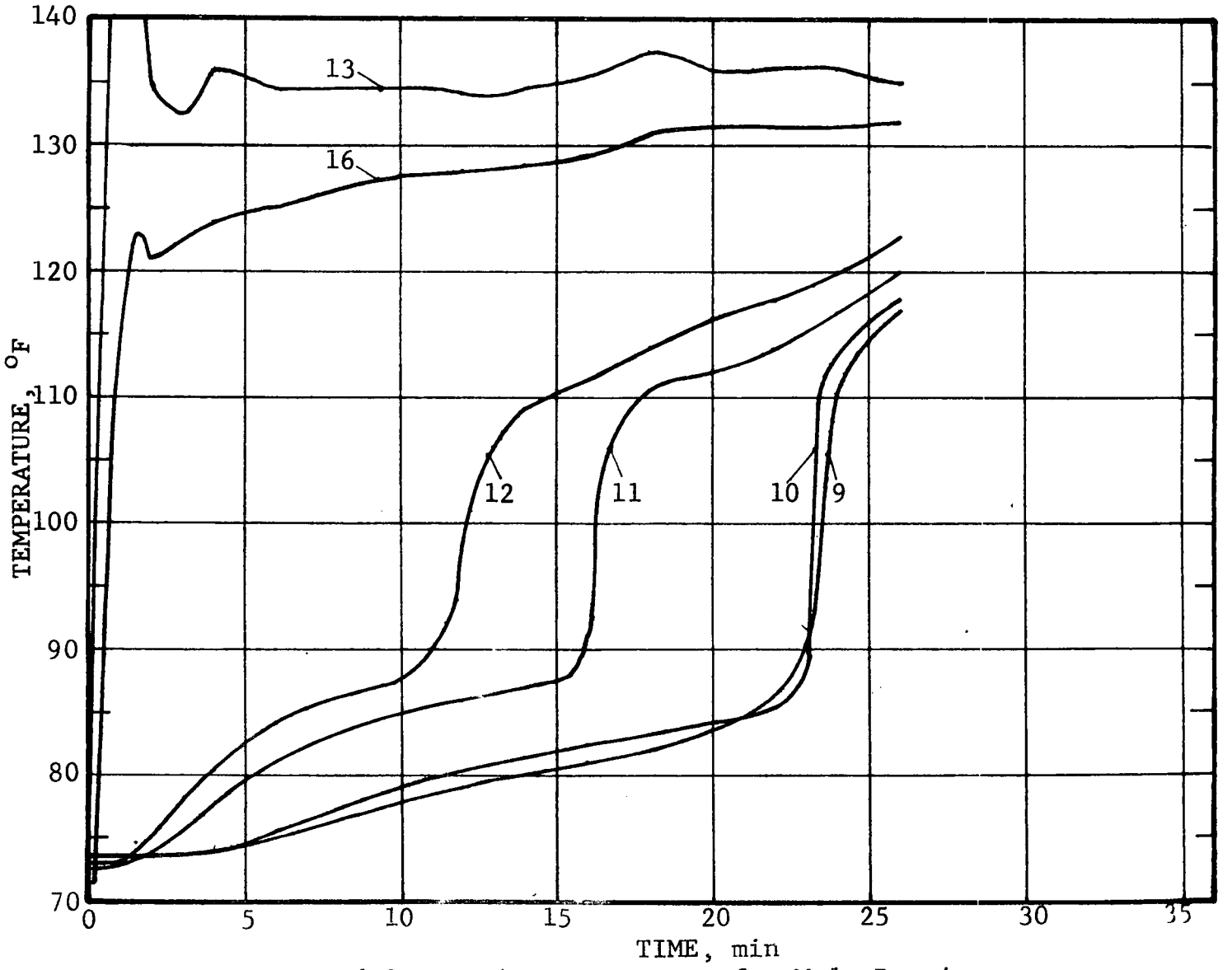


Figure 8 PCM and Coolant Temperatures for Melt Run 4

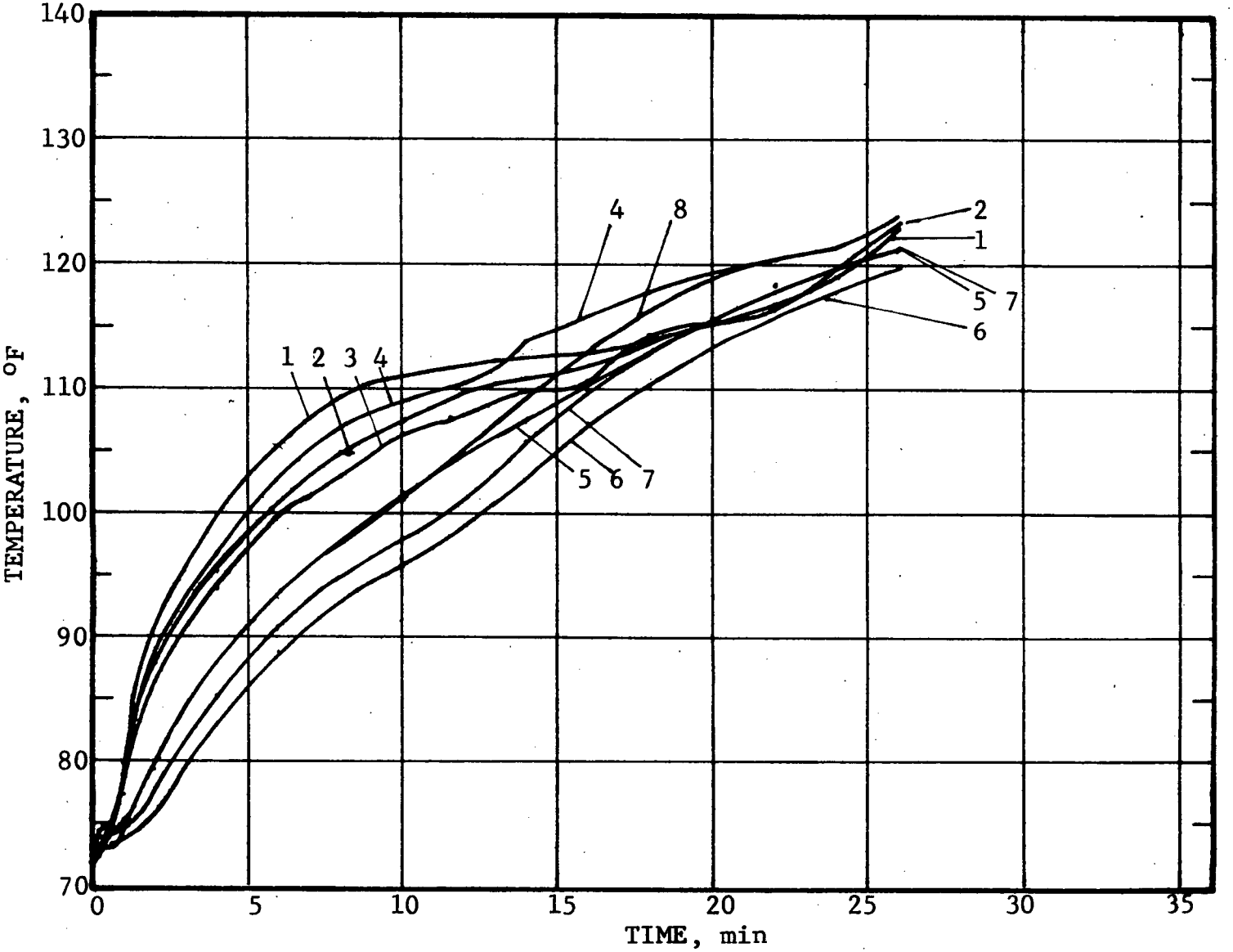


Figure 9 Surface Temperatures for Melt Run 4

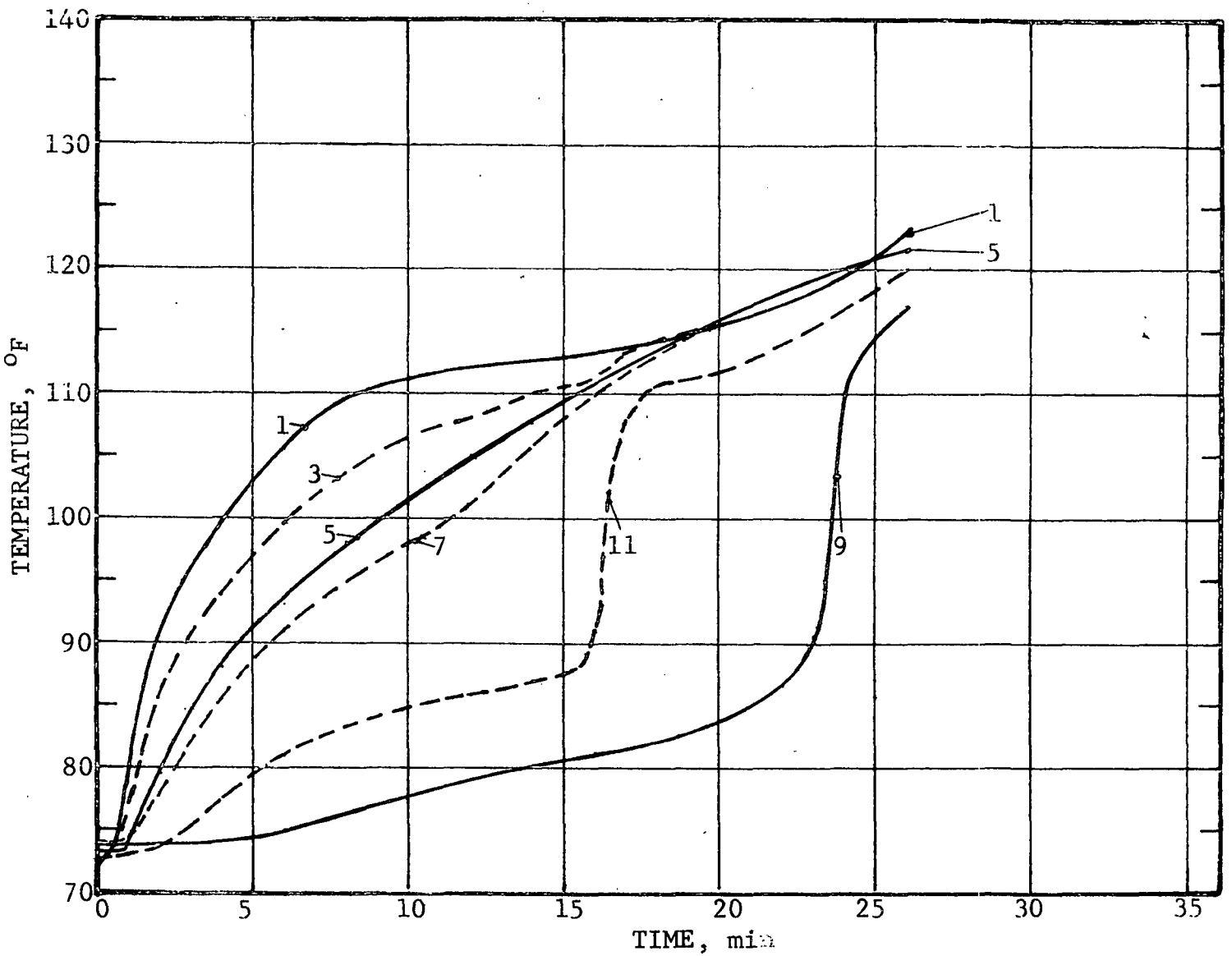


Figure 10 PCM and Surface Temperatures for 3/4 and 1/4 Inch Cell Locations for Melt Run 4

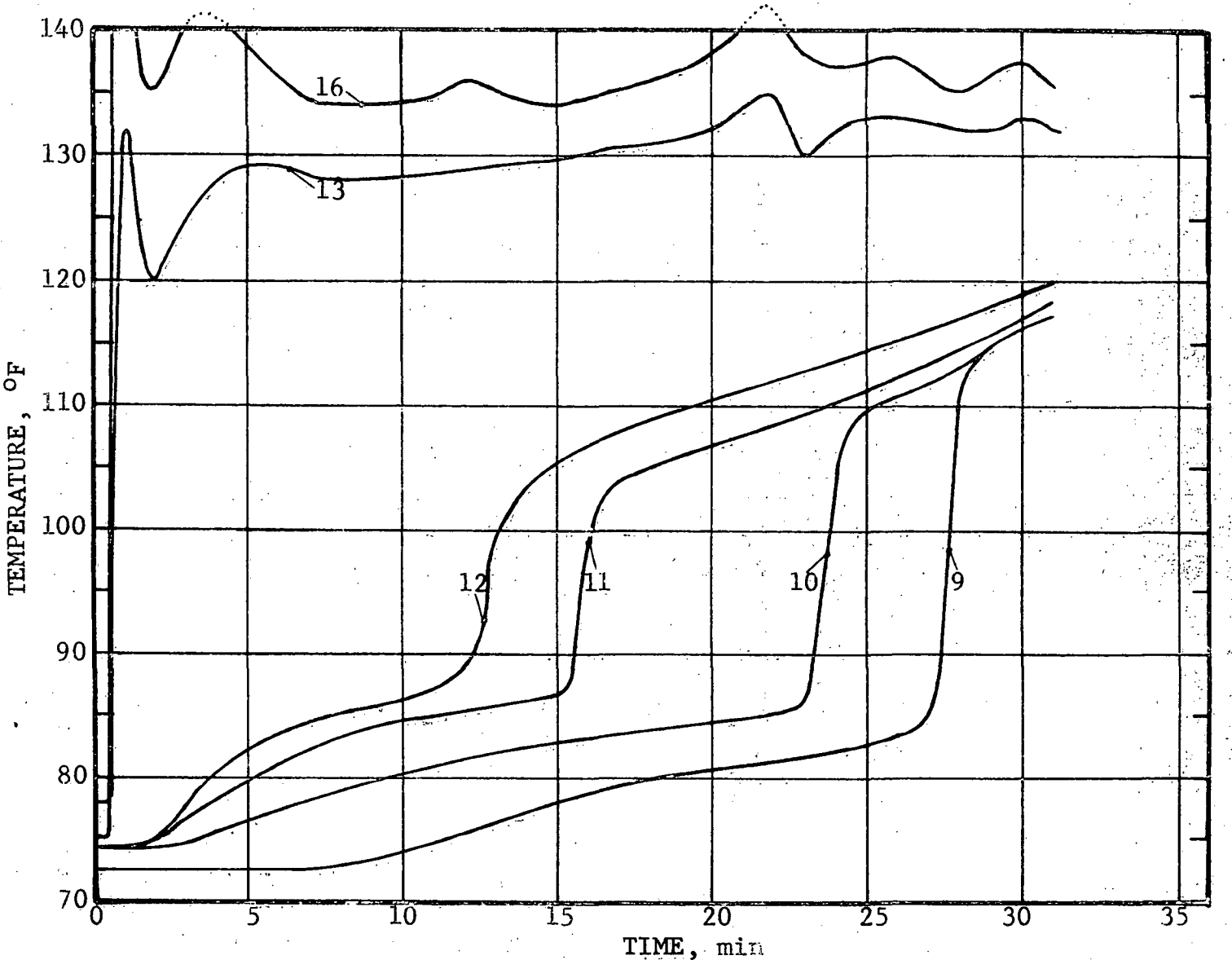


Figure 11 PCM and Coolanol Temperatures for Melt Run 10

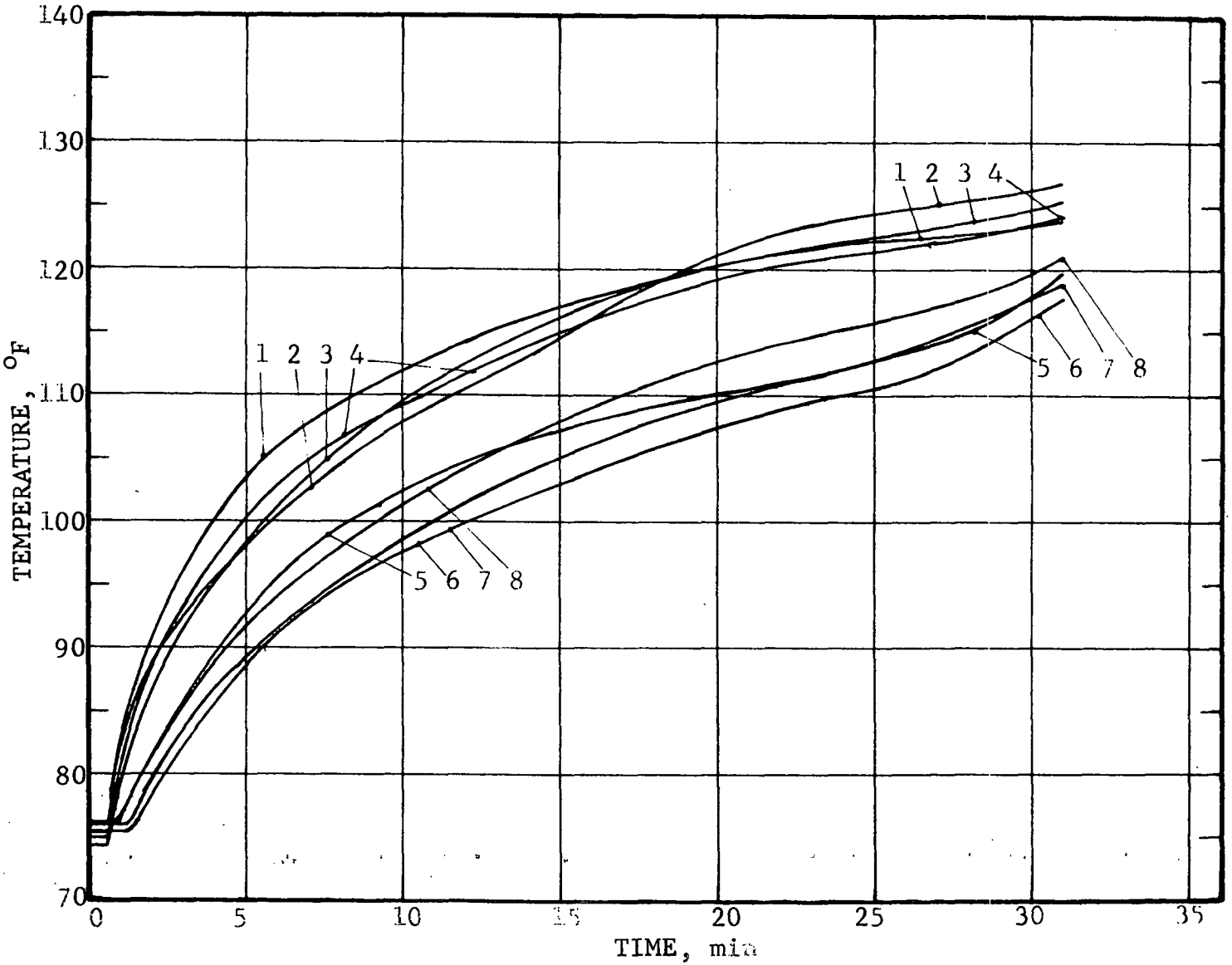


Figure 12 Surface Temperatures for Melt Run 10

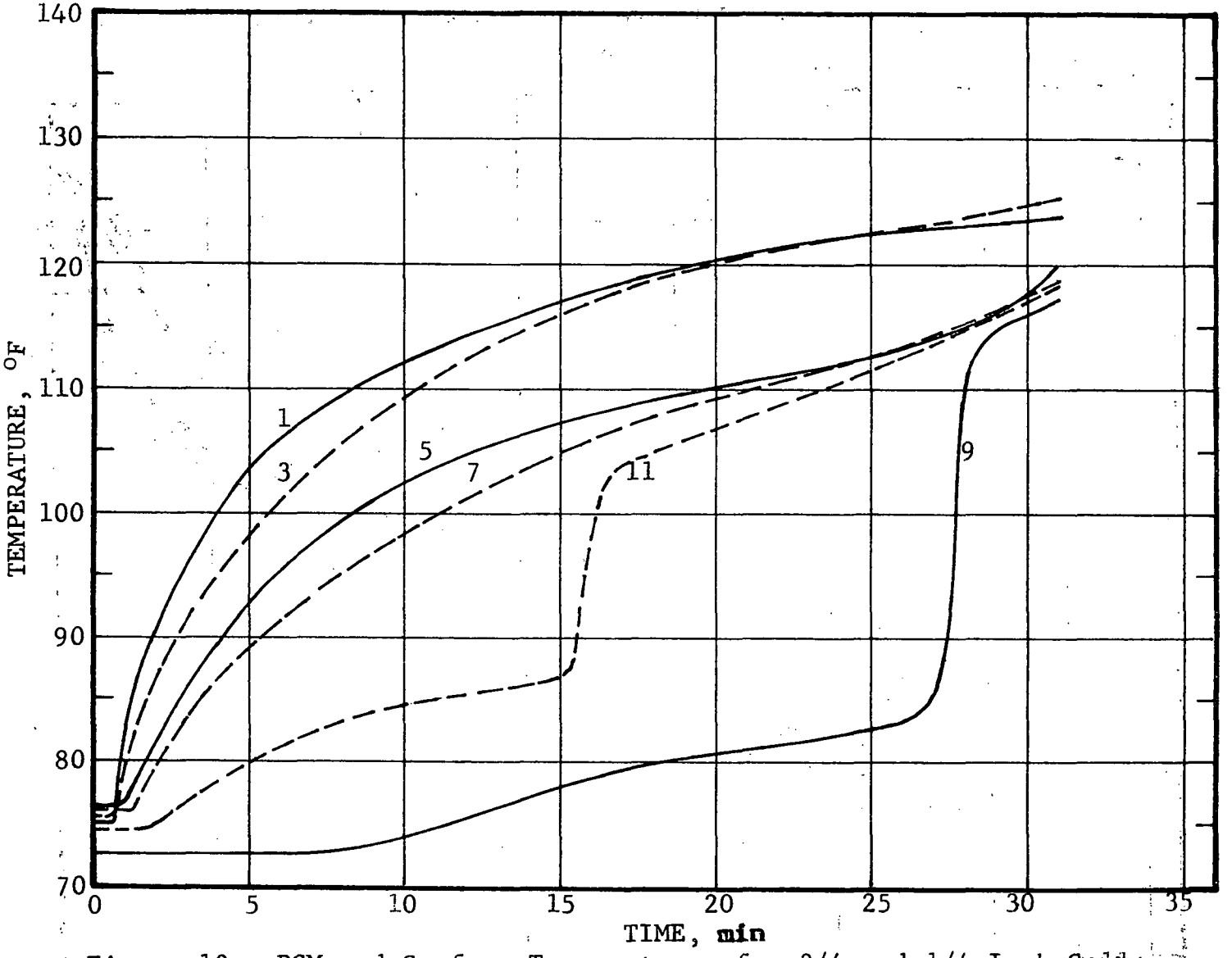


Figure 13 PCM and Surface Temperatures for 3/4 and 1/4 Inch Cell Locations for Melt Run 10

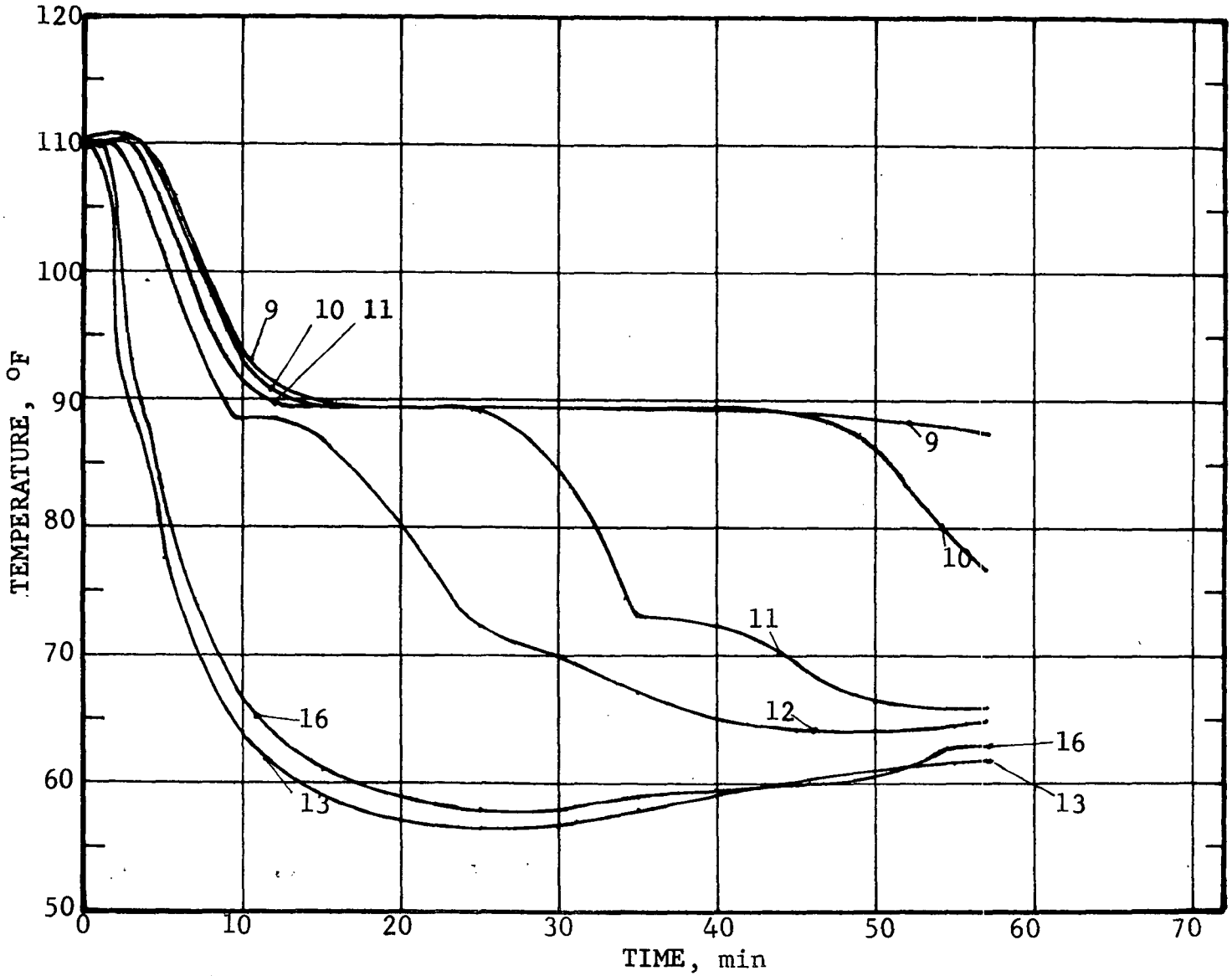


Figure 14 PCM and Coolanol Temperatures for Freeze Run 5

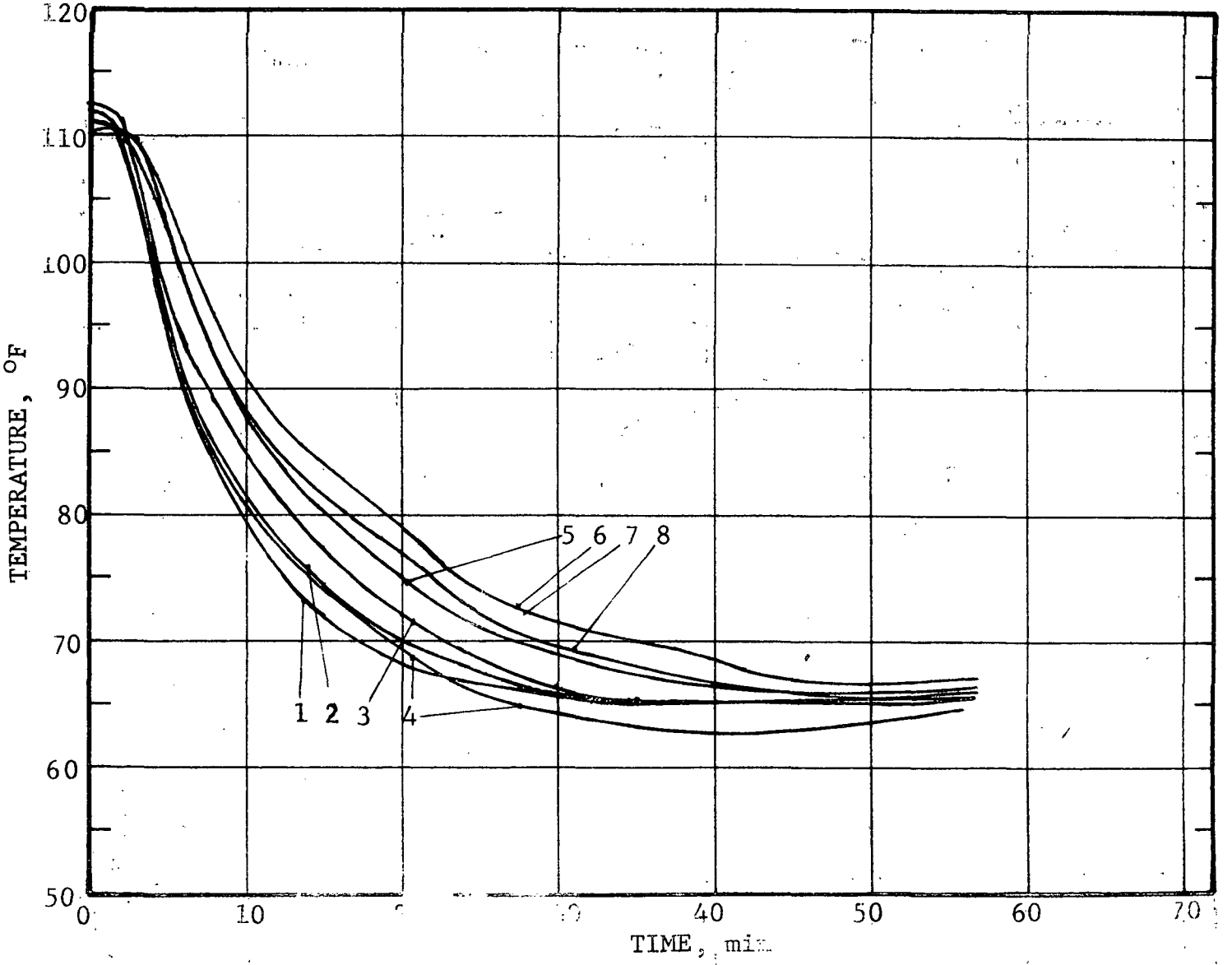


Figure 1[illegible] surfa[illegible]erat[illegible]s for Freeze Run 5

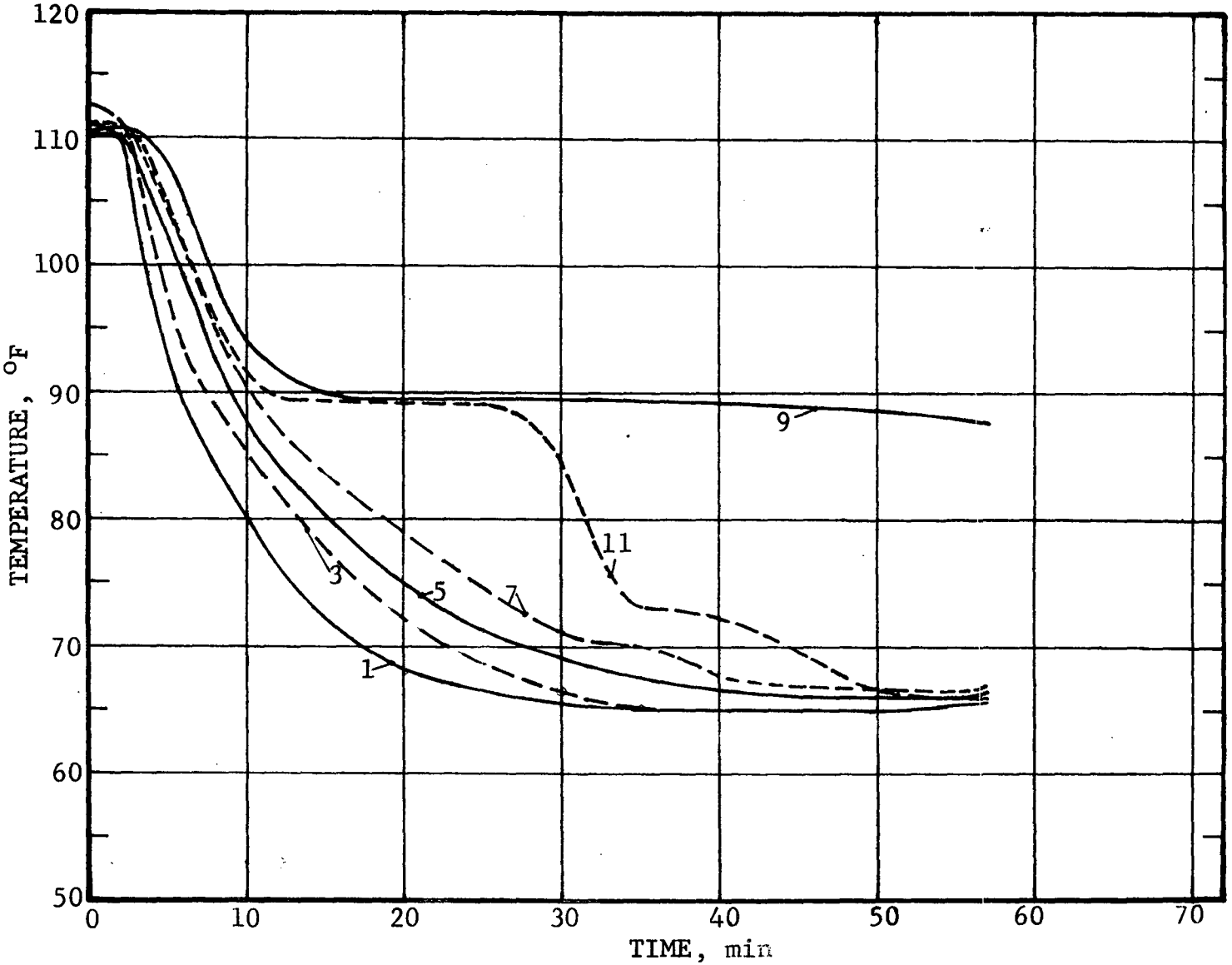


Figure 16 PCM and Surface Temperatures for 3/4 and 1/4 Inch Cell Locations for Freeze Run 5

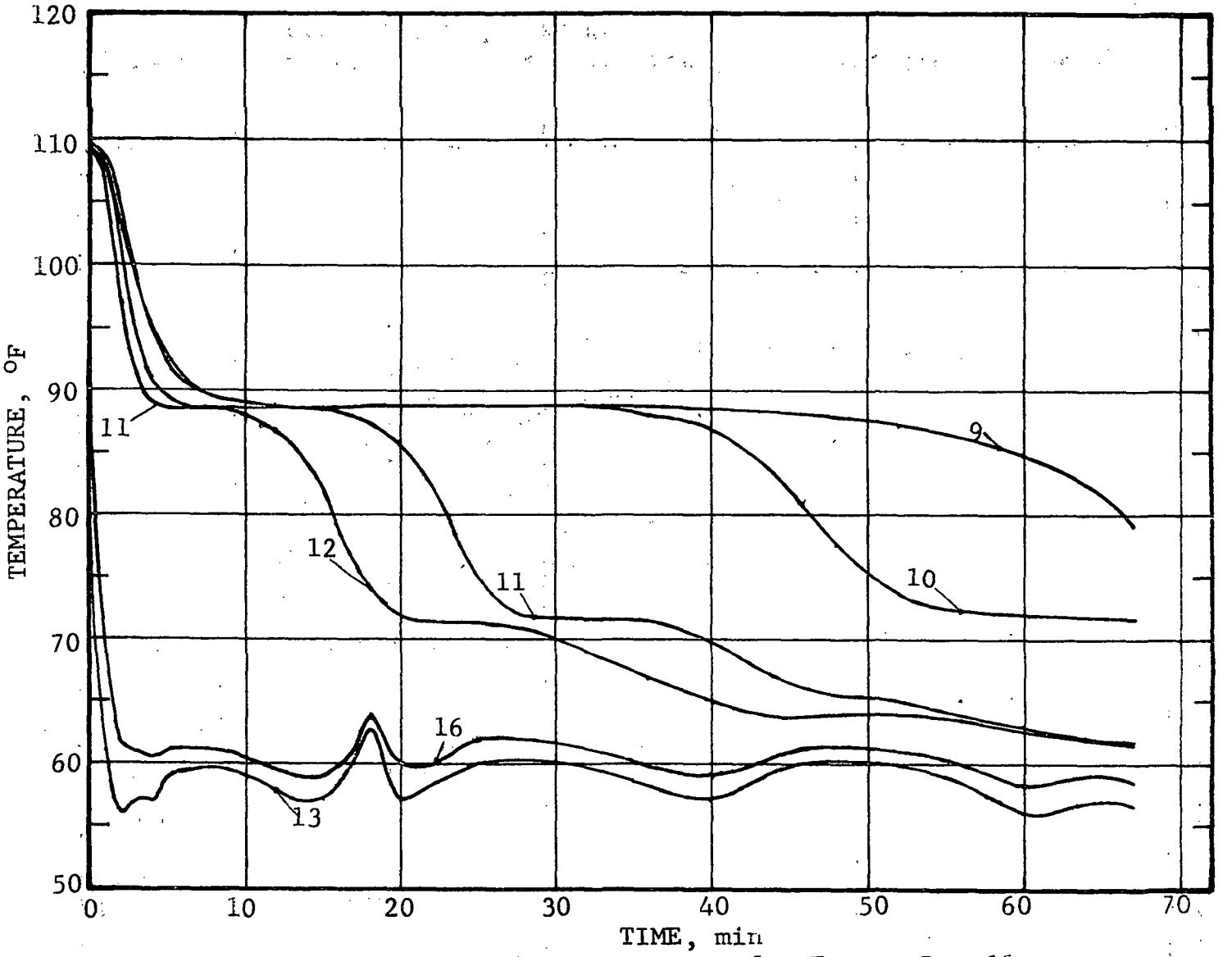


Figure 17 PCM and Coolanol Temperatures for Freeze Run 11

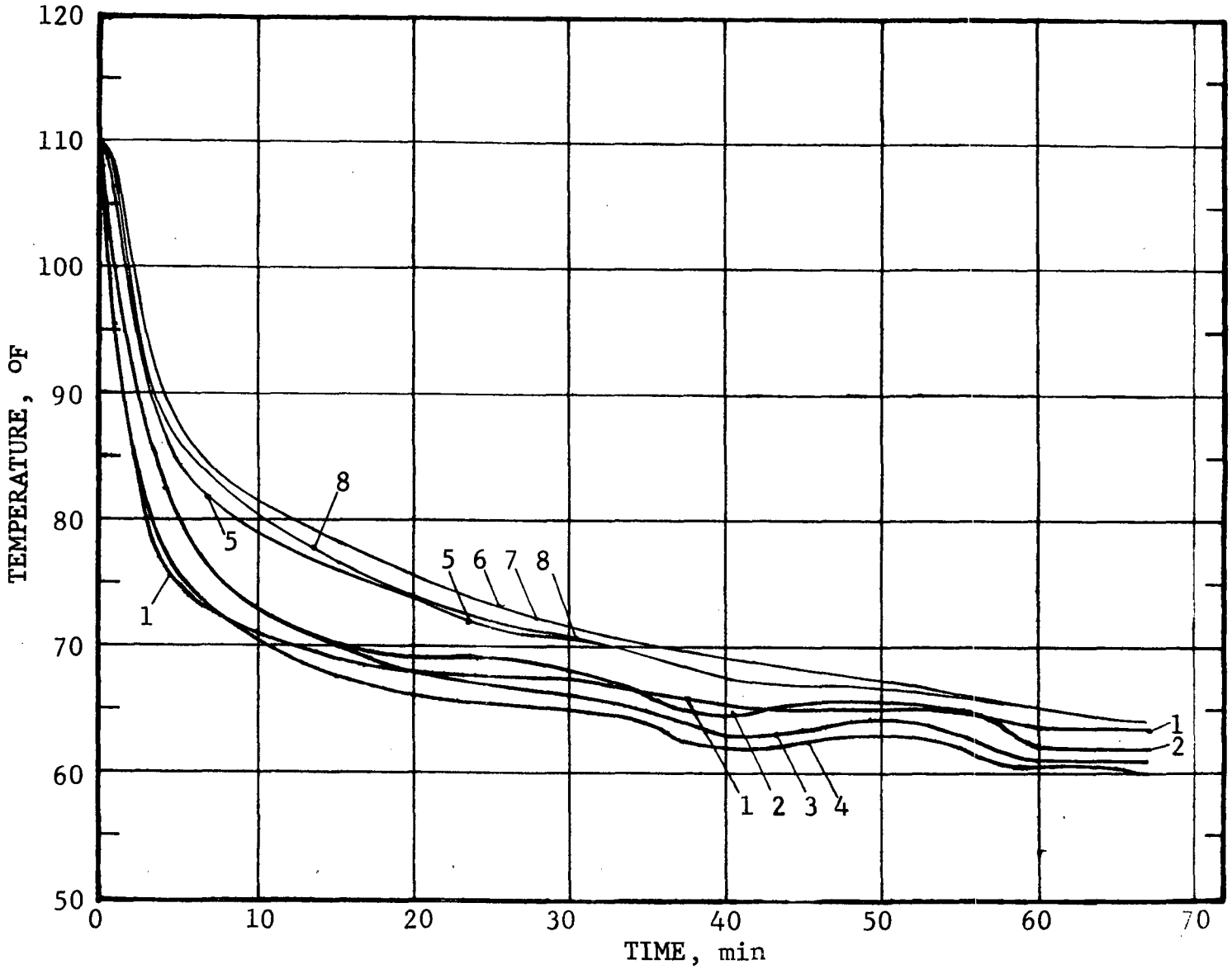


Figure 18 Surface Temperatures for Freeze Run 11

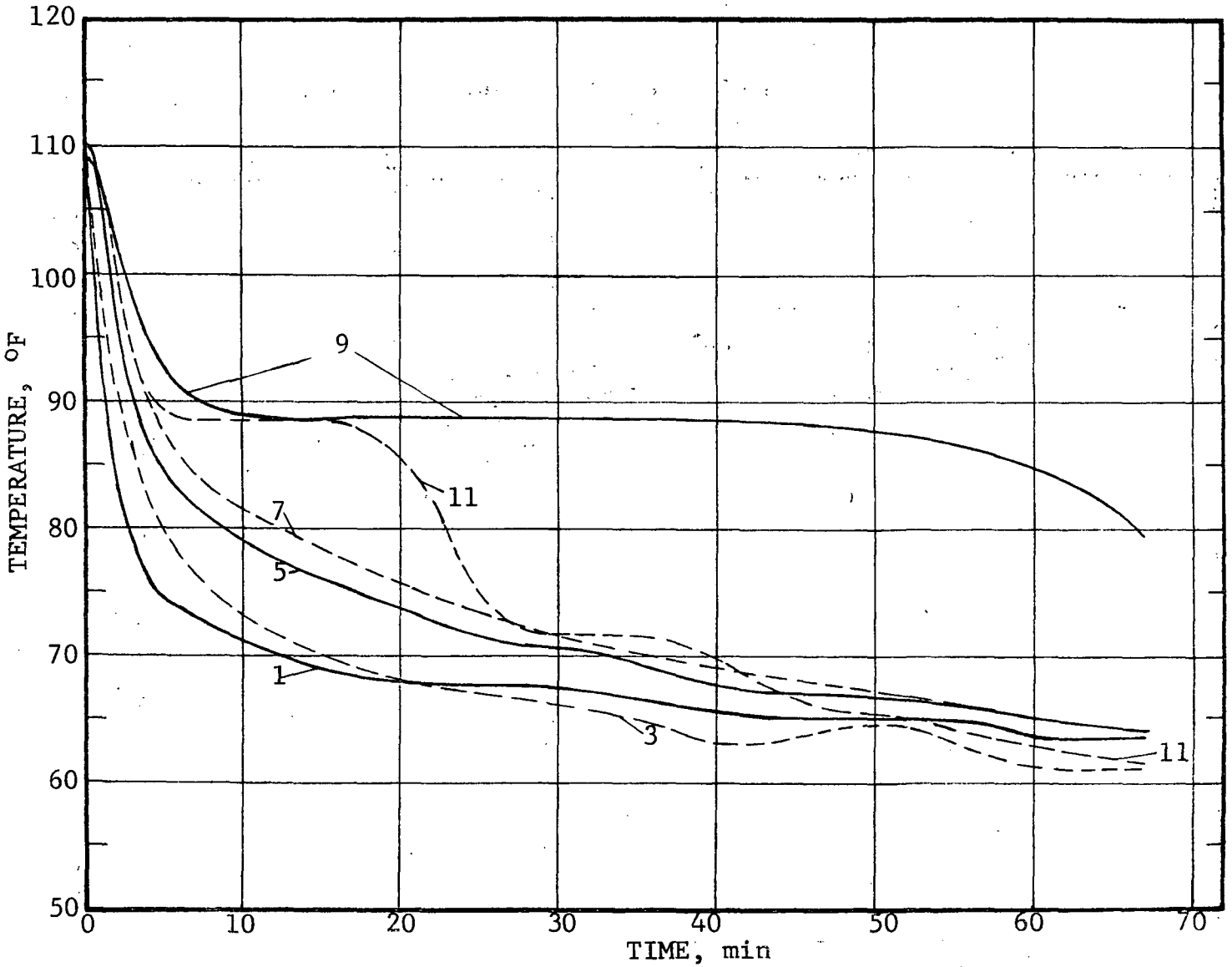


Figure 19 PCM and Surface Temperatures for 3/4 and 1/4 Inch Cell Locations for Freeze Run 11

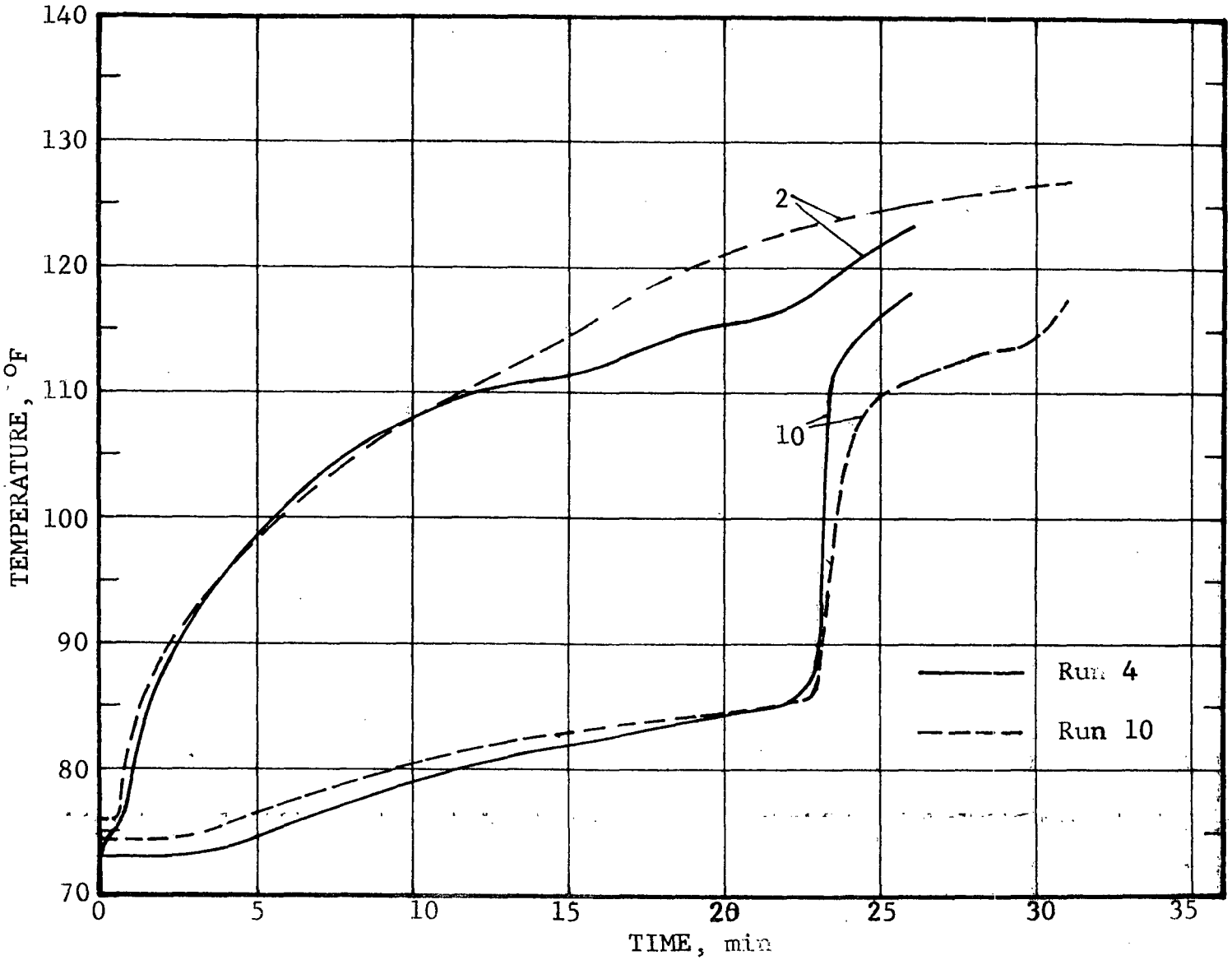


Figure 20 PCM and Bottom Surface Temperature Comparison Between Melt Runs 4 and 10 for 1/2 Inch Cell Location

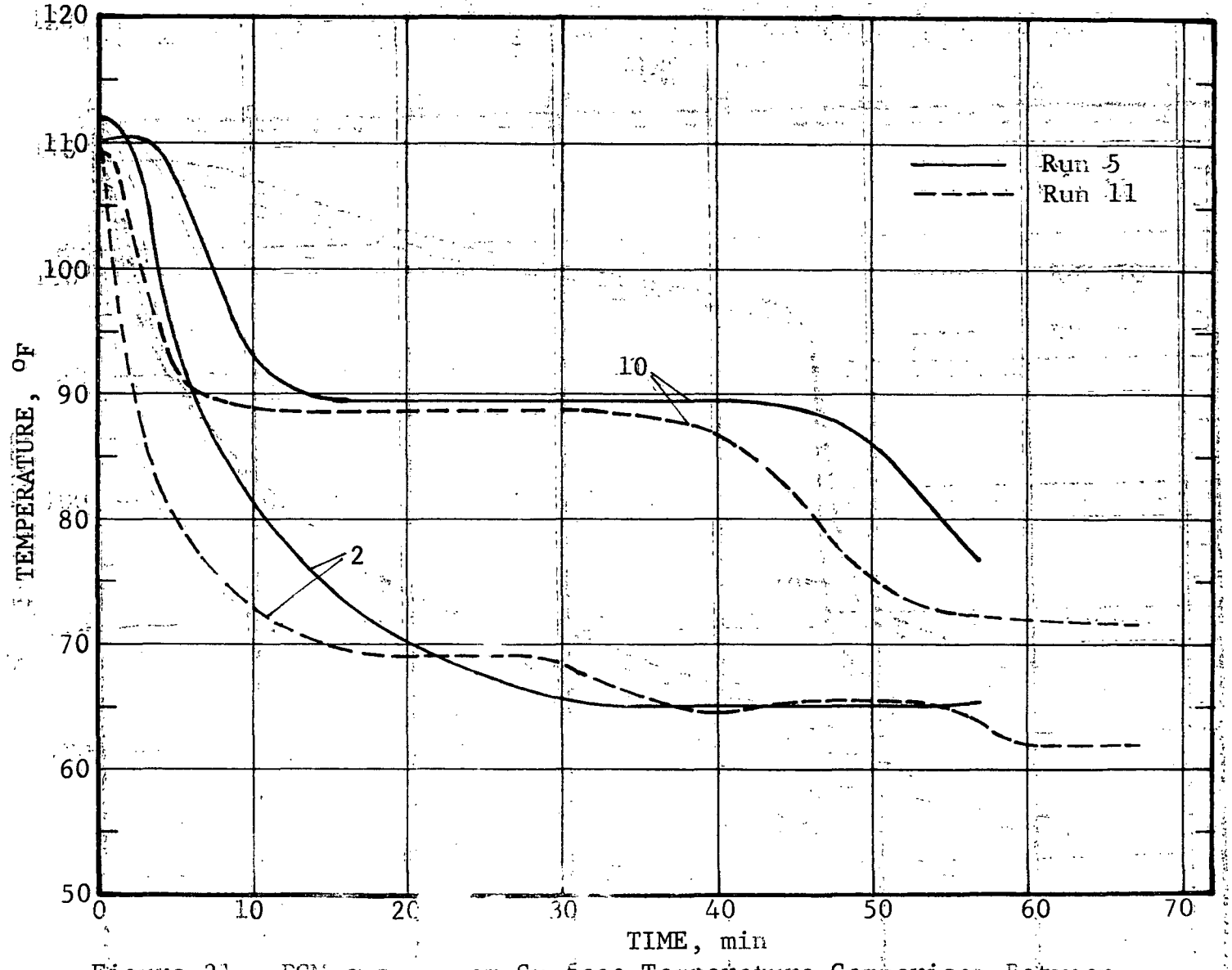


Figure 21. PCM and Room Surface Temperature Comparison Between Freeze Runs 5 and 11 for 1/2 Inch Cell Location

context of this discussion, referral to bottom and top surfaces are intended to correspond to Figure 4. The bottom surface implies that next to the coolanol flow passage. In the inverted orientation, this surface exists at the top of the arrangement.

A sequence of PCM photographs are shown for freeze run 14 in Figures 22 through 25, for melt run 2A in Figures 26 through 29, and for melt run 10 in Figures 30 through 37. In all cases, the coolanol flow is from the end with the 3/4 inch cells toward the end with 1/8 inch cells.

Discussion

Thermal inertia of the flow lines and equipment presented practical difficulty in rapidly acquiring high temperatures at the inlet to the capacitor. Consequently, most runs were conducted under the condition of a rising inlet temperature. The 160°F inlet temperature desired for melt runs 3 and 9 was not obtained. This problem could be overcome by installing a by-pass line around the capacitor and initially operating the system in a by-pass mode until the system temperature stabilizes at the desired level. The prime problem, however, which this difficulty presented to the desired objectives was that of making it difficult to operate normal and inverted configurations under identical conditions.

During the early freeze runs, several fins were observed to be bent. (See Figure 22) It is assumed that this occurred due to contraction of the housing which imposed a compression load on the very thin fins.

As melting progressed in the melt tests, the residual solid remaining in the larger cells eventually assumed a cylindrical shape having an elliptical cross section. These were completely surrounded by liquid. They were supported at the ends where there was a larger film of solid adhering to the plexiglass windows. This larger film blocked the view of the primary solid configuration within the cells. This blockage prevented the use of

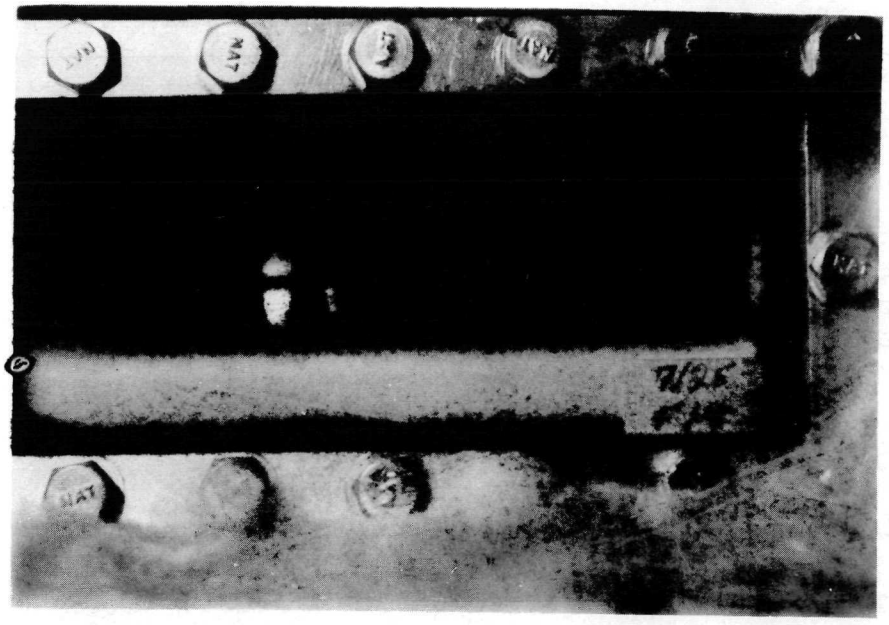


Figure 22 Freeze Run 14 Five Minutes After Start

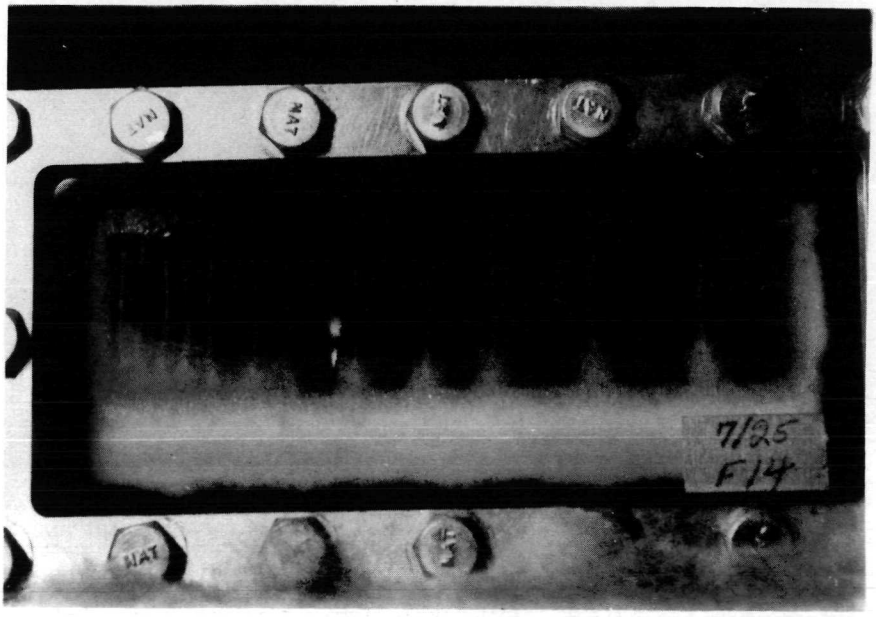


Figure 23 Freeze Run 14 Ten Minutes After Start

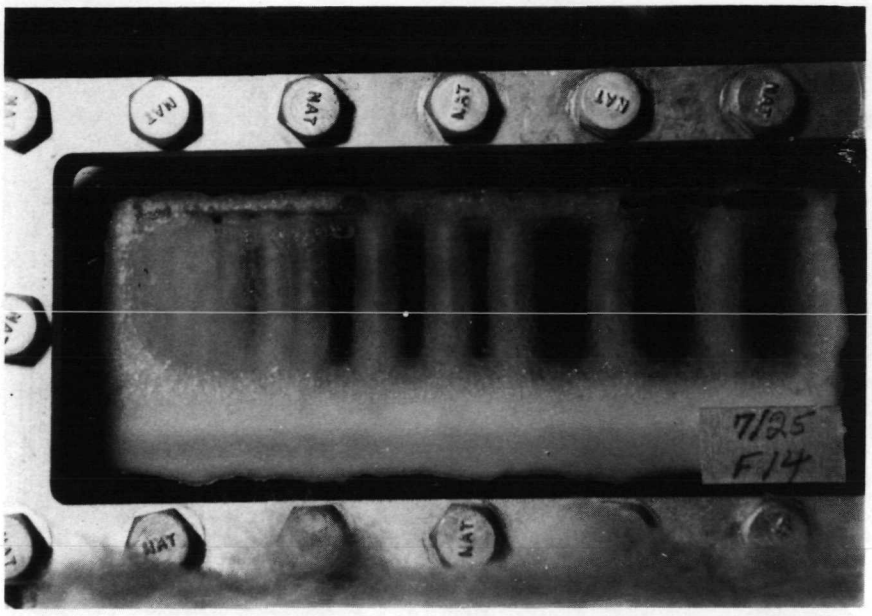


Figure 24 Freeze Run 14 Sixteen Minutes After Start

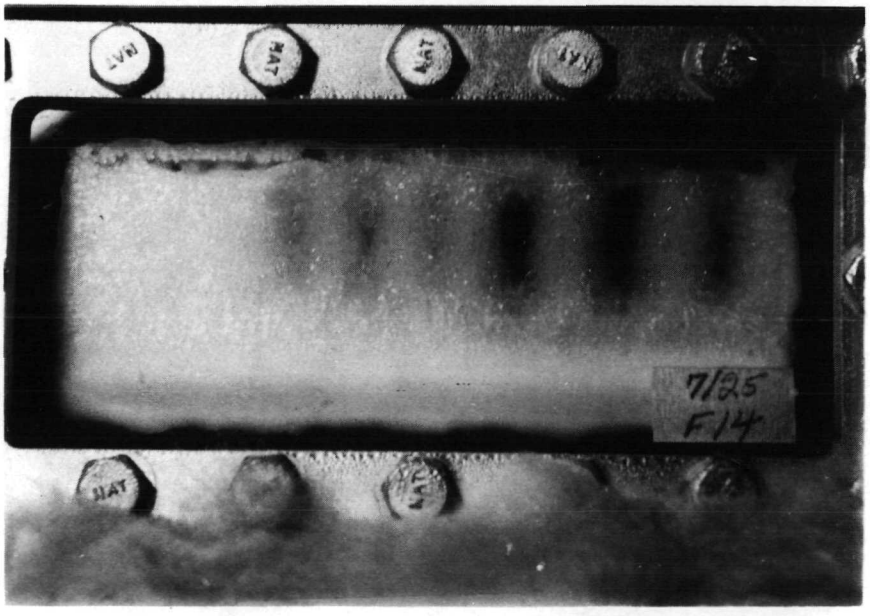


Figure 25 Freeze Run 14 Twenty Eight Minutes After Start

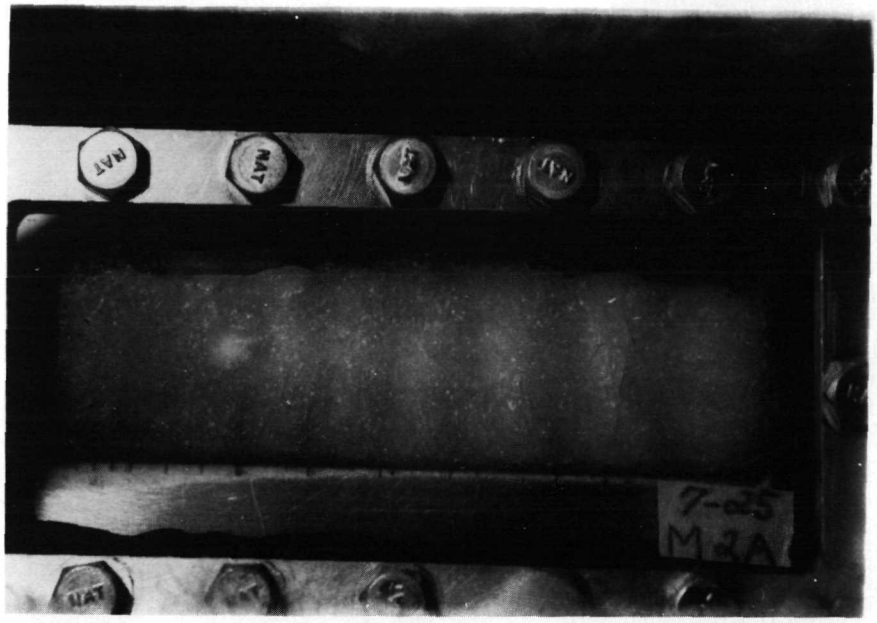


Figure 26 Melt Run 2A Five Minutes After Start

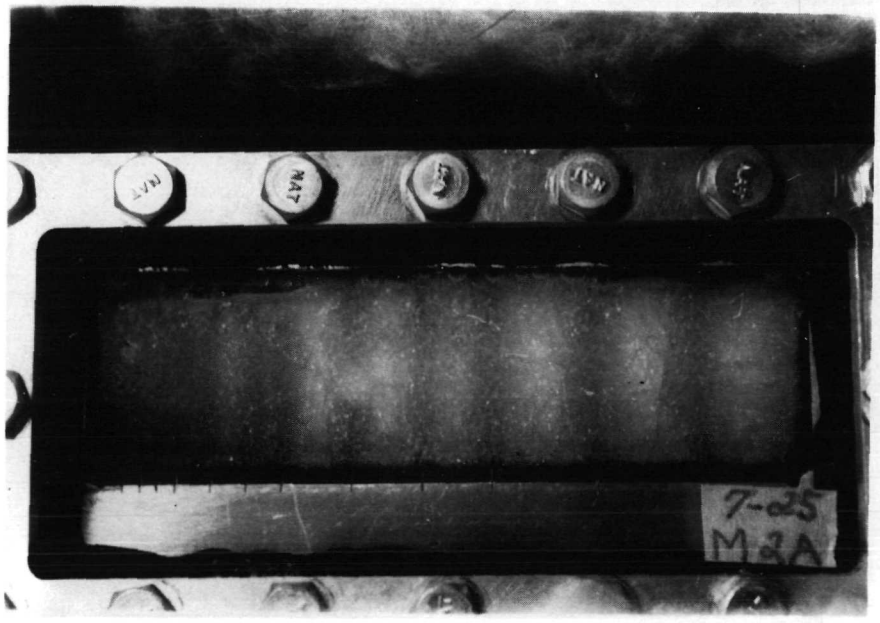


Figure 27 Melt Run 2A Eleven Minutes After Start

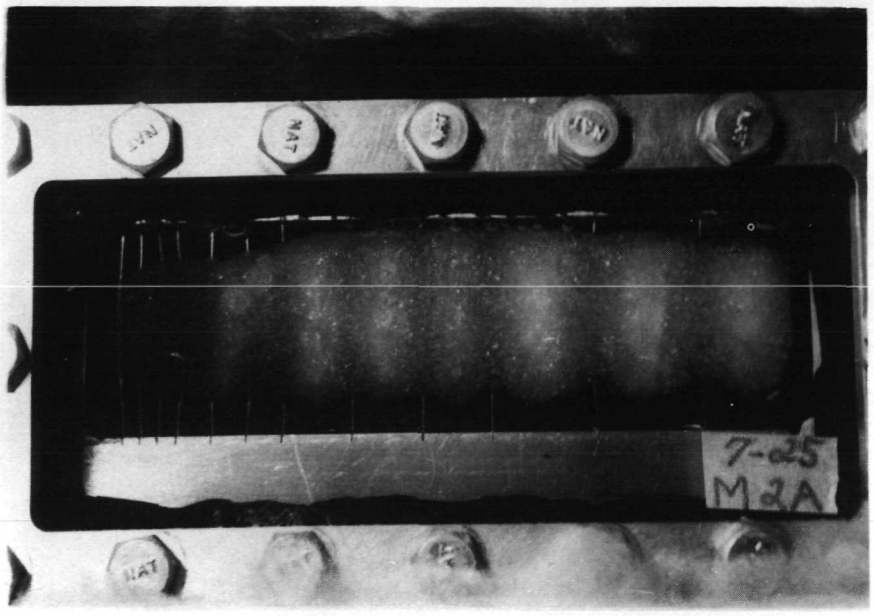


Figure 28 Melt Run 2A Seventeen Minutes After Start

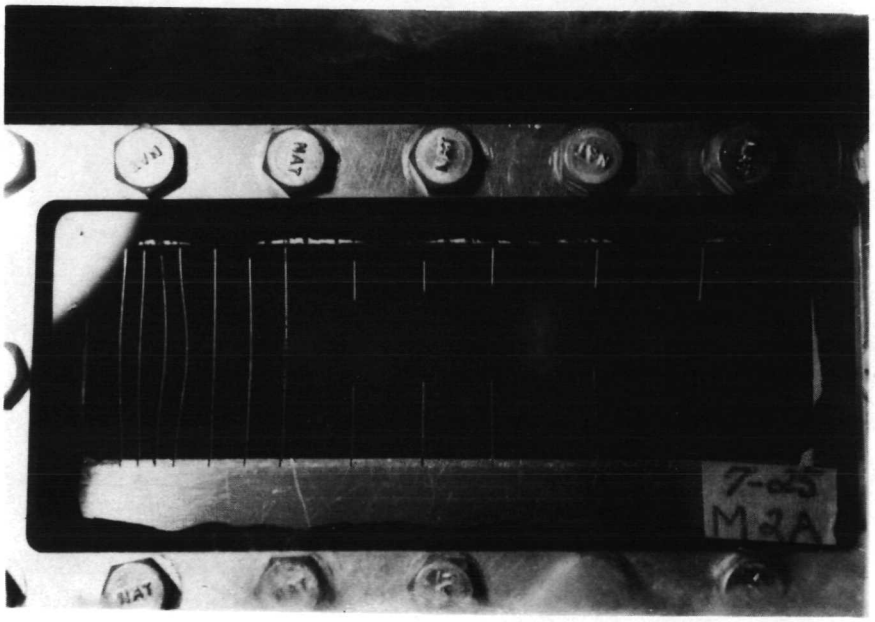


Figure 29 Melt Run 2A Twenty Three Minutes After Start

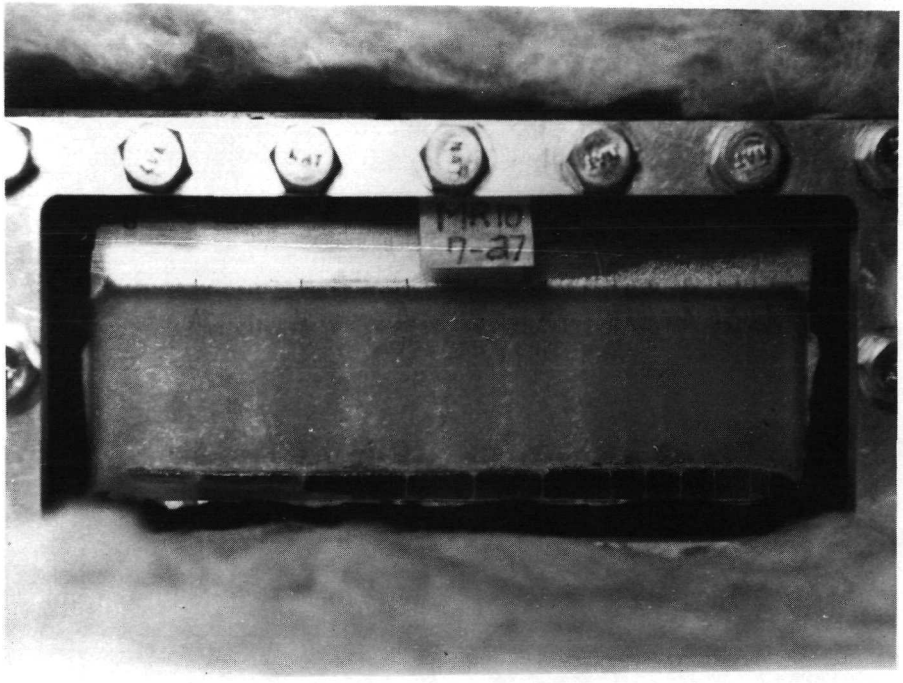


Figure 30 Melt Run 10 Four Minutes After Start

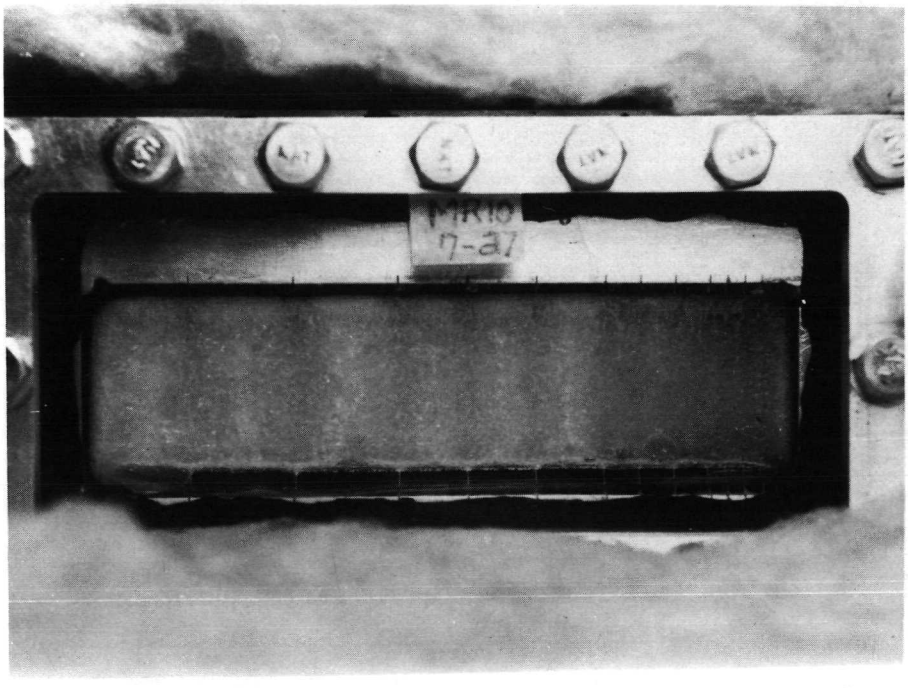


Figure 31 Melt Run 10 Eight Minutes After Start

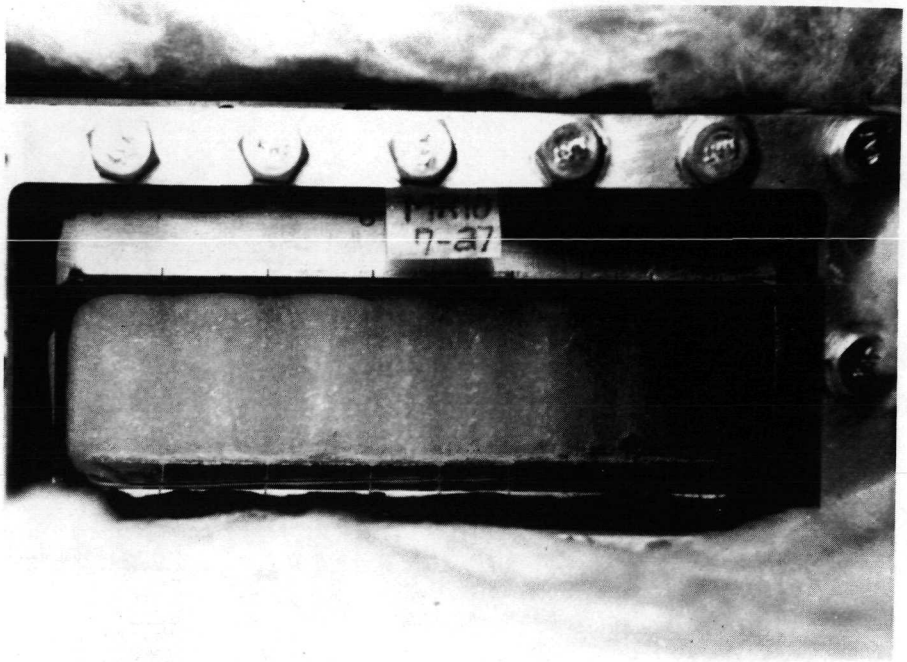


Figure 32 Melt Run 10 Ten Minutes After Start

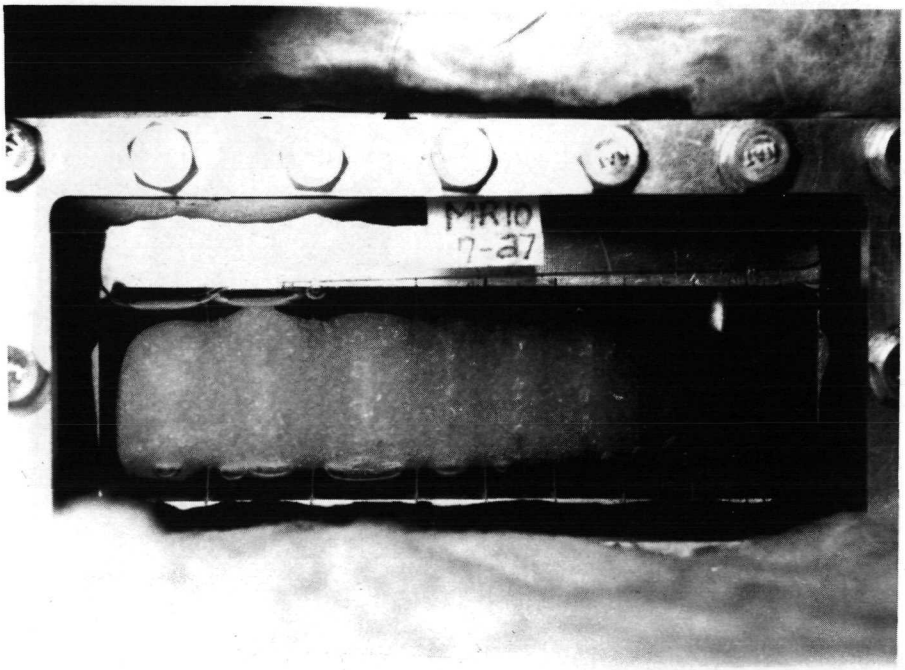


Figure 33 Melt Run 10 Fourteen Minutes After Start

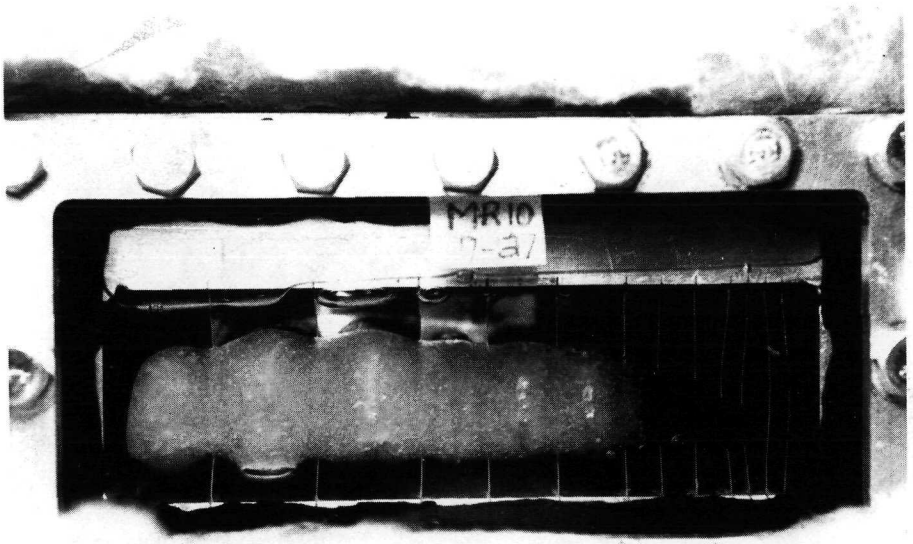


Figure 34 Melt Run 10 Nineteen Minutes After Start

Figure 35 Melt Run 10 Twenty One Minutes After Start

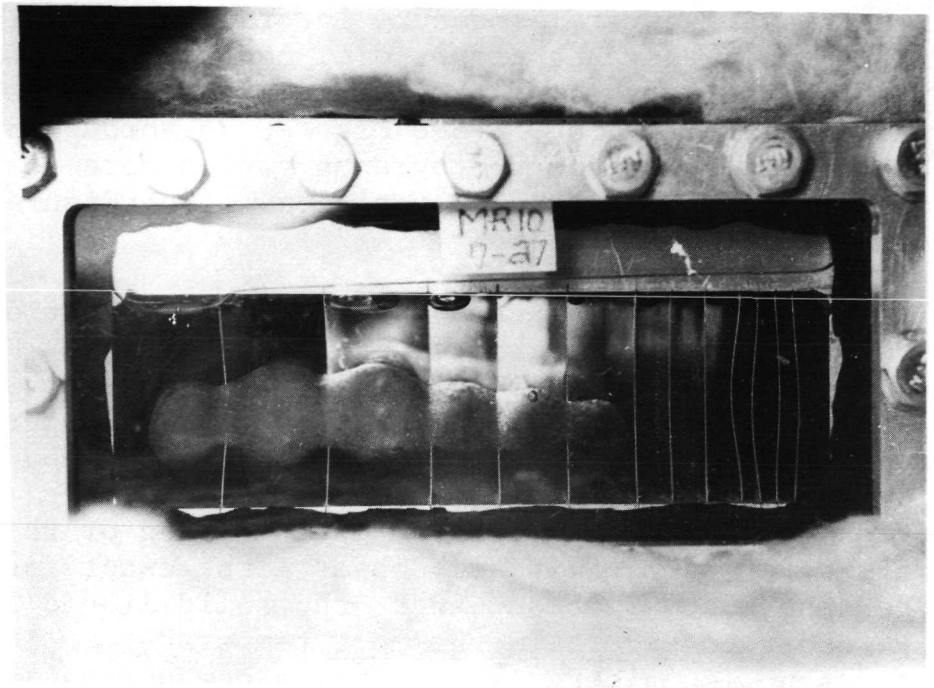


Figure 36 Melt Run 10 Twenty Four Minutes After Start

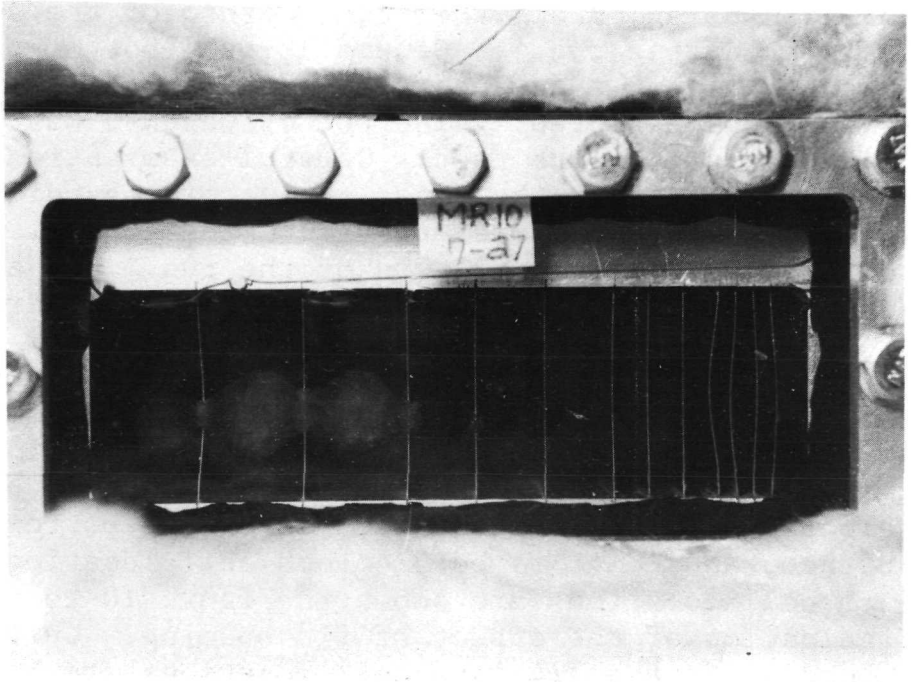


Figure 37 Melt Run 10 Twenty Six Minutes After Start

photographs to accurately measure interfacial position for the melt runs. The photographs did, however, yield good qualitative indications of the melting process and the influence of fin spacing. (See Figures 26 through 37) For the freeze runs, the photographs gave a more reliable indication of the interfacial position. (See Figures 22 through 25)

Melt run 4 was conducted in a normal orientation with the coolanol flow at the bottom of the capacitor. Figure 8 shows the coolanol inlet and outlet temperatures and the four PCM temperatures. The rapid rise in PCM temperatures occurred as the solid melted past the probe thereby exposing it to the hotter liquid. In almost every case, this occurred simultaneously with complete melting of the cell. Attempts were made to visually observe the exact instant when the thermocouple bead was at the solid-liquid interface in order to ascertain the melt temperature. This effort was not considered sucesssful since definition of when the bead was just at the interface was questionable within a time period comparable to that during which the temperature rose from below the reported melt temperature to one well above the reported value. Figure 9, while somewhat confusing, shows the relationship of the surface temperatures during run 4. Generally, the bottom plate was slightly higher than a corresponding location on the top as would be expected. Also the difference was higher in the early part of the test. Careful examination of the responses reveals that the phase change process did influence the surface temperatures since the spatial variation shifted during the duration of the run. For example, thermocouple number 1 shows the highest reading early as would be expected since it is nearest the coolanol inlet. The reading of thermocouple number 4, however, crosses that of number 1 and becomes higher at a time corresponding to complete melting in the region of the 1/8 inch cell. Similar observations about the other locations can be made upon close examination. Temperatures of the bottom surface, top surface, and PCM for the 3/4 inch cell locations and 1/4 inch cell locations are shown in Figure 10 to provide some indication of the effect of fin spacing. Curves corresponding to the 1/4 inch cell locations are broken.

Similar results for melt run 10 which was conducted in an inverted orientation are shown in Figures 11 through 13. There are two apparent differences. A slightly longer time delay between melting of the 1/2 inch and 3/4 inch cells existed for the inverted position than in the normal position. Also there was a larger temperature difference throughout the duration of the test between the two surfaces for the inverted case. Compare Figures 9 and 12. Direct comparisons of temperatures of the heated surface and PCM for 1/2 inch cell locations between melt runs 4 and 10 are shown in Figure 20. The comparison does not show much difference. Run 10 was conducted later in the series of tests than run 4. As the PCM was melted and frozen several times, the void occurring upon melting seemed to increase. The liquid PCM was frozen in the normal position prior to melt run 10. Upon inverting the capacitor for run 10, this void existed at the bottom as is evidenced by Figure 30.

Freeze run 5 was conducted in the normal orientation. Temperature responses for the same locations and comparisons discussed above are shown in Figures 14 through 16. Freeze run 11 was conducted in the inverted configuration. Similar results are shown in Figures 17 through 19. Figure 21 provides some direct comparisons of PCM and cooled surface temperatures corresponding to the 1/2 inch cell locations between freeze runs 5 and 11. The dashed lines correspond to run 11. The coolanol temperature dropped faster in run 11 than it did in run 5.

All freeze runs are characterized by a significant plateau in the outputs of the PCM thermocouples. The length of the plateau is shorter for the narrow cells than it is for the wider ones. Figure 21 indicates that between run 5 and run 11 the freeze temperature appeared to have shifted by about 1 F^{o}. Figures 14 and 17 also show some indications of a less pronounced plateau around 71.5 ^{o}F. This probably corresponds to a solid phase transition. The reported phase transition for nonadecane is 73.04 ^{o}F.

While there are some variations between the runs compared above, there appeared to be no pronounced differ-

ences. Since the total PCM housing was a good heat conducting path, the heat transfer occurred from both surfaces. This characteristic tends to overshadow the advantages of inverting the unit. Further reduction and evaluation of the remaining data will be continued.

NUMERICAL STUDY

Attention last summer as well as this summer has been devoted to a numerical analysis of the heat transfer and solid-liquid interfacial progression inside a PCM cell such as that illustrated in Figure 3a. The goal of the effort is development of a method to facilitate a parametric study encompassing a range of fin spacings, fin thickness, etc... . To be practical, the method should not require excessive computer time.

Last summer, a computer program was written for a nodal network applicable to a symmetrical half-section of a two-dimensional cell. The method utilized an explicit finite difference formulation. The phase change was handled by keeping an accumulative record of the stored energy for each node. When the node was undergoing phase change, the prediction technique was overridden by holding the temperature at the fusion value. For melting cases, an effective thermal conductivity based on the one-dimensional correlations of Reference 7 was used to account for convection. Physical and boundary conditions were used that corresponded to some existing experimental data. Predictions of interfacial position agreed well with the experimental data for melt tests. Since experimental fin temperatures were input as boundary conditions, general extension of the method to a parametric study is limited. The method could be modified to allow inclusion of fin temperatures as unknowns at the expense of excessive computer time. This penalty is imposed by the stability requirement inherent with an explicit formulation.

Consider the two-dimensional nodal arrangement shown in Figure 38. There are M x N PCM nodes. The letter I

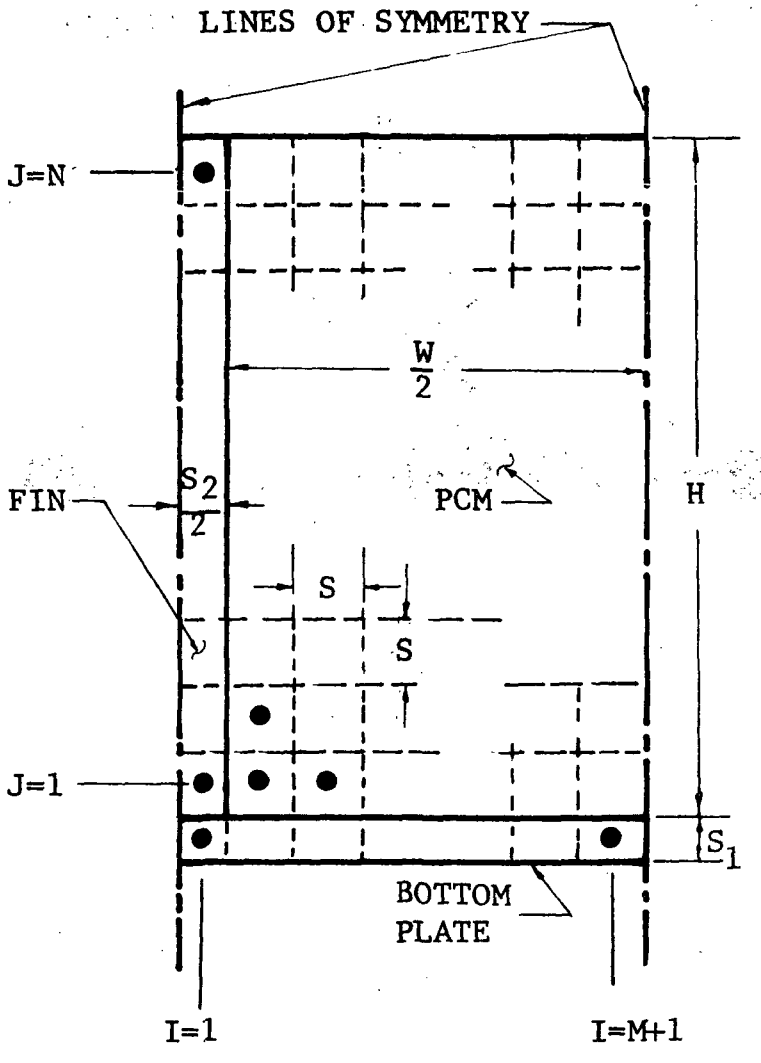


Figure 38 Two-Dimensional Nodal Network Used in the Numerical Study

designates the vertical column in which a node lies while J designates the horizontal row. Figure 39 represents a general node (I,J), its surrounding nodes, and the notational scheme for interconnecting thermal conductances.

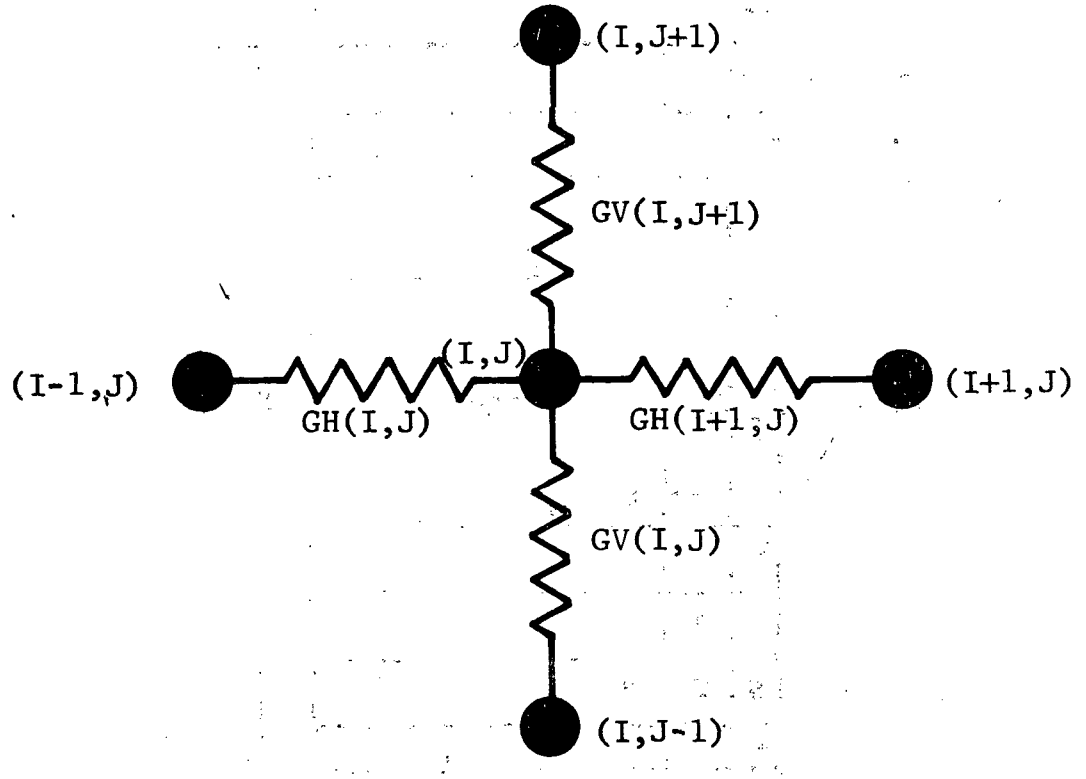


Figure 39 General Designations for Nodes and Connecting Conductances

The two-dimensional transient heat conduction equation for an isotropic, homogeneous medium in the absence or sources, sinks , or phase change is

$$\alpha\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) = \frac{\partial T}{\partial t} \qquad (1)$$

For numerical solutions of equation (1), the partial derivatives are approximated by finite difference analogs. A forward difference analog for the time derivative and central difference analogs for the spatial derivatives can by used. With regard to the latter, there are two possibilities for a transient problem (Ref.12). If all temperatures appearing in the spatial analogs are evaluated at the old time level, the formulation is explicit, and the resulting equation for a node involves only one unknown at the new time level. The solution for a system of nodes by this method is fairly simple and straightforward. If all temperatures appearing in the spatial analogs are evaluated at the new time level, the formulation is implicit. The resulting equation for a node involves five unknown temperatures at the new time level. Solution for a system of nodes involves solving a set of simultaneous equations equal to the number of nodes.

Physical arguments can be used to formulate the difference analogs discussed above. The summation of heat transfer to node (I,J) from all surrounding nodes should equal the product of the nodal capacitance and the difference in its old and new temperature. With reference to Figure 39, the explicit formulation based on this physical approach is

$$GH(I,J)[T(I-1,J) - T(I,J)] + GH(I+1,J)[T(I+1,J) - T(I,J)]$$
$$+ GV(I,J)[T(I,J-1) - T(I,J)] + GV(I,J+1)[T(I,J+1) - T(I,J)]$$
$$= \frac{C(I,J)[T'(I,J) - T(I,J)]}{\Delta t} \qquad (2)$$

The implicit formulation is

$$GH(I,J)[T'(I-1,J) - T'(I,J)] + GH(I+1,J)[T'(I+1,J) - T'(I,J)]$$
$$+ GV(I,J)[T'(I,J-1) - T'(I,J)] + GV(I,J+1)[T'(I,J+1) - T'(I,J)]$$
$$= \frac{C(I,J)[T'(I,J) - T(I,J)]}{\Delta t} \qquad (3)$$

The primes denote temperatures at the new time level.

Numerical solution of the explicit formulation given by equation (2) is subject to a restriction on the allowable time step given by

$$\Delta t \leq \frac{C(I,J)}{GH(I,J) + GH(I+1,J) + GV(I,J) + GV(I,J+1)} \quad (4)$$

For a nodal arrangement, the node having the smallest time step given by equation (4) governs the entire solution. Inclusion of fin temperatures as unknowns presents a severe time restriction. The thermal capacitance of a fin node is low while the vertical conductances along the fin are large. Both factors reduce the allowable time step. The governing fin node in Figure 38 is (1,1) and the governing PCM node is (2,1). As an illustrative example, consider nonadecane as the PCM (k=0.087 Btu/H_r-ft-F, ρ=47.2 lb_m/ft^3, C_p=0.5 Btu/lb_mF), aluminum fin (k=93 Btu/Hr-ft-F, ρ= 17[illegible] lb_m/ft^3, C_p=0.22 Btu/lb_mF), a cell geometry with W = 0.[illegible] inch and S_2 = 0.008 inch, and M = 4. For these numbers the time step given by equation (4) for node (1,1) is 0.96×10^{-5} Hr while that for node (2,1) is 2.8×10^{-3} Hr.

The implicit formulation is not restricted by a stability requirement, but the solution is not as simple. In general, it requires a matrix inversion routine and may be impractical for a large number of nodes.

Reference 13 has outlined an alternate method known as the alternating-direction-implicit method (ADI). The method involves writing the analogs to the space derivatives explicitly in one direction and implicitly in the other. This set of equations are solved for one time step, then the formulations are reversed. This second formulation is then solved for one time step and the resulting solution is taken to be correct. As an example pertinent to Figure 38, an implicit formulation in the horizontal direction (J = constant) and an explicit formulation in the vertical direction (I = constant) is

$$GH(I,J)[T'(I-1,J) - T'(I,J)] + GH(I+1,J)[T'(I+1,J) - T'(I,J)]$$
$$+ GV(I,J)[T(I,J-1) - T(I,J)] + GV(I,J+1)[T(I,J+1) - T(I,J)]$$
$$= \frac{C(I,J)[T'(I,J) - T(I,J)]}{\Delta t} \qquad (5)$$

When equation (5) has been written for all nodes along a horizontal row (J=constant), the system of equations constitute a tridiagonal coefficient matrix and the solution can readily be obtained by means of the Thomas algorithm described in Reference 13. The implicit formulation in the vertical direction and explicit in the horizontal direction is

$$GH(I,J)[T(I-1,J) - T(I,J)] + GH(I+1,J)[T(I+1,J) - T(I,J)]$$
$$+ GV(I,J)[T'(I,J-1) - T'(I,J)] + GV(I,J+1)[T'(I,J+1) - T'(I,J)]$$
$$= \frac{C(I,J)[T'(I,J) - T(I,J)]}{\Delta t} \qquad (6)$$

The ADI method consists of either solving equation (5) successively for each horizontal row and then reversing directions and successively solving equation (6) for each vertical row or vice-versa. For the first case, time is advanced by one time increment and equation (5) is successively solved for each horizontal row. The temperatures thusly predicted are not considered correct; they, however, are next used as old values in solving equation (6) successively for each vertical row. At this time, the resulting temperatures are considered correct. This alternating scheme is then continuously repeated to achieve the desired solution. It is stated in Reference 13 that this method has been shown to be stable for any ratio of time increment to space increment as long as the same time increment is used for the successive applications of equations (5) and (6).

Conceptually, the ADI method appeared attractive for the numerical problem considered in this work. Consequently, some effort during this summer has been devoted to studying its application. Pursuit along this line has been exploratory since the discussion of the method in Reference 13 pertained to application for a homogeneous and isotropic medium. The method was formulated for the arrangements depicted in Figures 38 and 39 and a Fortran computer program was written to facilitate computations.

For the computations, the model used was that shown in Figure 38. The upper surface was assumed to be adiabatic. Initially all temperatures were assumed to be uniform (72 ^{o}F or 73.5 ^{o}). The system was then perturbed by stepping the bottom surface temperature T_s to 110 ^{o}F. The physical properties and geometry were the same as those listed above in connection with the stability requirements.

For the preliminary exploration of the method, the problem was treated at first solely as a conduction problem without including any technique for handling the phase change. Since there appeared to no stated preference as to which direction should be solved first in the alternating scheme, solutions were examined for both cases.

After several computer runs, it became evident that the ADI method applied to the fin-PCM problem was not independent of the time increment. Some predictions for nodes (1,1) and (2,1) are compared in Table 2. For a time step of 0.001 Hr , the PCM temperatures for node (2,1) are obviously in serious error for the case when the method solves for unknowns in a horizontal row first. Also, the fin temperatures for node (1,1) demonstrate some overshoot. When the method solves for unknowns in a vertical row first, the PCM temperatures of node (2,1) appear reasonbly accurate. Those for the corner fin node (1,1) still demonstrate some overshoot. The program was also run with a time increment of 10^{-5} Hr. which is almost adequate to satisfy the stability requirement imposed by the explicit formulation as governed by node (1,1). For this small time increment, the results are not dependent on the direction pursued first. Consequently, these temperatures are considered to

be correct

Table 2 Comparisons of Some Numerically Predicted Values

ADI METHOD t= 0.001 Hr.					
Horizontal First			Vertical First		
TIME Hr.	T(1,1) °F	T(2,1) °F	TIME Hr.	T(1,1) °F	T(2,1) °F
0.002	98.39	233.53	0.002	103.09	82.18
0.004	106.34	212.01	0.004	110.44	91.06
0.006	108.09	166.14	0.006	110.74	96.79
0.008	108.06	134.24	0.008	109.60	100.15
0.010	107.74	117.53	0.010	108.63	102.11
ADI METHOD t= 10^{-5} Hr.					
Horizontal First			Vertical First		
TIME Hr.	T(1,1) °F	T(2,1) °F	TIME Hr.	T(1,1) °F	T(2,1) °F
0.002	105.47	84.84	0.002	105.49	84.85
0.004	105.96	92.27	0.004	105.98	92.28
0.006	106.31	96.73	0.006	106.33	96.74
0.008	106.58	99.54	0.008	106.59	99.54
0.010	106.78	101.40	0.010	106.79	101.41
MODIFIED ADI t= 0.001 Hr.					
Horizontal First			Vertical First		
TIME Hr.	T(1,1) °F	T(2,1) °F	TIME Hr.	T(1,1) °F	T(2,1) °F
0.002	104.11	82.57	0.002	104.12	81.01
0.004	105.35	89.82	0.004	105.25	88.92
0.006	105.83	94.53	0.006	105.73	94.00
0.008	106.16	97.68	0.008	106.08	97.34
0.010	106.42	99.84	0.010	106.34	99.62

As a matter of interest, a modified formulation was considered where all temperatures along a horizontal row or along a vertical row were considered as new values when the values along that respective row were being sought. The horizontally implicit, vertically explicit formulation is

$$GH(I,J)[T'(I-1,J)-T'(I,J)] + GH(I+1,J)[T'(I+1,J)-T'(I,J)]$$

$$+ GV(I,J)[T(I,J-1)-T'(I,J)] + GV(I,J+1)[T(I,J+1)-T'(I,J)]$$

$$= \frac{C(I,J)[T'(I,J)-T(I,J)]}{\Delta t} \qquad (7)$$

The horizontally explicit, vertically implicit formulation is

$$GH(I,J)[T(I-1,J)-T'(I,J)] + GH(I+1,J)[T(I+1,J)-T'(I,J)]$$

$$+GV(I,J)[T'(I,J-1)-T'(I,J)] + GV(I,J+1)[T'(I,J+1)-T'(I,J)]$$

$$= \frac{C(I,J)[T'(I,J)-T(I,J)]}{\Delta t} \qquad (8)$$

While there are no defendable physical arguments to support this modified approach as given by equations (7) and (8), their solution by the ADI technique yielded solutions which appeared relatively good. Some results for this case are also listed in Table 2 for comparison. The PCM temperatures appear to be slightly low as compared to the values given by the direct ADI method using the small time increment.

A technique for handling the phase change was included in the programs and a few runs were made. At the end of a time step, those nodes whose new temperatures exceeded the melt temperature were considered to be undergoing phase change. The fraction melted was computed by using the

nodal capacitance and the predicted rise above the melt temperature. Comparison of the results using the direct ADI method for a time step of 10^{-5} Hr. by solving equation (6) first and equation (5) second with those for the modified scheme using equation (8) first and equation (7) second for a time step of 10^{-3} Hr. showed reasonably good agreement although the latter predictions were still lower than the former.

There is a difficulty in using the implicit method for phase change. Since all nodal temperatures are calculated simultaneously, those nodes experiencing phase change are in reality overpredicted and these values accordingly affect the accuracy of the predicted temperatures of neighboring nodes. The explicit method does not exhibit this feature since a nodal temperature is predicted in a forward stepping process. When a node begins to melt, its temperature can be accordingly adjusted before proceeding to calculation of the temperatures for neighboring nodes. An iteration procedure to handle this problem with the implicit method may be time consuming. The accuracy of the modified ADI as well as the method for accounting for the phase change warrant further study.

CONCLUSIONS AND RECOMMENDATIONS

Experiments have been performed on an invertable thermal capacitor with nonadecane as the PCM. Part of the temperature data has been reduced and plotted to facilitate comparisons. Photographs of the melting and solidification processes were also obtained. These provided good definition of the solidification process but only qualitative information of the melting process because a film of solid which adhered to the viewing windows blocked direct view of the interfacial contour.

A numerical study of a finned thermal capacitor cell has been continued by considering application of the

alternating-direction-implicit (ADI) method described by Reference 13. Direct application of the method to a typical cell does not appear to offer any advantage over the explicit method because of the large difference in the thermal diffusivity of the fin and that of the PCM. A modified ADI method, which does yield solutions with considerably less computer time, was explored but its accuracy has not been ascertained.

There are some logical extensions of this work. Some specific recommendations are as follows:

- Reduce and Examine the Remaining Experimental Data
- Use the Experimental Data as Basis for Checking an Analytical Study of the Entire Unit
- Explore the Dependability and Accuracy of the Modified ADI Numerical Approach
- To Facilitate a Parametric Study, Pursue Possible Combination of the Explicit and Implicit Schemes using the Latter for the Fins

Some study also needs to be devoted to a method of analyzing a capacitor which incorporates honeycomb as the filler since this is presently planned for use in certain flight units. References 14 and 15 contain some information on effective conductivities of honeycomb. These, however, pertain to situations where radiation is important. Something along this line where the honeycomb is filled with PCM would be most helpful.

REFERENCES

1. Stafford, J.L., and M.G. Grote, "Thermal Capacitor, Liquid Coolant-To-Phase Change Material Heat Exchangers for the NASA Skylab I Airlock Module," AIAA Paper No. 71-429, Presented at the 6th AIAA Thermophysics Conference, Tullahoma, Tennessee, April 26-28, 1971.

2. Schelden, B.G., and J.O. Golden, "Development of a Phase Change Thermal Control Device," AIAA Paper No. 72-287, AIAA 7th Thermophysics Conference, San Antonio, Texas, April 10-12, 1972.

3. Hale, D.V., M.J. Hoover, and M.J. O'Neill, "Phase Change Materials Handbook," LMSC-HREC D225138, September 1971, Lockheed Missiles and Space Co., Huntsville, Ala.

4. Bentilla, E.W., K.F. Sterrett, and L.E. Karre, "Research and Development By Use of Fusible Material," Final Report, Northrop Corporation, NSL 65-16-1, MSFC Contract No. NAS 8-11163, April, 1966.

5. Hoover, M.J., P.G. Grodzka, and M.J. O'Neill, "Space Thermal Control Development," Final Report on MSFC Contract NAS 8-25183, HREC-5183-3, LMSC-HREC D 225500, Lockheed Missiles and Space Co., Huntsville, Ala., December, 1971.

6. Prensner, D.S., and G.R. Hsu, "Thermal Convection in Laterally Bounded Air Spaces," Proceedings of the Eight Southeastern Seminar on Thermal Sciences, Vanderbilt University, Nashville, Tennessee, March 23-24, 1972, pp. 339-368.

7. O'Toole, J.L., and P.L. Silveston," Correlations of Convective Heat Transfer in Confined Horizontal Layers," Chemical Engineering Progress Symposium Series, Vol. 57, No. 32, 1961, pp. 81-86.

8. Tien, C., and Y.C. Yen, "Approximate Solution of a Melting Problem With Natural Convection," "Chemical Engineering Progress Symposium Series," Vol. 62, No. 64, 1966, pp. 166-172.

9. Boger, D.V., and J.W. Westwater, "Effect of Buoyancy on the Melting and Freezing Process," J. Heat Transfer, February, 1967, pp. 81-89.

10. Catton, I., and D.K. Edwards, "Effect of Side Walls on Natural Convection Between Horizontal Plates Heated from Below," J. Heat Transfer, Trans. ASME, November, 1967, pp. 295-299.

11. Grodzka, P.G., C. Fan, and R.O. Hedden, "The Apollo 14 Heat Flow and Convection Demonstration Experiments-Final Results of Data Analyses." LMSC-HREC D 225333, MSFC Contract No. NAS8-25577, Lockheed Missiles and Space Company, Huntsville, Ala., September, 1971.

12. Chapman, Alan J. Heat Transfer , New York, The MacMillan Company, pp. 186-194, 1967.

13. von Rosenberg, Dale V., Methods for the Numerical Solution of Partial Differential Equations, American Elsevier Publishing Co., Inc. New York, 1969.

14. Swann, R.T., and C.M. Pittman, "Analysis of Effective Thermal Conductivities of Honeycomb-Core and Corrugated-Core Sandwich Panels," NASA TN D - 714, April, 1961.

15. Stroud, C.W., "Experimental Verification of an Analytical Determination of Overall Thermal Conductivity of Honeycomb-Core Panles". NASA TN D-2866, June, 1965.

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

THERMAL CONTROL OF THE SCIENTIFIC INSTRUMENT PACKAGE IN THE LARGE SPACE TELESCOPE

Prepared by:	Keith H. Hawks
Academic Rank:	Assistant Professor
University:	Purdue University
Laboratory: (Division) (Branch)	Program Development--Preliminary Design Propulsion and Mechanical Systems Division Thermal and Environmental Analysis Branch
Research Counterpart:	Carl G. Fritz
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author wishes to thank the ASEE/NASA program directors for a unique and stimulating summer experience; in addition, special thanks to Carl G. Fritz for his expert technical guidance and advice, and for the many rewarding and enlightening discussions. Thanks are also due to G. S. Sandhu, Brown Engineering, for his analog computer expertise; to Mr. Goerner and Mr. Laue for their support; to Harold Manning for his advice and explanations of the SIP cameras and the telescope optical system; to James Price for his assistance with the digital computer programs; and to Robert Nixon, George Kearns, William Mordan and Charles Haupt for their assistance.

THERMAL CONTROL OF THE SCIENTIFIC INSTRUMENT PACKAGE IN THE LARGE SPACE TELESCOPE

by

Keith H. Hawks

ABSTRACT

A study of the thermal control concepts suitable for the thermal control of the Scientific Instrument Package (SIP) in the Large Space Telescope (LST) was conducted. The data generated last summer by the author's parametric study of heat rejection concepts suitable for spacecrafts were used as a study guide.

The general thermal control system philosophy was to utilize passive control where feasible and to utilize active methods only where required for more accurate thermal control of the SIP components with narrow temperature tolerances.

A thermal model of the SIP and a concept for cooling the SIP cameras are presented. The model and cooling concept have established a rationale for determining a Phase A baseline for the thermal control of the SIP. A discussion of the involvement of the author's work in selecting the baseline design is given.

INTRODUCTION

One of the principal goals of NASA is the establishment of large astronomical observatories. Typical of such telescopes being considered for these observatories is a 3-meter (120-inch), diffraction-limited instrument that can observe over a wide portion of the electromagnetic spectrum, particularly those portions not visible to earthbound telescopes. A 3-meter telescope above the earth's obscuring atmosphere would have 10 times the resolving power of the 200-inch telescope on Mount Palomar operating under the best atmospheric conditions. In addition, the space telescope would be capable of detecting stars 100 times fainter than the faintest stars detectable from earth (1).

Marshall Space Flight Center (MSFC) has undertaken a Phase A study of a Large Space Telescope (LST). This was a logical "next step" for MSFC's contribution to optical stellar space astronomy after having gained engineering and management experience in the Skylab Apollo Telescope Mount (ATM) and High Energy Astronomy Observatory (HEAO) projects. The primary object of the LST project is to orbit a large, high resolution optical telescope system. A precursor LST is to be launched in 1979; a diffraction-limited LST is then to be launched in the early 1980's by the Space Shuttle.

The orbiting LST is composed of three functional elements entitled Optical Telescope Assembly (OTA), Scientific Instrument Package (SIP), and Support Systems Module (SSM) (Figure 1). The OTA contains the optics, the forward structure, the sunshade, baffles and systems for stabilization, telescope alignment and focus. The SIP consists of the scientific instruments, supporting optics and the structure located at the telescope focal region. The SSM includes the aft structure and systems required to support the OTA and SIP.

The OTA is a 3-meter Ritchey-Chretien Cassegrain two-mirror telescope. The primary mirror is a 120-inch mirror with a f/2.2 focal ratio. The overall telescope focal ratio is f/12, operates in the spectral range of 900 to 20000 Å with a spatial resolution of 0.04 arc-second at 5000 Å.

The primary function of the SIP is to convert the OTA focal plane energy into scientific information. Typical SIP instruments are field cameras, faint object spectrographs (FOS) and guidance instruments.

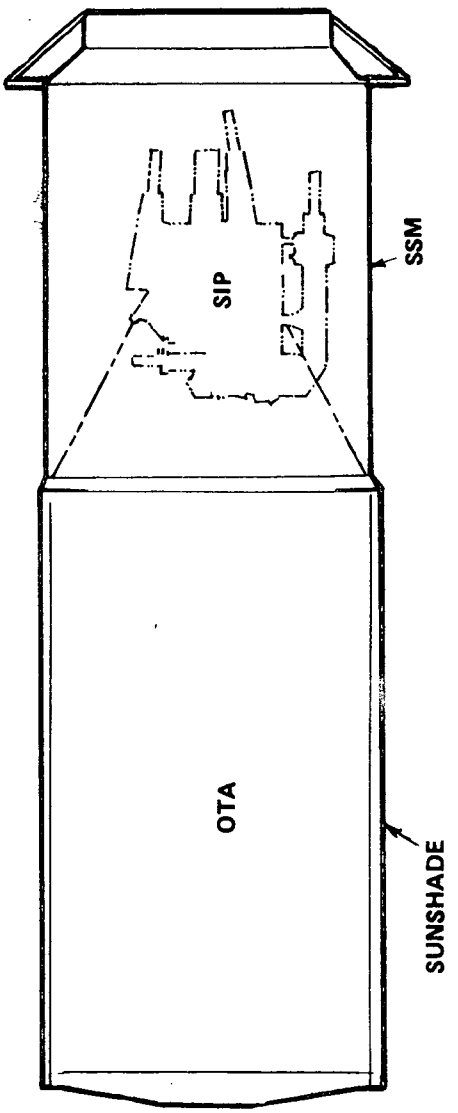


FIGURE 1. LST

The SSM is designed to structurally interface with the OTA. It provides the OTA and SIP with electrical power, communications, course attitude sensing and control, a launch vehicle interface structure and a docking structure for on-orbit servicing or retrieval by the Space Shuttle.

The conceptual design of the thermal control system for the LST was divided into the following three areas of responsibility: the OTA, the SIP and the SSM. The Optical Systems Division of the Itek Corporation was given the responsibility for the thermal control of the OTA and the SIP. MSFC was given the responsibility for the thermal control design of the SSM as well as interfacing and monitoring the thermal work of Itek. MSFC has been conducting a SSM thermal model design and also undertaking an in-house thermal control design study of the OTA and SIP.

For his summer project, the author was given the responsibility of conducting a thermal study of the heat rejection concepts suitable for the thermal control of the SIP. The results of the study would be used to establish a Phase A baseline for the thermal control of the SIP.

OBJECTIVES

The initial objective of this study was to establish a rationale for resolving the following questions:

1. Should the cameras be an integral part (thermally) of the structure or should they be separate from the structure?

2. What are the structural temperature gradients for different materials?

3. What is the effect on the structural temperature gradients of turning one camera off and another on?

4. Should the structure be an open truss design or a skin-stringer design?

5. Can the cameras be cooled to -20°C and what effect will this have on the SIP?

The final objective of the project was to establish the Phase A baseline for the thermal control of the SIP.

SIP THERMAL MODEL

Thermal design requirements have always had a considerable impact on the design of astronomy vehicles; however, with the advent of precise pointing requirements, the influence of the thermal design is even more pronounced. The LST must have a vehicle pointing accuracy of $\pm$ 1 arc-second, and the thermally induced structural dimensional changes cannot change more than $\pm$ 0.00008 inchs axially and $\pm$ 0.0004 inchs radially. Therefore, this study must also be capable of addressing itself to determining the thermal gradients induced in the SIP during nominal and abnormal operations.

To meet the objectives of this study, a thermal model of the SIP and a concept for cooling the cameras were developed. Figures 2 and 3 show a schematic of the SIP thermal model. The model is for both the SIP and SSM. The thermal control system philosophy was to assume a passive thermal control system until the study proved the assumption to be invalid. Thus, the SSM had to be included in the model of the SIP to account for radiation heat transfer blockage and reradiation. The only instruments depicted in Figures 2 and 3 are the F/96 cameras. The remaining SIP instruments will be discussed later in this report. Itek drawings No. 911333 and No. 911345 were used as the basis for forming the schematic. Figure 4 shows the dimensions that were used to size the SIP-SSM components.

The nodal model used to perform the transient orbital thermal analysis is presented in Figures 5, 6, 7 and 8. The model contains 238 nodes. Some of the node numbers are indicated on the figures. The two nodes on the primary support ring were taken to be constant temperature nodes (21.4°C:70°F). The thermal interface between the OTA and SSM has been a constant temperature primary ring. This was a requirement imposed by the primary mirror. The center node on the primary ring gives the model the capability of simulating the optic hole in the pressure bulkhead between the OTA and SSM. The SIP support struts were taken to be 2 inch diameter tubes. Each long and short strut was modeled by 3 and 2 nodes, respectively.

Table 1 gives a listing of the SIP instruments that have the largest heat generation rates. These are the instruments that have been included in the thermal model. The instruments listed in this report will not necessarily be the final choice of instrumentation for the LST. The LST instrumentation baseline will probably vary until the last possible moment before launch.

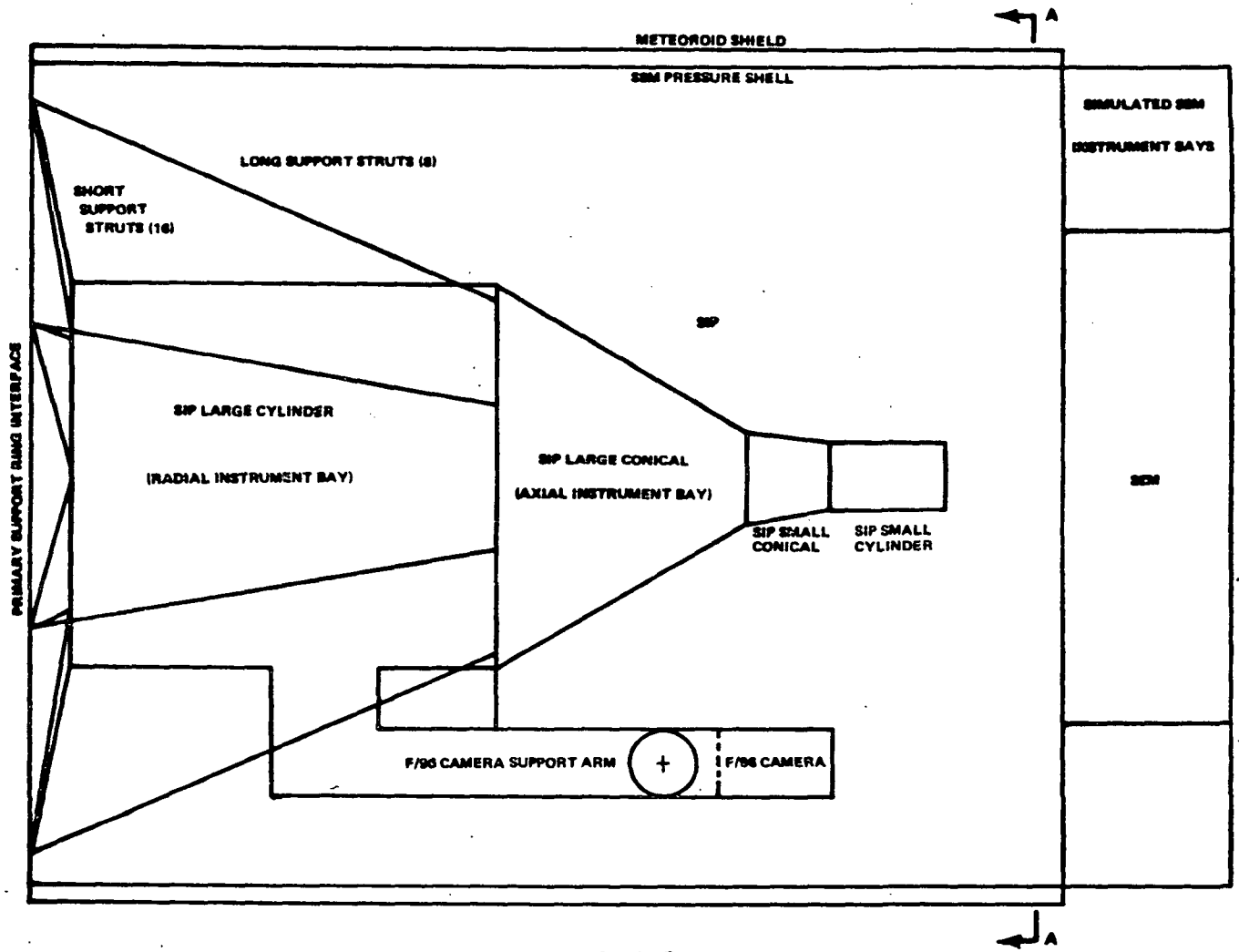


FIGURE 2. SCHEMATIC OF SIP THERMAL MODEL

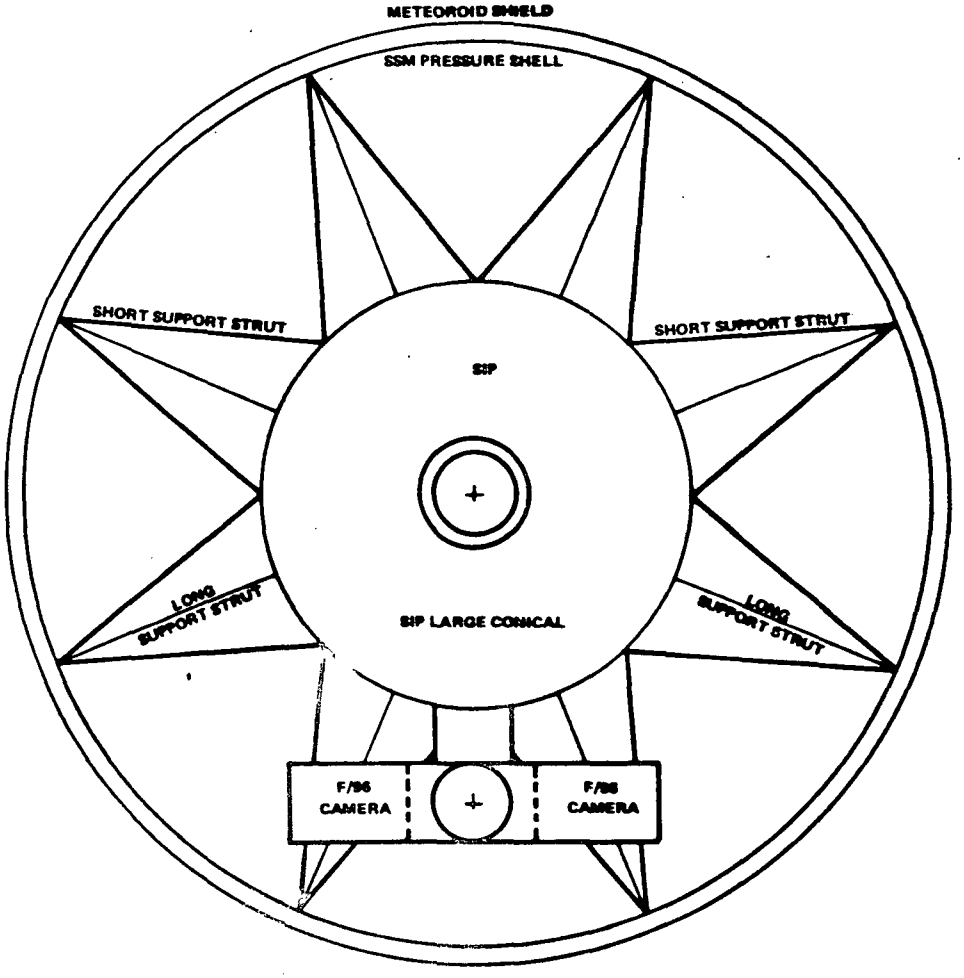


FIGURE 3. MATIC OF SIP THERMAL MODEL--SECTION AA

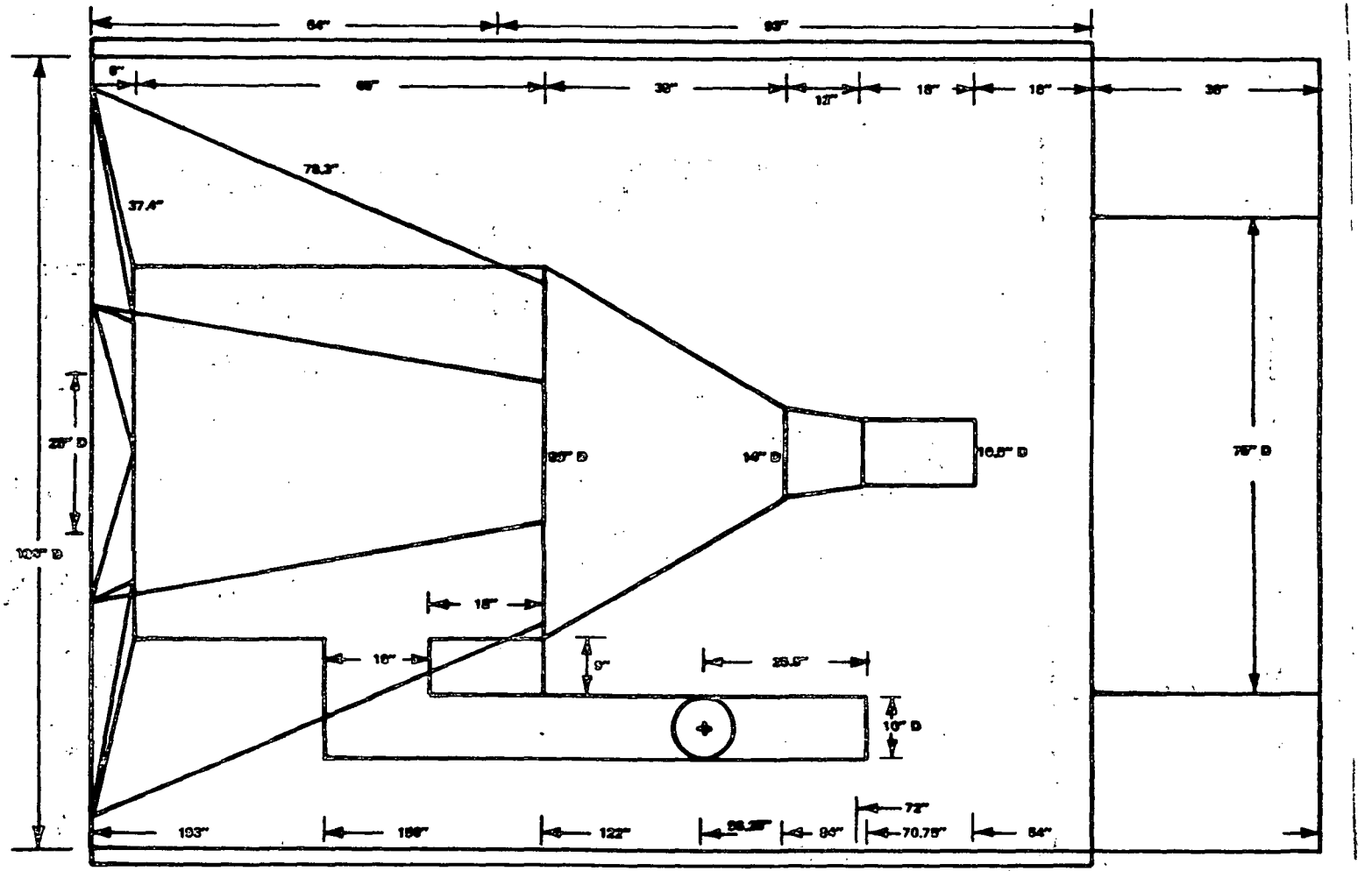
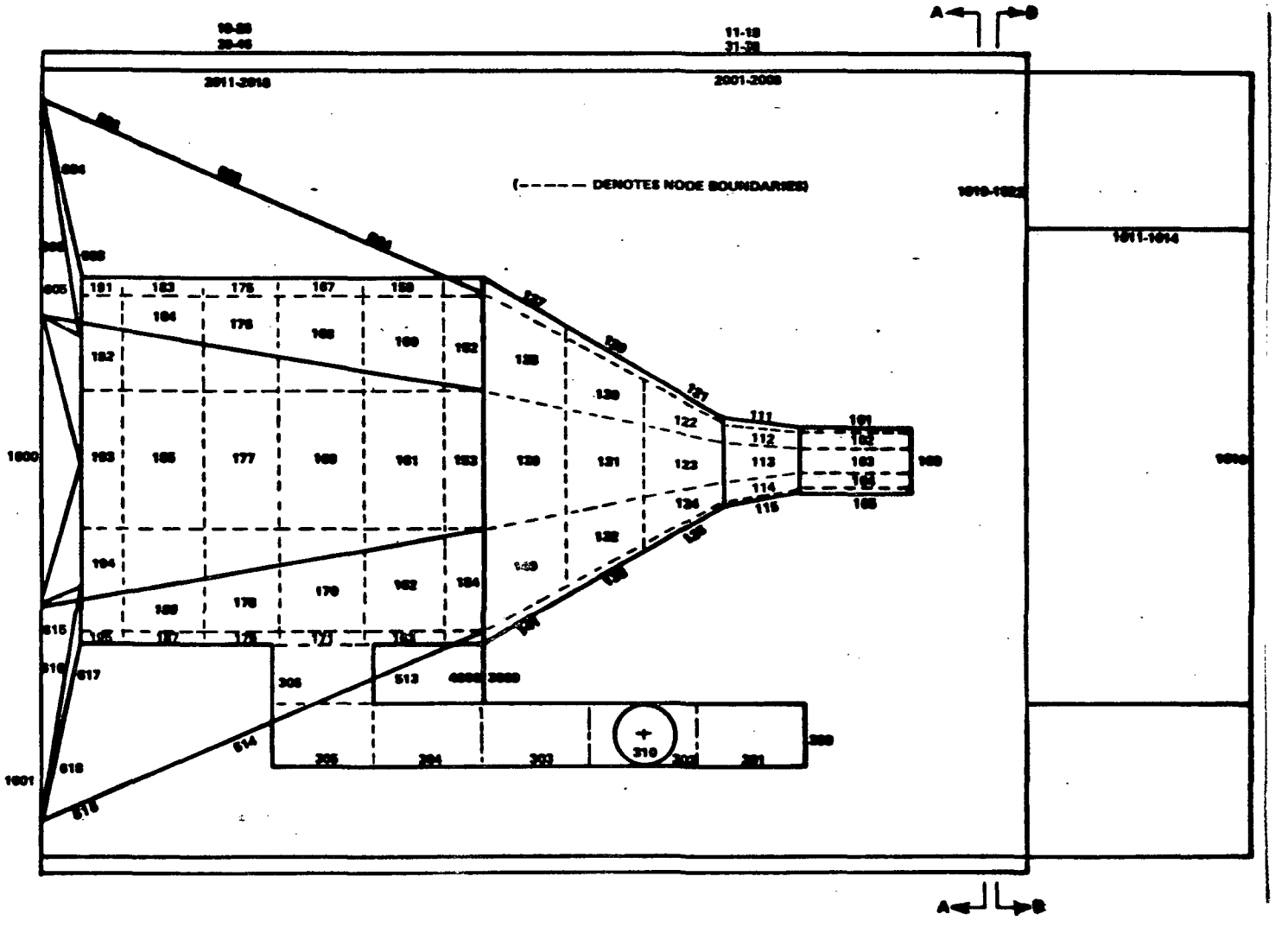


FIGURE 4. SIP THERMAL MODEL DIMENSIONS



GURE 5. SIP NODAL MODEL

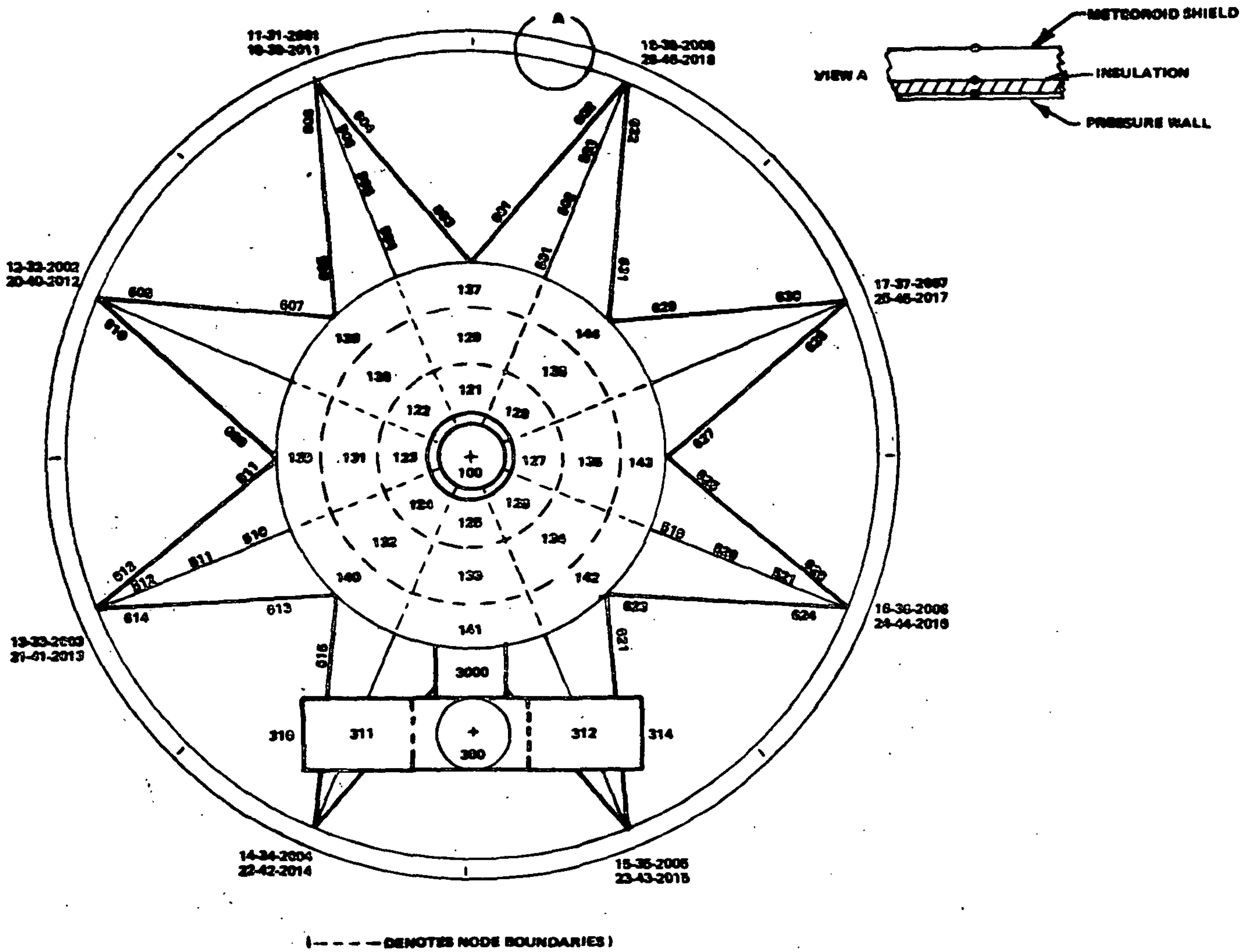


FIGURE 6. SIP NODAL MODEL--SECTION AA

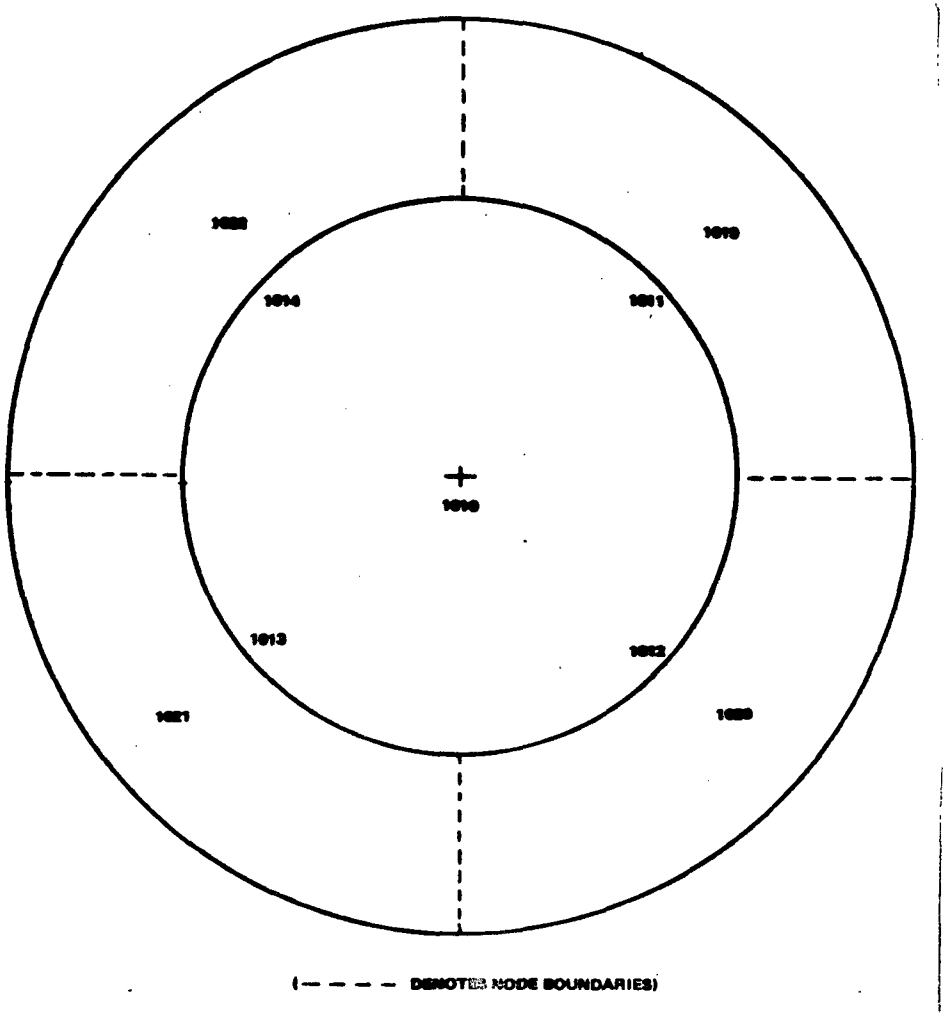


FIGURE SIP NODAL MODEL--SECTION BB

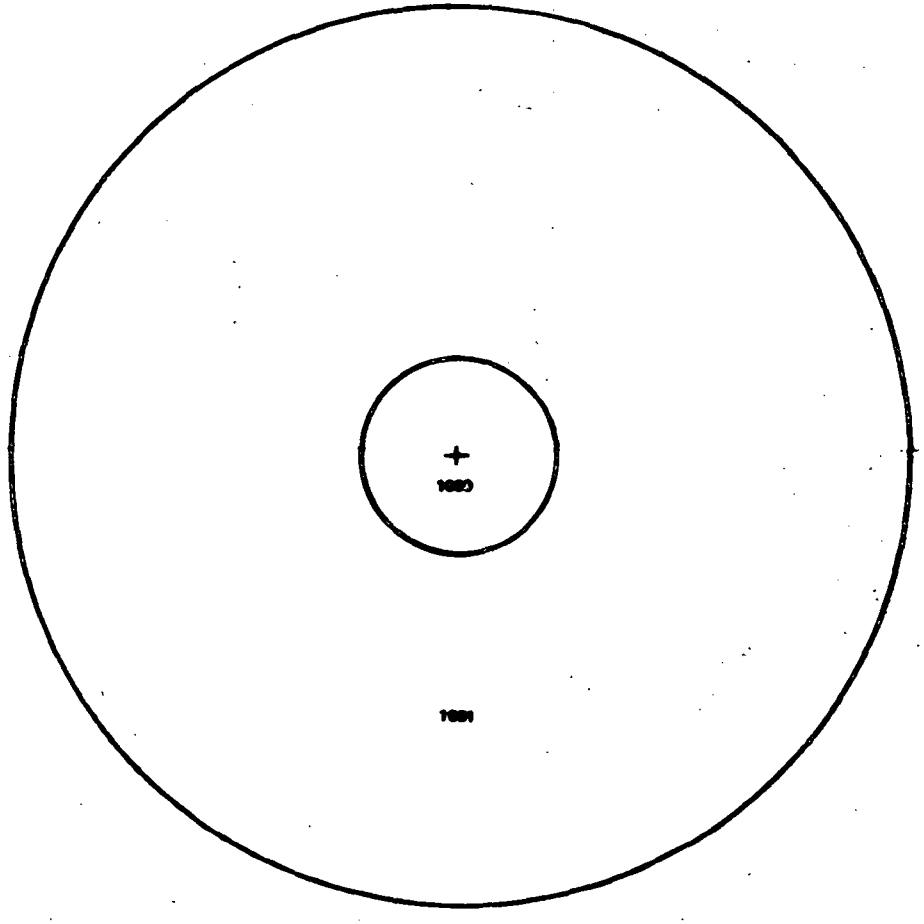


FIGURE 8. PRIMARY SUPPORT RING NODAL MODEL

TABLE 1. INSTRUMENT POWER REQUIREMENTS

Instrument	Power (Watts)
Field Cameras	
F/12	50
F/96 (3 Tubes), Ea.	50
High Spectral Resolution Spectrographs	
No. 1. Echelle (1100-1800 Å)	22
No. 2. Echelle (1800-3500 Å)	22
Faint Object Spectrographs	
No. 1. UV (1150-2200 Å) and Spare, (2 Tubes), Ea.	22
No. 2. Near UV (2200-3800 Å)	22
Visible (3800-6600 Å)	
No. 3. Near IR (6600 A-1.1μ)	22
No. 4 IR (1.1-4μ)	6
Fine Guidance	30
Diagnostic Instruments	
Figure Sensor	6
Focus Sensor	6

Figure 9 shows a plane view of the axial positions of the instruments. In this figure all of the instruments have been rotated into the plane of the paper for clarity. The hardest part of establishing the thermal model was in determining the position of the instruments relative to the SIP structure. The radial position of the instruments was determined by using Figures 10 and 11. Figure 10 was scaled from the Itek drawings. The "spare" instrument is a position for expansion; as yet no camera has been specified for this position. However, to simulate a heat input from a camera in that position, the spare was assumed to be identical to FOS No. 1. During the course of this study FOS No. 1 and FOS No. 3 were the only FOS cameras that were shown on the Itek drawings. Since FOS No. 2 and FOS No. 4 are similar in construction to FOS No. 3 and all three are instruments in the radial bay, they were given the same axial position relative to the SIP structure as FOS No. 3. Figure 11 is the telescope field of view (FOV) allocations for the instruments. The data for Figure 11 was obtained from Kollsman Instrument Corporation who was responsible for the instrumentation of the SIP. The radial positions of the FOV allocations for the radial instruments were assumed to be

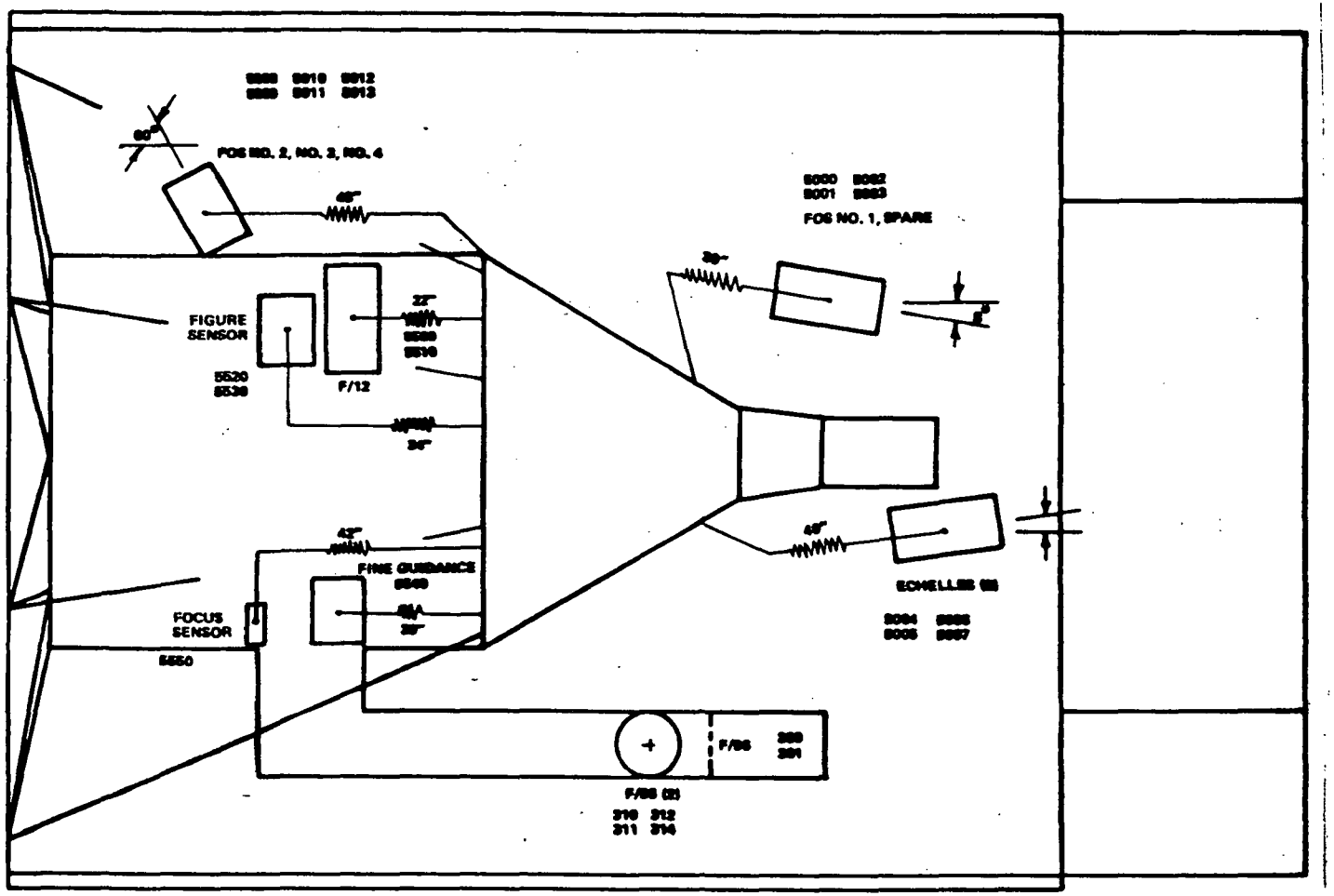


FIGURE 9. POSITION AND NODAL MODEL OF SIP INSTRUMENTS

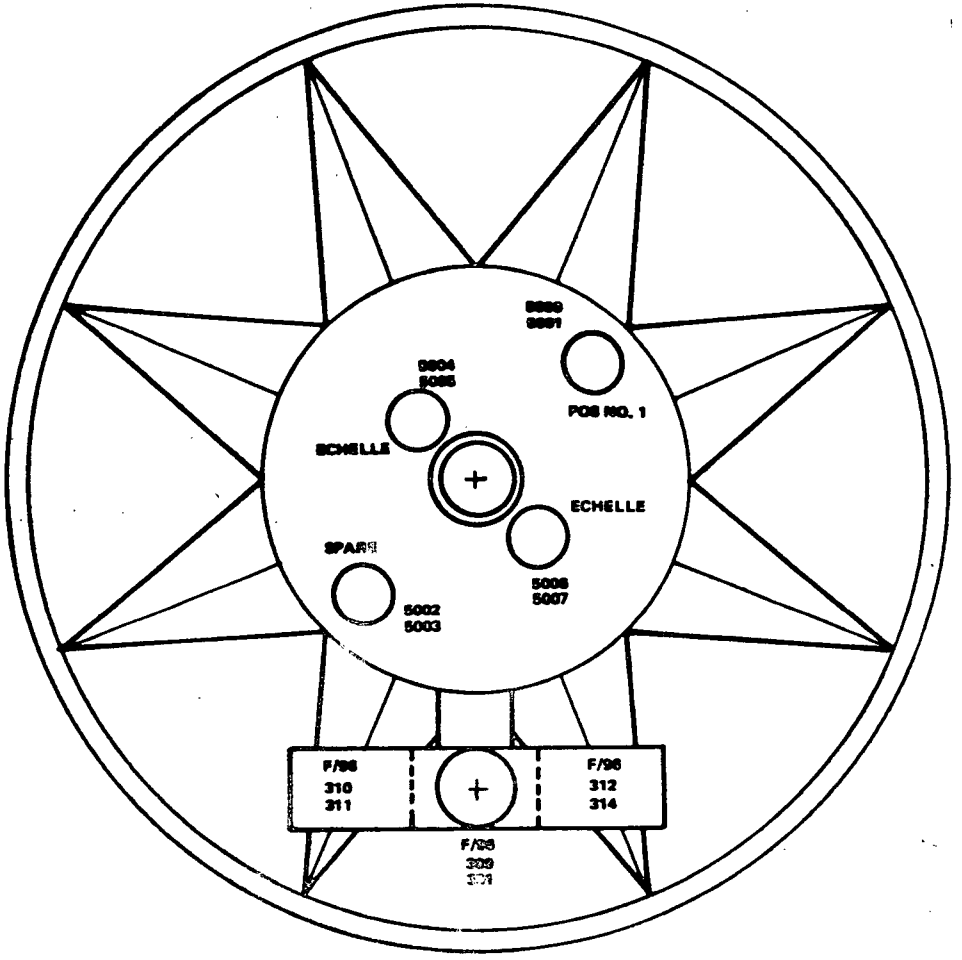


FIGURE POSITION OF AXIAL INSTRUMENTS

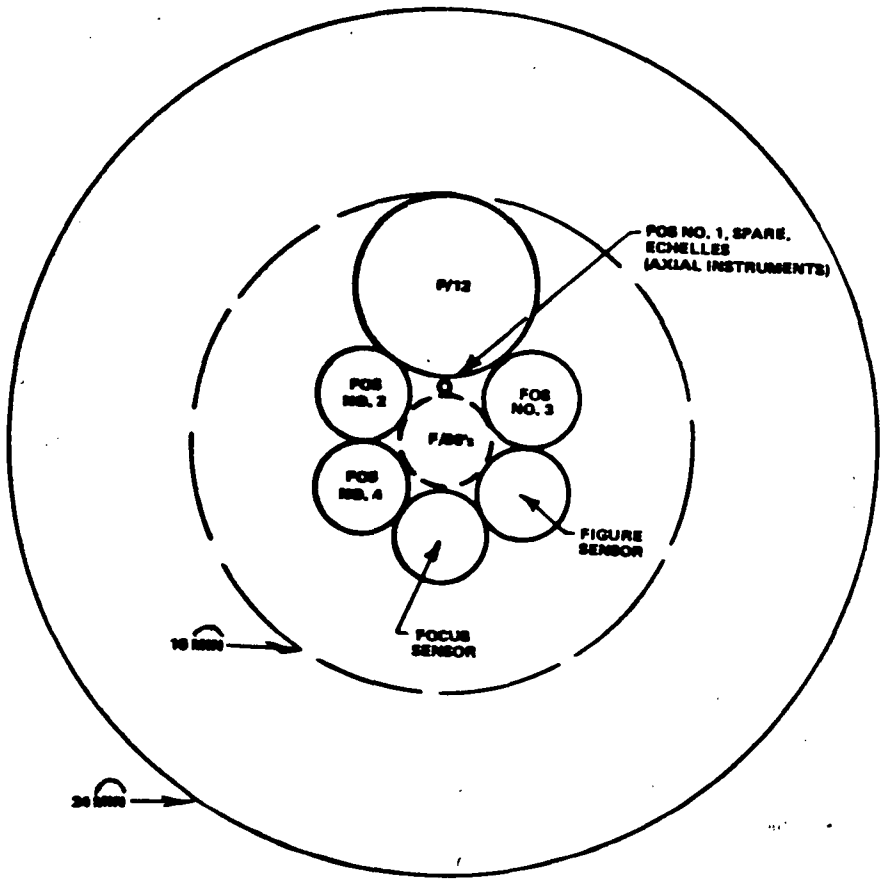


FIGURE 11. TELESCOPE FIELD OF VIEW ALLOCATIONS

the same positions for the instruments with respect to the SIP. The resistances shown in Figure 9 indicate the thermal conductance paths selected for the cameras. The thermal path lengths were chosen by selecting the most probable thermal conduction paths from the structural drawings. The F/96 cameras were positioned at the end of the F/96 camera support arm.

The geometrical shapes of the instruments were assumed to be cylinders since the camera tubes and coils are cylindrical by design. The length and diameter of the cylinders were chosen, after lengthy discussions with Kollsman as to the construction and operation of the cameras, to include that portion of the camera where the majority of the camera heat is generated. Two nodes were allotted to each instrument with the exception of the focus sensor and the fine guidance. This allowed radiation heat transfer from the sides and the exposed end of the instruments. The focus sensor and fine guidance were of a different design and were only given a cylindrical node.

The environmental heating rates of the external surface nodes were predicted using the Lockheed Orbital Heat Rate Program (2). The spacecraft was assumed to be launched into a 28.5 degree, 330 nautical mile circular orbit. One orbital orientation was considered ($\beta = 52^{\circ}$) with the LST longitudinal axis perpendicular to the solar vector.

The Chrysler Shape Factor Program (3) was utilized to calculate the geometric shape factors between all of the surfaces. The Chrysler program can only handle 200 nodes for a given problem. Thus, the model was divided into two parts while using the Chrysler program. Approximately 13 hours computer time was required to calculate all the view factors.

The thermal response of the SIP to the environmental heating rates and the internal heat sources was evaluated through utilization of the TRW "SINDA" digital computer program (4,5). In this solution, the physical system is first transformed into an equivalent thermal resistance-capacitance network which is input to the SINDA program. The program then employs finite difference techniques to solve for the transient temperature response of the nodal network. The SINDA program allows the programmer to easily remove and add conduction conductors and to vary the nodal heat loads during an orbit. These features were essential if the study objectives were to be met.

SIP THERMAL MODEL CAPABILITIES

The SIP nodal thermal model, as programmed for SINDA, has the following capabilities:

1. The SIP structure can be included as either a skin-stringer or an open truss body.

2. The model has been prepared for the following structure material models:
   A. Invar-titanium
   B. Aluminum
   C. Graphite/epoxy

3. The cameras can be conductively (thermal) isolated from the SIP structure.

4. The model can simulate the concept of switching cameras off and on.

5. The model can simulate the optical hole in the pressure bulkhead.

6. The SSM components are simulated in the model to include all radiation blockage and reradiation.

7. The model has the capability of simulating the concept of connecting the cameras to the pressure shell via heat pipes.

CAMERA TUBE COOLING CONCEPT

Kollsman Instrument Corporation had specified that the camera tubes would have to be cooled. Actually, the cooling of the photocathode and target of the tube is desired; but due to the difficulty in obtaining access to the interior of the tube, it is usually felt that cooling the tube would have to suffice. Since the final selection of cameras has not been made, the actual temperature to which the tubes must be cooled cannot be specified. But, for the candidate tubes being considered, the temperature usually being mentioned is -20°C.

Figure 12 shows the camera tube cooling concept developed by the author and his NASA counterpart. A BeO, or Al, shell is inserted over the tube. BeO, a ceramic, was selected because it has a thermal conductivity higher than Al at the same temperature and is a very poor electrical conductor. A layer of insulation is placed between the shell

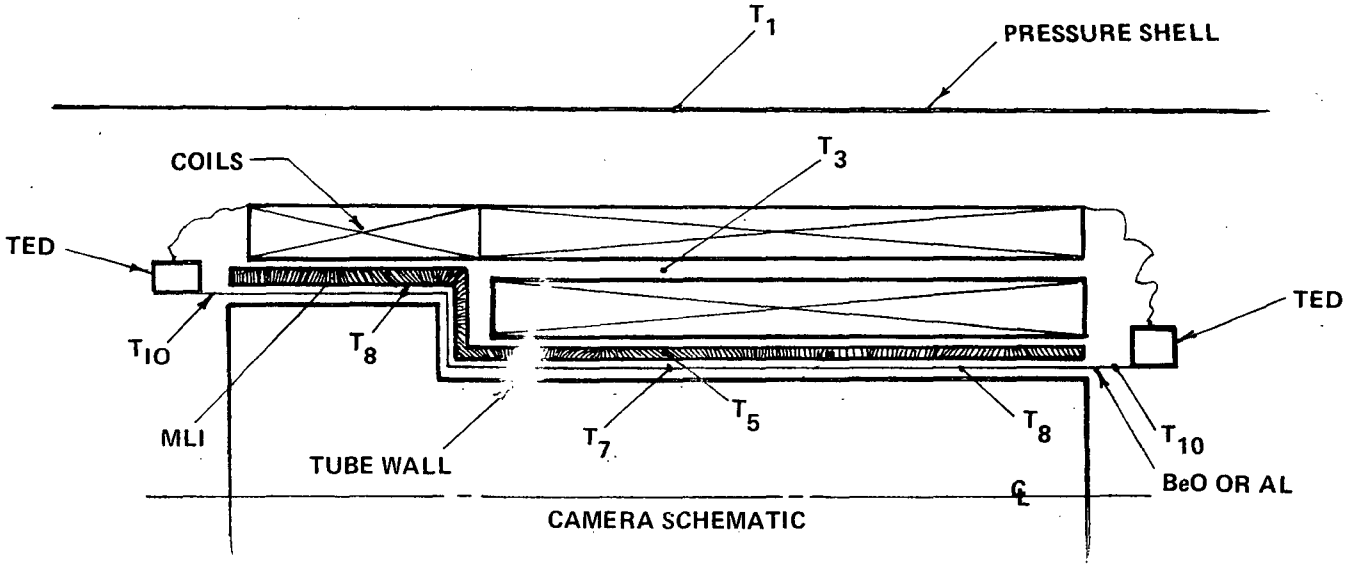


FIGURE 12. CAMERA COOLING CONCEPT

and the coils. The cold plate of a thermoelectric device is attached to the BeO shell at each end of the tube. The hot plate of the thermoelectric device is attached by way of a highly conductive path to the outside surface of the coils. Thus, the thermoelectric will reject heat to the surface of the coils.

Figure 13 shows some typical results of data obtained from the model developed for the tube concept. T_1 is the temperature inside the pressure shell of the LST wall. This temperature was obtained from an analog computer program for the case when the sun is incident normally to the meteoroid shield--outer skin of the LST. This would have the worst effect on the temperature of the tube. T_3, T_5 and T_7 are the temperatures of the coils, insulation and BeO, respectively. T_7 would be the hottest temperature imposed upon the tube wall by the shell. T_8 is the BeO temperature midway between the point where T_7 is determined and where the temperature of the thermoelectric coldplate is applied. It was assumed that the tube internal heat generation, consisting of radiant energy entering the ends of the tubes and the internal I^2R losses, would be 5.7 watts. The largest camera being considered generates 26 watts of heat in the coils. Data supplied by Ohio Semitronics Corporation showed that the thermoelectric devices would require 28 watts of power. Thus, for the case being considered in Figure 13, the coil surface will have to radiate about 60 watts of heat to the LST pressure wall.

The maximum temperature obtained for the coils was 29.6°C (84.7°F). This temperature is very moderate for the coils. The impressive result of the model is T_7. With the thermoelectric cold plate at -20°C, the hotest temperature imposed on the tube wall would be -19.3°C. Thus, these results indicate that the camera tubes can easily be cooled to -20°C. Another case was computed with the thermoelectric cold plate at -25°C; T_7 was -23°C, or 3° C below the desired temperature.

The communication and data management branch, PD-DO-EC, had expressed the desire to cool the photocathode to -20°C and to know the internal photocathode temperature distribution. Thus, a cylindrical model of the photocathode was constructed. In the model the edges of the photocathode were assumed to be cooled to -20°C using the BeO as a high conduction path to a thermoelectric device. The front surface was allowed to see, radiatively, a surface at 21.4°C (70°F). A partial differential equation was determined for the model, and a solution was obtained. The maximum temperature difference at the front surface of the photocathode was 10.3°C (18.5°F) and 4.7°C (8.4°F) at the rear surface which was assumed in the model to be an adiabatic surface.

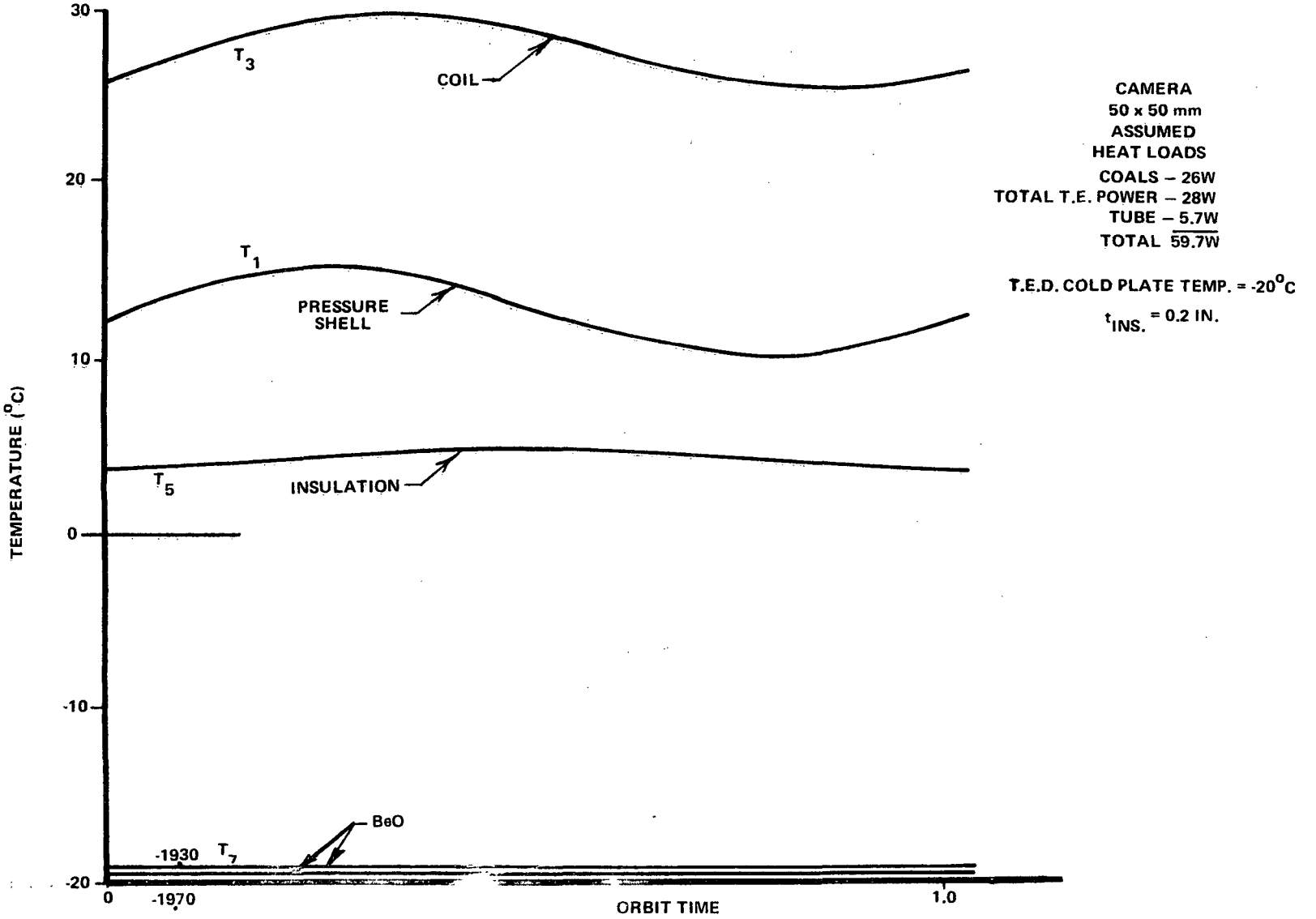


FIGURE 13. RESULTS OF CAMERA COOLING CONCEPT MODEL

CONCLUSION

The SINDA program with the SIP nodal model and the camera cooling concept have provided a means of meeting the objectives of this study. The results of the model of the camera cooling concept have completely revised the pre-Phase A SIP thermal control ideas. The camera heating loads can be separated from the SIP structure or, by utilizing the thermoelectric hot plate, the heat can be dumped into the structure. Additional SINDA runs are needed before a decision can be made. The basis for the final decision will be the thermal gradient changes within the structure during an orbit while switching cameras off and on.

The Phase A baseline for the LST is scheduled to be determined at the end of the month, August 1972. Even though the Summer Faculty Program will have ended before that time, the author will continue to work with his NASA counterpart until the thermal baseline for the SIP has been selected. At that time, all of the thermal results of this study will be published by the author and his NASA counterpart in a NASA publication. However, from the data on hand at the time of this writing, the Phase A baseline for the thermal control of the SIP in the LST will consist of the following passive and active concepts:

1. Passive concepts used will be insulation; control surfaces, paint; cold plates, camera surfaces, SIP surfaces.

2. Active concepts used will be heaters; thermoelectric cooling units.

REFERENCES

1. Optical Telescope Technology, NASA SP-233, A workshop held at MSFC, Huntsville, Alabama, April 29-May 1, 1969.

2. Lovin, J. K., L. W. Spradley, A. W. Lubkowitz, and J. E. Tyson, Lockheed Orbital Heat Rate Package, HREC-0514-1, LMSC/HREC D162217-1, Lockheed Missiles and Space Company, Huntsville, Alabama, April 1970.

3. Ponder, C. A., Computer Program for Calculating Geometric Shape Factors and Gray Body Interchange Factors Between Diffuse Gray Surfaces. Technical Report HSM-R588-70, Chrysler Space Division, Huntsville, Alabama, March 1970.

4. Ishimoto, T., and L. C. Fink, Systems Improved Numerical Differencing Analyzer--Program Manual, TRW, NASA 9-10435, June 1971.

5. Smith, J. P., Systems Improved Numerical Differencing Analyzer--User's Manual, TRW, NASA 9-10435, April 1971.

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

NONEQUILIBRIUM AND EQUILIBRIUM RADIATION TO THE SPACE TUG FROM THE SHOCK LAYER

Prepared by:	William L. Hendricks
Academic Rank:	Assistant Professor
University:	California Polytechnic State University
Laboratory: (Division) (Branch)	Aero-Astrodynamics Aerophysics Thermal Environment
Research Counterpart:	Bill R. Lake
Date:	August 18, 1972
Contract No.:	NGT-01-003-045

ACKNOWLEDGEMENTS

The author wishes to express his appreciation of the informative and stimulating technical discussions with the members of the Thermal Environment Branch of the Aerophysics Division and their contractors: Lockheed-Huntsville, Northrop-Huntsville and REMTECH-Huntsville. In particular, the conversations with Alan Forney, Carl Engel and Bill Lake were most rewarding.

LIST OF FIGURES

NOMENCLATURE

English:

A	=	Cross-sectional area
Bo	=	Boltzmann number [see Eq. (5)]
C_D	=	Drag coefficient
d_e	=	Effective relaxation distance
E_n	=	Integro-Exponential functions (n = 1, 2, 3)
h	=	Enthalpy
h_t	=	Total enthalpy $h + v_i v_i/2$
k	=	Radiation volumetric absorption coefficient
k_P	=	Planck mean absorption coefficient
k_R	=	Rosseland mean absorption coefficient
n	=	See Equation (8)
P	=	Pressure
q^R	=	Net radiation heat flux normal to the wall
q_i^R	=	Radiation flux vector
r	=	Distance of point M from axis of symmetry
R	=	Nose radius
T	=	Absolute temperature
u, v	=	Velocity vector component
v_i	=	Velocity vector
U, V	=	Velocity vector components for the Radiationless Case
$dV(P)$	=	Elementary volume around point P
W	=	Weight

x, z	=	Coordinates of point M (Fig. 1)
X, Z	=	Coordinates of point M (x, z) for the radiationless case

Greek:

β	=	See Equation (8)
γ	=	Specific heat ratio
Γ	=	Radiation-convection parameter [Eq. (3)]
Γ_β	=	Radiation-convection parameter for a thick gas Eq. (7)
Γ_n	=	Radiation-convection parameter for a thin gas Eq. (5)
δ	=	Shock layer thickness
Δ	=	Shock layer thickness for the radiationless case
η	=	Optical thickness between x and the wall ($\eta = k\,x$).
θ	=	See Figure 1
ρ	=	Density
σ	=	Stephan-Boltzmann constant
τ	=	Optical depth = $\int_0^x k(x)\,dx$
τ_{MP}^s	=	See Equation (1e)
τ_Δ	=	See Equation (23b)
$\tau_{d\varepsilon}$	=	See Equation (23c)

Subscripts:

c	=	Convective
Eq.	=	Equilibrium
i	=	Indicial notation
N. Eq.	=	Nonequilibrium
s	=	Immediately behind the shock

SL	=	At sea level
T	=	Total equilibrium and nonequilibrium
w	=	Conditions at the wall
∞	=	Upstream of shock
o	=	Radiationless (reference) case

NONEQUILIBRIUM AND EQUILIBRIUM RADIATION TO THE SPACE TUG FROM THE SHOCK LAYER

BY

William L. Hendricks

ABSTRACT

Preliminary estimates of the thermal radiation heat transfer to an aerobraking Space Tug are presented. The Tug is descending from a geosynchronous orbit to a low Earth orbit on one or more passes through the atmosphere. For the flight regime of the Tug, between a velocity of 22,000 ft./sec. and 34,000 ft./sec. with an altitude between 220,000 ft. and 340,000 ft., the nonequilibrium radiation can be more than 1,000 times larger than the value of the equilibrium radiation. Therefore, an analysis of the radiative heating for these altitudes must include the nonequilibrium radiation from the shock layer.

An uncoupled radiation-gas dynamics solution for an optically thin gas is obtained using a small perturbation analysis with the radiationless solution as a reference state. The nonequilibrium radiation model consists of two gaseous regions in the shock layer. The region immediately behind the shock is radiating at a temperature corresponding to translational and rotational equilibrium (γ = 1.4) while the gas behind this relaxation zone is radiating at the equilibrium temperature. This relaxation distance is assumed to be a constant fraction of the actual measured relaxation distance in order for the heat flux to closely approximate that from the actual exponentially decaying temperature profile. The analytical expression obtained from this model gives a quick estimate of the stagnation point radiation heat transfer including nonequilibrium radiation. An expression for the radiation heat flux distribution around a spherical body is also presented.

Preliminary estimates of the radiative and convective heating for a one pass aerobraking trajectory are given for several body radii. The radiative heating rates are essentially proportional to body radius while the convective heating is inversely proportional to the square root of the body radius. For a typical body radius of 7 ft. the maximum stagnation heating rate from radiation is nearly one half that from convection.

INTRODUCTION

The National Aeronautics and Space Administration is presently working on a preliminary design of a Space Tug capable of performing a round trip mission from low earth orbit to geosynchronous orbit. The Tug or AMOOS (Aeromaneuvering Orbit-to-Orbit Shuttle) is to be an unmanned craft which will be put into orbit via the Shuttle and after performing its mission returned to earth in the cargo bay of the Shuttle. Since the Tug is essentially another stage of the Shuttle, the need for a Tug is obvious. The baseline Tug is capable of separating from the Shuttle at a 100 n. mi., 28.5 degree inclined orbit with a 3000 lb., 8160 lb., or no payload; ascending to a geosynchronous orbit (19,300 n. mi. and zero degree inclination) to deploy the payload; retrieving a 3000 lb., 0 lb., or 4210 lb. payload, respectively, and returning it to the Shuttle (Ref. 1). It is feasible that the Tug could also be the main means of transporting equipment and supplies to build and furnish a moon colony.

During the return to the Shuttle, the apogee of the elliptical orbit is reduced to 270 n. mi. by drag dissipation of the orbital kinetic energy on one or more passes through the upper atmosphere (Ref. 2). The orbit is then circularized and a final Hohmann transfer is made for rendezvous with the Shuttle. A preliminary trajectory analysis (Ref. 2) indicates that for decay times between about 0.25 and 20 days, perigee altitudes range between 220 K ft. and 340 K ft. with a velocity between 22,000 ft./sec. and 34,000 ft./sec. for a W/C_DA range from 10 to 80 psf.

The Shuttle payload capability will determine the maximum weight and volume of the Space Tug. The Shuttle is capable of orbiting 65 K lbs. payload to a 100 n. mi. orbit. The clear cargo bay volume is able to accept a 15 by 60 ft. cylindrical module. Since the Tug is to have a 20 mission lifetime and less aerodynamic heating than the Apollo type vehicles, a nonablating heat shield is proposed. One of the proposed shapes is a 2:1 ellipsoid nose/cylindrical configuration with a nose radius of 14 ft. at a zero angle of attack. Another proposed shape is a spherical nose/cylindrical top/blunted bottom configuration with a blunted bottom radius greater than 7 ft. at a 40 - 60 degree angle of attack. The final configuration will depend on an accurate calculation of convective and radiative heat transfer which are dependent upon $1/\sqrt{R}$ and R respectively.

The heat transfer analysis required for the particular flight regime described previously is rather unique. The preliminary trajectory indicates that the flow field will vary from hypersonic, slip flow (slightly rarefied) at the lower altitudes to hypersonic, free-molecular flow (highly rarefied) at the higher altitudes. There has been a considerable amount of experimental and theoretical work done in the continuum flow regime. However, the meager amount of experimental radiation heat transfer data is large when compared to the theoretical radiation heat transfer analyses conducted for rarefied flow regimes.

OBJECTIVE

In this research a theoretical model which gives an accurate analytical expression for the equilibrium and nonequilibrium heat transfer is desired. This expression would be intended to give a quick, preliminary estimate of the radiation heat transfer to a vehicle such as the Space Tug. The expression should be in terms of the three following flight parameters: (1) velocity, (2) altitude and (3) nose radius.

I. Equilibrium Radiative Heating at the Stagnation Region of a Hypersonic Reentry Vehicle

A. Basic Equations of a Radiating Shock Layer

In order to simplify the basic equations the flow in the shock layer will be assumed to be inviscid and non-heat conducting. A perfect, gray gas will also be assumed. To isolate the influence of radiation and greatly simplify the analysis, nonequilibrium processes due to molecular transport, dissociation, ionization and vibration are neglected.

The conservation equations for a steady flow are:

Mass (neglecting the radiation of nuclear energy)

$$V_i \frac{d\rho}{dx_i} + \rho \frac{dV_i}{dx_i} = 0 \quad , \qquad (1a)$$

Momentum (neglecting body forces and the radiation stress tensor)

$$\rho V_i \frac{dV_i}{dx_i} + \frac{dP}{dx_i} = 0 \quad , \tag{1b}$$

Energy (neglecting chemical energy input and the radiation energy density)

$$\rho V_i \frac{dh_t}{dx_i} = - \frac{dq_i^R}{dx_i} \quad . \tag{1c}$$

If one assumes that the wall temperature is less than 0.4 of the temperature of the gas in the shock layer, the radiation from the wall can be neglected for the flight regime of the Tug (Ref. 3). This cold, absorbing wall assumption greatly simplifies the divergence of the heat flux at a point M (Fig. 1) in the shock layer to (Ref. 4)

$$\frac{dq_i^R(M)}{dx_i} = \sigma k(M) \left\{ 4\left[T(M)\right]^4 - \int_V \frac{k(P)\left[T(P)\right]^4 EXP\left(-\tau_{MP}^{S}\right)}{\pi \overline{MP}^2} dV(P) \right\} \quad . \tag{1d}$$

In Equation (1d) the first term on the right-hand side represents the energy emitted per unit time by the particles within a unit volume at M. The second term represents the energy absorbed per unit time from all other points P within the shock layer volume V. The energy absorbed from the gas outside the shock layer is neglected. The term τ_{MP}^{s} corresponds to the optical length between M and P,

$$\tau_{MP}^{s} = \int_{s=0}^{s=MP} k(s)\, ds \tag{1e}$$

The boundary conditions are that (1) at the wall, the velocity normal to the wall is zero and (2) at the shock the Rankine-Hugoniot relations apply. Thus the equations for the inviscid radiating shock layer are a complicated set of coupled, nonlinear integro-differential equations in terms of the independent variables x and z (Fig. 1).

B. Radiation Similarity Parameters

The important radiation parameters can be determined by nondimensionalizing the energy equation. Using the equilibrium state just behind the detached shock of the radiationless stagnation case as a reference state (subscript s) for the nondimensional variables:

$$\bar{\rho} \equiv \rho/\rho_s \quad , \quad \bar{v} \equiv v/v_s \quad , \quad \bar{h} \equiv h/h_s \quad ,$$

$$\bar{x}_i \equiv x_i/\Delta \quad \text{and} \quad \bar{q}_i^R \equiv q_i^R / q_0^R \quad ,$$

the energy equation becomes

$$\bar{\rho}\,\bar{V}_i\,\frac{\partial \bar{h}}{\partial \bar{x}_i} = -\Gamma\,\frac{\partial \bar{q}_i^R}{\partial \bar{x}_i} \quad . \tag{2}$$

Equation (2) contains the assumption that $h_t \cong h$ where Γ is defined as a ratio of the energy emitted by a gas at the reference state to the energy convected by the fluid in the shock layer, i.e.,

$$\Gamma = \frac{q_o^R}{\rho_\infty V_\infty h_s} = \frac{\sigma T_s^4}{\rho_s V_s h_s} \quad . \tag{3}$$

If Γ is small a small perturbation analysis will be sufficient to determine the radiating shock layer solution. The radiationless shock layer solution is the undisturbed state of the gas. Hence, a flight regime where $\Gamma << 1$ can be solved with an uncoupled radiation-gas dynamics solution, whereas a flight regime where $\Gamma \gtrsim 1$ requires a coupled radiation-gas dynamics solution.

It is also important to know the range of values of the optical depth ($\uparrow$) for a particular flight regime. If $\uparrow << 1$ or $\uparrow >> 1$ the expression for the divergence of the heat flux, Equation (1d), is further simplified.

It will now be advantageous to investigate the value of the right-hand side of Equation (2) for the flight regime of the Tug. For an optically thin gas Equation (2) can be written as (Ref. 5)

$$\bar{\rho}\,\bar{V}_i \frac{d\bar{h}}{d\bar{x}_i} = -\Gamma_n \bar{T}^4 \quad , \tag{4}$$

where

$$\bar{T} = T/T_s$$

and

$$\Gamma_n = \frac{4\tau_\Delta}{Bo} = \frac{4(k_P \Delta)\sigma T_s^4}{\rho_s V_s h_s} \tag{5}$$

is the radiation-convection parameter. Therefore, a perturbation solution for an optically thin gas requires that $\Gamma_n << 1$. For an optically thick gas Equation (2) can be written as (Ref. 5)

$$\bar{\rho}\,\bar{V}_i \frac{d\bar{h}}{d\bar{x}_i} = \Gamma_k \frac{d}{d\bar{x}_i}\left(\bar{T}^3 \frac{d\bar{T}}{d\bar{x}_i}\right) \quad , \tag{6}$$

where

$$\Gamma_k = \frac{16}{3\,Bo\,\tau_\Delta} = \frac{16}{3}\left(\frac{\sigma T_s^4}{\rho_s V_s h_s}\right)\left(\frac{1}{k_R \Delta}\right) \tag{7}$$

is the radiation-convection parameter. From Equation (6) it is apparent that a perturbation solution for an optically thick gas requires that $\Gamma_k << 1$.

C. Radiation Transfer Regimes of the Space Tug

To determine the values of Γ over the flight regime of the Tug, an ideal planetary entry atmosphere will be assumed. This atmosphere will be characterized by the following three criteria:

(1) T_∞ negligible compared to T_s

(2) Density ratio across radiationless, hypersonic shock is constant ($\simeq 18$). Therefore, the radiationless shock layer thickness-nose radius ratio Δ/R is also constant (Ref. 6).

(3) The temperature immediately behind the shock is proportional to the shock velocity, i.e., $T_s \propto V_\infty$.

At constant density we assume $C_p \propto T$, thus $h_s \propto T_s^2$ and, since $h_s \propto V_\infty^2$ one obtains $V_\infty \propto T_s$. This model is valid only for chemical equilibrium immediately behind the shock, but one should be aware that this model is used purely for the determination of Γ and τ in the flight regime of the Tug. Knowing values of Γ and τ one can then use various approximations for the radiation solution. The usual approximation for the absorption coefficient for air

$$k_s \propto \rho_s^{n+1} T_s^{\beta} \tag{8a}$$

becomes

$$k_s \propto \rho_\infty^{n+1} V_\infty^{\beta} \tag{8b}$$

for the model atmosphere. Substituting Equation (8b) and the criteria for the model atmosphere into the definitions of the various parameters Equations (5) and (7) one obtains:

$$\tau_\Delta \equiv k_s \Delta \propto \rho_\infty^{n+1} V_\infty^{\beta} R = b_1 \rho_\infty^{n+1} V_\infty^{\beta} R \quad , \tag{9a}$$

$$Bo \equiv \frac{\rho_\infty V_\infty h_s}{\sigma T_s^4} \propto \rho_\infty V_\infty^{-1} = b_2 \rho_\infty V_\infty^{-1} \quad , \tag{9b}$$

$$\Gamma_n \equiv \frac{4\sigma T_s^4 k_s \Delta}{\rho_\infty V_\infty h_s} = \frac{4\tau_\Delta}{Bo} \propto \rho_\infty^{n} V_\infty^{\beta+1} R = b_3 \rho_\infty^{n} V_\infty^{\beta+1} R \quad , \tag{9c}$$

and

$$\Gamma_k \equiv \frac{16\sigma T_s^4}{3\rho_\infty V_\infty h_s}\left(\frac{1}{k_s \Delta}\right) \propto \rho_\infty^{-n-2} V_\infty^{1-\beta} R^{-1} = b_4 \rho_\infty^{-n-2} V_\infty^{1-\beta} R^{-1} \quad . \tag{9d}$$

The values for n and β are chosen such that they will represent the parameters as closely as possible over a flight regime. The b's are determined, after n and β have been selected, for a good representation of the atmosphere at a flight condition of interest. After the constants have been calculated a plot of the flight regime with lines of constant τ and Γ will show the radiation transfer regime.

Two sets of values of n and β will be considered. Goulard (Ref. 4) suggests that one use $n = \frac{1}{2}$ and $\beta = 6$ for high temperature air behind a shock wave. Applying the flight condition (Ref. 7) of a V_∞ = 34,000 ft./sec., an altitude of 220,000 ft. and a body radius of 10 ft. to find the constants (b's), one obtains

$$\frac{1}{Bo} = 1.30 \times 10^{-6} \frac{(V_\infty / 34,000)}{(\rho_\infty / \rho_{SL})} , \qquad (10a)$$

$$\leftarrow_\Delta = 1.52 \times 10^{3} \left(\frac{\rho_\infty}{\rho_{SL}}\right)^{1.5} \left(\frac{V_\infty}{34,000}\right)^{6} \left(\frac{R}{10}\right) , \qquad (10b)$$

$$\Gamma_n = 1.30 \left(\frac{\rho_\infty}{\rho_{SL}}\right)^{0.5} \left(\frac{V_\infty}{34,000}\right)^{7} \left(\frac{R}{10}\right) , \qquad (10c)$$

and

$$\Gamma_k = \frac{1.30 \times 10^{-12}}{\left(\frac{\rho_\infty}{\rho_{SL}}\right)^{2.5} \left(\frac{V_\infty}{34,000}\right)^{5} \left(\frac{R}{10}\right)} \qquad (10d)$$

where R is in ft. and V_∞ in ft./sec. Pai (Ref. 8) suggests that in the temperature range from 7,000°K to 12,000°K, which is the range of interest for the present flight regime, one should use n = 0 and β = 4.4. This particular power representation gives

$$\frac{1}{Bo} = 1.30 \times 10^{-6} \frac{(V_\infty / 34,000)}{(\rho_\infty / \rho_{SL})} , \qquad (11a)$$

$$\tau_\Delta = 1.52 \times 10^{1} \left(\frac{\rho_\infty}{\rho_{SL}}\right) \left(\frac{V_\infty}{34,000}\right)^{4.4} \left(\frac{R}{10}\right) , \qquad (11b)$$

$$\Gamma_n = 1.30 \times 10^{-2} \left(\frac{V_\infty}{34,000}\right)^{5.4} \left(\frac{R}{10}\right) , \qquad (11c)$$

and

$$\Gamma_R = \frac{1.30 \times 10^{-10}}{\left(\frac{\rho_\infty}{\rho_{SL}}\right)^2 \left(\frac{V_\infty}{34,000}\right)^{3.4} \left(\frac{R}{10}\right)} . \qquad (11d)$$

Figure 2 shows the radiation transfer regime for the flight of the Space Tug using a body radius of 14 ft. Notice that the gas is optically thin throughout the flight range, since $\tau_\Delta \leqq 0.00229$. The radiation-convection parameter for an optically thin gas (Γ_n) is always less than 0.01865 (see Fig. 2). These results are valid for either selection of n and β . It is now apparent that a perturbation solution for $\Gamma_n << 1$ will allow a simplified procedure for finding the equilibrium radiation heat transfer.

D. One-Dimensional Model

1. Nonradiating Shock Layer

The numerical solutions of the two-dimensional nonradiating shock problem, Equations (1a) - (1e), show that near the stagnation region the flow properties are essentially independent of the distance r from the body axis (see Fig. 1). Thus, a one-

dimensional analysis will be used. The following assumptions are also utilized:

(1) The shock layer thickness Δ is much less than the body radius R.

(2) The density, temperature and enthalpy behind the shock are nearly constant in the stagnation area. The enthalpy is assumed much greater than the kinetic energy.

(3) The incompressible, potential flow values of velocity are assumed to be the actual velocity in the shock layer, that is

$$U = \frac{1}{2}\left(\frac{\rho_\infty V_\infty}{\rho_s \Delta}\right) r \tag{12a}$$

and

$$V = -\left(\frac{\rho_\infty V_\infty}{\rho_s \Delta}\right) Z \tag{12b}$$

(4) The pressure is assumed constant across the shock layer.

2. Radiating Shock Layer

Since $\Gamma_n < < 1$ the radiating shock layer problem is amenable to a one-dimensional perturbation solution. The solution is obtained by relating the radiating compressible shock layer to the nonradiating incompressible case. Characteristics of the radiating flow model are listed below:

(1) The radiation cooling increases the density of the shock layer causing a reduction of the shock layer thickness.

(2) The density and thermal properties are functions of z only and do not vary with respect to r.

(3) The velocity component v in the z direction is related to the incompressible velocity V by the expression

$$\rho v(z) = \rho_s V(Z) \tag{13a}$$

where the coordinate z is related to the coordinate Z for the corresponding radiationless case by the relation

$$\rho^{1/2}\, dz = \rho_s^{1/2}\, dZ \quad . \tag{13b}$$

(4) The density and temperature variations through the shock layer are related by the expression

$$\frac{\rho}{\rho_s} \simeq \frac{T_s}{T} \quad . \tag{14}$$

Goulard (Ref. 4) has solved the radiating shock layer for an optically thin gas ($\tau \ll 1$) and $\Gamma_n \ll 1$. His solution for the temperature variation is in terms of the nonradiating coordinate Z (measured from the body), the equilibrium temperature immediately behind the shock T_s and the radiation-convection parameter Γ_n,

$$\frac{T}{T_s} = 1 + \Gamma_n \ln\left(\frac{Z}{\Delta}\right) \quad . \tag{15}$$

Note from Equation (15) that the small disturbance hypothesis will tend to lose its validity when Z approaches zero. Hence, the values for the radiative heat flux

$$\frac{q_0^R}{2\sigma T_s^4 \tau_\Delta} = \int_0^1 \left(\frac{T}{T_s}\right)^{(\beta - n + 3.5)} dZ_1 \qquad (16a)$$

$$\cong 1 - (\beta - n + 3.5)\Gamma_n \qquad (16b)$$

and the ratio of the radiating shock layer thickness to the nonradiating shock layer thickness

$$\frac{\delta}{\Delta} = 1 - \frac{\Gamma_n}{2} \qquad (17)$$

are slightly overestimating the radiation cooling effect; they are a lower bound to the actual values of q^R and δ . Equations (16a) and (16b) are the expressions for the radiating shock layer heat flux from a small perturbation analysis ($\Gamma_n << 1$) of the nonradiating shock layer solution. Hence, Equation (16b) is within 10% of Equation (16a) for values of $\Gamma_n \leq .04$. However, Equation (16a) compares well with viscous and inviscid numerical solutions of the radiating shock layer up to values of $\Gamma_n \approx 10$ (Ref. 9). The above relations can also be extended up to $\tau_\Delta = 0.3$ for an error less than 5% if one replaces Γ_n with $\Gamma_n E_2(\tau_\Delta/2)$.

Using Equation (16b) an expression for the equilibrium radiation heat transfer can be written as

$$q^R = 2\sigma T_s^4 \tau_\Delta \left[1 - (\beta - n + 3.5)\Gamma_n\right] \quad . \qquad (18)$$

The term in the rectangular brackets adjusts the heat transfer from its isothermal, transparent value ($\Gamma_n = 0$) to a value ($\Gamma_n << 1$) which accounts for radiation cooling.

It will be more convenient to write Equation (18) in terms of altitude (density ratio), shock velocity and body radius. Applying the criteria described in Section I.C.:

$$T_s \propto V_\infty \quad , \qquad (19a)$$

$$\Delta \propto R \quad , \qquad (19b)$$

and

$$k_s \propto \rho_\infty^{n+1} V_\infty^{\beta} \qquad (19c)$$

where

$$\tau_\Delta \equiv k_s \Delta \quad , \qquad (19d)$$

Equation (18) becomes

$$q^R \propto R\left(\frac{\rho_\infty}{\rho_{SL}}\right)^{n+1}\left(\frac{V_\infty}{10^4}\right)^{4+\beta}\left[1 - (\beta - n + 3.5)\Gamma_n\right] \quad . \qquad (20)$$

In order to determine the constant of proportionality in Equation (20) and an appropriate value for n and β , a correlation with known equilibrium data was performed. The best correlation for a value of $\Gamma_n = 0$ is

$$q^R = 25\, R \left(\frac{\rho_\infty}{\rho_{SL}}\right)^{1.5} \left(\frac{V_\infty}{10^4}\right)^{10} \tag{21}$$

which corresponds to a value of ½ for n and 6 for β . Figure 3 - Figure 6 shows a comparison of the experimental data (Ref. 10, 11, 12, 13), Equation (21) and a correlation proposed by Nerem for a much larger range of velocities (Ref. 14). The deviation from the equilibrium value is due to the nonequilibrium radiation which is significant for high altitudes. These figures show that the nonequilibrium radiation is as much as 1000 to 10,000 times the equilibrium radiation. For the flight regime of the Tug the nonequilibrium radiation is much more significant than the equilibrium radiation. The stagnation point equilibrium radiation is obtained by inserting the constant of proportionality for n = ½ and β = 6 into Equation (20), and replacing Γ_n with its equivalent Equation (10c) to yield

$$q^R = 25\, R \left(\frac{\rho_\infty}{\rho_{SL}}\right)^{1.5} \left(\frac{V_\infty}{10^4}\right)^{10} \left[1 - 2.245 \times 10^{-3} \left(\frac{\rho_\infty}{\rho_{SL}}\right)^{0.5} \left(\frac{V_\infty}{10^4}\right)^{7} \left(\frac{R}{10}\right)\right]. \tag{22}$$

II. Nonequilibrium Radiative Heating at the Stagnation Region of a Hypersonic Reentry Vehicle

As mentioned earlier, the nonequilibrium radiation is the primary contribution to the radiation heat transfer for the Space Tug. This was determined from the experimental data shown in Figure 3 - Figure 6 . From the literature it is evident that a satisfactory theory capable of predicting the magnitude of the non-equilibrium radiation has not yet been developed. Present numerical methods require a great deal of computer time and the results, even after adjustments, are good only within an order of magnitude of the data. The problem is very complex because one must consider the transport of energy from the translation energy of the molecules, immediately behind the shock front, to the energies associated with vibration, electronic excitation, dissociation, ionization and the formation of new species over a relaxation distance.

The radiation model to be employed consists of two gaseous regions. The gas molecules immediately behind the shock are assumed to be in translational and rotational equilibrium. Thus, its temperature for this region ($T_{N.\ Eq.}$) corresponds to the value obtained from the normal shock relations for a gas of $\gamma = 1.4$. For the flight regime of the Space Tug, this temperature may be on the order of 50,000^{o}K. For simplicity, this gas is assumed to be at a constant temperature over an effective relaxation distance (d_e). This effective relaxation distance will be less than that measured in a laboratory in order for the heat flux of this model to simulate

that of the actual case where the temperature drops exponentially with distance behind the shock. The region of gas aft of this relaxation zone is assumed to be at the equilibrium temperature T_s. The optical depth $\tau_\Delta \equiv k_s \Delta$ is evaluated at the equilibrium temperature since the major portion of the shock layer will be in equilibrium. The radiation model is depicted in Figure 7. If the major portion of the shock layer is in equilibrium, the flow solution can be uncoupled from the radiation solution as in Section I. This seems to be a good approximation for velocities less than Earth parabolic velocity, approximately 36,700 ft./sec. (Ref. 15). Therefore, the problem reduces to adjusting the expression for the heat flux, Equation (22), to account for a portion of the gas being at a nonequilibrium temperature ($T_{N.\ Eq.}$).

The one-dimensional radiative heat flux to the cold, absorbing wall is given by (Ref. 5)

$$q^R_{N.Eq.} = -2\sigma \left(\int_0^{\tau_\Delta - \tau_{d_e}} T^4(\eta)\, E_2(\eta)\, d\eta + \int_{\tau_\Delta - \tau_{d_e}}^{\tau_\Delta} T^4(\eta)\, E_2(\eta)\, d\eta \right) \tag{23a}$$

where the optical depths are given by

$$\tau_\Delta \equiv k_s \Delta \tag{23b}$$

$$\tau_{d_e} \equiv k_s d_e \tag{23c}$$

and $E_2(\eta)$ is the exponential function. Equation (23a) the energy emitted by the gas in front of the shock layer is neglected. Inserting the temperature profile for the proposed nonequilibrium radiating model (see Fig. 7)

$$T(\eta) = T_s = T_{EQ.} \quad , \qquad 0 < \eta < (\tau_\Delta - \tau_{de}) \tag{24a}$$

$$T(\eta) = T_{N.EQ.} \quad , \qquad (\tau_\Delta - \tau_{de}) < \eta < \tau_\Delta \tag{24b}$$

into Equation (23a) one obtains

$$q^R_{N.EQ.} = \sigma T^4_{EQ.} \left\{ 1 - 2E_3(\tau_\Delta - \tau_{de}) \right.$$

$$\left. - 2\left(\frac{T_{N.EQ.}}{T_{EQ.}}\right)^4 \left[E_3(\tau_\Delta) - E_3(\tau_\Delta - \tau_{de}) \right] \right\} . \tag{25}$$

Since the gas is optically thin, the exponential function can be approximated by the relation

$$E_3(\tau) \approx \frac{1}{2} - \tau \quad , \qquad \tau << \tag{26}$$

Substituting Equation (26) into Equation (25) yields

$$q^R_{N.EQ.} = 2\sigma T^4_{EQ.} \tau_\Delta \left\{ 1 + \left[\left(\frac{T_{N.EQ.}}{T_{EQ.}} \right)^4 - 1 \right] \frac{\tau_{de}}{\tau_\Delta} \right\} . \tag{27}$$

The first term in Equation (27) is the equilibrium radiation heat transfer for a transparent gas ($\Gamma_n = 0$) while the second term accounts for the nonequilibrium radiation heat transfer. The total radiation heat transfer can now be written for small amounts of radiation cooling (see Equation (18) and note that $T_s = T_{Eq.}$) as

$$q_T^R = 2\sigma T_{Eq.}^4 \tau_\Delta \left[1 - (\beta - n + 3.5)\Gamma_n\right] \cdot$$

$$\cdot \left\{1 + \left[\left(\frac{T_{N.Eq.}}{T_{Eq.}}\right)^4 - 1\right] \frac{\tau_{de}}{\tau_\Delta}\right\} \quad . \quad (28)$$

The first portion of Equation (28) can be written in terms of the flight parameters (see Equation (22)) so that Equation (28) becomes

$$q_T^R = 25R\left(\frac{\rho_\infty}{\rho_{SL}}\right)^{1.5}\left(\frac{V_\infty}{10^4}\right)^{10}\left[1 - 2.245 \times 10^{-3}\left(\frac{\rho_\infty}{\rho_{SL}}\right)^{0.5}\left(\frac{V_\infty}{10^4}\right)^{7}\left(\frac{R}{10}\right)\right] \cdot$$

$$\cdot \left\{1 + \left[\left(\frac{T_{N.Eq.}}{T_{Eq.}}\right)^4 - 1\right] \frac{d_e}{\Delta}\right\} \quad . \quad (29)$$

For a particular body radius the relaxation distance (Ref. 11) is seen to be inversely proportional to the density ratio times the velocity ratio to the 3.8 power. Since the nonequilibrium effects are obviously independent of body radius, the ratio of the effective relaxation distance to the standoff distance is

$$\frac{d_e}{\Delta} = \frac{K}{\left(\frac{\rho_\infty}{\rho_{SL}}\right)\left(\frac{V_\infty}{10^4}\right)^{3.8}} , \qquad (30)$$

where the constant K is picked so that the nonequilibrium expression for the heat flux fits the data as closely as possible. Inserting Equation (3) into Equation (29) one finally obtains an analytical expression for the total radiative heat flux to the stagnation region of the Tug, i.e.

$$q_T^R = 25\, R\left(\frac{\rho_\infty}{\rho_{SL}}\right)^{1.5}\left(\frac{V_\infty}{10^4}\right)^{10}\left[1-2.245\times10^{-3}\left(\frac{\rho_\infty}{\rho_{SL}}\right)^{0.5}\left(\frac{V_\infty}{10^4}\right)\left(\frac{R}{10}\right)\right]\cdot$$

$$\cdot\left\{1+\left[\left[\frac{4.31\times10^{-5}}{\left(\frac{\rho_\infty}{\rho_{SL}}\right)\left(\frac{V_\infty}{10^4}\right)^{3.8}}\right]\right]\cdot\left[\left(\frac{T_{N,Eq.}}{T_{Eq.}}\right)^4-1\right]\right\} \qquad (31)$$

where

$$\left[\left[X \right]\right] = \begin{cases} X & X \leqq 1 \\ 1 & X \geqq 1 \end{cases} , \qquad (32)$$

R is the body radius in ft., V_∞ is the shock velocity in ft./sec., $T_{N.\ Eq.}$ is the temperature behind the shock corresponding to a $\gamma = 1.4$ and $T_{Eq.}$ is the equilibrium temperature immediately behind a shock. The temperatures, just described, can be found from Reference 16 and Reference 17 respectively for a particular altitude and velocity.

A comparison of Equation (31) with calculations from experimental data is shown in Figure 8. For the case of V_∞ = 20,000 ft./sec., which is smaller than that expected for the Tug, Equation (31) underestimates the total radiative heat flux. For the case of V_∞ = 36,000 ft./sec., which is larger than that expected for the Tug, Equation (31) overestimates the total radiative heat flux. For a velocity (25,000 ft./sec.) and radii (3.28 and 10 ft.) in the range of the Tug, Equation (31) is seen to be a good estimate of the total heat flux until the effective relaxation distance approaches the shock layer thickness. When $d_e \cong \Delta$ the radiative heat flux is overestimated but the radiation becomes negligible at this high altitude due to the dissipation of the shock wave.

III. Distribution of Radiative Heat Flux Around a Sphere

The total radiative heat flux at a stagnation region can be determined with the use of Equation (31). In this section an expression for the radiative heat flux distribution around a spherical body is given.

The ratio of the radiative flux at some angle (Θ), measured from the stagnation point, to the flux at the stagnation point $q^R(0)$ is approximately

$$\frac{q^R(\Theta)}{q^R(0)} = (\cos \Theta)^P \qquad (33)$$

where P is independent of body radius and altitude (Ref. 3). From Reference 18 the viscous and inviscid distributions are essentially

the same. Therefore, using the calculated distribution for the inviscid solution from Reference 3 one can approximate P to give

$$q^R(\theta) = q^R(0)\,(\cos\theta)^{\left[36.9 - 22.9\left(\frac{V_\infty}{30{,}000}\right)\right]} \tag{34}$$

where V_∞ is in ft./sec. The value for $q^R(0)$ comes from Equation (31).

IV. Radiative Heating of the Space Tug

In this section Equations (31) and (34) are used to give preliminary estimates of the total radiative heat flux to the Space Tug.

The radiative heat flux, Equation (31), is shown for the flight regime of the Space Tug in Figures 9, 10 and 11 for body radii of 14 ft., 7 ft., and 1.84 ft., respectively. These body radii are from preliminary design concepts. The radiative heat flux shows an increase as the body radius increases, velocity increases and as the altitude decreases. By comparing the curves for a Γ_h of zero (transparent shock layer) with those including a small Γ_h, one notices that there is an overestimation of the heat flux by as much as 16%, 6% and 1% for body radii of 14 ft., 7 ft., and 1.84 ft., respectively. The larger estimates are only for the larger velocities and the lower altitudes. This implies that an uncoupled solution overestimates the radiative heat flux. In Figure 12 a Γ_h of zero is assumed in order to plot q^R/R independent of the body radius. Therefore a quick estimate of q^R/R can be found from Figure 12 for a particular altitude and velocity; but, it should be remembered that these values are in error for a combination of large body radii, high velocities and low altitudes.

Values of the convective heat flux times the square root of the body radius are also shown in Figure 12. These values are from the expression of Kemp & Riddell for a sphere in a continuum gas (Ref. 19),

$$q_c \sqrt{R} = 7.80 \times 10^2 \left(\frac{\rho_\infty}{\rho_{SL}} \right)^{0.5} \left(\frac{V_\infty}{10^4} \right)^{3.25} . \qquad (35)$$

Knowing a body radius and a trajectory one can use Figure 12 to give a quick estimate of the stagnation point convective and radiative heat flux for a blunt body including the effects of radiation non-equilibrium.

By using the preliminary trajectory for the Tug, which is shown in Figures 9 - 11, along with Equations (31) and (35) one can determine the radiative and convective heat flux versus time for a particular body radius. Figures 13, 14 and 15 show the effects of non-equilibrium radiation, continuum convection from Kemp & Riddell and the slip flow convection (Ref. 20). For a body radius of 14 ft. (Fig. 13) the radiative heat flux peak is about the same as the convective slip flow peak. For a body radius of 7 ft. (Fig. 14) the radiative heat flux peak is nearly half the convective heat flux peak. For a body radius of 1.84 ft. (Fig. 15) the radiation is negligible when compared to the convection.

Although the stagnation point radiative heat flux is significant, an examination of Figure 16 shows that the radiative heat flux decreases rapidly as one examines points around the body. This radiation distribution, Equation (34), is proportional to the cosine of Θ to a

power between 10 and 20 while the convection distribution is proportional to the cosine of Θ to a power of about 1.5 (Ref. 21). Hence the radiative heat flux is much more concentrated at the stagnation point than the convective heat flux.

CONCLUSIONS AND RECOMMENDATIONS

From a preliminary analysis of the nonequilibrium radiative heat flux for the flight regime of the Space Tug, it is apparent that the radiative heat flux is a significant portion of the total stagnation point heat flux for a blunt body.

To further enhance the accuracy of predicting the nonequilibrium radiation heat transfer, I suggest that one:

(1) Perform a numerical analysis of the coupled radiation-gas dynamic problem to determine the amount of overestimation assuming an uncoupled solution.

(2) Use a theoretical model that includes the concentrations of the various components such as atoms, molecules, ions and electrons during a relaxation of the translational degrees of freedom, molecular rotations, molecular vibrations, dissociation and chemical reactions, ionization and electronic excitation.

(3) Include other geometric shapes, such as cones, wedges and cylinders.

(4) Include the effects of precursor radiation since the nonequilibrium temperatures are large. By including precursor radiation one would expect a higher stagnation point heating (Ref. 22).

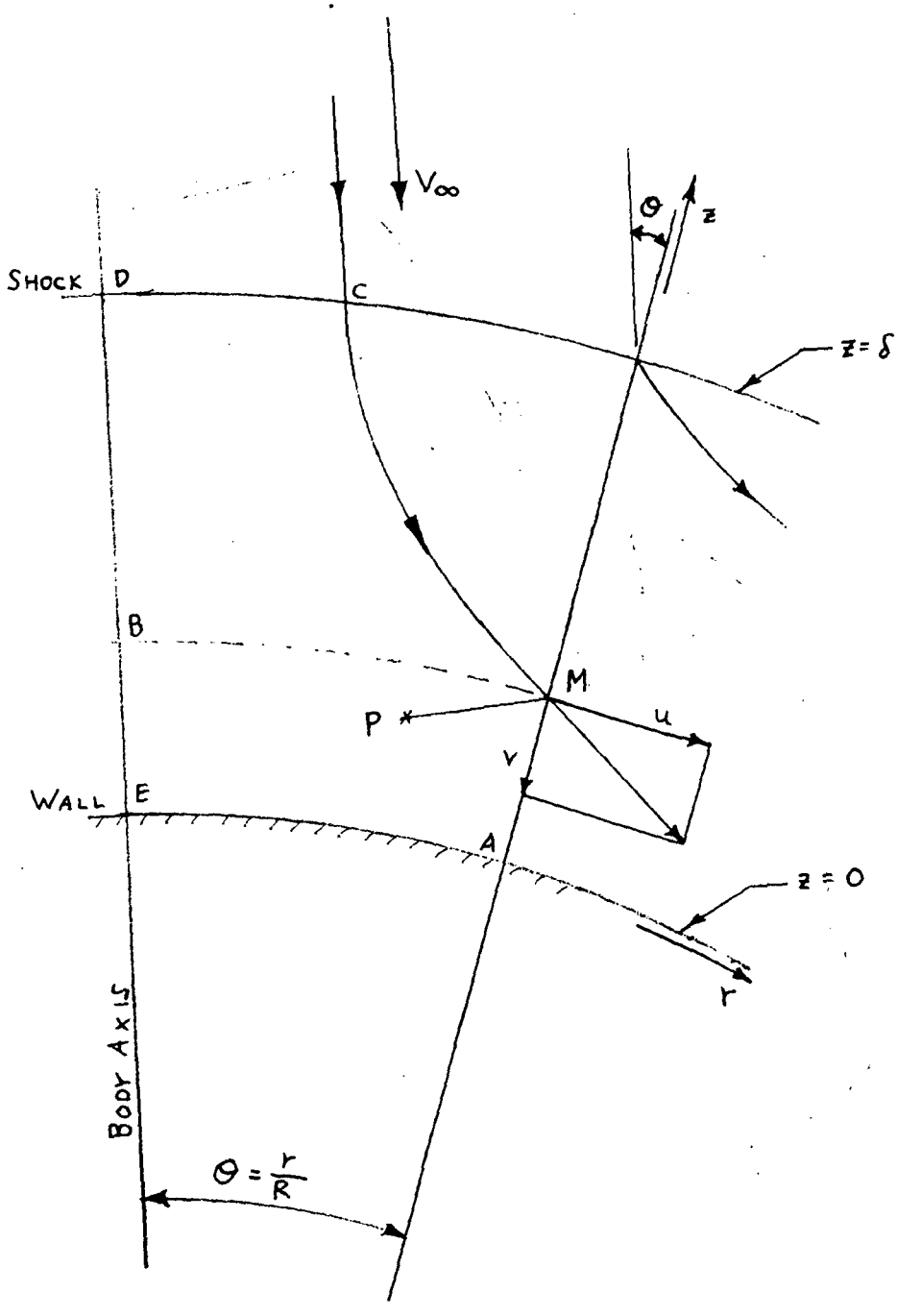


Fig. 1 Detached Shock Layer Nomenclature

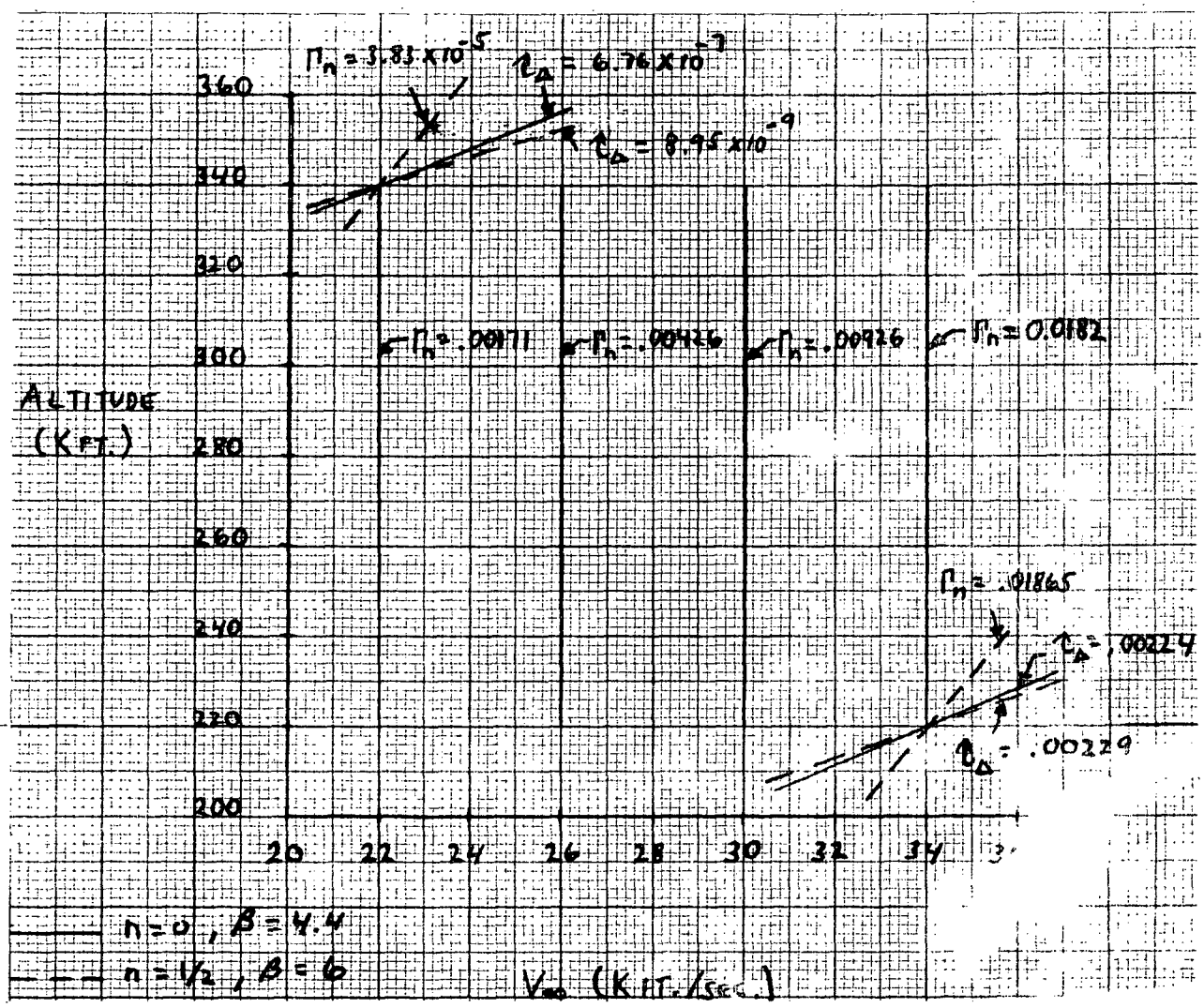


Fig. 2 Radiation Similarity Parameters for the Space Tug Flight Regime

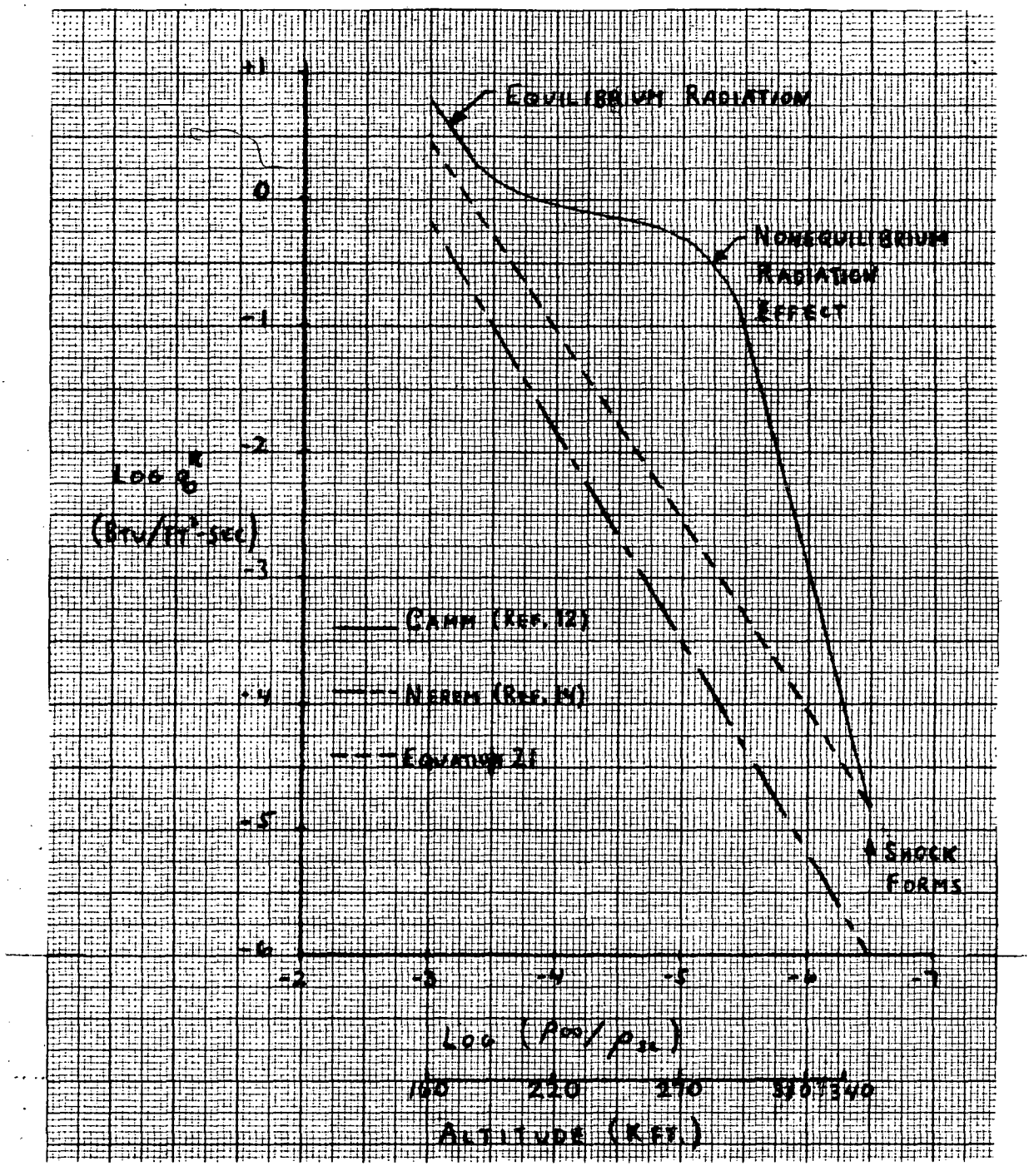


Fig. 3 Comparison of the Total Radiative Heat Flux with Two Equations for the Equilibrium Radiative Heat Flux for V_∞ = 20,000 ft./sec. and R = 3.28 ft.

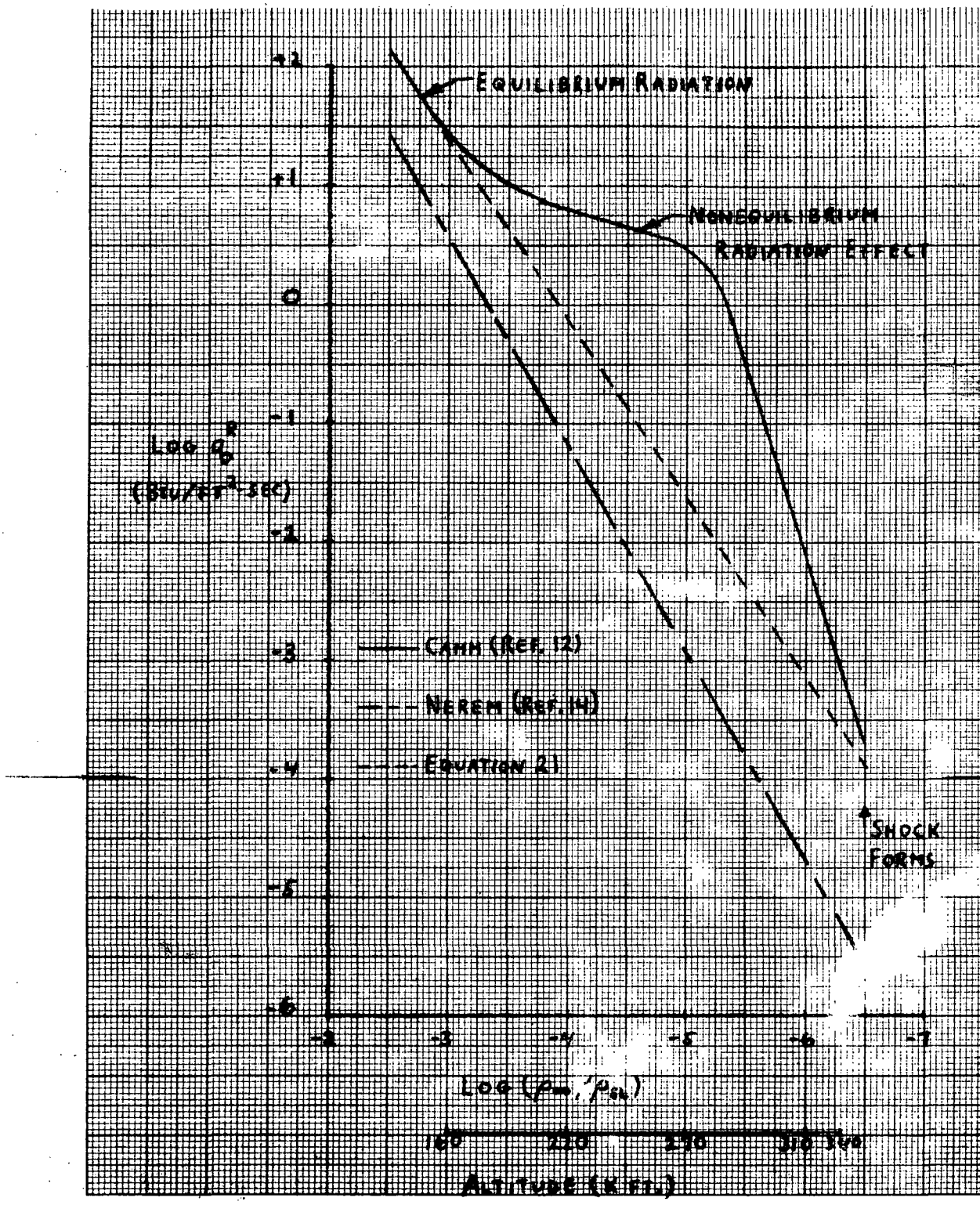


Fig. 4 Comparison of the Total Radiative Heat Flux with Two Equations for the Equilibrium Radiative Heat Flux for V_∞ = 25,000 ft./sec. and R = 3.28 ft.

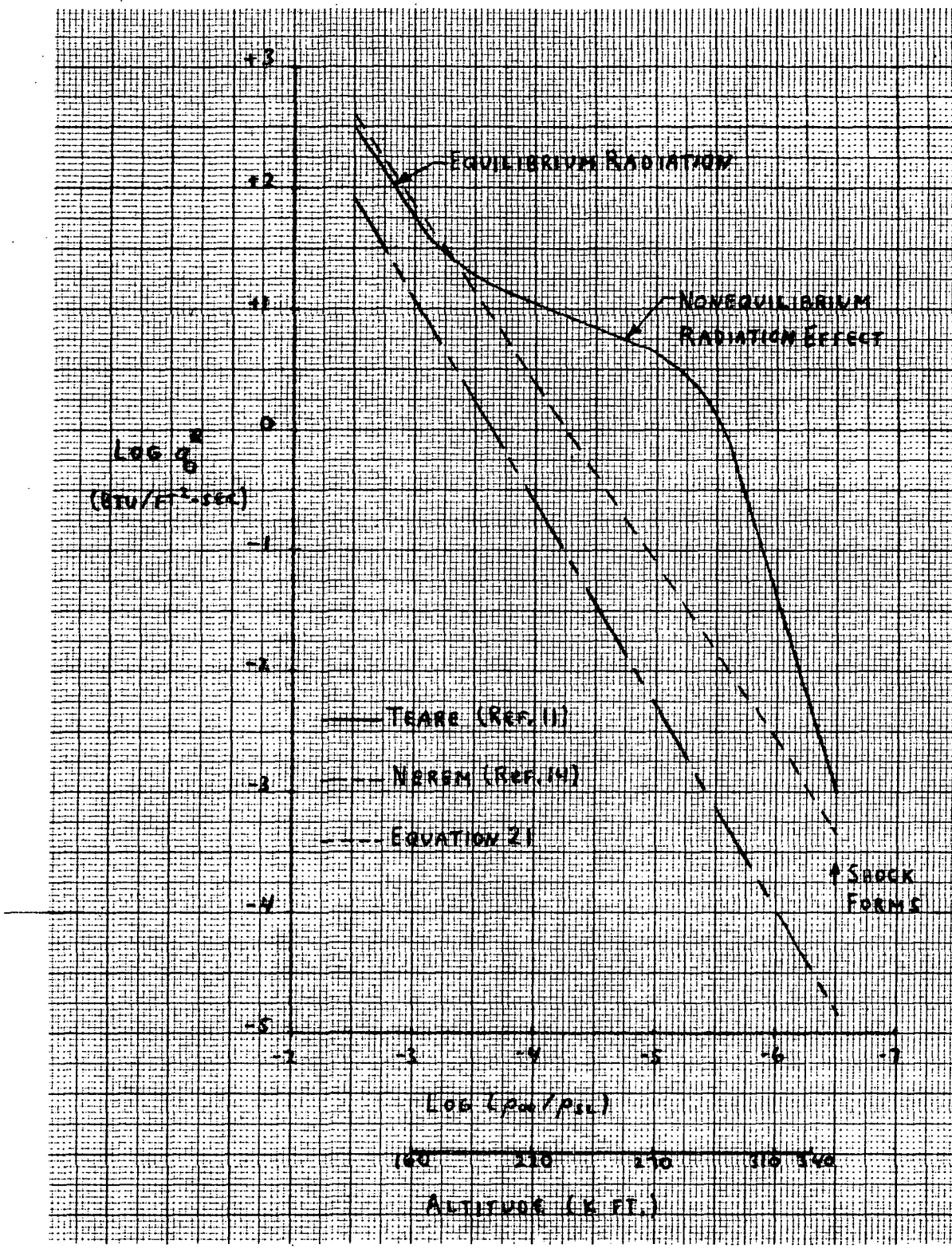


Fig 5. Comparison of the Total Radiative Heat Flux with Two Equations for the Equilibrium Radiative Heat Flux for V_∞ = 25,000 ft./sec. and R = 10 ft.

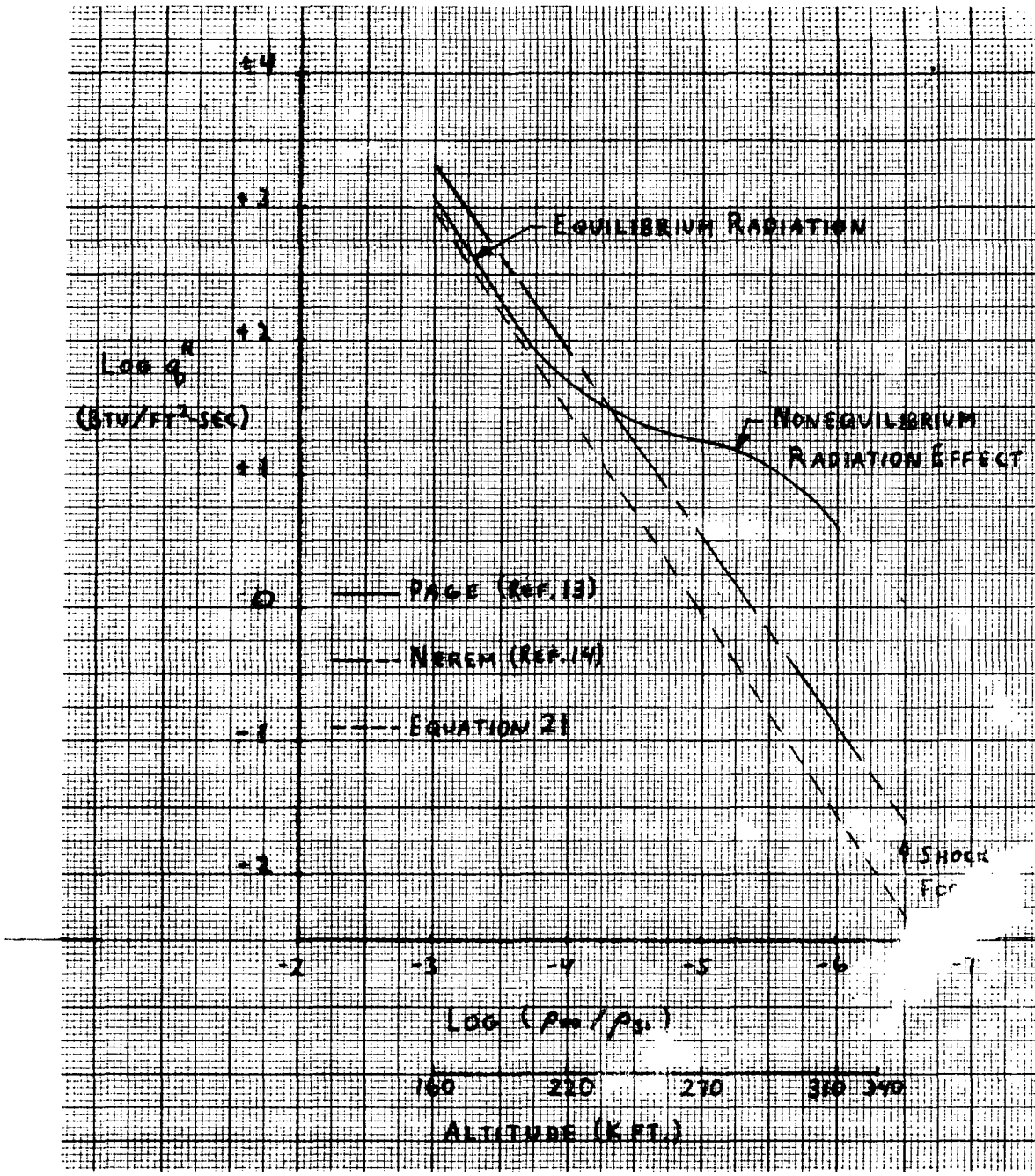


Fig. 6 Comparison of the Total Radiative Heat Flux with Two Equations for the Equilibrium Radiative Heat Flux for V_∞ = 36,000 ft./sec. and R = 3 ft.

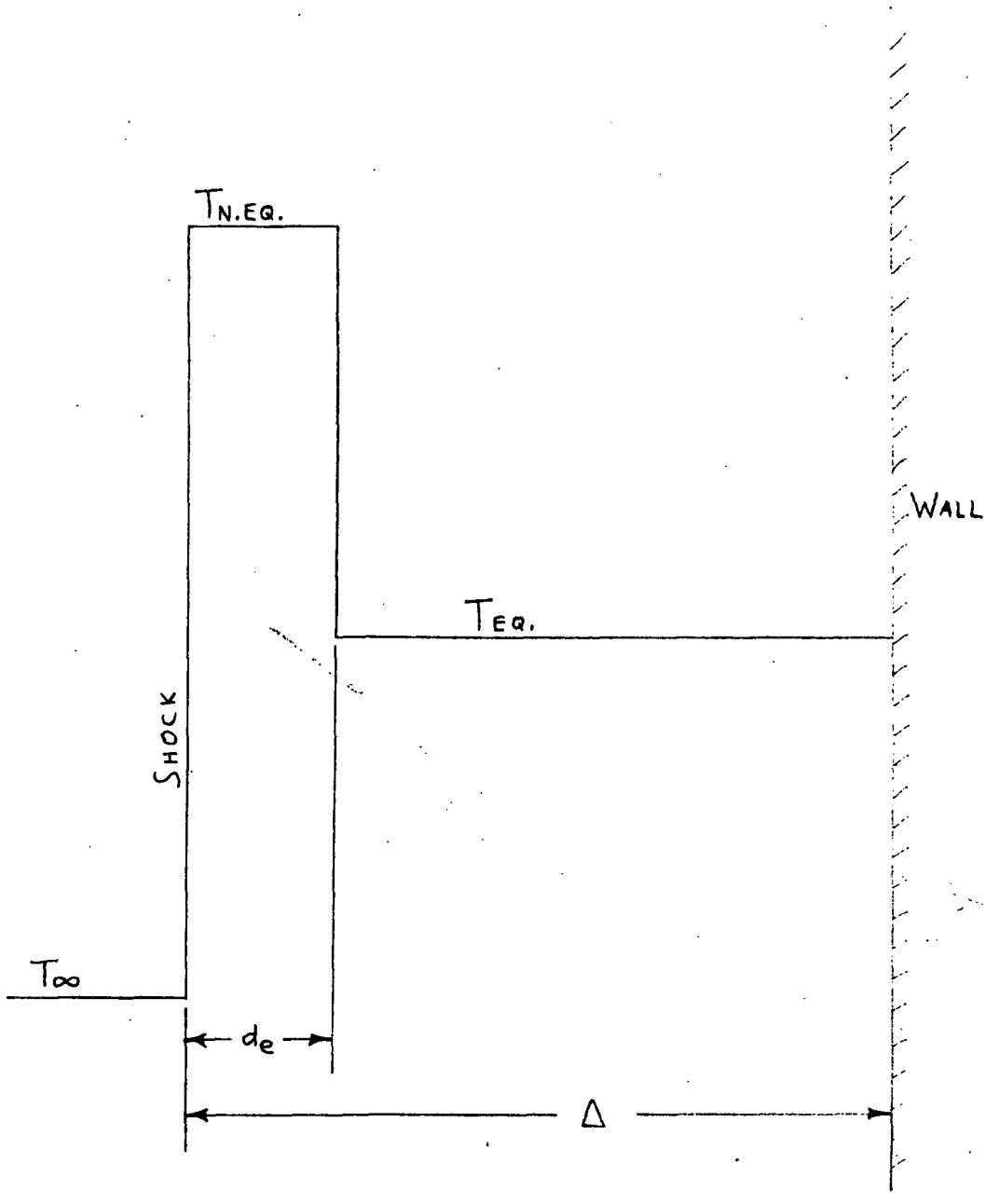


Fig. 7 Nonequilibrium Radiation Model

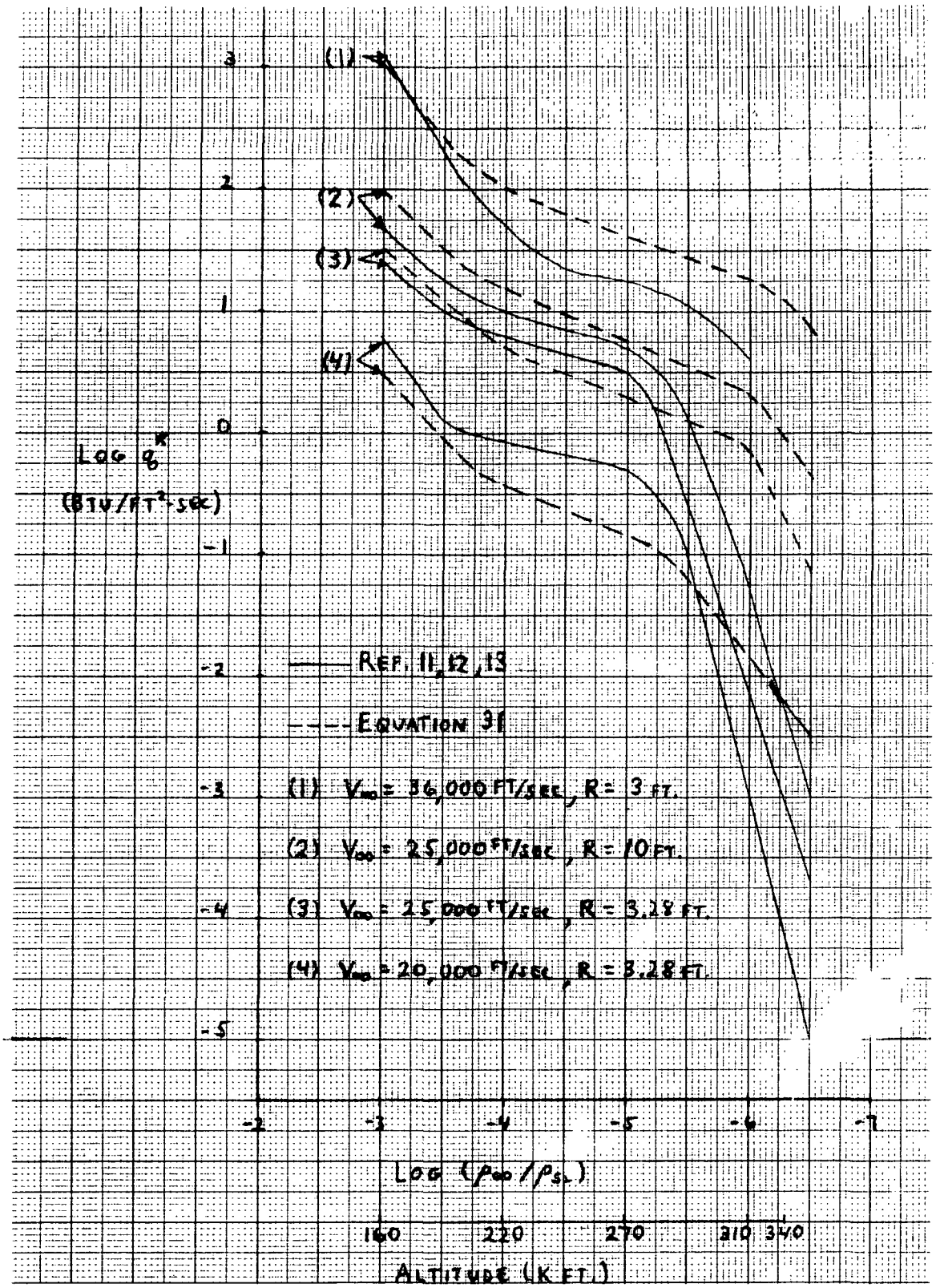


Fig. 8 Comparison of Data with Nonequilibrium Radiation Calculations (Eq. 31).

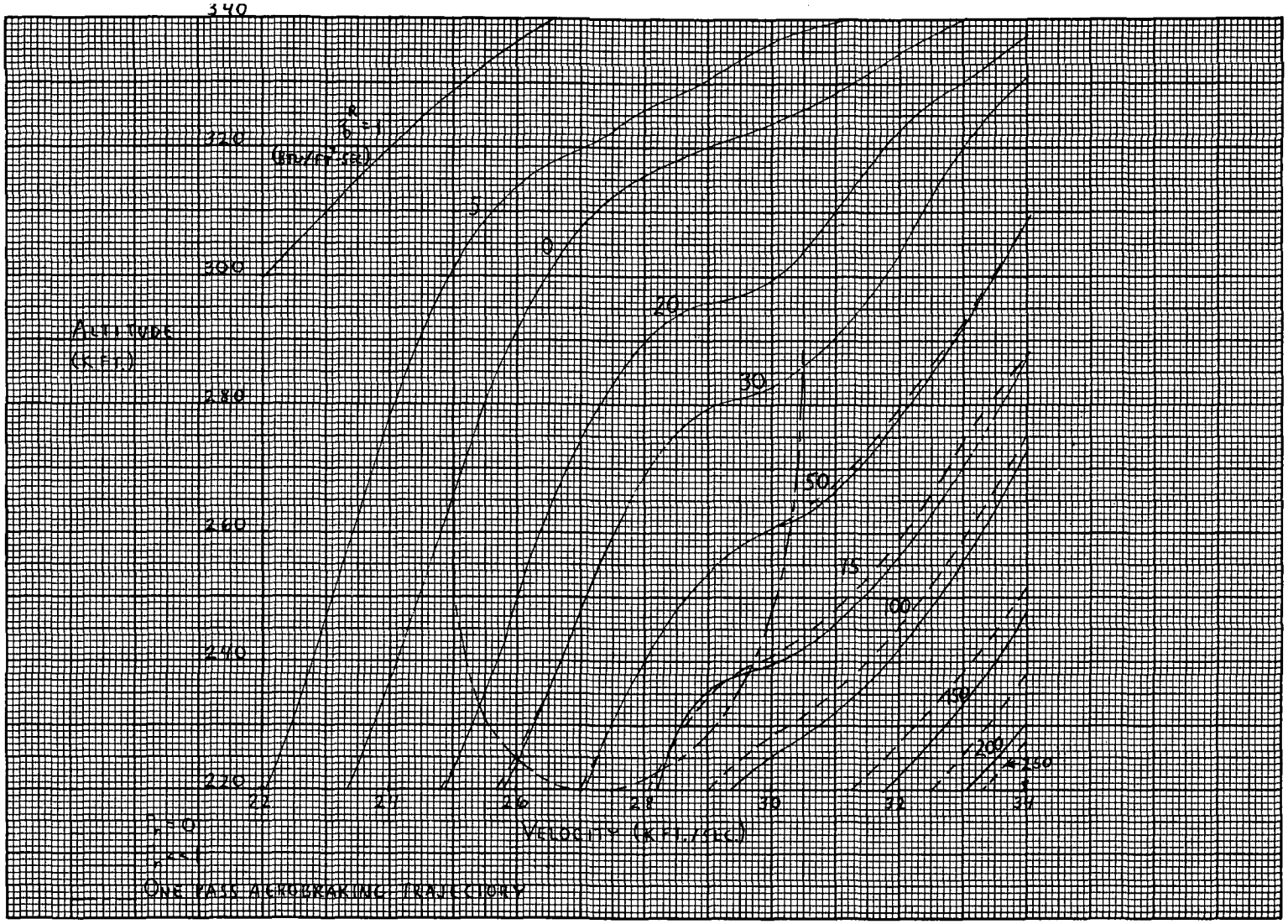


Fig. 9 Radiative Heat Flux with Self-Absorption ($\Gamma_n \neq 0$) for the Space Tug Flight Regime, R = 14 ft.

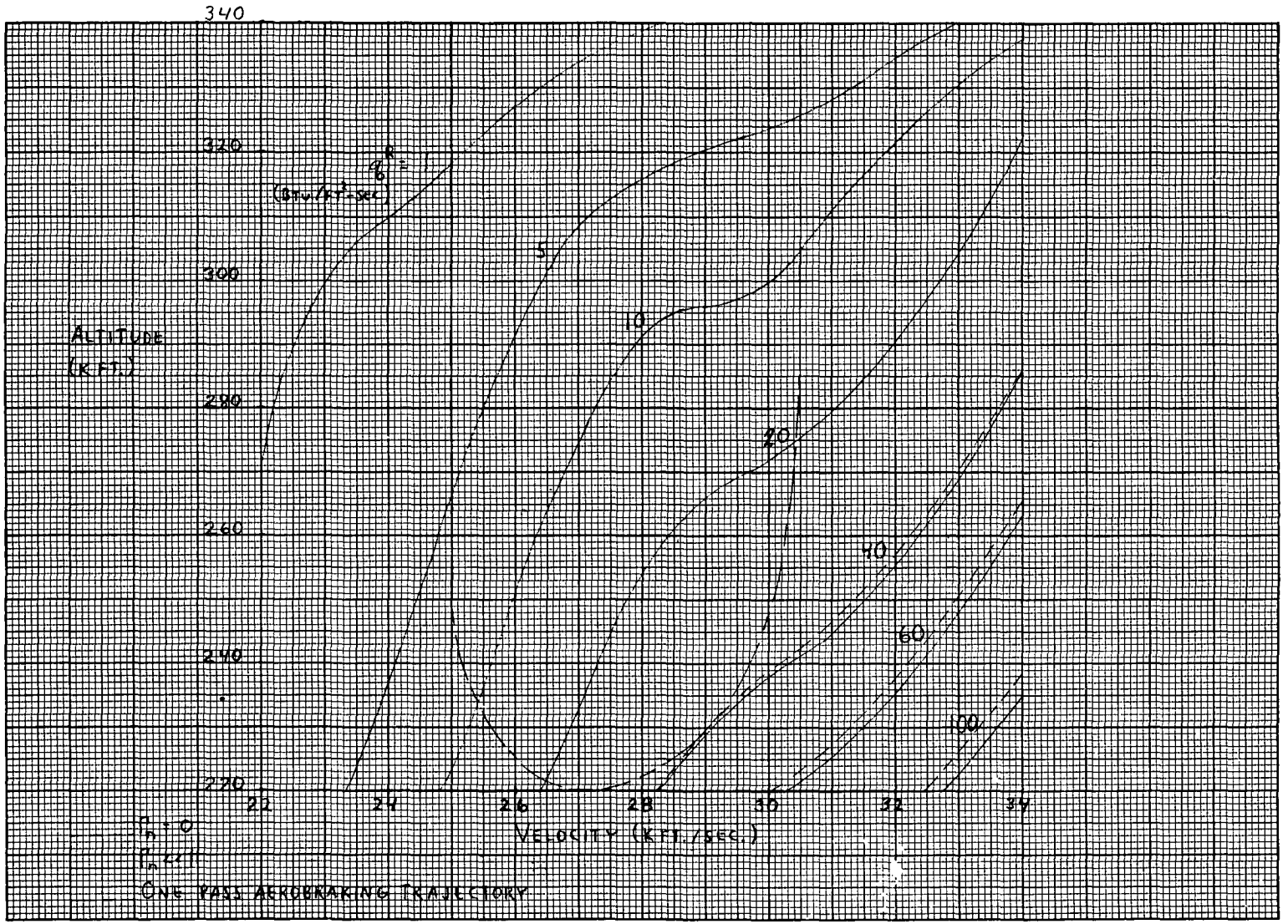


Fig. 10 Radiative Heat Flux with Self-Absorption ($\Gamma_n \neq 0$) for the Space Tug Flight Regime, R = 7 ft.

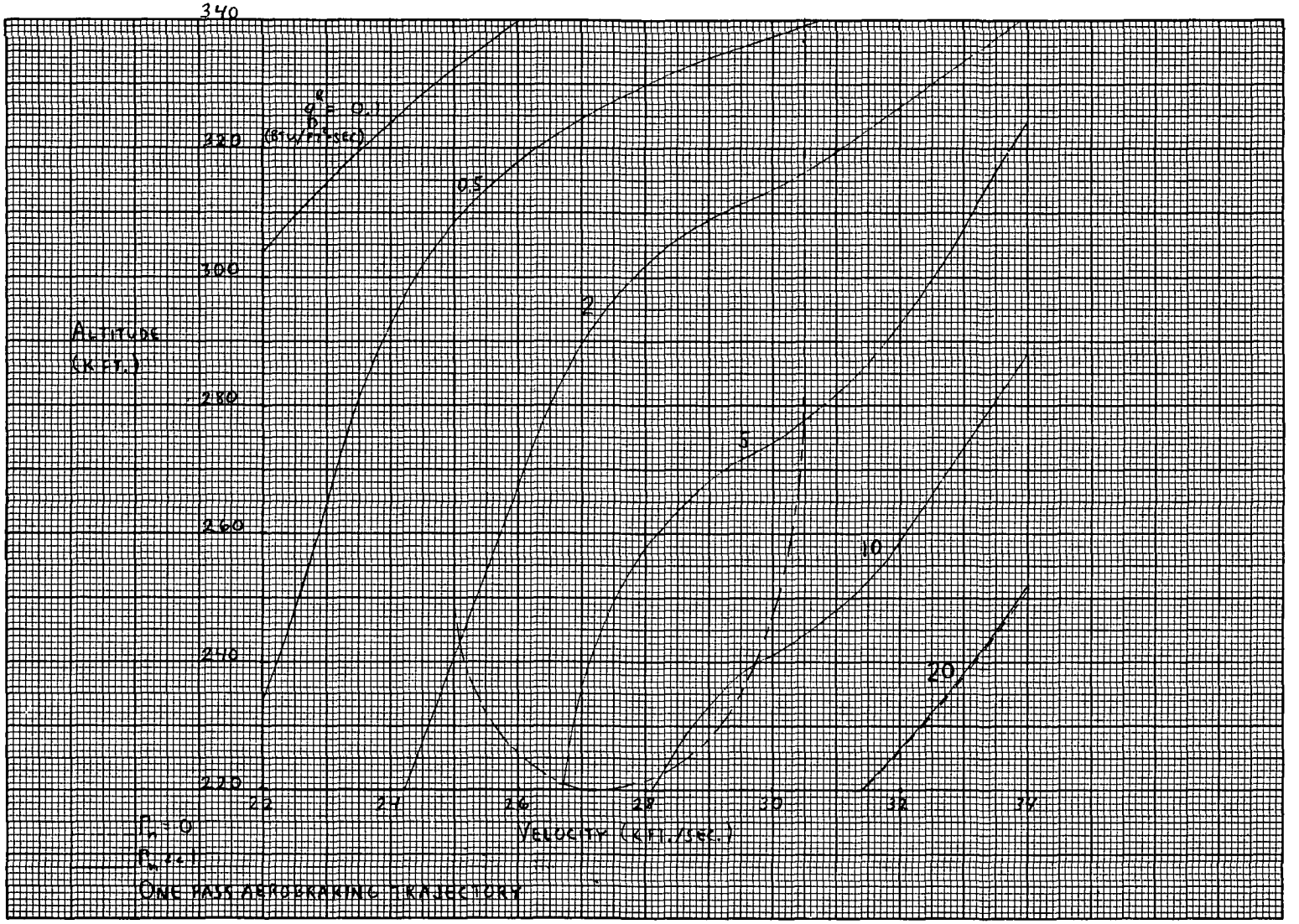


Fig. 11 Radiative Heat Flux with Self-Absorption ($\Gamma_n \neq 0$) for the Space Tug Flight Regime, R = 1.84 ft.

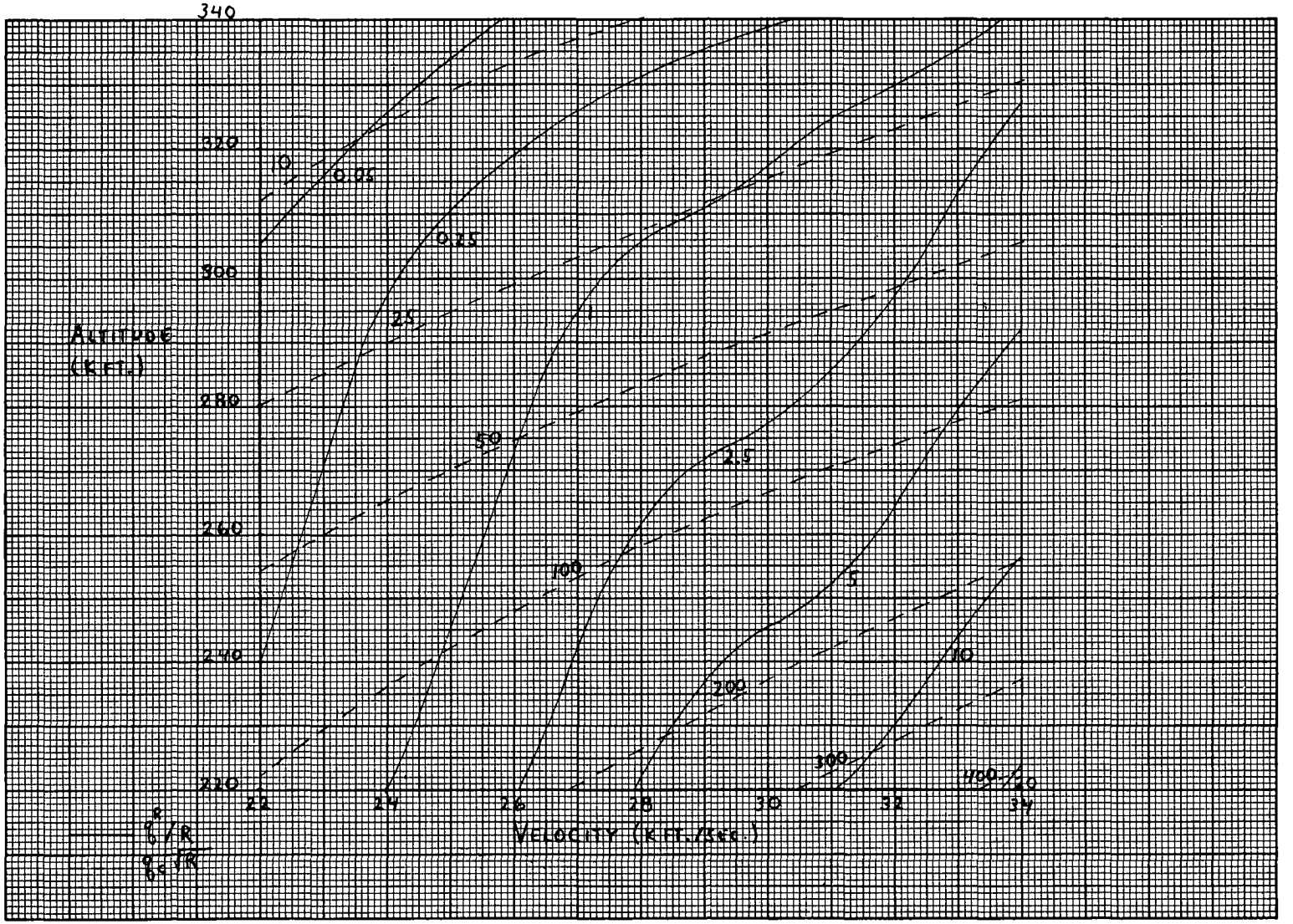


Fig. 12 Radiative and Convective Heating at the Stagnation Point for the Space Tug Flight Regime with $\Gamma_n \cong 0$

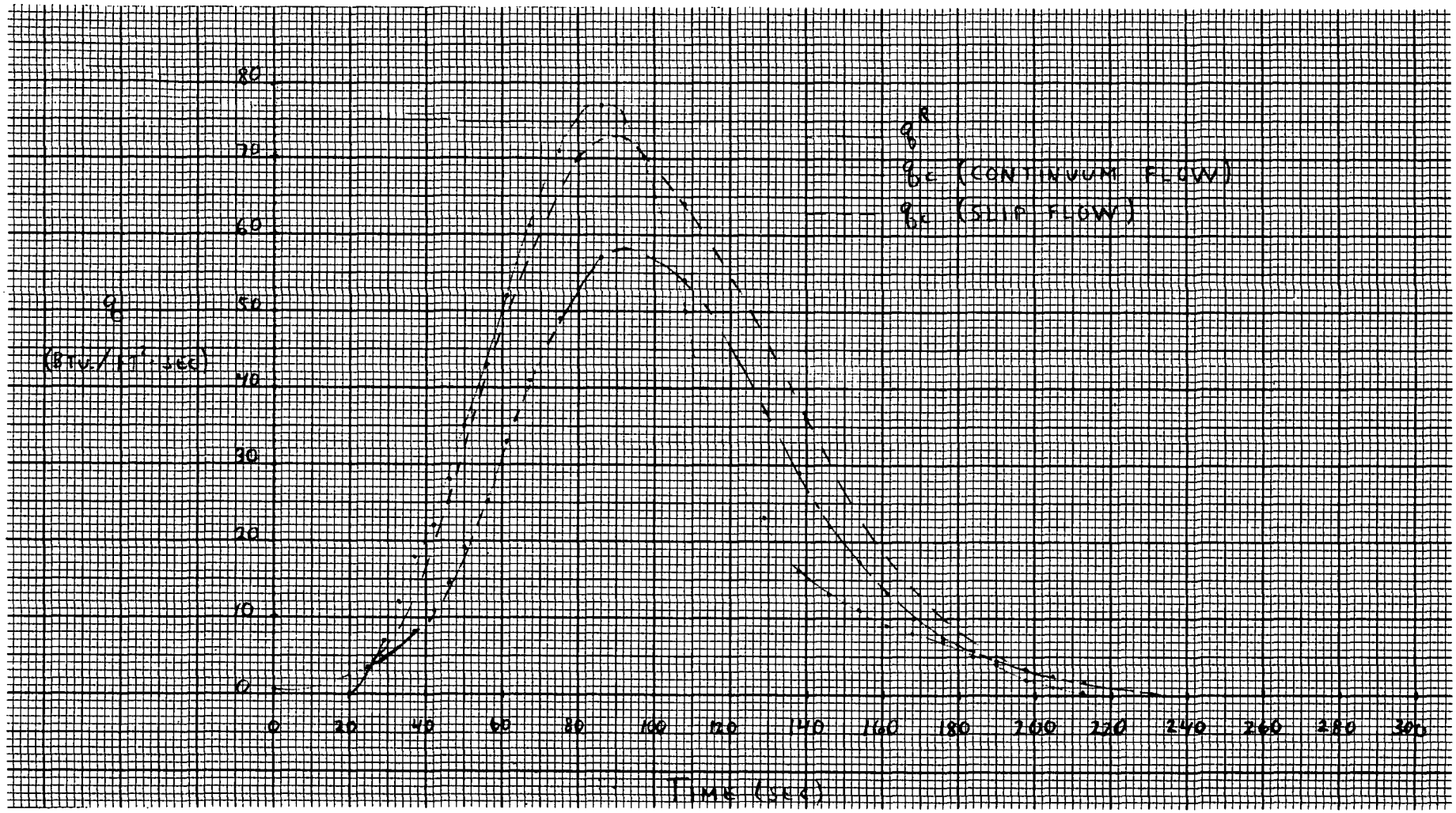


Fig. 13 Radiative and Convective Heat Flux for a One Pass Aerobraking Trajectory with $\Gamma_n \neq 0$, R = 14 ft.

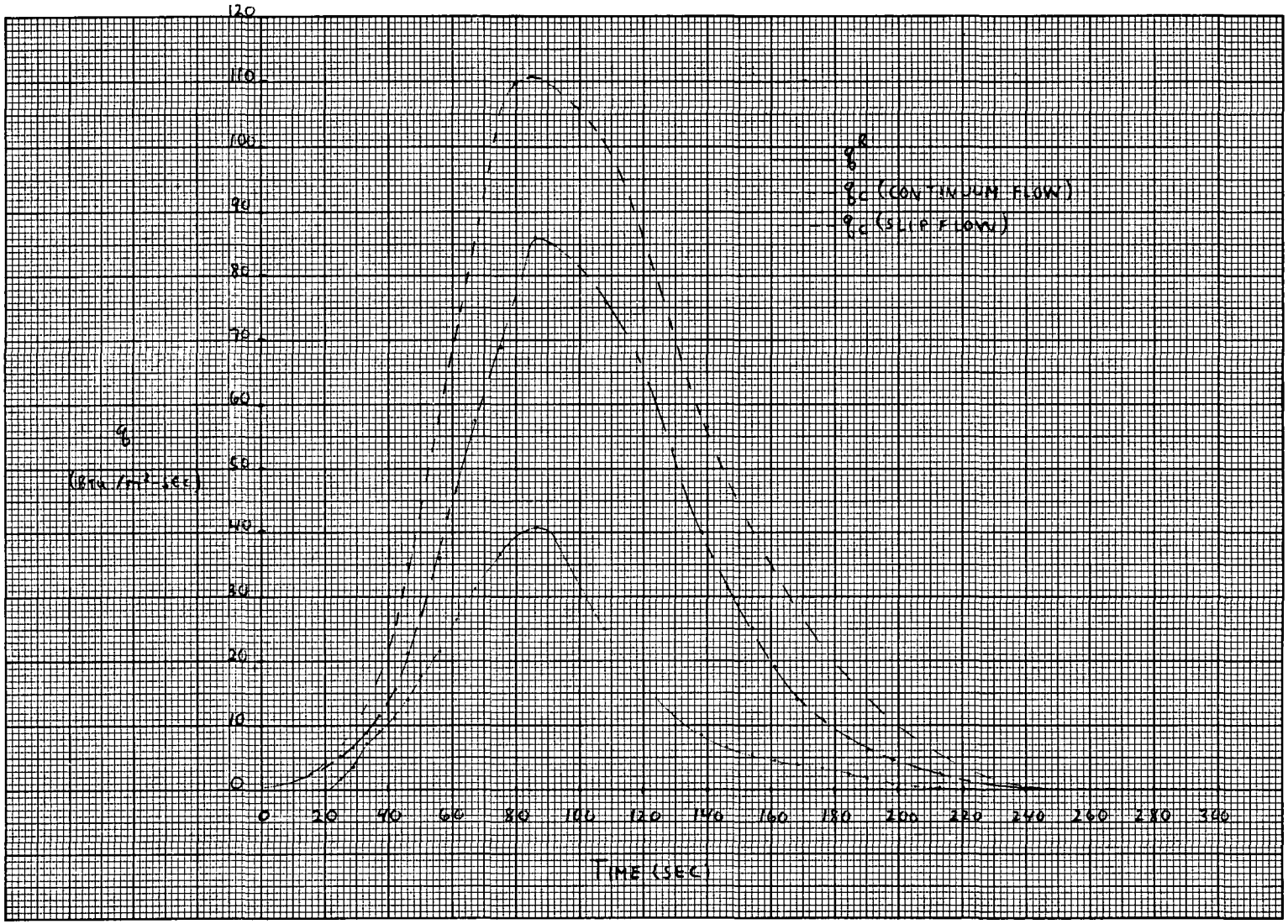


Fig. 14 Radiative and Convective Heat Flux for a One Pass Aerobraking Trajectory with $r_n \neq 0$, R = 7 ft.

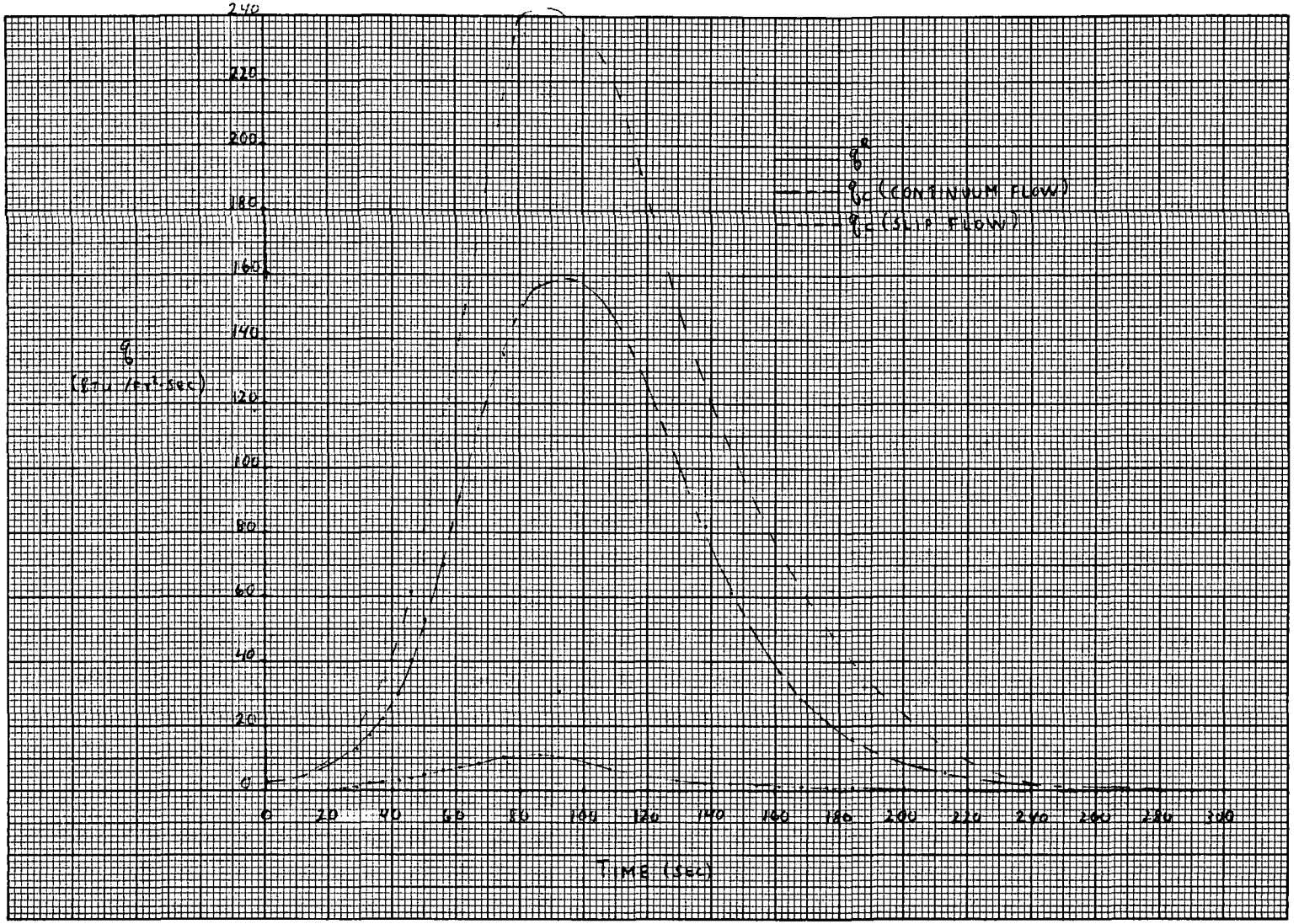


Fig. 15 Radiative and Convective Heat Flux for a One Pass Aerobraking Trajectory with $r_n \neq 0$, R = 1.84 ft.

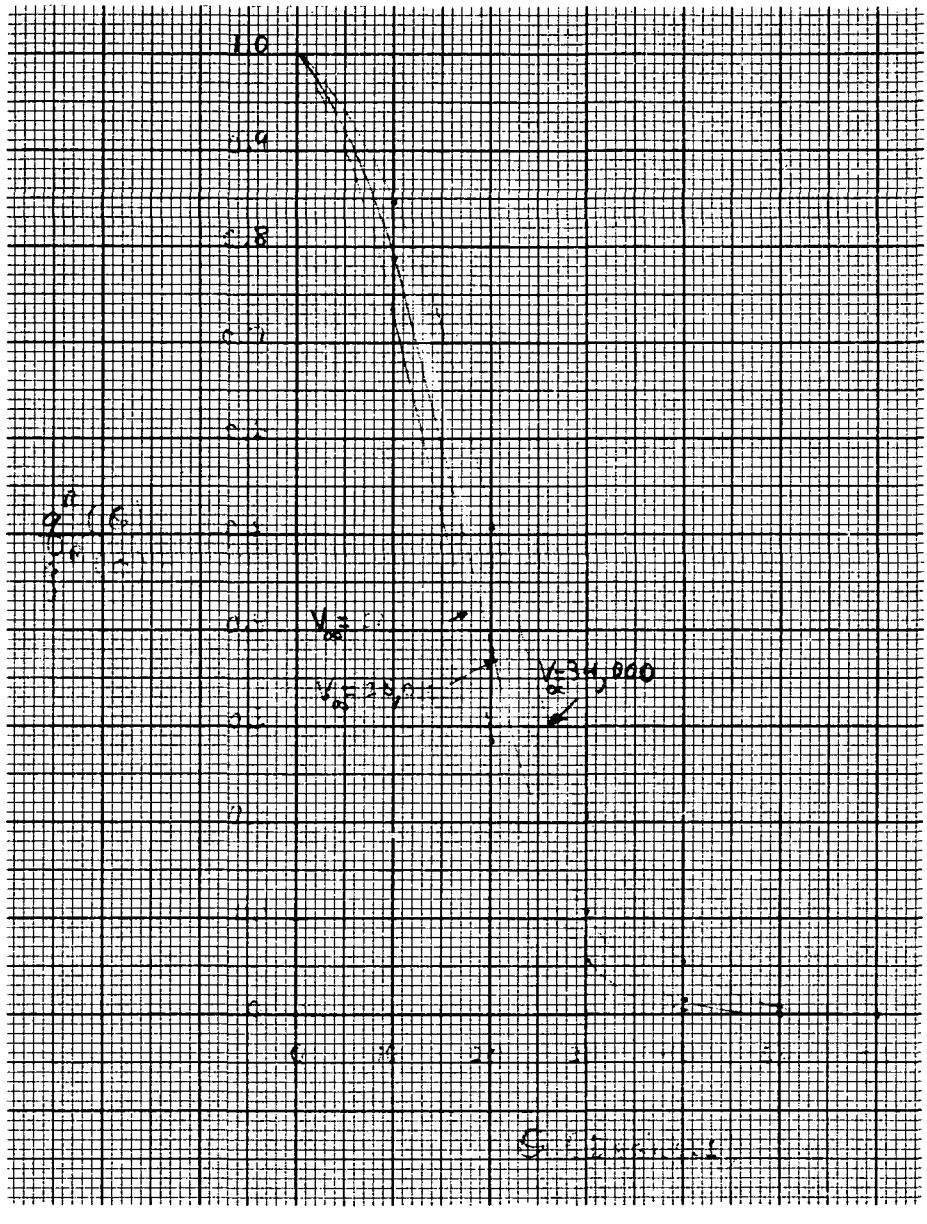


Fig. 16 Radiative Heat Flux Distribution Around A Sphere

REFERENCES

1. Preliminary Design Office Program Development;"Baseline Tug Definition Document," George C. Marshall Space Flight Center, Huntsville, Alabama, March 1972.

2. The Boeing Company, Huntsville, Alabama; "Space Tug Aerobraking Study," Volume II, Boeing Document No. D5-17142, Prepared under Contract NAS8-27501 for NASA/George C. Marshall Space Flight Center, Huntsville, Alabama, April 1972.

3. Olstad, Walter B., "Blunt-Body Stagnation-Region Flow with Nongray Radiation Heat Transfer — A Singular Perturbation Solution," NASA TR R-295, November 1968.

4. Goulard, Robert, "Preliminary Estimates of Radiative Transfer Effects on Detached Shock Layers," AIAA Journal 2, 494 - 502, March 1964.

5. Vincenti, Walter G. and Kruger, Charles H., Jr.; Physical Gas Dynamics, John Wiley, 1965.

6. Hayes, Wallace D. and Probstein, Ronald F., Hypersonic Flow Theory, Academic Press, 1959.

7. Yoshikawa, K. K. and Chapman, D. R., "Radiation Transfer An Absorption Behind A Hypersonic Normal Shock Wave," NASA TN D-1424, September 1962.

8. Pai, Shih - I, Radiation Gas Dynamics, Springer - Verlag, January 1966.

9. Chin, J. H. and Hearne, L. F., "Shock-Layer Radiation For Sphere-Cones With Radiative Decay," AIAA Journal 2, 1345, July 1964.

10. Allen, R. A., Rose, P. H. and Camm, J. C., "Nonequilibrium and Equilibrium Radiation At Super-Satellite Reentry Velocities," IAS 63-77, 1963.

11. Teare, J. D., Georgiev, S. and Allen, R. A., "Radiation From The Nonequilibrium Shock Front," AVCO - Everett Res. Rep. 112, 1961.

12. Camm, J. C., Kivel, B., Taylor, R. L. and Teare, J. D., "Absolute Intensity of Nonequilibrium Radiation in Air And Stagnation Heating at High Altitudes," AVCO - Everett Res. Rep. 93, 1959.

13. Page, W. A. and Arnold, J. O., "Shock-Layer Radiation of Blunt Bodies at Reentry Velocities," NASA TR R-193, April 1964.

14. Nerem, R. M., "Shock Layer Radiative Emission During Hypervelocity Reentry," Dissertation at the Ohio State University, 65-1210, 1964.

15. Anderson, J. D., Jr., "An Engineering Survey of Radiation Shock Layers," AIAA Journal 7, 1665 - 1675, September 1969.

16. Huber, P. W., "Hypersonic Shock-Heater Flow Parameters for Velocities to 46,000 Ft./Sec. and Altitudes to 323,000 Ft.," NASA TR R-163, 1963.

17. Ames Research Staff; "Equations, Tables, and Charts for Compressible Flow," NACA Rep. 1135, 1958.

18. Chou, Y. S. and Blake, L. H., "Energy-Momentum Coupling in Radiating Shock Layers About a Blunt Body," AIAA Journal 8, 1680 - 1686, September 1970.

19. Kemp, N. H. and Riddell, F. R., "Heat Transfer to Satellite Vehicles Re-entering the Atmosphere," Jet Propulsion, February 1957.

20. Engel, C. D., Private Communication, REMTECH Corporation, Huntsville, Alabama.

21. Corning, G., Aerospace Vehicle Design, Braun-Brumfield, Inc., Ann Arbor, Michigan, 1964.

22. Lasher, L. E. and Wilson, K. H., "Effect of Shock Precursor Heating on Radiative Flux to Blunt Bodies," NASA CR-1265, February 1969.

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

THE PROSPECTS OF UTILIZING REMOTELY SENSED DATA IN THE PREPARATION OF ENVIRONMENTAL-SOCIOLOGICAL MODELS

Prepared by:	Robert B. Honea
Academic Rank:	Assistant Professor
University:	East Tennessee State University
Office:	Environmental Applications
Research Counterpart:	James C. Derington
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

LIST OF FIGURES

THE PROSPECTS OF UTILIZING REMOTELY SENSED DATA IN THE PREPARATION OF ENVIRONMENTAL-SOCIOLOGICAL MODELS

By

Robert B. Honea

ABSTRACT

Many of the problems presently confronting society are of such inordinate complexity that it will take most, if not all, of our technological abilities to assemble data, to analyze and effect solutions. To force society to accept the proposition that nature is a process, that it is interacting, that it responds to laws, and that it represents values and opportunities for human use with limitations and even prohibitions, is one of our major objectives. Beyond this realization, however, less evangelical concepts must be considered. This research deals with such concepts.

With the establishment in 1969 of a new national policy to protect the environment, the various agencies of Federal, State and local governments and other concerned public and private organizations, charged with the responsibility of preparing environmental impact statements, became acutely aware that present data acquisition systems either lacked the necessary detail or were so costly and tedious as to negate their use in impact studies.

The use of remote sensing to acquire needed data concerning soils, water conditions, micro-climatic conditions, flora, land use, recreational potential and cultural conditions appears to hold great promise in solving data needs. This report demonstrates the procedures and means whereby such data can be acquired and processed into an information format.

INTRODUCTION

Many of the problems presently confronting society are of such inordinate complexity that it will take most if not all of our technological abilities to assemble data, to analyse and to effect solutions. Yet, man continues to change the earth. By political process or by pragmatic management decisions, he may exercise a certain degree of control. However, our industrial society forces him to build dams, excavate tunnels, build highways, and develop residential tracts somewhere. If population is allowed to increase and the same population aspires to the western "high standard of living" then society is committed to the intensive use of the earth and consequently must prepare for it.

Despite evangelical attempts to convince society to moderate population growth and cultural development, we must remember that progress is "charging on" and rarely waits for anyone. Nevertheless, to force society to accept the proposition that nature is a process, that it is interacting, that it responds to laws, and that it represents values and opportunities for human use with limitations and even prohibitions, is one of society's major obstacles (61). The purpose of this research, does not consider this problem directly, but rather, how the necessary data to assess the proper relationship between man and nature might be acquired.

The impact of population growth on the quality of the environment and the attending problems thus created, depends largely on the present and future patterns of land use. Seventy-five per cent of our population now reside in urban areas and each year, an estimated 420,000 acres of land are consumed by urban sprawl (22; p. 1). If present trends continue, 90 per cent of our population will live in a few mammoth conurbations. These concentrations will heavily tax electrical power, transportation, sanitation and other public services. We must develop planning programs which will make urban life fulfilling rather than frustrating. At the same time we must make rural living more attractive to prevent migration to the cities.

OBJECTIVES

With the establishment of the National Environmental Policy Act of 1969 (18), various agencies of federal, state and local governments and other public and private organizations were charged with the responsibility of preparing environmental impact statements. These agencies soon became acutely aware that present data acquisition systems either lacked the necessary detail or were so costly and tedious as to negate their use in impact studies. Since all development or redevelopment begins from a knowledge of the present position, the two immediate and most obvious essentials to properly assess environmental impact are: (1) an inventory of the present position and (2) an understanding of the reasons for that position.

Accurate, up-to-date land use information is a necessity in making wise decisions on matters relating to the physical development of a community. These decisions are related to such matters as housing needs and urban renewal; location of schools, parks, playgrounds, and cultural facilities; transportation and parking needs; opportunities for industrial expansion; zoning; as well as the guidelines for environmental protection.

The use of remote sensing to acquire needed land use data concerning agriculture, soils, water conditions, micro-climatic conditions, flora, fauna, recreational potential and cultural conditions appears to hold great promise. This report considers the various means whereby land data in its numerous forms may be extracted from multi-spectral photography and imagery, and processed into an information format for land classification and environmental assessment.

Before discussing these various systems, it is worth considering the historic and methodological background of land classification and environmental assessment. Indeed, the bulk of this summer's research has involved a search of the literature on the subjects of land classification, land use mapping and environmental assessment. The following section is included as background material pertinent to the research effort.

TAXONOMY, LAND USE CLASSIFICATION

AND

ENVIRONMENTAL ASSESSMENT

The types of phenomena which science considers can be classified in many ways. Our concern here is with methodology of classification and the end to which classification schemes are put. This sets the stage for the discussion which follows.

The Methodology of Classification

Classification is an initial step in ordering knowledge and the deriving generalizations in any scientific methodology. It is necessary "to study and remember facts, discover relationships and develop fundamental principles"(69; p. 4). This basic dictum of all inquiry probably stems from the Aristotelian principle: Every object must ultimately be placed in a class on the basis of its relationship to other objects (7; p. 421).

The measure of success for any classification is determined by the clarity of definitions, selection of pertinent and measureable attributes, and the use of the product restricted to predetermined purposes. The importance of purpose cannot be emphasized enough. I can not think of a better example to bring this point home than a story related by Russell Ackoff concerning the importance of definitions (1).

Ackoff was once involved in determining the accuracy of results obtained in an urban survey designed to determine the number of persons per room in certain dwelling units. The survey had been conducted without an explicit definition of "room." Ackoff met with the designers of the survey and asked what definition they had used implicitly. They were impatient with the question, observing, "Everyone knows what a room is." Ackoff persisted, and one of those present offered: "A room is a place enclosed by four walls, a floor, and a ceiling." Having maneuvered the designers into this position, Ackoff delightfully proceeded as follows:

> "Can't a room be triangular?"
> "Sure. It can have three or four walls."
> "What about a circular room?"
> "Well, it can have one or more walls."
> "What about a paper carton?"

"A room has to be large enough for human occupancy."
"What about a closet?"
"It must be used for normal living purposes."
"What are normal living purposes?"
"Look, we don't have to go through this nonsense; our results are good enough for our purposes."

Ackoff refused to let go.

"What are your purposes?"
"To get an index of living conditions by finding the number of persons per room in dwelling units."
"Doesn't the size of the room matter?"
"Yes, we probably should have used square feet of floor space, but that would have been too hard to get."
"Doesn't the height of the room matter?"
"I guess so. Ideally, we should have used volume."
"Would a room with 10 square feet of floor area and 60 feet high be the same as one with 60 square feet of floor space 10 feet high?"
"Look, the index is good enough for the people who use it."
"What do they use it for?"
"I'm not sure, but we've had no complaints."

(1; pp. 47-48)

This amusing conversation makes clear that without an explicit statement of the purposes of the inquiry there are no criteria for defining.

In defining objects for the purpose of classification it is necessary to define the class of objects which are to be classified. This means specifying the essential properties necessary for determining whether or not a particular object is a member of the class of interest. The "essential properties" of an object are properties which are individually and collective sufficient for the inclusion of the object in the class to be counted (1; p. 154).

"The definitions of a class of objects, then, should consist of a specification of properties, each of which is necessary and all of which are sufficient for differentiating the class of interest from all other classes. The selection of these properties should be dictated by the research objectives, and the name by which the class is identified should be chosen with past and present terminology in mind" (1; p. 154).

The elementary essence of the above statements is that ultimately all classes of objects are defined by the objects contained therein. This may be taken one step further by saying that any object may be identified as belonging to a class by an adequate predicational statement. For example:

"A field of corn is a member of the set of agricultural land uses identified as cropland."

For the purposes of mapping agricultural land use this statement may be sufficient such that anyone will understand that a corn field will be classified as cropland. Inevitably, however, some uncertainty will arise which will necessitate further refinement of the definition and we are "off again" on a circulur argument which may not end until we have reduced the definition to a long series of monadic statements which may include a definition of every subject and adjective contained in the original definition offered. We can take confort, however, from the assurance of Frederick Schiller, that "A single, unmistakeable absolute definition . . . is not extant" (75; p. 71).

Land Use Classification

In light of the previous discussion, it would appear that the most logical places to begin discussion would be to define "land use" or more specifically "land." Unfortunately this is not easy as there are probably as many definitions for land as there are attributes of the land. The most common picture envisioned is the "stuff of the land," i.e.,the aggregate of soil, rocks and vegetative matter that make up the solid surface of the earth. Yet, the soil scientist is likely to have his definition, the economist his, and the geographer his "pet" definition, all of which are slightly different. Perhaps the practical way to avoid difficulty is to cite the previous discussion and simply state that the definition of land depends largely upon the research objectives.

"Land use" has been studied and the term used so many different ways that some specification as to its meaning is desirable. Clawson (14; pp. 14-28) lists nine concepts about land which are often considered as land use:

1. Location
2. Activity on the land
3. Natural qualities of the land
4. Improvements to and on the land
5. Intensity of land use
6. Land tenure
7. Land prices, land market activity and credit

8. Inter-relations in land use between different tracts of land or accessibility

9. Inter-relations between activities on the land and other inclusive social developments.

Clawson and the Committee on Land Use Statistics prefer the definition of "land use" to refer to "man's activities on the land which are directly related to the land" (14; p. 29). This concept provides a surrogate to other relevant characteristics of the land such as land quality, intensity of use, and improvements on the land.

"Land cover" is often thought of in the context of land use but the two should not be equated. The term "land cover" is generally interpreted to refer to basic objects as they appear on the earth's surface. For example, a field strewn with rusting hulks of wrecked automobiles would be identified as such, and not as a "junk yard." The term encompasses a more pure taxonomic concept where a separate category is provided for every land use characteristic of the earth's surface. Several investigators have proposed the adoption of such a system (13 and 18). This philosophy would permit the land surface to evolve the land use classification system, rather than the researcher.

In 1941, the National Resources Planning Board identified five types of land classification:

1. Land classification in terms of inherent characteristics;

2. Land classification in terms of present use;

3. Land classification in terms of use capabilities;

4. Land classification in terms of recommended use; and

5. Land classification in terms of program effectiveness

(65; p. 5)

Although some dispute arose over whether the fourth and fifth types should be included (63 and 69; p. 5) the important information to note is that land classification encompasses more than land use classification. Nunn suggests three types of land classification: physical classification, economic classification and institutional classification, whereby land use would be basic to each (69; p. 6).

It is doubtful that any universal land classification system will be developed. The development and use of any land classification scheme must initially be dictated by the purposes of the study. Broadly,

the objectives of land classification are to serve as an effective tool for resource study, evaluation and determination of uses and producing maximum net income to the individual and society (69; p. 7). In later discussion some examples of land classification will be cited which will illustrate this point.

The Role of Land Classification in Environmental Assessment

Environmental assessment, land classification, land use mapping or whatever the organized activity may be called seems to be cyclical with periodic "surges in activity too commonly occurring when land problems are accute" (69; p. 1). Current interest stems from the realization that man's total environmental system includes not only the biosphere but also his interaction with his natural and man-made surroundings.

Recent alarm has stimulated federal response with President Nixon calling for the "development of a National Land Use Policy to be carried out by an effective partnership of Federal, State and local governments" (22; p. xiii and 51; p. 6). The upsurge of interest in land use manifests itself in obvious changes around us, i.e., the encroachment of suburban areas upon what was predominantly rural land, the demand for better water and recreation sources, urban renewal of slum areas and redevelopment of central business districts.

Although the desire to intelligently classify land stimulated many early studies, the first effort in the United States may have occurred with the geographical and geological expeditions of Hayden, Wheeler, King and Powell between 1867 and 1879 (76 and 69; p.2). Upon his return, Major Powell attempted to secure national legislation which would zone the use of large agricultural areas in the southwest. Had his ideas been accepted, some of the major problems of the"dust bowl"area may have been avoided.

Since that time, a series of national acts have continued to spur interest in land use. In 1879 the U. S. Geological Survey was created to collect and classify "land information useful to geologists, miners, engineers, farmers and lumbermen" (69; p. 2). Other Congressional legislation stimulating land classification activity include: the Carey Act (1894), the Reclamation Act (1902), the Enlarged Homestead Act (1909), the Congressional Act of March 3, 1891 (first forest preserves), the Stockraising Homestead Law (1916), the Taylor Grazing Act (1934) and most recently, the Environmental Policy Act (1969).

Aside from the Hayden, Wheeler, King and Powell expeditions, many of the earliest studies involving the study of land use were undertaken by geologists, and geographers who were motivated by a methodological

interest in accurately classifying and recording the phenomena observed in the field. Included in this group were well known geographers such as W. D. Jones, C. C. Colby, D. S. Whittlesey, and C. O. Saver, and V. C. Finch to name a few.

Among the more significant contributions made at this time was that of V. C. Finch in his Montfort Study (29). Finch introduced the system that is now known as the "fractional code" method for recording complex field data. Although evolved before 1930, Finch's work was not published until 1933.

The significance of this contribution lies in an understanding of the techniques one was forced to employ in field work during that period. Any intensive study of an area meant tedious field mapping which included the preparation of base maps, because aerial photography and detailed planimetric maps were not widely available at that time. The basic system employed a composite fraction in which the numerator was employed to reflect the cultural and economic qualities of the land parcel, such as land use; while the denominator was used to record the physical characteristics of the land such as slope, soil type or drainage conditions. Delineations of individual parcels were usually governed by any significant change in any one or more of the elements noted in the fraction. The smallest unit delineated, however, was about two acres.

Subsequent land classification work by the Tennessee Valley Authority also employed the fractional code method of recording. The majority of this work was carried out under direction of G. Donald Hudson in the period between 1932 and 1936.

Hudson made some significant improvements in the fractional code method. For the first time surveyors were provided with aerial photographs upon which they could record field data. In addition, TVA was simultaneously producing planimetric maps of the entire valley which greatly facilitated the mapping effort. More important, however, was the fact that the detailed field data was incorporated into a scheme to produce a classification of land in the Valley in terms of five categories based on the economic conditions of the people and the physical conditions of the land (40). Using a minimum of 200 acre units, parcels of rural land were rated from Class I, which included land with no significant agricultural problems, to Class V land, which in most cases was suitable only for forest land, recreational purposes, or game preserves (40; p. 1).

Hudson lists four major contributions the land classification program could make: (1) It could aid in specific departmental decisions of the Authority such as the evaluation of one reservoir site as contrasted with another; (2) With respect to the regional development program, it could furnish data essential to the solution of such

problems as the location of land that should be put in public ownership; (3) It could furnish information helpful to local, state and federal agencies in such problems as employment relief, highway construction and public works; and (4) It can make available information helpful to private commercial and industrial organizations (40; p. 8). Unfortunately the land classification program was halted in the late 1930's because of more pressing problems. Hudson left to become the Chairman of the Department of Geography at Northwestern University where he influenced another land classification program, the Rural Land Classification of Puerto Rico.

It is worth mentioning at this point several other significant land classification programs. One such program is the soil and land classifications conducted by the Soil Conservation Service and cooperating State Experiment Stations. Modern soil surveys were conducted as early as 1917 but it was not until the erosion and production control programs of the 1930's that soil surveys were incorporated into land classification schemes. These actions brought into being the well-known Land-Capability Classification (69; p. 6).

Soil surveys are conducted in order to understand the behavior and production capabilities of various soils under certain cultivation practices and physical conditions. Fine sandy-loam soils might perform very well when cultivated on flat to gentle slopes but will perform quite differently on undulating to hilly slopes.

The Land-Capability Classification scheme recognizes eight land-capability classes. The first four classes are considered safe for cultivation with varying degrees of conservation practices; the second four are considered to be suitable only for varying amounts of grazing, forestry; while the eighth class of land is suitable only for wildlife, recreation, or watershed purposes.

In addition to the Soil Conservation Service, other federal agencies have been involved in selected land classification studies, two of which are the U. S. Forest Service and the U. S. Bureau of Reclamation. The objectives of the Forest Service primarily involves estimating: (1) the location, extent, kind, and availability of timber supplies; (2) the present and potential productivity of forested areas; (3) the depletion and replenishment rates of timber stands; and (4) information administratively useful for using, protecting and developing forest resources and industries (69; p. 14).

The Bureau of Reclamation has a more specialized charter. With the Fact Finders Act of 1924, Congress specified "that lands proposed for irrigation development must be classified with respect to their capacity under a proper agricultural program to support a farm family and pay water charges" (64; p. 21). Six classes of land were recognized ranging from Class 1 land which was considered to have a high

potential return capacity for irrigation and grading to Class 6 land, which was considered to have no potential for irrigation at all.

Numerous Land Grant Universities as well as other agricultural schools have evolved economic land classification schemes which are designed to measure income potential of farm operations over long-time periods and thus provide a relative success/failure appraisal of any particular farm. One notable system was developed by Cornell University and is similar to the TVA land classification system except that it is intended to be applied to specific farms (20).

There have been numerous land use mapping or land inventory projects undertaken in the last four decades, some of which represent major undertakings. It is not possible to discuss each except to mention a few: Derwert Whittesey's Major Agricultural Regions of the Earth (93); Francis Marshner's Map of the Major Land Uses in the United States (60); Clarence F. Jones' Rural Land Classification Program of Puerto Rico (48); Sir Dudley Stamp's Land Utilization Survey of Great Britain (15) (currently being redone for the second time); the International Geographical Union's magnanimous World Land Use Survey (23, 44, and 45); The Canada Land Inventory (25); and the New York State Land Use and Natural Resources Inventory. (There are many other studies which are discussed in various publications cited in the bibliography.)

In order to provide insight into the recent state-of-the-art of land inventoring, the New York State Land Use and Natural Resources Inventory has been chosen for more thorough discussion. In 1966, the Center for Aerial Photographic Studies at Cornell University was directed to conduct a state-wide land use and natural resources survey by the Office Planning Coordination of the State of New York. This inventory is now complete. The main objective of this inventory was to produce land use maps for the entire state. These maps consist of overlays depicting land use data on standard 1:24,000 series, U.S.G.S. topographic maps.

The unique aspect of these maps is that they were produced almost entirely by aerial photographic interpretation of black and white photography utilizing a classification of 120 land uses. To accomplish the task a team approach was employed to support the photo interpretation section. One field team was designated to check interpretation results, while a second group was designated to supply supplemental data. A third team was responsible for drafting and geocoding the data for computer storage (53; p. 101).

Kreig, in discussing the project, makes a point in distinguishing between photo interpretation and photo reading or photo analysis (53; p. 103). The distinction in essence is that the decision time involved

in photo reading is very small while photo interpretation involves inductive and deductive reasoning which takes considerably longer. By emphasizing photo reading rather than photo interpretation the photographic area a person can cover in a given length of time is significantly increased. In addition, the skill requirements of the photo interpreter are kept to a minimum. This rapid interpretation ability is further augmented by providing "easily accessible, useful supplemental data" and designing "the classification so that categories can be easily separated by the interpreter" (53; p. 103).

The classification system was considered the most important factor in success of the land use mapping project. "It must not only enable the interpreter to keep decision time down in utilizing it, but the classification system must also group land uses in such a manner as to be useful to planners and others. . ."(53; p. 104).

The classification scheme consisted of three levels; each first level was broken down into "area" data and "point" data. A sample for agricultural land use is shown below:

Agriculture

Areas:

Orchards

Vineyards

Horticulture, floriculture

Specialty farms

High-intensity cropland

Cropland and cropland pasture

Permanent pasture

Inactive agricultural lands

Other inactive lands

Lands under construction

Point data:

Specialty farms (Ay): Type Present
Mink (y-1)
Pheasant and game (y-2)
Aquatic agriculture (y-5)
Horse farms (y-6)

Dairy operations (d): number

Poultry operations (e): number

Active farmsteads (f): number

In order to effectively utilize the land inventory data for sub-regional and state-wide purposes, the land use data was generalized and geocoded by kilometer grid cells for computer storage and retrieval. This permits greater flexibility of the data and allows one to calculate complex spatial corelations.

Summary

In summarizing the above discussions, some obvious points should be noted. Firstly, taxonomic considerations and operational definitions are necessary to develop any classification scheme. Lucid and precise definitions for objects under consideration as well as a clearly defined purpose are absolutely essential or the research may be to no avail.

Secondly, land classification will always be a timely subject. The surface of the earth is ever changing and man's perception of the world around him is constantly evolving. Therefore, we will be constantly faced with the problem of reappraisal of "where do we stand?" and "where do we go from here?"

Finally, because our society and the manifestations of our culture are changing rapidly, we must develop new and sophisticated procedures to constantly evaluate our position, and to make long range land use plans.

PROCEDURES AND TECHNIQUES

In recent years, NASA has sought to broaden its Earth Resources Program to include the development of potential users of satellite data. In order to accomplish this task, Marshall Space Flight Center (MSFC) has initiated several "demonstration projects" which will serve to illustrate how NASA's technology may be applied toward the solution of many environmental and resource-use problems.

Several of these "demonstration projects" involve the mapping of land use and related data from multi-spectral aerial photography. The acquisition and evaluation of the multi-spectral photography is accomplished with a Mark I camera and MINI-ADDCOL viewer produced by International Imaging Systems. The Mark I Camera is a modified version of a K-22 Airforce Reconnaissance Camera (Figure 1). Four spectrally filtered 3.5 by 3.5 inch images are recorded simultaneously in a 9 by 9 inch film format. The lens cone provides mounts for four Schneider Xenotar lenses with 150 mm (6 inch) focal lengths. In normal configuration blue, green, red and infrared images are recorded on black and white infrared film. The infrared film utilized is Kodak emulsion 2424 which possesses the unique characteristic of being sensitive in the visible and near infrared portions of the electromagnetic spectrum (Figure 2). Wratten filters [#47B (Blue); #57A (Green); and #25 (red)] are used in conjunction with infrared rejection filters to produce pass bands for the visible region records and a Wratten infrared filter [#88A (infrared)] is used to produce the IR record (Figure 3).

After exposure and development, contact positive transparencies can be made for viewing and interpreting in the MINI-ADDCOL viewer. The MINI-ADDCOL viewer is a four-channel, additive color projector designed to superimpose up to four individual spectral images in registration on a rear projection screen. Each 3.5 by 3.5 in. image is enlarged 2.56 X by the projector system such that the interpreter views a 9 by 9 inch image on the screen (Figure 4).

The Mark I Camera and MINI-ADDCOL viewer offer specific advantages which are worth noting. The system utilizes black and white film which is considerably cheaper and simpler to process than color and color IR photography. Composite images in natural and false color can be reconstructed on the viewing screen by projecting the appropriate individual images through a blue, green, or red filter. Details, particularly tonal variations which are difficult to distinquish on normal black and white photography and color photography, can be enhanced by varying filter combinations and illumination intensities within the projection system.

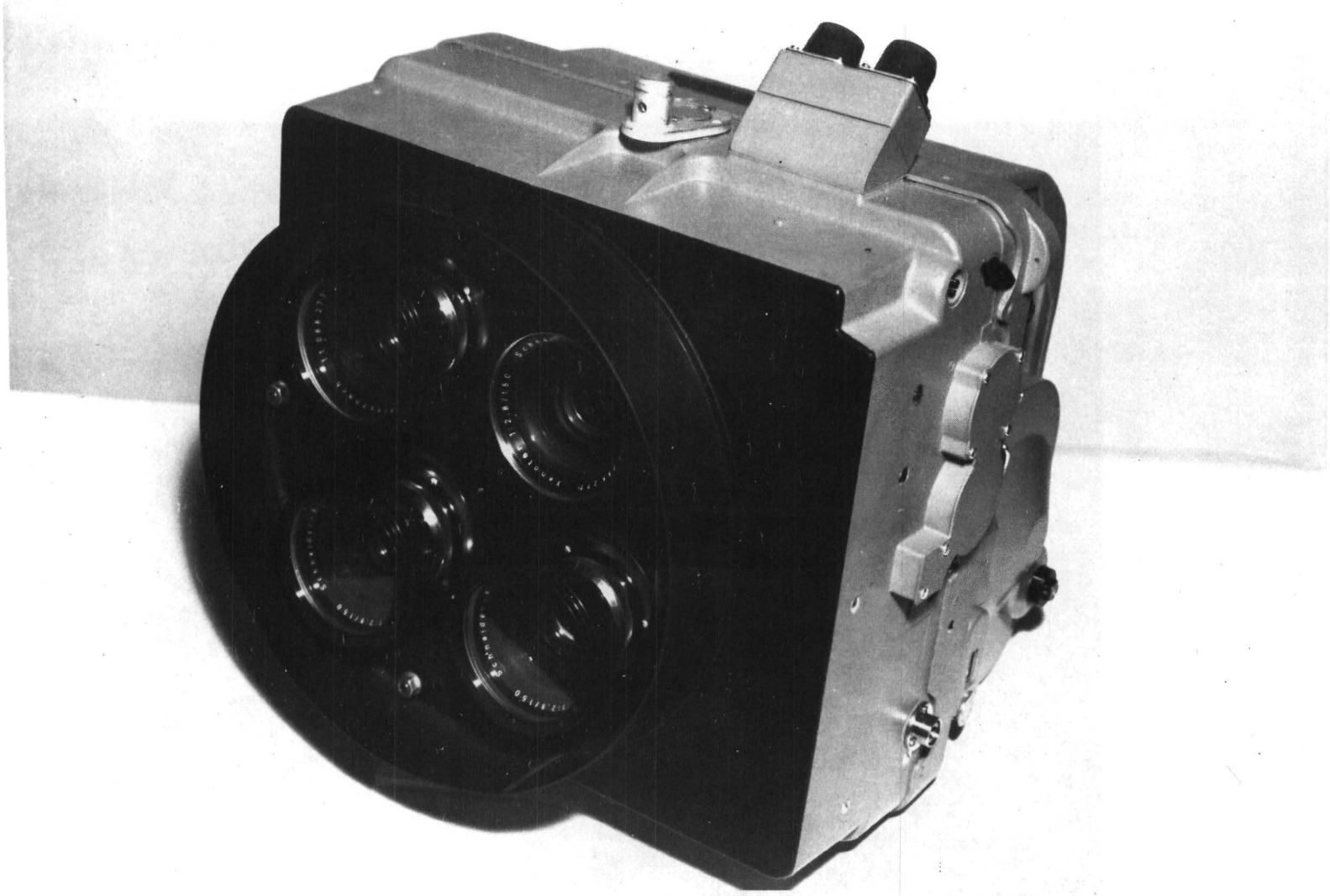


Figure 1. The Mark I Camera

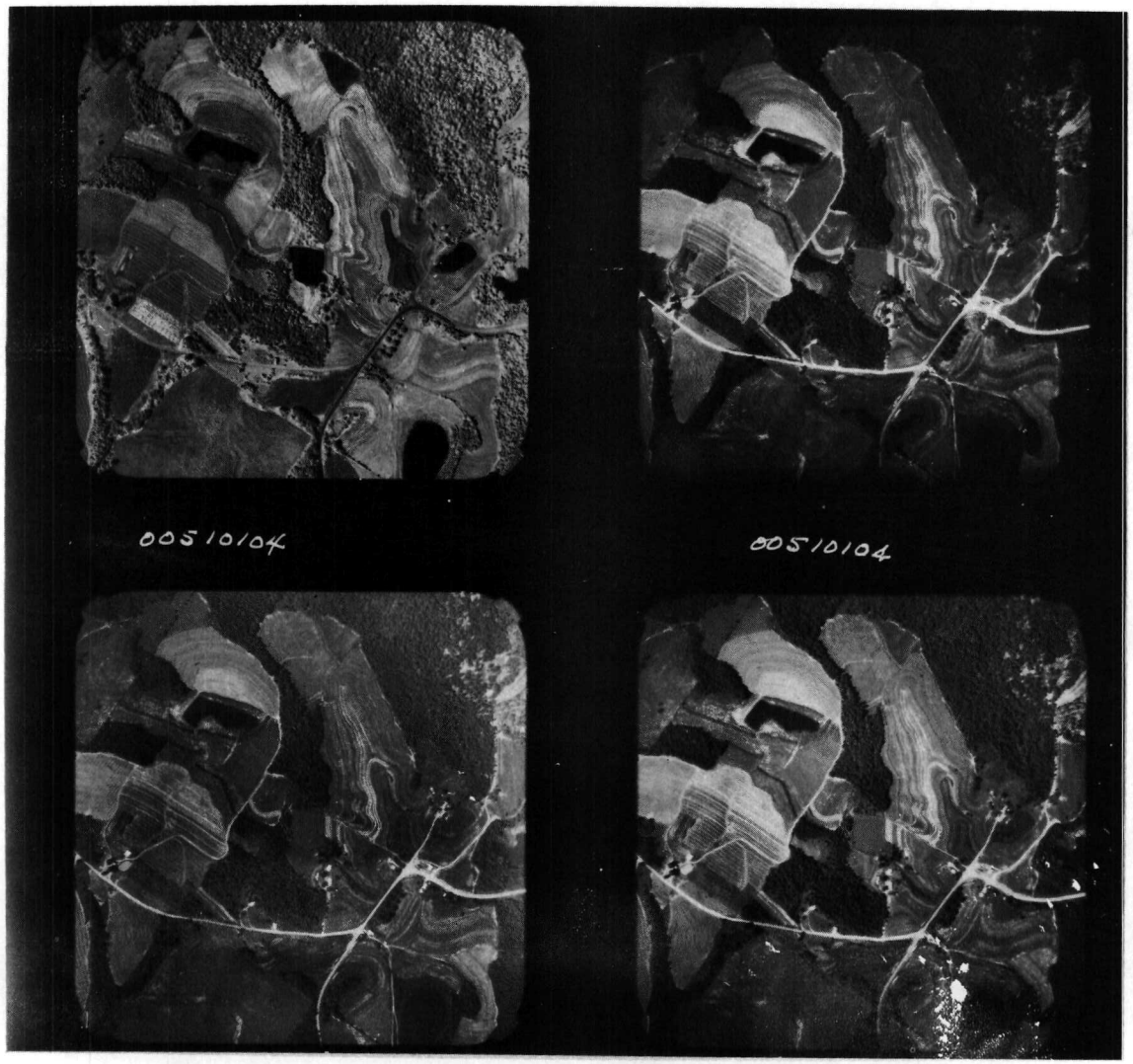


Figure 2. Sample of Kodak 2424 Film

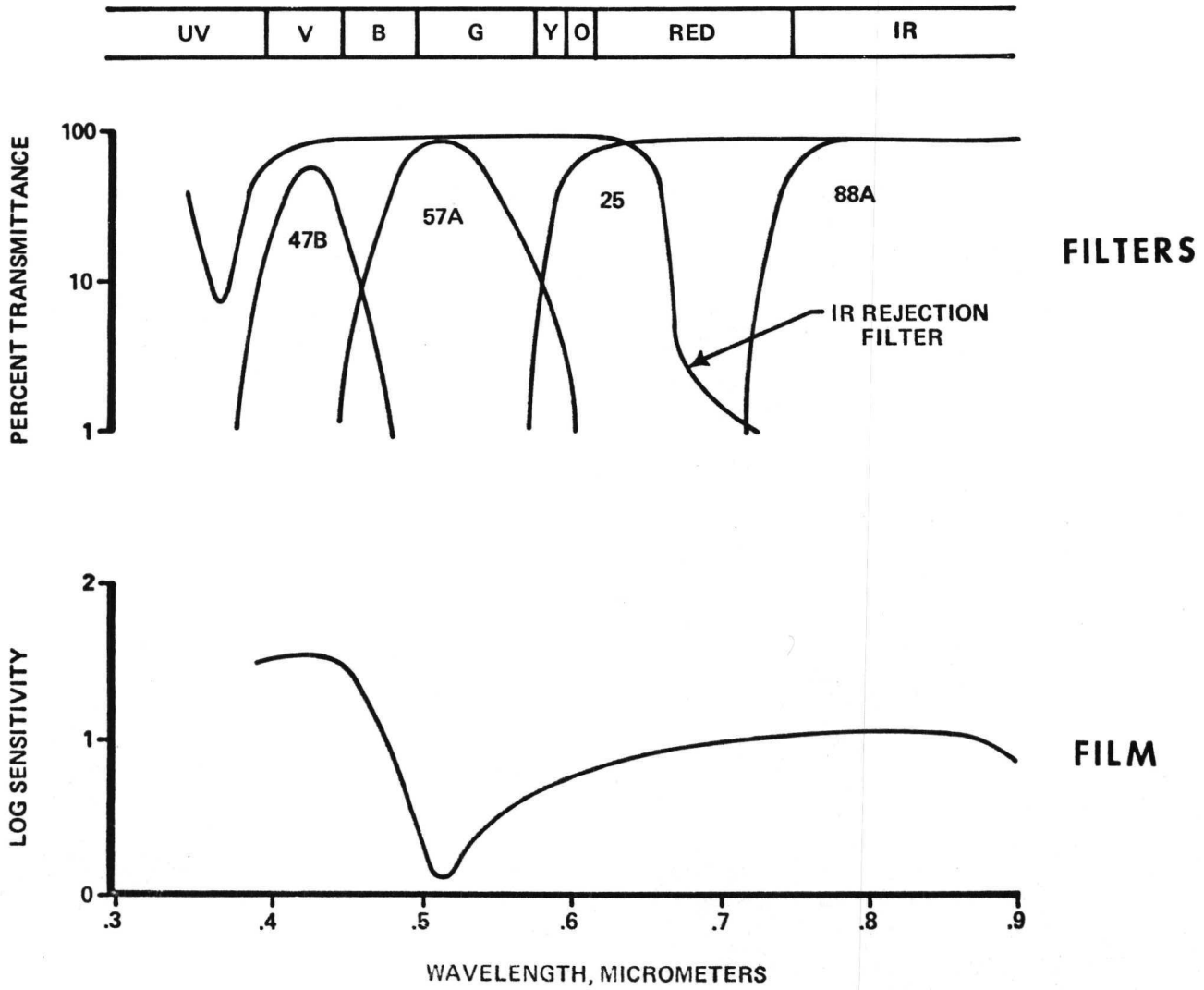


Figure 3. Spectral Curve of Wratten Filters and Film

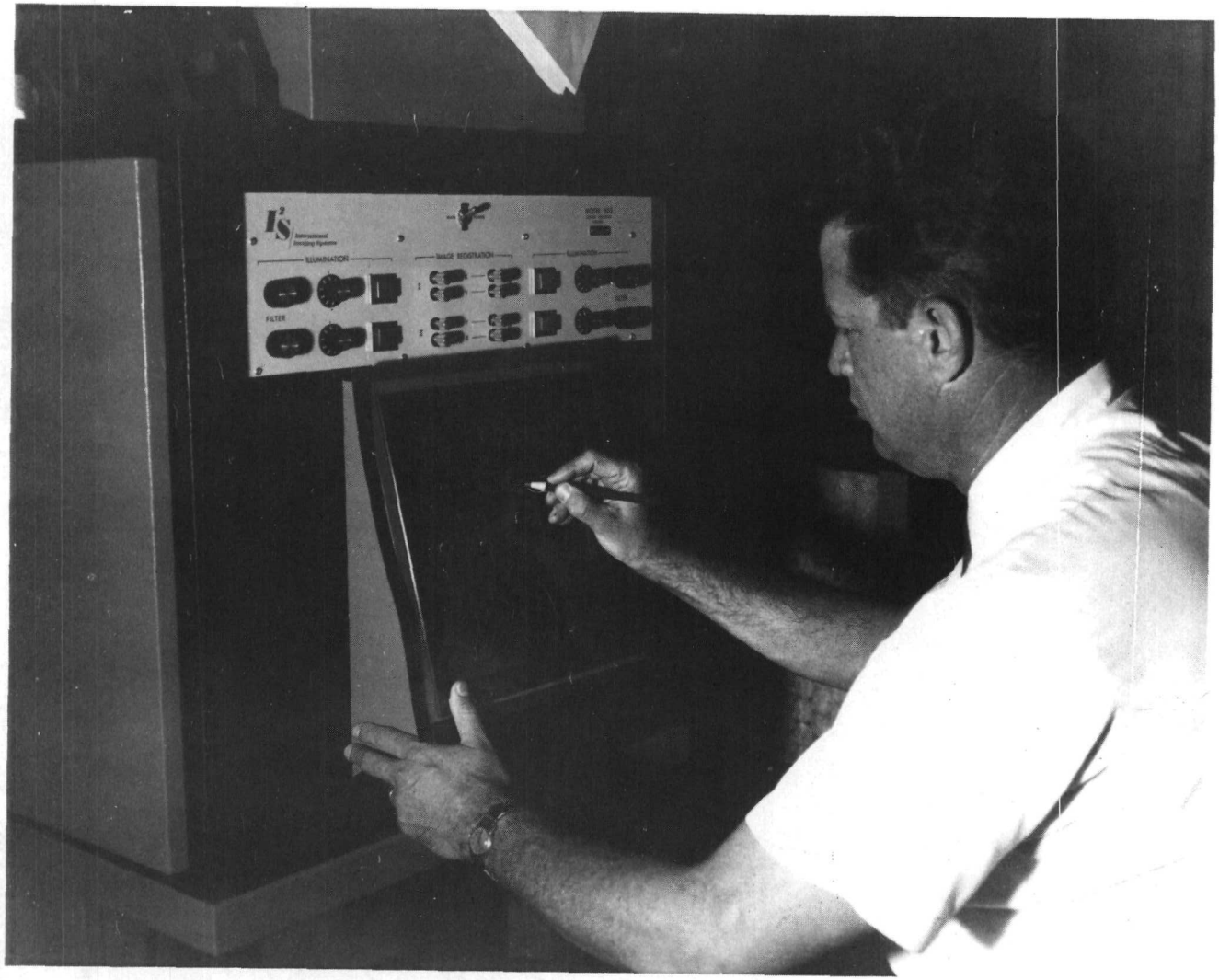


Figure 4. Tracing on Transparent Sheet
from Multispectral Projector

During the summer of 1971, this investigator was fortunate to participate in the development of a series of techniques to enable one to extract necessary land use data from the photography via the MINI-ADDCOL viewer and transfer these data to planimetric maps. The procedures can hardly be described as revolutionary when compared to some of the sophisticated automated interpretation techniques being attempted today. However, when one investigates the development status of these automated interpretation procedures (several of which are on-going at MSFC), it is soon realized that in terms of present data requirements, we are still far from a workable automatic interpretation system. This is not to say that automated techniques of interpreting are not encouraging. This investigator is convinced, however, that any immediate improvement in the data acquisition, data processing, and information retrieval system must occur at the man-machine interface not the machine-machine interface.

The procedure for producing the land use maps in outlined as follows:

1. The MINI-ADDCOL viewer normally employs frosted glass for the viewscreen upon which the images are projected from the rear. In this procedure clear plexiglass in substituted for the frosted glass and then overlain with matte acetate. The image is then projected onto the matte acetate which may be drawn on with any suitable drawing instrument. In this case color pencils were found to be most desirable in that various multiple characteristics about the land surface could be denoted or delimited in separate colors.*

2. Areas to be mapped are projected onto the matte acetate with the proper image and filter combinations for interpretation. The infrared and red images projected through the green and red filters respectively are generally the most commonly used for interpreting land use. Use of this combination enhances healthy vegetation which is best recorded in the near infrared portion of the electromagnetic spectrum and contrasts well with the red clay soils of this area, best displayed by the red portion of the spectrum.

*It should be noted that this step represents an improvement over the system developed last year in that previously clear acetate and "grease pencil" were employed which were considerably more difficult to handle. This improvement must be credited to a suggestion made by Bill Scarbrough, the aerial photographer for the Environmental Applications Office.

3. Land use boundaries are drawn in with colored pencils and the proper land use identification code inserted. Other characteristics of the land surface are noted at the same time.

4. The acetate overlay is then removed from the screen and inserted into a Kail Projector (Figure 5). The Kail Projector is also a rear projection device which employs a matte surface to display the image. Normally, a material such as tracing paper or matte acetate is used for drafting from the projector but in the search for a optimal means of recording the land use data in map manuscript form, another cartographic inovation was employed. In related discussions with Tennessee Valley Authority officials, it was discovered that the TVA Maps and Survey Branch could provide composite reproductions of the cultural and hydrologic separation sheets of a TVA 1:24,000 series topographic map. This could be reproduced on a transparent matte material called Cronaflex and would provide an excellent base map upon which one could project the land use interpreted from the viewer.

5. Using the Cronaflex copies of the TVA maps as a base, a draftsman inks the land use boundaries and corresponding codes direct onto matte sheets (Figure 6). Other data may be simultaneously recorded on the same sheet or recorded on a separate overlay.

Completion of these procedures will produce a black and white land use map at a scale of 1:24,000, incorporating the cultural and hydrologic data found on TVA quadrangle maps. The production of this map represents a major accomplishment which is all too often met with the comment "so what?" Enbodied in one topographic map or land use map is more readily available information for human use than any other storage system heretofore developed, computer oriented or otherwise. The topographic contour information on one U.S.G.S. or TVA map alone will consume most of the storage capacity of present day computers if detailed data is required. Yet, there are several present-day information requirements which dictate that we improve our ability to store and manipulate data of this type.

Although certain generalized information may be derived through the use of several map overlays of various spatial data, the acquisition, storage, display, and cartographic processing of these data is formidable and time consuming. Any quantitative assessment of the relationships between spatial variables such as soil, vegetation, cropland, pastureland, cultural elements, and sociological elements

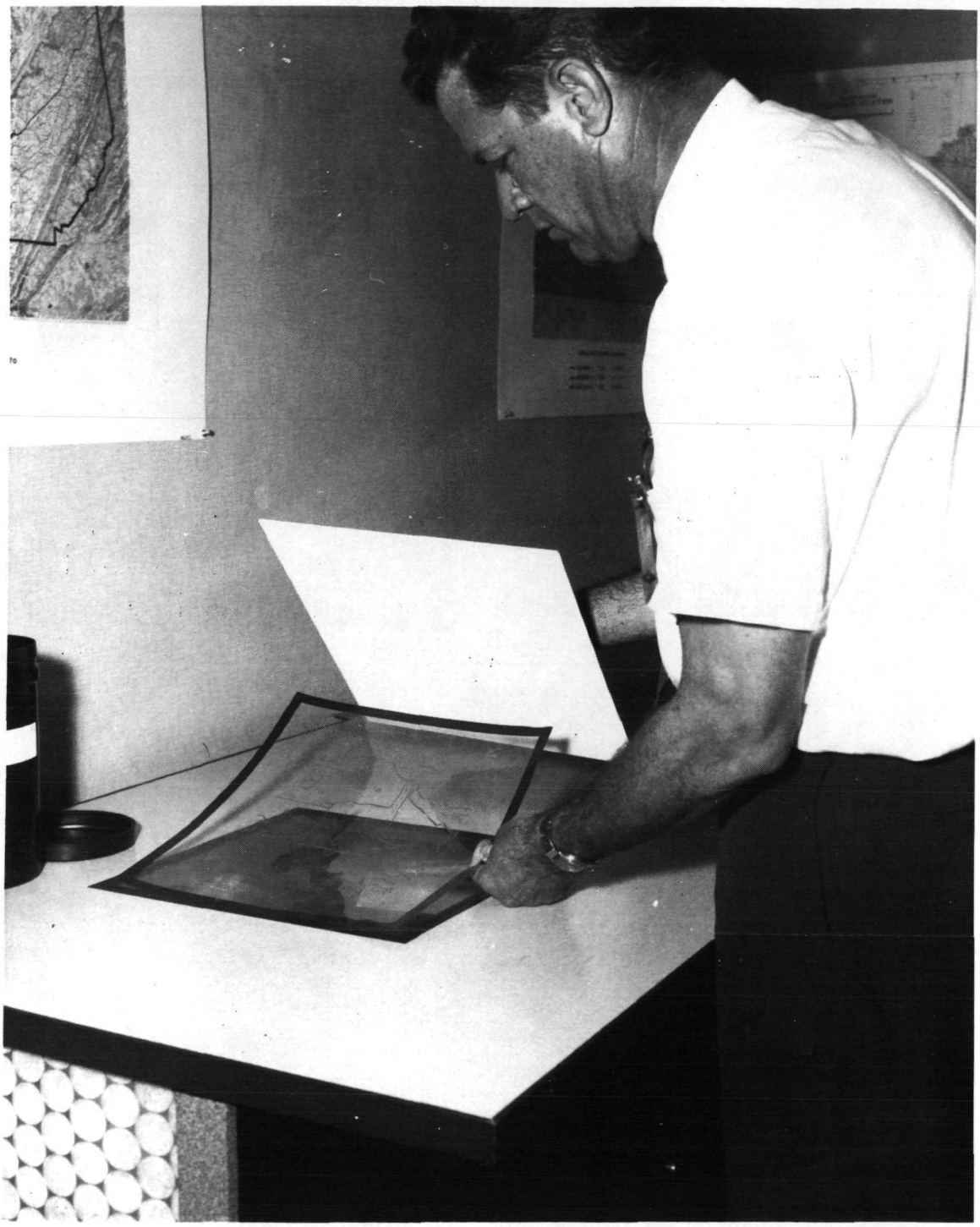


Figure 5. Placing Tracing of Photo Interpretation on the Kail Projector

Figure 6. Tracing Map on Kail Projector

further increase the time element and is practical only over a limited area.

The development of new possibilities and techniques for data handling and development of "information systems" which facilitate the flow and manipulation of data and their presentation for use appears to hold great promise in solving this problem (25; p. 105). It is in this light that some discussion has been included as to how the above system might be improved.

SUGGESTED IMPROVEMENTS

As mentioned above, the best system one could possibly expect at present for the acquisition, reduction, storage and display of data from multi-spectral photography or imagery is a semi-automated procedure. Perusal of the literature and selected discussions with knowledgeable persons indicates that the necessary "software" (computer programs and techniques) and "hardware" (machines) are available and the only problem is to integrate the separate units into a functioning system. This is no small task and is likely to be costly. It would appear that much of the problem that is presently holding up the development of such a system, is the belief that fully-automated procedures are "just around the corner." It is highly likely that automatic interpretation procedures will be effective with small scale, low-resolution photographs covering large areas. But with large and medium scale mapping, I am convinced we are going to have to rely on human interpretation..

Therefore, I suggest a concentrated effort to develop a semi-automated system which would facilitate the extraction of data from the imagery and process it into a usuable form. It is not inconceiveable that a device might be developed which will permit one to outline and label various land characteristics on a view screen while a "slave mechanism" coupled to a computer encodes the data in digital form and produces a map manuscript. Several auto-servo mechanisms of this type are presently available and one is on order by the Computation Laboratory at MSFC. Most of these systems commonly use a tracing-type digitizer and a set of coordinate points which are used as index points for the data sets (23; p. 107). These data could then be recalled in several forms for human evaluation (i.e., via line printer, cathode ray tube screen or coordinatograph).

Inherent in this concept are some problems which must be recognized. At the present, the computer has no way of knowing where any particular data point is located relative to the real world except in terms of its relationship to surrounding data points. A system will have to be devised which will permit the computer to relate pre-selected points to know geographic coordinates on the earth surface. This is no small task as any cartographer can attest. However, progress in this area is being made (i.e., the U. S. Department of Commerce's DIME - Geocoding System and the Canada Geographic Information System (23; p. 107).

RECOMMENDATIONS AND CONCLUSIONS

The recommendations made in this section regard suggestions of future activities and emphasis areas for the Environmental Applications Office at MSFC. It is anticipated that some of these activities may form the basis for research activities for this investigator the following summer.

In past years, the primary public role of NASA has been space oriented. These activities have over-shadowed the development programs in the field of remote sensing sponsored by NASA. Many earth scientists and social scientists have been quietly waiting the arrival of an earth oriented space program. With the launch of the ERTS-1 satellite, the possibility of studying the earth from space is no longer a scientist's dream. There is a tremendous potential for the use of ERTS imagery for regional environmental studies.

One such region is the area administered by the Tennessee Valley Authority. Included in the TVA region are sub-regions such as that administered by TARCOG (Top of Alabama Regional Council of Governments). Both of these agencies are users of large area-small scale data, which is studied in a generalized form. The ERTS imagery fits the above criteria perfectly and provides the added advantage of repetitive coverage every 18 days.

It is recommended, therefore, that a pilot demonstration project be initiated to incorporate the use of ERTS data into the broad environmental assessment needs of TVA and TARCOG. Ideally, the pilot study should encompass the TARCOG region for several reasons:

1. ERTS photography of the area should be available, which will provide the broad overview data suitable for use at a 1:250,000 scale level of generalization.

2. High altitude RB-57 photography taken at a scale of 1:120,000 is also available for the North Alabama area, which is suitable for data needs at scales around 1:50,000 to 1:100,000.

3. MSFC has acquired multi-spectral photography of the entire TARCOG region at a scale of 1:24,000, which is suitable for most detailed land mapping.

4. In discussion with all groups concerned, a strong desire and willingness to cooperate was visably evident.

Such a demonstration project appears to have tremendous potential. Establishing a cooperative arrangement with a user such as TVA or TARCOG would not only provide the expertise and assistance these agencies lack internally but also would provide an excellent demonstration of NASA's existing technology.

The second recommendation regards a pragmatic project which involves a micro approach to environmental studies rather than the macro project mentioned above.

With the enactment of an Environmental Policy Act of 1969, any public works or private project significant enough to alter the environment must be accompanied by an environmental impact statement. Most of the present impact statements prepared to date have been extremely generalized and provide little or no proof regarding environmental protection.

The basic problem revolves around the lack of adequate, up-to-date detailed data. Most of the data sources collected at the State and Federal level are at county scale or at best, census tract scale. There is simply no means available, at present, to collect and process data for small drainage sheds or short road corridors except by costly and tedious field work.

Current expertise exists at MSFC to assist in these perplexing problems. It is recommended that MSFC develop one or two selected research projects specifically designed to study and demonstrate data acquisition and processing procedures necessary for the preparation of environmental impact statements. These projects could be conducted locally and with "in-house" capabilities available in the Computation Laboratory, the Environmental Applications Office and other MSFC laboratories.

In the opinion of this research fellow, the above two projects would be relevant to the needs of today's society. The potential for inter-governmental agency co-operation should not be considered lightly. Public opinion often dictates the success or failure of any agency's programs. These programs have the potential to display publicly the relevance of NASA's programs.

SELECTED BIBLIOGRAPHY

1. Ackoff, Russell L. Scientific Method: Optimizing Applied Research Decisions. New York: John Wiley and Sons, Inc. 1962.

2. Amiran, D. H. K. "Introduction" pp. 11-12 in Land Use in Semi-Arid Mediterranean Climates. UNESCO/International Geographical Union Symposium, Iralion, Greece, September 19-26, 1962. Belgium: United Nations Educational, Scientific and Cultural Organization, Place de Fontenoy, Paris. 1964.

3. Anderson, James R. "The Dilemma of Idle Land in Mapping Land Use," The Professional Geographer, Vol. XIV,No. 3 (May, 1962), pp. 15-18.

4. Anderson, James R. "Land Use Classification Schemes," Photogrammetric Engineering (June, 1971), pp. 379-387.

5. Anderson, James R. "Toward More Effective Methods of Obtaining Land Use Data in Geographic Research," The Professional Geographer, Vol. XIII, No. 6 (November, 1961), pp. 15-18.

6. Anderson, James R., Hardy, Ernest E., and Roach, John T. A Land Use Classification for Use with Remote Sensing in the United States: A Proposed Definitional Structure. Preliminary Report to U. S. Geological Survey, May, 1972.

7. Berry, Brian J. L. "A Synthesis of Formal and Functional Regions Using a General Field Theory of Spatial Behavior," in Berry and Marble (Eds.). Spatial Analysis: A Reader in Statistical Geography. Englewood Cliffs, New Jersey: Prentice-Hall, Inc. 1968, pp. 419-428.

8. Billings, W. D. "The Environmental Complex in Relation to Plant Growth and Distribution," Quarterly Review of Biology, Vol. 27, No. 3 (1952), pp. 251-265.

9. Birch, J. W. "Rural Land Use: A Central Theme in Geography" in Land Use and Resources: Studies in Applied Geography. A Memorial Volume to Sir Dudley Stamp; London: Institute of British Geographers, 1968, pp. 13-28.

10. Bond, A. D. and Atkinson, R. T. "An Integrated Feature Selection and Supervised Learning Scheme for Fast Computer Classification of Multi-spectral Data," in Remote Sensing of Earth Resources, Vol. I. Selected technical papers from the Conference on Earth Resources Observation and Information Analysis Systems, Tullahoma, Tennessee, The University of Tennessee Space Institute, 1972, pp. 645-671.

11. Brooner, William C. and Nichols, David A. "Considerations and Techniques for Incorporating Remotely Sensed Imagery into the Land Resource Management Process," in Remote Sensing of Earth Resources, Vol. I. Selected technical papers from the Conference on Earth Resources Observation and Information Analysis Systems, Tullahoma, Tennessee, The University of Tennessee Space Institute, 1972, pp. 1-25.

12. Buchanan, R. Ogilvie. "The Man and His Work," Land Use and Resources: Studies in Applied Geography. A Memorial Volume To Sir Dudley Stamp; London: Institute of British Geographers, 1968.

13. Burley, Terence M. "Land Use or Land Utilization?" Professional Geographer. Vol. 12 (November, 1961).

14. Clawson, Marion and Stewart, Charles L. Land Use Information. Baltimore: The John Hopkins Press, 1969.

15. Coleman, Alice. Land Use Survey Handbook. Strand, London: King's College, Director of the Second Land Use Survey, 1968.

16. Colvocoresses, Alden P. "A Unified Plane Coordinate Reference System," Surveying and Mapping, Vol. 27, No. 4 (December, 1967), pp. 621-624.

17. Colwell, Robert N. "Spectrometric Considerations Involved in Making Rural Land Use Studies with Aerial Photography," Photogrammetria, Vol. 20 (1965), pp. 15-33.

18. The Congressional Record. January 1, 1970, Public Law 91-190, 91st Congress, s.1075, National Environmental Policy Act of 1969.

19. Conklin, H.E. "The Classification of Agricultural Land." Mimeographed discussion, Cornell University, September, 1956, p.5.

20. Conklin, H.E. "The Cornell System of Economic Land Classification," Farm Economics No. 198. Ithaca, New York: Cornell University, January, 1955.

21. Conklin, H.E. and Broder, F. Zuaus. An Economic Classification of farms areas in Lureis County, New York. Land Classification Leaflet No.4; Ithaca, New York: Cornell University, May, 1954, p.3.

22. Council on Environmental Quality. Environmental Quality. The First Annual Report of the Council on Environmental Quality together with the President's Message to Congress; Washington, D.C.: U.S. Government Printing Office, 1970.

23. Cox, Ian H. (editor). New Possibilities and Techniques for Land Use and Related Surveys. Papers presented at an international symposium, London, 21-23 April, 1970; Occational Papers No. 9, World Land Use Survey Commission; Tonbridge, Kent: Tonbridge Printers Limited, 1970.

24. Crawford, Roger J. Utility of an Automated Geocoding System for Urban Land Use Analysis. Research Report No.3; Seattle, Washington: University of Washington, Urban Data Center, 1967.

25. Department of Regional Economic Expansion. The Canada Land Inventory. Report No.1/1965 second edition, 1970, Objective Scope and Organization; Ottawa: Queen's Printer for Canada, 1970.

26. Ely, R.T. "A National Policy for Land Utilization." National Real Estate Journal, Vol. 23 (1922), pp. 11-14.

27. The Federal Register, Vol. 36, No.79 (Friday, April 23, 1971), pp. 7724-7729.

28. Finch, V.C. "Geographic Surveying," Bulletin of the Geographic Society of Chicago, IX (1933), pp. 3-11.

29. Finch, V.C. "Geographic Surveying, and Morrfort, A Study in Landscape Types in Southwestern Wisconsin," Bulletin No.9; The Geographic Society of Chicago. Chicago: University of Chi· Press, 1933, pp. 2-40.

30. Flawn, Peter T. Environmental Geology: Conservation, Land Use Planning and Resource Management. New York: Harper and Row, Publishers, 1970.

31. Frazee, C.J., Myers, V.I. and West. "Remote Sensing for Detection of Soil Limitations in Agricultural Areas," Proceedings of the Seventh International Symposium on Remote Sensing of Environment, Vol. I, Ann Arbor, Michigan: University of Michigan, May, 1971, pp. 327-343.

32. Garrison, W.L. and Marble, D.F. "The Spatial Structure of Agricultural Activities," Annals of the Association of American Geographers, Vol. 47 (1957), pp. 137-144.

33. Goodsell, O.E. and Rudolph, W.M. Land Classification: A Selected Bibliography. Agricultural Economics Bibliography No.83; Washington, D.C.: U.S. Department of Agriculture, March, 1940.

34. Gould, P. "A Bibliography of Space Searching Procedures for Geographers," Research Note. Department of Geography, Pensylvania State University, September, 1965. (University Park, Pa.)

35. Gwyn, A.P. Aerial Photo Interpretation and Land Use. HUD Project MD. P-61; Baltimore, Maryland: Regional Planning Council

36. Harris, R.C. "Historical Geography in Canada," Canadian Geographer, Vol. II (1967), pp. 235-250.

37. Harvey, David W. "Theoretical Concepts and the Analysis of Agricultural Land-Use Patterns in Geography," Annals of the Association of American Geographers, Vol. 56, No. 2 (June, 1966), pp. 361-374.

38. Henshall, Janet D. "Models of Agricultural Activity" in R.J. Chorley and P. Haggett (eds.), Models in Geography. London: Methuen Ltd., 1967, pp. 425-460.

39. Higbee, Edward. American Agriculture: Geography, Resources, Conservation. New York: John Wiley and Sons, Inc. 1958.

40. Hudson, G. Donald. The Rural Land Classification Program: A Summary of Techniques and Uses. Land Classification Section, Division of Land Planning and Housing, Tennessee Valley Authority, December, 1935.

41. Hudson, G. Donald. "The Unit-Area Method of Land Classification," Annals of the Association of American Geographers, XXVI (1936), pp. 99-112.

42. Institute of British Geographers. Land Use and Resources: Studies in Applied Geography. A Memorial Volume To Sir Dudley Stamp; London: Institute of British Geographers, 1968.

43. Inter-County Regional Planning Commission. Land Use Classification Manual. Denver, Colorado: Land Use and Mapping Committee, Inter-County Regional Planning Commission, 1966.

44. International Geographical Union. Report of the Commision on World Land Use Survey for the Period 1949-1952. Worcester: Published with the cooperation of the United Nations Educational, Scientific, and Cultural Organization (UNESCO), 1952.

45. International Geographical Union. Report of the Commission on Inventory of World Land Use. New York: Office of the Secretary-Treasurer, International Geographical Union, Twentieth Century Fund, August 9-18, 1956.

46. James, Preston E. "Mapping at Chorographic Scales," Proceedings, Eight General Assembly and Seventeenth International Congress. Washington, D.C.: The United States National Committee of the International Geographical Union, National Academy of Sciences - National Research Council, August 8-15, 1952, pp. 175-178.

47. Johnson, K.E., Adams, C.H., Doyle, F.L. and Childers, L.P. Draft Environmental Impact Statement for Riverfront Industrial Park Madison County, Alabama. Huntsville, Alabama: University of Alabama in Hunstville, Division of Graduate Programs and Research, Center for Environmental Studies, February, 1972.

48. Jones, Clarence F. (ed.). The Rural Land Classification Program of Puerto Rico. Evanston, Illinois: Northwestern University, 1952.

49. Jones, W.D. and Finch, V.C. "Detailed Field Mapping in the study of the Economic Geography of an Agricultural Area," Annals of the Association of American Geographers, Vol. 15 (1925), pp. 148-157.

50. Kariel, H.C. and Kariel, P.E. Explorations in Social Geography. Reading, Massachusetts: Addison-Wesley Publishing Company, 1972.

51. Karmin, Monroe W. "Nixon Asks 'National Land Use Policy' Giving States Wide Control Under HUD," The Wall Street Journal, Wednesday, January 20, 1971, p. 6.

52. King, Leslie J. Statistical Analysis in Geography. Englewood Cliff, N.J.: Prentice-Hall, Inc., 1969.

53. Kreig, R.A. "Aerial Photographic Interpretation for Land Use Classification in the New York State Land Use and Natural Resources Inventory," Photogrammetria, Vol. 26 (1970), pp. 101-111.

54. Land Classification Advisory Committee of the Detroit Metropolition Area. Land Use Classification Manual. Chicago, Illinois: Public Administration Service, 1962.

55. Land Classification, United States Bureau of Reclamation Manual, Vol. V, pt. 2, Washington D.C.: April 9, 1953. Chapter 2.13 and 2.5.

56. Leopold, Luna B., Clarke, Frank E.,Hanshaw, Borvee B., and Bolsley, James R. A Procedure for Evaluating Environmental Impact. Geological Survev Circular 645; Washington, D.C.: U.S. Government Printing Office, 1971.

57. Lockheed Electronics Company, Inc., Houston Aerospace Systems Division. 1970 Land-Use Inventory. Houston, Texas: Lockheed Electronics Company, December, 1971.

58. Marble, Duane F. and Horton, Frank E. "Extraction of Urban Data From High and Low Resolution Images," Proceedings of the Sixth International Symposium on Remote Sensing of the Environment. Vols. I, II and III, Ann Arbor: The University of Michigan, 1969, pp. 807-817.

59. Marsden - Jones, E.M. and Turrill, W.B. "Sixth Report of the Transplant Experiments of the British Exological Society at Potterne, Wittshire," Journal of Ecology, Vol. 33 (1945), pp. 57-81.

60. Marshner, Francis J. Land Use and Its Patterns in the United States. Agricultural Handbook No. 153; Washington, D.C.: Agricultural Research Section, U.S. Department of Agriculture, 1959; and Major Land Uses in the United States. Washington, D.C.: U.S. Department of Agriculture, 1950.

61. McHarg, Ian L. Design with Nature. Garden City, New York: The Natural History Press, 1969.

62. Mair, R.L. A Communications Theory of Urban Growth. Cambridge, Mass.: M.I.T. Press, 1962.

63. Mushbach, W.F. and Johnson, V.W. "National Resources Planning Board on Land Classification," Journal of Land and Public Utility Economics, Vol. 17, No.4 (November, 1941).

64. National Resources Planning Board. Area Analysis - A Method of Public Works Planning. Technical Paper No. 6; Washington: U.S. Government Printing Office, 1943.

65. National Resources Planning Board. Land Classification in the United States. Washington, D.C.: U.S. Government Printing Office, 1941.

66. Nelson, H.J. "A Service Classification of American Cities," Economic Geography, Vol. 31 (1955), pp. 189-210.

67. Nephew, E.A. "Healing Wounds," Environment, Vol. 14, No. 1 (January-February, 1972), pp. 12-21.

68. Nunnally, Nelson R. and Witner, Richard E. "Remote Sensing for Land Use Studies," Photogrammetric Engineering. (May, 1970), pp. 449-453.

69. Nunns, Frederick K. Systems, Procedures and Research Needs of Rural Land Classification in the United States. Committee Working Paper; Farm Economies Research Division, U.S. Department of Agriculture, 1957.

70. Ogden, C.K. and Richards, I.A. The Meaning of Meaning. New York: Harcourt, Brace and Co., rev. ed., 1947.

71. Pico, Rafael. "Rural Land Classification - Its Use in Planning in Puerto Rico," Proceedings, Eight General Assembly and Seventeenth International Congress. Washington, D.C.: The United States National Committee of the International Geographical Union, National Academy of Sciences - National Research Council, August 8-15, 1952, pp. 175-178.

72. Polunin, N. Introduction to Plant Geography. New York: McGraw Hill, 1960.

73. Prunty, Merle C. Jr. (ed.). Festschrift: Clarence F. Jones. Evansten, Illinois: Northwestern University, May, 1962.

74. Salmon, Wesley C. Logic. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963.

75. Schiller, F.C.S. Formal Logic. London: McMillan and Co., 1912, p.71.

76. Schmeckebler, L.F. Catalogue and Index of the Publications of King, Hayden, Wheeler, and Powell Surveys, U.S. Geological Survey Series G., Miscellaneous 26 Bull. 222, Washington, 1904.

77. Schmertz, Micheal F. "Open Space for People," Architectural Record, (July, 1972), pp. 131-140.

78. Shantz, A.L. "Plants as Soil Indicators," Soil and Men. USDA Yearbook of Agriculture; Washington D.C.: U.S. Government Printing Office, 1938, pp. 835-860.

79. Shelly, M.W. and Bryan, G.L. (eds.). Human Judgements and Optimality. New York: John Wiley and Sons, Inc., 1964.

80. Simon, N.A. Models of Man. New York: John Wiley and Sons, Inc., 1957.

81. Simonson, Roy W. "What Soils are," Soils. USDA Yearbook of Agriculture. Washington, D.C.: U.S. Government Printing Office, 1957, pp. 17-31.

82. Simpson, Robert B. Production of a High Alititude Land Use Map and Data Base for Boston. Interagency Report USGS-235; Washington, D.C.: Prepared by the U.S. Geological Survey for the National Aeronautics and Space Administration, December, 1970.

83. Soileau, John M., Smith, Wesley G. and Matson, Roger A. Tennessee Valley Land and Its Changing Use. Bulletin Y-27; Muscle Shoals, Alabama: National Fertilizer Development Center, Tennessee Valley Authority, July, 1971.

84. Stevenson, Walter B., Jr. Planner's Mapping and Classification Guide. Montgomery, Alabama: Alabama Development Office, 1972.

85. Su, M.Y., Jayroe, R.R., Jr. and Cummings, R.E. "Unsupervised Classification of Earth Resources Data," Remote Sensing of Earth Resources, Vol. I. Selected technical papers from the Conference on Earth Resources Observation and Information Analysis Systems; Tullahoma, Tennessee: The University of Tennessee Space Institute, 1972, pp. 673-693.

86. Toebes, Gerritt H. (ed.). Natural Resource Systems Models in Decision Making. Proceedings of a 1969 Water Resources Seminar; Lafayette, Indiana: Purdue University, Water Resources Research Center, 1969.

87. Tomlinson, R.F. (ed.). Environmental Information Systems. Proceedings of the UNESCO/IGU First Symposium in Geographical Information Systems, Ottawa, September, 1970; A publication of the International Geographical Union Commission on Geographical Data Sensing and Processing, 1971.

88. United Nations. Land Use in Semi-Arid Mediterranean Climates. UNESCO/International Geographical Union Symposium, Irdlion, Greece, September 19-26, 1962. Belgium: United Nations Educational, Scientific and Cultural Organization, Place de Fontenoy, Paris, 1964.

89. U.S. Bureau of the Budget. Standard Industrial Classification Manual. Washington, D.C.: U.S. Government Printing Office, 1957, p. 433.

90. USGS. Conference on Land Use Information and Classification: Conference Papers Department. Washington, D.C.: Department of the Interior, U.S. Geological Survey. June, 28-30, 1971.

91. Urban Renewal Administration Housing and Home Finance Agency and Bureau of Public Roads of the Department of Commerce. Standard Land Use Coding Manual. Washington, D.C.: U.S. Government Printing Office, January, 1965.

92. Vegas, Paul L. A Procedure for the Use of Small Scale Photography in Land Use Mapping. NASA Technical Report, Earth Resources Laboratory at Mississippi Test Facility, March 1, 1972.

93. Whittlesey, Derwent. "Major Agricultural Regions of the Earth," Annals of the Association of American Geographers, Vol. 26, (1936), pp. 199-240.

94. Wieslander, A.E. and Retzer, J.L. Soil - Vegetation Surveys: There Value and Use in Wild Land Management. California Forest and Range Experiment Station, November, 1955.

95. Wilsie, Carroll P. Crop Adaptation and Distribution. San Francisco: W.H. Freeman and Company, 1962.

96. Wray, James R. "Census Cities Project and Atlas of Urban and Regional Change," Paper presented at NASA, Third Annual Earth Resources Program Review, December 1-3, 1970. NASA Manned Spacecraft Center, Houston Texas.

Page Intentionally Left Blank

1972

NASA - ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

STUDY OF ICE PARTICLE FORMATION AND LIFETIME IN SPACE ENVIRONMENT

PREPARED BY:	DUEN-REN JENG
ACADEMIC RANK:	ASSOCIATE PROFESSOR
UNIVERSITY:	UNIVERSITY OF TOLEDO
LABORATORY: (Division) (Branch)	ASTRONAUTICS PROPULSION & THERMODYNAMICS THERMAL ENGINEERING
RESEARCH COUNTERPART:	CHARLES COTHRAN
DATE:	AUGUST 11, 1972
CONTRACT NUMBER:	NGT-01-003-045

STUDY OF ICE PARTICLE FORMATION AND LIFETIME IN SPACE ENVIRONMENT

By

Duen-Ren Jeng

ABSTRACT

The ice particles are formed when liquid and/or humid gases vent to the space. These submicroscopic ice particles are potential contamination sources of the environments during Skylab operations.

The analysis is made for predicting the critical size of ice particle formed and its nucleation rate based on the theory of homogeneous and heterogeneous nucleation by sublimation.

The equations which are pertinent for studing the groth and evaporation of the ice particles are formulated. The mechanisms affecting the lifetime of ice particle are discussed.

The gasdynamic techniques for experimental study of ice particle formation are proposed.

ACKNOWLEDGEMENTS

The author wishes to thank NASA/ASEE Summer Faculty Fellowship Program for providing him the opportunity to initiate this study. He also wishes to thank Charles Cothran with whom the author has held helpful discussion of certain aspects of the problems examined here. A special word of appreciation is extended to Jerry Vaniman and Farouk Huneidi for their interest and encouragement.

INTRODUCTION

In the debriefings of all astronauts of the Gemini and Mercury missions, there has been a continual insistence of the observation of "Space fireflies" and the inability to see stars fainter than second magnitude [1]*. This phenomena has strongly suggests that astronomical measurements from manned spacecraft must be made in an optical environment produced not by the ambient medium but by the spacecraft itself. The scattering of sunlight by the cloud of debris which accompanies the vehicle in its orbit is of greatest concern.

It is clearly known to us that waste products expelled from an orbiting spacecraft form a bebris cloud about the vehicle. If the waste products are expelled in a single dump, this will be swept away by a combination of aerodynamic drag and solar radiation pressure. Clearance times range from seconds to hours, depending on the orbital altitude, velocity, mass, and composition of material in the dump. If the waste products are continuously expelled, then a permanent comet-shaped debris cloud will follow the vehicle. The major sources of contamination of the space environment from space vehicle are known to be cabin gas leakage, venting of waste materials, exhaust products from thruster firings, and outgassing of nonmetallic materials.

The Skylab design includes several overboard vent systems, namely Waste Tank Nonpropulsive Vents, Contingency Condensate System, Molecular Sieve, and various experiment vents. When liquid and/or humid gases are vented through these Nonpropulsive vent into the space environment, the condensation process will occurto form a submicroscopic ice particle or water droplet. Newkirk [2], N. S. Kovar, R. P. Kovar and G. P. Bonner[3] have investigated the problems of contamination of exposed optical surfaces, and the scattering of light by such ice particles under assumed size distribution for the ejecta, and an efficiency of 100 percent for the conversion of water and water vapor into ice droplets. It is thus essential that the knowledge of the size distribution of the ice particles and density, a prior to the study of the optical environment.

The objective of the present research is centered at the review of the literature and the study of ice particle formation in space environments and to study the lifetime

* Numbers in parentheses refer to references at the end of the report.

of these particles parametrically. Although the discussion presented in this report concerns with the Orbital Workshop (OWS) Waste Disposal System [4] , hopefully the theory developed herein, is applicable to all other modes of contamination.

The OWS waste disposal system, shown schematically in Fig. 1, consists of three heated probes for free liquid waste, two unheated probes for free gaseous waste, and a trash airlock for dumping bagged liquids, wet waste and dry trash. Volatile waste is vented overboard through the NPV ducts, but the bagged waste and particles greater than approximately 5 microns are retained in the waste tank by screens as shown in Fig. 2.

The NPV system, shown schematically in Fig. 3, consists of two short straight ducts (1.5 in. Dia. and 13 in. long) that are sealed until payload orbital insertion to retain internal pressure that is required for structural considerations during boost. After orbital insertion, the sealing caps are removed by pneumatic actuators. Thereafter, the waste tank is vented continuously.

The waste management subsystem contributes material to the waste tank through the urine collection and disposal system and the fecal collection and disposal system. The food management subsystem's food and beverage utilization modes also contribute bagged trash to the waste tank. The environmental control system condensate and wash cloth squeezer water is periodically discharged into the waste tank. The waste tank pressure approaches the external ambient pressure (10^{-8} Torr) at the beginning of the first manned mission. When waste material is dumped into the waste tank some of the volatile constituents flash to a vapor, increasing the tank pressure and causing flow out through the NPV ducts. The chemical composition and physical state of this effluent must be known to evaluate its contribution to external contamination and ice formation.

PHYSICAL MODEL OF EXPANSION THROUGH NPV

The analysis of waste tank performance was given in Ref. [4] . It is known from the phase diagram for water that the pressure below the triple point pressure (4.58 mm Hg) water can not exist in a liquid state. For this low pressure regime, water exists as a vapor at temperature above

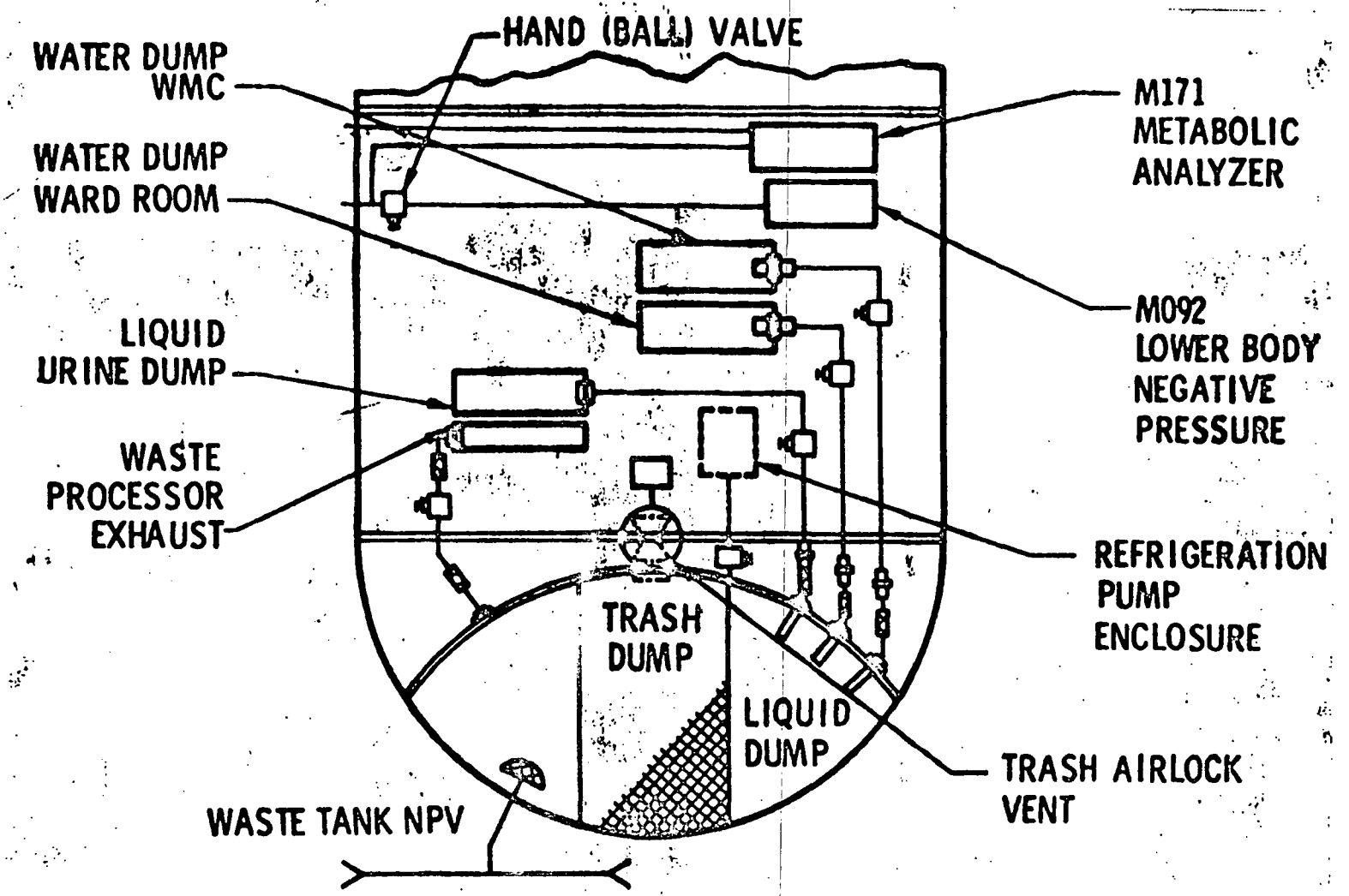


Fig.1 Orbital Workshop Waste Provisions-Schematic

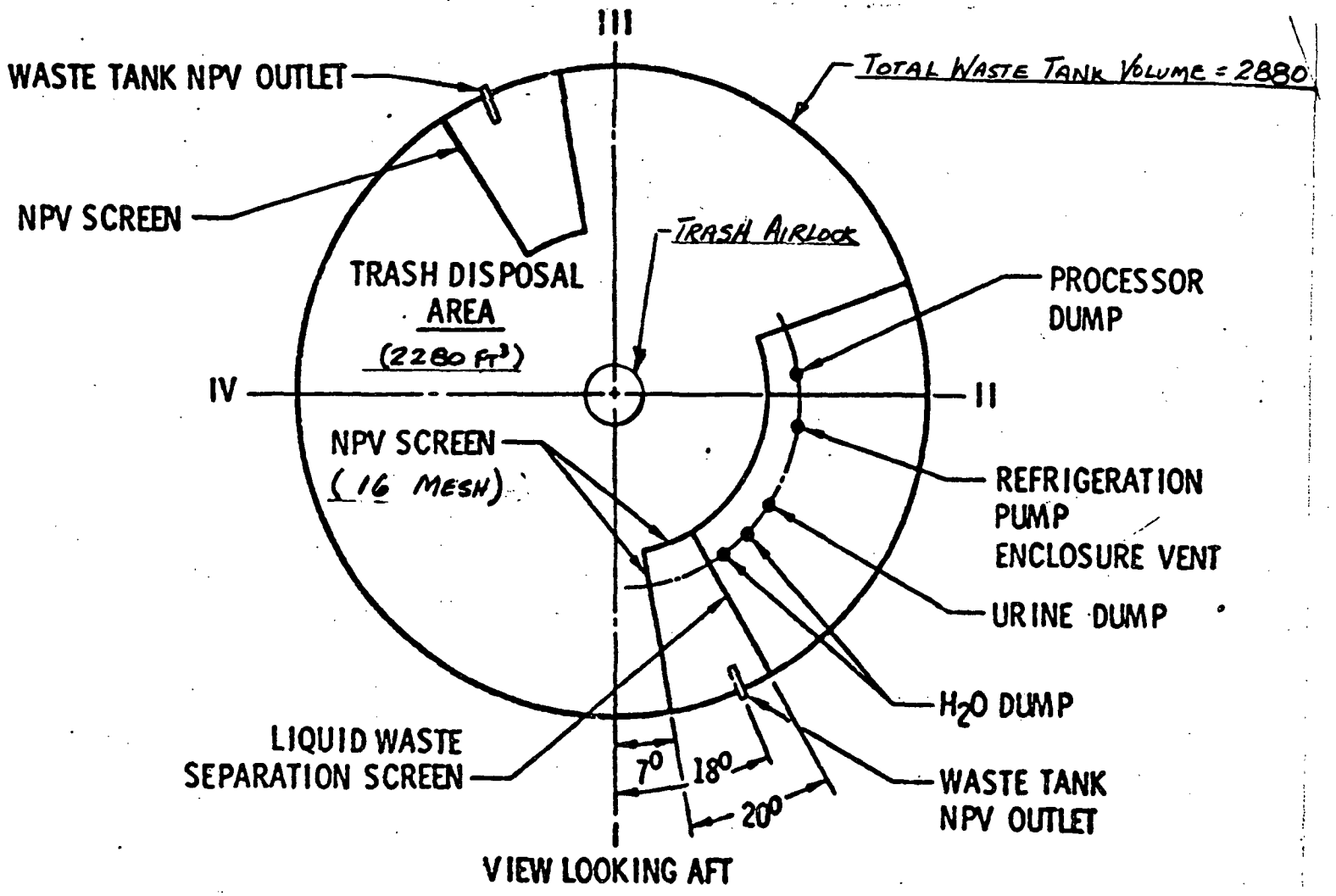


Fig.2 Orbital Workshop Waste Tank Screen Arrangement

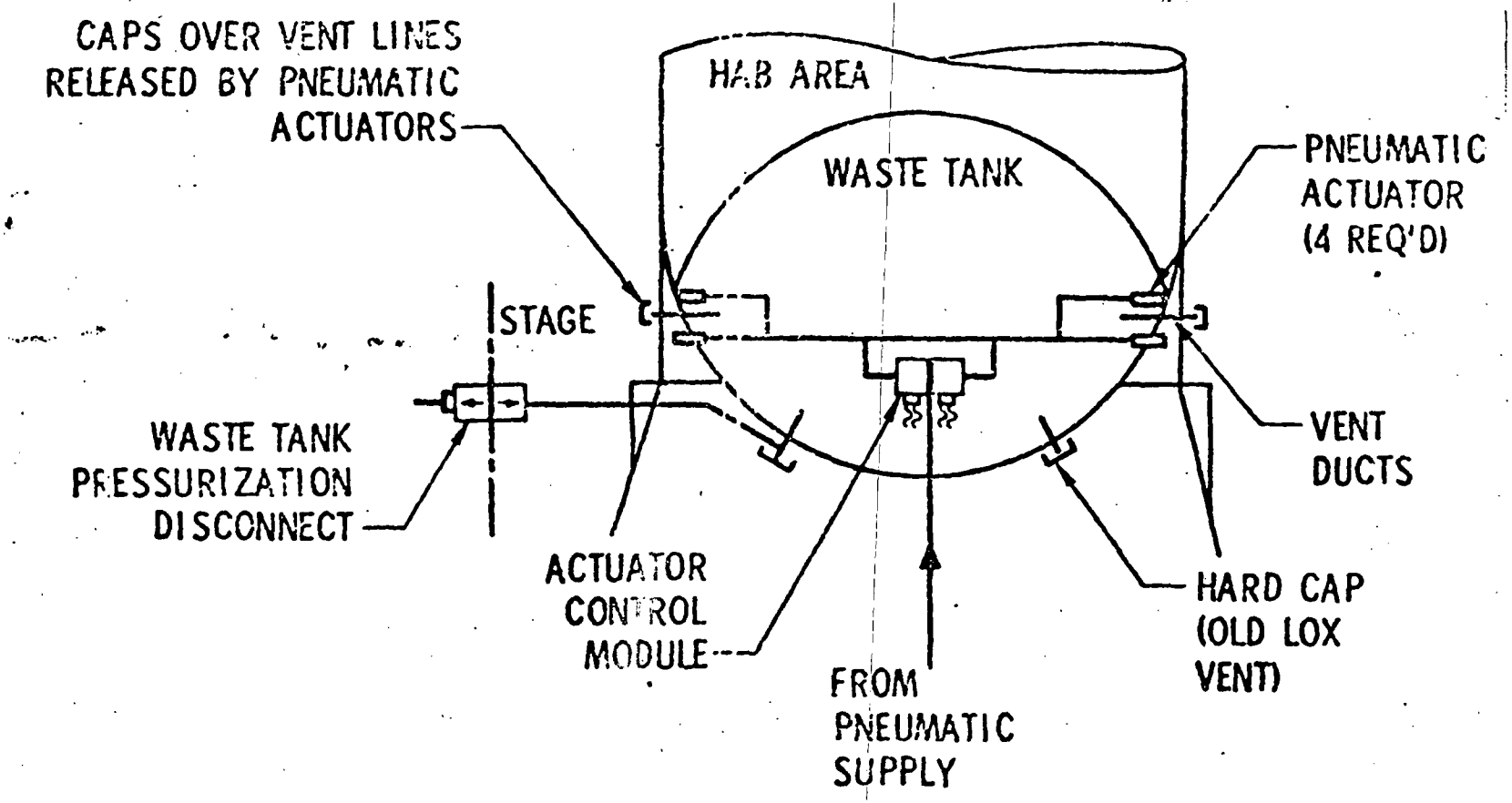


Fig. 3 Waste Tank Launch Pressurization and Venting

0.01°C. At temperatures below this point ice and vapor will coexist. Vaporization in the waste tank increases the tank pressure and subsequent venting decreases the tank pressure. The waste tank pressure P_{wt} and its temperature T_{wt} are nonsteady and depend on the water dump into the tank and out flow rate through NPV.

A typical waste tank pressure and temperature histories obtained from computer analysis is shown graphycally in Fig. 4. To make an analysis feasible, let's consider a simple physical model of the waste tank as shown in Fig. 5. The water vapor with P_{wt} and T_{wt} is expanding through the NPV to ambient pressure P and temperature T with possible heat addition $\dot{Q}_{cv}$ by heater

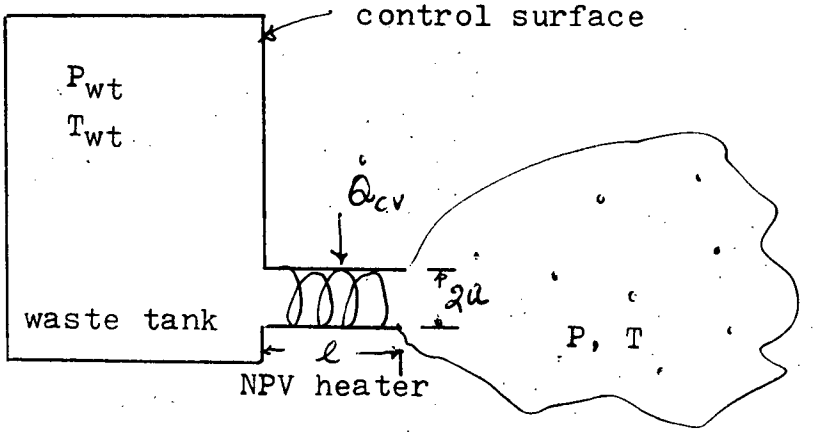


Fig. 5 Physical Model of Waste Tank

installed at NPV. The temperature T is defined as a temperature of the water vapor at local equilibrium (The atmospheric temperature is in different temperature). Therefore the first low of thermodynamics for this system may be written as

$$\dot{Q}_{cv} = \frac{d}{dt}\int_v u\rho dV + \int_A (u + Pv)\rho V_n \cos\alpha \, dA \qquad (1)$$

Eq. (1) can be integrated over time t to give the total energy changes that occur during this period.

$$\int_0^t \dot{Q}_{cv}\, dt = \int_0^t \left(\frac{d}{dt}\int_v u\rho dV\right) dt + \int_0^t \left[\int_A (u + Pv)\rho V_n dA\right] dt \qquad (2)$$

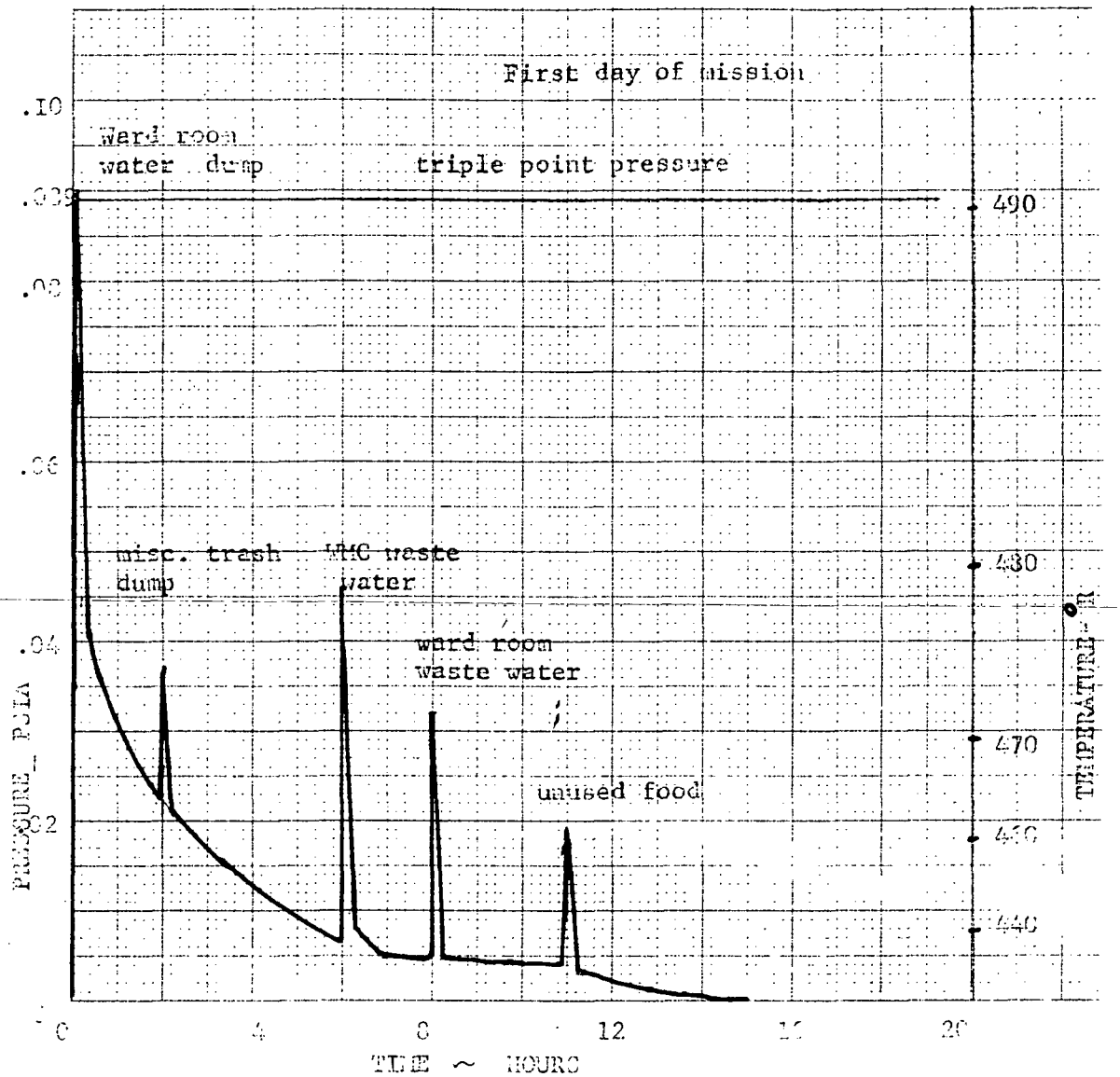


Fig. 4 OWS waste tank pressure and temperature

and for uniform state, uniform-flow process (2) becomes

$$Q_{c.v} = mh + [m_2 u_2 - m_1 u_1]_{c.v} \qquad (3)$$

In the above, u denotes specific internal energy, h specific enthalpy, ρ density, V_r relative velocity, V volume, A area. For this period of time t, we can write the continuity equation for the uniform state, uniform flow process as

$$(m_2 - m_1)_{c.v} + m = 0 \qquad (4)$$

where m the mass leaving the control volume during time t, m_1 and m_2 are the initial and the final mass of water in the tank respectively.

To evaluate m, we may consider a free molecula flow (since P_{wt} and P are low) through a cylindrical tube of length l and radius of a. An approximate equation for m may be obtained for the case ℓ/a is very large as [5]

$$m \simeq \frac{4}{3} \frac{2-f}{f} \left(\frac{2\pi}{RT_{wT}}\right)^{1/2} \frac{a^3}{\ell} \int_0^t \frac{d(P_{wT} - P)}{dt} dt \qquad (5)$$

where f denote Maxwell's reflection coefficient and for diffuse surface, the value of f = 1. With the known value of P_{wt}, T_{wt}, P and $Q_{c.v.}$, the temperature T after expansion may be calculated by using equations (3), (4) and (5).

HOMOGENEOUS NUCLEATION OF ICE PARTICLE BY SUBLIMATION

When the water vapor is expanded through NPV, the water vapor is cooled and the ice particles can be formed by direct sublimation from the vapor in the same way as droplets are formed by condensation. Condensation is a process which does not occur easily in a pure environment. For pure water vapor at room temperature the vapor pressure must be about four times its saturation value before appreciable condensation occurs. However, the environment around the spacecraft is not a pure environment but contains numerous small dust-particles which may be neutral or electrically charged, droplets of various solutions. The influence of these duct particles and/or electric charges on the formation of ice particle exemplify heterogeneous nucleation, i.e., the formation of a droplet about a nucleus foreign to the system of condensable vapor and carrier gas.

The thermodynamic theory of condesation of supersaturated vapors is given in detail by Dufour and Defay [8] and its application to heterogeneous nucleation will be discussed in the following. The treatment given here is based on the methods of Fletcher [9]. Before the discussion of heterogeneous nucleation, we shall discuss the nucleation process in the absence of any foreign particles. This homogeneous nucleation process is not, of itself, of any interest from the point of view of practical application, but we shall discuss it in detail for two reasons. The first of these is that the basic precesses are the same, whether or not foreign particles are present, and the basic physical mechanisms are more easily understood for the simpler homogeneous case. The second reason is that all the concepts and explicit results developed for homogeneous nucleation carry over to the heterogeneous case with only minor modification, so that our detailed examination will be worth while.

Let's consider a system at absolute temperature T, containing vapor at pressure P, and one particle of ice, radius r, consisting of g molecules. Let μ_v denote the free energy of a water molecule in the vapor phase and μ_s in the solid phase. Imagine a spherical ice particle to have been formed by the condensation of g molecules of vapor. The total change in the free energy is

$$\Delta G = (\mu_s - \mu_v) g + 4\pi r^2 \sigma_{sv} = (\mu_s - \mu_v) g + \alpha g^{2/3} \quad (6)$$

α is a constant such that $\alpha g^{2/3} = 4\pi r^2 \sigma_{sv}$

where σ_{sv} is the free energy per unit area of ice-vapor interface. We shall find later that the knowledge of the surface free energy is important in the theory of nucleation, thus, we shall examine this in more detail. The surface energy of ice particle is poorly known, because the lack of mobility of the surface makes direct measurements impossible. There is however a mobility which can be observed, that of a capillary edge which is displace along a solid surface. Consider, for example, the edge A (Fig. 6). Let σ_{Lv} and σ_{sL} denote the surface energy of liquid/vapor and solid/liquid interface. At equilibrium, we have

$$\sigma_{Lv} \cos\phi + \sigma_{sL} = \sigma_{sv} \quad (7)$$

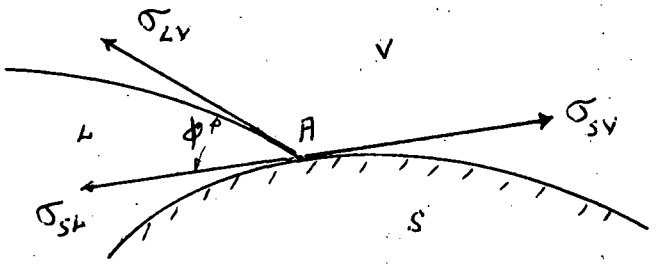


Fig. 6 Equilibrium of an edge A at the surface of solid. V=vapor phase, L=liquid phase, ϕ =contact angle

$$\text{and} \quad \cos\phi = \frac{\sigma_{SV} - \sigma_{SL}}{\sigma_{LV}} \qquad (8)$$

Note that the values of σ_{SV} and σ_{SL} are not accessible to experiment, only their difference $(\sigma_{SV} - \sigma_{SL})$ is an experimental quantity. Various theoretical estimates[11, 12, 13] place σ_{SV} value in the range 100-120 erg/cm^2. It is the author recommendation that more study should be done in this area. In the following analysis, it is assumed that the surface energy is independent of size of the particle. The effect of the size on the surface energy is investigate theoretically by Tolman[13].

The way in which the free energy ΔG, varies with the size of the ice particle, as measured by the value of g, is shown in Fig.7. Curve 1 is for a saturated vapor where $\mu_S = \mu_V$, while curve 2 is for a supersaturated vapor for which $\mu_V > \mu_S$. In the latter case, there is a maximum in the free energy of the system when g has a certain critical value g*, when the embryo has this critical size it is in equilibrium with the vapor, but the equilibrium is unstable. If the particle is greater than the critical size it tends to grow larger with a decrease of free energy. If, however, the vapor phase is thermodynamically stable, embryos of the new phase reach only a relatively small size and then decay.

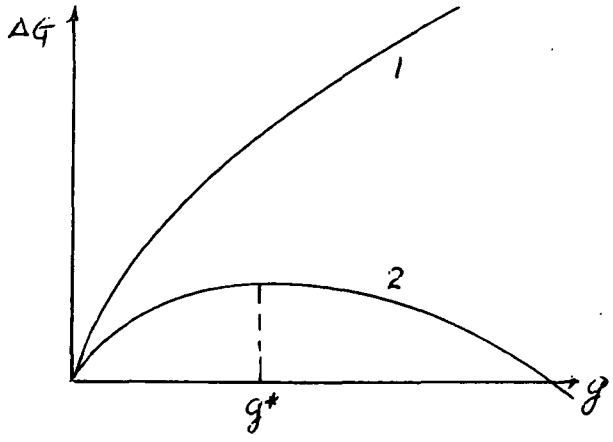


Fig. 7 The free energy ΔG of a particle containing g molecules; curve 1 in saturated vapor, curve 2 in supersaturated vapor, g^* is the critica size.

The condition for equilibrium between the particle and the vapor can be obtained by differentiating (6). It is found that

$$\frac{d\Delta G}{dg} = 0$$

when

$$\mu_s - \mu_v = -\frac{2}{3} \propto g^{-\frac{1}{3}} = -\frac{8\pi\sigma_{sv} r^2}{3g} \qquad (9)$$

but

$$g = \frac{4}{3}\pi r^3 \frac{N\rho_s}{M} \qquad (10)$$

where N is Avogadro's number, ρ_s is the density of the ice, and M its molecular weight. Therefore, combining (9) and (10), we obtain for the equilibrium condition

$$\mu_v - \mu_s = \frac{2M\sigma_{sv}}{N\rho_s r^*} \qquad (11a)$$

where r^* is the radius of the critical nucleus containing g^* molecules and the corrsponding ΔG^* is given by

$$\Delta G^* = \frac{16\pi\sigma_{sv}^3 M^2}{3[\rho_s RT \ln P/P_\infty]^2} \qquad (11b)$$

Now consider the form of μ_s and μ_v. If two phase are in equilibrium, then the chemical potential is the same for molecules in each phase. In the present case ice will be in equilibrium with a saturated vapor with pressure P_∞ over a plane ice surface. The vapor we are considering is at an arbitrary pressure P, so that

$$\mu_s - \mu_v = \mu_v(P_\infty) - \mu_v(P) \qquad (12)$$

Suppose that the pressure of the vapor is changed by the small amount dP at constant temperature, and that the corresponding change in the value of μ_v, the free energy per molecule, is $d\mu_v$. Then if v_v is the volume occupied per molecule in the vapor phase,

$$d\mu_v = + v_v dp \qquad (13a)$$

For the same change of pressure in the solid phase, $d\mu_s$, the change in free energy per molecule is

$$d\mu_s = v_s dp \qquad (13b)$$

Therefore, we have

$$d(\mu_v - \mu_s) = (v_v - v_s) dp \approx \frac{kT}{p} dp \qquad (14)$$

as $v_s \ll v_v$; and k is Boltzmann's constant. Integration of (14) over a vapor pressure range P_∞ to P yields

$$\mu_v - \mu_s = kT \ln \frac{p}{p_\infty} = kT \ln S \qquad (15)$$

where S is the saturation ratio. Substituting for $(\mu_v - \mu_s)$ from (15) into (11) gives

$$\ln \frac{p}{p_\infty} = \frac{2M\sigma_{sv}}{\rho_s R T r^*} \qquad (16)$$

which is Kelvin's formula (R= Universal Gas Constant).

EMBRYO PARTICLES IN STATISTICAL EQUILIBRIUM WITH THE VAPOR

Although, according to(6) and the curves given in Fig.7, the system is in equilibrium only when g=0. For small finite value of ΔG, there is always a small probability, $\exp\{-\Delta G/kT\}$, that the system will be found in a non-equilibrium state because of microscopic density and temperature fluctua-tions. If, in a system containing n_o molecules, there are n_g embryoes each consisting g molecules, then the most probable distribution n_g is shown by Frenkel[14] to be given by

$$n_g = n_o \exp\{-\Delta G/kT\} = n_o \exp\left\{g \ln S - \frac{\alpha g^{2/3}}{kT}\right\} \qquad (17)$$

The way in which the number of drops of a given size n_g varies with the number of molecules g in each is shown in Fig. 8, in which ln n_g is plotted against g. Curve 1 corresponds to a system in which the vapor is

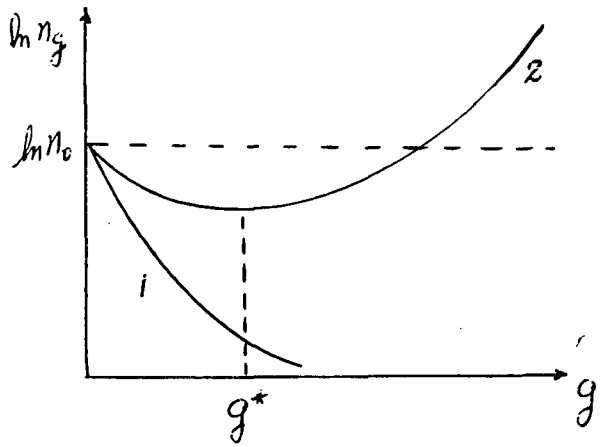


Fig. 8 The equilibrium distribution of droplets; curve 1; a saturated system, curve 2, supersaturated system. Ref. [15]

just saturated; n_g decreases contineously as g is increased. But, for a supersaturated system, represented by curve 2, n_g has a minimum value when the number of molecules in the embryo has a critical value g^* corresponding to the critical radius r^* in (16). For values of $g > g^*$, n_g increases and becomes greater than n_0 when g is very large. Clearly, this part of curve 2 cannot correspond to reality, so that for a supersaturated vapor, (17) breaks down at large values of g. In order to determine the true distribution curve, it is necessary to examine the equilibrium conditions between embryos of various sizes in more detail [15].

NUCLEATION RATE OF ICE PARTICLES

It is now, of course, our main interest to evaluate the rate at which embryos of the critical size are generated within the vapor. This calculation was first made by Volmer and Weber [16], and their approach was subsequently improved by Becker and Doring [17] and by more recent workers [15, 18]. We shall not go into the argument in detail, but sketch its outline. Let's

suppose that the equilibrium distribution given by (17) has been set up, and examine the microscopic balancing conditions for dynamic equiblibrium between embryos of size g, embryos of size g-1, and the vapor. It is assumed that g is a moderately large number so that the equations can be expressed in differential form. The number of vapor molecules β_g striking an embryo of size g, radius r, per second is

$$\beta_g = 4\pi r^2 N P (2\pi M R T)^{-1/2} \qquad (18)$$

It is assumed that all the vapor molecules which hit the embryo are captured by it, i.e. the accommodation coefficient is unity; (Koros [19] reported the value of 0.83 ± 0.12 for the temperature range of -115°C to -140°C for the vapor captured by a flat ice surface). Therefore, the number of embryos of size g-1 becoming embryos of size g per second is $\beta_{g-1} n_{g-1}$. Similarly, if γ_g is the number of molecules evaporating from an embryo of size g per second, then the number of embryos of size g becoming embryos of size g-1 per second is $\gamma_g n_g$. For dynamic equilibrium we have

$$\beta_{g-1} n_{g-1} = \gamma_g n_g \qquad (19)$$

In practice, the distribution given by n_g (see Fig. 17) cannot be set up because it requires an infinite number of very large embryos and these are not available. The detailed balance condition (19) is therefore not satisfied in the supersaturated vapor. Suppose then, that at any time, the number of embryos of size g is f_g instead of n_g. Eq.(19) is no longer valid for the new distribution and we have instead

$$I_g = \beta_{g-1} f_{g-1} - \gamma_g f_g \qquad (20)$$

where I_g is the net number of drops passing from size g-1 to size g per second. At first f_g changes with time, $\frac{\partial f_g}{\partial t} = I_g - I_{g+1}$, but eventually a steady distribution of embryos is reached such that $I_g = I$, a constant independent of g.
Rewrite (20)

$$I_g = n_{g-1}\beta_{g-1}\left[\frac{f_{g-1}}{n_{g-1}} - \frac{f_g}{n_g}\right] = -\beta_g n_g \frac{d}{dg}\left(f_g/n_g\right) \qquad (21)$$

When the distribution f_g is steady, I_g has the constant value I. It is also assumed that β_g is constant (= β). Under these conditions (21) may be integrated to give

$$\frac{f_g}{n_g} = -\frac{I}{\beta}\int \frac{dg}{n_g} + A \qquad (22)$$

Where A is a constant. Substituting n_g by (17) into (22) gives

$$\frac{f_g}{n_g} = -\frac{I}{\beta n_0}\int \exp\left(\frac{\alpha g^{2/3}}{kT}\right)\exp(-g\ln S)\,dg + A$$

$$\approx \frac{I}{\beta n_g}\ln S + A \qquad (23)$$

and as for large values of g, $f_g \propto \frac{1}{\beta}$, A=0 and (22) may be written

$$\frac{f_g}{n_g} = \frac{I}{\beta}\int_g^{\infty}\frac{dg}{n_g} \qquad (24)$$

From the condition $g \to 0$ $\frac{f_g}{n_g} \to 1$, we obtain

$$I = \beta \Big/ \int_0^{\infty}\frac{dg}{n_g} \qquad (25)$$

To calculate I, the number of droplets formed per second, it is necessary to evaluate the integral (25) in which we can substitute for n_g from (17) to give

$$\frac{1}{n_g} = \frac{1}{n_0}\exp\left\{\frac{\alpha g^{2/3}}{kT} - g\ln S\right\} = \frac{1}{n_0}\exp\left\{\frac{\Delta G}{kT}\right\} \qquad (26)$$

Now $1/n_g$ has a sharp maximum value $1/n^*$ when $g=g^*$, so in this region we can replace ΔG by its expansion with respect to the difference $g-g^*$ viz.:

$$\Delta G = \Delta G^* + \frac{1}{2}\left(\frac{\partial^2 \Delta G}{\partial g^2}\right)_{g=g^*}(g-g^*)^2 + \cdots = \Delta G^* - \frac{4\pi}{9g^{*2}}\sigma_{sv} r^{*2}(g-g^*)^2$$

giving

$$\frac{1}{n_g} = \frac{1}{n_0}\exp(\Delta G^*/kT)\exp\left[-\frac{4\pi\sigma_{sv} r^{*2}}{9kTg^{*2}}(g-g^*)^2\right] \qquad (27)$$

since

$$n^* = n_0 \exp\{-4\pi r^{*2}\sigma_{sv}/3kT\} \qquad (28)$$

$$\int_0^{\infty} \frac{dg}{n_g} = \frac{1}{n^*} \int_0^{\infty} \exp\left[-Y(g-g^*)^2\right] dg = \frac{1}{n^*}\left(\frac{\pi}{Y}\right)^{1/2} \qquad (29)$$

where

$$Y = \frac{4\pi\sigma_{sv} r^{*2}}{9kTg^{*2}}$$

and

$$I = \frac{2P}{(2\pi MRT)^{1/2}} \frac{M}{\rho_s} \left(\frac{\sigma_{sv}}{kT}\right)^{1/2} n^* \qquad (30)$$

with

$$r^* = \frac{2M\sigma_{sv}}{RT\rho_s \ln S} \qquad (31)$$

Eqs. (28), (30) and (31) suffice for a numerical calculation of I. Barnard[20] derived a equation similar to (30) but it is valid even for small value of g*.

HETEROGENEOUS NUCLEATION OF ICE PARTICLES

When the water vapor is vented through the NPV, it is most likely to accompany foreigh particles and ions. The ice formation by sublimation may be aided by the presence of the these foreigh particles and in this section, we shall discuss the heterogeneous nucleation by sublimation in the presence of aerosol particles. The treatment now to be given follows that of Fletcher [21, 22]. Refering to Fig. 9, we denote the vapor phase by v, the ice embryo by s and the nucleus by c. If we denote volume by V and surface area by A, then the free energy of formation of an embryo of radius r on a nucleus of radius R is

$$\Delta G = n_s(\mu_s - \mu_v)\bar{V}_s + \sigma_{sv}A_{sv} + (\sigma_{cs} - \sigma_{cv})A_{cs} \qquad (32)$$

where we denote volume of ice by V_s and surface area by A. From the mechanical equilibrium of the line common to these three phases we have the following relation

$$m = \cos\phi = \frac{\sigma_{cv} - \sigma_{cs}}{\sigma_{sv}}$$

From the geometry of Fig. 9, we see that

$$S_{sv} = 2\pi r^2(1-\cos\psi)$$

$$S_{sc} = 2\pi R^2(1-\cos\phi)$$

$$V_s = \frac{1}{3}\pi r^3(2-3\cos\psi+\cos^3\psi) - \frac{1}{3}\pi R^3(2-3\cos\phi+\cos^3\phi)$$

and

$$\cos\phi = (R - r\cos\theta)/d = (R - rm)/d \qquad (34)$$

$$\cos\psi = -(r - R\cos\theta)/d = -(r - Rm)/d$$

with

$$d = (R^2 + r^2 - 2rRm)^{1/2}$$

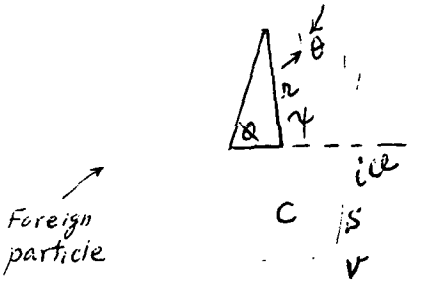


Fig. 9 Ice embryo on nucleating particle in water vapor

The size of the critical embryo can now be determined as before by setting $\frac{\partial \Delta G}{\partial r} = 0$, which lead to the result

$$r^* = \frac{2M\sigma_{sL}}{\rho_s RT \ln P/P_\infty} \qquad (35)$$

which is exactly the same result as for homogeneous nucleation. The free energy of formation of critical embryo is

$$\Delta G^* = \frac{16\pi\sigma_{sv}^3 M^2}{3[\rho_s RT \ln P/P_\infty]^2} f(m,x) \qquad (36)$$

where

$$x = \frac{r}{r^*} = \frac{r\rho_s R \ln P/P_\infty}{2M\sigma_{sv}}$$

and

$$f(m,x) = 1 + \left[\frac{1-mx}{g'}\right]^3 + x^3\left[2 - 3\left(\frac{x-m}{g'}\right) + \left(\frac{x-m}{g'}\right)^3\right] + 3mx^2\left[\frac{x-m}{g'} - 1\right]$$

with $g' = (1+x^2-2mx)^{1/2}$. When m=1.0, the nucleus is completely wet by water and behaves as a droplet of ice of radius R. The nucleation rate per particle is analogously to (30)

$$I \simeq 4\pi 10^{25} R^2 \exp\{-\Delta G^*/kT\} \qquad (37)$$

The effect of electrical charge and surface structure on the nucleation is remained to be examined.

EXPERIMENT ON ICE PARTICLE FORMATION

The Skylab Contamination Ground Test Program [23] has been designed to analyze and test the major vents in a simulated vacuum environment. The Skylab Contamination Ground Test Program (SCGTP) will be conducted utilizing the Martin Marietta Corporation Space Simulation Laboratory Vacuum Chamber to provide a simulated operational environment to measure the anticipated Skylab contamination plumes and particle impingement effects. NPV tests will utilize a simulated waste tank volume with flight-configured NPV installed within the vacuum chamber. Contamination monitor sensors, physical geometry and distances, contamination sources, and typical Skylab dumping and venting sequences will be selected to simulate and record the contamination phenomena. The SCGTP will be designed to accomplish the following objectives;

a. Provide quantitative data about the particle size, distribution, charge distribution, mass flow characteristics, surface contamination effects, and plume effects during a condensate, molecular sieve, fecal processor dump;

b. Determine the OWS Waste Tank pressure and temperature profile, ice accumulation and constituent behavior during a biocide/urine flush, condensate dump, and urine bag rupture. Determine characteristics of the discharged effluent at the two NPV. With some values of parameter chosen, ice partical size, density and effluent may be calculated and compared with SCGTP data.

In the following, a fundamental experimental method by using supersonic nozzle [24, 25] will be proposed for studying the condensation of water vapor by heterogeneous nucleation which is not only interested to the present problem but also important to the cloud physics.

Although the theoretical work of heterogeneous nucleation has been developed by Fletcher; his Eq.(36, 37) has not been fully verified by experiment.

Laboratory techniques for studies of condensation by homogeneous nucleation with different vapors include the use of cloud chambers, turbulent jet mixing apparatus and supersonic nozzles. The use of converging - diverging nozzles is attractive particularly for water vapor condensation studies using moist air. Cooling rates of 10^5 to 10^6 °C/sec can readily be achieved in steady flow. Qualitatively, the nozzle condensation process for homogeneous nucleation in steady flow may be described as follows. An unsaturated water vapor air mixture enters the nozzle supply as shown schematically in Fig. 10. Pressure and temperature drop isentropically and saturation is reached ahead if the sonic throat at x_s. Owing to the absence of condensation surfaces the mixture becomes supersaturated and condensation nuclei are formed. However, their number is small at first and practically no condensate appears. At some point in the supersonic part of the nozzle, the nucleation frequency rises steeply and soon afterwards measurable condensation

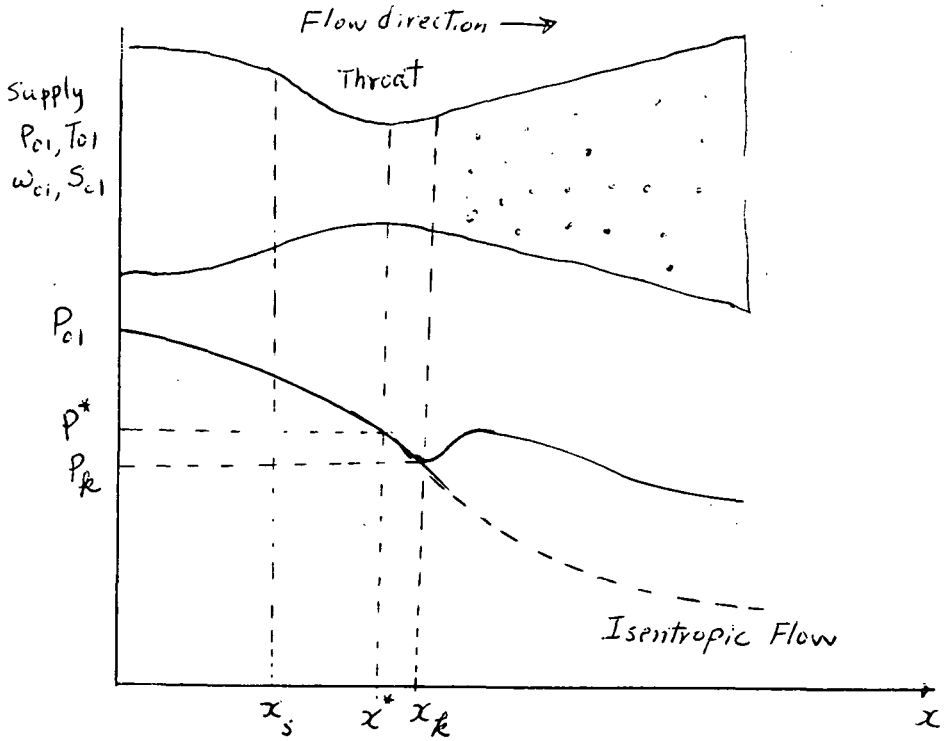


Fig. 10 Sketch of the condensation process of water vapor in nozzles.

sets in at x_k. This point is comparable to the condition of "critical" supersaturation in cloud chamber experiments. It was found that at temperature from 200 to 220°K the condensate was ice. If we can design a sensor to detect the particle size r* or density of the condensate at x_k, (use Cloud Intensity Monitor, Scattering Spectrometer system etc.), the quantities of σ_{sv} and condensation rate I may be calculated by virtue of eqs. (16) and (30). This gasdynamic technique is likely to be extended to study the heterogeneous nucleation by injecting the foreign particles of known size through the nozzle.

LIFETIME OF THE ICE PARTICLES

In Ref.[2] , Newkirk presented a method of calculating various mechanisms, such as aerodynamics drag, radiation pressure, solar wind, charge drag and Lorentz force, orbital dispersion which can accelerate the particles away from the spacecraft. In addition to the above mention mechanism, the evaporation might be expected to be an important removal mechanism. Only the last mechanism will discuss in this report. When the water vapor is vented from NPV, the ice particals of critical size are formed and they may grow to become larger particles or evaporate. If the water vapor pressure surrounding the particle is larger than the equilibrium vapor pressure of the particle, condensation and growth occur. If the surrounding vapor pressure is smaller than the equilibrium vapor pressure of the particle, evaporation occurs. From the molecular theory of effusion, an equation for the rate of change of mass of the particle due to molecules striking the particle of radius r is,

$$\frac{dm_1}{dt} = \alpha 4\pi r^2 (2\pi M R T_\infty)^{-1/2} M P_\infty(T_\infty) \qquad (38)$$

where α is accomodation coefficient and T_∞ is the ambient temperature. In similar manner, one must account for the number of molecules being evaporated, i.e.,

$$\frac{dm_2}{dt} = \alpha 4\pi r^2 (2\pi M R T_r)^{-1/2} M P_r(T_r) \qquad (39)$$

where we have assumed the accomodation coefficient for the evaporation process to be α . Conservation of mass yield the rate of change of radius of the particle[26]

$$\frac{dr}{dt} = \alpha \frac{M}{\rho_v} (2\pi M R)^{-\frac{1}{2}} \left[T_\infty^{-\frac{1}{2}} S P_\infty(T_\infty) - T_r^{-\frac{1}{2}} P_r(T_r) \exp\left(\frac{2M\sigma_{sv}}{R T_r \rho_v r} \right) \right] \qquad (40a)$$

If the classical diffusion theory is used, (40a) may be replaced by

$$4\pi r^2 \frac{dr}{dt} = \frac{4\pi r^2 m}{\rho_s} D \left.\frac{\partial n}{\partial r}\right|_r \qquad (40b)$$

where m is the mass of a water vapor molecule and D is the diffusion constant. The analogous heat balance equation is

$$4\pi r^2 L_{sv} \rho_s \frac{dr}{dt} = -4\pi r^2 K_c \left.\frac{\partial T}{\partial r}\right|_r + \frac{4\pi r^3}{3} \rho_s C_s \frac{\partial T}{\partial t} + 3\pi \sigma_{sv} r \frac{dr}{dt}$$
$$+ \frac{4}{3}\pi r^3 \int \kappa_\lambda F_\odot(\lambda) d\lambda + Q_y \qquad (41)$$

where κ_λ is the absorption coefficient in units of cm^{-1} and $F_\odot(\lambda)$ is the specific flux of sunlight at wavelength λ, L_{sv} is the latent heat of sublimation per unit mass, C_s is the specific heat of ice, K_c is the thermal conductivity of vapor. $\frac{\partial n}{\partial r}$ and $\frac{\partial T}{\partial r}$ satisfy

$$D\nabla^2 n = \frac{\partial n}{\partial t} \qquad (42)$$

$$K_c \nabla^2 T = \frac{\partial T}{\partial t} \qquad (43)$$

The solution of Eqs.(40) through (43) with proper boundary conditions would seem to constitute a proper analysis of the particle growth and evaporation. Some thermal properties required for its calculation may be found in [27] and [20]. The experimental value of absorption coefficient κ_λ is also given in [20].

CONCLUSION AND RECOMMENDATIONS

Analysis is made for the prediction of ice particle formation when the water vapor are vented thru OWS waste tank. With the known pressure and temperature of the tank, ambient pressure and the heat input to the vapor through NPV, the following quantities may be numerically evaluated;

(1) ambient vapor temperature
(2) flow rate of vapor
(3) escape velocity of vapor
(4) critical size of ice particle
(5) rate of formation of critical size particle (for heterogeneous nucleation, the size of foreign particles should be known)

Although no numerical calculation has been made at the present time, the author hopes that this will be done in the future and made some comparison with the experimental data obtained from the SCGTP performed by Martin Marietta Co., Denver, Colorado.

In all the treatments so far mentioned, aggregates of the ice are regarded as well-defined solid spherical continum particle having the thermodynamic and physical properties of macroscopic ice particle. The macroscopic concept become very vague when the aggregates become very small and contain as few as 20 molecules. Under such condition, the macroscopic properties are probably meaningless. A statistical-mechanical treatment of condensation is necessary. Concerning the lifetime of ice particles, Eq.(41) is valid only for a single particle and in practice, the cloud of ice shall be considered as a whole. This cloud of ice may be treated as a system of gas and the information about the absorptivity, transmitivity and reflectivity of the cloud is essential for the future treatment of the problem. Further research is necessary in this direction.

The gasdynamic techniques is recommended for studying the homogeneous and heterogeneous nucleation of ice particle. This techniques provides a mean of detecting the onset of sublimation without recourse to the difficult and uncertain direct observation of particles. With the availability of the sensors to detect the particle size and its density at critical location, the surface energy of particle, and nucleation rate can be obtained indirectly. It is hoped that some future endeavors will be done along this line.

REFERENCES

1 Ney, E. P. and Huch, W. F. "Optical Environment in Gemini Space Flights" Science Vol. 153 pp297 (1966)

2 Newkirk, G. Jr. "The Optical Environments of Manned Spacecraft" Planet Space Sci. Vol. 15, pp 1267 - 1285 (1967)

3 Kovar, N. S., Kovar, R. P. and Bonner, G. P. , "Light Scattering by Manned Spacecraft Atmospheres" Planet Space Sci. Vol. 17 pp 143 - 154 (1969)

4 Samuel, H. D. "Evaluation of Contamination Generated by Waste Disposal Through the Waste Tank Nonpropulsive Vent Ducts" Saturn System Engineering Study Report, Report No. MDC G2168-P (1971) Marshall Space Flight Center, Huntsville, Ala.

5 Kennard, E. H., Kinetic Theory of Gases, McGrall-Hill 1938 pp 306

6. Mason, B. J., The Physics of clouds, Oxford at the Clarendon, Press 1957

7 Fletcher, N. H., The Physics of Rainclouds, Cambridge University Press 1962

8 Dufour, L., and Defay, R., Thermodynamics of Clouds Academic Press, 1963

9 Fletcher, N. H., "Size Effect in Heterogeneous Nucleation" J. Chem. Phys. Vol. 29 pp 572 (1958)

10 Mason, B. J., "The Spontaneous Crystallyation of Supercooled Water", Quart. J. R. Met. Soc. Vol. 78 pp 22 (1952)

11 McDonald, J. E., "Homogeneous Nucleation of Supercooled Water Drops", J. Met. Vol. 10, pp 416 (1953)

12 DeRenck, A. V. S., "The Surface Free Energy of Ice", Nature, Lond., Vol. 179 pp 1119 (1957)

13 Tolman, R. C., "The Effect of Droplet Size on Surface Tension", J. Chem. Phys. Vol. 17, pp 333 (1949)

14 Frenkel, J., Kinetic Theory of Liquids, Clarendon Press, Oxford pp 382 (1946)

15 Farley, F. J. M., "The Theory of the Condensation of Supersaturated Ion-free Vapor", Proc, Roy. Soc. A , Vol. 212 pp 530 (1952)

16 Volmer, M. and Weber, A., "Keimbildung in Ubersattigen Gebilden", Z. Phys. Chem. Vol. 119 pp 277 (1926)

17 Becker, R. and Döring, W., "Kinetische Behandlung der Keimbilung in Ubersattigen Dampfen" Ann. Phys., Lpz., Vol. 24 pp 719 (1935)

18 Barnard, A. J.,"The Theory of Condensation of Supersaturated Vapors in the Absence of Ions" Proc. Roy. Soc. A220 pp132 (1953)

19 Korous, R. M., Sticking Probability of Water Vapor Molecules on An Ice Surface, Ph. D. Thesis, Princeton University, (1961)

20 Riehl, N., Bullemer, B. , Engelhadt, H. Editor, Physics of Ice, Plenum Press, pp271,(1969)

21 Fletcher, N. H., "Size Effect in Heterogeneous Nucleation" J. Chem. Phys. Vol.29, No 3 pp572(1958)

22 Fletcher, N. H., "On Ice Crystal Production by Aerosol Particles", J. Met. Vol. 16 pp 173 (1959)

23 Skylab Contamination Ground Test Program Plan NASA George C. Marshall Space Flight Center Report Report No. 10M33110 Rev. B March 1972

24 Wegener, P. P. "Water Vapor Condensation Process in Supersonic Nozzles" J. of Appl. Phys. Vol. 25, pp 1485 (1954)

25 Wegener, P. P. and Pouring, A., "Experiments on Condensation of Water Vapor by Homogeneous Nucleation in Nozzles" The Physics of Fluids Vol. 7, No. 3, pp 352 (1964)

26 Schuster, B. G., "Observation of Homogeneous Nucleation and Droplet Growth in a Wilson Cloud Chamber" Ph.D. Thesis New Mexico State University (1967)

27 Smith, J. G., "Re-evaporation Nuclei and Evaporation in a Wilson Cloud Chamber" Ph.D. Thesis, University of Missouri at Rolla (1966)

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

SATELLITE AERODYNAMICS AND DENSITY DETERMINATION FROM SATELLITE DYNAMIC RESPONSE

Prepared by:	Gerald R. Karr
Academic Rank:	Assistant Professor
University:	University of Illinois at Urbana-Champaign
Laboratory: (Division) (Branch)	Aero-Astrodynamics Laboratory Aerospace Environment Division Space Environment Branch
Research Counterpart:	R. E. Smith
Contract No.:	NGT-01-003-045

SATELLITE AERODYNAMICS AND ATMOSPHERIC DENSITY DETERMINATION FROM SATELLITE DYNAMIC RESPONSE

by

Gerald R. Karr

ABSTRACT

A method for determining satellite aerodynamic properties and upper atmospheric density from observed satellite dynamic response has been successfully developed and tested.

The aerodynamic drag and lift properties of a satellite are first expressed as a function of two parameters associated with gas-surface interaction at the satellite surface. The dynamic response of the satellite as it passes through the atmosphere is then expressed as a function of the two gas-surface interaction parameters, the atmospheric density, the satellite velocity, and the satellite orientation to the high speed flow. By proper correlation of the observed dynamic response with the changing angle of attack of the satellite, it is found that the two unknown gas-surface interaction parameters can be determined. Once the gas-surface interaction parameters are known, the aerodynamic properties of the satellite at all angles of attack are also determined. The atmospheric density may then be accurately calculated once the true aerodynamic properties are known.

Employing accelerometer data from the OVI-15 satellite, analysis was successfully made of the aerodynamic properties of that satellite and a determination was made of the absolute value of atmospheric density near the orbit perigee. These results constitute the first successful application of the proposed method of analysis. These results also serve to illustrate the potential of the technique in the analysis and prediction of satellite orbit decay in the atmosphere and the accurate determination of upper atmospheric density from satellite dynamic response.

1. Introduction

The problem of satellite orbit decay prediction and the problem of upper atmospheric density determination have encountered a common source of unknown error which can be traced to a lack of knowledge of satellite aerodynamics. The basic equation employed in both orbit decay and density determination is the familiar drag equation:

$$\text{Drag} = 1/2\ \rho U^2 C_D \bar{A}$$

where ρ is the density, U is the velocity of the satellite with respect to the atmosphere, C_D is the drag coefficient, and $\bar{A}$ is a suitable reference area. In most applications of this equation to satellites the value of C_D is considered to have a constant value. Generally, however, the assumption of constant drag coefficient is not valid and the use of such an assumption can lead to considerable error (see Karr, 1972). The uncertainty in satellite aerodynamics has prevented the assignment of even an approximate value of drag coefficient with a known range of uncertainty. A value of C_D of 2.0 or 2.2 is often used in satellite drag studies and in the determination of atmospheric density. These values of C_D are likely too small and, combined with the fact that C_D is not constant, have resulted in an overestimation of upper atmospheric densities (see Karr and Smith 1972).

A more accurate treatment of satellite aerodynamics has obvious benefit to the determination of upper atmospheric density and the prediction of satellite orbit decay. Satellites traveling in the earth's upper atmosphere experience the aerodynamic flow regime termed the free molecular flow regime. In this flow regime, the collision of atmospheric gas molecules with the satellite surface dominate the flow and collisions of gas molecules with other gas molecules may be ignored. Satellite aerodynamic properties are then the result of the interaction of high speed gas molecules with the solid satellite surface. Unfortunately, very little is known about the gas surface interaction at satellite velocities and this lack of information is the basic source of uncertainty in satellite aerodynamic properties.

In the interest of developing a more accurate treatment of satellite aerodynamics for application to orbit decay prediction and density determination, a model of the gas surface interaction has been developed which utilizes two parameters to describe the interaction (see Karr, 1969 and Karr and Yen, 1970). The advantage in this treatment of satellite

aerodynamics is that no a priori assumptions of the aerodynamic properties need be made. The gas surface interaction parameters are considered as unknowns to be determined from the observed dynamic response of the satellite as it travels through the atmosphere. The determination of the gas surface interaction parameters serves as the key to the subsequent determination of both the aerodynamic properties and the atmospheric density.

To test the proposed method of analysis, accelerometer data from the OVI-15 satellite is used. The accelerometer data provide an accurate, instantaneous measure of the level of aerodynamic force and the attitude of the satellite with respect to the flow. As is pointed out in the paper, accurate satellite attitude informatin is essential to the analysis. The OVI-15 satellite, although less than ideal in shape for an aerodynamic study, provided a good basis for the test of the proposed method of analysis. The results serve to illustrate the potential that this method of analysis has to future determinations of aerodynamic and atmospheric properties.

2. Satellite aerodynamics and the gas surface interaction.

Consider a local satellite surface element in which the high speed flow of molecules is incident at an angle of as shown in Figure 1. Associated with the incident flow is the incident momentum which gives rise to the incident force $\vec{F}_i$. This force is colinear with the satellite velocity, U, with respect to the atmosphere. Assume for now that the speed ratio is infinite where the speed ratio is defined as the satellite velocity divided by the thermal velocity of the gas molecules. The thermal velocity of the gas molecules is taken to be equal to $\sqrt{R\ T/M}$ where R is the gas constant, T is the temperature, M is the mean molecular weight.

The molecules reflected from the surface cause a net reaction force $\vec{F}_r$ which is colinear with the mass-motion velocity vector $\vec{U}_j$ of the molecules leaving the surface. The direction of U_j is given by the angle θ_j.

Modeling of the interaction is performed by providing relationships between the incident and reflected quantities. The relationships are given by

$$U_j = \sqrt{1 - \alpha_j}\ U$$

$$\theta_j = \frac{\pi}{2} P_j + (1-P_j)\ \theta$$

where α_j and P_j are the parameters of the interaction. The subscript j is used if more than one set of interaction coefficients is to be considered in the analysis. Unitl more is learned of the interaction, these linear relationships provide a useful first approximation to the interaction that occurs at satellite velocities. The parameters α_j and P_j are capable of describing a much wider range of possible interactions than other models. The development of this model and the capabilities are described in detail in Karr 1969 and Karr and Yen, 1970.

The model described above is particularly useful in the determination of forces acting on the satellite surface. The total vector force acting on the element of surface shown in Figure 1 is given by

$$d\vec{F} = -(\vec{U}-\sigma_j\vec{U}_j)\,\rho\,\vec{U}\cdot\vec{n}\,dA$$

where σ_j is employed if more than one gas surface interaction is employed in the analysis. In order to conserve mass at the surface, the sum of the σ_j values must be unity.

The magnitude and direction of U_j is determined by the parameters α_j and P_j. The vector force acting on the local element of surface is then expressed as a function of α_j, P_j, U, ρ, and θ. For a given satellite shape (assumed to be convex) the total aerodynamic forces and torque acting on the satellite are found by integration of $d\vec{F}$ and $\vec{R}\times d\vec{F}$ over the surface exposed to the flow. In general, the results will be of the form

$$\text{Drag} = 1/2\,\rho\,U^2\,C_D\,(\alpha_j,\ P_j, \beta)\,\bar{A}$$

$$\text{Lift}_{1,2} = 1/2\,\rho\,U^2\,C_{L_{1,2}}\,(\alpha_j,\ P_j, \beta)\,\bar{A}$$

$$\text{Torque}_{1,2,3} = 1/2\rho U^2\,C_{T_{1,2,3}}(\alpha_j,\ P_j, \beta)\,\bar{A}\bar{L}$$

where θ is an angle of orientation and the subscripts on C_L and C_T are to indicate that there are two components of lift and three components of torque. The six aerodynamic properties are found to be a strong function of the gas surface interaction parameters. For non-spherical objects, the angle of orientation, β, also has a strong influence on the drag, lift and torque properties (see Karr and Yen, 1970).

3. Aerodynamics of the OVI-15 satellite.

The approximated shape of the OVI-15 satellite is a cylinder with spherical ends as shown in Figure 2. The satellite spin axis was normal to the longitudinal axis of the cylinder with a spin rate of about 10 rpm. Near the center of the satellite a three axis accelerometer detected the forces acting on the satellite. Since the data to be used in the subsequent analysis has been filtered and averaged over a number of spin cycles, the aerodynamic properties averaged over a spin cycle are developed.

The total instantaneous vector force acting on the satellite is given by

$$\vec{F} = 1/2\ \rho\ U^2\ \bar{A}\ (C_D\ D + C_{L_1}\ L_1 + C_{L_2}\ L_2)$$

where D, L_1, and L_2 are mutually orthogonal unit vectors in the drag and lift directions. The L_1 and L_2 directions are defined with respect to the instantaneous orientation such that L_1 is perpendicular to the cylinder axis and the velocity vector. The direction of L_2 is perpendicular to both L_1 and D. Due to symmetry the lift force in the L_1 direction is zero.

From Karr, 1969, C_D and C_{L_2} are obtained for the infinite speed ratio case, given by

$$C_D = 2 + 4\sqrt{1-\alpha_j}\ \left(1 - \cos\frac{\pi}{2}P_j\right)\ P_j\ (4 - P_j)$$
$$+ 2A_R \cos\theta_s$$
$$+ A_R\sqrt{1-\alpha_j}\int_{\pi}^{2\pi}\left[\cos\theta_s \sin\xi\, C_j - \cos^3\theta_s \sin^3\xi\,(C_j + S_j)\right]d\xi$$

$$C_{L_2} = -\sqrt{1-\alpha_j}\ A_R\ \cos^2\theta_s\ \sin\theta_s\int_{\pi}^{2\pi}\sin^3\xi(C_j + S_j)d\xi$$

$$C_{L_1} = 0$$

where the first two terms in C_D are due to the spherical ends and the remaining terms are due to the cylindrical section. The angle ξ is a cylindrical surface-integration angle. The quantity A_R is the area ratio of the cylinder to the sphere given by

$$A_R = \bar{A}_{cyl}/\bar{A}_{sph} = 2rL/\pi r^2 = 4L/\pi D$$

The quantities C_j and S_j contain the parameter P_j where

$$C_j = \frac{\cos\left(\pi/2\, P_j + (1 - P_j)\,\theta\right)}{\cos\theta}$$

$$S_j = \frac{\sin\left(\pi/2\, P_j + (1 - P_j)\,\theta\right)}{\sin\theta}$$

$$\theta = \sin^{-1}\left(-\cos\theta_s \sin\xi\right)$$

The angle θ_s is an angle of instantaneous orientation defined as the angle between the velocity vector and the longitudinal axis of the cylinder.

Since the OVI-15 spin axis is perpendicular to the axis of the cylinder, the angle θ_s is a function of the angle the spin axis makes with the velocity vector, γ, and a spin angle, λ, which changes from 0 to 2π every spin cycle (see Figure 3).

For certain values of P_j the surface integrals over the angle ξ are easily performed. For $P_j = 0$, which corresponds to specular type reflection,

$$C_j\Big|_{P_j=0} = S_j\Big|_{P_j=0} = 1$$

For $P_j = 1$ which corresponds to diffusive type reflection,

$$C_j\Big|_{P_j=1} = 0 \quad ; \quad S_j\Big|_{P_j=1} = -1/\cos\theta_s \sin\xi$$

For $P_j = 2$ which corresponds to perfect backscatter type reflections

$$C_j\Big|_{P_j=2} = -1 \quad ; \quad S_j\Big|_{P_j=2} = 1$$

The values of C_D and C_{L_2} at the three values of $P_j = 0, 1, 2$, were used to obtain an polynomial approximation for C_D and C_L as a function of P_j.

Since the accelerometers were body fixed, the output of the accelerometers were a function of both drag and lift forces given by

$$F/\frac{1}{2}\rho U^2 \bar{A} = \left[-C_D \sin\gamma \cos\lambda + C_L \frac{1 - \sin^2\gamma \cos^2\lambda}{\cos\theta_s}\right] i$$

$$+\left[-C_D \sin\gamma \sin\lambda - C_L \frac{\sin^2\gamma \cos\lambda \sin\lambda}{\cos\theta_s}\right] j$$

$$+\left[-C_D \cos\gamma - C_L \frac{\sin\gamma \cos\gamma \cos\beta}{\cos\theta_s}\right] k$$

where i, j, and k are unit vectors in the directions of the three axis of the accelerometers with i along the axis of the cylinder, k is in the direction of the nominal spin axis and j is orthogonal. Since the data being used in this analysis has been averaged over spin cycles, the component of force as expressed above were integrated over the angle λ. The results after averaging over one spin cycle were

$$\bar{F}_y = \sin\gamma \frac{1}{2} \rho U^2 \pi r^2 C_F(\alpha_j, P_j, \gamma)$$

$$\bar{F}_z = \cos\gamma \frac{1}{2} \rho U^2 \pi r^2 C_F(\alpha_j, P_j, \gamma)$$

where C_F is the integrated force coefficient.* These results show that $\bar{F}_y$ and $\bar{F}_z$ measure the identical forces except for the factor $\sin\gamma$ and $\cos\gamma$. This property was used by Fess and Young to obtain the angle γ which the spin axis makes with the velocity vector.

$$\gamma = \cos^{-1}\left[\bar{F}_z / \sqrt{\bar{F}_y^2 + \bar{F}_z^2}\right]$$

The force coefficient C_F was found by fitting a 3rd order polynomial to the values of C_F and C_{L_2} at the three values of $P_j = 0$, 1, and 2.

$$C_F = A + \sqrt{1 - \alpha_j}\,(G + H P_j + Q P_j^2 + P P_j^3)$$

where

$$A = -2 - 4 A_R E(\gamma, \tfrac{\pi}{2}) / \pi$$

$$G = -4 A_R E(\gamma, \tfrac{\pi}{2}) / 3\pi$$

$$H = 4F - 2G - 2A$$

$$Q = -4F + 5G/4 + 11A/4$$

$$P = F - G/4 - 3A/4$$

$$F = -4/3 - \pi A_R/2$$

where $E(\gamma, \frac{\pi}{2})$ is a complete elleptic integral of the second kind resulting from the average over one spin cycle.

* The accelerometer in the x direction did not function so only $\bar{F}_y$ and $\bar{F}_z$ are treated in the analysis.

4. Method of analysis.

The objective of the analysis is to find the best values of α_j, P_j and density which explains the observed accelerometer output of the OVI-15. Although α_j, P_j, and ρ are considered as unknowns in the analysis, certain assumptions on the characteristic variation of ρ are made to facilitate the analysis. The assumption made is that the density variation is symmetric with respect to the perigee of the orbit. The absolute value of density is still treated as an unknown quantity.

4.1 Least squares fit.

Assuming symmetrical density variation about perigee, differences in forces measured at points equal distance from perigee must be due to changes in the aerodynamic force coefficient, C_F. Since the density is equal at these two points, we can write

$$\frac{1}{2} U_1^2 \rho_1 \quad (T-\Delta t_i) = \frac{1}{2} U_2^2 \rho_2 (T+\Delta t_i)$$

where the subscript 1 indicates approach to perigee, subscript 2 indicates recession from perigee and T is the perigee passage time. The aerodynamic properties and forces measured at these two points must then satisfy the following relationship

$$\frac{\overline{F}_{i1}}{C_F(\alpha_j, P_j, \gamma_{i1})} = \frac{\overline{F}_{i2}}{C_F(\alpha_j, P_j, \gamma_{i2})}$$

where γ_{i1} and γ_{i2} are the angles of orientation at $T-\Delta t_i$ and $T+\Delta t_i$ respectively, and $\overline{F} = \sqrt{\overline{F}_z^2 + \overline{F}_y^2}$. In the analysis, the quantity DEL_i is found from the preceeding relation, defined as,

$$DEL_i = \overline{F}_{i1} C_{F_{i2}} - \overline{F}_{i2} C_{F_{i1}}$$

where i is used to indicate a comparison made at $T \pm \Delta t_i$. A solution in the least squares sense is obtained by finding the values of α_j and P_j which provide a minimum to the sum-of-DEL_i-squared for a number of observations near perigee

$$SUM = \sum_{i=1}^{n} (DEL_i)^2$$

The best values of α_j and P_j are those which satisfy

$$\frac{\partial(SUM)}{\partial(\sqrt{1-\alpha_j})} = 0 \quad ; \quad \frac{\partial(SUM)}{\partial P_j} = 0$$

in the region of $0 \leq P_j \leq 2$, and $0 \leq \sqrt{1-\alpha_j} \leq 2$.

4.2 Perigee passage time.

The analysis requires first that the aerodynamic properties be different at the comparison points used in the analysis. If the C_F were not different, then the last equations would be satisfied for all values of α_j and P_j. Although the OVI-15 satellite was designed to maintain a $\gamma = 90^o$ thoughout the orbit, considerable uncontrolled drift in the spin axis was found to occur. This malfunction was desirable for purposes of this analysis since the angle γ changed considerably during a given orbit. This factor resulted in a changing value of C_F over the orbit which provided a good sampling of C_F and $\bar{F}$ values for a wide range of angles of orientation. The analysis also requires that the perigee passage time be accurately known since the comparisons are made at equal points on each side of this time. Since the report of Fess and Young did not provide a perigee passage time, it was necessary to calculate that time from the accelerometer output. Since the aerodynamic properties are changing during a perigee pass due to the changing angle of orientation, the peak in the $\bar{F}$ curve is shifted in time from the peak in dynamic pressure. The derivative of $\bar{F}$ during a perigee pass is

$$\bar{F}' = S\, C_F' + S'\, C_F$$

where S is the dynamic pressure. At the perigee passage time, the dynamic pressure is maximum and $S' = 0$. Therefore, at the perigee passage time

$$\bar{F}'(T) = \frac{F(T)}{C_F(T)}\; C_F'(T)$$

This equation was employed to find the true value of T for each data set employed. The quantities C_F and C_F' are a function of α_j and P_j in addition to the angle γ. The analysis was able to take into consideration the expected shift in $\bar{F}$ output which was found to vary from 3 to 15 seconds depending upon α_j, P_j and the rate of change in the angle γ.

4.3 Speed ratio effect.

As discussed by Karr 1972 and Karr and Smith 1972, changes in speed ratio with altitude result in a systematic increase of C_D with altitude. The amount of increase was found to be a function of the satellite shape and orientation. This factor was taken into consideration in the analysis of data of the OVI-15 by approximating the expected change in aerodynamics properties with respect to speed ratio. Using information from Karr and Smith 1972 and taking into account the average over a spin cycle, the following speed ratio correction factor was obtained.

$$COF = 1 + .682/S + .56148/S^2 + .4 \sin^2 \gamma\, (1.66/S - .1528/S^2)$$

The correction factor is a function of the angle of orientation which must be taken into account in the determination of α_j and P_j. For the altitudes of interest for the OVI-15, speed ratios of about 10 are obtained which result in an increase in the aerodynamic force coefficient of about 5%. The determination of density is then strongly influenced by the speed ratio effect.

5. Results.

5.1 Data used in the analysis.

An example of the data employed in the analysis is shown in Figure 4. Data from orbits number 890, 893 and 896 were employed in the least squares fit to α_j and P_j. These orbits had significant changes in γ during perigee passage and experienced approximately the same atmospheric conditions. These nearly polar orbits occurred in mid-September 1969 with perigees at 150 km, perigee latitude at 10 °S latitude with a local perigee time at about 1930. Since the sun declination was near +3 °, the variation in density with latitude near perigee is approximately zero according to the Jacchia 1971 model of the atmosphere.

The choice of orbits used in the analysis contributed to the reduction of errors resulting from any nonsymmetry of density variation. Further reduction in this type error was made by using only data within ±10 ° of perigee. In addition, since data from three orbits was used, the errors due to wave motion or other short duration density disturbances would contribute only to the random error.

The values of accelerometer output $\bar{F}$, angle of orientation γ, and time in seconds from the beginning of data transmission are given in Table I. The units on $\bar{F}$ is in counts in which 5.3 counts equal 10^{-6} of acceleration. The angle γ is given in degrees. The data is seen to cover an angle of attack range of about 28 degrees. All the data falls within 150 seconds of perigee which for these orbits means that the data is taken within ± 10 ° of perigee. Since the true perigee is always within ± 20 seconds of the peak $\bar{F}$, corrected perigee times will not change the range of data significantly.

5.2 Gas Surface interaction parameters.

Using the data given in Table I and taking into account the perigee passage time correction and the speed ratio correction, the results of the sum-of-the-squares-of-DEL_i are given in Figure 5. These results show a unique minimum of the sum-of-the-squares-of-DEL_i at $P_j = .44$ and $\sqrt{1-\alpha_j} = .6$. These values for α_j and P_j mean that the reflection is between a specular and diffusive type reflection in direction and has moderate accomadation of energy.

5.3 Aerodynamic properties.

Using $\alpha_j = .64$ and $P_j = .44$, the aerodynamic force properties of the OVI-15 satellite may be found using the equations already derived. The results are shown in Figure 6 which gives C_F as a function of angle of orientation. The level of C_F is dependent directly on the reference area, $\bar{A}$, choosen to represent the satellite. The plot is given for two acceptable areas (1) the maximum cross sectional area seen by the flow, $\pi r^2 + 2rL$ and (2) the minimum cross sectional area seen by the flow, πr^2. These results are for the infinite speed ratio case and must be modified according to speed ratio influence.

At angles of 0, 90 and 180°, the quantity C_F is equal to the drag coefficient. At all other angles, C_F is influenced by both drag and lift coefficients. These results show that the drag coefficient is higher than the value of 2.2 normally assumed in drag analysis. The speed ratio correction will cause these values to be increased by about 5% to 10% for the altitudes at which the data was taken.

The atmospheric density values may be obtained since the values of C_F throughout the orbit have been calculated. Assuming an orbit of e = .113 and perigee at 150 km, the value of U at the data points are obtained and density values are given by

$$P(I, J) = 2amF(I, J)/U^2(I, J)\bar{A}C_F(I, J)\ COF$$

where COF is the speed ratio correction factor dependent upon the angle $\gamma(I, J)$, a is the conversion factor needed to convert accelerometer counts into accelerometer values ($a = 10^6 g/5.3$ counts), and m is the satellite mass = 214 kg. The speed ratio correction requires an estimate of the speed ratio. For the date, time, and region of the atmosphere for which the data cooresponds, an estimate was made of the exospheric temperature from information provided by Smith, 1972. In this region of the atmosphere the exospheric temperature remains essentially constant and was taken to be 1100°. Using the Jacchia 1971 model atmosphere, a value of T/M versus altitude were fitted to a polynomial over the altitude range of interest from 140 km to 220 km. The speed ratio is then given at each data point by

$$S = U/\sqrt{2RT/M}$$

Where R is the universal gas constant. Values of density calculated in this manner are about 5 to 15% less than those predicted by the Jacchia 1971 model. A more complete discussion of these results will be made at a later date.

6. Discussion of results.

The results of this analysis are important for a number of reasons. First, the analysis illustrates a new method for the analysis of satellite

dynamic response. Second, the results obtained for α_j and P_j are the most accurate of measurements of the gas surface interaction parameters at satellite velocity. Third, the values of C_F obtained in the analysis are the most accurate of measurements of satellite aerodynamic properties. Finally, the results obtained for density are the most accurate of measurements of absolute values of upper atmospheric density.

Past drag analysis have required that critical assumptions be made on C_D or ρ or some of the gas surface interaction parameters. The analysis presented here on the other hand has employed very few assumptions in comparison. The assumption of symmetrical density variation about perigee is most subject to error. However, the possible error introduced by nonsymmetry is expected to be much less than the errors committed in past drag studies. The method presented here is far superior in terms of error than previous methods.

Improvement of the errors in the present analysis could be made by a more accurate treatment of aerodynamics and more accurate measurement of accelerations and angles of orientation. The accuracy of angle measurements was about $\pm 1^o$ while the force measurements were accurate to $\pm 5\%$ for the data used. The aerodynamic description of the OVI-15 could be improved by employing a more accurate expression for the variation in C_F with P_j. The polynomial approximation employed in the analysis could be improved or the exact expressions could be employed at the expense of computer tim

The results obtained for C_F as a function of angle of attack f OVI-15 is of special interest because of the many drag analysis wh have been performed on the satellite. For example, Champion, Marcos, and McIsaac, 1970; Marcos and Champion, 1972; Marcos, Champion, and Schweinfurth, 1971; have analyzed the accelerometer data of the OVI-15 to reveal a number of properties of the upper atmosphere. In these analyses, accelerometer data was used only when the satellite was broadside into the flow. This instantaneous attitude would correspond exactly to the 0^o or 180^o spin axis orientation of the OVI-15. At this attitude C_D is equal to C_F as given in Figure 5 and would have a value of 2.5396 for the case of infinite speed ratio and $\alpha_j = .64$ and $P_j = .44$. This value of C_D is 15.4% greater than the value of 2.2 which was employed in these analysis. Additional correction would have to be made if speed ratio effects were taken into account. These corrections would result in substantial decrease in densities reported using a C_D of 2.2.

OVI-15 orbital data has been analyzed by Ching 1971 and King-Hele and Walker 1969. The King-Hele and Walker analysis employed a constant C_D of 2.2 independent of the satellite orientation. The reference area used by King-Hele and Walker was midway between the maximum of $2.578 \pi r^2$ and the minimum of πr^2 shown in Figure 5. On the basis of the reference area employed by King-Hele and Walker, a C_D value of between 4.543 and 3.889 would correspond to the values of α_j and P_j found in the analysis.

The orbital decay analysis reported by Ching 1971 includes a factor to represent the changing aerodynamic drag properties of the OVI-15. The factor is based on the changing cross sectional area seen by the flow. The drag coefficient is considered constant while the reference area used in the analysis is changed by as much as 25%. A 25% correction factor is too large in view of the results of Figure 5 which show that the maximum change in $C_D \bar{A}$ would be 16%.

In addition, the effect these results have on past analysis of OVI-15 data in particular, the results indicate that the assumption of C_D used in most drag studies have been too low. The value of $C_D = 2.0$ or 2.2 which has been used for most past drag analysis is lower than could be expected for most shapes with $\alpha_j = .64$ and $P_j = .44$. A sphere for example would have a $C_{D_{sph}} = 2.352$ which is 7% higher than the 2.2 value often used. It should be noted, however, that the results obtained here are for one satellite surface and one atmospheric composition. It is expected that other surfaces and other compositions should change the values of α_j and P_j and result in a change in aerodynamic properties. More data must be collected before a firm value of α_j and P_j can be assigned to a given gas and surface combination. Future work should be directed towards this goal.

TABLE I

Orbit Number = 890 Time of Peak $\overline{F}$ = 672

I	F(i, 1)	GAMA(i, 1)	F(i, 2)	GAMA(i, 2)	TIME(i, 1)	TIME(i, 2)
1	317.5	137.75	320.0	134.8	647	697
2	307.0	139.50	306.0	133.25	622	722
3	291.5	141.75	291.0	132.25	597	747
4	270.0	142.50	268.0	130.0	572	772
5	251.0	144.0	244.0	129.0	547	797
6	225.0	145.75	214.0	127.5	522	822

Orbit Number = 893 Time of Peak $\overline{F}$ = 681

I	F(i, 1)	GAMA(i, 1)	F(i, 2)	GAMA(i, 2)	TIME(i, 1)	TIME(i, 2)
1	292.0	141.75	292.0	137.25	650	712
2	282.0	142.75	283.5	135.0	625	737
3	264.0	144.50	263.0	132.75	600	762
4	243.0	146.20	239.5	132.0	575	787
5	218.0	147.50	215.0	130.0	550	812
6	199.0	148.20	190.5	128.0	525	837

Orbit Number = 896 Time of Peak $\overline{F}$ = 675

I	F(i, 1)	GAMA(i, 1)	F(i, 2)	GAMA(i, 2)	TIME(i, 1)	TIME(i, 2)
1	306.0	133.5	302.5	129.5	650	700
2	302.0	135.0	291.0	127.7	625	725
3	289.0	136.25	275.0	126.75	600	750
4	270.5	137.5	256.0	125.0	575	775
5	250.0	138.0	231.0	123.0	550	800
6	226.5	141.0	200.5	120.5	525	825

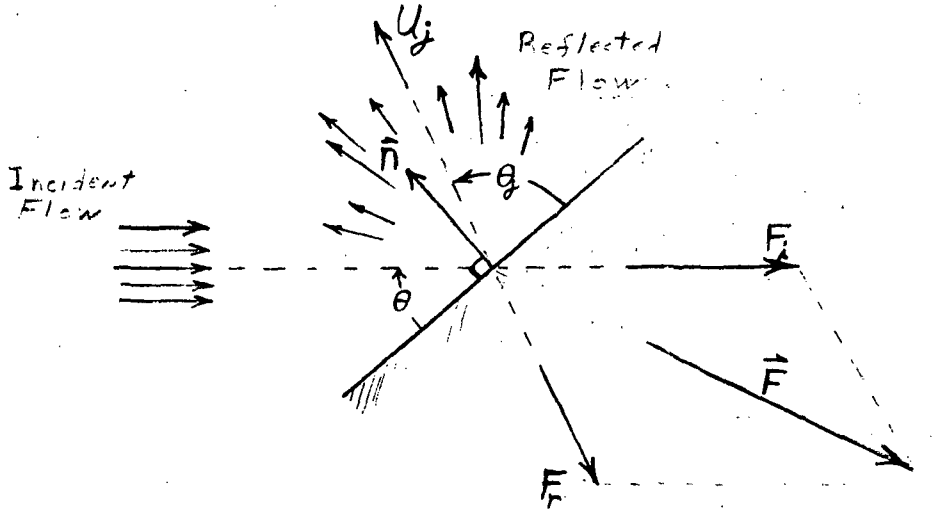


Figure 1. The gas surface interaction and the forces of the interaction.

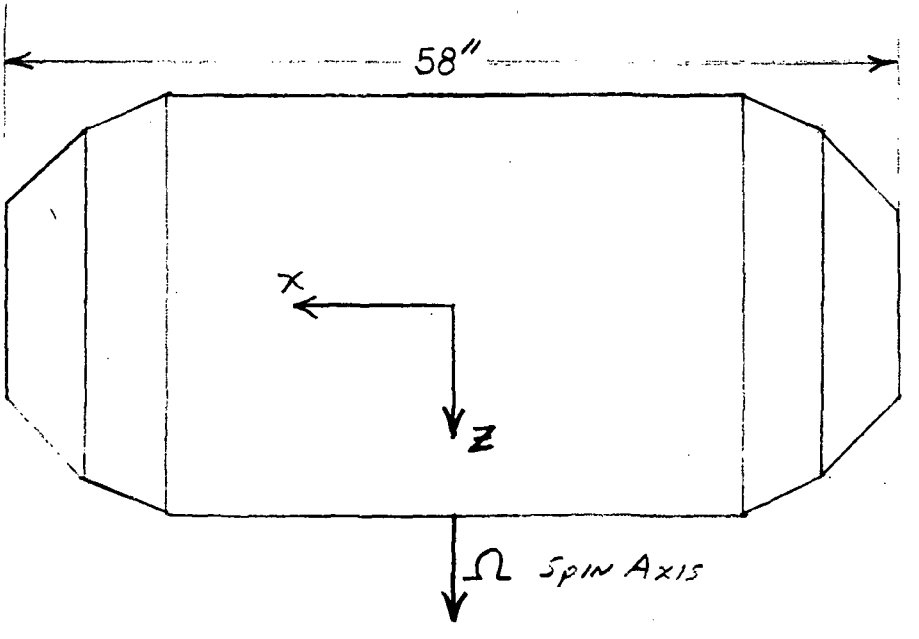


Figure 2. Configuration of the OVI-15 satellite.

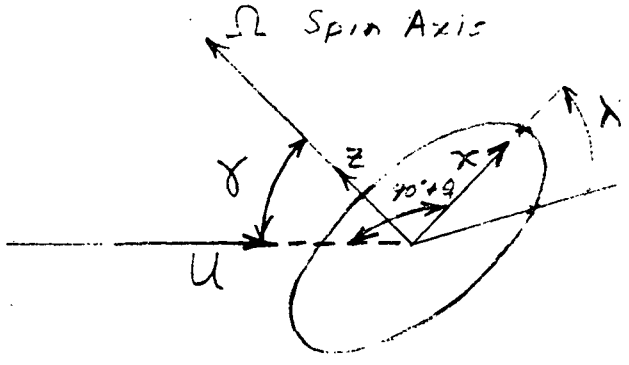


Figure 3. Orientation of OVI-15 with respect to flow velocity U.

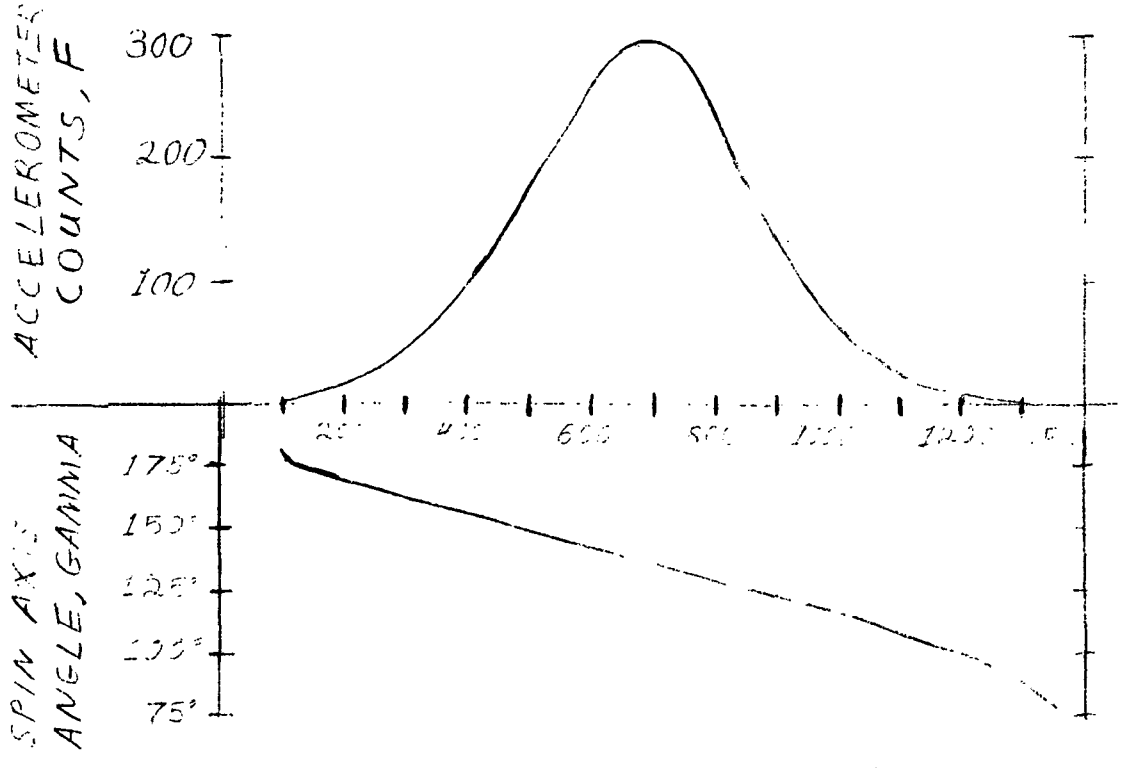


Figure 4. Accelerometer output and angle of orientation gamma for orbit number 893. (From Fess and Young, 1969).

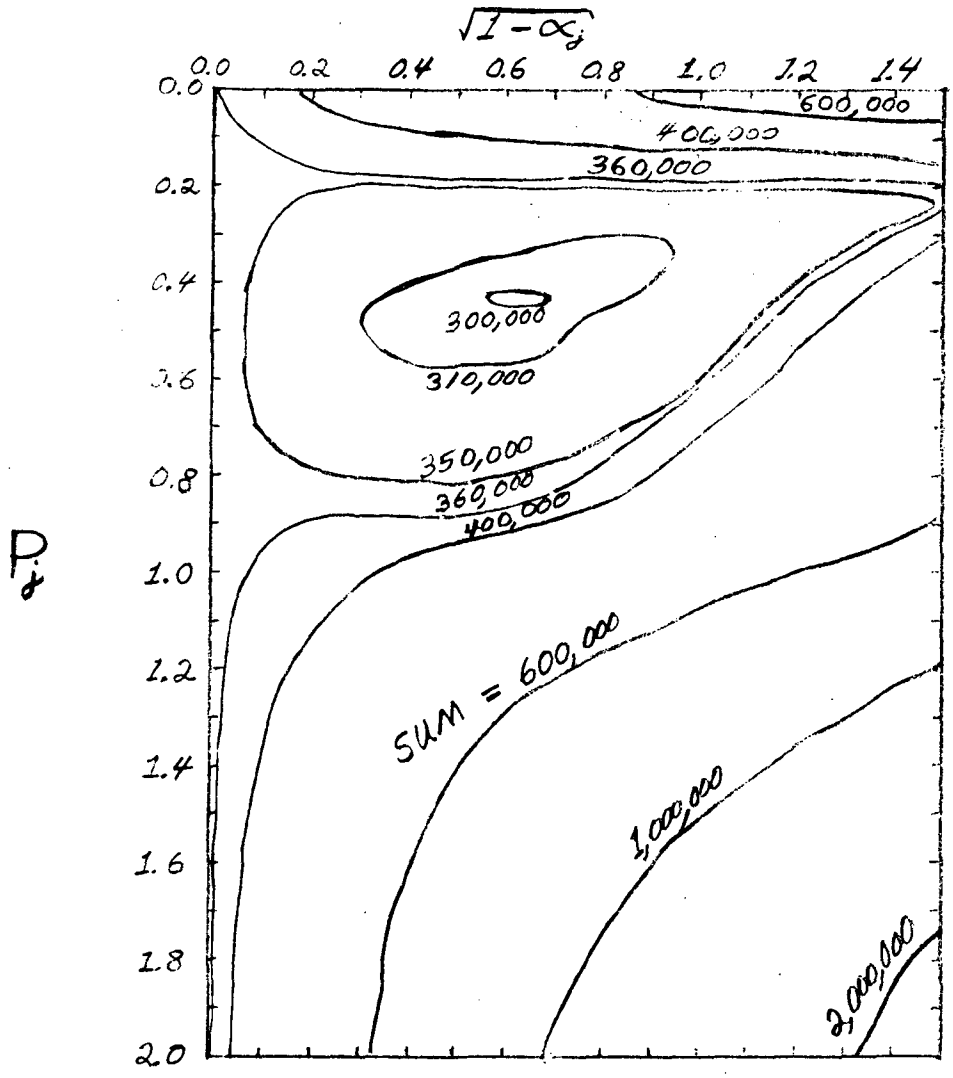


Figure 5. Constant values of the sum-of-the-squares-of-DEL_i over an acceptable range of α_j and P_j for data from orbits number 890, 893, and 896.

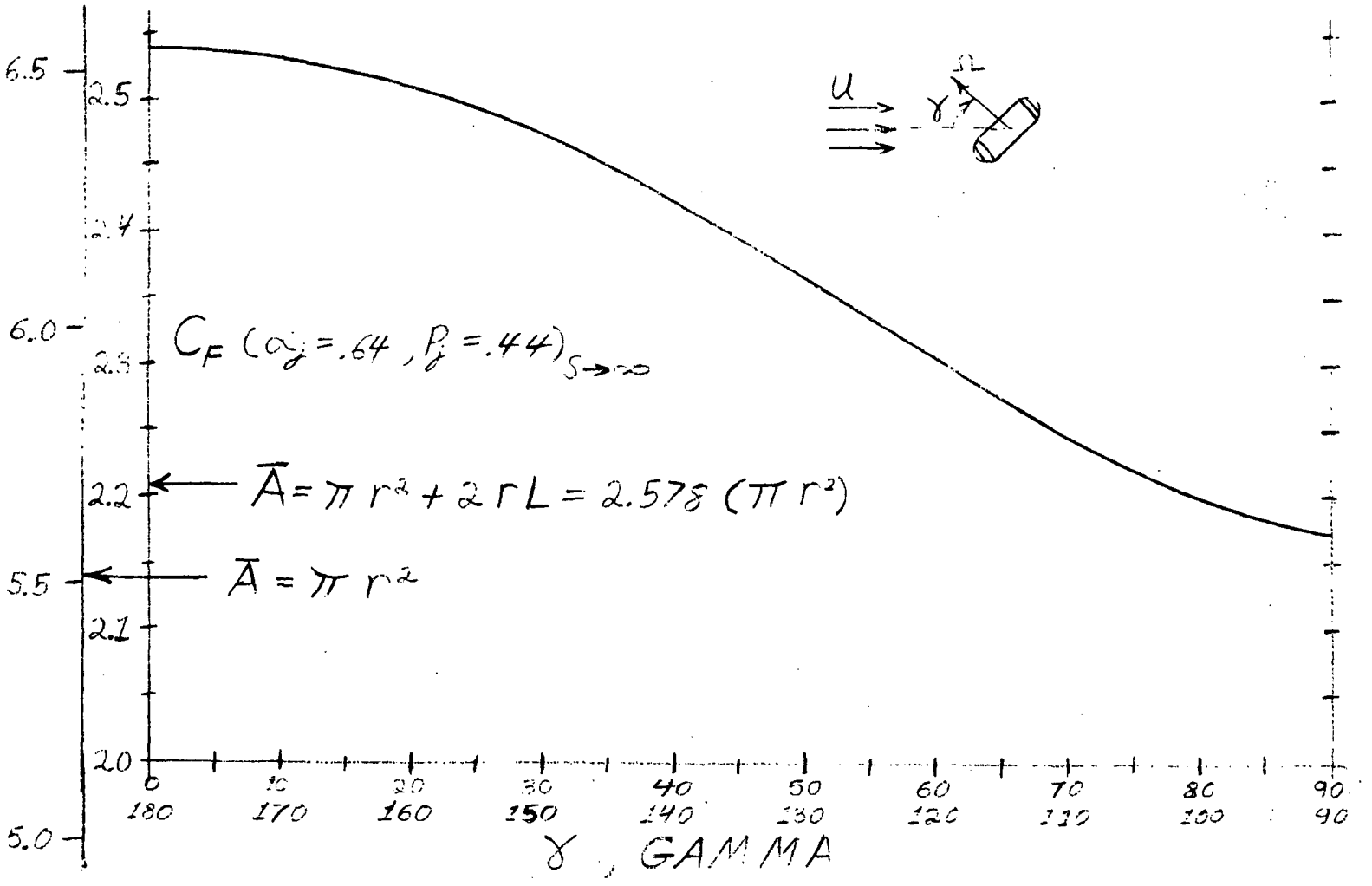


Figure 6. Force coefficient as a function of spin axis orientation of the OVI-15.

REFERENCES

Champion, K.S.W., F.A. Marcos, and J.P. McIsaac, 1970: Atmospheric Density Measurements by Research Satellite OVI-15. Space Research X, North-Holland Publishing Co., Amsterdam, 450-458.

Ching, B. K., 1971: Atmospheric Density and Rotation Below 195 km From a High Resolution Drag Analysis of the Satellite OVI-15 (1968-059A), Journal of Geophysical Research, Vol. 76, No. 1, 197-201.

Fess, W. A. and K. R. Young, 1969: OVI-15 (SPADES) Low G Accelerometer Data Processing. The Aerospace Corporation Report No. TOR-0066(5306)-7.

Jacchia, L. G., 1971: Revised Static Models of the Thermosphere and Exosphere with Empirical Temperature Profiles. Smithsonian Astrophysical Observatory, Special Report 332.

Karr, G. R., 1969: A Study of Effects of the Gas-Surface Interaction on Spinning Convex Bodies with Application to Satellite Experiments. Ph.D. Thesis, Coordinated Science Laboratory Report R-435, University of Illinois, Urbana, Illinois.

Karr, G. R., 1972: Satellite Aerodynamics as a Function of Atmospheric Properties. Paper presented at AIAA 5th Fluid and Plasma Dynamics Conference, Boston, Mass., AIAA Paper No. 72-659, 10 pages.

Karr, G. R. and R. E. Smith, 1972: Influence of Satellite Aerodynamics on Atmospheric Density Determination. Paper presented at International Conference on Aerospace and Aeronautical Meteorology, Washington, D. C., 6 pages.

Karr, G. R. and S. M. Yen, 1970: Aerodynamic Properties of Spinning Convex Satellites, Seventh Rarefied Gas Dynamic Symposium, (in press).

Marcos, F. A. and K. S. W. Champion, 1972: Gravity Waves Observed in High Latitud Neutral Density Profiles. Space Research XII, Akademie-Verlag, Berlin, 941-946.

Marcos, F. A., K. S. W. Champion, and R. A. Schweinfurth, 1971: More Accelerometer and Orbital Drag Results from SPADES (OVI-15) and Cannon Ball I (OVI-16) Satellites. Space Research XI, Akademie-Verlag, Berlin, 941-946.

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

LARGE SILVER HALIDE SINGLE CRYSTALS AS CHARGED PARTICLE TRACK DETECTORS

Prepared by:	John H. Kusmiss
Academic Rank:	Associate Professor
University:	Western Michigan University
Laboratory: (Division) (Branch)	Space Sciences Laboratory Nuclear and Plasma Physics Division Physical Analysis Branch
Research Counterpart:	Dr. Thomas A. Parnell
Date:	August 11, 1972
Contract Number:	NGT-01-003-045

LARGE SILVER HALIDE SINGLE CRYSTALS AS TRACK DETECTORS FOR HEAVY CHARGED PARTICLES

By

John H. Kusmiss

ABSTRACT

Large silver halide single crystals can be used to record the passage of heavy charged particles. The trajectory of the particle is made visible under a microscope by the accumulation of metallic silver at regions of the lattice damaged by the particle. This "decoration" of the particle track is accomplished by exposure of the crystal to light. The decoration of normally present lattice imperfections such as dislocations can be suppressed by the addition to the crystal of less than ten parts per million of a suitable polyvalent metal impurity. In fact, careful control of the concentrations of metallic impurities is absolutely necessary because they play important roles in the details of the track decoration process.

Large single crystals of the silver halides present certain immediate advantages over photographic emulsions as track detectors for heavy charged particles. The decoration of tracks in large single crystals requires only a few hours, and the distortion of tracks associated with the development of nuclear emulsions is not encountered. Because the mass density of a large silver halide single crystal is roughly twice that of a photographic emulsion, a 0.5 cm thick crystal corresponds to a 1.0 cm thick emulsion stack. Moreover, the possibility apparently exists of being able to decorate the path of a particle at the time it passes through an impurity-doped AgCl crystal, but not at later times, thus providing a means of using these crystals as "triggered" detectors. Another attractive possibility for these detectors is being able to vary the threshold for the rate of ionization necessary to produce a track; it has been claimed that this can be done by setting the level of impurity doping. The potential utility of silver halide single crystals as heavy ion detectors and dosimeters in space applications is obvious.

This report presents a brief review of the published work pertaining to silver halide single crystals as charged particle detectors and some comments as to the course which future work should take. An account of some preliminary attempts to grow thin single crystals of AgCl is given also, and suggestions for a more refined technique are offered.

INTRODUCTION

The silver halides tend to decompose with the absorption of light, the result being metallic silver and a gaseous halogen. This fact accounts in a crude way for the use of the silver halides in photography, although the exact details of the conversion of a microcrystal of silver halide into metallic silver make up a very complicated story in solid state physics. The literature of the photographic process is voluminous and will not concern us here except as it relates to the use of large (of the order of $1x1x0.5cm^3$) single crystals as detectors of charged particle tracks. For the sake of concreteness, and because most of the work has been done with it, we will refer for the most part to silver chloride.

It is of some historical interest that AgCl and AgBr were among the materials that received attention as "crystal counters" in the 1940's before the scintillation detector gained pre-eminence. The idea was to collect the charge pulse that resulted when a single crystal of an insulating solid was traversed by an ionizing particle while the crystal was maintained at the temperature of liquid nitrogen and subjected to an electric field of several kilovolts per centimeter (1). The present interest in silver chloride as a recorder of charged particle tracks dates from the pioneering work of Childs and Slifkin (2) in 1960. They demonstrated that charged particle tracks as well as lattice dislocations could be observed with a microscope throughout the interior of large crystals by a method of decoration combining the techniques of Hedges and Mitchell (3) and of Haynes and Shockley (4). In a later paper (5) Childs and Slifkin summarize four years of work on the decoration of particle tracks and dislocations in AgCl.

TRACK DECORATION IN AgCl

When light in the near ultra-violet is absorbed in the surface region of a thick AgCl crystal, it produces photoelectrons with a quantum efficiency close to unity. If the surface region is sufficiently strain-free, the electrons can be swept further into the interior of the crystal by the application of an electric field. At room temperature the conduction electrons in AgCl have a mobility of 50 cm^2/volt-sec and a lifetime of the order of microseconds. In that time an applied field of 10^4 volts/cm can move an electron about one centimeter. Two other facts are needed to explain the decoration process - first, that interstitial silver ions are the predominant type of point defect and are relatively mobile in AgCl at room temperature, and second, that photoelectrons can be trapped at a variety of sites in the interior of the crystal. The combination of trapped photoelectrons and migrating silver ions gives neutral silver atoms. Each trap is then free to capture a second photoelectron, and after

many repetitions of the cycle a speck of metallic silver of the order of a micron in size can be formed. In practice the decoration process must be pulsed because the dark conductivity of AgCl with several ppm of metallic impurity is so high that a static internal electric field would decay in about 100 μsec. Haynes and Shockley (4) devised a simple method of synchronizing the incident actinic light flashes with the applied electric field. The specimen is placed between two electrodes, of which the positive one is semi-transparent. A surface charge builds up because of ionic conductivity in the crystal, and the applied field is then discharged through a mercury vapor lamp, producing a light flash and leaving an internal field which sweeps electrons into the specimen. Typically, with an ultra-violet photon flux of 10^9 per cm^2 from each flash, a pulse repetition rate of 10^3 per second results in adequate decoration after one or two hours.

The addition of certain impurities, monovalent copper for example, extends the tail of the light absorption versus wavelength curve for AgCl, making possible the decoration of thin crystals without the use of an electric field. In that case, the cuprous ions are known to function effectively as hole traps (6).

ELECTRON TRAPPING SITES ALONG PARTICLE TRACKS

The decoration process depends in an essential way on the creation of electron trapping sites that remain long after the charged particle has moved through the lattice. The energy deposited in the form of ionization and vibrational excitation is dissipated relatively soon after the passage of the particle, and the number of displaced silver ions also returns fairly rapidly to its thermal equilibrium value. The chlorine ions or atoms displaced by the particle are possible candidates for electron trapping, since presumably they would be relatively immobile. However, it seems most likely that the majority of trapping sites are provided by the tangle of dislocations produced by the thermal spike associated with the passage of the particle (7). AgCl has a low thermal conductivity and is softer than lead, so that modest transient changes in temperature can easily lead to persistent plastic deformation localized along the particle track.

COMPARISON WITH OTHER TYPES OF TRACK DETECTORS

The decoration of tracks in large crystals of AgCl has been seen from the above to be a simple and rapid process; this represents an immediate advantage over photographic emulsion stacks and plastic dielectric detectors. The development of emulsion stacks introduces distortions in track geometry, while the etching procedures for tracks in plastic dielectrics are rather laborious. Tracks in AgCl are stably embedded in a rigid, transparent material which is insensitive to humidity. Since the mass density of AgCl in bulk is almost

twice that of a nuclear emulsion, the equivalent of a one-centimeter emulsion stack is afforded by a crystal which is only half a centimeter thick.

AgCl track detectors have a potential versatility that neither emulsions nor plastic dielectrics posses, namely the ability to be triggered. This possibility was experimentally realized by Breuer et al. (8) and more recent results have been reported by Schopper et al. (9). An additional feature also reported by Schopper et al. is the variation in the detection threshold by control of the doping level of cadmium. Thus it is possible to set the minimum rate of energy loss by a particle which will result in a track. Finally, AgCl detectors are apparently not sensitive to singly charged particles, such as relativistic protons, so that they can be used in experiments in which a large unwanted background of protons is present.

POSSIBLE USES OF SILVER CHLORIDE DETECTORS

Silver chloride track detectors show considerable promise for use in experiments to measure the heavy ion component of cosmic radiation, particularly in the presence of a flux of electrons and protons which is often many orders of magnitude more abundant. Because they can be either triggered or "turned on" for longer periods of time, and because their sensitivity threshold can be controlled by proper impurity doping, large AgCl single crystals offer great flexibility for use in space flight applications.

It has been estimated that the radiation hazard to humans in space from heavy charged particles is significant (10) AgCl track detectors could be used as heavy ion dosimeters for personnel on extended space missions.

IMPURITY EFFECTS IN AgCl DETECTORS

Metallic impurity concentrations are known to be of the utmost importance in determining the track detection characteristics of AgCl. In addition to direct effects which impurities may exert on the formation of photolytic silver along tracks and dislocations, they may also give rise to silver specks at random locations in the crystal, thus contributing to the "background." In crystals containing one ppm of either iron or copper, line dislocations can be decorated, but radiation tracks cannot. On the other hand, in crystals with 4 ppm lead, tracks can be decorated but no dislocations are observed. One explanation of this difference in behavior is given in reference (11). Although a few ppm are necessary for good track decoration, the solubility limit of about 8 ppm should not be exceeded because the excess lead precipitates out and increases the background.

Using 200-500 micron thick sheets of AgCl crystals, Schopper et al. (9) found that particles produce latent tracks which can later be decorated only if the crystal is illuminated with weak visible light during the time the particle passes through the crystal. The later decoration is accomplished by irradiation with ultra-violet light of wavelength less than 4100 Angstroms. To the best of the present writer's knowledge, no explanation of the mechanism of stabilizing latent tracks has yet appeared in print.

DIRECTIONS FOR FUTURE INVESTIGATIONS

Three goals for future work would be to (1) elucidate the exact nature of the damage to the crystal lattice which results in a decorated track; (2) explain the mechanism whereby latent tracks can be "fixed" in doped crystals by irradiation with visible light; and (3) determine quantitatively the relationship of track width and silver speck density versus particle energy loss.

APPENDIX

GROWING THIN SINGLE CRYSTALS OF AgCl

Herein is described briefly a crude attempt to grow large thin single crystals of AgCl. Starting with nominally pure precipitated AgCl powder, sandwiches of the molten material were made with quartz microscope slides in air by heating with a hotplate. The slides were separated by platinum wire 0.015 inch in diameter which was bent into various planar shapes having a constriction at one end and being open at the other end. After the AgCl solidified between the slides, the resulting sandwiches were placed in a horizontal quartz tube which was then evacuated. A heating coil of high resistance wire mounted on wheels fitted closely around the cylindrical tube. The movable heating coil was attached to a clock motor drive which pulled the coil horizontally at a constant speed of 5 cm/hr. This horizontal version of the Bridgman technique was used to grow polycrystalline specimens containing crystallites which had areas typically of the order of a few mm^2. Since AgCl and quartz have quite different coefficients of expansion, the sandwiched AgCl is under strain, and the consequent dichroism causes individual crystallites to show up when viewed between crossed polarizing filters.

Detailed procedures for growing both pure and doped thin single crystals of the alkali halides are given in references (12) and (13). In the present writer's opinion the best way to get very pure starting material for crystal growing would be to zone refine nominally pure material. According to Moser et al. (14), the passage of 70 zones at a rate of 3 inches/hr through an AgCl sample in a chlorine atmosphere

results in optimum distributions for Cu, Pb, Ni, and Fe. Improvements of the crude procedure described above would include: (1) working completely in red light rather than weak white light; (2) never allowing molten AgCl to come in contact with oxygen; and (3) eliminating contact between molten AgCl and platinum, since there are indications of reaction with platinum, according to reference (14).

ACKNOWLEDGEMENTS

I should like to thank Dr. Thomas Parnell (MSFC) for suggesting the topic covered in this report and for his continuing interest and collaboration. I am grateful to the various members of his group for their cooperation and encouragement, in particular Dr. Peter Eby for interesting conversations.

I am indebted to Dr. Hans Walter for the use of his laboratory facilities at UAH, and for helpful discussions.

Finally, I should like to thank Charles Childs of the Materials Research Center of the University of North Carolina for providing some experimental supplies and for informative discussions.

REFERENCES

1. R. Hofstadter, Nucleonics, April 1949, p. 2; Nucleonics, May 1949, p. 29.

2. C. Childs and L. Slifkin, Phys. Rev. Letters 5, 502 (1960).

3. J. M. Hedges and J. W. Mitchell, Phil. Mag. 44, 223 (1953a).

4. J. R. Haynes and W. Shockley, Phys. Rev. 82, 935 (1951).

5. C. B. Childs and L. M. Slifkin, Brit. J. Appl. Phys. 16, 771 (1965).

6. F. Moser, N. R. Nail, and F. Urbach, J. Phys. Chem. Solids 9, 217 (1959).

7. F. Seitz and J. S. Koehler, Solid State Physics, Vol. 2, edited by F. Seitz and D. Turnbull, (Academic Press), p. 377.

8. K. Breuer, G. Haase, and E. Schopper, Brit. J. Appl. Phys. 18, 1824 (1967).

9. E. Schopper, G. Haase, G. Henig, J.-U. Schott, and F. Zoergiebel, Volume 6 of the Conference Papers of the 12th International Conference on Cosmic Rays held at Hobart, Tasmania in August 1971.

10. P. Todd: "Biological Effects of Heavy Ions," Second Symposium on Protection Against Radiation in Space, NASA SP-71, 1965.

11. C. B. Childs and T. A. Parnell, Proceedings of the National Symposium on Natural and Manmade Radiation in Space, held at Las Vegas, Nevada in March 1971, NASA TMX-2440.

12. C. Berry, W. West, and F. Moser, Chapter 12 of The Art and Science of Growing Crystals, edited by J. J. Gilman (Wiley).

13. F. Zoergiebel, G. Haase, G. Henig, E. Schopper, and J.-U. Schott Zeitschrift fuer angewandte Physik 30, 316 (1970).

14. F. Moser, D. C. Burnham, and H. H. Tippins, Journal of Applied Physics 32, 48 (1961).

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

DETERMINATION OF OPTICAL CONSTANTS OF In Bi FROM REFLECTANCE MEASUREMENTS

Prepared by:	Frank E. Martin
Academic Rank:	Professor
University:	Central Missouri State University
Laboratory: (Division) (Branch)	Space Sciences Laboratory Space Thermophysics Division Thermal Environment Physics Branch
Research Counterparts:	Roger L. Kroes Roger C. Linton
Date:	August 11, 1972
Contract No:	NGT-01-003-045

ACKNOWLEDGMENTS

The author is greatly indebted to Mr. Roger C. Linton and Dr. Roger L. Kroes for proposing the investigation, and for guidance and aid as it progressed. He is also indebted to Dr. H. U. Walter, of the University of Alabama Research Institute at Huntsville for growing the InBi single crystal used in the present work, Mr. Robert Andrews for technical assistance in the laboratory, Mr. William Watson for preparing the figures for publication, and Miss Virginia Diane Bush for typing the text. Support by NASA, ASEE, Marshall Space Flight Center, Auburn University and the University of Alabama, which together generously granted him a second summer faculty research fellowship and provided the necessary equipment and services for the work to be carried out, is gratefully acknowledged.

DETERMINATION OF OPTICAL CONSTANTS OF In Bi FROM REFLECTANCE MEASUREMENTS

By

Frank E. Martin

ABSTRACT

Electromagnetic theory makes possible the determination of optical constants of a substance as functions of wavelength from reflectance data measured for different angles of incidence. These constants are the real index of refraction n and the absorption coefficient k, and they are related by

$$N = n + ik$$

where N is the complex index of refraction.

Reflectance measurements were made on samples of the intermetallic compound In Bi. A fall in reflectance was noted from values obtained for a freshly cleaved sample to values obtained after periods of exposure of the surface to air. The decrease in reflectance with increased exposure time is attributed to formation of a surface layer insoluble in Freon or ethyl alcohol; it is therefore assumed to be a tarnish layer. Computer processing was employed to calculate optical constants form reflectance data.

INTRODUCTION

The purpose of this report is to provide background material for determining optical constants, the real index of refraction n and the absorption coefficient k, from reflectance measurements, to describe the experimental set-up used in making the measurements, and to discuss results obtained with the intermetallic compound In Bi. An intermetallic compound is formed from metallic elements that combine in a definite stoichiometric ratio. Its physical behavior is metallic.

Maxwell's equations may be used to determine these optical constants for a given wavelength from reflectance data taken at that wavelength[1-6]. They comprise the real and imaginary parts of the complex index of refraction

$$N = n + ik \tag{1}$$

which determines the optical behavior of a homogeneous isotropic medium.

Measurements were made with ultraviolet (UV) light, with the aid of a vacuum monochromator, since at wavelengths less than 200 nm, the vacuum ultraviolet (VUV) spectral range, atmospheric absorption cannot be tolerated. The present writer's 1971 NASA research fellowship report contains a brief historical account[6] of work with reflectance measurements in UV and VUV ranges[6]. Madden[7] gives further historical details.

THEORY

The real index of refraction of a medium is defined by

$$n = \frac{c}{v} \tag{2}$$

c is the speed of light in a vacuum and v is the speed in the medium. The absorption coefficient k is defined by

$$I = I_0 e^{-\frac{4\pi k x}{\lambda}} \tag{3}$$

I_0 is the intensity of incident radiation of wavelength and I is the intensity of radiation that has traveled a distance x into the medium.

In the literature and in the writer's earlier report[6] derivations of the Fresnel equation for reflectance are given in detail. In the report a geometric approach to the derivations minimized the use of vector formalism. A briefer statement will be given here, leading to forms of the Fresnel relations which will be combined with Snell's law to yield reflectance equations related to and useful in the present work.

Maxwell's equations for a homogeneous medium at rest, of dielectric constant

$$\kappa_e = \frac{\epsilon}{\epsilon_0}$$

and magnetic permeability

$$\kappa_m = \frac{\mu}{\mu_0}$$

and having no charge and only ohmic current may be written

$$\nabla \times \underline{E} = -\mu \frac{\partial \underline{H}}{\partial t}, \quad \nabla \cdot \underline{H} = 0$$
$$\nabla \times \underline{H} = \sigma \underline{E} + \epsilon \frac{\partial \underline{E}}{\partial t}, \quad \nabla \cdot \underline{E} = 0 \qquad (4)$$

where σ is conductivity in the Ohm's law expression

$$\underline{J} = \sigma \underline{E}$$

Plane wave solutions of these equations for the electric and magnetic field intensity vectors may be written

$$\underline{E} = \underline{E}_0 e^{i(\omega t - \underline{K} \cdot \underline{r})}$$
$$\underline{H} = \frac{\underline{K} \times \underline{E}}{\omega \mu} \qquad (5)$$

where $\underline{K}$ is the wave propagation vector, having the directions of a ray i.e. normal to the plane wave fronts, and magnitude .
At a plane interface between two homogenous isotropic media(with a vacuum considered as a limiting case), reflection and refraction occur as shown in Figue 1. The interface is the XY-plane, with incident and reflected rays above it in a medium of permeability , dielectric constant , index of refraction n and absorption constant k; and with the refracted ray in the lower medium of constants ', ', n', and k'. The propagation vectors $\underline{K}$, $\underline{K}'$, $\underline{K}''$ for the three rays lie in the plane of incidence, the XZ-plane, as shown, with the angles of incidence and reflection having the value i, and the angle of refraction the value r. The direction of the unit normal vector is as shown.

For, static fields it is easily shown by Stokes' theorem that the tangential components of $\underline{E}$ and $\underline{H}$ must be continuous at the boundary. Moreover, for a time-varying field of form

$$\underline{H} = \underline{H}_0 e^{i(\omega t - \underline{K} \cdot \underline{r})}$$

it must be true for tangential components at the boundary not only that

$$H_0 + H'' = H'_0$$

but also that

$$\underline{H} + \underline{H''} = \underline{H'} \tag{6}$$

because for continuity of the fields the exponentials must be the same over the boundary. When an equation such as (6) is written for tangential components at a boundary, it may be considered to apply either to the amplitudes of the vector fields or to the time-varying fields themselves.

In general, the direction of $\underline{E}$ or $\underline{H}$ is arbitrary, but at a boundary plane

(1) the tangential components of the electric vectors are continuous, or

$$\hat{n} \times (\underline{E} + \underline{E''}) = \hat{n} \times \underline{E} \tag{7}$$

(2) the tangential components of the magnetic vectors are continuous, or

$$\hat{n} \times (\underline{H} + \underline{H''}) = \hat{n}\, \underline{H'} \tag{8}$$

from which

$$\hat{n} \times \frac{\underline{k} \times \underline{E} + \underline{k''} \times \underline{E''}}{\mu} = \hat{n} \times \frac{\underline{k'} \times \underline{E'}}{\mu'}$$

since ω is the same for both media. Use of the vector identity

$$\underline{A} \times (\underline{B} \times \underline{C}) = \underline{B}(\underline{A} \cdot \underline{C}) - \underline{C}(\underline{A} \cdot \underline{B})$$

with the final term in the left hand member of the equation gives

$$\hat{n} \times (\underline{k} \times \underline{E}) = \underline{k}(\hat{n} \cdot \underline{E}) - \underline{E}(\hat{n} \cdot \underline{k})$$

where the final term in the right hand member vanishes because $\hat{n}$ and E are orthogonal. Similar relations hold for other terms, with the result that

$$-\frac{\underline{E}(\hat{n} \cdot \underline{k})}{\mu} - \frac{\underline{E''}(\hat{n} \cdot \underline{k''})}{\mu} = -\frac{\underline{E'}(\hat{n} \cdot \underline{k'})}{\mu'} \tag{9}$$

$\underline{E}$ may always be written as the sum of two orthogonal components. Two cases of polarization are therfore considered, for which

(1) the electric intensity vectors are perpendicular to the plane of incidence:

Then (7) requires

$$E_y + E''_y = E'_y \tag{10}$$

and (9) requires for non-magnetic materials

$$-KE_y \cos i + KE''_y \cos i = -Kn E'_y \cos r$$

since

$$K = K'' \text{ and } K' = K\frac{v}{v'} = n$$

where v and v' are velocities in the two media, and n is the relative index of refraction.

Then

$$(E_y - E''_y)\cos i = n E'_y \cos r \tag{11}$$

Eliminating E'_y from equations (10) and (11) gives one of the Fresnel wave reflectance equations

$$\frac{E''_y}{E_y} = -\frac{n\cos r - \cos i}{n\cos r + \cos i}$$

Eliminating r by use of Snell's law of refraction

$$n = \frac{\sin i}{\sin r}$$

gives

$$\frac{E''_y}{E_y} = -\frac{\sqrt{n^2 - \sin^2 i} - \cos i}{\sqrt{n^2 + \cos^2 i} + \cos i}$$

For an absorbing medium(as for a metal with its large conductivity), let

$$n \longrightarrow N = n + ik$$

The reflectance in the sense of the present work is the intensity reflectance, in this case for the perpendicular or senkrecht light, which is defind by

$$R_s = \left| \frac{E''_y}{E_y} \right|^2$$

which becomes in terms of the above relations

$$R_s = \left| \frac{\left[(n+ik)^2 - \sin^2 i\right]^{\frac{1}{2}} - \cos i}{\left[(n+ik)^2 - \sin^2 i\right]^{\frac{1}{2}} + \cos i} \right|^2 \tag{12}$$

(2) the electric intensity vectors are parallel to the plane of incidence: A similar analysis leads to an expression for the parallel reflectance

$$R_p = \left| \frac{\left[(n+ik)^2 - \sin^2 i\right]^{\frac{1}{2}} - (n+ik)^2 \cos i}{\left[(n+ik)^2 - \sin^2 i\right]^{\frac{1}{2}} + (n+ik)^2 \cos i} \right|^2 \tag{13}$$

For unpolarized light, such as that from the normal incidence grating monochromator in the present work, the reflectance is given by

$$R = \frac{1}{2} \left(R_s + R_p \right) \tag{14}$$

An expression is also available for the reflectance of partially plane-polarized light[5] if the intensities of the parallel and senkrecht polarized incident light are known.

The above relations give

$$R = R(n, k, i)$$

Calculation of reflectances from this equation is very time-consuming. In determining n and k experimentally measurements of reflectance are made for two angles of incidence, if the polarization of this incident light is known, yielding

$$R_1 = R(n, k, i_1) \tag{15}$$

$$R_2 = R(n, k, i_2)$$

These simultaneous equations may not be solved directly for n and k as simple functions. Graphical methods for finding n and k were developed by Tousey[8]. Computer methods are available[9,10] for calculating R from n, k, and i, and for determining n and k from the measured values of R and i.

EXPERIMENTAL PROCEDURE

Figure 2 shows a block diagram of the equipment. The Hinteregger hydrogen-discharge light source was operated in the window-less mode. A McPherson 225 vacuum monochromator was used chiefly in the range 100 to 360 mm. The grating was blazed for maximum efficiency in the VUV.

A McPherson dual-beam reflectometer was used as a single-beam instrument. Use of this instrument in the dual-beam mode is discussed by Linton[9]. The effect of using it as a single beam instrument was to allow monochromatic light reflected from the mirror to fall either directly on a sodium salicylate phosphor at position 2 if the sample was removed from the light path; or on the same phosphor after being reflected from a sample in the light path. The sample was moved out of the light path without removal from the reflectometer. Light from the phosphor was passed though a totally reflecting lucite rod to a photomultiplier tube. The reflectance for a given wavelength and angle of incidence was found as the ratio of the intensity measured with the phosphor in position 1 to the intensity for position 2.

The monochromator and reflectometer were evacuated by a fore pump and duffusion pump to a vacuum of 10^{-4} n/m^2. Differential pumping was employed to maintain proper pressure in the open hydrogen capillary tube of the Hinteregger lamp, and to minimize hydrogen entrance into the monochromator. An Ultek sorption-pump system was used to pump down the reflectometer after it was opened for sample replacement or adjustment.

Reflectance data were taken on two samples of single-crystal InBi, grown in the physics laboratory of the University of Alabama at Huntsville. Sample 1 was freshly cleaved about May 1, and had been exposed to air for two months when the first reflectance data were taken with it. Sample 2 was freshly cleaved just before a sequence of reflectance measurments covering a five week period was begun with it. Data were taken over a range of wavelengths at two different angles of incidence. In principle it makes no difference what values of i_1 and i_2 in equation (15) are used. Actually $(i_2 - i_1)$ should be large enough to afford well-spread values for R_1 and R_2. In this work i_1 was 20° and i_2 was 70°.

Reflectance data processed on the Rand Univac 1108 computer system yielded values of n and k. A Fortran V program of iterative nature was used[10]. The program proposed, for each wavelength used, a sequence of n and k values for which reflectances are calculated. Comparison of calculated and measured reflectance is made until a satisfractory match is obtained for a given n, k pair.

EXPERIMENTAL RESULTS

Figure 3 shows the reflectances versus wavelength, for angles of incidence of 20^{o} and 70^{o}, for sample 1 of InBi, for which the surface was cleaved about May 1, 1972, with the reflectance data taken after the sample had been exposed to air for some eight weeks.

Three runs were made, and the data for both curves are composite data. Separate points indicate the reflectance data for some wavelengths, for which results are drawn from single runs; and bars show the spread of data for wavelengths for which the data were drawn from more than one run.

Figure 4 shows the reflectances for the same sample after an additional five weeks of exposure to air. These data are drawn from a single run. Within the limits of experimental error and within the limitations imposed by the fact that one of them is composite, the data sets for Figures 3 and 4 show practically no differences.

Figure 5 shows reflectances versus wavelength for angle of incidence 20^{o}, for sample 2 for various exposure times to air after the surface was cleaved on June 28. The uppermost curve comprises data from three runs. The sample was cleaved on June 28 and immediately placed in the reflectometer, which was evacuated until the three runs had been completed. Figure 6 shows reflectances for angle of incidence 70^{o} for the same sample and air-exposure times.

Figure 7 shows the optical constants n and k versus wavelength, for sample 1, calculated from the reflectance data of Figure 3. Because the reflectance of Figure 4 is parctically the same as that for Figure 3, the values of n and k in Figure 7 may be considered those for both sets of reflectance data.

Figure 8, 9 and 10 show the optical constants versus wavelength for sample 2, calculated for the various air-exposure times after it was cleaved.

DISCUSSION

Results obtained with sample 2 will be considered first, since differences in reflectance and in n and k were noted with this sample for increased air-exposure times. The greatest reflectances, both for incidence angles of 20° and 70° were noted for miminal air-exposure time. The 20° reflectance starts at about 10 precent at wavelength 110 nm and gradually increases until a 250 nm wavelength is reached, and thereafter remains nearly constant for increasing wavelength. The 70° reflectance starts at about 40%, increases until a 200 nm wavelength is reached, and then remains nearly constant for increasing wavelength. From these data it is clear that most of the light energy is not reflected at low wavelengths: it must be either absorbed or scattered. At longer wavelengths most of the light energy is reflected, with a correspondingly small amount being absorbed or scattered.

The effect of an air-exposure time of two weeks was to reduce both the 20° and 70° reflectance. Immediately after these effects were noted. the sample surface was washed in Freon and ethyl alcohol to remove any grease or other soluble foreign contaminant that might have deposited on the surface during the period of exposure to air. A check-run was made which indicated no change of consequence in the reflectances. The effect of an additional exposure-time of three weeks was to reduce the reflectances still futher. The decrease in reflectance with increased exposure time is attributed to formation of a surface layer insoluble in Freon or ethyl alcohol; it is therefore assumed to be a tarnish layer.

For sample 1, the first reflectance data were obtained after an air-exposure time of eight weeks. Both the 20° and the 70° reflectances are lower than those for the freshly cleaved sample 2, with the 20° reflectances markedly lower at short wavelengths. The 20° reflectances of sample 1 for the eight-week air exposure-time are nearly the same as those obtained with sample 2 after five weeks of exposure time; the 70° reflectances, for sample 1 are a little higher(about 4%)than those for sample 2 after the five week, period. A possible reason for this state of affairs will be given later.

Exposure of sample 1 to air for an additional five-week period failed to lower the reflectances. Accordingly, it is felt that specimen 1 had reached the maximum surface degradation at the end of eight weeks of exposure to air, and that sample 2 had reached essentially this same stage of surface degradation at the end of five-weeks of exposure.

Reflectances for the incidence angle of 70° should exceed those for the incidence angle of 20°. This was found true in the data taken for sample 1, which had large length and width. But for sample 2, a narrower sample than 1, it was noted that at longer wavelengths 20° reflectances sometimes slightly exceeded the 70° reflectances. Experimental errors could contribute in part to this state of affairs, but the deviation though small was systematic. It is considered possible that all of the light beam fell on both samples for the 20° angle of incidence(nearly normal incidence), but that some of the light beam escaped past the edges of the narrower sample 2 for the 70° angle of incidence(nearly grazing incidence)because of the increase in cross-sectional area of the beam for a section that is not normal to the direction of propagation. At longer wavelengths there is less difference between 20° and 70° reflectances, so that loss of reflectance due to escaped light would reveal itself more clearly at these wavelengths. Also there would be more spreading of the beam at longer wavelengths so that more light would tend to escape past the sides of a narrow sample. The escaped light hypothesis could also account for the slightly lower 70° reflectance for sample 2 after 5 weeks air-expasure time as compared with that of sample 1 for eight-weeks(or thirteen-weeks) exposure.

This hypothesis tends to find confirmation in that

(1) the 70° reflectance for the broader sample 1 always exceeds the 20° reflectance

(2) if sample 2 is rotated 90° in the reflectometer sample-holder though both 20° and 70° reflectances are slightly reduced because of escaped light past the lower edge of the narrow sample, the values of the 70° reflectance always exceed the 20° values, as should be the case if there were no lateral light escape.

For sample 2 as cleaved, the n values show first a slight decrease to a value of about 0.70, then a slight increase and finally a sharp rise with increasing wavelength. The k values are low at low wavelength, corresponding to a tendency toward transparency in metals for low wavelength, then increasing almost linearly, but apparently at a slightly greater rate, with increasing wavelength. More time would be needed for analysis before trying to account for the shapes of the curves.

The significance of the "n" and "k" values calculated for sample 1 and for sample 2 for longer elapsed periods after cleaving is in doubt. The computations for n and k were based on the assumption of a homogeneous isotropic crystal. If the sample changes with elapsed time turn out to affect it uniformly throughout its volume, then the values obtained should be true n and k values, which should afford insight into the changes taking place. If, as appears more probable,

a composite sample with a surface tarnish layer is obtained, the only true n and k values are those for sample 2 as cleaved. In this case the divergence of the calculated n and k values from those for sample 2 as cleaved may afford some measure of the practical departure from the assumption of a homogeneous isotropic crystal. This divergence is greatest for sample 1, which had the longest air-exposure periods, as might be expected if the surface-tarnish theory is correct.

RECOMMENDATIONS

Attempt can be made to understand the shapes of the n and k curves for the as cleaved sample 2 in terms of the classical or quantum theories of anomalous dispersion. Possibly the rise in the k curve (and perhaps also of the n curve) can be understood in terms of approach to a resonant frequency of bound electrons with increasing wavelength (decreasing frequency). A second computer program is available that can calculate the optical constants of a uniform film of constant thickness from reflectance data at different wavelengths, if the thickness of the film and the optical constants of the substrate are known. Ellipsometric determination of film thickness could be sought, and/or the program can be used to try to guess a thickness of the film. If time is available, consideration can be given to polarization effects due to reflection of light from the reflectometer mirror (needed for operation in the dual-beam mode) before reflection from the sample. The effect appears largest at the lowest end of the wavelength range in this investigation, and then not so large that valuable information cannot be obtained without compensation for polarization.

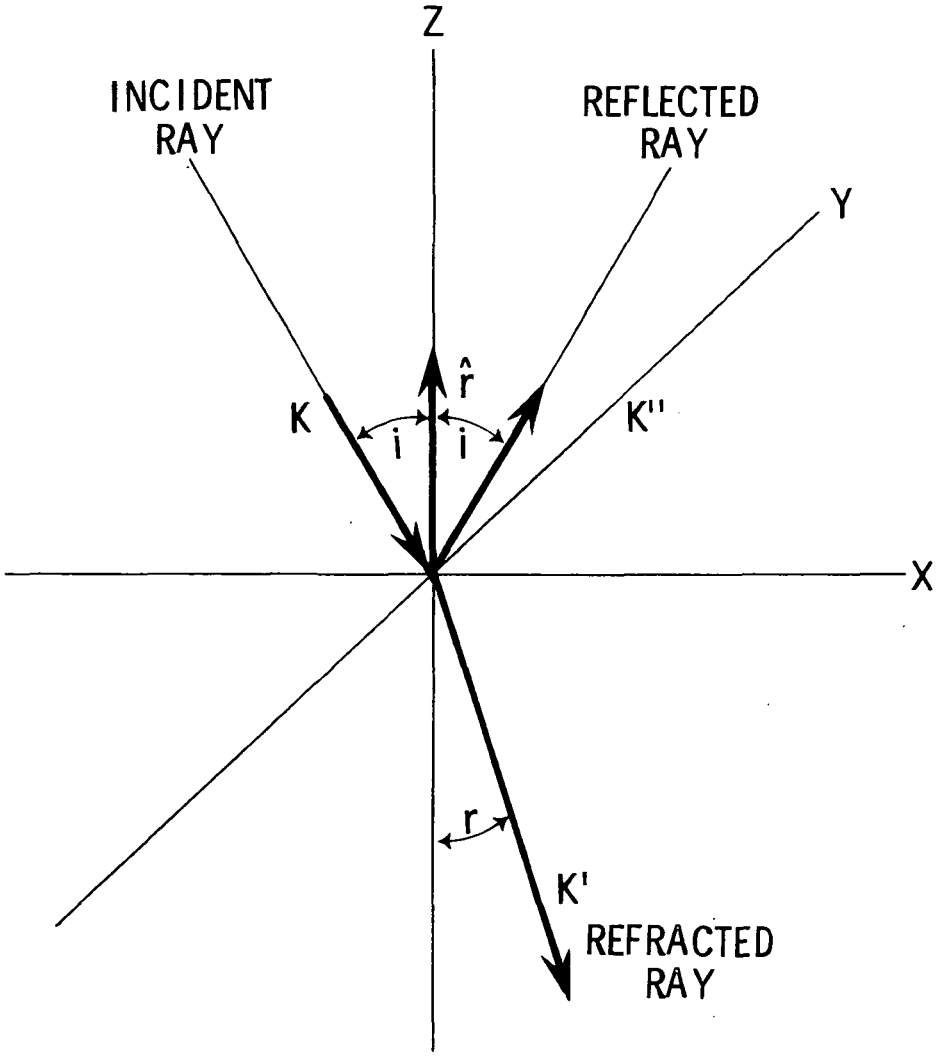


Figure 1. Reflection and refraction at a plane surface.

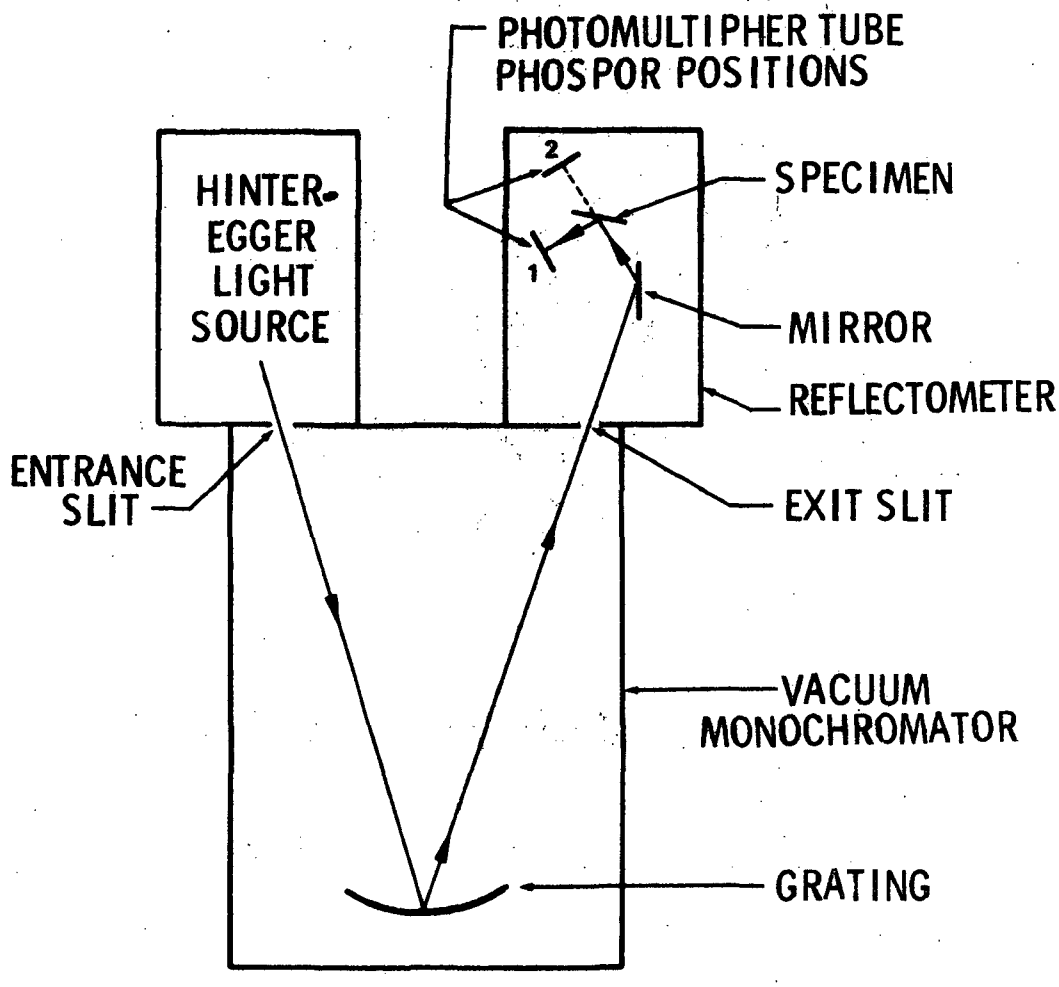


Figure 2. Block diagram of equipment.

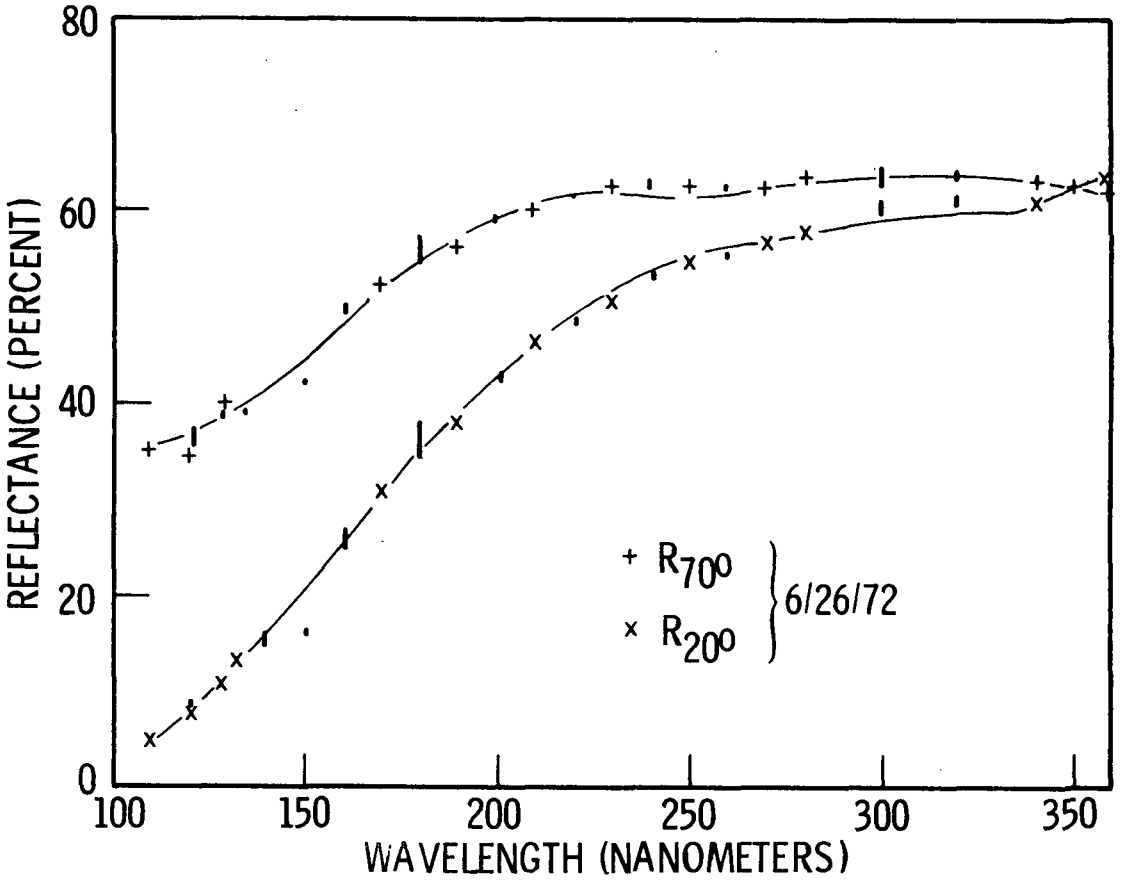


Figure 3. Reflectance of InBi versus wavelength, sample 1, surface exposed to air for eight weeks after it was cleaved.

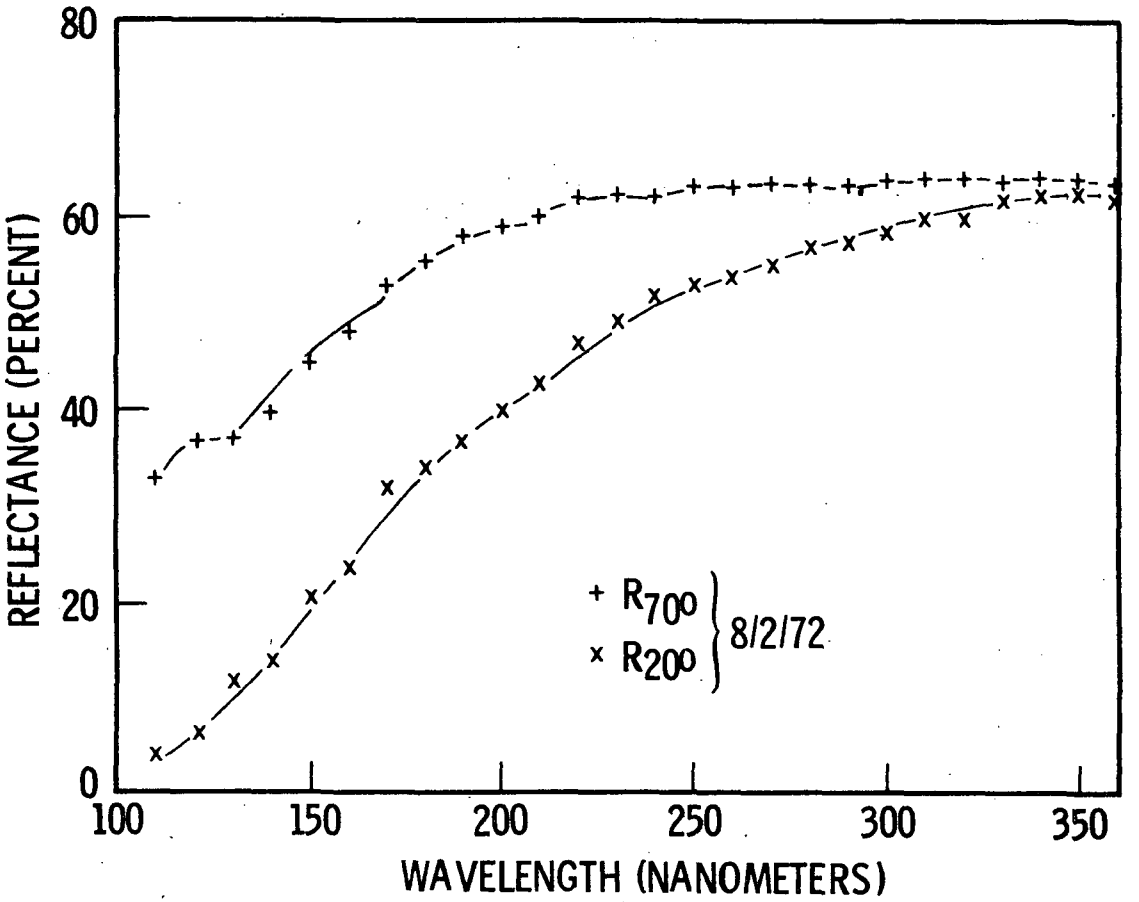


Figure 4. Reflectance of InBi versus wavelength, sample 1, surface exposed to air for thirteen weeks after it was cleaved.

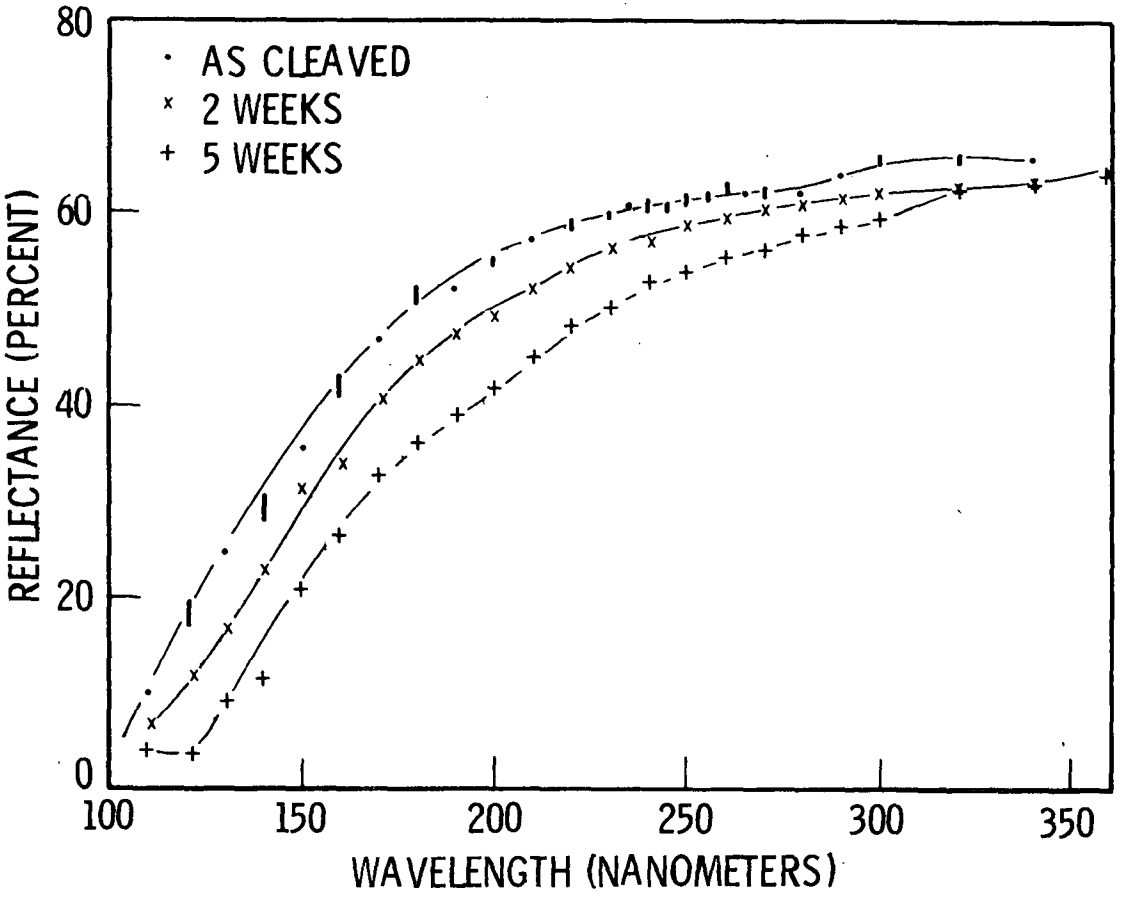


Figure 5. Reflectance of InBi versus wavelength, sample 2, angle of incidence 20°, various air-exposure times after the surface was cleaved.

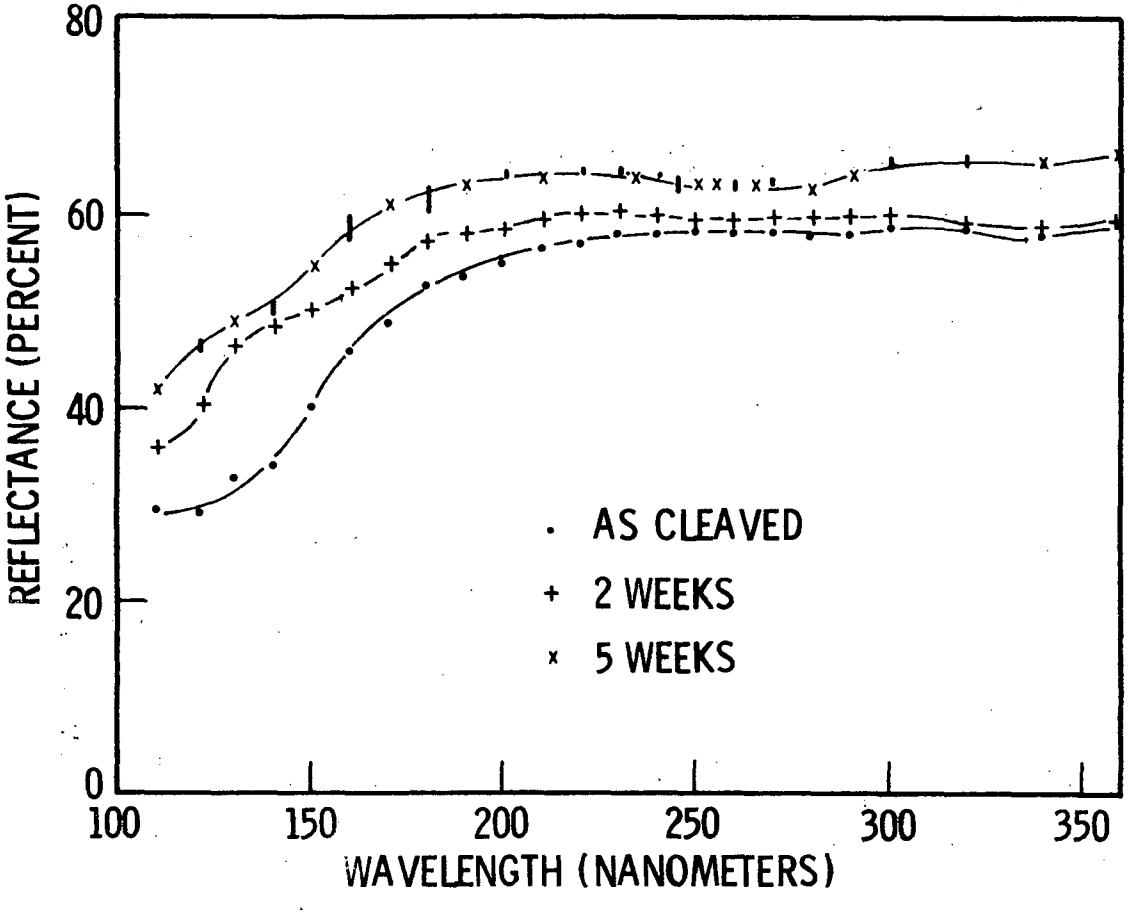


Figure 6. Reflectance of InBi versus wavelength, sample 2, angle of incidence 70°, various air-exposure times after the surface was cleaved.

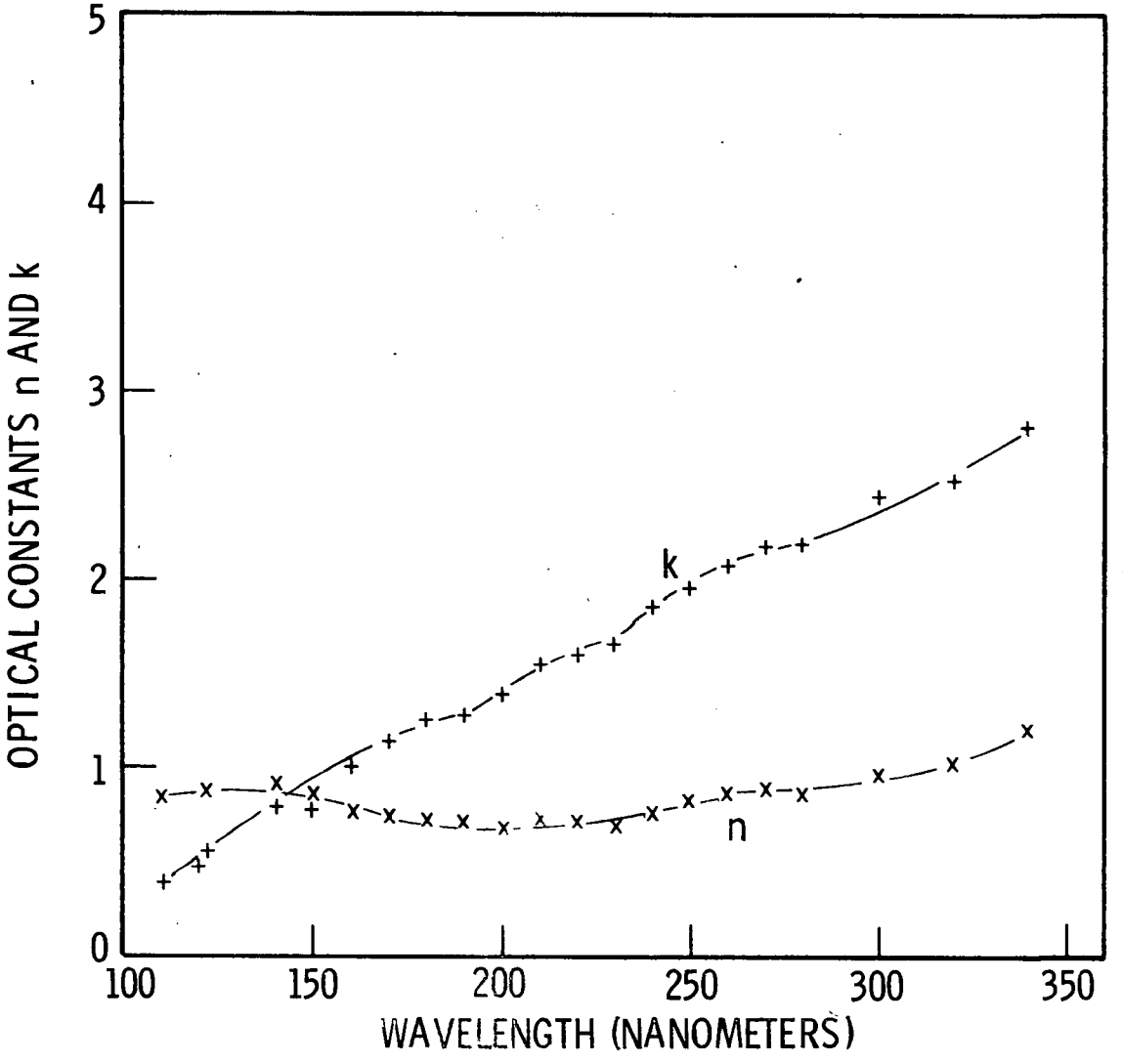


Figure 7. Optical constan of InBi versus wavelength, sample 1.

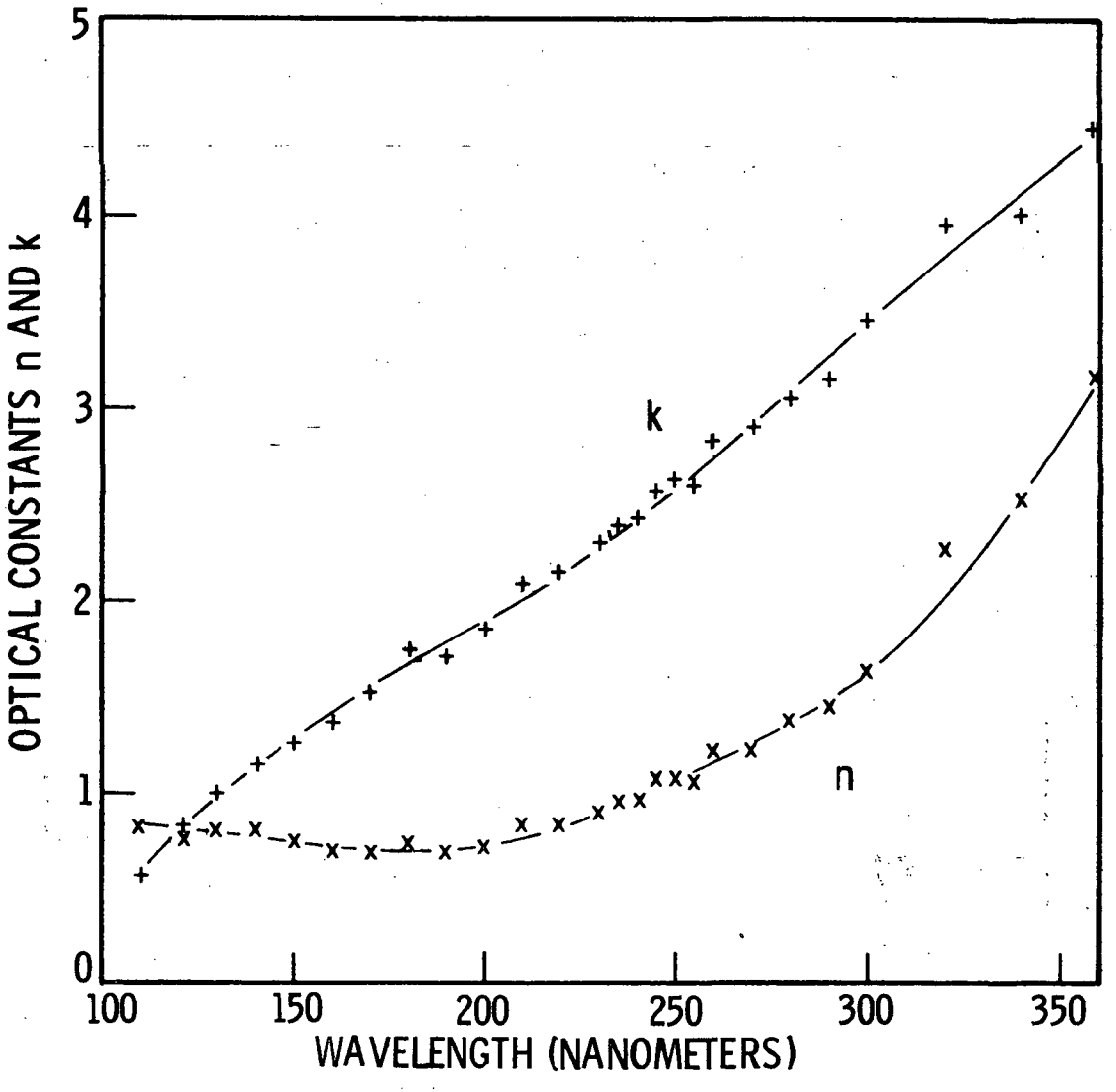


Figure 8. Optical constants of InBi versus wavelength, sample 2, as cleaved.

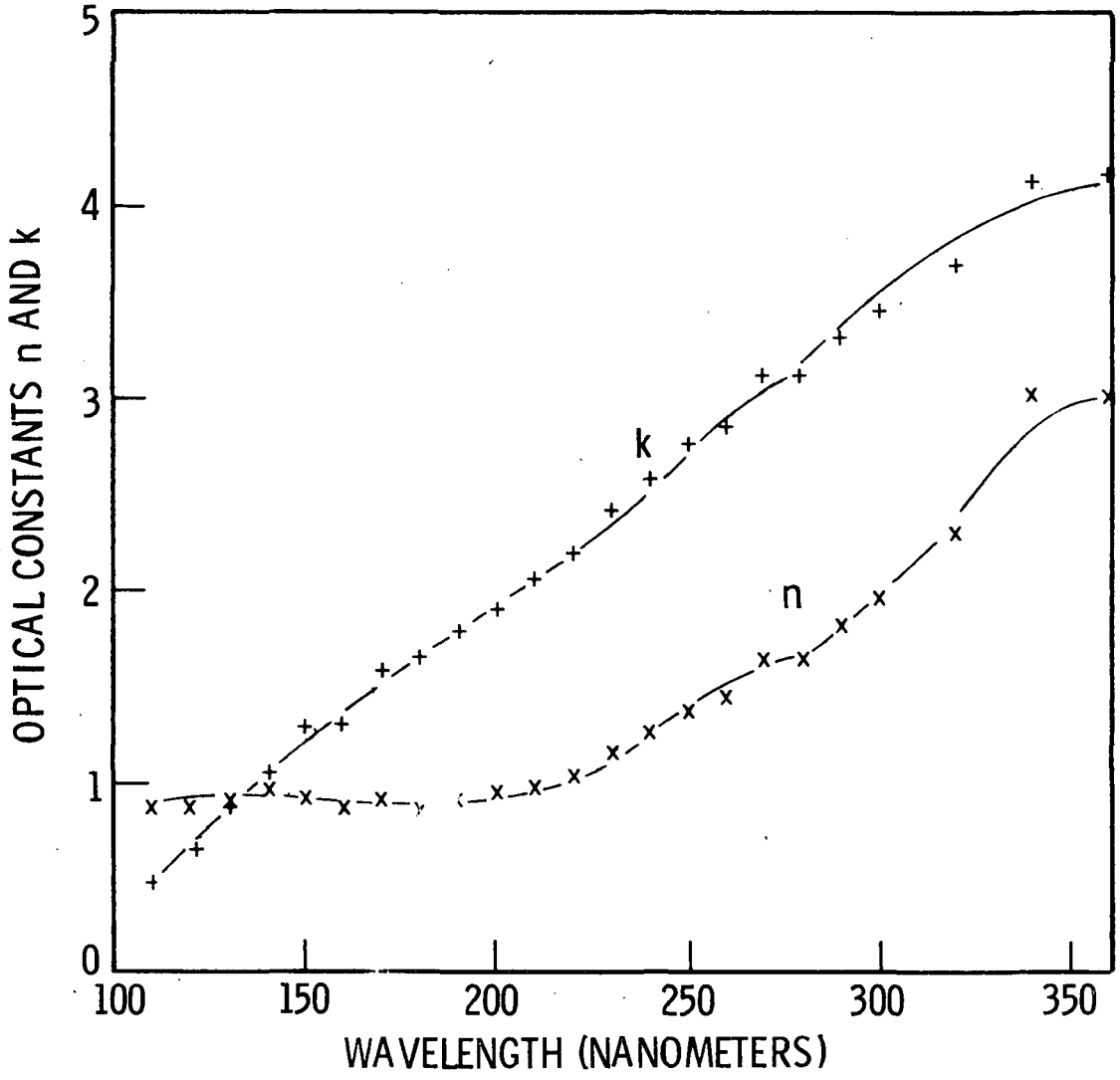


Figure 9. Optical constants of InBi versus wavelength, sample 2, air-exposure time of two weeks after the surface was cleaved.

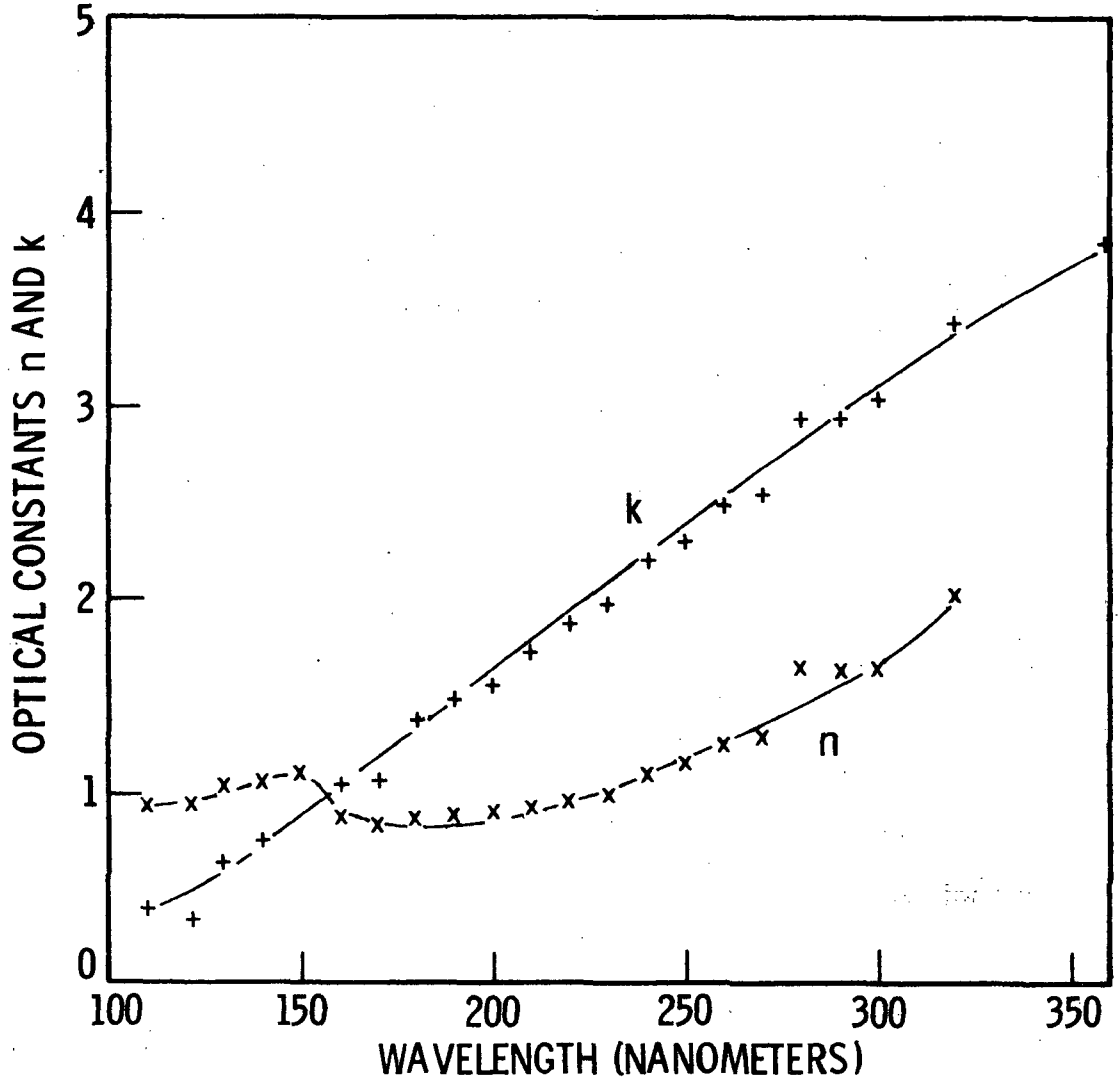


Figure 10. Optical constants of InBi versus wavelength, sample 2, air-exposure time of five weeks after the surface was cleaved.

REFERENCES

1. Newell, Jr., H. E.: Vector Analysis, McGraw-Hill Book Co., New York, 1955.

2. Slater, J. C. and Frank, N. H.: Electromagnetism, McGraw-Hill Book Co., New York, 1947.

3. Panofsky, W. K. H. and Phillips, Melba: Classical Electricity and Magnetism, Second Edition, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1962.

4. Jenkins, F. A. and White, H. E.: Fundamentals of Optics, Third Edition, McGraw-Hill Book Co., New York, 1957.
   E. T. Arakawa, M. W. Williams
5. Hayes, Jr., J. D.: The Optical Properties of Vacuum-Evaporated Films of Tellurium and Amorphous Selenium (Thesis), ORNL-TM-2023, Oak Ridge National Laboratory, January, 1968.

6. Martin, F. E.: Determination of Optical Constants by Reflectance Measurements in the Vacuum Ultraviolet, Research Reports, 1971 NASA-ASEE Summer Faculty Fellowship Research Program, MSFC-Alabama/Auburn Program, September, 1971.

7. Madden, R. P.: Preparation and Measurement of Reflection Coatings for the Vacuum Ultraviolet. Physics of Thin Films, Advances in Research and Development, George Hass, ed., Academic Press, New York, 1963.

8. Tousey, R.: On Calculating the Optical Constants from Reflection Coefficients, J. Opt. Soc. Am., Vol. 29, pp. 235—239, June 1939.

9. Linton, R. C.: Measurement and Application of Contaminant Optical Constants. AIAA 6th Thermophysics Conference, Tullahoma, Tenn., April 1971.

10. Meeker, D. S.: Computer Programs for the Analysis of Reflectance Data. Working Paper, S&E-SSL-TR-WP-1-70, NASA, George C. Marshall Space Flight Center, January 1970.

11. Binnie, W. P.: The Structural Crystallography of Indium Bismuthide, Acta Crystallographica, 9, 686, 1956.

12. Irani, G. H., Huen, T., and Wooten, F.: Optical Constants of Silver and Gold in the Visible and Vacuum Ultraviolet, J. Opt. Soc. Am., 61, 128, 1971.

13. Dekker, A. J.: Solid State Physics, Prentice-Hall Inc., Englewood Cliffs, N. J., 1957.

14. Kittel, Charles: Introduction to Solid State Physics, Fourth Edition, John Wiley & Sons, Inc., New York, 1971.

15. Richtmyer, F. K., Kennard, E. H., Lauritsen, T.: Introduction to Modern Physics, Fifth Edition, McGraw-Hill Book Co., New York, 1955.

1972

ASEE - NASA FACULTY FELLOWSHIP PROGRAM

ALABAMA - AUBURN PROGRAM

MARSHALL SPACE FLIGHT CENTER

DATA TRANSMISSION SIGNAL DESIGN AND ANALYSIS

Prepared by:	Jerry D. Moore
Academic Rank:	Assistant Professor
University:	Louisiana Tech University
Laboratory:	Astrionics Laboratory Instrumentation and Communication Division Telemetry and Data Technology Branch
Research Counterpart:	F. H. Emens
Date:	August 11, 1972
Grant Number:	NGT-01-003-045

DATA TRANSMISSION SIGNAL DESIGN AND ANALYSIS

by

Jerry D. Moore

ABSTRACT

The error performances of several digital signaling methods are determined as a function of a specified signal-to-noise ratio. This consolidates some of the results presented in the literature and provides a consistent method for a comparative performance analysis. Results are then obtained for Gaussian noise and impulse noise. Performance of a receiver for differentially encoded Biphase signaling is obtained by extending the results of differential phase shift keying. The analysis presented obtains a closed-form answer through the use of some simplifying assumptions. The results give an insight into the analysis problem, however, the actual error performance may show a degradation because of the assumptions made in the analysis. Bipolar signaling decision-threshold selection is investigated. The optimum threshold depends on the signal-to-noise ratio and requires the use of an adaptive receiver. The specifications for a minimax receiver are provided for the case of a fixed threshold level, constant noise power, and variable signal power. These conditions are possible for a data bus where the transmission media changes. An orthogonal signal set is proposed as a new data bus signal candidate. The frequency spectrum for the signal has a discrete component at the bit rate and a continuous portion that resembles the Biphase spectrum.

I. INTRODUCTION

The emphasis on digital data transmission technology has significantly increased in the past few years. Presently, the data bus projects are implementing hardward for transmissions up to 10^7 bits/sec. The choice of a signaling method to transmit the data between modems has been based on several items, viz., (1) Bandwidth required (2) Synchronization properties (3) Detection methods (4) Intersymbol interference characteristics (5) Noise penalty (6) Transmitter and Receiver complexity (7) Error performance and (8) Error detection and correction features. These considerations and the requirements of a given system are used to select acceptable signaling candidates. Typically, several sets are possible and the final choice may involve trade-offs, e.g., bandwidth is usually sacrificed for synchronization ease.

The following sections deal with a few of the problems encountered in digital data transmission analysis and design. Sections II and III are closely related for they attempt to clarify the performance criteria of digital signaling in the presence of Gaussian and impulse noise. Specifically, the probability of bit error is determined as a function of a specified signal-to-noise ratio (SNR). There are several SNRs used in the literature. These sections should consolidate some of the results and give a consistent method for comparison. Section IV presents the analysis of differentially encoded Biphase signaling. An exact performance analysis of this signaling method is not known, but the use of results obtained for DPSK are extended to give an approximation. A comparison of the analytical results to experimental results has not been attempted. Threshold selection for a Bipolar receiver is considered in Section V. The effect of a fixed-voltage decision threshold is illustrated. Under the further conditions of fixed noise power and variable signal power, a minimax receiver design is specified. Section VI gives a brief discussion of binary signal power density spectra and presents an additional signal set candidate.

II. PERFORMANCE CRITERIA FOR GAUSSIAN NOISE

The performance of a digital system is usually modelled as a random variable or some statistic of the random variable. For example, the random variable may be the number of errors that the receiver makes in the bit decisions. An important statistic of this random variable is the probability that a particular bit is in error and is designated P [E]. Almost all theoretical analyses attempt to derive P [E] as a function of the ratio of signal power (or energy) and the noise power. The concept of a signal-to-noise ratio is useful, but care must be exercised when the term is encountered for it is difficult to interpret the various forms of signal-to-noise ratios. The inherent problem is to define the points in the digital system where the signal power and noise power are to be measured. One form used quite often is

$$\text{SNR} = \frac{\text{Signal power at output of transmitter}}{\text{Noise power in some specified bandwidth}} \tag{2.1}$$

In order to illustrate how the SNR effects the error performance, the synchronous single-sample receiver is analyzed when white Gaussian noise is the disturbance.

Single-Sample Receivers (SSR) - Fundamentals

The receiver samples the received signal once per each bit period and forms bit decisions by comparing each sample value to a threshold value. The conditional decision statistic (i.e., sample value) for the k^{th} bit is

$$D_j^k = r(kT_b + \Delta t) \quad , \quad k = 0, 1, 2, \cdots \tag{2.2}$$
$$= s_j(kT_b + \Delta t) + n(kT_b + \Delta t)$$

where T_b is the bit period and Δt represents the shift into the bit period at which the sample occurs. A sampling instant Δt is chosen such that the voltage level sample from the data signal will characterize the data, i.e.,

$$s_j(kT_b + \Delta t) = V_j \quad , \quad j = 0, 1 \tag{2.3}$$

Assuming that the noise is zero-mean Gaussian, the conditional decision statistic will be Gaussian with mean value of V_j, i.e.,

$$p_{D_j}(a) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(a-V_j)^2}{2\sigma^2}} \tag{2.4}$$

where σ^2 is the variance of the noise and corresponds to noise power. For binary transmission with equal message probabilities, the decision threshold is chosen at the midpoint between the voltage levels, i.e.,

$$V_{th} = \frac{V_1 + V_0}{2} \tag{2.5}$$

This allows the probability of error for the k $\underline{\text{th}}$ bit to be formulated as

$$\begin{aligned} P[\varepsilon_k] &= \sum_{j=0}^{1} P[\varepsilon_k | m_j]\, P[m_j] \\ &= \sum_{j=0}^{1} P[(-1)^j D_j^k > V_{th}]\, P[m_j] \end{aligned} \tag{2.6}$$

After some simplification it follows that

$$P[\varepsilon_k] = \frac{1}{\sqrt{2\pi}} \int_{\frac{V_1-V_0}{2\sigma}}^{\infty} e^{-\alpha^2/2}\, d\alpha \tag{2.7}$$

$$= Q\left[\frac{V_1 - V_0}{2\sigma}\right]$$

$$= \frac{1}{2}\,\mathrm{erfc}\left[\frac{V_1 - V_0}{2\sqrt{2}\,\sigma}\right]$$

where the Q function and complementary error function are two of the tabulated forms encountered in the literature. The value of P [E_k] will decrease as the argument $(V_1-V_0)/\sigma$ increases. For specific situations, the ratio of the signal voltage to noise variance can be related to the traditional signal-to-noise ratio given by (2.1). This is accomplished in the following sections.

For the case of zero mean AWG noise and an ideal nonband limited channel, the probability of error for the SSR will be 1/2 for all physically realizable signal powers. This follows since the noise power is infinite, i.e.,

$$\sigma^2 = \int_{-\infty}^{\infty} S_N(f)\,df = \int_{-\infty}^{\infty} \frac{N_0}{2}\,df \qquad (2.8)$$

$$= \text{infinite}$$

where $S_N(f) = \frac{N_0}{2}$ is the power density spectrum of the white noise. It follows that $(V_1 - V_0)/\sigma$ will approach 0 and

$$P[\varepsilon_R] = Q[0] = \frac{1}{2} \qquad (2.9)$$

This demonstrates the need for a bandlimiting filter on the input of the receiver.

Ideal Filtered Input SSR

A filter at the input of the receiver can improve the performance. The SSR samples are taken at the output of the filter. Assuming that an ideal filter with lower cutoff frequency f_L and upper cutoff frequency of f_U is capable of passing the data signal without affecting the voltage level at the sample time, i.e., the bandwidth is defined as

$$B = f_u - f_L \qquad (2.10)$$

The noise variance (power) of the filter output is

$$\sigma^2 = \int_{f_L}^{f_u} N_0\,df = N_0 B \qquad (2.11)$$

Combining (2.11) into the error expression of (2.7) yields

$$P[\varepsilon_R] = Q\left[\frac{V_1 - V_0}{2\sqrt{N_0 B}}\right] \qquad (2.12)$$

Table 1 relates V_1-V_0 to the average transmitted signal power for several common binary signal sets. These results assume that the filter does not alter the peak signal voltage. Thus

$$P[\epsilon_R] = Q\left[\sqrt{\frac{SNR}{2}}\right] \quad \text{for Unipolar-NRZ, RZ, Biphase} \tag{2.13}$$

$$= Q\left[\sqrt{SNR}\right] \quad \text{for Polar-NRZ, RZ}$$

Practical Filtered Input SSR

The ideal filter cannot be realized. Instead of attempting to approximate the ideal filter, another approach to the problem is used. The receiver filter can modify the received signal prior to sampling, but the constraint of no intersymbol interference is imposed at the sample time. This leads to the raised cosine spectra introduced by Nyguist. The terminology "raised cosine" applies to the amplitude spectra at the output of the receiving filter. Some of the mathematics of this analysis are given in Appendix A. The performance results are obtained for Unipolar and Polar signaling as

$$P[\epsilon_R] = Q\left[\sqrt{\frac{P_{AVG}}{N_o f_s}}\right] = Q\left[\sqrt{SNR}\right] \quad \text{for Unipolar} \tag{2.14}$$

$$P[\epsilon_R] = Q\left[\sqrt{\frac{2P_{AVG}}{N_o f_s}}\right] = Q\left[\sqrt{2 \cdot SNR}\right] \quad \text{for Polar}$$

P_{AVG} is the average transmitted power and f_s is the signaling rate, i. e., $f_s = 1/T_b$. It is interesting to note that these results are the same as obtained for the theoretically optimum correlation (or matched filter) receiver (cf. Appendix B).

TABLE I

Signal Set	V_1	V_o	P_{AVG}	$\frac{V_1 - V_o}{\sigma}$
Unipolar -NRZ	$\sqrt{P_1}$	0	$P_1/2$	$\left(\frac{2P_{AVG}}{N_o B_{NRZ}}\right)^{1/2} = \sqrt{2 \cdot SNR}$
Unipolar - RZ	$\sqrt{2P_1}$	0	$P_1/2$	$2\left(\frac{P_{AVG}}{N_o B_{RZ}}\right)^{1/2} = \sqrt{2 \cdot SNR}$
Polar - NRZ	$\sqrt{P_1}$	$-\sqrt{P_o}$	$P_1 = P_o$	$2\left(\frac{P_{AVG}}{N_o B_{NRZ}}\right)^{1/2} = \sqrt{4 \cdot SNR}$
Polar - RZ	$\sqrt{2P_1}$	$-\sqrt{2P_o}$	$P_1 = P_o$	$2\left(\frac{2P_{AVG}}{N_o B_{RZ}}\right)^{1/2} = \sqrt{4 \cdot SNR}$
Biphase	$\sqrt{P_1}$	$-\sqrt{P_o}$	$P_1 = P_o$	$2\left(\frac{P_{AVG}}{N_o B_{BP}}\right)^{1/2} = \sqrt{2 \cdot SNR}$

Where

$$SNR = \frac{P_{AVG}}{N_o B_{NRZ}}$$

and

$$B_{RZ} = 2B_{NRZ} = B_{BP}$$

III. PERFORMANCE CRITERIA FOR IMPULSE NOISE

The calculation of bit error probabilities for additive impulse noise requires difficult and extensive measurements on the noise statistics. The use of bounding techniques can reduce the effort to practical limits. The answers obtained are extremely conservative, i. e., the actual P[E] will be considerably less than the bound answer for a given signal-to-noise ratio or the actual-signal-to-noise ratio required to establish a given P[E] will be less than the bound answer. The following calculations are based on the results of Houts and Moore [1].

The impulse noise model used is the GIN (Unique Waveform) model given by

$$n(t) = \sum_{k=1}^{NB} I_k f(t-T_k), \quad 0 \le t \le T_n \tag{3.1}$$

It has been shown that for a correlation receiver the expected number of bits in error given that a noise burst has occurred is

$$E\{N\} = \sum_{m=1}^{K} P_1[\mathcal{E}_m] \tag{3.2}$$

where $P_1[E_m]$ is the probability of error of the m^{th} bit which the noise waveform f(t) overlaps. Thus the bit error ratio is

$$P[\mathcal{E}] = \frac{E\{N\}\, P[NB=1]}{\#\text{ of bits transmitted}} \tag{3.3}$$

$$= \frac{E\{N\}\, P[NB=1]}{K}$$

If each $P_1[E_m]$ term is bounded by the same expression, i.e.,

$$P_1[\mathcal{E}_m] \le P_{BO}[\mathcal{E}], \quad m = 1, 2, \cdots, K, \tag{3.4}$$

it follows that

$$P[\mathcal{E}] \le P_{BO}[\mathcal{E}]\, P[NB=1] \tag{3.5}$$

Houts and Moore [1] have shown in Equation (3.69) that

$$P_1[\mathcal{E}_m] \leq \int_{\sqrt{\frac{(1-\rho)E_{AVG}}{2\sigma^2}}}^{\infty} p_N(n)\,dn \leq \frac{\sigma^2}{(1-\rho)E_{AVG}} \tag{3.6}$$

where E_{AVG} is the average signal energy, $p_N(n)$ is the normalized probability density function for the weighting factor, i. e., $N = I/\sigma$, σ^2 is the variance of N, and ρ is the correlation coefficient for the signal set.

The ratio of E_{AVG}/σ^2 can be expressed in a form that involves the ratio of average transmitted signal power to average impulse noise power at the channel output. If a receiving filter is used, the noise waveform f(t) would be changed and the noise power would be measured at the filter output. An expression for the noise power is given in Equation (A. 9) of reference 1 as

$$P_{AVG\text{-}NOISE} = \int_{-\infty}^{\infty} \nu_B \sigma^2 |F(f)|^2 df \tag{3.7}$$

where ν_B is average occurrence rate and F(f) is the Fourier transform of the noise waveform f(t). The noise statistics are determined based on the assumption that $\int f^2(t)\,dt = 1$. It follows from Parseval's theorem that

$$P_{AVG\text{-}NOISE} = \nu_B \sigma^2 \tag{3.8}$$

The average signal power is given by

$$S = \frac{E_{AVG}}{T_b} \tag{3.9}$$

Thus

$$\frac{E_{AVG}}{\sigma^2} = \frac{T_b S}{P_{AVG\text{-}NOISE}/\nu_B} = \nu_B T_b \cdot SNR \tag{3.10}$$

The bound relations of (3.6) becomes

$$P_1[\varepsilon_m] \le \int_{\sqrt{\frac{(1-\rho)\upsilon_B T_b \cdot SNR}{2}}}^{\infty} p_N(n)\,dn \le \frac{1}{(1-\rho)\upsilon_B T_b \cdot SNR} \qquad (3.11)$$

The probability of a noise burst occurring is usually not known, but will be related to the average occurrence rate by the following approximation,

$$P[NB=0] = e^{-\upsilon_B T_b} \qquad (3.12)$$

$$P[NB=1] = 1-P[NB=0] = 1-e^{-\upsilon_B T_b} \approx \upsilon_B T_b \quad \text{for } \upsilon_B T_b << 1 \qquad (3.13)$$

Using (3.13) and the last part of (3.11) in (3.5) yields

$$P[\varepsilon] \le \frac{1}{(1-\rho)\,SNR} \qquad (3.14)$$

An example of using this result is the calculation of an SNR required for a specific error rate. To establish a given error rate of less than 10^{-6} requires

$$10^{-6} = \frac{1}{(1-\rho)SNR}$$

$$SNR = \frac{10^6}{(1-\rho)} \approx 60\ dB \qquad (3.15)$$

This represents an extremely pessimistic viewpoint. If the SNR is known, the error rate can be bounded. The measurements of typical impulse noise of a simulated data bus [2] provides a means of obtaining a predicted SNR.

Again the pessimistic approach is used. The noise has been modeled as a damped sine wave with maximum observed peak value of 250 mv. An average power is not known, but will be approximated by assuming a continuous sinusoid with peak of 100 mv.

$$P_{AVG\text{-}NOISE} = \frac{.01}{2} = .005 \qquad (3.16)$$

The signal power can be approximated by

$$S = .5 \text{ watts to } 5 \text{ watts} \tag{3.17}$$

Thus

$$\text{SNR} = 100 \text{ to } 1000 \tag{3.18}$$

and

$$P[\varepsilon] \leq 10^{-2} \text{ to } 10^{-3} \tag{3.19}$$

Using the SNR = 17.15 dB $\approx$ 20dB $\Rightarrow$ 100 as in reference [2], yields the same approximate results. The results of (3.19) do not indicate satisfactory system performance. Equation (3.11) indicates that a tighter bound is available. If the p.d.f. description for the weighting factor is known, the results of (3.11) could be used as

$$P_1[\varepsilon_m] \leq \frac{\int_{\sqrt{(1-\rho)\nu_B T_b \cdot SNR/2}}^{\infty} p_N(n)dn}{} \tag{3.20}$$

A worst case value for ν_B from Ref. 2 is

$$\nu_B = \frac{10}{1.5 \times 10^{-3}} = 666 \tag{3.21}$$

Thus

$$\nu_B T_b = .666 \times 10^{-3} \tag{3.22}$$

For SNR = 10^2 to 10^3 then

$$\frac{(1-\rho)\nu_B T_b \cdot SNR}{2} = (1-\rho)(.0333) \text{ to } (1-\rho)(.333) \tag{3.23}$$

Let $\rho = 0$, then a lower limit of 0.182 to 0.58 should be used and

$$P_1[\varepsilon_m] \leq \int_{.182 \text{ to } .58}^{\infty} p_N(n)dn = 1 - \int_{-\infty}^{.182 \text{ to } .58} p_N(n)dn \tag{3.24}$$

For a Gaussian p.d.f.,

$$P_1[\varepsilon_m] \leq 1 - .572 = .428 \text{ or } \leq 1 - .719 = .281 \tag{3.25}$$

and

$$P[\varepsilon] \leq .428\,(.666 \times 10^{-3}) = .285 \times 10^{-3} \text{ OR } \leq (.281)(.666 \times 10^{-3}) = .187 \times 10^{-3} \tag{3.26}$$

If it is required that $P[E] \leq 10^{-6}$ then

$$P_1[\varepsilon_m] \approx \frac{10^{-6}}{.666\times10^{-3}} = 1.5\times10^{-3} \tag{3.27}$$

From which

$$\sqrt{\frac{(1-\rho)\nu_B T_b \cdot SNR}{2}} \approx 2.97 \tag{3.28}$$

$$SNR = \frac{8.82\,(2)}{(1-\rho)\,(.666\times10^{-3})} = \frac{26.4\times10^{3}}{(1-\rho)}$$

or

$$SNR \approx 44 \text{ dB} \tag{3.29}$$

The results of (3.29) and (3.26) show an improvement of the results obtained in (3.19) and (3.15).

IV. ANALYSIS OF PHASE DETECTOR FOR DIFFERENTIALLY ENCODED BIPHASE SIGNALING

The determination of the error performance of differential Biphase using a phase detector involves many assumptions and approximations. It does not appear feasible to compare the predicted results to actual results without an extensive test program. In the following sections, the mathematics of the DPSK (differential phase shift keying) analysis are presented. These results are then extended to Biphase signaling.

DPSK Analysis

The signal set is specified as

$$s_j(t) = \sqrt{2S}\cos(2\pi f_0 t + \theta_j), \quad j = 0, 1 \tag{4.1}$$

where S is the signal power, f_0 the frequency and $\theta_j = (-\pi)^j$. The narrowband AWG noise is modeled as

$$\begin{aligned} n(t) &= x(t)\cos(2\pi f_0 t) - y(t)\sin(2\pi f_0 t) \\ &= r(t)\cos(2\pi f_0 t + \varphi(t)) \end{aligned} \tag{4.2}$$

where

$$\begin{aligned} r(t) &= \sqrt{x^2(t) + y^2(t)} \\ \varphi(t) &= \tan^{-1}\left(\frac{y(t)}{x(t)}\right) \end{aligned} \tag{4.3}$$

The signal plus noise is

$$\begin{aligned} v(t) &= \sqrt{2S}\cos(2\pi f_0 t + \theta_j) + x(t)\cos(2\pi f_0 t) - y(t)\sin(2\pi f_0 t) \\ &= \left[\sqrt{2S}\cos(\theta_j) + x(t)\right]\cos(2\pi f_0 t) - y(t)\sin(2\pi f_0 t) \\ &= x'(t)\cos(2\pi f_0 t) - y(t)\sin(2\pi f_0 t) \\ &= r'(t)\cos(2\pi f_0 t + \Theta(t)) \end{aligned}$$

where (4.4)

$$r'(t) = \sqrt{(x'(t))^2 + y^2(t)}$$

and

$$\theta(t) = \tan^{-1}\left(\frac{y(t)}{x'(t)}\right)$$

Following the procedure shown by Schwartz [7] and assuming that $x'(t)$ and $y(t)$ are statistically independent, allows the joint probability density function $p(r, \theta)$ to be determined. Integrating over all values of r' gives

$$p(\theta|j) = \frac{e^{-\rho^2}}{2\pi} + \frac{(-1)^j}{2}\sqrt{\frac{\rho^2}{\pi}}\cos(\theta)\, e^{-\rho^2\sin^2(\theta)}\left[1 + \operatorname{erf}\left((-1)^j \rho\cos(\theta)\right)\right] \tag{4.5}$$

where

$$\theta' = \theta + (-\pi)^j, \quad \rho^2 = \frac{S}{N}, \quad N = \text{variance of noise} \tag{4.6}$$

For large signal-to-noise ratio, i.e., $\rho^2 >> 1$ this can be approximated [7] by

$$p(\theta|j) \approx \begin{cases} \dfrac{\cos(\theta + (-\pi)^j)}{\sqrt{\pi/\rho^2}}\, e^{-\rho^2 \sin^2(\theta + (-\pi)^j)} & \text{for } |\theta + (-\pi)^j| \\ 0 & \text{for } |\theta + (-\pi)^j| > \frac{\pi}{2} \end{cases} \tag{4.7}$$

An approximation proposed by Cahn [8] for (4.7) is

$$p(\theta|j) \approx \begin{cases} \dfrac{1}{\sqrt{\pi/\rho^2}}\, e^{-\rho^2\left[\theta + (-\pi)^j\right]^2} & \text{for } |\theta + (-\pi)^j| \le \frac{\pi}{2} \\ 0 & \text{for } |\theta + (-\pi)^j| > \frac{\pi}{2} \end{cases} \tag{4.8}$$

It is obvious that this expression is not exact and will introduce error in the answers obtained from its application.

The result of (4.7) is similar to the results of Rice for a sinusoidal wave plus Gaussian noise. The primary difference is the inclusion of the phase of the sinusoid in (4.7). No attempt was made to justify the approximation of (4.8), however, this would have been a straight forward conclusion for small values of Θ.

The performance of DPSK will be determined using (4.8). Numerous simplifying assumptions are made throughout the derivation. The main significance of the results is to compare to answers stated by Cahn and others [8, 9, 6].

The probability of error for the differential phase detector depends on the two data bits involved, i.e., 00, 01, 10, or 11. The phase of the first bit θ_1 and the second bit θ_2 are compared by the detector as $\Delta\theta = \theta_2 - \theta_1$. Formulations of the error expressions are

$$P[\varepsilon|10] = P[\varepsilon|00] = P[|\Delta\theta| > \pi/2] \tag{4.9}$$
$$P[\varepsilon|01] = P[\varepsilon|11] = 1 - P\left[\frac{\pi}{2} < \Delta\theta < \frac{3\pi}{2}\right]$$

Thus to determine the probability of error it is necessary to find the expression for $p(\Delta\theta)$. This is possible by the following steps.

$$\Delta\theta = \theta_2 - \theta_1 \tag{4.10}$$

The characteristics function of $\Delta\theta$ is

$$M_\Delta(u) = E\left[e^{ju\Delta\theta}\right] = E\left[e^{ju\theta_2} e^{-ju\theta_1}\right] \tag{4.11}$$

Assuming statistical independence of the phases

$$M_\Delta(u) = E\left[e^{ju\theta_2}\right] \cdot E\left[e^{-ju\theta_1}\right] \tag{4.12}$$

Thus

$$p(\Delta\theta) = p_{\theta_2}(\theta) * p_{\theta_1}(-\theta) = p(\theta) * p(-\theta)$$
$$= \int_{-\infty}^{\infty} p(\theta)\, p(\theta+\Delta\theta)\, d\theta \tag{4.13}$$

Applying (4.13) to the error expression $P[E|00]$, etc., and observing that each error term can be evaluated by using the basic form of $p(\theta|j)$ with $j=0$ yields

$$p(\Delta\theta) = \int_{-\pi/2}^{\pi/2} p(\theta|j=0)\, p(\theta+\Delta\theta|j=0)\, d\theta \qquad (4.14)$$

Substituting (4.8) into (4.13), completing the square in the exponent and then approximating the integral yields

$$p(\Delta\theta) = \frac{e^{-\rho^2 \frac{\Delta\theta^2}{2}}}{\sqrt{2\pi}\,\frac{1}{\rho}} \int_{-\pi/2}^{\pi/2} \frac{1}{\sqrt{2\pi}\,\frac{1}{\rho}}\, e^{-\frac{(\theta+\frac{\Delta\theta}{2})^2}{2\left(\frac{1}{4\rho^2}\right)}}\, d\theta \qquad (4.15)$$

$$= \frac{1}{\sqrt{2\pi}\,\frac{1}{\rho}}\, e^{-\rho^2\frac{\Delta\theta^2}{2}} \qquad -\pi < \Delta\theta < \pi$$

The probability of error from (4.9) become

$$P[\varepsilon] = 1 - P\left[-\tfrac{\pi}{2} < \Delta\theta < \tfrac{\pi}{2}\right] = 2\int_{-\pi/2}^{\infty} p(\Delta\theta)\, d\Delta\theta$$

$$= 2\,Q\left[\frac{\rho\pi}{2}\right]$$

The Q function can be bounded by two different functions (cf., Wozencraft and Jacobs [10] pp. 83, 84) to yield

$$P[\varepsilon] < e^{-\pi^2\rho^2/8} \qquad (4.17)$$

or

$$P[\varepsilon] < \frac{2\sqrt{2}}{\rho\pi\sqrt{\pi}}\, e^{-\pi^2\rho^2/8} \qquad (4.18)$$

These can be contrasted to the result of Lucky, Salz and Weldon [6]

$$P[\varepsilon] = e^{-\rho^2} \tag{4.19}$$

and Cahn [8] and Bennett and Davey [4]

$$P[\varepsilon] = \frac{1}{2} e^{-\rho^2} \tag{4.20}$$

and later results by Cahn [9]

$$P[\varepsilon] = \frac{1}{\sqrt{\pi \rho^2}} e^{-\rho^2} \tag{4.21}$$

The optimum coherent receiver result is

$$P[\varepsilon] = \frac{1}{2\sqrt{\pi \rho^2}} e^{-\rho^2} \tag{4.22}$$

Observe that the exponent of (4.18) exhibits a factor of $\pi^2/8$ improvement over the optimum case. This impossibility results from the many approximations made during the derivation. The other results stated in (4.19), (4.20), and (4.21) could be acceptable, but there are few comparisons of experimental data to these equations. Plots of the various answers are shown in Fig. 4-1.

Extension of DPSK Analysis To Differential Biphase

Differential phase detection of Bipolar signaling has received limited analysis in the literature. If the channel and receiving filter act to smooth out the Biphase pulses to sinusoidal form, it appears that the DPSK could be used directly. This is apparently the procedure followed by Perry [2]. The signal-to-noise ratio used for DPSK must be modified for Biphase. Assuming the same peak signal level, the average signal power on the channel is increased by a factor of 2. The noise power is changed by the receiving filter by a factor of $\pi^2/8$. The new value of ρ^2 is thus

$$\rho_{B\phi}^2 = \frac{4}{\pi^2} \rho_{DPSK}^2 \tag{4.23}$$

as per Perry [2]. This could be substituted into (4.19) to obtain Perry's result or in (4.20) to obtain a similar result.

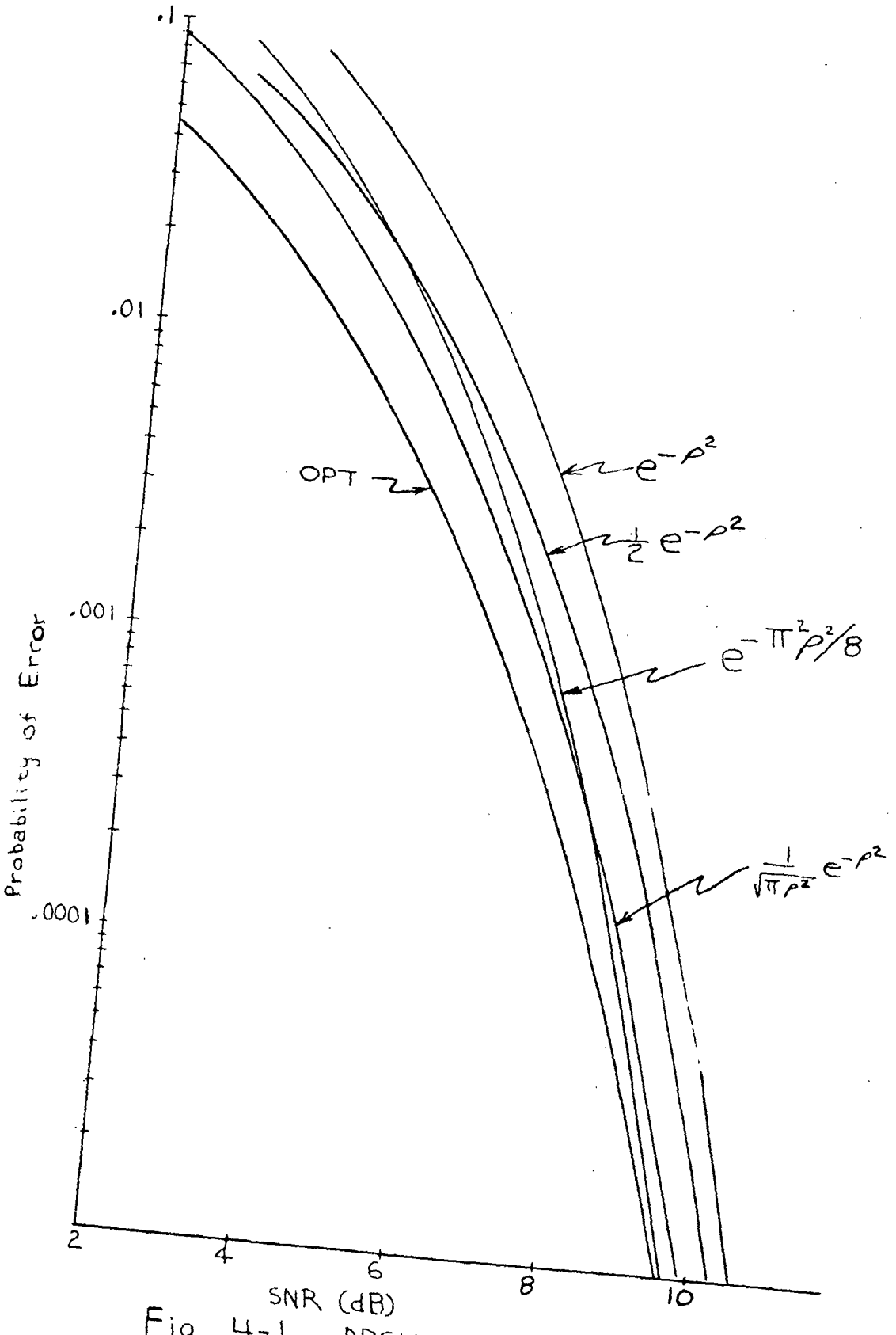


Fig. 4-1. DPSK Performance

V. THRESHOLD SELECTION FOR BIPOLAR SIGNALING

In Bipolar signal transmission, the binary symbol "ZERO" is represented by no signal on the line, and the binary symbol "ONE" is represented alternately by positive and negative standard pulses. A representation of the data train waveform is

$$s(t) = \sum_{n=-\infty}^{\infty} b_n \, g(t - nT_b) \tag{5.1}$$

$$b_n = \begin{cases} 0 & \text{with prob. } (1-p) \\ +1 & \text{with prob. } p/2 \\ -1 & \text{with prob. } p/2 \end{cases}$$

If g(t) = A throughout the bit period, the power density spectrum of the signal is given by (cf. Eq. 19.124 of Bennet & Davey [4]),

$$W_s(f) = 2A^2 T_b \, \mathrm{sinc}^2(fT_b) \sin^2(\pi f T_b) \tag{5.2}$$

This function is plotted in Fig. 5-1. Note that a discrete DC component is not present and that amplitudes at frequencies near DC are low. These features indicate why Bipolar is a prime candidate for use in data bus applications.

A single-sample receiver (SSR) could be used to detect Bipolar. The sample value is compared to a threshold voltage of +E volts and -E volts. A sample level between +E and -E is output as a ZERO, while a level greater than +E or less than -E is output as a corresponding ONE. A simple SSR would not take account of the alternating properties of the ONE symbols. For such a case, Perry [2] has shown that the optimum threshold value is

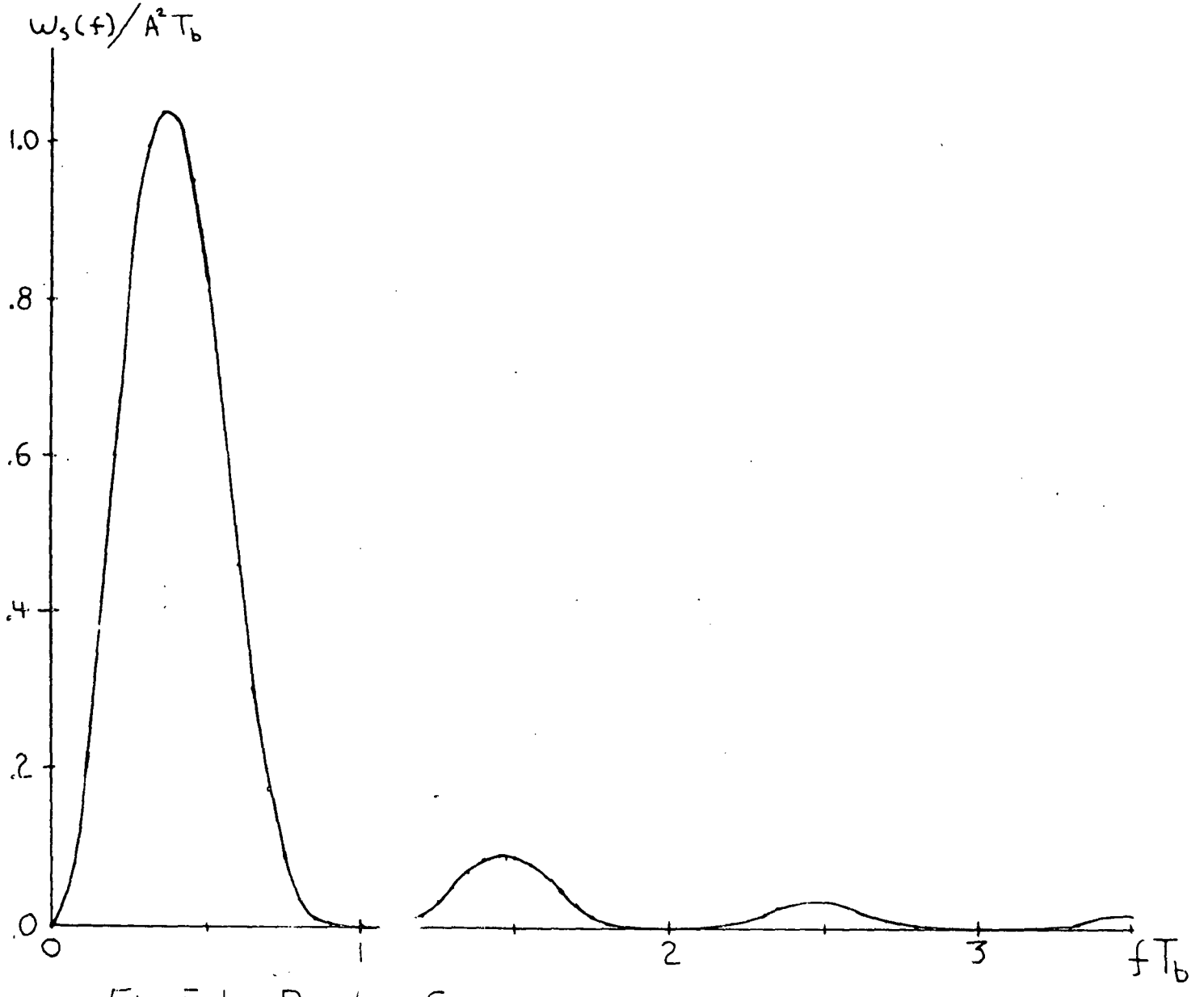


Fig. 5-1. Bipolar Spectrum

$$E_{OPT} = \sqrt{\frac{S}{2}}\left[\frac{\cosh^{-1}\left(e^{S/N}\right)}{S/N}\right] \qquad (5.3)$$

where S is average signal power, N is noise power and equal transmission probabilities have been assumed.

Note that if the magnitude of the sample level is greater than E, the output is ONE without regard to polarity. A slightly more complex SSR could utilize the alternating property for error checking. Such a receiver would have an optimum threshold of

$$E_{OPT} = \sqrt{\frac{S}{2}} + \frac{N}{\sqrt{2S}}\ln(2) \qquad (5.4)$$

This follows from an extension of the result given by Wozencraft & Jacobs [10] (cf. , Eq. (2.115) and Fig. 2.35 b).

A large signal-to-noise ratio would yield low probabilities that a given polarity ONE would be received as the other polarity ONE. For such conditions the threshold of (5.3) should approach that of (5.4). This can be shown mathematically by expanding the $\cosh^{-1}$ term of (5.3) into a power series, viz.,

$$\cosh^{-1}(x) = \ln(2x) - \frac{1}{2}\frac{1}{2x^2} - \frac{1}{2}\cdot\frac{3}{4}\cdot\frac{1}{4x^4} - \cdots \qquad (5.5)$$

Letting $x = e^{S/N}$ and neglecting all but the first term allows the result of (5.3) to become (5.4).

The significant feature is that an optimum threshold is a function of the signal-to-noise ratio. Implementation would require an adaptive receiver. Most receivers would not use the adaptive threshold, but would have a fixed value set for E. The performance of this fixed threshold type system which does not check for polarity is described by the curves shown in Fig. 5-2. The solid line curves represent the performance for the various ratios of E/A, i.e., .2,.3, ···, .8. A fixed value for E and E/A means that A most remain constant, and that the noise power is what must change to change the SNR value. Another interesting condition is for fixed value of E and noise power, but a variable value for A. This might result when transmission media with different characteristic impedances are used in the data channel. For such a set of conditions, the original value of E/A has a strong control over the performance as A (or SNR) changes. A numerical example illustrates the relationships. Let the minimum value of A be 1.4 volts. Table 5.1 shows the values of A required to establish given E/A values based of original values for E/A of .5, .6 and .7. The right band columns of the table indicate the dB increase in SNR as the A value changes. These data are plotted in Fig. 5-2 as the dashed lines and with an original SNR of 11 dB. A best choice for the threshold depends on the criteria used. The smallest maximum P[E] (i.e., the minimax) is obtained by using E = .7 volts. Actually the E_{OPT} value for A = 1.4 volts should be used, but E_{OPT} will approach .5A for large SNR. A significant improvement in P [E] could be obtained as A increases if a large P [E] can be tolerated at the value of A = 1.4. Such a case is illustrated by curve C.

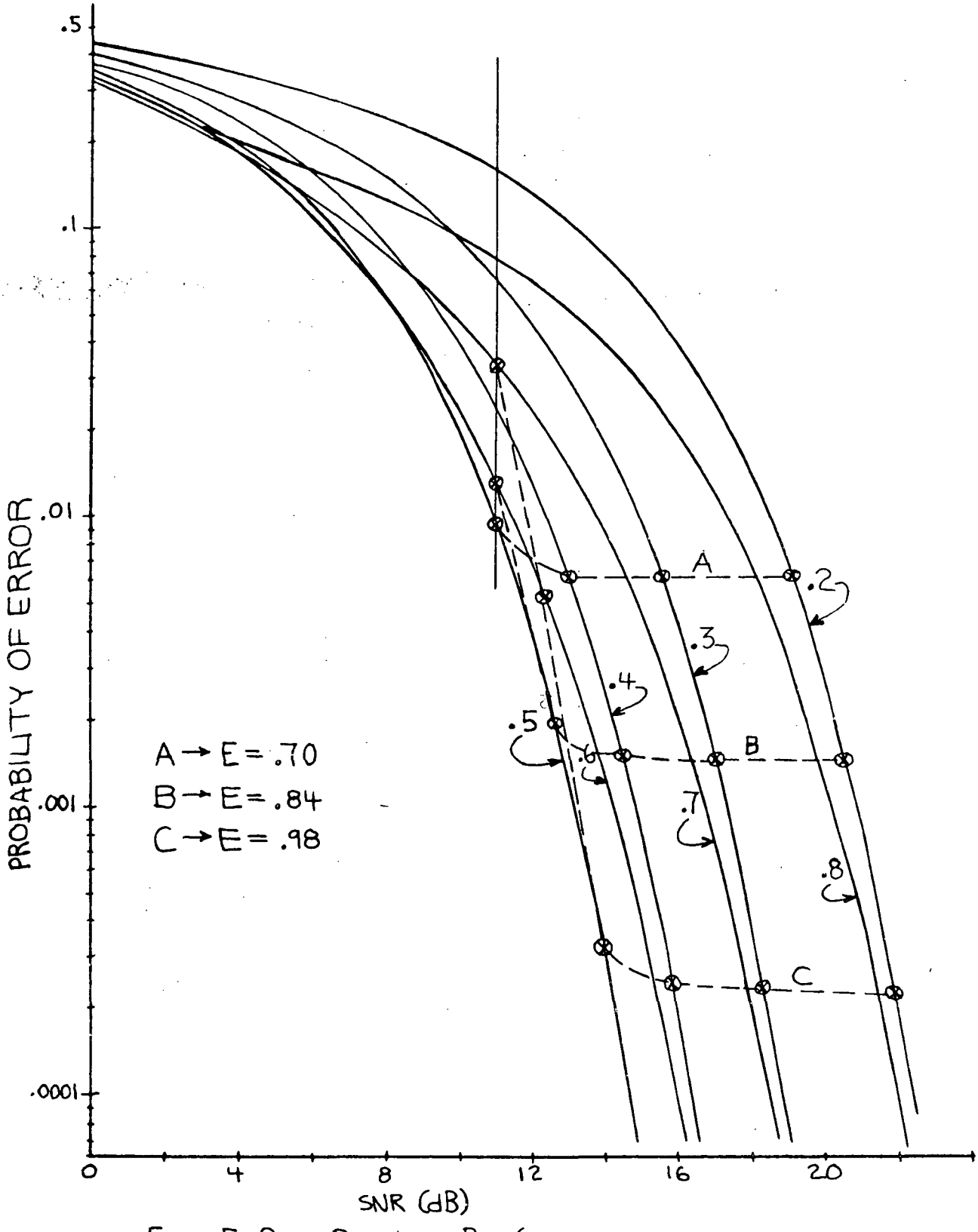


Fig. 5-2. Bipolar Performance

TABLE 5.1

E = .7

E/A	A	20 log $\frac{A}{1.4}$
.5	1.40	0
.4	1.75	2.0
.3	2.34	4.5
.2	3.50	8.0

E = .84

E/A	A	20 log $\frac{A}{1.4}$
.6	1.40	0
.5	1.68	1.6
.4	2.10	3.5
.3	2.80	6.0
.2	4.20	9.6

E = . 98

E/A	A	20 log $\frac{A}{1.4}$
.7	1.40	0
.6	1.63	1.3
.5	1.96	3.0
.4	2.45	4.8
.3	3.26	7.4
.2	4.90	10.9

VI. SPECTRAL PROPERTIES OF AN ORTHOGONAL SIGNALING SET

One of the important considerations in data transmission signal design is the spectrum occupancy. Several signaling sets have been considered in references 2 and 3. A general expression for the spectra of one class of binary signal sets has been given by Bennett and Davey [3]. If a train of data signals are represented by a summation of two signals and translations, i. e.,

$$s(t) = \sum_{n=-N}^{N} s_n(t) \tag{6.1}$$

with

$$s_n(t) = \begin{cases} g_1(t - nT_b) & \text{with prob. } p \\ g_2(t - nT_b) & \text{with prob. } (1-p) \end{cases} \tag{6.2}$$

then the power density spectrum is

$$\begin{aligned} W_s(f) &= 2f_s\, p(1-p)\, \left| G_1(f) - G_2(f) \right|^2 \\ &\quad + f_s^2 \left[p\, G_1(0) + (1-p)\, G_2(0) \right]^2 \delta(f) \\ &\quad + 2f_s^2 \sum_{m=1}^{\infty} \left| p\, G_1(mf_s) - (1-p)\, G_2(mf_s) \right|^2 \delta(f - mf_s) \end{aligned} \tag{6.3}$$

where $f_s = 1/T_b$ and G(f) is the Fourier transform of g(t).

An addition to the signal sets considered in references 2 and 3 is the Orthogonal set shown in Fig. 6-1. The spectrum for this signaling is shown in Fig. 6-2. The significant features of this spectrum are

1. A discrete component at the bit rate $1/T_b$

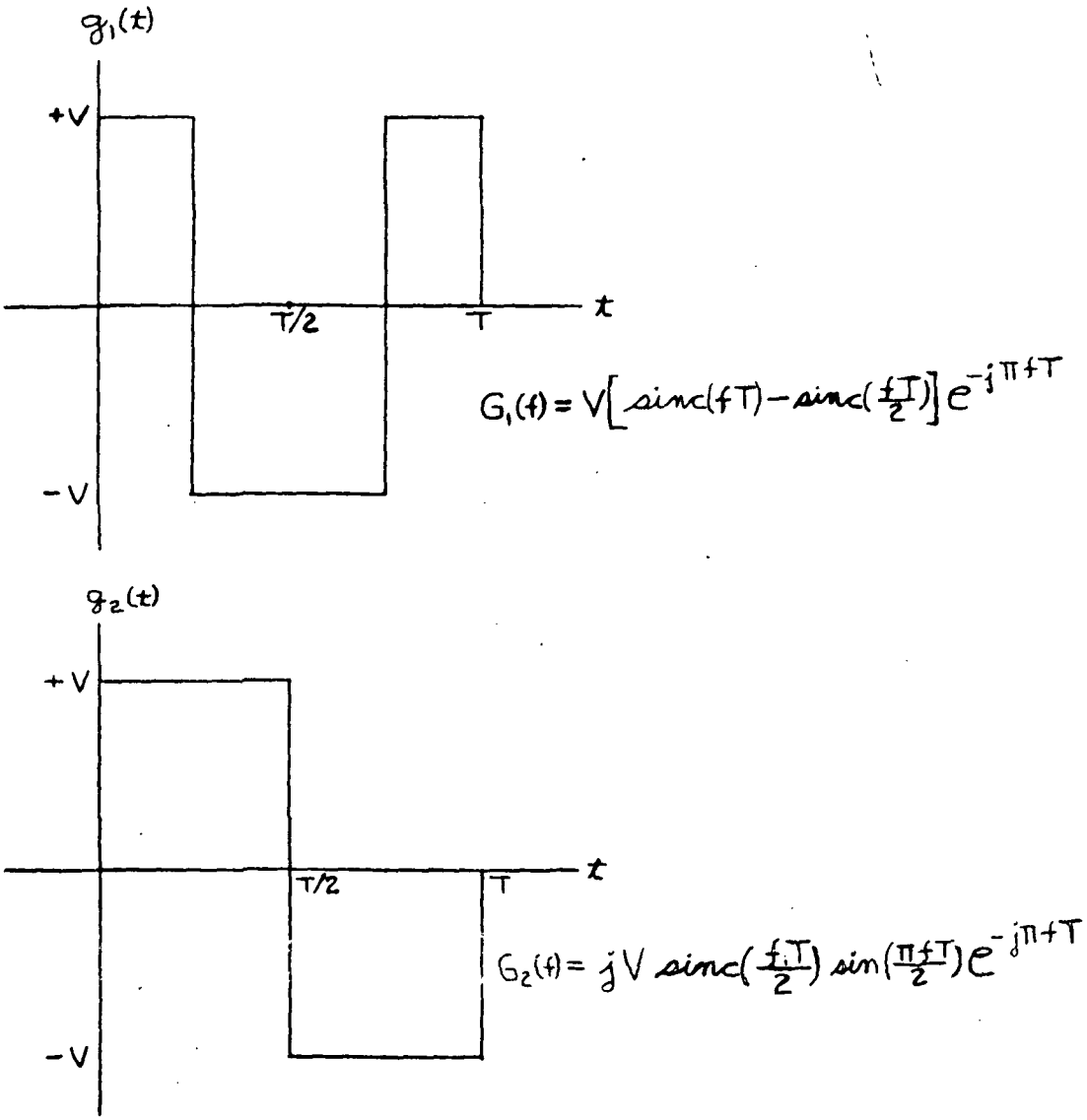


Fig. 6-1. Orthogonal Signal Set

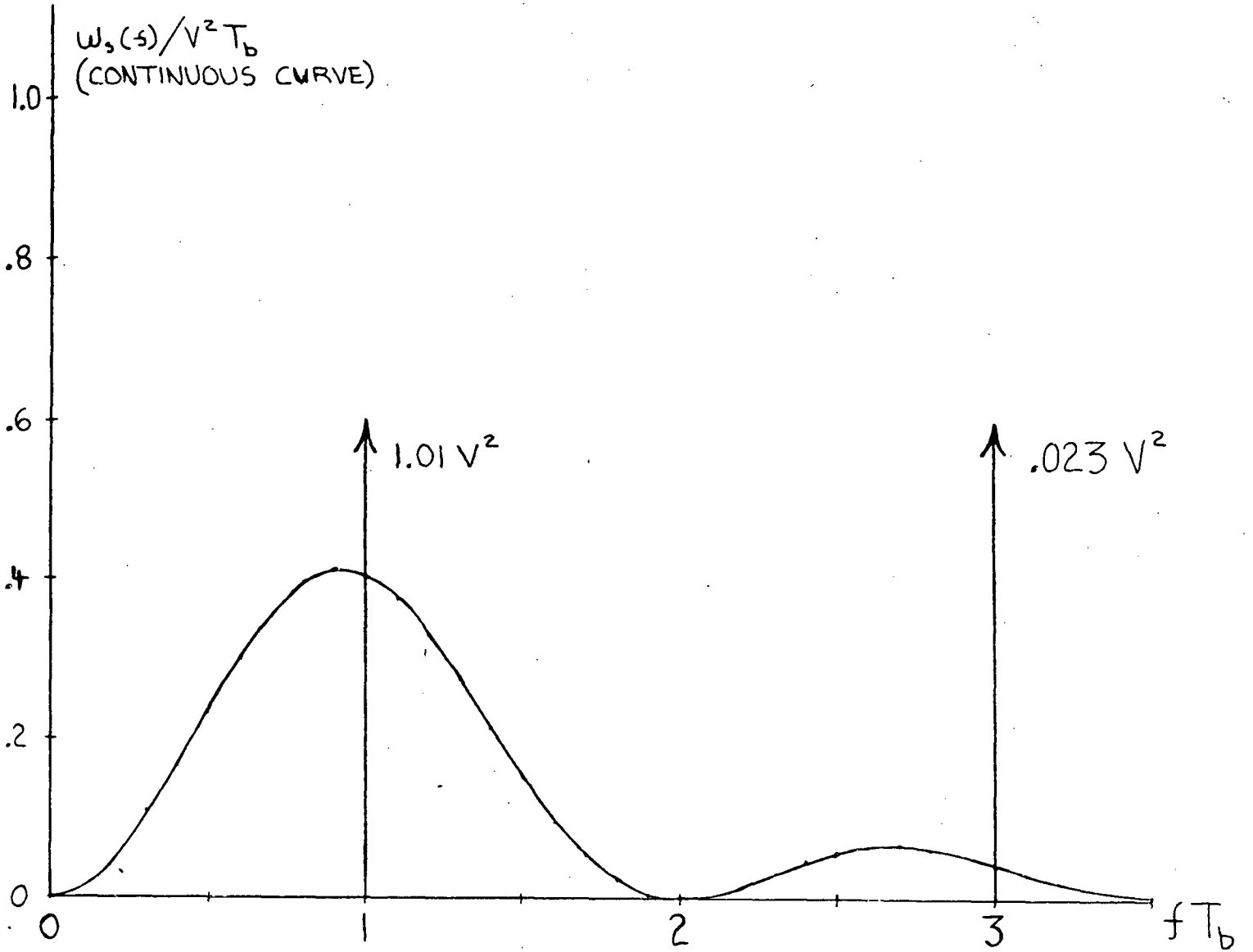


Fig. 6-2. Orthogonal Signal Spectrum

2. No discrete component at DC.

3. Continuous spectrum has low values at low frequencies approaching DC.

Items 2 and 3 are generally the same as that obtained for Biphase; however, the discrete component of Item 1 is not present in Biphase. A disadvantage of Orthogonal as compared to Biphase is the 3dB penalty in SNR that appears in the error performance equation for the optimum correlation receiver.

APPENDIX A

NYQUIST CRITERIA EFFECT ON SNR

An actual system containing sending and receiving filters, a channel model, and the SSR is shown in Fig. A-1. The main design objective associated with this system is to produce a received signal r(t) such that there is no intersymbol interference between the samples taken from the signal. Neglecting noise, the input-to-output relations for the system are

$$s(t) = \sum_{n=-N}^{N} b_n \, q(t - nT_b) \quad ; T_b = 1/f_s \tag{A.1}$$

where each b_n has a value depending on the data, e.g., (0,1) or (-1, +1).

$$S(f) = \sum_{n=-N}^{N} b_n \, G(f) e^{-j 2\pi n f T_b} \tag{A.2}$$

$$R(f) = Y_1(f) \, H(f) \, Y_2(f) \, S(f)$$

$$= \sum_{n=-N}^{N} b_n \, Q(f) \, e^{-j 2\pi n f T_b} \tag{A.3}$$

$$Q(f) = Y_1(f) \, H(f) \, Y_2(f) \, G(f) \tag{A.4}$$

$$r(t) = \sum_{n=-N}^{N} b_n \, q(t - nT_b) \tag{A-5}$$

To minimize intersymbol interference requires that

$$q(mT_b) = A \, \delta_{mo} \tag{A.6}$$

where $\delta_{oo} = 1$ and $\delta_{mo} = 0$ for $m \neq 0$.

Several functions could satisfy this requirement, e.g., a retangular pulse of width T_b or a function of the form A sinc ($f_s t$). The form of q (t) proposed by Nyguist [3, 4, 5, 6] is

$$q(t) = \frac{A \cos(\rho \pi f_s t)}{1 - 4\rho^2 f_s^2 t^2} \operatorname{sinc}(f_s t) \tag{A.7}$$

where the rolloff factor ρ is

$$\rho = 1 - \frac{2f_1}{f_s} \tag{A.8}$$

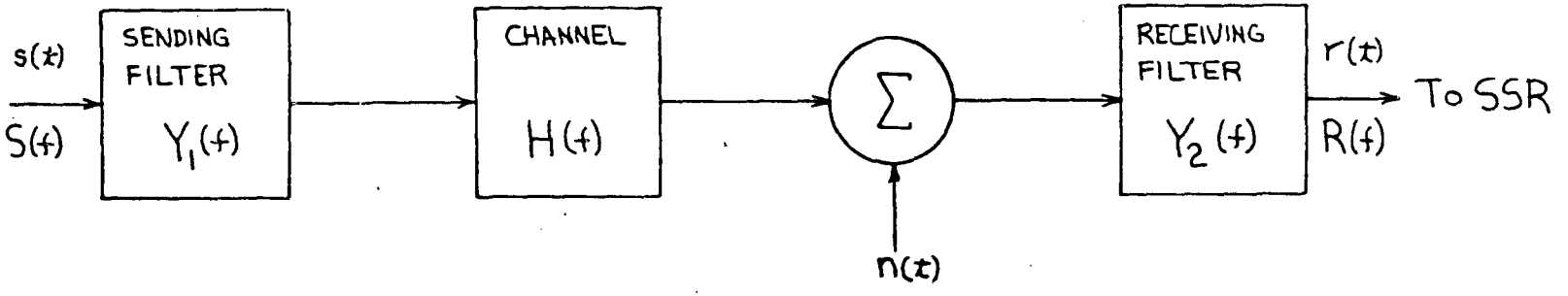


Fig. A-1. Digtal System Block Diagram.

The transform of q(t) is

$$Q(f) = \begin{cases} AT_b & |f| < f_1 \\ \dfrac{AT_b}{2}\left[1 + \cos\left(\dfrac{\pi(|f| - f_1)}{f_s - 2f_1}\right)\right] \\ \text{or} \\ \dfrac{AT_b}{2}\left[1 + \cos\left(\dfrac{\pi(|f| - f_1)}{\rho f_s}\right)\right] \\ \text{or} \\ \dfrac{AT_b}{2}\left[1 - \sin\left(\dfrac{\pi(|f| - f_s/2)}{\rho f_s}\right)\right] & f_1 < |f| < f_s - f_1 \\ 0 & |f| > f_s - f_1 \end{cases} \tag{A.9}$$

A special case of interest is when $\rho = 1$. It follows that $f_1 = 0$ and

$$q(t) = \frac{A\cos(\pi f_s t)}{1 - 4f_s^2 t^2}\,\mathrm{sinc}(f_s t) \tag{A.10}$$

(see Fig. A-2.)

$$Q(f) = \begin{cases} \dfrac{AT_b}{2}\left[1 + \cos\left(\dfrac{\pi|f|}{f_s}\right)\right] & 0 \le |f| \le f_s \\ 0 & |f| > f_s \end{cases}$$

Note that

$$q(0) = A = \int_{-\infty}^{\infty} Q(f)\,df = 2\int_0^{f_s} Q(f)\,df \tag{A.11}$$

The next step in the development is to recognize that the error rate depends on the ratio of A/σ. This follows from (2.7) and (A.5), i.e.,

$$P[\varepsilon_k] = Q\left[\frac{(b_1 - b_0)A}{2\sigma}\right] \tag{A.12}$$

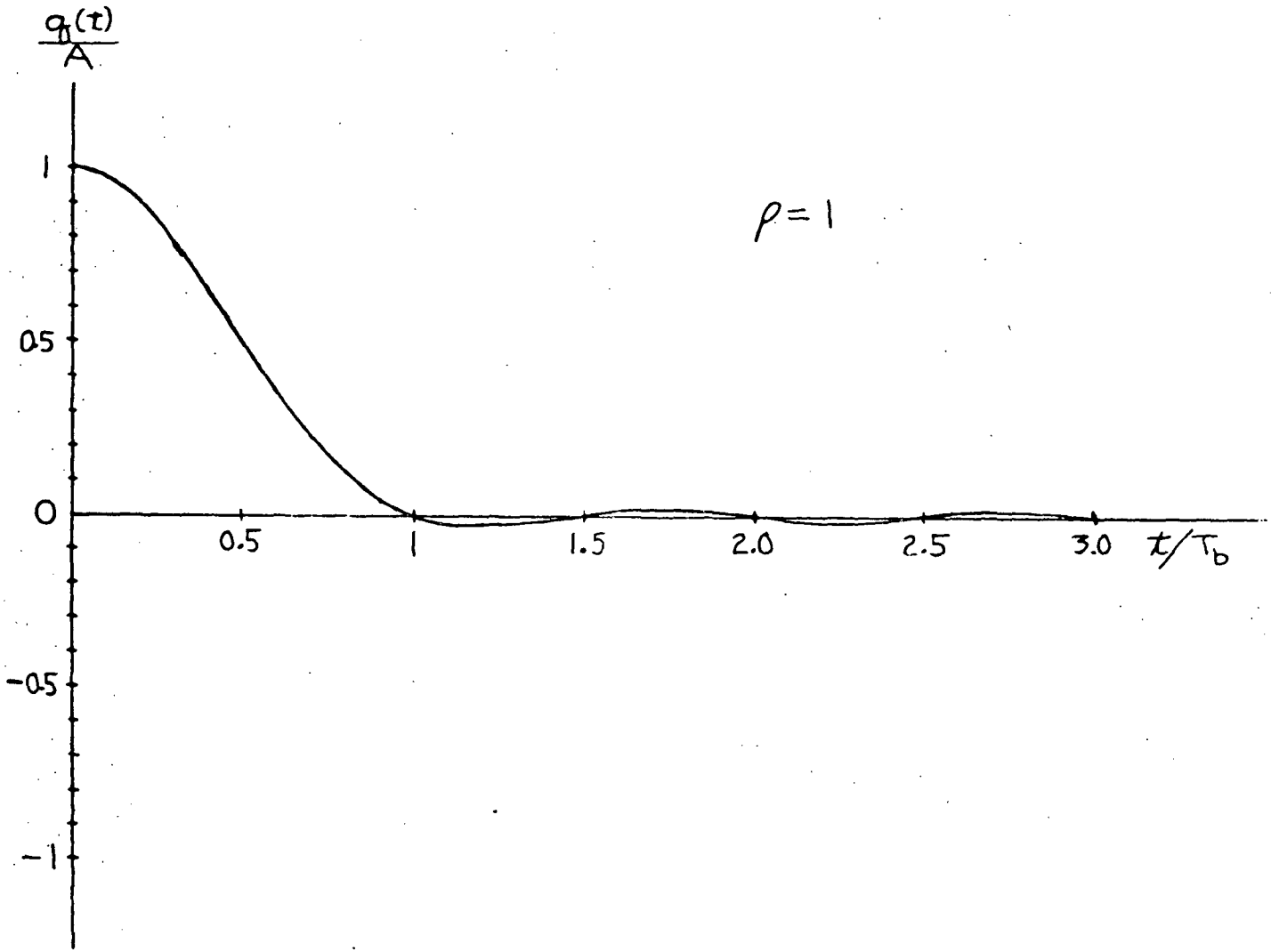


Fig. A-2. Nyquist Pulse Function

where b_1 and b_0 are the ONE and ZERO representations that b_n assumes. The probability of error will be a minimum when A/σ is maximized. An expression for σ^2 can be obtained in terms of the other system parameters. For white noise with two-sided power density spectra $\frac{N_0}{2}$, it follows that

$$\sigma^2 = \int_{-\infty}^{\infty} |Y_2(f)|^2 \frac{N_0}{2}\, df = N_0 \int_0^{\infty} |Y_2(f)|^2 df \tag{A.13}$$

Thus

$$\frac{A^2}{\sigma^2} = \frac{4\left[\int_0^{f_s} Q(f)\, df\right]^2}{N_0 \int_0^{\infty} |Y_2(f)|^2 df} \tag{A.14}$$

The expression can be maximized with respect to $|Y_2(f)|$ and solved [4, 5] to yield

$$Y_2(f) = \sqrt{\frac{A f_s\, Q(f)}{H_0^2\, W_1}} \left[\int_0^{\infty} Q(f)\, df\right]^{1/2} \tag{A.15}$$

where

$$H(f) = H_0$$

$$W_1 = f_s \int_{-\infty}^{\infty} |Y_1(f)\, G(f)|^2 df \tag{A.16}$$

Thus

$$\frac{A^2}{\sigma^2} = \frac{2 H_0^2 W_1}{N_0 f_s} \tag{A.17}$$

The term W_1 is related to the average transmitted power, viz.,

$$\text{Avg. Power in } s(t) = \frac{b_1^2 + b_0^2}{2T_b} \int_{-T_b/2}^{T_b/2} g^2(t)\,dt \qquad \text{(A.18)}$$

$$= f_s \frac{(b_1^2 + b_0^2)}{2} \int_{-\infty}^{\infty} |G(f)|^2 df$$

$$\text{Avg. Power at Channel Input} = \frac{f_s(b_1^2 + b_0^2)}{2} \int_{-\infty}^{\infty} |Y_1(f) G(f)|^2 df \qquad \text{(A.19)}$$

$$= \frac{b_1^2 + b_0^2}{2} W_1$$

Also (A.20)

$$W_1 = f_s \int_{-\infty}^{\infty} \left| \frac{Q(f)}{Y_2(f) H_0} \right|^2 df$$

$$= 2 f_s \int_{0}^{\infty} \left| \frac{Q(f)}{Y_2(f) H_0} \right|^2 df$$

The ratio of (A.17) can be written in terms of the SNR

$$\frac{A^2}{\sigma^2} = \frac{4H_o^2}{b_1^2 + b_o^2} \cdot SNR \qquad (A.21)$$

Thus

$$P[\varepsilon_k] = Q\left[\frac{b_1 - b_o}{\sqrt{b_1^2 + b_o^2}} \sqrt{SNR}\right] \qquad (A.22)$$

For On-Off signaling $b_1 = 1$, $b_o = 0$ and

$$P[\varepsilon_k] = Q\left[\sqrt{SNR}\right] \qquad (A.23)$$

For antipodal $b_1 = 1$, $b_o = -1$ and

$$P[\varepsilon_k] = Q\left[\sqrt{2 \cdot SNR}\right] \qquad (A.24)$$

APPENDIX B

MATCHED FILTER ANALYSIS

The impulse response of the filter, h(t) is related to the transmitted waveform g(t) by

$$h(t) = g(T_b - t) \tag{B.1}$$

Assuming pulse signaling

$$g(t) = \begin{cases} V & 0 \le t \le T_b \\ 0 & \text{elsewhere} \end{cases} \tag{B.2}$$

and

$$G(f) = V T_b \ \text{sinc}(f T_b) e^{-j\pi f T_b} \tag{B.3}$$

Thus

$$h(t) = V \qquad 0 \le t \le T_b \tag{B.4}$$

and

$$H(f) = V T_b \ \text{sinc}(f T_b) e^{-j\pi f T_b} \tag{B.5}$$

The filter output y(t) due to g(t) is

$$y(t) = g(t) * h(t) \tag{B.6}$$

Thus

$$Y(f) = G(f) \cdot H(f) \tag{B.7}$$
$$= V^2 T_b^2 \ \text{sinc}(f T_b) e^{-j 2\pi f T_b}$$

and

$$y(t) = \begin{cases} V^2 t & 0 \le t \le T_b \\ V^2 T_b \left(2 - \frac{t}{T_b}\right) & T_b \le t \le 2T_b \end{cases} \tag{B.8}$$

Sampling the output at $t = T_b$ yields

$$y(t)\Big|_{t=T_b} = V^2 T_b \quad = \text{Energy in basic pulse} \tag{B.9}$$

The output of the filter due to the noise cannot be calculated exactly but will be Gaussian with variance

$$\sigma^2 = \int_{-\infty}^{\infty} \frac{N_o}{2} |H(f)|^2 df = \frac{N_o V^2 T_b}{2} \tag{B.10}$$

For Polar signaling either g(t) or -g(t) is transmitted thus, $P_{AVG} = V^2$ and the performance equation (2.7) becomes

$$P[E_R] = Q\left[\frac{P_{AVG} T_b}{\sqrt{\frac{N_o P_{AVG} T_b}{2}}}\right] = Q\left[\sqrt{\frac{2 P_{AVG}}{N_o / T_b}}\right] \tag{B.11}$$

$$= Q\left[\sqrt{2 \cdot SNR}\right]$$

REFERENCES

1. R. C. Houts and J. D. Moore, "Characterization of Impulse Noise and Analysis of Its Effects Upon Correlation Receivers," Technical Report No. 135-102, Communication Systems Group, Bureau of Engineering Research, University of Alabama, October 1971.

2. J. L. Perry, "Signal Design and Detection," Report No. 2635-M6-01, SCI Electronics Inc., undated.

3. H. Nyguist, "Certain Topics in Telegraph Transmission Theory," Trans. AIEE, vol. 47, pp. 617-644, April 1928.

4. W. R. Bennett and J. R. Davey, Data Transmission. New York: McGraw-Hill, 1965.

5. W. R. Bennett, Introduction to Signal Transmission. New York: McGraw-Hill, 1970.

6. R. W. Lucky, J. Salz, and E. J. Weldon, Principles of Data Communication, New York: Mc Graw-Hill, 1968.

7. Mischa Schwartz, Information Transmission, Modulation, and Noise. New York: Mc Graw-Hill, 1959.

8. Charles R. Cahn, "Performance of Digital Phase-Modulation Communication Systems," IRE Trans. Commun. Systems, vol. CS-7, pp. 3-6, May 1959.

9. Charles R. Cahn, "Combined Digital Phase and Amplitude Modulation Communication Systems", IRE Trans. Commun. Systems, vol. CS-8, pp. 150-155, September 1960.

10. J. M. Wozencraft and J. M. Jacobs, Principles of Communication Engineering. New York: John Wiley and Sons, 1967.

1972

NASA - ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

OPTICAL PRODUCTION AND DETECTION OF ULTRASONIC WAVES IN METALS FOR NONDESTRUCTIVE TESTING

PREPARED BY:	RICHARD A. MORRISON
ACADEMIC RANK:	ASSISTANT PROFESSOR
UNIVERSITY:	TALLADEGA COLLEGE
LABORATORY:	ASTRONAUTICS: MATERIALS BRANCH
RESEARCH COUNTERPART:	W. N. CLOTFELTER
DATE:	AUGUST 11, 1972
CONTRACT NUMBER:	NGT-01-003-045

Abstract

Optical Production and Detection of Ultrasonic Waves in Metals for Nondestructive Testing

by

Richard A. Morrison

Ultrasonic waves have been produced by striking the surface of a metal with the focused one-joule pulse of a Q-switched ruby laser. Rayleigh (surface) waves, and longitudinal waves have been detected with conventional transducers.

Optical methods of detection have been tested and developed. Rayleigh waves were produced with an oscillator and transducer. They have been optically detected on curved polished surfaces, and on unpolished surfaces. The technique uses a knife edge to detect small angle changes of the surface as the wave pulse passes the illuminated spot. Optical flaw detection using pulse echo and attenuation is demonstrated.

Optical detection methods are used to study the Q-switched ruby pulse. It is concluded that focusing of the ruby pulse with an inexpensive cylindrical lens leaves shallower scars on the surface, and produces better directed waves. It is strongly suggested that the Q-switched pulse is slow and limits the frequency response of the system.

INTRODUCTION

Ultrasonic source waves are widely used for the nondestruction detection of flaws in materials. These sound waves are usually generated and detected with piezoelectric transducers. It should also be possible to generate and detect these waves using laser beams. When the optical technique is convenient, it will have certain advantages over the technique using transducers. The surface must be flat and reasonably smooth in order to couple a transducer to a metal surface. In contrast to this, a well-focused, Q-switched laser pulse will induce ultrasonic waves in a surface of arbitrary shape. The laser beam also eliminates the necessity of making electrical convections to the specimen. It is also possible to detect ultrasonic waves optically . Optical detection requires a polished surface, but hopefully does not require a flat surface. Optical detection also is inherently a broad band detection technique as contrasted to piezoelectric techniques, which detect the waves by coupling a resonant device to the metal. The resonant detector suffers from uncertain losses in the couplant and unknown phase effects. It therefore detects a distorted, filtered, version of the original pulse. Optical detection techniques require no couplant and should detect a more broad band of frequencies than the usual narrow-band detector. It should be possible to reproduce the sonic pulse more accurately using optical detection.

Recently, accurate determinations of the velocity of ultrasonic surface (Rayleigh) waves have been used to nondestructively detect stresses in metals. Optical techniques hold the promise for this test technique, especially on a specimen without convenient flat surfaces.

This experiment was divided into three parts. In the first part, a Q-switched laser has induced ultrasonic waves in aluminum specimens. The Rayleigh waves were first detected with quartz Y-cut transducers. Longitudinal waves were also detected. An attempt was made to detect transverse waves, with negative results. In the second part, Rayleigh waves were generated with a transducer and were detected optically on various surfaces using a He-Ne continuous wave laser. Optical flaw detection was demonstrated. Surface waves were successfully detected on an unpolished surface. In this part, careful study was also made of the instrumentation required to make very accurate velocity measurements, optically, for the nondestructive evaluation of residual stress in metals. In the third part, the optical detection techniques were used to study the characteristics of the Rayleigh-wave pulses induced by the ruby laser pulse.

Experiments

To determine if the acoustic waves induced by the Q-switched pulse of a ruby laser could be useful for nondestructive testing, several transducers were placed on a specimen on which the laser pulses were focused. Rayleigh waves were detected (Fig. 1) with a Y-cut quartz transducer, and a wave velocity measured (Fig. 2). The velocity was consistent with published velocities for these waves in aluminum. A longitudinal wave detector was then placed on a surface perpendicular to a line between the detector and the source. Waves were detected. The computed velocity agreed well with published values for longitudinal waves in aluminum.

A transverse-wave detector was mounted, with a viscous couplant, on the specimen. A wave was detected. However the velocity for this wave was identical to the Rayleigh wave velocity. To verify that this wave was actually the Rayleigh wave, both a Rayleigh wave detector and a transverse wave detector were mounted next to each other. (Fig. 3) The travel distance between the source and the two detectors was then varied by moving the source spot away from the detectors. As the distance between the source and the detector increased, the travel time to both detectors increased by exactly the same amount. Transverse waves in aluminum are expected to have a velocity about 10% less than the velocity of surface waves. A 10% difference in velocity would have been easily measurable with this experiment.

If the major effect of the laser pulse is a sudden local heating of the specimen, it is reasonable to assume that shear waves would be less strongly produced than Rayleigh waves and longitudinal waves. The lack of detectable shear waves is consistent with this hypothesis.

Part II

In this section, an ARulab PG650 C pulsed oscillator with a Y-cut quartz transducer was used to generate Rayleigh waves in specimens, and the waves were detected optically. Two methods were tried. In the first method (Fig. 4) the beam from the laser is reflected at normal incidence from the specimen back into the laser. The beam from the other end of the laser was incident on a photodiode, amplified and displayed on an oscilloscope. With this method the pulse could be detected, but with a fluctuating amplitude and phase. It was not possible to prevent the fluctuations. In fact, in subsequent attempts, it was important to prevent the reflected beam from returning to the laser in order to stop this unpredictable oscillation from interfering with the desired signal. The method finally adopted, used a knife edge. In this method (Fig. 5) the beam reflected from the specimen is partially intercepted by a knife edge (a razor blade). The beam that passes the knife edge is focused onto a photodiode. As the wave pulse passes the spot at which the laser beam is reflected,

there will be a small angle change of the reflecting surface causing a small fluctuation in the amount of light blocked by the knife edge. In order to enhance the effect, the photodiode signal was filtered, amplified, and filtered again before entering the oscilloscope (Fig. 6). Using this signal handling 7 MHZ signals of several volts were observed, with an electronic noise of less than 5 mv. A 15 mw He-Ne laser was used. As mentioned before, care was taken to prevent the reflected beam from re-entering the laser since the large uncontrolled oscillations that resulted drove the first preamplifier out of its linear range. If the surface was well polished, no lenses were required to produce a strong signal. If the surface was not well polished, it was helpful to use a converging lens to bring the reflected rays to a spot small enough to reach the photodiode. Good signals have been received with lenses of focal lengths from 7.5 cm to less than 1/2 cm. The latter were microscope objectives. These were used with diffusely - scattering specimens to gather enough of the scattered light. They were placed to focus the incident light on the surface and collimate the reflected beam, (Fig. 5). With a microscope objective, however, it was difficult to prevent some of the reflected light from re-entering the laser. The technique is simple to set up and is sensitive enough to detect small scratches. (Fig. 7) It was detected waves on the curved surface of a 3-inch diameter cylinder (with a flat portion on which the transducer was mounted.) (Fig. 8) A 5 cm focal length converging lens was used to collect the reflected rays into a small area on the knife edge. The ease with which this experiment was accomplished strongly suggests that this technique can detect signals on polished surfaces of odd shapes.

An attempt was made to detect acoustical waves in an unpolished specimen. The difficulty here is that little of the reflected light is reflected speculatively. A 10X microscope objective focused the light on the plate and collected some of the scattered light. A beam splitter reflected part of the return beam onto the knife edge. Acoustical waves were detected, but with only a 3:1 signal-to-noise ratio. A more powerful laser will increase the amount of light (and, therefore, the signal amplitude) that reaches the photodiode. Also more careful placement of mirror can replace the beam splitter with full mirrors, increasing the amount of light by a factor of at least 4 times.

Flaw detection was demonstrated on a surface with a scratch. The pulse echo from the scratch was detected in one experiment (Fig. 7a). The attenuation of the surface wave was detected in a second experiment. (Fig. 7b). In the attenuation experiment, the wave is detected before crossing the scratch, and after. A clear amplitude change was noted. The two measurements were done in surfaces of similar reflectivity.

An attempt has been made to detect the applied stress on an aluminum specimen with this method. This technique involves the measurement of the velocity between two points on the surface of a specimen. The specimen is then stressed, and the new travel time is noted. A tensile stress is expected to decrease the velocity of these waves by a very small percentage. The change in travel time expected is approximately 1 nsec/KSI/inch of travel distance. That is, for each thousand pounds per square inch of pressure and each inch of distance, a time shift of about 1 nsec can be expected. The total travel time/inch is about 9 μsec. Thus a relative change of $1.10^{-3}/9 \approx 10^{-4}$ must be measured. In order to do this the laser beam was split into two parts, which were reflected from two places on the specimen (Fig. 9). Both reflected beams were adjusted to hit the knife edge and enter the photodiode. A 7-Mhz transducer was used. The pulse was detected first in one reflected beam and then in the other (Fig.11). Delayed triggers from the oscillater and from another oscilloscope were used to provide successive trigger pulses for the display. With the proper timing provided by the variable delay circuitry, it was possible to display the pulses simultaneously on an expanded time scale (Fig. 11). Microscope objectives were used to provide a small spot size on the specimen. The small spot size was necessary to define the distance between the two spots unambiguously. The specimen was mounted between two large bolts, which were mounted in a frame. After the variable delays were set to line up the two pulses on the display, the specimen was stressed. Strain was measured by resistive strain gauges mounted between the beam spots, above and below them. Using a modulus of elasticity of 10.6 KSI (μinch/inch.), the stress could be calculated. Stresses up to 12 KSI were used. As the specimen was stressed, the position and angle of the surface changed. This changed the angle of the reflected beam. Any re-alignment of the mirrors, however, would change the spot position on the specimen. Therefore, the knife edge photodiode was moved to best intercept the individual beams, and separate exposures were taken of each beam on the same photograph. (Fig. 11). In this way, the relative travel time could be determined. This experiment did not show consistent results. Several reasons suggest themselves:

(1) The beam spot size perhaps was too big. If this is so, then as the angle changes, the knife edge intercepts the light from different positions on the specimen. This problem is easy to test for and eliminate. Simply move the knife edge around. If, near maximum, the pattern always has the same timing, then the spot is small enough. Microscope objectives with ≈1 cm focal length provided a small enough spot when positioned correctly.

(2) When the plate is stressed, it moves slightly out of focus and the beam spot becomes too large for accuracy. This

possibility was rejected by the method suggested in #1.

(3) The beam did not hit the plate at normal incidence (if it did, the reflected beam would travel back along its original route, reentering the laser. This must be avoided). Therefore, as the plate moved during stressing, the spot position moved. An attempt was made to minimize this error by permitting only a vertical displacement of the light beam. Since the surface wave travels in a horizontal direction, a vertical displacement of the sensing point should not change the travel time. However, it was not possible to determine if all horizontal displacements were eliminated. This may have been the source of error.

(4) Electronic noise. The cleanness and reproducibility of results - as long as the optics was not touched - shows that problems in the electronics were not the source of error.

Part III

In this part acoustic waves were both generated and detected optically. Transducers and oscillators were not used. In order to take advantage of the broadband nature of this method, the tuned filters were first discarded and the signal was observed at the output of the preamplifier (after an RC high-pass filter with a 500 KHz time constant). The ruby laser was located ≈ 4 m from the impact point. Without some focusing of the ruby light, no pulse was observed. With focusing by a 5 cm lens, a pulse was observed (Fig. 12). The shape varied with the size of the beam spot and the power output from the laser. A rapid (≤ 10 nsec) risetime for this pulse was expected. Focusing did not effect the rise time critically. The electronics are capable of a 30 nsec risetime. (See App 1) Despite this, the pulse risetime was 100 nsec. In order to do the accurate timing, a shorter risetime is desired. The reason for the long risetime is not well understood. Preliminary tests, however, suggest that the risetime of the ruby pulse is $\gtrsim$ 50 nsec. These tests were done with two photodiodes, both nominally capable of 4 nsec. risetime. They observed the Q-switched pulse. Their outputs went directly to a Tektronics type 7704 oscilloscope. Their signals agreed, thus verifying that the photodiodes were not at fault. It may be possible to tune the pockels cell supply voltage, timing, and alignment to give a faster rise time.

In order to demonstrate the sensitivity of the purely optical system, the tuned amplifier was reinserted after the amplifier and tuned for maximum response. The resulting oscilloscope pattern showed considerable detail, due probably to the complex echoes coming from all directions (Fig. 13). Two experiments were tried to simplify the pattern. A metal comb was inserted in the beam. This comb, which

alternated absorbtion and transmission in intervals of 1/2 nm was designed to direct waves of 10 MHZ. No signal was observed. In a second attempt, an inexpensive cylindrical lens with a 3 cm focal length was used to focus the ruby light. This produced a major simplification of the echo pattern, presumably because of the better defined direction of the wave pulse. (Fig. 14)

Conclusions

It is necessary to determine if these optical techniques are useful in a program of nondestructive testing of metals.

The basic detection technique is easy to set up and lens placement is not critical, if lenses are necessary at all. The only requirement of a lens is that it concentrate as much light as possible at the knife edge so that light not intercepted by the knife edge enters the photodiode.

Suggestions for Increased Sensitivity:

(a) Mount the 150 pf. capacitor in the preamp chassis so as to minimize capacity of ground. Connect the preamp directly to the Lite-Mike. Then mount the Lite-Mike preamp combination on a plate that can be moved from side to side with a fine screen - without an angle change. Adjust this position for maximum signal. Be sure the beam is cut off by the knife edge, not the edge of the photodiode. Use a sharp knife edge.

(b) Increase the power of the He-Ne laser, and use a large area photodiode that can safely withstand the extra power Do not focus the light to a point on the photodiode; merely reduce it in size so that the light does not intercept the edges of the diode. The less-concentrated light is safer to the diode. If the laser is made more powerful, this method should work on unpolished surfaces.

This detection technique is, I believe, usable now for flaw detection on curved polished surfaces. The polish need not be perfect.

The ruby laser does slightly scar the surface of the metal. With an ordinary lens it leaves a small pin hole. As the ruby continues to enlarge the pin hole, the wave amplitude decreases. To increase the amplitude, move to a fresh spot. The cylindrical lens is more effective in directing the wave where it is wanted and leaves a much shallower line on the metal. If the beam is deliberately kept slightly defocused, the wave amplitude is not degraded too much, and

the scar can be avoided. This lens is highly recommended.

On a curved surface, the ruby may be the best method for inducing surface waves in the material. A reliable, simple, laser is recommended for this purpose.

If optical detection is done with the ruby, considerable attention must be paid to light leaks. All leaks in the body of the Lite-Mike must be sealed and a cover should be placed around the detection optics. This can be jury-rigged of cardboard for the test being done. Small holes usually admit a large flash from the Q-switched pulse anyway. This is a convenient trigger for the oscilloscope. It is also more accurate than the sync pulse from the Pockels cell power supply. The cover must be sufficiently good to eliminate any of the smaller laser flashes, after the big one. These show up at varying times and confuse the data.

For the stress measurement, several recommendations can be made:

(a) The He-Ne laser power should be doubled or tripled.

(b) Two lite mikes, with the associated electronics, should be used.

(c) A flexible 3-dimensional optical structure should be purchased, one capable of mounting lenses, etc., in various places. It should be possible to move lenses along the line of sight without other displacement or angle change. It should be possible to displace the knife edge across the beam line without other changes.

(d) Some provision must be made for accurately measuring the wave travel distance between the two detection points.

(e) The object must be stressed - without effecting the other optical components - with a minimum of surface movement of angle change.

This technique is of interest, but more work is required to make it of operational use.

Appendix I

Methods of Amplifying and Filtering Used in These Experiments:

The smallness of the signals emitted from the photodiode required low-noise amplification. An ORTEC 109A preamplifier with an FET front end was used. This preamp must be protected from the large, low frequency oscillations due to vibrations of the various optical parts, and due to oscillations of the laser output. Therefore, a

small capacitor (150 pf) was inserted, in a small BNC insertion box, between the photodiode and the preamp (Fig. 15). In order to eliminate excess capacity to ground, the ORTEC was connected directly to the insertion box. In a more permanent setup, the 150 pf capacitor should be connected inside the ORTEC chassis to eliminate to capacity of the insertion box. The rationale behind (a) the 150 pf high pass filter and, (b) the low capacity to ground are as follows:

(a) The preamp input resistance to ground, several hundred megohm, with the 150 pf capacitor form a high pass filter.

(b) In its most sensitive setting the photodiode (EG&G Lite-Mike # 560B) has an effective impedance of 5 KZ. This, with the stray capacity to ground of cabling, etc., form a low-pass filter, that seriously limits the frequency response of the electronics.

To determine if the frequency response of the lite-mike is the source of problems, switch from X100 to X10. The signal should be reduced by a factor of 10, if the Lite-Mike frequency response is not the limitation. To find if the X10 scale is giving adequate frequency response, switch from X10 to X1 and see if the signal goes down by a factor of 10. If so, then the X10 scale is OK.

When a broad band signal was desired, the preamp was connected to the oscilloscope through a high pass filter. When highest sensitivity was desired, the preamp was connected to an Arulab model PA620-SN tuned amplifier. The oscilloscope then observed the output of the amplifier.

An acceptable preamplifier, which slightly sacrifices frequency response and noise level for convenience, is the Panametrics Acoustic Emission Preamp.

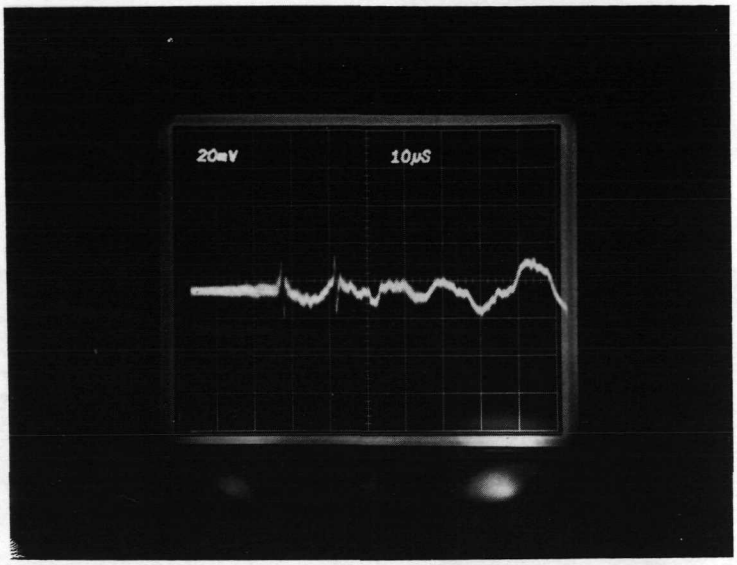


Fig. 1

Detection of Rayleigh Waves. The first sharp pulse is the original shock pulse from the laser. The second, smaller one is an echo from the far edge of the specimen. The slower noise are the body

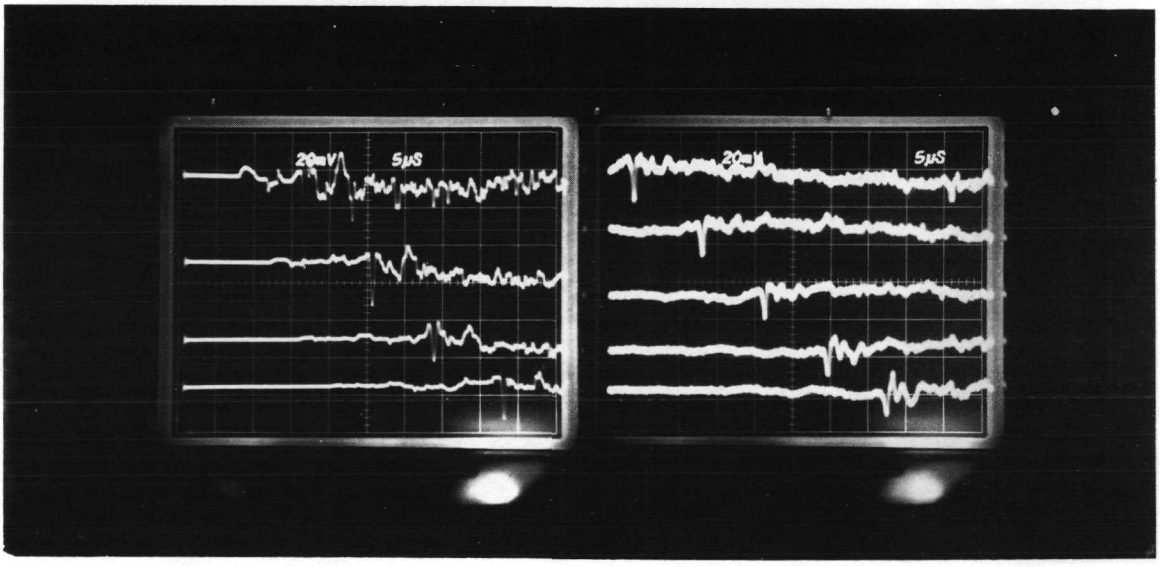


Fig. 2

Measurement of the Rayleigh wave velocity in an aluminum specimen. The detector was moved away from the impact point in 2.54 cm intervals. Successive traces show the time of arrival of the pulse at successive detector positions. The photograph on the right is a continuation of the photograph on the left. The oscilloscope sweep for this photograph was delayed by 47.2 sec to bring the pulses on scale.

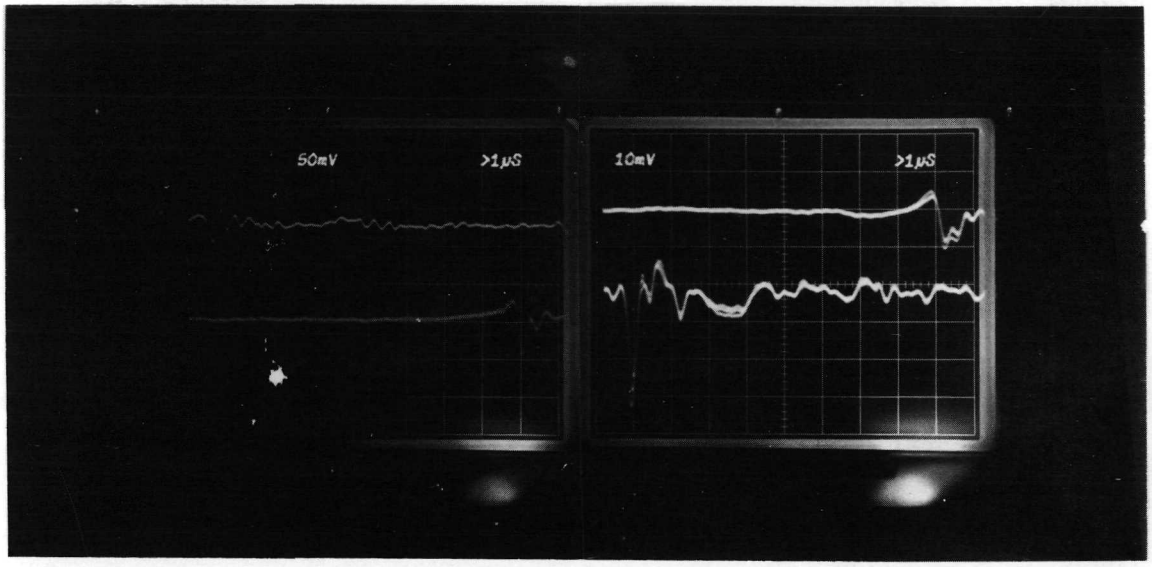


Fig. 3

Demonstration that the waves detected by the shear wave detector are Rayleigh Waves. The top trace in the left-hand photograph is the pulse in the Rayleigh Wave detector and the bottom one is from the shear-wave detector. When the impact point was moved ≈15 cm away from the detector, the presweep delay was increased, but the time difference between arrival times remained the same. Therefore, the waves traveled at the same velocity. The right-hand photograph shows (top and bottom interchanged) the time separation after the time increase. The sweep speed in the two photographs is the same.

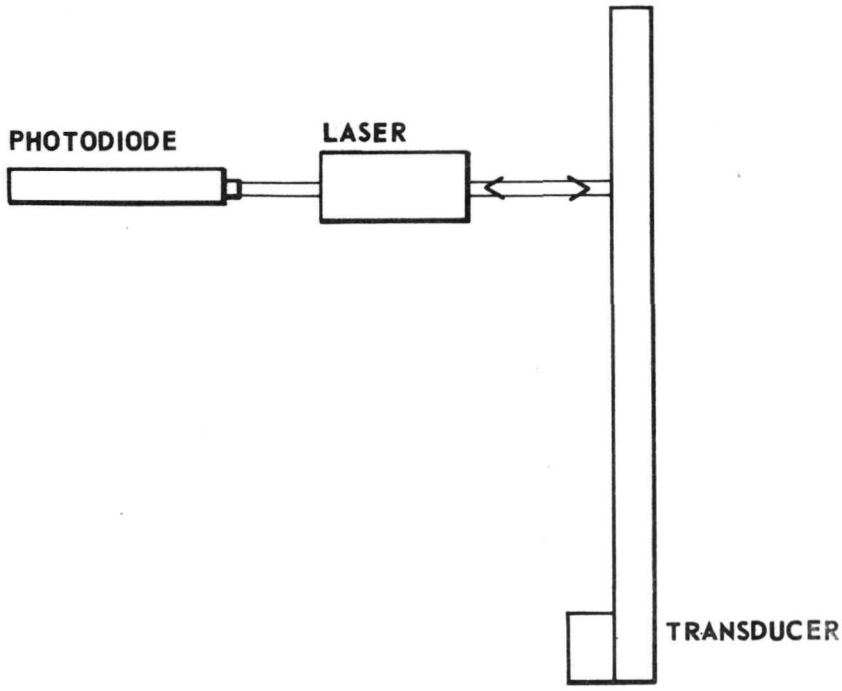


Fig. 4

Preliminary Optical Setup

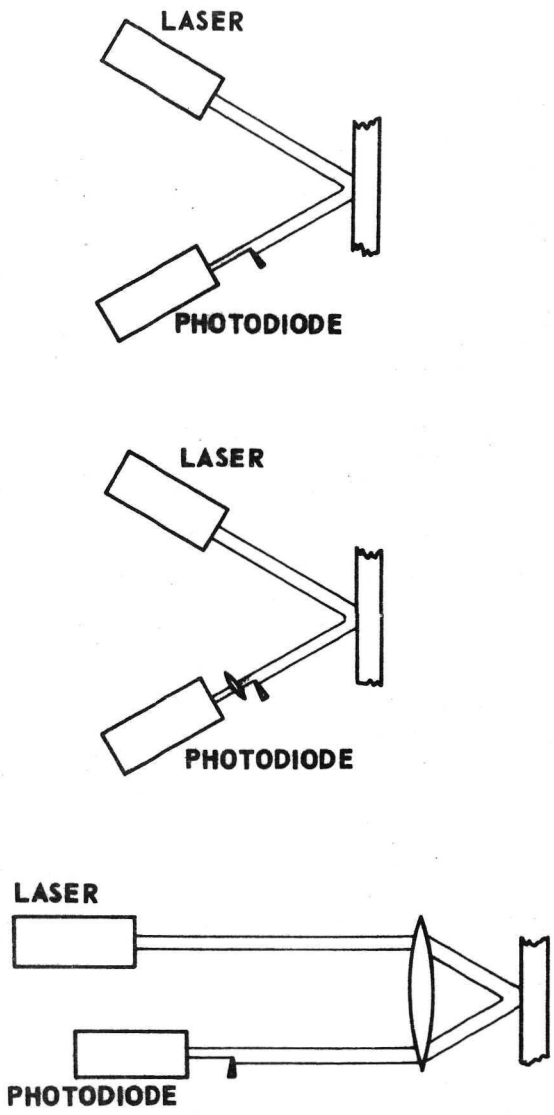


Fig. 5

Several forms of the knife-edge technique used in this experiment

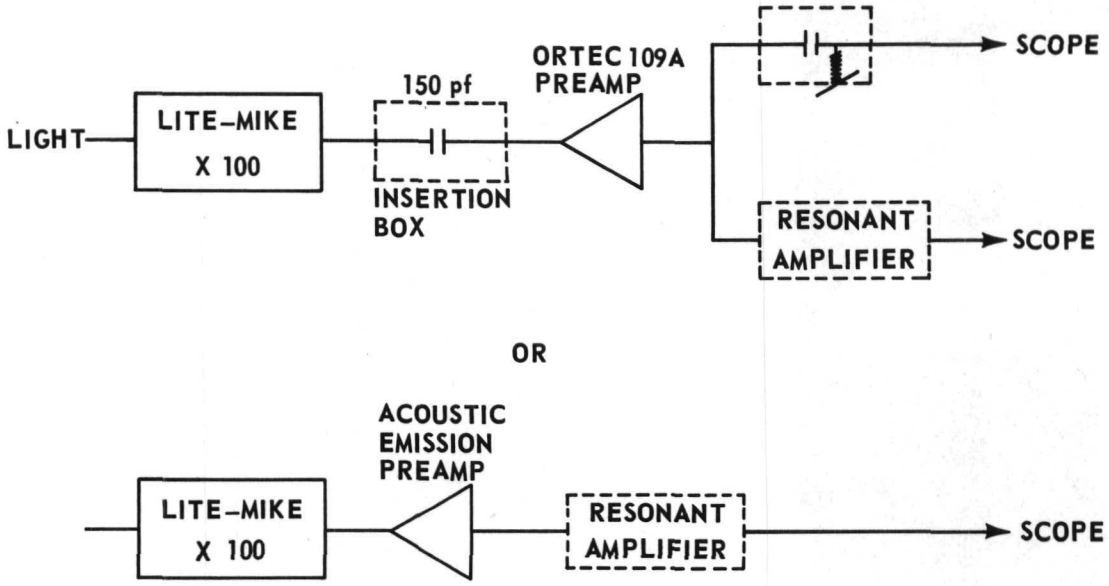


Fig. 6

Amplifiers and filters used in processing the optical signal

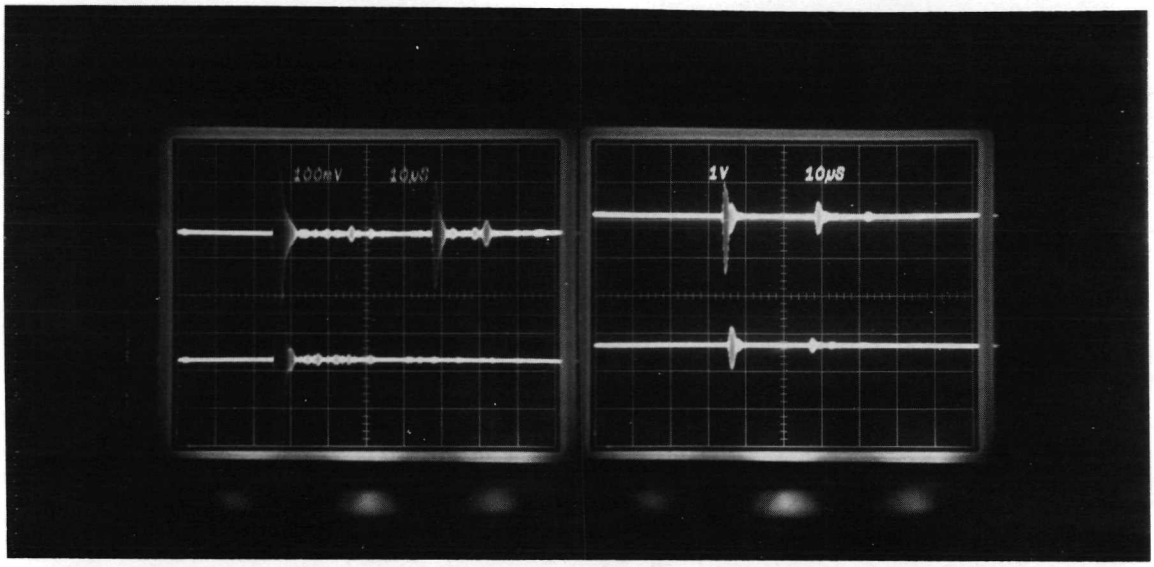


Fig. 7a - Fig. 7b

Demonstration of Optical Scratch Detection

Fig. 7a Pulse echo technique. Top trace. The first signal (at 25-30 μsec) is the original pulse. The second large pulse (at 68-70 sec) is the echo returning from the far edge of the specimen. The pulse at 46 sec is the echo from the scratch, which was about halfway between the sensing point and the far edge. Other, smaller pulses are due to other flaws on the surface. The pulse at 82 sec is the echo from waves that turn the far corner and reflect from scratches on the edge. Bottom trace: Half-inch-wide masking tape barely covered the scratch and severely attenuated penetration of waves beyond it.

Fig. 7b Attenuation method. The top trace comes from just before the wave reaches the scratch, the bottom trace from just after. This method must be used with caution. Points of good polish in both places should be used, otherwise the signal may be attenuated by lack of polish.

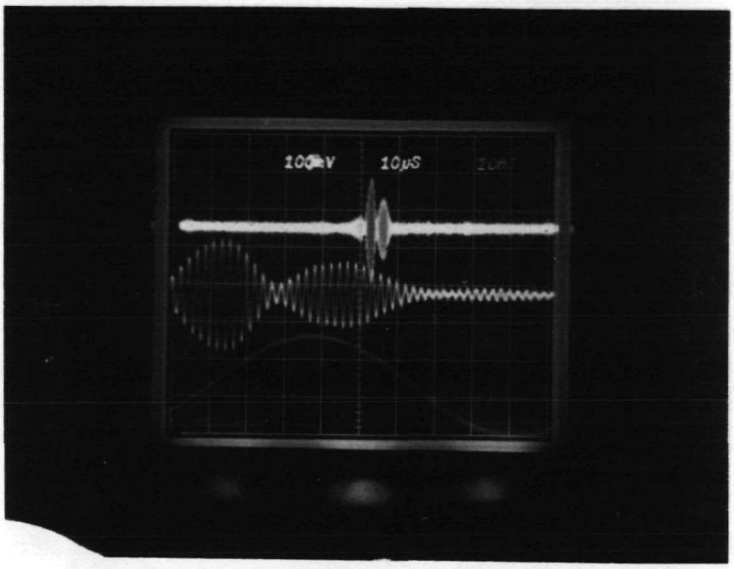


Fig. 8

Fig. 8a Several traces showing a signal from a curved specimen. First trace is at 10 μsec/division, second at 1 μsec/division, and third at 20 nsec/division.

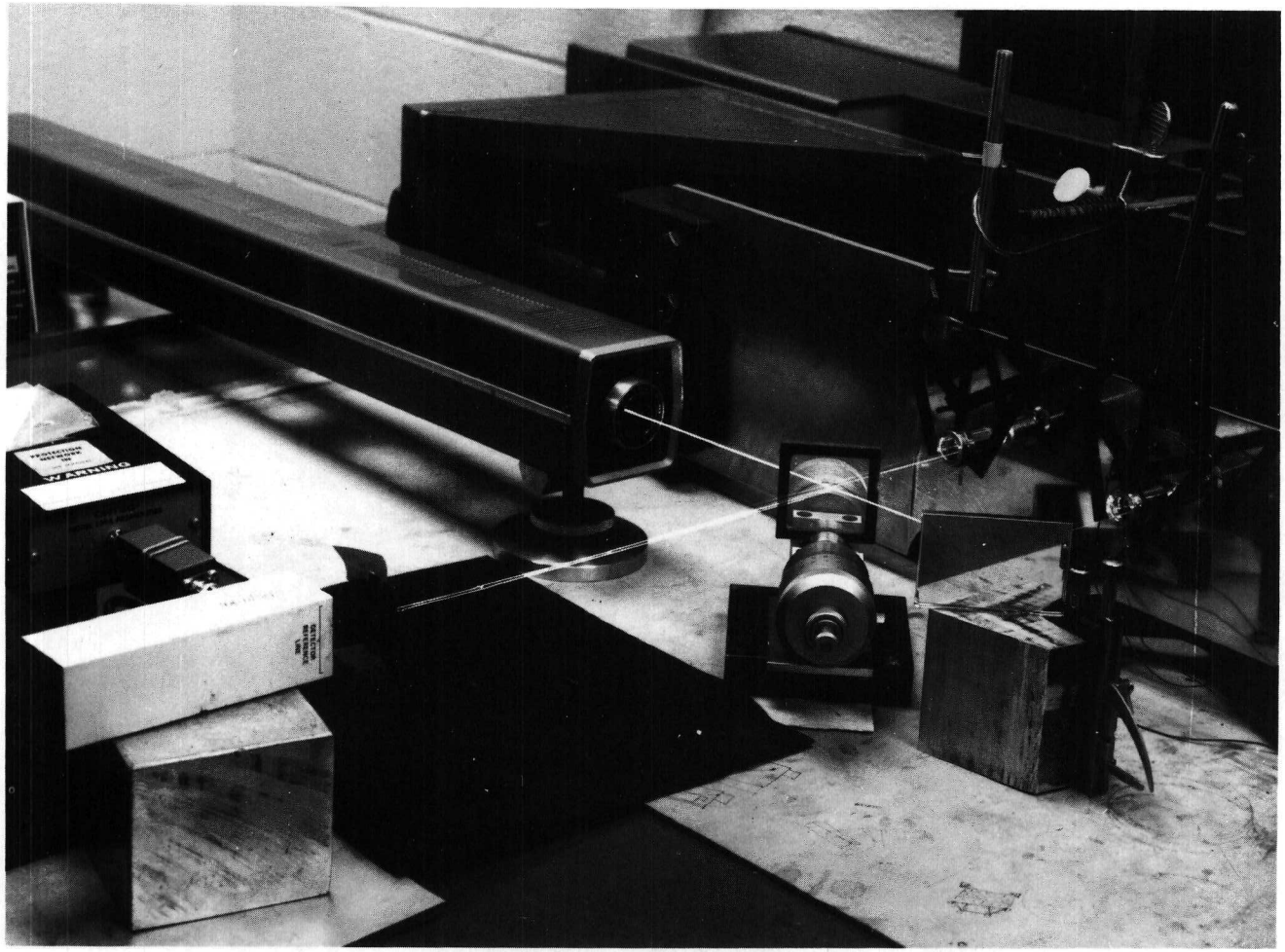


Fig. 9

A photograph of the intrinsic stress experiment. The razor blade is taped to the lens-holder for the photodiode.

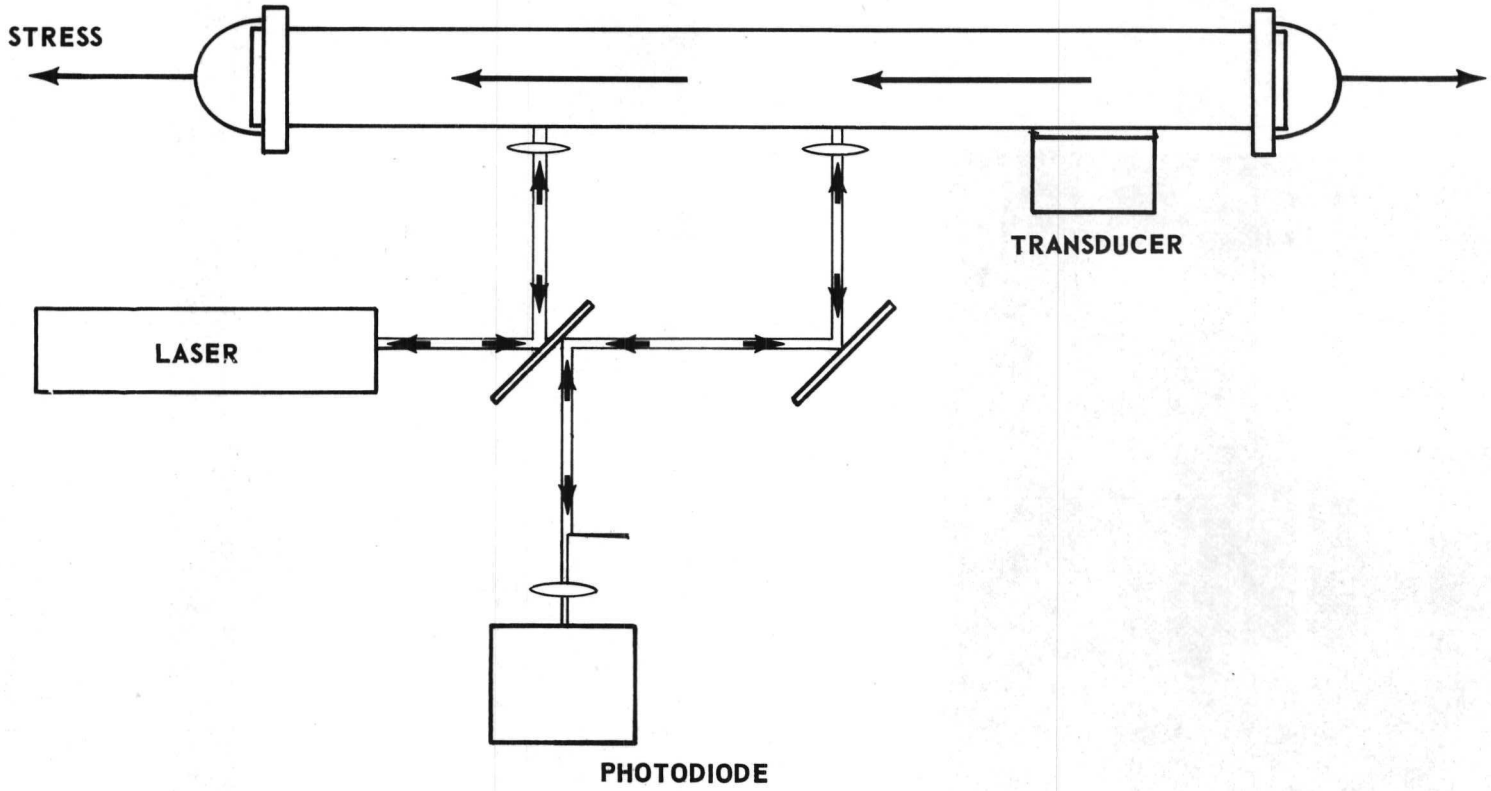


Fig. 10

Diagram of the optics of the intrinsic stress experiment

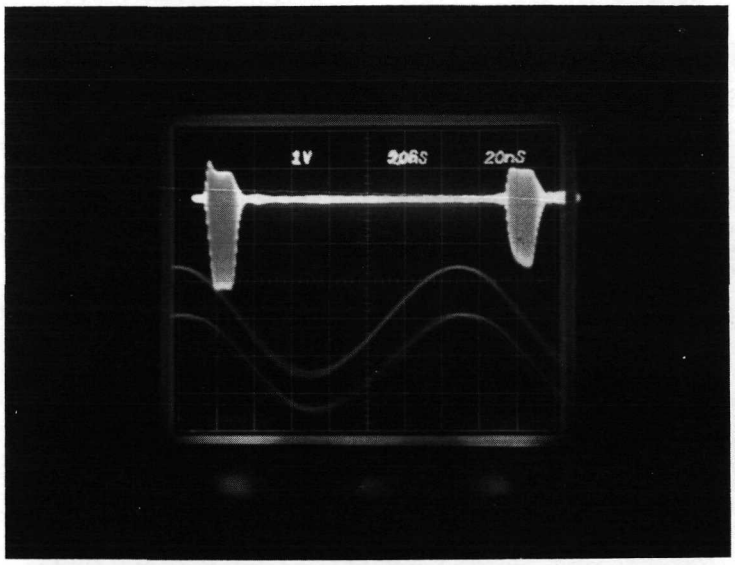


Fig. 11a Oscilloscope Trace

Showing an example of the first and second signals. Time scales are 5 μsec and 20 nsec per division

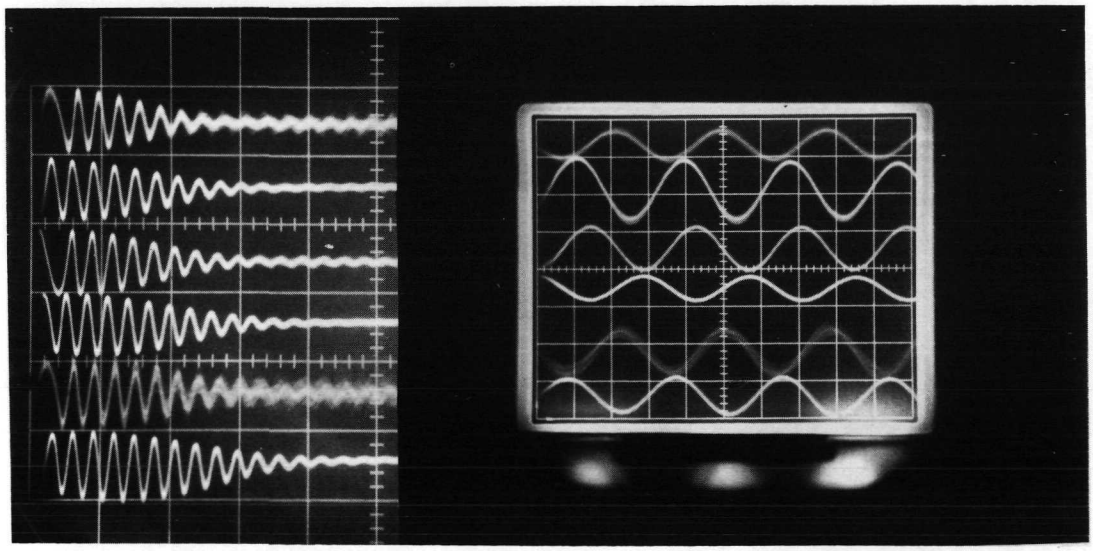


Fig. 11b

Scope traces for this experiment. These pictures were the data. They were taken with two oscilloscopes operating at the same time. The traces alternate between first beam traversed by the sound wave and second beam, the top pair at 3 KSI, the middle pair at 6 KSI, and the bottom pair at 9 KSI. The left and right photographs show the same signals, the left at 500 nsec and the right at 50 nsec per scale division.

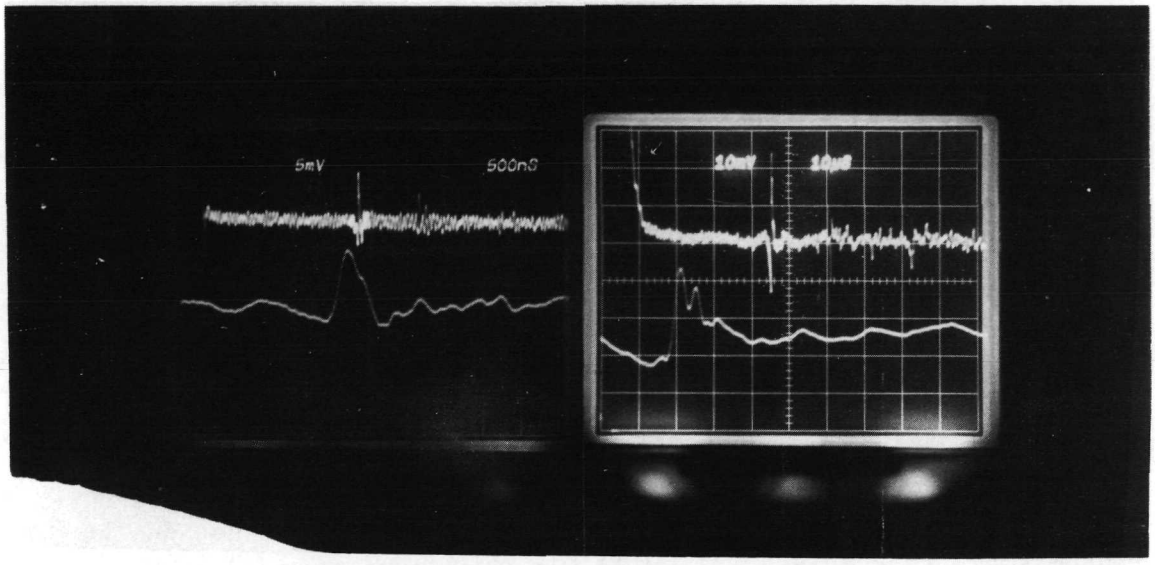


Fig. 12

The shape of the acoustic pulse produced by the Q-switched laser. The difference between these two traces occurred when the laser power was increased, the right one at higher power in the ruby.

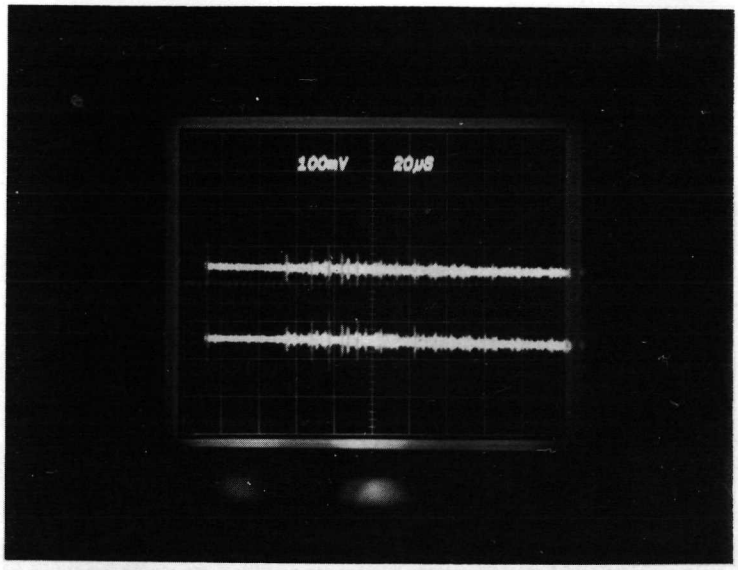


Fig. 13

The acoustic pattern received after the focused beam is fired. Two traces were taken to demonstrate that the peaks are not due to random small ruby pulses following the main burst.

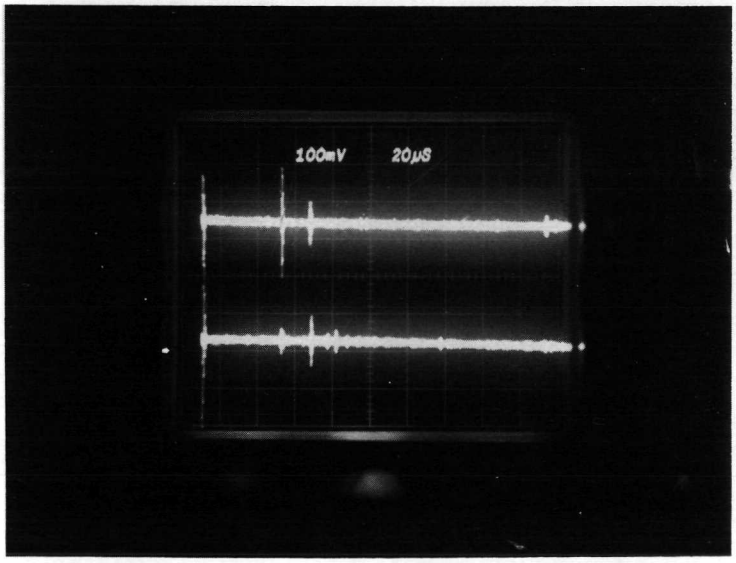


Fig. 14

The echo pattern received after the cylindrically focused beam is fired. This picture is easier to interpret. The two traces show the effect of small changes in knife-edge position. The pulse at 26 sec is the direct shock, the one 8 μsec later, the first echo from the far corner.

1972

ASEE-NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERISTY – UNIVERSITY OF ALABAMA)

AN ANALYSIS OF TEXTURE ON LUNAR GROUND PHOTOS

Prepared by:	Judy Olson
Academic Rank:	Assistant Professor
University:	University of Georgia
Laboratory: (Division) (Branch)	Aero-Astrodynamics Laboratory Aerospace Environment Division Environments Research Office
Research Counterpart:	Otha H. Vaughan, Jr.
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

LIST OF FIGURES

ACKNOWLEDGMENTS

Special thanks are extended to the NASA/ASEE Summer Faculty Fellowship Program and this year's Director, Mr. J. Fred O'Brien, for facilitating this project; to Mr. O.H. Vaughan, Jr., my NASA counterpart; to Mr. Doug Thomas, who performed the photographic scanning; and to Ellen Williams for her invaluable help with problems in programming.

AN ANALYSIS OF TEXTURE ON LUNAR GROUND PHOTOS

By

Judy Olson

ABSTRACT

Texture is one of the major aspects of a single lunar ground photo giving information about the distance to features. If texture is measurable and its relation to distance known, it could perhaps be incorporated into simulations of views from particular ground locations.

An experiment was performed in an attempt to evaluate autocorrelation as an indicator of texture using five small patches representing varying distances on each of four lunar photos. Each patch was scanned at 50 micron increments on a 64-level gray scale. Several problems were involved in using the ordinary autocorrelation value but a rougher "autocorrelation" measure, the percentage of neighboring pairs (at a given lag distance) which fell within one graylevel of one another, yielded very encouraging results. As the distance to the feature decreased, initial slope of the graph of percentage against lag increased and the general level of the graph decreased. When a crestline was present, the graph tended to continue to decrease at higher lags as well, rather than leveling off after the initial steep slope.

Further study is needed to relate such characteristics to more precisely known distances, to look into the effects of variables such as lighting conditions, and to find a feasible method of incorporating the feature into simulated views.

AN ANALYSIS OF TEXTURE ON LUNAR GROUND PHOTOS

INTRODUCTION

This study of texture on lunar ground photos developed from an interest in realistic, or visually effective, representations of the lunar surface.[1] There are many different ways of graphically representing an undulating surface to convey to the viewer its essential characteristics – its relief and slopes, the spatial arrangement of high and low points, and so on. The choice of method is dependent upon the purpose for which the representation is intended. Visual representations, i.e., those which quickly convey the general nature of the surface, are not substitutes for the highly accurate and versatile topographic map but supplement it and allow for efficient communication of aspects which are not readily conveyed by these more abstract devices. Generally speaking, visual representations are but transformations of the information contained on the topographic map.

One type of realistic representation of a surface is the view from a particular ground location, as it would be seen by an observer or recorded on photographic film. This type of representation is useful in becoming acquainted with the look of an area, in recognizing location, and in identifying specific features. It would be impossible to generate a view in its exact detail simply from the information on a topographic map of the area. In examining actual ground photos, however, there are basically three characteristics which combine to convey information to the viewer about the general nature of the area. These three characteristics are (1) the arrangements of crestlines, (2) the general light and shade conditions, and (3) the texture (see Figure 1).

It was the last of these, texture, which was chosen for study in this project. The purpose of the study was to examine the autocorrelation function as a measure of texture and to determine whether the apparent change in texture on the photo with increasing distance from the point of observation can be detected through the use of the measured autocorrelation values. Thus, the problem was one of automatic pattern recognition (not recognition of objects but of a certain aspect of the scene, distance). One of the implications of a highly patterned behavior of the autocorrelation function over changing distances from the point of observation would be the possibility of simulating texture in generating views from particular points, since texture would then be a function of the information contained on the topographic map of the area. Another possible implication would be that distance on actual photos might be more predictable from this measure; it is well known by those who have worked with the Apollo missions that distance to visible features is very difficult to judge on the lunar surface and on lunar ground photos. The actual attainment of either of these goals was, of course, beyond the scope of this limited project. It was hoped, however, that the results would at least be a step along the way and that they would contribute to an understanding of the relationship between a measurable aspect of the photograph and a visual aspect.

THE PROPOSED TEXTURE INDICATOR – AUTOCORRELATION

"Texture" is a concept and is not easily defined nor is it easily measured. Although the concept will not be discussed in detail here[2], it is useful to note that certainly texture on the photograph has something to do with the spatial arrangement of the graytones, i.e., with the relationship between the values of graytones to the values of neighboring graytones. Since autocorrelation is a measure of relationship between values and their neighbors at some defined distance (or lag), it seemed quite plausible that this technique would be useful for measuring that characteristic of the photograph which differs as the distance to the feature (from the observation point) changes.

There are, of course, an infinite number of lag distances on the flat photo for which the autocorrelation coefficient can be calculated, and even if only a part of the autocorrelation function is calculated, it is no longer a simple (one-valued) indicator of texture. It could also be measured in any direction or combination of directions. It seems reasonable, however, that just a few characteristics associated with autocorrelation might be highly related to distance from the viewer. These characteristics include the general level of the autocorrelation values over short lag distances; the rate of decay of the function; the lag distance at which the autocorrelation value falls to a certain value; and the lag distance at which the function begins to repeat if, indeed, it is a repeating function.

It was difficult to predict just what the relationship between these characteristics and distance from the camera would be, and thus, the study was intended as a somewhat exploratory one. It would be expected, of course, that the level of autocorrelation and the smoothness of the function would increase with increasing distance from the camera, since the photo scale would be decreased and hence the graytone over any finite area on the photo would represent an aggregation, and thus a smoothing out, of the various tones on the corresponding portion of the lunar surface. The exact nature of these changes, however, was not so easily predictable. There was also the possibility that the many uncontrolled aspects and those whose effects simply had not been considered would greatly affect results. With an exploratory approach, an open mind would be maintained and adjustments in the texture analysis could be made as the project progressed.

In the literature, autocorrelation and similar techniques (including Fourier transformations) have been utilized in a number of situations which are somewhat related in nature to the present problem. In pattern recognition they have been used in distinguishing between alphanumeric digits[3] as well as for experiments in automatic terrain classification[4] and for analyzing precipitation patterns.[5] In psychology, they have been used in studying pattern discrimination[6] and texture.[7] With the evidence at hand and with the facilities available for scanning the photos and recording the graytone values,[8] it was felt that such an approach to analyzing this major photo characteristic, texture, would be a useful one.

THE EXPERIMENT

Four lunar ground photos were chosen for studying including two from Apollo 15 and two from Apollo 16. These photos are shown in Figure 2 and are described briefly in the caption. They were chosen primarily for the variety of distances represented on them.

Prints of these photos were available at MSFC, from which approximately 4 × 5-inch negatives were made for the purpose of scanning (this reduced scale was the same as represented in Figure 2). On each negative, five patches ½ × ½ cm. in size, representing various distances from the camera, were chosen for analysis. Some of these patches were deliberately chosen such that a crestline passed through for the purpose of observing the effects of this rather abrupt transition of distances. Each of these 20 patches was then scanned at an increment of 50 microns producing a 100 × 100 matrix of values on a 64-level graytone scale. Due to a minor problem, it was necessary to eliminate the last 6 rows, resulting in a 94 × 100 matrix of data for each patch.

The autocorrelation values for lags one through 15 (representing lag distances of 50 microns to 0.75 mm on the 4 × 5-inch negatives) were computed for each patch. Simple product-moment correlation was applied to the data for each lag, the data being paired lists of values, one value being the graytone at a given point, its paired value being the graytone of its neighbor in a given direction at the given lag distance. Pairings were set up in several ways (horizontally, vertically, both, etc.) so that possible direction bias could be noted.

The results of this first attempt to measure autocorrelation were disappointing to say the least. The difficulties included the following: (1) The amount of data involved was not suited to the specific method used to calculate the coefficients. Values obtained intermediate to final results were extremely large and created handling problems in the programming. (2) The method was also unsuitable in that it was extremely slow and required far more computer time than anticipated. (3) Apparently (though not yet verified) the method was also unsuitable for the types of value variations which occur on such photos. For example, in cases where there is very little variation in value and thus a visually smooth texture (neighboring values highly related visually) the measured coefficient can be very low.

The first two of these problems would probably be lessened considerably with more efficient programming. The third, however, suggested that an alternative approach might be useful and an examination of previous output suggested an extremely simple one. Rather than measuring the relationship between neighboring values with the product-moment correlation procedure, the relationship was measured simply as the percentage of pairings (vertical pairings were used here) which were within one graylevel of one another. This "autocorrelation" measure was also calculated for lags one through 15. The same program was used as previously with the addition of one short loop and instructions to by-pass the unneeded procedures. The large-number problem and the computer-time problem were eliminated and the results were considerably more encouraging than previous ones.

Figure 3 presents the basic resulting information in graphic form. Each graph represents the results for one patch, with lag on the horizontal axis and percentage of pairs whose values were within one graylevel of one another on the vertical axis. Each column corresponds to one photo. The letters and numbers in the figure correspond to those in Figure 2 and Cr designates a crestline.

The behavior of the function represented here is quite well-patterned. Looking at the results in graphs 2, 3, and 5 for Photograph A there is a marked increase in the initial steepness of slope (i.e., the slope of the line over the first few lags) and a general decrease in

the level of the curve as the distance represented decreases (the top graph representing the most distance, the lower graph the least). Graphs 1 and 4 are from patches through which there was a crestline. The effect of the crestline seems to be that it distinctly descreases the similarity of values at the higher lags; there is less leveling off of the function after the initial steep slope.

For Photographs B and C, the same general characteristics are found including that graphs 6, 8, 11, and 13 (crestline patches) again show decreasing values for the higher lags. The crestline in the patches corresponding to graphs 11 and 1 was the horizon line; these graphs do seem to differ slightly from other crestline graphs in that they are high and the initial steep slope is very short.

The graphs for photograph D, with 17 being a crestline graph, also illustrate the same characteristics though it is immediately noticeable that the changes from one graph to another are weaker. The differences in distances for the patches represented, however, are not as great as those in the other photos and, thus, weaker contrast from graph to graph would be expected.

The examination of these curves seems to indicate that the values for lags approximately through the fourth yield information concerning distance, with the differences between values at these lags and those at higher lags yielding information as to whether a distinct crestline is present. Unfortunately, since exact distances are unknown, a precise comparison between photos is precluded. Comparing rough estimates with the values graphed, there seems to be some systematic behavior but not as much as within each photo separately. Probably lighting conditions and camera focus would need to be taken into account in comparing from photo to photo.

IMPLICATIONS FOR FURTHER RESEARCH

The results of this study suggest some very concrete steps for further inquiry. As a pattern recognition study it suggests that this simple "autocorrelation" measure might be useful for recognizing distances, crestlines, and possibly other features. Some of the variables which need attention include general lighting conditions (analyzing the interactive effects of lighting and texture) and film development effects. More immediate, however, are the needs (1) to relate precisely-known distances to the characteristics of this measurement, (2) to find those aspects of the function which best indicate the desired information and determine how reliably and accurately they perform, (3) to examine the effects of the scanning increment size on the usefulness of the measure, and (4) to explore the possibilities of a slightly revised measurement (such as the percentage of neighbors within two graytone values of one another or perhaps a standardized difference value for the pairs, the inclusion of pairs in all directions rather than just one, etc.). In addition, the assumption was made here that the surface characteristics on the moon were homogeneous and the effects of blockiness, distinct changes of composition, etc., should also be studied.

Concerning the implications for the inclusion of texture in transforming topographic information into the view from a particular point, the results suggest that, indeed, texture (a measurable expression of it) is a function of topographic information. With further study it

would seem quite possible to develop a method by which photographic manifestations of these functions could be generated. One of the greatest problems involved would probably be the simultaneous maintenance of the proper functions in more than one direction.

SUMMARY

This study has attempted to measure the varying textures on lunar ground photos which change with varying distances to the features in view. Autocorrelation was used as this measure but its calculation in the ordinary manner proved less useful than expected. A rougher but more easily calculated "autocorrelation" measure, the percentage of neighboring values within one graytone level of one another on a 64-level scale, yielded very encouraging results. Graphs of this function over fifteen lags for patches representing various distances from the viewing point showed increasing initial slopes and a lowering in general level as distances decreased. When crestlines passed through the analyzed area there tended to be a more noticeable decrease in values at higher lags.

The study suggests that this measure may be a useful one in automatic recognition of such aspects as distance and crestlines and that it may be quite possible to incorporate texture into simulated views of the lunar surface from a given ground location.

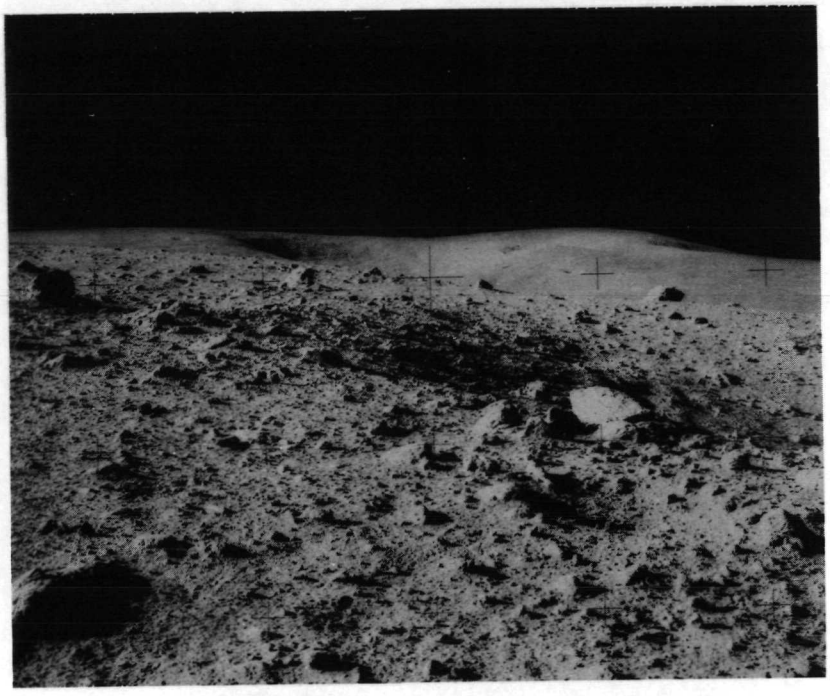


FIGURE 1. (FIGURE 1 FROM REFERENCE 1) A LUNAR GROUND PHOTOGRAPH ILLUSTRATING THE CRESTLINES, GENERAL LIGHT AND SHADING, AND TEXTURE EFFECTS. (POLAROID COPY OF A SECTION OF AS14-64-9117, APOLLO 14)

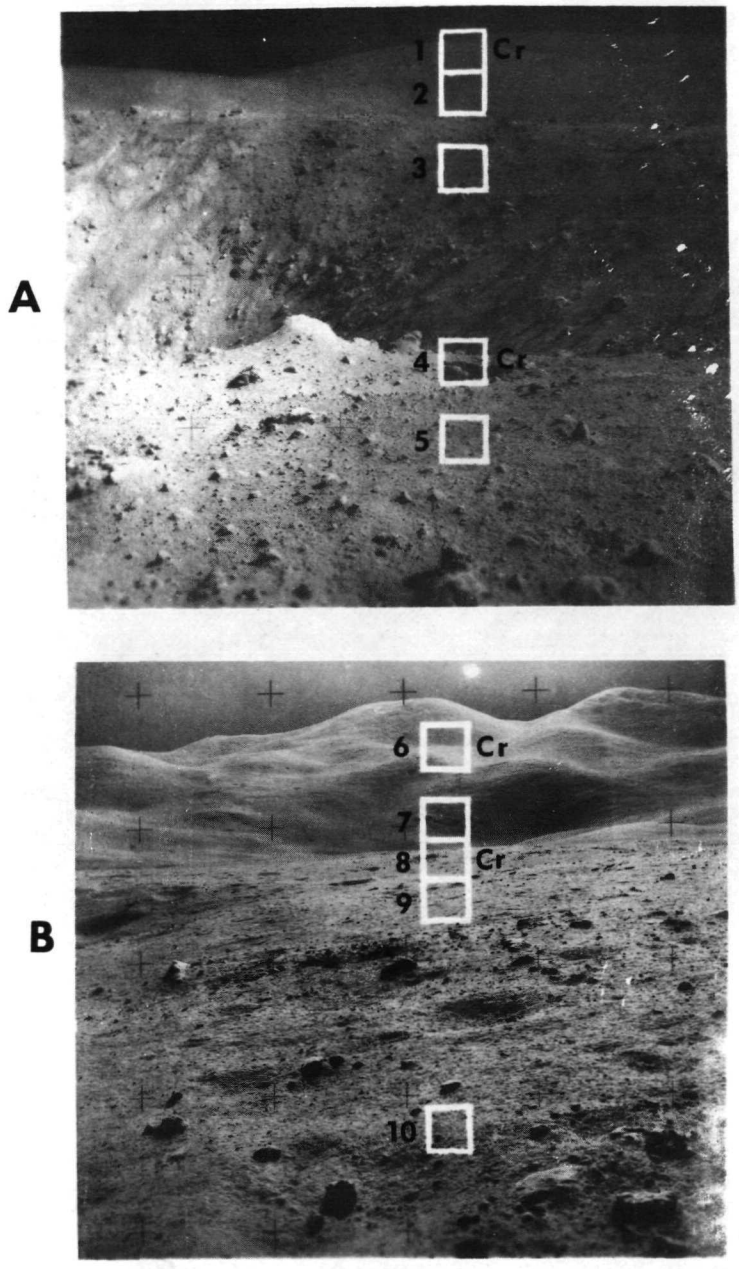


FIGURE 2. PHOTOGRAPHS USED IN THE STUDY, WITH THE ANALYZED PATCHES DESIGNATED. PHOTO A IS NUMBER 106-17239 FROM APOLLO 16. THE LARGE CRATER IS NORTH RAY. PHOTO B IS AS15-90-12247 FROM APOLLO 15 AND SHOWS THE MOUNTAINS TO THE SOUTH OF HADLEY IN THE BACKGROUND. (CONTINUED – NEXT PAGE)

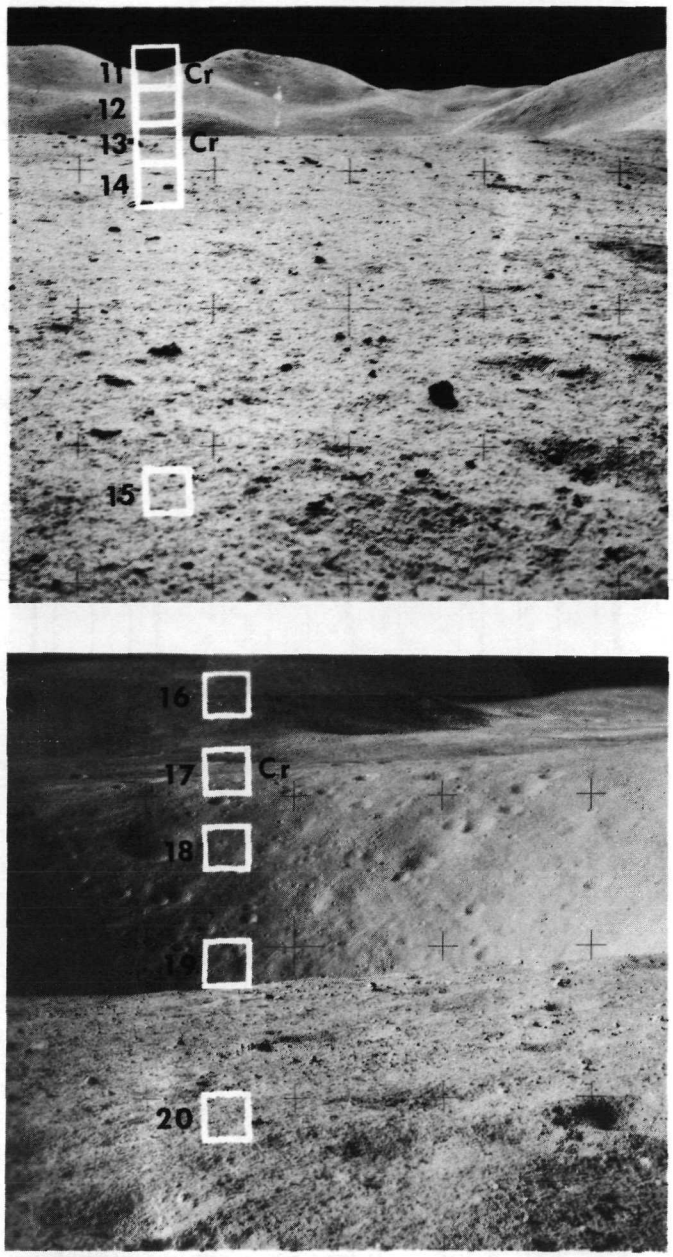


FIGURE 2 (CONTINUED). PHOTO C IS AS15-82-11119 FROM APOLLO 15. THIS PHOTO ALSO LOOKS TOWARD THE SAME AREA BUT FROM A DIFFERENT LOCATION. PHOTO D IS NUMBER 109-17823 FROM APOLLO 16 AND SHOWS SPOOK CRATER WITH STONE MOUNTAIN IN THE BACKGROUND.

PATCH NUMBERS (1-20) ARE INDICATED; Cr DESIGNATES THAT A DISTINCT CRESTLINE PASSES THROUGH.

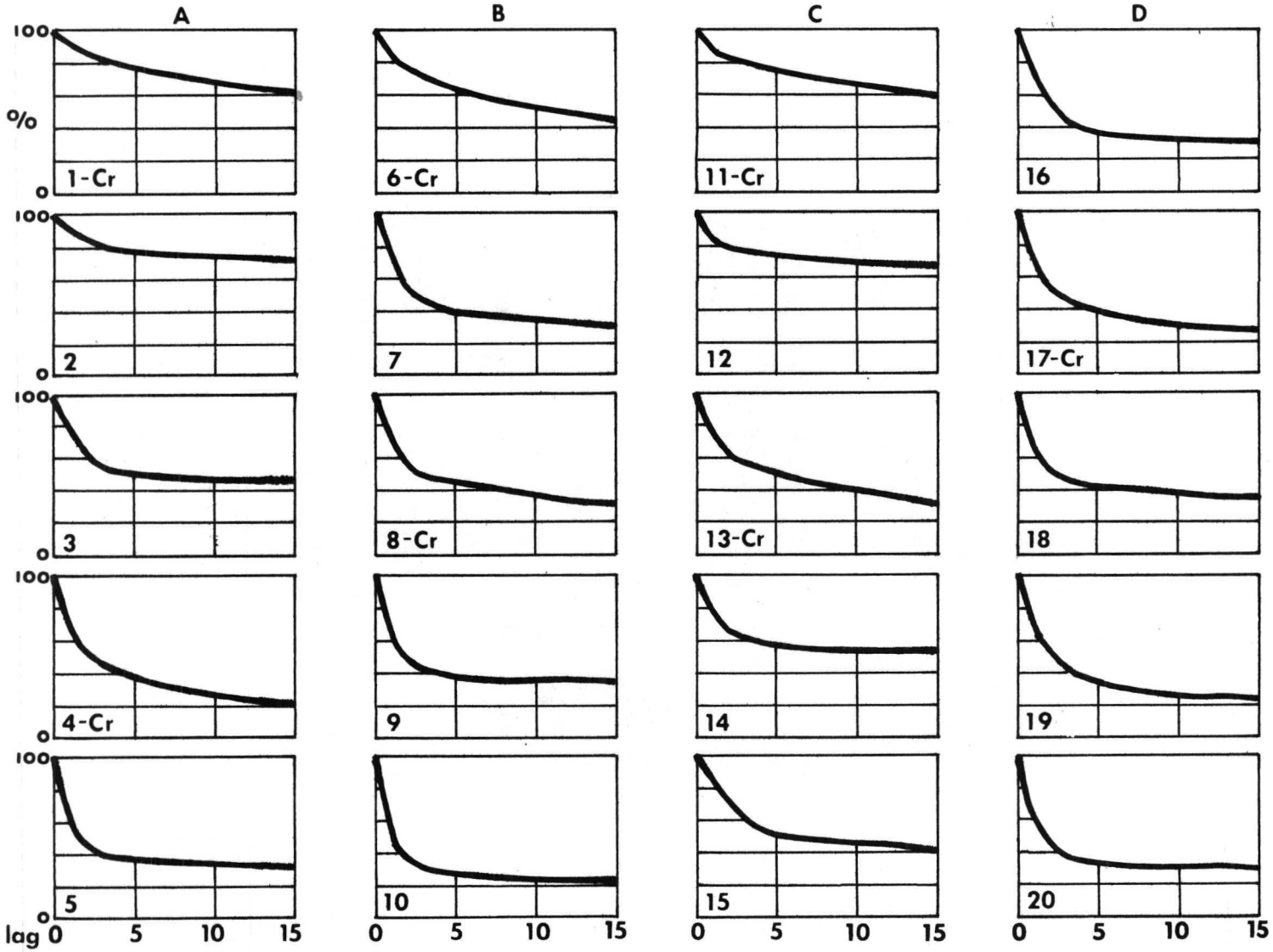


FIGURE 3. GRAPHS OF THE BASIC RESULTS. (SEE TEXT)

TABLE I

BASIC RESULTS. THE PROPORTIONS OF PAIRS WITHIN ONE GRAYLEVEL VALUE FOR 15 LAGS. PATCH NUMBERS CORRESPOND TO THOSE IN FIGURE 2.

Lag \ Patch	1*	2	3	4*	5	6*	7	8*	9	10	11*	12	13*	14	15	16	17*	18	19	20
1	.900	.905	.768	.662	.603	.823	.733	.697	.588	.543	.857	.829	.732	.743	.811	.737	.692	.688	.656	.614
2	.839	.842	.617	.506	.442	.732	.537	.557	.452	.357	.826	.788	.629	.661	.695	.547	.542	.539	.471	.640
3	.816	.821	.550	.445	.394	.690	.464	.503	.416	.298	.805	.771	.574	.619	.625	.452	.467	.476	.402	.371
4	.787	.803	.522	.407	.372	.659	.420	.470	.408	.276	.779	.757	.534	.595	.568	.407	.423	.439	.363	.334
5	.762	.793	.509	.380	.368	.630	.393	.453	.392	.263	.766	.743	.507	.576	.525	.381	.400	.429	.335	.332
6	.740	.785	.497	.350	.361	.612	.390	.437	.381	.262	.751	.732	.472	.569	.498	.364	.375	.411	.315	.322
7	.722	.779	.493	.324	.354	.597	.386	.427	.369	.252	.729	.729	.451	.560	.487	.348	.356	.410	.293	.318
8	.706	.765	.482	.304	.359	.573	.383	.408	.367	.257	.708	.719	.428	.550	.478	.345	.340	.409	.289	.316
9	.686	.766	.473	.287	.357	.554	.372	.386	.368	.246	.693	.718	.413	.547	.468	.344	.322	.397	.280	.305
10	.671	.762	.473	.268	.350	.532	.350	.368	.370	.234	.676	.708	.402	.548	.459	.341	.318	.386	.269	.303
11	.657	.747	.468	.257	.343	.515	.341	.355	.372	.242	.656	.696	.395	.536	.446	.339	.306	.372	.260	.307
12	.637	.742	.464	.244	.334	.499	.330	.332	.370	.246	.639	.673	.375	.533	.430	.329	.296	.359	.257	.314
13	.629	.731	.464	.228	.325	.480	.324	.319	.371	.239	.620	.665	.356	.540	.417	.325	.287	.359	.250	.306
14	.621	.726	.467	.217	.325	.465	.313	.320	.370	.243	.600	.664	.337	.531	.405	.315	.285	.353	.244	.299
15	.612	.711	.466	.206	.327	.452	.305	.314	.364	.235	.582	.650	.322	.531	.397	.312	.276	.354	.238	.299

*Indicates a crestline.

REFERENCES AND FOOTNOTES

1. Olson, Judy, "Visualization of the Lunar Surface," Final Reports, NASA-ASEE Summer Faculty Fellowship Program (University of Alabama – Auburn University), 1971.

2. See, for example: Carlucci, Luigia, "A Formal System for Texture Languages," Pattern Recognition, v. 4 (1972), pp. 53-55.

3. See, for example: McLaughlin, J.A., and J. Ravin, "Nth-Order Autocorrelations in Pattern Recognition," Information and Control, v. 12 (February, 1968), 121-142.

4. Idelsohn, J.M., "A Learning System for Terrain Recognition," Pattern Recognition, v. 2 (1970), 293-301.

5. Zawadzki, I.I., "Autocorrelation of Weather Patterns by an Incoherent Optical Method," Proceedings, American Meteorological Society, Radar Meteorology Conference, 14th, Tucson, Arizona, November, 1970, 421-424.

6. Julesz, Bela, "Visual Pattern Discrimination," IRE (Institute of Radio Engineers) Transactions on Information Theory, February, 1962, 84-92.

7. Kaizer, Hubert, A Quantification of Textures on Aerial Photographs, Technical Note No. 121, Optical Research Lab., Boston University, Mass., June, 1955. (Abstract available only.) It is interesting to note that in this study, the decorrelation distances (lag distance at which the autocorrelation fell to 1/e) were correlated 0.99 with the pooled subject's ranks (according to texture) of a series of targets.

8. A High Speed Rotating Drum Microdensitometer (Optronics International).

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

SIMULATION, MANUAL AND COMPUTERIZED

(SMAC)

Prepared by:	Murl Wayne Parker
Academic Rank:	Associate Professor
University:	Mississippi State University
Laboratory:	Astronautics Laboratory
(Division)	Mechanical and Crew Systems Integration Division
(Branch)	Systems Operations Branch
Research Counterpart:	R. A. Schlagheck
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

LIST OF FIGURES

LIST OF TABLES

SIMULATION, MANUAL AND COMPUTERIZED (SMAC)

By

Murl Wayne Parker

ABSTRACT

Monte Carlo simulation of discrete stochastic real world systems has become a major problem solving technique in recent years with the development of high speed, high capacity computers. Projections indicate that this trend will continue in the foreseeable future.

As in the use of any new tool, acceptance of simulation results is limited by the understanding of the technique. To overcome this limitation, Simulation, Manual and Computerized (SMAC) has been developed. The guiding concept was to develop a simulation modeling method which can be quickly taught and learned, yet can be used to solve significant real world problems.

SMAC consists of two manual simulation methods and one computer program. One of the manual methods is designed to impart an understanding of simulation; the other is designed to be a practical simulation method.

The computer program will simulate queueing situations with up to five parallel servers or up to three series servers. The program is built around GASP IIA routines and retains the 8K capability of GASP.

INTRODUCTION

With the development of high speed, high capacity computers, Monte Carlo simulation has become a major problem solving technique. With the continued progress in computer technology, it seems likely that simulation will become an even more powerful tool.

As in the use of any new tool, acceptance of simulation results is limited by the understanding of the technique. This limitation is especially acute for simulation. This is true for at least two reasons; first, because simulation is such a powerful technique, it produces results which are difficult to believe. Second, there is some conceptual difficulty in understanding how some numbers on a piece of paper, or a set of computer commands, can really represent or model some real world system to the extent of being able to predict the results of system operations.

OBJECTIVES

To overcome this conceptual difficulty, Simulation, Manual and Computerized (SMAC) has been developed. SMAC is a simulation modeling method which consists of two manual procedures and one computer program along with instructions for the use of these.

MANUAL PROCEDURES

The first manual procedure (SMAC I) is intended for use as a teaching technique. It is designed to impart an understanding of the modeling process and is setup so that activities in the real world system are reflected one-to-one by a simulation activity. This correspondence is intended to overcome the conceptual difficulty mentioned above.

SMAC I uses three forms:

a. Entity Activity Forms

b. Queue Sheets

c. Master Timesheets

These three forms are shown in figures 1, 2, and 3, respectively.

ENTITY NAME _____

ACTIVITIES	SYMBOL
1.	
2.	
3.	
4.	
5.	

ACTIVITY	BEGIN	END	TIME	ACTIVITY	BEGIN	END	TIME	

Figure 1: Entity Activity Form

QUEUE NAME _____

ENTER		EXIT		LAPSED	ENTER		EXIT		LAPSED	
NO.	TIME	√	TIME	TIME	NO.	TIME	√	TIME	TIME	
1					26					
2					27					
3					28					
4					29					
5					30					
6					31					
7					32					
8					33					
9					34					
10					35					
11					36					
12					37					
13					38					
14					39					
15					40					
16					41					
17					42					
18					43					
19					44					
20					45					
21					46					
22					47					
23					48					
24					49					
25					50					

Figure 2: Queue Sheet

MASTER TIME SHEET

SHEET _____ of _____

TIME _____ to _____

TIME	EV.	TIME	EV.	TIME	EV.	TIME	EV.	TIME	EV.	TIME	EV.	TIME	EV.	TIME	EV.
1		26		51		76		101		126		151		176	
5		30		55		80		105		130		155		180	
10		35		60		85		110		135		160		185	
15		40		65		90		115		140		165		190	
20		45		70		95		120		145		170		195	
25		50		75		100		125		150		175		200	

Figure 3: Master Timesheet (SMAC I)

The second manual procedures (SMAC II) is intended for use as a practical manual simulation method and uses only one form, a Master Timesheet, shown in figure 4.

Three sets of random number tables are supplied for use with SMAC I and SMAC II. The first set consists of uniform random deviates (00 to 99). The second set of numbers is composed of exponential random deviates with means ranging from 0.25 to 20.0. The third set is composed of Poisson random deviates with means ranging from 0.25 to 20.0. Samples of these sets are shown in tables 1, 2, and 3, respectively.

Uniform random numbers are supplied because random deviates from any distribution can be generated using uniform deviates as a basis. Exponential and Poisson deviates are supplied because these distributions are often needed in modeling queueing situations, a common use of simulation.

SMAC I and II Definitions

The following definitions are needed in order to use SMAC I and II.

a. Permanent system entities -- Persons, machines, or facilities which are integral to system operation as opposed to being a transient element in the system.

b. Entity activities -- Distinguishable categories of entity behavior. Enough categories should be clearly defined so that activity at any time falls into one and only one category.

c. System queues -- Waiting lines. If a queue is defined prior to some sort of service or service facility, all entries to the service are considered to join the queue, even if the service is immediately available and no waiting is required. When no waiting is required, the entry is called a zero entry.

d. Queue discipline -- The rules under which the queue operates. For our purposes, all queues will operate under the following standard discipline unless otherwise defined.

SHEET ____ of ____

MASTER TIME SHEET

TIME ____ to ____

TIME								TIME							
1								51							
2															
3															
4															
5								55							
6															
7															
8															
9															
10								60							
								65							
15															
20								70							
25								75							
30								80							
35								85							
40								90							
45								95							
50								100							

Figure 4: Master Timesheet (SMAC II)

71	52	21	57	92	18	97	28	72	72	41	31
55	78	58	25	15	98	8	64	95	69	70	55
73	48	47	03	82	28	39	86	05	37	33	29
79	49	45	05	98	79	71	27	57	64	49	32
71	55	15	97	20	80	84	14	51	51	27	87
23	32	76	78	74	30	20	13	36	55	05	36

Table 1: Sample of Uniform Random Deviates

17	1	5	8	2	26	2	1	2	13	9	1
7	0	17	0	2	0	13	4	2	0	5	0
4	9	0	3	2	14	1	1	5	0	10	4
1	0	2	3	1	0	10	0	7	3	3	8
12	3	7	5	15	5	1	5	2	5	1	4
4	0	13	2	6	2	1	6	5	0	9	2

(Mean = 5.0)

Table 2: Sample of Exponential Random Deviates

3	6	0	4	6	7	5	6	7	7	4	6
6	3	4	13	6	6	4	5	5	3	4	6
1	7	5	8	4	6	5	5	5	7	2	13
6	3	5	2	8	4	5	4	0	2	5	3
7	6	8	4	7	3	10	8	8	6	4	6
7	3	5	7	1	4	5	0	5	3	5	3

(Mean = 5)

Table 3: Sample of Poisson Random Deviates

(a) First come, first served.

(b) No balking (leaving if the line is too long).

(c) Poisson arrivals (the number of arrivals into the system is a random variable distributed according to the Poisson distribution).

(d) Exponential service times (the service times are random and are distributed according to the exponential distribution).

(e) Single or parallel service facilities are served by a single queue, with no facility preference.

e. System events -- Arrival to a queue, departure from a queue, or an entity changing from one activity to another.

f. RN -- Random number.

SMAC I Simulation

Preparation for simulation.

a. For each permanent system entity, fill in the heading of an entity sheet. List the activities for each sheet.

b. For each queue in the system, fill in the heading of a queue sheet. Define the queue discipline if different from the standard discipline.

c. List the system events and assign a mnemonic letter to each.

d. Select the proper RN table or generate a set of appropriate RN's.

e. Enter system independent events onto the master timesheets. Use sequential numbers for queue entries and exits.

Simulation

a. Use the master timesheets as a guide and move the system through time. As the system is updated on the master timesheet, make sure that each queue and entity sheet is also updated. Follow the specific procedure given below.

(1) Find the earliest completion time for the last entry on any entity sheet.

(2) On the queue sheets, enter a ✔ for the arrivals up to the time found in (1). Enter these arrivals (as numbers) on the timesheets if they are not already entered.

(3) Determine the next activity to occur after the time found in (1). Enter the mnemonic(s) for the starting event(s) on the master timesheet.

(4) Determine the length of time of the activity and enter the mnemonic for the ending event on the master timesheet. Enter the starting time, ending time, and elapsed time on the entity sheet.

(5) Enter the service beginning time and elapsed time for any customer leaving the queue due to the events entered on the master timesheet.

(6) Repeat steps (1) to (5) until the simulation is complete.

b. No procedure for analyzing the results of a SMAC I simulation is given since SMAC I is intended as a training procedure only. Once the process of simulation and modeling is understood, SMAC II should be used as a more efficient procedure.

SMAC II Simulation

Preparation for simulation.

a. List the permanent system entities. For each, list the entity activities.

b. List the system queues. For each, define the queue discipline.

c. List the system events, and assign a mnemonic letter to each.

d. Enter appropriate headings on the columns in each timesheet, one column for each queue and one for each entity.

e. Select the proper RN table or generate a set of approximate RN's.

f. Enter system independent events on to the master timesheets (use sequential numbers for queue entries and exits).

Simulation

Beginning with time one, enter the mnemonic for each event on to the master timesheets (use sequential numbers for queue entities and exits).

Analysis

a. For each entity sum the times spent in each activity. Calculate the percentage for each activity.

b. For each queue,

(1) Tally the zero entries.

(2) Note the time spent in the queue by each non-zero entry.

(3) Note the total duration of each queue length experienced.

c. From this information, the following queue statistics may be calculated:

(1) Percent zero entries.

(2) Percent non-zero entries.

(3) Average queue length.

(4) Maximum queue length.

(5) Average waiting time (all entries).

(6) Average waiting time (non-zero entries).

COMPUTERIZED PROCEDURES

In order to allow simulations of more complicated situations over longer periods of simulated time, a computer program has been developed (SMAC III). The program has the capability of simulating up to three series queues or up to five parallel service stations for a single queue.

SMAC III is based on the GASP IIA language developed by Pritsker and Kiviat (1). GASP IIA is a set of Fortran subroutines developed to facilitate simulation of complicated systems. Essentially, the routines provide convenient recordkeeping, sequencing, and timekeeping capabilities.

The use of GASP as a base for SMAC III served several useful purposes. First, because of the convenience mentioned above, the routine was easier to write. Second, GASP is written to stay within an 8K core limitation so that it can be used on small machines such as the IBM 1130. This 8K capability has been retained in SMAC III. Third, users of SMAC will automatically have access to GASP IIA routines. This allows the use of these routines to perform more complex simulations as the capability is developed.

Use of GASP requires that the user write a main program, an EVNTS subroutine, an OTPUT subroutine, and other subroutines structured to update the GASP files as the simulation occurs.

These other subroutines are the main part of any GASP simulation and determine the effectiveness and scope of the simulation. SMAC uses three of these subroutines; ARV, START, and FIN. ARV deals with arrivals to the system, START deals with beginning service for a customer, and FIN deals with completion of service for a customer. Flow charts for these three subroutines are shown in figures 5, 6, and 7, respectively. From the figures it is evident that the logic is only slightly different for series and parallel queues.

Built into SMAC III are the assumptions of first come - first served, Poisson arrivals, and exponential service times. Balking is allowed, and a minimum service time can be specified.

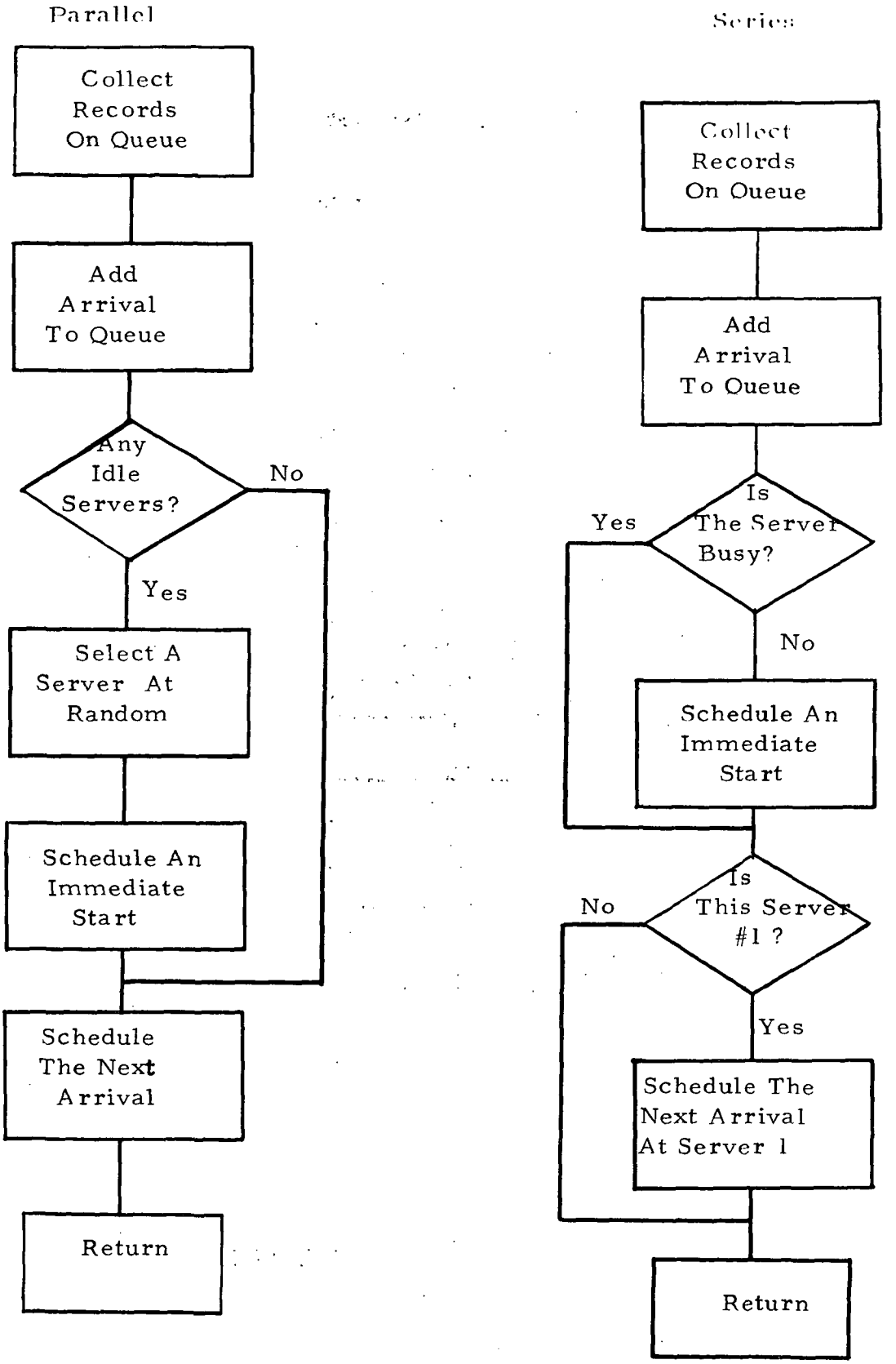


Figure 5: Subroutine ARV

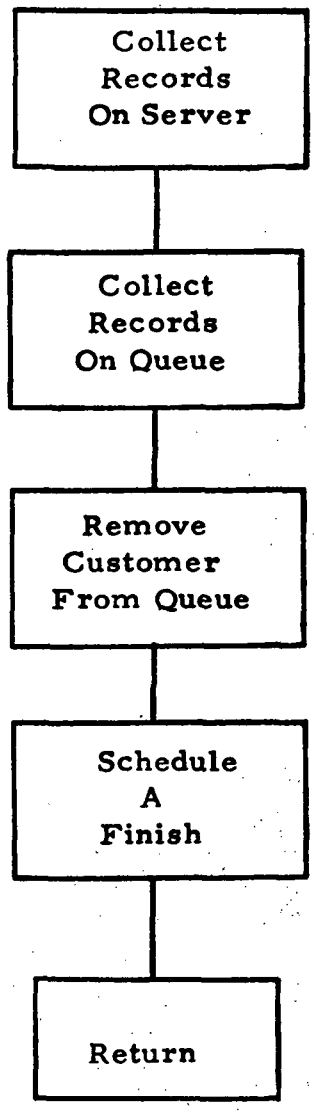


Figure 6: Subroutine START

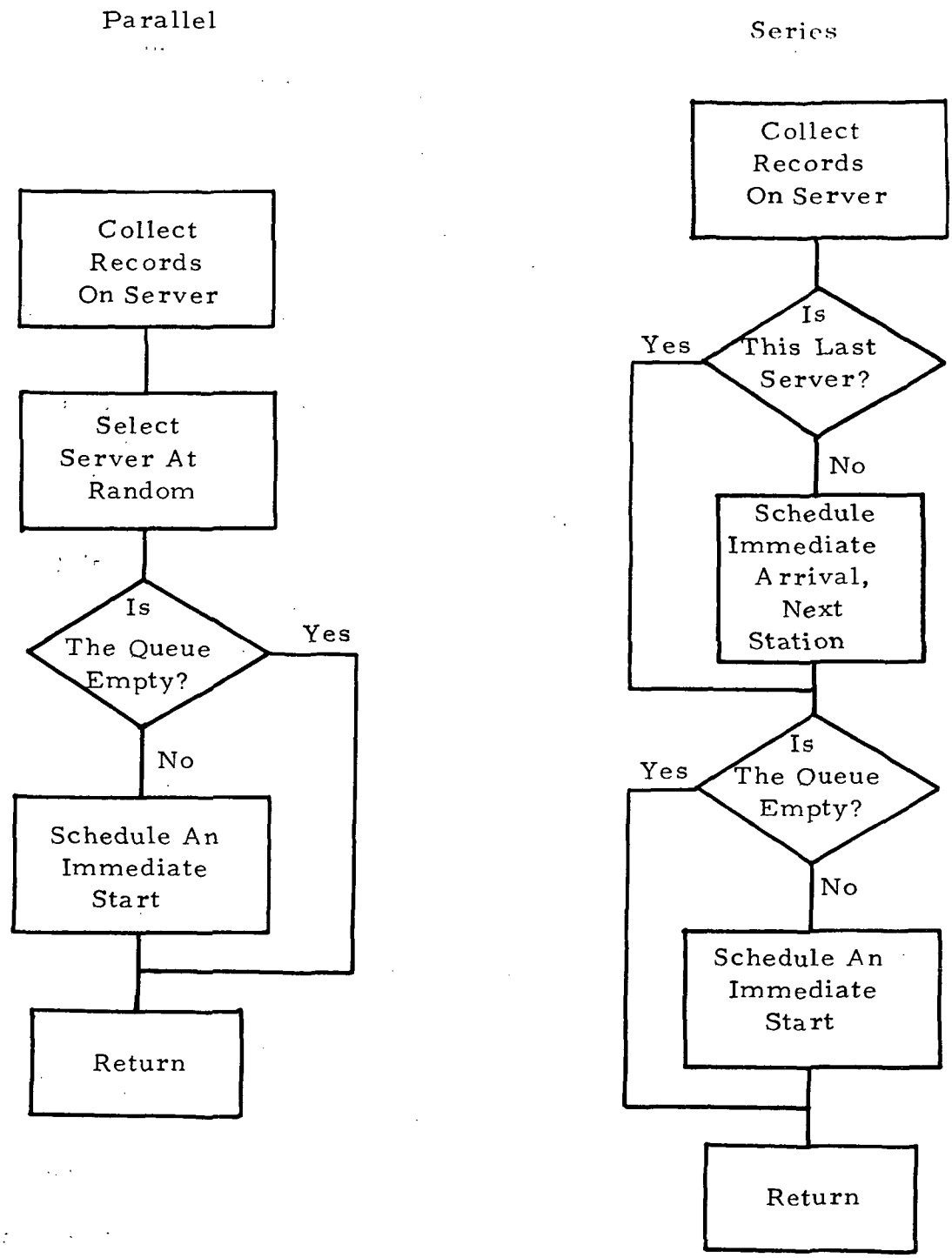


Figure 7: Subroutine FIN

To use SMAC III, it is only necessary to use two data cards. The format for these two cards is shown in table 4.

Restrictions

There are several constraints on the operation of SMAC III, as listed below.

a. All parallel servers are considered to have the same average service time.

b. The total number in all queues at any one time is restricted to approximately 10. If too many customers are in the system at any one time, GASP error 87 occurs (too many entries in a file).

c. Error messages are GASP standard errors which may not be too meaningful to a SMAC user.

d. All arrivals and services are considered to be Poisson and exponential, respectively.

In addition to these operating constraints, it has been necessary to modify some of the GASP routines. These include DATAN, SUMRY, ERROR, and PRINTQ. These modifications took two primary forms. The first was removal of standard GASP output, as not meaningful to SMAC users; and second, reduction of program size to permit operation within the 8K limitation.

CONCLUSIONS AND RECOMMENDATIONS

SMAC offers a vehicle whereby Monte Carlo simulation can be easily understood. In addition it offers the capability of simulating systems which do not call for the application of more powerful languages. This simulation may be manual or computerized. SMAC III users also have access to GASP IIA, a powerful Fortran based language which can be used to simulate complicated systems.

In order to make SMAC III more useful, the ERROR routine should be modified to print error messages meaningful to SMAC users. As a software addition, a set of notes for the user (learner) must be written.

Card 1

Column	Data
1-12	Name of user
13-16	Project Number *
17-18	Month (Number) *
19-20	Day (Number) *
21-24	Year *
25-28	Run Number *
29-30	Simulation Type * (0 for parallel, 1 for series)
31-40	Mean time between arrivals
41-50	Number of Servers
51-60	Number of simulated Time units the simulation is to run

Card 2

Column	Data
1-10	Queue length at which an arrival will balk (if zero, no balking will occur)
11-20	Minimum service time
21-30	Average time to service (all parallel servers or the first series server)
31-40	Average time to service (second series server)
41-50	Average time to service (third series server)

* Data must be right justified.

Table 4: SMAC Data Cards

REFERENCES

Pritsker, A. A. B., and Kiviat, P. J., Simulation with GASP II: A FORTRAN Based Simulation Language, Englewood Cliffs, NJ, Prentice-Hall, 1969.

1972

NASA - ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

COHERENT OPTICAL PROCESSING OF PHASED ARRAY RADAR DATA

PREPARED BY:	JERRY W. ROGERS
ACADEMIC RANK:	ASSOCIATE PROFESSOR
UNIVERSITY:	MISSISSIPPI STATE UNIVERSITY
LABORATORY:	ASTRIONICS INSTRUMENTATION & COMMUNICATION DIV. ADVANCED STUDIES & ANALYSIS OFFICE
RESEARCH COUNTERPART:	J. H. KERR
DATE:	AUGUST 11, 1972
CONTRACT NUMBER:	NGT-01-003-045

ACKNOWLEDGMENTS

The author wishes to thank the ASEE/NASA program directors for a very interesting and stimulating summer. Thanks is also due NASA counterpart, Mr. J. H. Kerr, for excellent support.

COHERENT OPTICAL PROCESSING OF PHASED ARRAY RADAR DATA

By Jerry W. Rogers

ABSTRACT

This study deals with interfacing a coherent optical processor, which utilizes an electron-beam addressed KD*P crystal modulator, with a linear phased array. For completeness,an abbreviated development of typical radar signals from a linear array is included.

A plan for formating the spatial modulator with linear array signals is presented. The theoretical expectations which include target angle and doppler are derived.

A simulated set of M signals which are typical of a linear array of M elements was devised. This set of signals was used to modulate the wave front of collimated laser light via the KD*P crystal according to the format presented. The results are compared with those which theory predicts.

INTRODUCTION

Phased arrays have the capability of producing narrow radiation/receiving patterns and as a consequence high spatial resolution. With the addition of controllable individual-element phase shifters, wide angle beam control can be achieved by simply controlling the phase shifts. These two properties have served as impetus to widespread interest in phased arrays. The price one pays, however, is that the processing electronics is substantial. In order to alleviate the electronic processing problem, coherent optical processors are being considered. One such optical processor is the subject of this paper.

RESEARCH OBJECTIVE

The research objective is as follows:

1. Identify the problem of interfacing the "on board" coherent optical processor with a phased array.

2. Propose a plan for analyzing radar returns in order to extract radar information.

3. Conduct feasibility experiments on simulated radar returns.

OPTICAL PROCESSING SYSTEM

The heart of an optical processor for radar returns is the device which converts the returns into an optical analog. A device capable of performing this conversion was constructed for Marshall Space Flight Center under contract, patterned after the ones previously reported by other researchers [1, 2]. This device, bearing the name of Image Forming Light Modulator (IFLM) followed the general cascade arrangement of polarizer, electro-optic cell, and analyzer. A Potassium Dideuterium Phosphate (KD*P) crystal arranged as a longitudinal-field modulator served as the electro-optical cell. The modulator was constructed in such a manner that it was addressable via an electron beam, and had the capability of operating at standard television rates. The crystal had 3.15 x 3.15 cm^2 in active face area with a resolution (uncooled) of approximately 200 lines.

The block diagram of the proposed system is shown in figure 1. In order to observe the nature of signals which the optical processor will be required to accommodate, a brief development of the pertinent characteristic equations will follow.

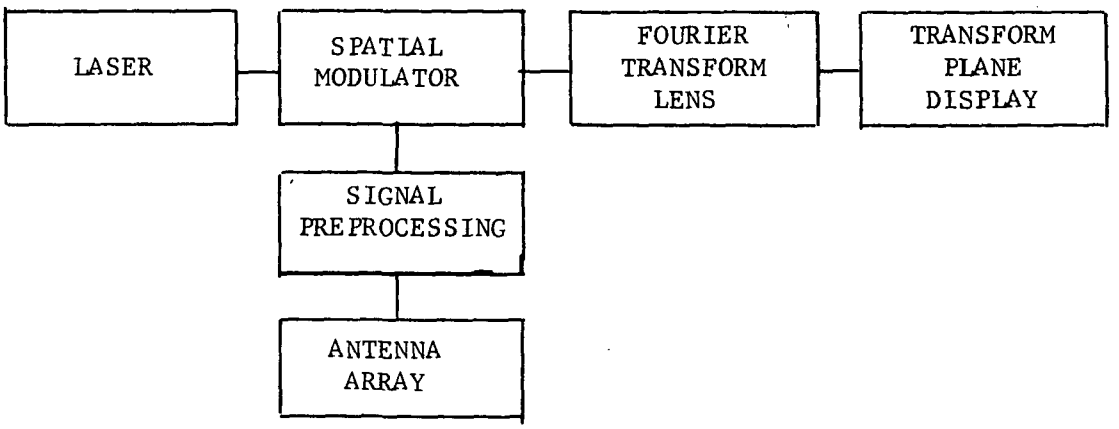


Fig. 1. Coherent optical processor for phased array radar data.

Because of the 200 line resolution limit on the available IFLM, only a linear array capability exists; thus, the development is for a linear array. Higher resolution can be achieved on the IFLM if the crystal is operated at Curie point. With greater resolution, a planar array would become practical, and the formating extention would be straight forward.

Consider the linear array of M identical radiators uniformly spaced d_m apart as shown in figure 2.

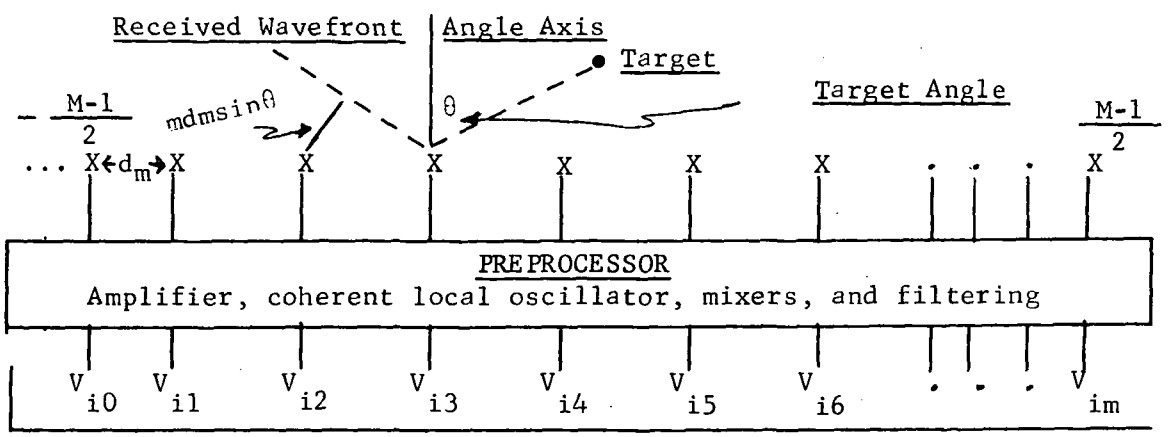


Fig. 2 Linear array antenna and preprocessor

For a signal received from target as shown in figure 2, the output from the mth channel of the preprocessor is of the form [3].

$$V_{im} = P_{T_P}(t - m\tau) \cos 2\pi [(fi + fd)t - m\tau fr] \quad (1)$$

where

fi = intermediate frequency
fd = doppler frequency
fc = carrier frequency
fr = fc + fd
t = t' $-\tau R$
$\tau R = 2R/C$
R = range
c = velocity of light
$\tau = \frac{dm}{c} \sin\theta$

If the conditions are met that $|m\tau|_{max}$ is small compared to pulse length, T_P; i.e., $|m\tau|_{max} << T_P$ and fc $>>$fd, then equation (1) can be written in its approximate form as:

$$V_{im} = P_{T_P}(t) \cos 2\pi [(fi + fd)t - m\alpha] \quad (2)$$

where

$$\alpha = \frac{dm}{\lambda c} \sin\theta \quad (3)$$

Note that θ is the target location angle and λc is the carrier wavelength. The problem now becomes that of processing M voltages whose forms are given by equation (2). Conceptually, one may either electronically switch from one signal to another and display or create a progressive delay, T_D, in the intermediate frequency signals and sum, thereby producing a single signal (time multiplex).

If this delay technique is employed, T_D must satisfy the condition

$$T_D = T_P + |\tau|_{max} \quad (4)$$

in order that the M signals after summation do not overlap. The length of the summed signal would then be MT_K where $T_K = T_D + \tau$.

IMAGE FORMING LIGHT MODULATOR SIGNAL FORMAT

The formating of the rf radar returns on the spatial modulator will be identical to that suggested by Casasent [5], and parallel that reported by Lambert et al [3] even though the latter was related to a Debye-Sears effect modulator. In the linear array format, there will be one line displayed on the IFLM for each antenna; i.e., there will be M lines. The mth line will be the spatial analog of the heterodyned signal from the mth array element. Each line must be full and must display completely overlapping signals in a specified range span. These two conditions can be best described as

$$T_P \geq Ts \tag{5}$$

and

$$\tau_R - \left(\frac{T_P - Ts}{2}\right) \leq T_R \leq \tau_R + \left(\frac{T_P - Ts}{2}\right) \tag{6}$$

where

τ_R = range delay of interest
T_P = pulse length
T_s = write time for one line
T_R = 2R/C = target range span from which all returns will be completely processed
R = target range

Thus, it can be observed that the range capabilities relate both to pulse duration and the write time for one line.

Implementation of the desired charge distribution over the face of the KDP crystal consists of a linear sweep of the electron beam across the crystal face with simultaneous sine wave intensity modulation of the electron source.

In general, the space variable, x, relates to the time variable, t, by the expression

$$x = At + B \tag{7}$$

Subject to the end condition that x = 0 @ t = 0 and x = D @ t = Ts, then equation (7) becomes

$$x = \frac{D}{Ts}\, t \tag{8}$$

where Ts is the write time for one line and D is the aperture dimension in the x direction. If the transmission function, $T(\nu)$, of the light

modulator is assumed to be $T(\mathcal{V}) = 1 + \mathcal{V}$, then the spatial variations, T(x,y), at the output plane of the modulator is

$$T(x,y) = \sum_{m=-\frac{M-1}{2}}^{m=\frac{M-1}{2}} \mathcal{S}(y-m\mathcal{l}) \left\{1+\mu\cos 2\pi\left[(fi+fd)\frac{Ts}{D}x - m\alpha_m\right]\right\} \text{rect}\left(\frac{x}{D}\right) \quad (9)$$

for M lines of sine wave amplitude modulation. In equation (9), $\mathcal{l}$ is the line separation, μ is the modulation index ($0 \leq \mu \leq 1$) and $\mathcal{S}(y)$ is the approximate transmission function of the line cross section. Note that equation (9) is the space analog of equation (2).

If the crystal is illuminated with a well-collimated coherent light source, configured such that one thin spherical lens is placed on focal length, f, behind the crystal, the output plane which appears two focal lengths, 2f, behind the crystal contains the Fourier transform of T(x,y) [4]. Subject to the condition that $Ts(fi+fd) \gg 1$, and omitting the mathematical details of transformation, the intensity distribution in the output plane is

$$I(u,v) = K_o^2 \frac{\sin^2 M\pi(\mathcal{l}v)}{\sin^2\pi(\mathcal{l}v)} \bullet \text{sinc}^2 Du + K_1^2 \frac{\sin^2 M\pi(\alpha_m+\mathcal{l}v)}{\sin\pi(\alpha_m+\mathcal{l}v)} \bullet \text{sinc}^2 D\left[(fi+fd)\frac{Ts}{D} - u\right] \quad (10)$$

where

$$[u,v] = \left[\frac{x_o}{\lambda f}, \frac{y_o}{\lambda f}\right]$$

x_o, y_o = output plane coordinate
λ = wavelength of illumination
f = focal length of spherical lens

In the vicinity of the first order fringe, I(u,v) has a maximum which occurs at

$$v = -\frac{\alpha_m}{\mathcal{l}}, \qquad u = (fi + fd)\frac{Ts}{D} \quad (11)$$

From measured coordinate, u and v, both the unknowns, fd and α_m, can be calculated since Ts, D, fi, and $\mathcal{l}$ are design constants.

EXPERIMENTAL VERIFICATION

The IFLM has peripheral electronics which make it compatible with standard television display; thus, one imposed condition was that the simulated rf must have synchronization pulses properly spaced and other standard video characteristics. Furthermore, because the resolution of the KD*P crystal is 200 lines or less, it would be necessary to process only every 4th or 5th line. Unfortunately, not all the necessary electronic equipment for synthesizing a composite video was available so an alternate was devised. The alternate plan consisted of sine-wave intensity modulating an oscilloscope with a staircase voltage wave applied to the vertical amplifier. Figure 3 is a block diagram of the experimental arrangement.

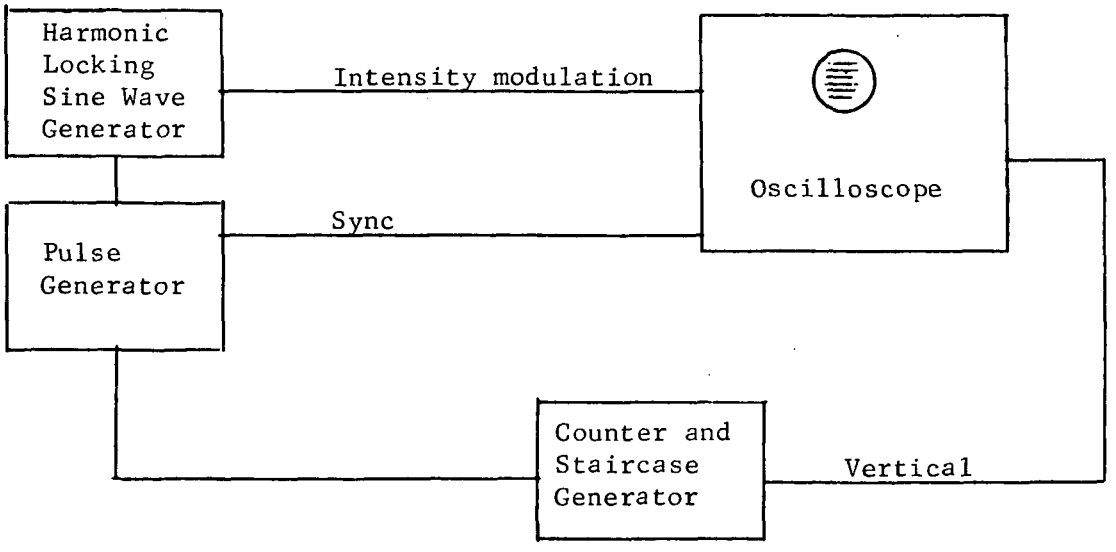


Fig. 3 Intermediate frequency simulation for the IFLM.

With synchronization pulses furnished to the horizontal sweep, a pattern of evenly spaced lines was the result. Each line was sine-wave intensity modulated along the x axis. Progressive phase shifts were produced by varying the phase of the harmonic locking frequency generator very slightly.

The first experiment was that of making a transparency of the oscilloscope pattern, and processing the transparency in order to verify the signal format. Several transparencies of oscilloscope patterns were made. The transparencies were made of patterns which had 18 lines (18 antennas) and include progressive angles of 180°,

120°, 90°, 45°, and 0°. In addition to the progressive phase angles, several different spatial frequencies in the object plane were selected for experiments. All transparencies were processed in an optical arrangement as shown in figure 4.

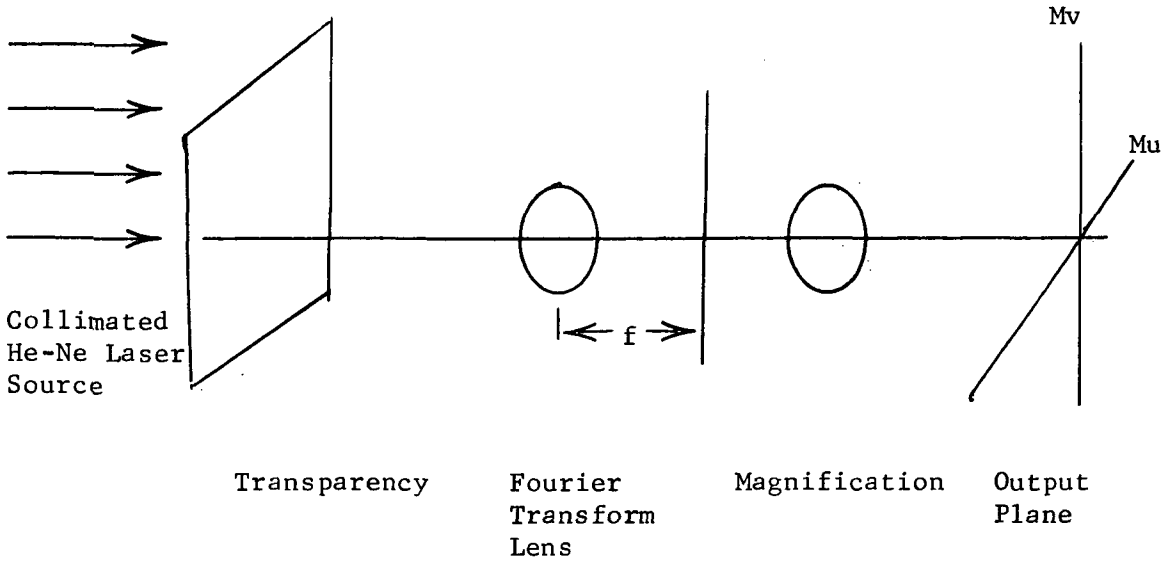


Fig 4 Fourier transform and magnification arrangement for processing simulated array transparency.

The transparencies were made with Polaroid projection film, and no great care was exercised in establishing that the lines were precisely sine-wave modulated. Each transparency was viewed under the microscope and measurements were made of line separation, spatial frequency, and progressive phase displacements. From these data and knowledge of the system parameters shown in figure 4, the expected distribution at the output plane was calculated.

The comparison of the calculated results and the experiments was very good; thus, establishing that the formating arrangement is sound.

The next experiment which involves the IFLM will be simply to view the oscilloscope pattern with a television camera. The composite video from the camera will be used via the peripheral electronics to place the image on the IFLM. The processing of the IFLM image will be identical that used on the transparencies. Unfortunately, the IFLM is not presently operational and the experiment was not done, but it is expected to be repaired very soon at which time the work will be continued.

CONCLUSIONS AND RECOMMENDATIONS

There are at least four properties of a system which utilizes a coherent optical processor that can be listed. These four properties are as follows:

1. Target angle can be obtained unambigously for each target in range span.

2. Target angle resolution is identical to resolution inherent in the array.

3. All antenna beam pointings appear simultaneously.

4. Angle and doppler frequency appear at the output completely associated in orthogonal coordinates.

Not any of the above properties are proven as such within the test of this report; however, all can be shown with very little additional effort.

Based on experimental work, the prospects of the development of a viable radar system which incorporates the coherent optic analyzer is uncertain. Even though the initial work is very encouraging, some of the limiting parameters have not been defined or fully explored. The effect of noise in the electronics and the coupling of the noise into the optical processor are examples of problems which are unresolved. In addition, the comparison of amplitude and phase modulation of the light source is of interest. Furthermore, a study of the electronics for implementation for cost, complexity, and reliability is in order, and the results of this study should be compared with existing conventional systems. Since radar systems are often objective oriented, it could be that the optical processing techniques will compare favorably for some types of applications and not for others. In either event, a recommendation that effort continue in this area is offered in order to help resolve some of the unknowns.

REFERENCES

1. M. vonArdenne, U. S. Patent 2,276,359 (1939).

2. D. Casasent, "An On-Line Electro-Optical Video Processing System," Report Number 331, Department of Computer Sciences, University of Illinois, May 1969.

3. L. B. Lambert, M. Arm, and A. Aimette, Optical and Electro-Optical Information Processing, Tippet et al., Eds, MIT Press, Cambridge, Mass., 1965, pp 715-747.

4. M. Born and E. Wolf, Principles of Optics, Pergamon, NY, 1959.

5. D. Casasent and Jack Stephenson, "Electro-Optical Processing of Phased-Array Antenna Data," Applied Optics, Vol. II, No. 5, May 1972.

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY _ UNIVERSITY OF ALABAMA)

AVOIDANCE OF TRAILING VORTEX HAZARD BY AIRPORT WARNING SYSTEM

Prepared by:	Robert S. Rudland
Academic Rank:	Instructor
University:	California State Polytechnic College
Laboratory: (Division) (Branch)	Aero-Astrodynamics Mission Planning and Analysis Division Orbital Mechanics Branch
Research Counterpart:	Harold B. Jeffreys
Date:	August 25, 1972
Contract No.:	NGT-01-003-045

LIST OF ILLUSTRATIONS

AVOIDANCE OF TRAILING VORTEX HAZARD BY AIRPORT WARNING SYSTEM

By

Robert S. Rudland

ABSTRACT

Over the years, increased air traffic has stimulated a number of new innovations, such as radar control for the airport operators to handle the increased traffic, and the new 'Jumbo Jet' to carry more passengers and cargo. This combination of large aircraft and heavy traffic has exposed a unique problem in air traffic operations, that of "Trailing Vortex Turbulence". The air traffic operator avoids positioning aircraft too close to these 'Jumbo Jets' during landing and takeoff. This tends to slow down airport traffic.

In an effort to improve airport traffic operation near 'Jumbo Jets' an Airport Trailing Vortex Warning System (ATVWS) has been developed at NASA to predict "Trailing Vortex Turbulence". The ATVWS consists of a computer simulator which is compatible with a vortex sensor such as a Laser Doppler Velocimeter (LDV) or an Acoustic Radar (AR).

A study of the traffic in one of the ten busiest airports, Atlanta Municipal Airport, provided information on traffic patterns of a statistical nature. With this information, airport traffic patterns can be studied and hazard statistics developed. Based on a careful study of this simulation model, requirements for the developement and testing of an ATVWS can be accurately specified. With the introduction of the ATVWS will come improved traffic as well as increased passenger safety.

INTRODUCTION

In order to understand "Trailing Vortex Turbulence" it is necessary to become familiar with aircraft "Wakes". Every aircraft generates a wake in flight although the cause for this wake was thought to be a result of "Prop Wash" or "Jet Blast". Although propellors and jet engines do create sizeable air disturbances close to the aircraft, a pair of counterrotating vortices can be found far behind most aircraft. These vortex streams have been named "Wake Turbulence" by Federal Aviation Administration (FAA)[1*].

The vortex streams or wake turbulence is created by air trying to roll around and over the wing tips in an effort to equalize the pressure on the top and bottom of the wing. This results in air leaving the wing tips with a strong swirling motion or vortex. On modern swept wing aircraft this vortex starts even before the air leaves the wing tips. The intensity of the swirling motion (vortex strength) is directly related to the aircraft weight and the wing loading across the span. The Boeing 747 Jumbo Jet has been observed to have its greatest vortex strength when flying slowly with its landing flaps retracted (clean).

The reason the FAA has taken time to study wake turbulence is because of the danger to aircraft that might encounter this wake before it has had time to break up into harmless eddies. The greatest hazard occurs in the airport where aircraft have only limited capability for recovery. Since wake turbulence is not visible to the pilot, he can not see that a hazard might exist and thereby avoid it.

A recent accident at Fort Worth[3] involving a Douglas DC-9 and a DC-10 Jumbo Jet has been attributed to wake turbulence. It was felt that a trailing vortex remained swirling above the runway due to the DC-10 and lasted several minutes. The DC-9 encountered this vortex left by the DC-10 and was unable to prevent its wingtip from hitting the ground. To prevent further mishaps of this type the FAA set up guidelines for air traffic controllers until more extensive studies can ensure against wake turbulence accidents.[2] These guidelines have increased the seperation distance behind Jumbo Jets from three to five miles.

* Superscripts refer to references found in the bibliography.

An examination of aircraft accidents over a ten year period from 1960 through 1969 [5] showed that of over 100 million air operations less than one in a thousand were involved in accidents of any type. Of all accidents, only a few percent could be attributed to wake turbulence. Of all wake turbulence accidents, only one air carrier accident was reported in ten years. Another way of putting this would be; 10 light aircraft out of 10 million operations are involved in wake turbulence accidents each year. Statistically this may be expressed as a one in a million chance of being in a wake turbulence accident if you are in a private aircraft.

Since 1969 the Jumbo Jet has placed not only private aircraft but also air carriers in a position of needing to avoid wake turbulence. Even large passenger aircraft are no longer immune from wake turbulence as was once the case. The Fort Worth accident mentioned earlier makes one aware that the lives of over 100 passengers could be endangered if one aircraft encountered wake turbulence of the nature of that which the DC-9 encountered. The need for a wake turbulence avoidance has now come of age much like the radar systems were developed to avoid mid air collisions.

Wake Turbulence Avoidance Systems

The FAA recently announced [3] a new program which will warn aircraft of hazardous wake turbulence conditions. One of the systems under study utilizes acoustic radar to locate the wake turbulence. This system sends out "beeps" from a loud speaker and receives a reflected signal from the wake turbulence. Although this system is limited to short range operations it can track the wake behind an aircraft from the point in time 30 seconds behind the aircraft to 90 second. [6] A Laser Doppler Velocimeter (LDV) using a design free from eye hazard has been studied by NASA Marshall Space Flight Center as an alternate system for detecting the wake turbulence. [9,10] This system is operational and is not limited by the range and sensitivity constraints of acoustic radar.

Other parts of this program involve experimental studies of aircraft encounters with actual aircraft at the National Aviation Facilities Experimental Center (NAFEC) [7] and vortex break up studies by modification of wing design is being performed at NASA. [8]

The Department of Transportation (DOT) has assigned the responsibility for system integration to the Transportation Systems Center (TSC) under the direction of the FAA. The use of this new facility, TSC, to perform system integration is unique in the FAA. Moreover, the FAA must shift to systems studies and simulation models more and more as time goes on due to the extreme complexity of the air traffic operation systems. It is no longer practical to simply build a new instrument and try it out in an airport to see if it works. The air traffic operations centers are so filled with instruments that a new instrument must be designed to fit appropriately with many other instruments in such a way that it can be effectively utilized. New computer simulators and modeling techniques can prove to be extremely useful in determining design requirements of the system.

System Simulation

A system simulation model has been designed and operated by Harold Jeffreys at NASA[22] Marshall Space Flight Center. This model is capable of determining the velocity field in the wake of an aircraft, calculating the roll rate of an encountering aircraft, evaluating aircraft seperation distances, and evaluating an LDV sensor operating in an airport environment. By carefully developing this model in modules, it is possible to interchange or modify these modules to meet a variety of ATVWS configurations. A flowchart is used to describe a possible simulator configuration (See Figure 1). The LDV detector used in this simulation is based upon an operational system used by R. M. Huffaker at NASA[9,10].

An examination of this flowchart identifies the areas where data is required as input for the aircraft and the airport. From this information a velocity field is calculated. Airport scan planes are used to evaluate the influence that this flow field will have upon encountering aircraft. These same scan planes are used to evaluate detector sensitivity, location, system compatibility, and comparitive advantages with other systems.

At the present time the modules are designed as follows (subject to day to day modifications): The vortex module uses Newman's vortex model[11] with subroutines to account for ground effects as

well as vortex break up[12]. The encounter module uses methods similar those used by Condit[7] to calculate roll rates. Subroutines include considerations for stall, maximum roll rate, and ground clearance. The detector module uses tested laser designs to predict resolution, range, sensitivity, signal to noise, and power. The probability module is still being developed and may use Monte Carlo methods or probability tables to evaluate hazard conditions.

The selection of scan planes as a means for modeling the flow field is an excellent technique which lends itself to an evaluation of scan plane locations for a detector system. Once the scan plane locations have been evaluated a detector system can be studied as a function of scan plane location. The simulator can perform many of the system studies in evaluating the detector, and scan planes. The simulator provides quick and easy system integration studies and allows each module to be optimized with respect to the system requirements. The simulator allows the systems analyst an opportunity to specify precise requirements for detector design as well as providing insight into the requirements for the predictor design.

After the simulator has practiced flying aircraft in and out of airports with a variety of seperation distances and wind conditions, unusual aircraft take-off and landing patterns may be studied. Since the wind plays such an important roll in altering the vortex streams, regions of wind rose data will prove more hazardous than others. Contours can be drawn on the wind rose data to show various degrees of hazard. Critical aircraft operations, such as the rotation of the rear engined, swept wing, jet aircraft, demand closer attention and must be identified during simulation studies.

The simulator has the capacity for predicting the degree of hazard for all possible flight conditions. It is now possible to determine the likelihood that this hazard might occur. If an airport were studied carefully, the probability that a given aircraft would takeoff at a given point on the runway, rotate, and fly out at a given climb rate could be statistically determined as a function of weather and aircraft loading. Were this information available, it would be possible to calculate the probability of hazard to be used to warn aircraft of a hazard.

Case Study of Atlanta Airport

The prediction of the probability of hazard relies upon the detailed study of each aircraft in the airport. In an effort to help identify the important variables that are necessary to define the takeoff distance, climb rate, and side slip of an aircraft during both take-off and landing, a case study was attempted at the Atlanta Municipal Airport (one of the ten busiest air carrier centers in the United States). Arrangements were made with both NASA and the FAA, Flight Standards Office, to spend a few days at the airport trying to measure the aircraft operations. Mr. Estep of the FAA and Mr. Burch of Tower Operations spent several hours in discussing ways to measure the aircraft in a meaningful way. Many variables were felt to contribute to the operation of an aircraft. These are:

a. Takeoff weight (Function of destination, cargo, CG, engines, . . .)
b. Location of center of gravity (CG)
c. Flap setting (Function of engines, t-o weight, CG, . . .)
d. Air temperature
e. Wind direction and velocity
f. Visibility and weather
g. Runway length and condition
h. Aircraft type and model
i. Preceeding aircraft
j. Safety and emergency regulations
k. Noise abatement
l. Passenger comfort
m. Destination

At first an effort was made to make measurements without regard to the influence of the above variables on the statistical averages. Two methods were used to gather this data. The first method for taking data was with a movie camera located at the top of the tower. (See Figure 2) This turned out to be too close to the runway for good viewing after take-off. Difficulty in operating the camera made it difficult to reduce the data. Poor visibility caused additional difficulty in exposure settings. A study of 20 aircraft from 10:30 AM on July 26, 1972, to 11:30 that same day (AM) with poor visibility and a 10 mph wind from the West creating a tailwind for all aircraft, revealed the following statistics. Most aircraft completed their landing (all wheels down) at Point (A) which is located 2,300 feet from the beginning of the runway. Of the nine aircraft landings

that were studied, they all completed their landing within 300 feet of Point (A). The departing aircraft attempted to rotate close to Point (B) which is 5,400 feet from the beginning of the runway. Of the seven departing aircraft observed rotation was accomplished from 300 feet before Point (B) to 1,200 feet past Point (B). The takeoff angle was determined for three aircraft and ranged from 3° to 6° with the smaller angle corresponding to the aircraft with the longest takeoff distance.

Another technique was also employed to study the average operating characteristics. This method provided an opportunity to examine aircraft flight records of take-off weights. By combining the aircraft weights with the type of aircraft and its takeoff performance curves, a takeoff distance could be determined. By collecting a sample of these for July 14, 1972, from Delta and Eastern a sample could be calculated for a typical day. Upon examining this data it was noted that takeoff weight was highly dependent upon the aircraft destination, since the landing weight was usually the factor that determined the takeoff weight. An aircraft scheduled for a long flight usually took-off heavier than an aircraft scheduled for a short flight. This could easily make a differance of 2,000 feet in take-off distance. In general these calculations show that most aircraft are scheduled to take-off at a point between 5,000 to 6,000 feet from the beginning of the runway.

Results of the Case Study

Figures 3, 4, and 5 illustrate the types of data that was available from air carrier records, tower operation records and FAA records. In addition, wind data from the weather bureau is shown in Figures 6, 7, and 8. Figure 5 can also be used to set up a typical days operation in the Atlanta Airport for computer simulation.

A summary of the observations made July 26, 1972, can be written as:

a. Typical take-off distance ... 5,500 feet
b. Standard deviation (T-O) ... 500 feet
c. Typical take-off angle ..... 4.5°
d. Standard deviation (T-O) ... 1.5°
e. Typical Landing distance ... 2,300 feet
f. Standard deviation (L) ..... 300 feet

g. Landing angle (typical) ..... 3°
h. Standard deviation (L) ..... 1/2°
i. Typical wind ..... 7.3 mph, SouthWest
j. Standard deviation ..... 4.4 mph, 50°

Calculation of the Probability of Hazard

The calculation of the likelihood of a dangerous wake turbulence encounter actually occuring can be calculated at this time if the radius of hazard is known and a scan plane is selected for any two aircraft. A typical scan plane configuration is illustrated in Figure 9 with the vortex and radius of hazard shown. If a detector system is used in conjunction with the predictor system, then the position of the vortex can be located with a high degree of certainty. The location of the second aircraft (encountering aircraft) can not be determined with any real certainty until after it has left the runway. By measuring the location of the first aircraft (generating aircraft) as it passes through the scan plane, and also measuring the wind conditions locally, the position of the vortex can be accurately calculated for some future time when the encountering aircraft will pass through the scan plane.

With the help of the airport statistics, a probability of the encountering aircraft passing through any point in the scan plane can be calculated. Now for the encountering aircraft to have better than one chance in a million to be in contact with any part of the hazard radius, the vortex core and the center of the encountering aircrafts probability distribution should be seperated by a distance of

$$d = 5.28\,\sigma_E + R_H \quad ,$$

where R_H is the hazard radius around the vortex and σ_E is the standard deviation for the encountering aircraft. (See Appendix I for developement).

When the probability distribution is included in a more detailed calculation useing circular distribution functions similar to those used to calculate the probability for hitting a target[14] the the distance d is reduced for the same one in a million probability. A typical value of σ_E of 233 feet and R_H of 150 feet reduces the distance d to

$$d = 4.8\sigma_E + R_H$$

The calculations are described in Appendix II for this method.

The need to improve the accuracy in measuring the Standard Deviations will show up directly in the predictions for the Probability of Hazard. Similarly the accuracy with which the vortex location is determined will directly affect the predictions for the Probability of Hazard. There is reason to believe that as more data is collected on the aircraft operations the region of the airport where the aircraft rotate on take-off will turn out to be an extremely dangerous region because of the limited space to maneuver. For this reason the radius of hazard will be large at rotation. Similarly the standard deviation would be expected to be small at rotation and grow as the aircraft climbs out. Figure 10 illustrates this concept.

Monte Carlo Method

In an effort to model actual airport conditions, it is possible to use the standard deviations to set up aircraft operations and wind conditions. By using a normally distributed random number generator to select the flight conditions, the computer simulator can fly sample operations in a manner similar to an actual airport operation. This model can be studied in an effort to evaluate aircraft seperation times in the airport. Appendix III describes a typical random number generator which may be used on the IBM 1108.

The computer simulation model is now in the advanced stages of its developement, and can be efficiently applied to the evaluation of the system design requirements prior to component testing. During testing questions concerning system design can be quickly answered on the computer simulator. The simulator will provide a first cut at the design of the predictor system which will act as the heart of the ATVWS package. Future FAA design and integration requirements can benefit from a study of the computer simulation developed for the ATVWS and may be used as a model for other systems simulation of airport systems.

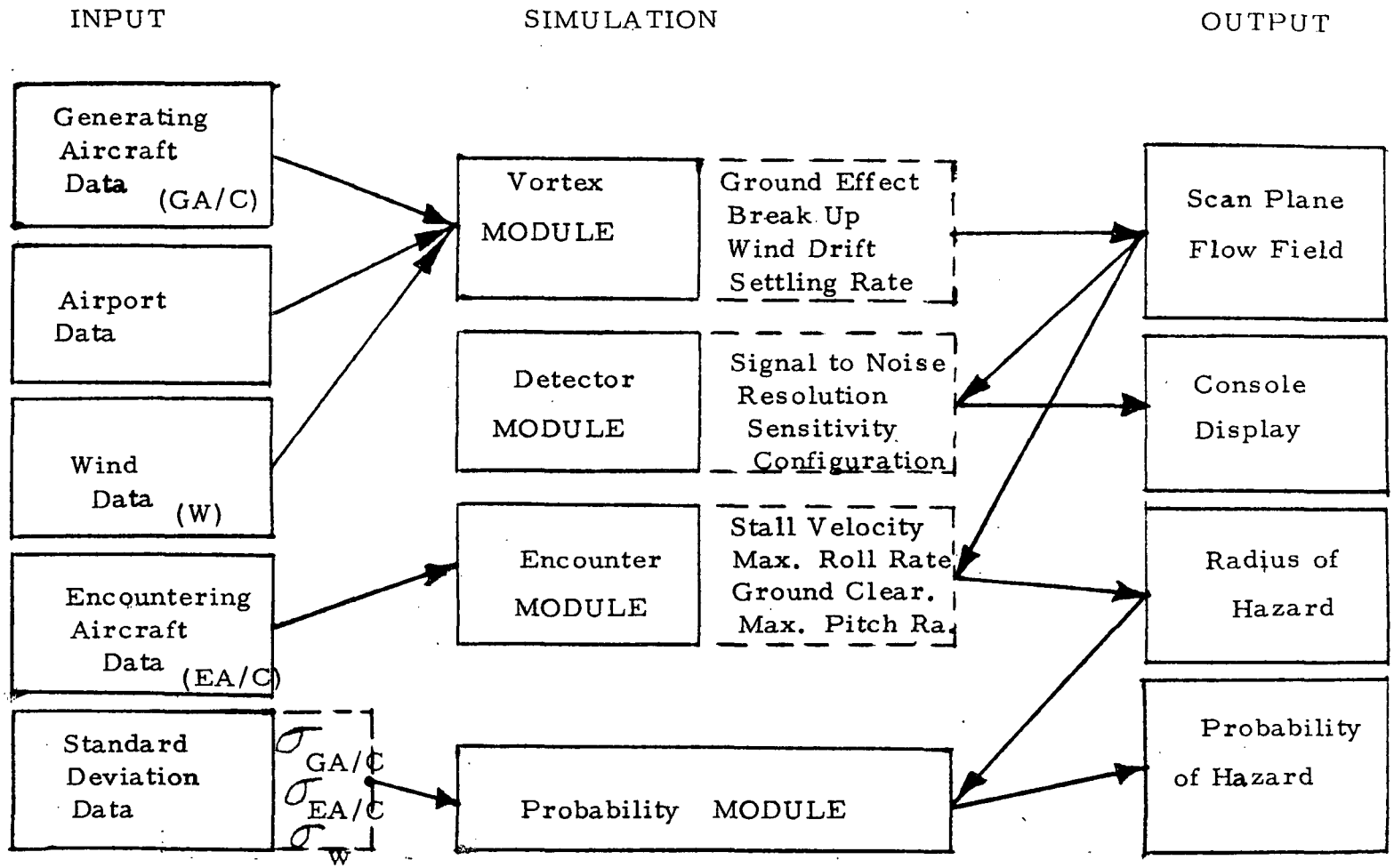


Figure 1. System Simulation Model

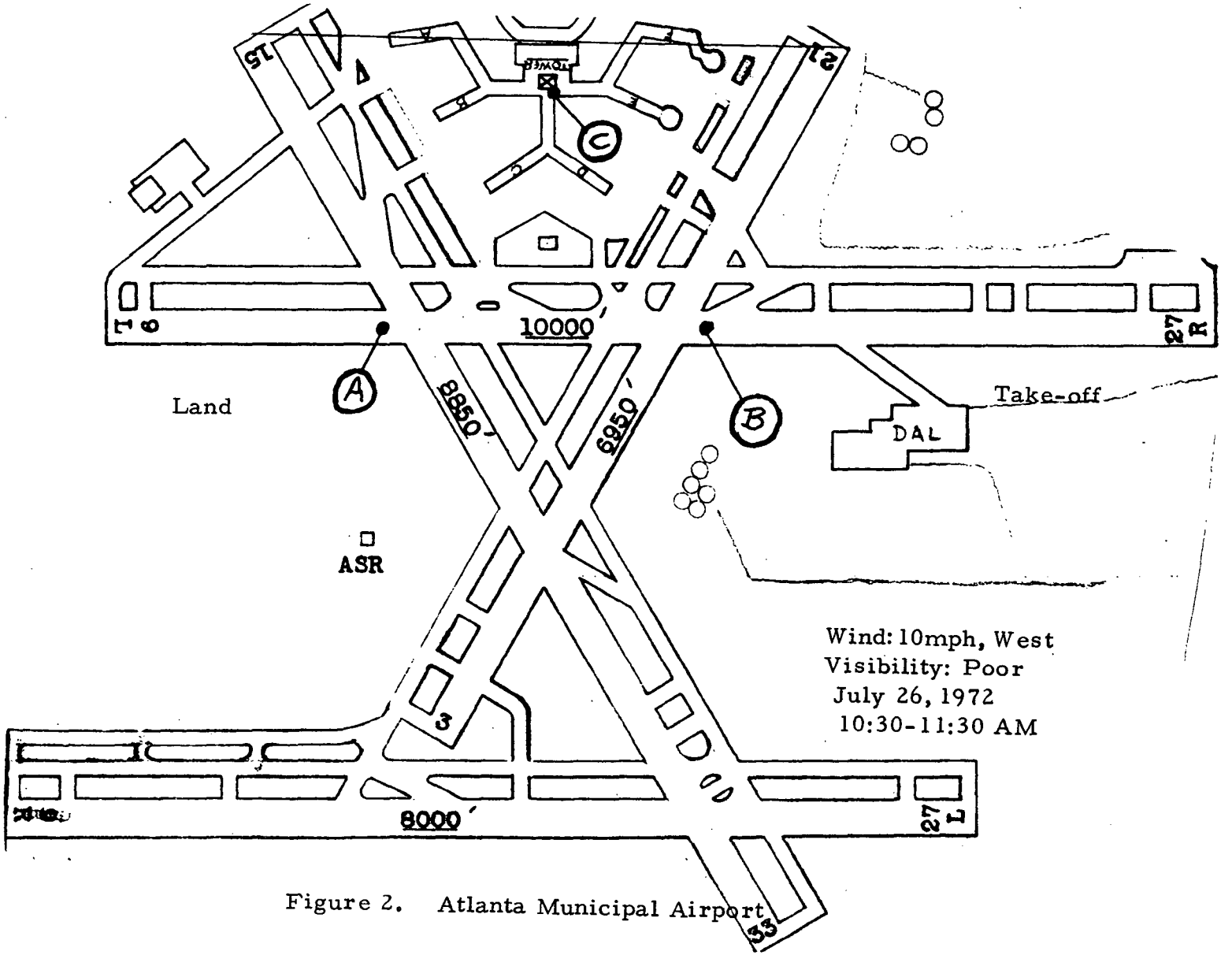


Figure 2. Atlanta Municipal Airport

**Typical Take-off Weights on July 14, 1972. Atlanta Mun. Apt.

TYPE	Takeoff speed	Takeoff Flap	Landing Speed	Landing Flaps	Max. Land Weight	Max. TO Weight	MxTypical T.O.Wt.	Actual Avg T.O. Wt.	Standard Deviatn.
	kt		kt		x1000	x1000	x1000	x1000	x1000
Douglas									
DC-4			97kt		63.5	73.	70.		
DC-6			110		88.2	106	100		
DC-7						143			
*DC-8	162	25°	133	50°	207	315			
*DC-8, Cargo	163		140		210	325	217	190	5
*DC-8, Stretch	163	23°	150	50°	275	355	250	218	6
DC-9	144	20	134	50°	81.7	90.7	86	81	4
DC-9, Stretch			126		95.3	98	96	88	6
*DC-10						453			
Boeing									
*B 707	170	14°	137	50°	190	327			
B 720			131		175	229			
B 727	138	25°	133	40°	142.5	160	145	133	6
B 727, Stretch						169	158	140	5
B 727, Cargo						169	158	147.7	2
B 737						93.5			
*B 747	170	1°	142	40°	564	710	600	495	20
Convair									
880			140		155	193	170	150	5
990						253			
Lockheed									
*L-1011						411			
*C5-A	140	10°	131	41°	635	728			
Learjet	132	20°	127	40°	11.9	13			

Figure 3. Typical Aircraft Weights

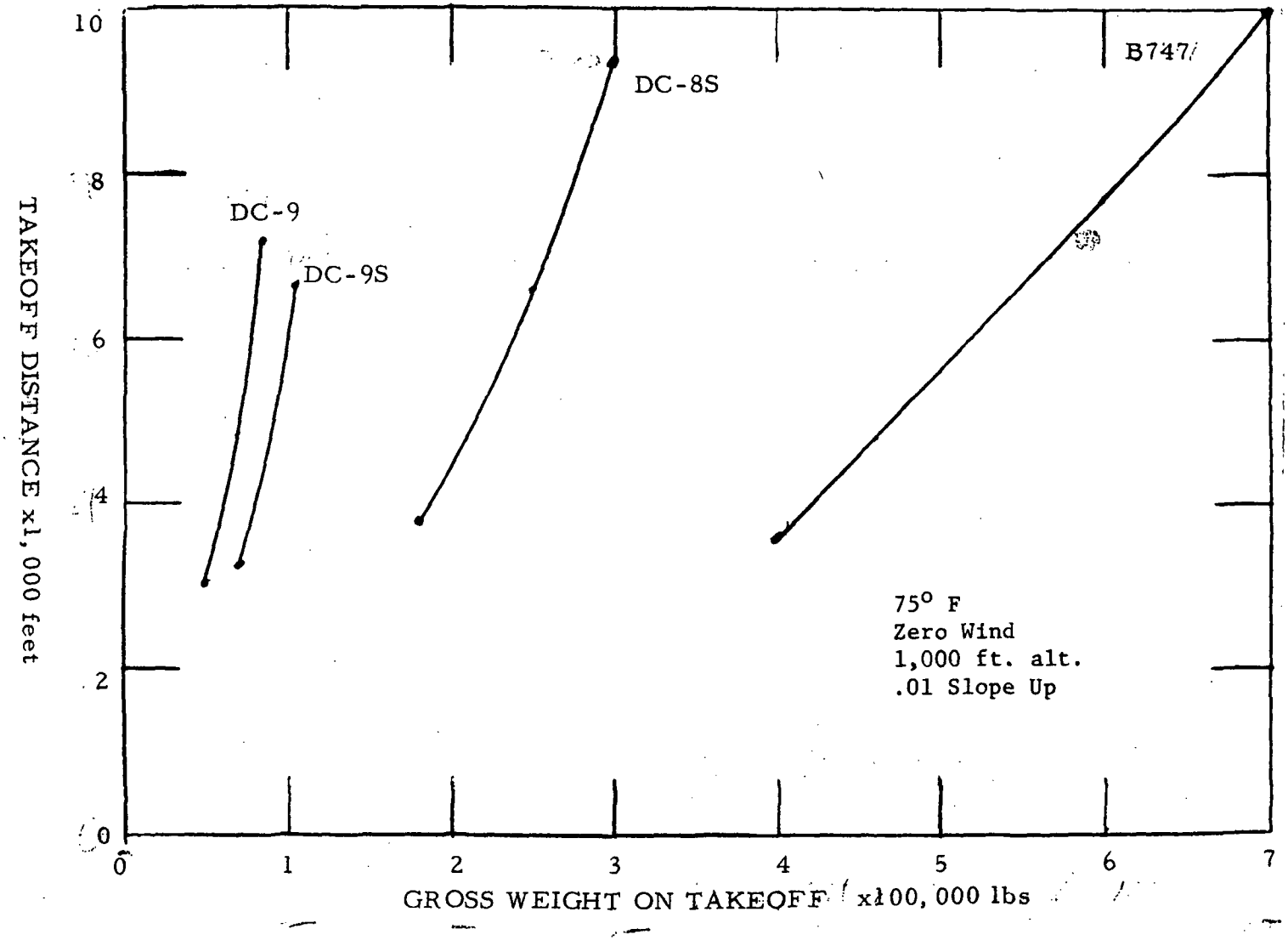


Figure 4. TakeoffWeight versus Takeoff Distance

TIME: AM	FLIGHT #	A/C	TIME: AM	FLIGHT #	A/C
10:20	E9710	B727 Q	10:35	EA651	DC9
10:20	DL791	DC9S	10:35	EA+(&	B727
10:20	EA252	DC9S	10:37	DL617	DC9S
10:20	EA511	DC9S	10:45	*DL41	B747
10:21	DL788	DC9S	10:45	EA266	DC9S
10:21	EA9131	B727 Q	10:45	EA447	DC9S
10:22	DL397	DC9	10:45	EA485	DC9S
10:23	DL627	DC9S	10:45	EA491	DC9
10:23	DL783	DC9/S	10:45	EA663	B727
10:24	DL883	DC8S/F	10:45	SO141	DC9
10:25	* DL44	B747	10:45	SO161	DC9
10:26	DL592	DC9S	11:00	EA542	DC9S
10:27	DL616	DC9S	11:00	EA675	B727
10:28	DL308	DC9	11:00	EA715	DC9
10:29	DL796	DC9S	11:05	SO65	DC9
10:30	DL897	DC8F	11:10	EA665	DC9
10:30	EA441	DC9S	11:10	PI242	DC9
10:30	EA467	DC9	11:20	EA725	DC9S
10:30	EA582	B727	11:25	E9060	B727
10:30	EA642	DC9S	11:30	EA704	B727
10:30	EA997	B727			
10:30	PI4	DC9			
10:31	DL670	DC9S			
10:32	DL690	DC9S			
10:34	DL894	DC8F			
10:35	DL797	DC9S			
10:35	EA649	DC9S			

July 1972 Schedule 10:20 to 11:30 AM

Figure 5. Schedule of Flights

Month	Temperature: Normal: Daily (a)	Daily minimum (b)	Monthly (b)	Extremes: Record highest	Year	Record lowest	Year	Wind: Mean speed	Prevailing direction	Fastest mile: Speed	Direction (a)	Year
(a)		(b)	(b)	11		11		33	14	29	29	
J	52.0	37.3	44.7	72	1967	-3	1966+	10.6	NW	54	NW	1965
F	53.7	38.4	46.1	79	1962	8	1970+	11.1	NW	59	W	1966
M	60.3	42.5	51.4	83	1967	21	1970+	11.0	NW	69	SE	1969
A	70.1	50.2	60.2	88	1970	31	1970+	10.2	NW	68	SW	1956
M	78.9	59.2	69.1	93	1962	37	1971	8.7	NW	65	SW	1971
J	85.7	67.5	76.6	98	1969+	49	1966	7.9	NW	70	NE	1953
J	87.0	70.7	78.9	98	1970	53	1967	7.4	SW	56	SE	1954
A	86.6	69.8	78.2	98	1968	56	1964	7.1	NW	49	NW	1941
S	81.8	64.3	73.1	93	1970+	36	1967	8.0	ENE	49	N	1952
O	72.4	52.4	62.4	88	1971+	29	1968+	8.4	NW	47	NW	1954+
N	60.9	41.5	51.2	84	1961	14	1970	9.2	NW	46	NE	1947
D	52.4	37.1	44.8	77	1971	1	1962	9.8	NW	63	W	1942
YR	70.2	52.6	61.4	98	JUL. 1970+	-3	JAN. 1966+	9.1	NW	70	NE	JUN. 1953

70 Year Averages

Month	Direction	Velocity
Jan	ESE	11.5mph
Feb	NW	11.9
Mar	SW	11.8
Apr	SW	10.8
May	SW	9.3
June	NW	8.2
July	SW	7.8
Aug	SSE/NW	7.4
Sept	ENE/E	8.3
Oct	NW	9.5
Nov	NW	10.7
Dec	NW	11.1

Station ATLANTA, GEORGIA Latitude: 33° 39' N

Month	Temperature: Averages: Daily maximum	Daily minimum	Monthly	Extremes: Highest	Date	Lowest	Date	Wind: Resultant: Direction	Speed	Average speed	Fastest mile: Speed	Direction (a)	Date
JAN	52.2	33.5	42.9	67	14	15	20	29	3.4	8.6	33	NW	26
FEB	55.4	33.2	44.3	72	20+	12	10	28	1.5	8.6	32	NW	13
MAR	58.2	36.8	47.5	72	15	25	9+	30	3.6	9.3	42	SW	19
APR	73.6	48.0	60.8	86	20	32	3	30	3.9	8.9	23	NW	14+
MAY	77.9	55.4	66.7	86	27+	37	4	28	3.6	9.4	65	SW	12
JUN	87.4	67.0	77.2	93	26	57	1	28	2.4	7.2	44	W	16
JUL	84.4	68.2	76.3	91	14	65	17+	23	2.5	7.7	34	NE	14
AUG	84.7	68.7	76.7	89	20+	63	29	09	2.3	7.3	24	SW	22
SEP	81.9	65.4	73.7	87	30+	57	14	08	2.9	8.1	22	NW	11
OCT	76.0	57.5	66.8	88	1	39	11	04	3.3	8.3	22	NW	10
NOV	61.7	39.7	50.7	79	2	26	22	35	2.3	10.0	29	NW	21
DEC	61.1	43.5	52.3	77	16+	28	19+	12	1.3	10.1	34	E	3
YEAR	71.2	51.4	61.3	93	JUN. 26	12	FEB. 10	31	1.2	8.6	65	SW	MAY 12

Figure 6. Atlanta Wind Averages

7/49—6/54

HOUR OF DAY	CLOUDS SCALE 10			WIND SPEED MPH						
	0-3	4-7	8-10	0-3	4-7	8-12	13-18	19-24	25-31	OVER 31
00	55	15	30	3	28	43	21	4		1
01	54	17	29	7	25	40	25	3		1
02	48	21	32	5	27	40	26	2		
03	54	15	32	4	24	45	25	2		
04	48	19	33	6	25	39	26	3		
05	45	19	36	4	28	39	28	1		
06	30	26	44	5	23	42	27	3		
07	34	24	43	3	21	33	39	5		
08	35	22	43	1	15	34	45	5		
09	33	26	41	2	11	29	46	10	1	
10	28	34	38	3	8	25	50	14		
11	23	37	40	2	9	30	39	17	3	
12	22	41	37		9	26	44	21		
13	22	41	37	1	6	27	46	19	1	
14	19	47	34	1	5	24	50	20	1	
15	22	47	31	1	5	27	38	28	1	
16	23	38	39		3	28	49	17	2	1
17	24	30	46		5	26	50	17	2	1
18	30	25	45	2	6	32	46	12	2	1
19	25	21	54	1	9	47	37	5	1	
20	24	27	49	3	20	43	29	5	1	
21	30	27	43	6	22	35	32	6		1
22	32	34	34	8	23	39	26	6	1	1
23	46	25	30	8	23	34	32	1	2	
AVG	34	28	38	3	16	34	37	9	1	+

TOTAL OBSERVATIONS 3720

B PERCENTAGE FREQUENCIES OF WIND DIRECTION AND SPEED

7/49—6/54

DIRECTION	HOURLY DISTRIBUTION OF WIND SPEED IN MILES PER HOUR												AV SPEED
	0-3	4-7	8-12	13-18	19-24	25-31	32-38	39-46	47-54	OVER 54	TOTAL		
N	1	1	1	+							4		6.3
NNE	1	1	+	+	+						3		5.1
NE	1	2	2	+							5		6.5
ENE	1	2	3	1	+						7		8.6
E	1	2	3	1	+						7		7.9
ESE	1	2	3	1							6		7.9
SE	1	2	2	+							5		6.2
SSE	2	4	2	+		+					8		6.5
S	2	2	1	+	+	+					5		6.2
SSW	1	2	2	1							6		7.5
SW	1	3	4	1	+	+					10		8.1
WSW	1	2	3	1	+						6		8.7
W	1	3	4	1	+	+					8		8.3
WNW	+	1	3	1	+						5		8.7
NW	1	3	3	1	+	+					8		7.9
NNW	1	2	2	1	+						5		8.2
CALM	3										3		
TOTAL	19	34	37	9	1	+					100		7.3

TOTAL OBSERVATIONS 3720

Figure 7. Wind Statistics for July

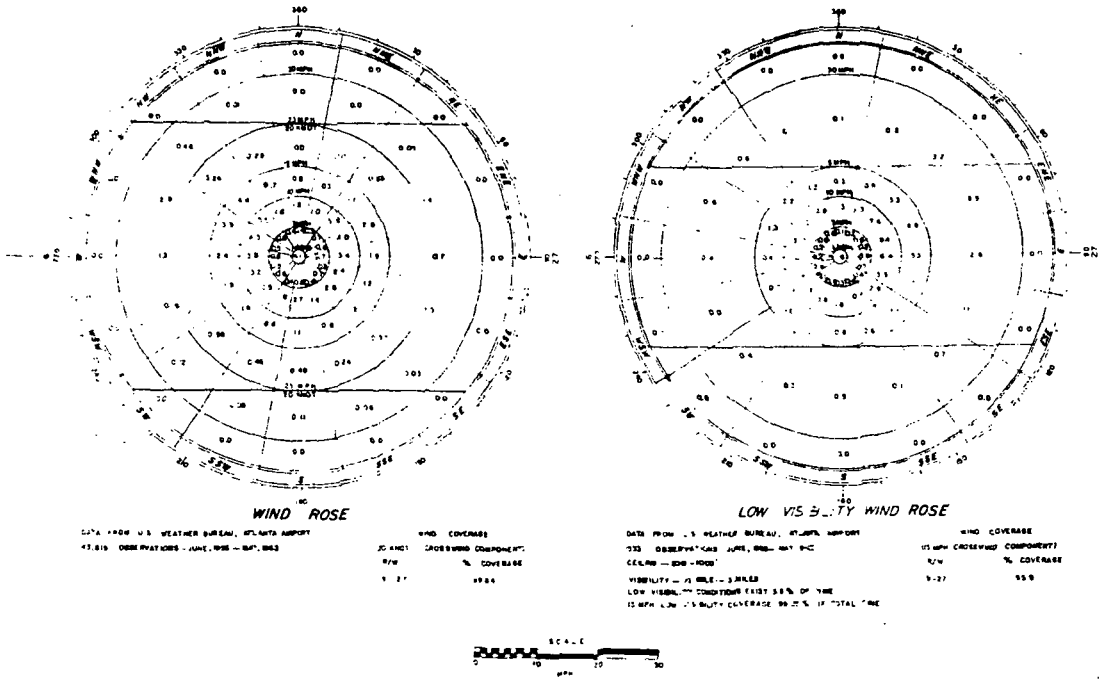


Figure 8. Wind Rose Data For Atlanta

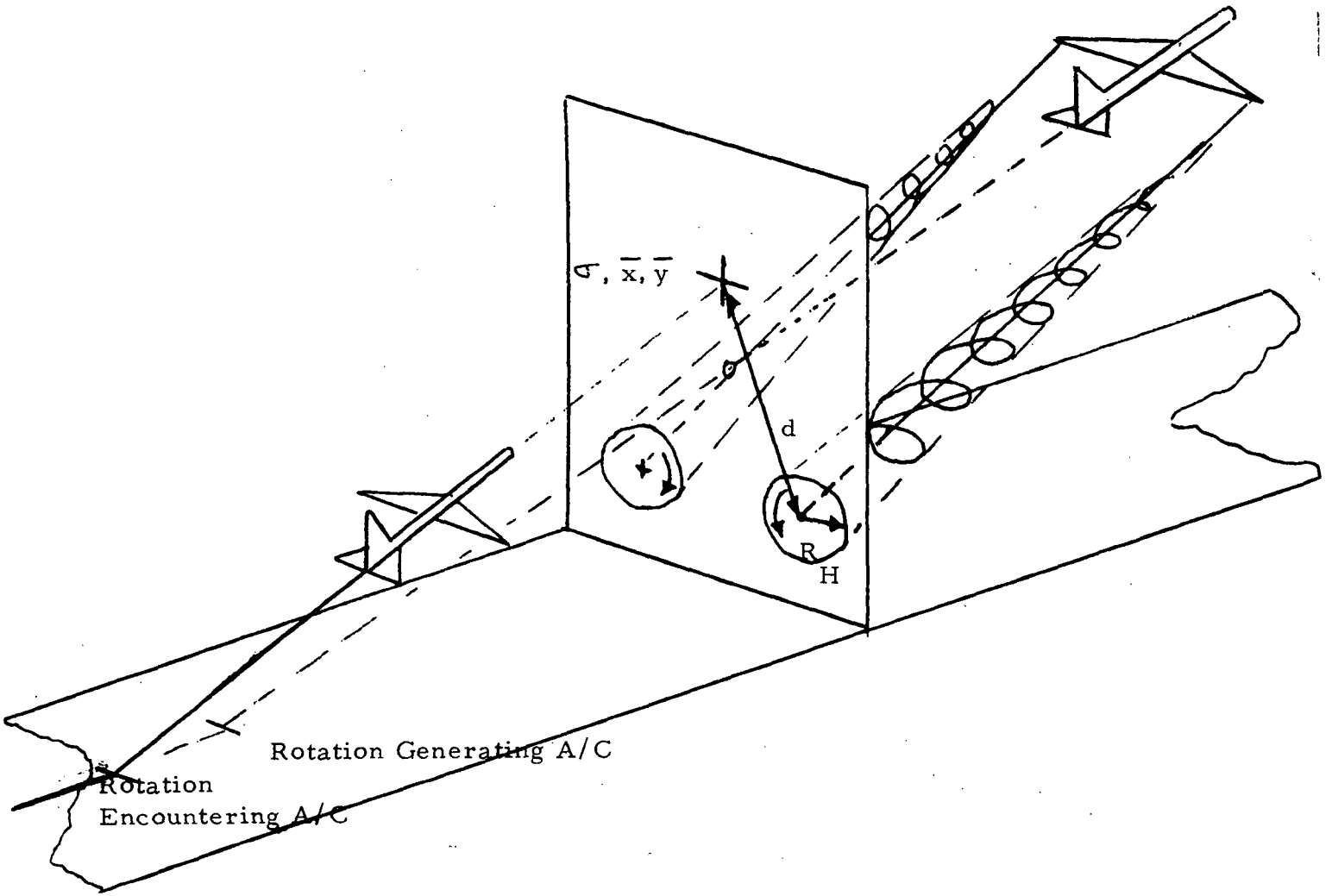


Figure 9. Scan Plane with Vortex Encounter

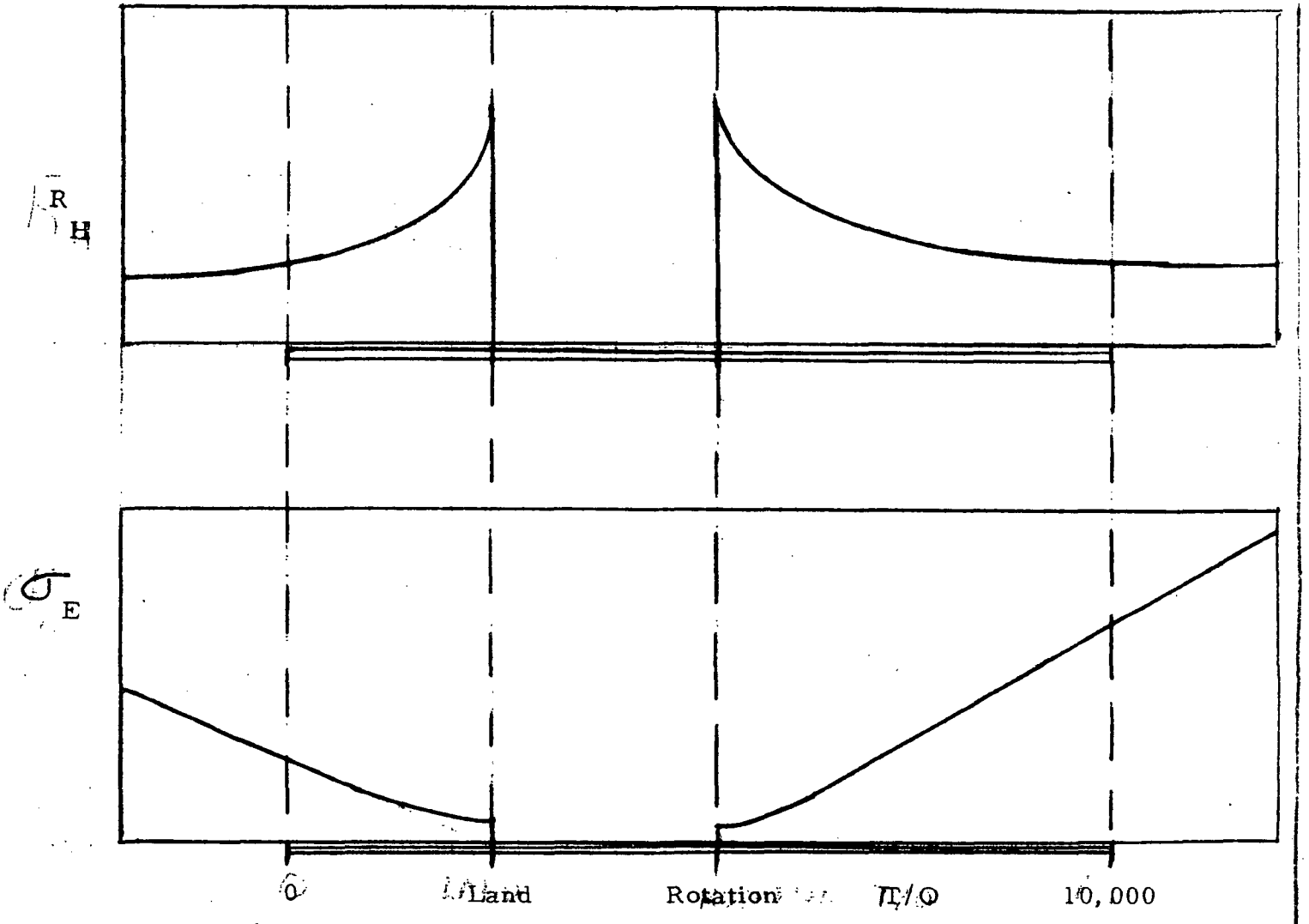
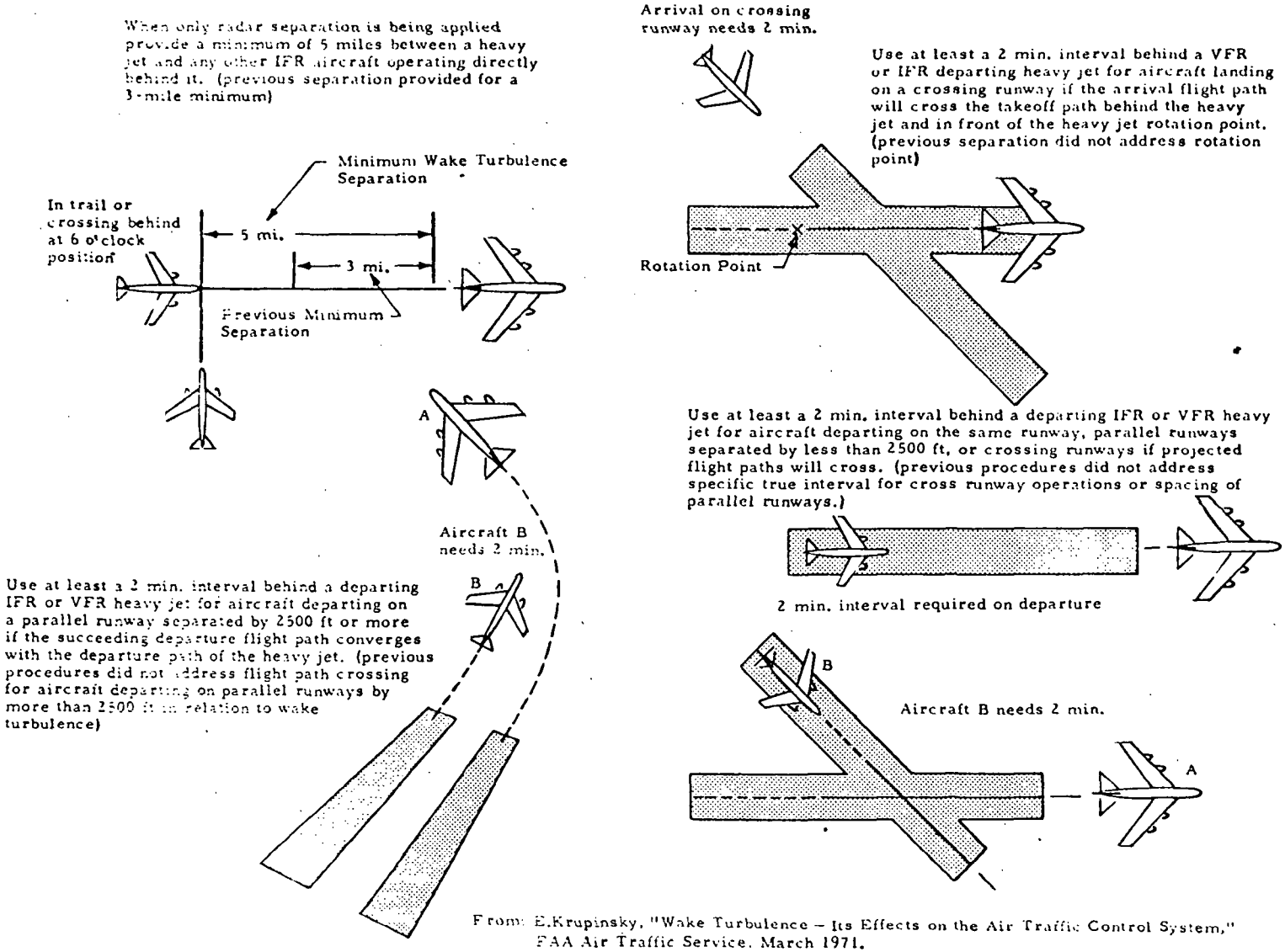


Figure 10. Proposed Standard Deviation and Hazard Radius versus Runway Position



From: E.Krupinsky, "Wake Turbulence – Its Effects on the Air Traffic Control System," FAA Air Traffic Service, March 1971.

Figure 11. REFERENCE: Seperation Distances

APPENDIX I

Circular distribution function

A symmetrical, normal, circular distribution function can be used to calculate the probability that a normally distributed event lies within a radius R. The probability is defined as:

$$P = \frac{1}{2\pi} \int_0^{2\pi} \int_0^R \mathrm{Exp}\left(\frac{-r^2}{2}\right) r \, dr \, d\phi$$

Upon integrating this becomes:

$$P = 1 - \mathrm{Exp}\left(\frac{R^2}{2}\right)$$

If the probability that the event lies outside of radius R and has a value of 10^{-6}, then:

$$P_{outside} = 1 - P = \mathrm{Exp}\left(\frac{-R^2}{2}\right) = 10^{-6}$$

Therefore :

$$R = 5.28$$

Since R in these calculations is dimensionless, and is made dimensionless by dividing by the standard deviation, then the actual radius may be written as:

$R = 5.28\,\sigma$ to the edge of the hazard radius

or

$R = 5.28\,\sigma + R_H$ to the vortex core.

APPENDIX II

Circular distribution function including the hazard radius

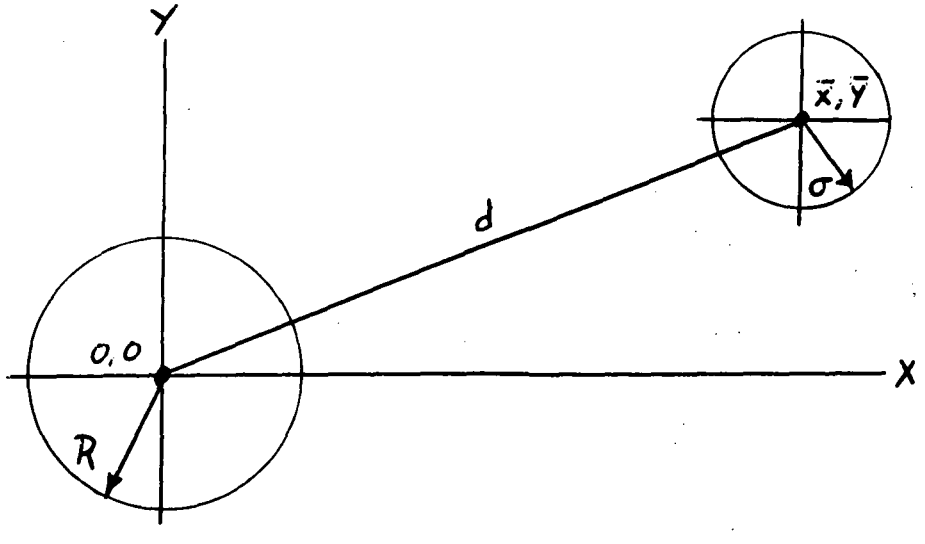
A symmetric, normal, circular distribution function has its center located away from the x, y origin at $\bar{x}, \bar{y}$. The Probability that an event lies inside of a circle about the origin becomes:

$$P = \frac{1}{2\pi} \int_0^R \int_0^{2\pi} \mathrm{Exp}\left(-1/2\left(r^2 + d^2 - 2rd\,\mathrm{Cos}\theta\right)\right) r\, dr\, d\theta$$

where all variables are made dimensionless by division by σ.

Choose the origin to be located at the center of vortex and take R to be the hazard radius. The standard deviation of the encountering aircraft, σ, may now be used to calculate the probability that the aircraft will fly inside of the hazard radius: (See NASA TN-1819)[14].

It can be seen from inspection of the tables that the value of d will be less than the value in Appendix I for all values of interest.



APPENDIX III

Random number generator

A method for generating random numbers that are normally distributed from 0 to 1 is shown here for use on an IBM 1108. This method depends upon the overflow technique of the computer and may not prove suitable for other machines.

The use of this routine depends upon the card:

```
CALL NORMAL(EX, STDX, N, X)
```

where;

EX is the mean value(input)
STDX is the standard deviation (input)
N is the seed number (input only once) :1234567891
X is the required random number (output)

The subroutines to go with this CALL card appear as:

```
      SUBROUTINE NORMAL(EX, STDX, N, X)
      SUM = 0.0
      DO 4 I = 1, 12
      CALL RAND(N, R)
    4 SUM = SUM + R
      X = STDX*(SUM-6.) + EX
      RETURN
      END

      SUBROUTINE RAND(N, R)
      N = N*316227
      R = ABS(N)
      R = R/34359738367.
      RETURN
      END
```

For further reference see Reference 15.

BIBLIOGRAPHY

1. "Wake Turbulence", Advisory Circular AC No 90-23B, Department of Transportation, Federal Aviation Administration, 19 Feb. 1971.

2. Krupinsky, E., "Wake Turbulence- Its Effects on the Air Traffic Control System", FAA Symposium on Turbulence, Wash. D. C., March 22-24 1971.

3. Miller, Barry, "Wake Turbulence Sensors Tested", Aviation Week and Space Technology, pp40-44, August 14, 1972.

4. Basic Aircraft, by W. Green & D. Punnett, Published by Frederick Warne & Co, Ltd., London, England, 1967.

5. Thomas, David D., "Turbulence Related Accidents-World Wide Synopsis", FAA Symposium on Turbulence, Wash. D. C., March 22, 71.

6. Beran, Donald, "Acoustical Radar Detection Techniques", FAA Symposium, Wash. D. C., March 22-24, 1971.

7. Vortex Wake Turbulence, FAA Report No. FAA-FS-71-1, Feb, 71.

8. Macready, P. B., Jr., "An Assessment of Dominant Mechanisms in Vortex-Wake Decay", Symposium on Aircraft Wake Turbulence, Seattle, Wash., Sept 1-3, 1970 (Available from Plenum Press, N. Y., Aircraft Wake Turbulence and Its Detection, Olden, J. H., Goldberg, A., and Rogers, Milton, 1971).

9. Huffaker, R. M., Jelalian, A. V., Keene, W. H., and Sonnenschein, "Applications of Laser Doppler Systems to Vortex Measurement and Detection", Symposium on Aircraft Wake Turbulence, Wash, 70.

10. Huffaker, R. M., Jelalian, A. V., & Thomson, J. A. L., "Laser-Doppler System for Detection of Aircraft Trailing Vortices", Proceedings of IEEE, Vol 58, No 3, March, 1970.

11. Newman, B. G., "Flow in a Viscous Trailing Vortex", Aeronautical Quarterly, 10, 149, June, 1959.

12. Crow, S. C., "Stability Theory for a Pair of Trailing Vortices", AIAA Paper 70-53, January 1970.

13. Handling the Big Jets, Davis, D. P., Air Registration Board, Brabazon House, England, 1971.

14. Weaver, W. C., Wicker, K. C., "Tables for the Integral of the Circular Bivariate Normal Frequency Function", NASA TN D-1819 July, 1963.

15. Naylor, T. H., Computer Simulation Techniques, 1966.

Additional References

16. Madden, S. J., and Harlan, R. R., "A Program to Analyse and Model Trailing Vortices in Airports", MIT, Measurements Systems Lab, Report RN-69, Oct, 1971.

17. Thomson, J. A. L., "Study of Conceptual and Operational Feasibility of Laser Doppler Detection Systems", Research Institute for Engineering Science, Wayne State University, Final Progress Report of Contract NAS 8-24810, DCN 1-9-75-10090(1F), Nov, 30, 1970.

18. McGowan, W. A., "Avoidance of Aircraft Trailing Vortex Hazards", AGARD Flight Mechanics Panel, Baden-Baden, Germany, Oct 20, 1970.

19. Wilson, D. J., Schrider, K. R., and Lawrence, T. R., "Feasibility of Wake Vortex Monitoring System for Air Terminals", Contract NAS 8-26668, Lockheed Report LMSC-HREC, D225936, June, 1972.

20. Goff, R. W., "Study of the Performance Requirements of a Vortex Monitoring System", Contract DOT-TSC-141, Raytheon Report ER72-4197, March, 1972.

21. Jeffreys, Harold B., "Probability of Hazards in the Airport Vicinity Due to Wing Tip Vortices", NASA Marshall Space Flight Center, Memo:S&E-AERO-MM-45-72, August 18, 1972.

22. Jeffreys, Harold B., "ATVWS Simulator . . . ", (To be published) Ref:NASA MSFC S&E AERO MMA, 1972.

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

ORBIT PERTURBATIONS DUE TO SOLAR RADIATION PRESSURE

Prepared by:	Gary A. Sawyer
Academic Rank:	Assistant Professor
University:	South Dakota State University
Laboratory: (Division) (Branch)	Aero-Astrodynamics Laboratory Mission Planning & Analysis Division Orbital Mechanics Branch
Research Counterpart:	Dr. L. D. Mullins
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

ORBIT PERTURBATIONS DUE TO SOLAR RADIATION PRESSURE

By

Gary A. Sawyer

ABSTRACT

This report presents the results of an investigation into the effects of solar radiation pressure upon the orbit of an artificial satellite. This disturbing force will be important for satellites with a large area to mass ratio and also for those whose orbits are high enough that atmospheric drag is not the more dominate force.

The procedure used for the analysis is to represent the radiation force as the gradient of a scalar function to be compatible with existing procedures for studying perturbations due to earth's oblateness. From this analysis, solar radiation pressure appears not to be responsible for any secular or long-periodic variations in the semi-major axis of the orbit nor does it provide any secular changes in the eccentricity of the orbit or the angle of inclination of the osculating plane. Solar radiation pressure does produce secular effects in the other orbital elements, but these are in the opposite sense of secularities caused by the gravitational attraction of the sun and therefore tend to slightly reduce the total secularity.

I. INTRODUCTION

Satellite motion in earth orbit cannot be completely described unless the equations of motion include the various perturbing forces that are influencing the orbiting body. These perturbing forces are small compared to the inverse-square two-body force, but can have a considerable effect on the satellite's orbit. The resulting differential equations of motion are very complicated and analytical solutions are generally unobtainable and "solutions" are found only after making a number of simplifying approximations.

For some applications, it is desirable to be able to construct an approximate analytical solution which is a good representation of the exact solution over a long period of time. It appears that the two-variable asymptotic expansion technique (1) can be utilized for the construction of such an uniformly valid representation of the satellite's motion. Perturbations to the satellite's orbit that are due to forces which can be represented as the gradient of some scalar function can be easily incorporated within the framework of the two-variable technique. The forces on the satellite being additive and the gradient operation having a distributive property, allows the perturbed motion to be determined using superposition techniques. This is desirable since the effects of the various forces can be somewhat isolated and that the problem does not have to be completely resolved whenever another perturbing force is to be considered.

II. DESCRIPTION OF PERTURBATION

This report contains the results of an investigation to include the effects of solar radiation pressure to the existing solution of an artificial satellite moving around an oblate planet. The perturbing effects of solar radiation pressure will be especially important for satellites that have a large ratio of area to mass. The force exerted by radiation impinging upon the satellite depends on the intensity of the radiation, the presented normal surface area, and the reflectivity of the surface. This can be shown by the following discussion:

Assuming that the radiation travels outward from the sun with spherical symmetry, the expression for the amount of radiant energy intercepted by a unit area per unit time, E_R, is given by

$$E_R = \frac{K}{4\pi r^2} \quad \text{joul/m}^2 \text{ sec} \qquad (2\text{-}1)$$

where r is the distance from the sun and K is a constant to be determined by the fact that at the earth's distance from the sun, E_R has a value of about 1.35×10^3 joules/m^2 sec. The momentum of the impinging photons can be determined by considering the radiant energy, the area of the surface normal to the radiation and the velocity of the photons. The amount of momentum transferred to the satellite is determined by the reflectivity of the surface and the time rate of change of momentum is the force on the satellite. Hence, the magnitude of the force on a body in sunlight can be represented by

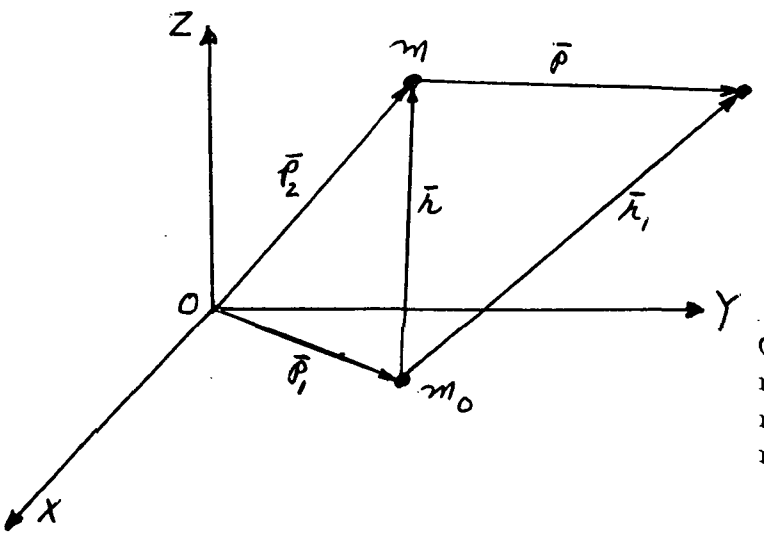
$$\frac{k\,(1+\beta)\,A}{r^2} \qquad (2\text{-}2)$$

where k = constant ($\approx 10^{17}$ nts.)
β = reflectivity ($0 \leq \beta \leq 1$)
A = effective cross-sectional area (m^2)
r = distance from the sun (m)

The direction of the force is radially outward from the sun. The usual assumption that is made is to consider that the force is a constant in both magnitude and direction (since a vector from the sun to the satellite does not change appreciably during one satellite revolution); however, this simplication is not made in the analysis reported here.

III. ANALYSIS OF SOLAR RADIATION PRESSURE

To begin the preliminary discussion, consider Figure 1 where OXYZ is an inertial frame of reference. (It is attempted in this section to follow the notation of Chen (2)). At any instant of time, the three bodies may be located by the vectors shown in the figure.



OXYZ = Inertial Frame
m_0 = earth
m = satellite
m_1 = sun

Considering gravitation forces plus solar radiation pressure, the force equations become

$$M_0 \ddot{\bar{P}}_1 = \frac{G M_0 M}{r^2} \hat{r} + \frac{G M_0 M_1}{r_1^2} \hat{r}_1 - \frac{k (1 + \beta_0}{r_1^2} \hat{r}_1 \qquad (3\text{-}1)$$

$$M \ddot{\bar{P}}_2 = -\frac{G M_0 M}{r^2} \hat{r} + \frac{G M M_1}{p^2} \hat{p} - \frac{k (1 + \beta) A}{p^2} \hat{p} \qquad (3\text{-}2)$$

where G is the gravitational constant and a caret over r, r_1, p represents an unit vector in that direction. It is easy to solve for the acceleration of the satellite with respect to the earth

$$\ddot{\bar{r}} = \ddot{\bar{p}}_2 - \ddot{\bar{p}}_1 = -\frac{G M_0}{r^2} \hat{r} + [G M_1 - k (1 + \beta) \frac{A}{M}] \frac{1}{p^2} \hat{p} +$$

$$-\frac{G M}{r^2} \hat{r} - [G M_1 - k (1 + \beta_o) \frac{A_o}{M_o}] \frac{1}{r_1^2} \hat{r}_1 \qquad (3\text{-}4)$$

let $\mu = G (M_0 + M)$, $\alpha_1 = G M_1 - k (1 + \beta) \frac{A}{M}$, $\alpha_2 = G M_1 - k (1 + \beta) \frac{A_o}{M_o}$

then $$\ddot{\bar{r}} + \frac{\mu}{r^2} \hat{r} = \nabla R \qquad (3\text{-}5)$$

where $R = \alpha_1 (\frac{1}{p} - \frac{1}{r_1}) - \alpha_2 \frac{\bar{r} \cdot \bar{r}_1}{r_1^3}$

and the gradient is determined with respect to the coordinates of the satellite. Taking advantage of the fact that $r/r_1 << 1$, the law of the cosines yields,

$$p^{-2} = [r^2 + r_1^2 - 2 r r_1 \cos \phi]^{-1} \quad \text{or}$$

$$\frac{1}{p} \approx \frac{1}{r_1} + \frac{r}{r_1^2} \cos \phi + \frac{1}{2} \frac{r^2}{r_1^3} (3 \cos^2 \phi - 1) \qquad (3\text{-}6)$$

where ϕ is the angle between r_1 and r. Combining (3-5) and (3-6) gives

$$R = \alpha_1 \left[\frac{r}{r_1^2} \cos \phi + \frac{r^2}{2r_1^3} (3 \cos^2 \phi - 1)\right] - \alpha_2 \frac{r \cos \phi}{r_1^2}$$

$$R = (\alpha_1 - \alpha_2) \frac{r}{r_1^2} \cos \phi + \alpha_1 \frac{r^2}{2r_1^3} (3 \cos^2 \phi - 1) \qquad (3\text{-}7)$$

Once the force has been expressed as the gradient of the scalar function R, the desired form of the Lagrange planetary equations is as follows (3):

$$\frac{da}{dt} = \frac{2}{na} \frac{\partial R}{\partial M}$$

$$\frac{de}{dt} = \frac{(1 - e^2)}{na^2 e} \left(\frac{\partial R}{\partial M} - \frac{1}{(1 - e^2)^{1/2}} \frac{\partial R}{\partial W}\right)$$

$$\frac{di}{dt} = \frac{\cot i}{na^2 (1-e^2)^{1/2}} \left(\frac{\partial R}{\partial W} - \frac{1}{\cos i} \frac{\partial R}{\partial \Omega}\right)$$

$$\frac{d\Omega}{dt} = \frac{1}{na^2 (1-e^2)^{1/2} \sin i} \frac{\partial R}{\partial i} \qquad (3\text{-}8)$$

$$\frac{dw}{dt} = \frac{(1 - e^2)^{1/2}}{na} \left(\frac{1}{e} \frac{\partial R}{\partial e} - \frac{\cot i}{1-e^2} \frac{\partial R}{\partial i}\right)$$

$$\frac{dm}{dt} = n - \frac{1}{na} \left(2 \frac{\partial R}{\partial a} + \frac{1 - e^2}{ae} \frac{\partial R}{\partial e}\right)$$

n = mean motion = $(\mu a^{-3})^{1/2}$
m = mean anomaly

The six orbital parameters defined by the solution of (3-8) determines the position of the satellite. The motion is shown in Figure 2. The position of the osculating plane is specified by the argument of the ascending node Ω and by the inclination i. The orbit orientation in the osculating plane is determined by the argument of perigee ω, and the size and shape of the orbit is defined by the semi-major axis a and the eccentricity e. The position of the satellite is defined by the argument of latitude u, measured between the lines of nodes ON, and the radius vector $\overline{r}$.

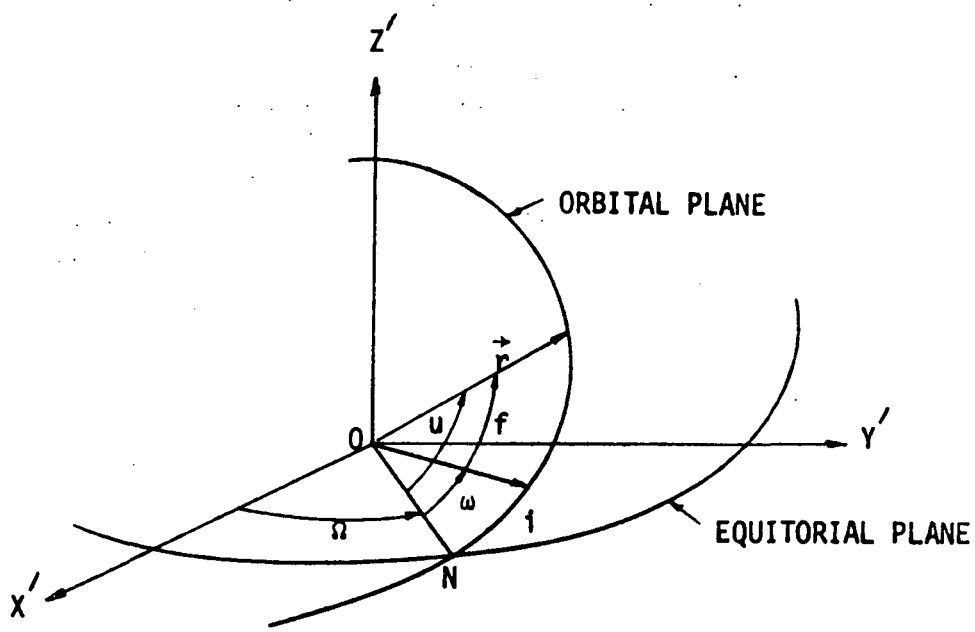


Figure 2.

To perform the partial differentiation indicated in (3-8), it is desirable to express cos $\emptyset$ as a function of the orbital elements. The expression for cos $\emptyset$ is as follows (2),

$$\cos \emptyset = A \cos (\omega + \nu) + B \sin (\omega + \nu) \qquad (3\text{-}9)$$

and $A = \cos u_1 \cos (\Omega - \Omega_1) + \sin u_1 \sin (\Omega - \Omega_1) \cos i_1$

$B = - \cos u_1 \sin (\Omega - \Omega_1) \cos i + \sin u_1 \cos (\Omega - \Omega_1) \cos i \cos i_1 +$
$\sin u_1, \sin i \sin i_1$

where the subscript 1 denotes orbital elements of the sun and ν is the true anomaly. With the aid of (3-9), and some trigonometric identities, (3-7) becomes

$$R = (\alpha_1 - \alpha_2) \frac{a}{r_1^2} \left(\frac{r}{a}\right) [(A \cos \omega + B \sin \omega) \cos \nu + (B \cos \omega - A \sin \omega) \sin \nu] +$$

$$\alpha_1 \left(\frac{a^2}{r_1^3}\right) \left(\frac{r}{a}\right)^2 \left[\frac{3}{2} (A \cos \omega + B \sin \omega)^2 \cos^2 \nu + \frac{3}{2} (B \cos \omega - A \sin \omega)^2 \sin^2 \nu +\right.$$

$3\,(A\cos\omega + B\sin\omega)\,(B\cos\omega - A\sin\omega)\cos\nu\sin\nu - 1/2]$

For this study, it is desirable to focus attention on long term effects and short periodic terms can be neglected. For this purpose, the disturbing function R can be averaged over a satellite period. Define $\bar{R} = (1/2\pi)\int_0^{2\pi} R\,dM$ then considering all elements except r and ν as constants, the indicated integration can be performed. The following results can be easily obtained and are used to get the desired form for $\bar{R}$.

$$\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{r}{a}\right)\cos\nu\,dM = -\frac{3e}{2}$$

$$\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{r}{a}\right)\sin\nu\,dM = 0$$

$$\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{r}{a}\right)^2\cos^2\nu\,dM = 1/2 + 2e^2$$

$$\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{r}{a}\right)^2\sin\nu\cos\nu\,dM = 0$$

$$\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{r}{a}\right)^2\sin^2\nu\,dM = 1/2 - \frac{e^2}{2}$$

$$\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{r}{a}\right)^2 dm = 1 + \frac{3e^2}{2}$$

$$\bar{R} = -(\alpha_1 - \alpha_2)\frac{3ae}{2r_1^2}(A\cos\omega + B\sin\omega) + \frac{3\alpha_1 a^2}{2r_1^3}\left\{(1+\frac{3e^2}{2})(\frac{A^2}{2}+\frac{B^2}{2}-\frac{1}{3}) + \right.$$

$$\left.\frac{5e^2}{2}\,[AB\sin 2\omega + (\frac{A^2}{2} - \frac{B^2}{2})\cos\omega]\right\}$$

For this disturbing function, the differential equations to be solved are

$$\dot{a} = 0 \qquad (3\text{-}11)$$

$$\dot{e} = \frac{3K_2 (1 - e^2)^{1/2}}{2\,na} \left(\frac{a_1}{r_1}\right)^2 (B \cos \omega - A \sin \omega) - \frac{15K_1 e(1-e^2)^{1/2}}{4n}$$

$$\left(\frac{a_1}{r_1}\right)^3 [2\,AB \cos 2\omega - (A^2 - B^2) \sin 2\omega]$$

$$(i)^{\circ} = - \frac{3K_2\, e\, c \cos \omega}{2\,na\,(1 - e^2)^{1/2}} \left(\frac{a_1}{r_1}\right)^2 + \frac{3K_1\, C}{4\,n\,(1 - e^2)^{1/2}} \left(\frac{a_1}{r_1}\right)^3 [(2+3e^2)\, A + 5e^2$$

$$(B \sin 2\omega + A \cos 2\omega)]$$

$$\Omega^{\circ} = - \frac{3K_2\, e\, c \sin \omega}{2\,na\,(1 - e^2)^{1/2} \sin i} \left(\frac{a_1}{r_1}\right)^2 + \frac{3K_1\, C}{4\,n\,(1 - e^2)^{1/2} \sin i} \left(\frac{a_1}{r_1}\right)^3$$

$$[(2 + 3e^2)\, B + 5e^2\, (A \sin 2\omega - B \cos 2\omega)]$$

$$\dot{\omega} = - \frac{3K_2 (1 - e^2)^{1/2} (B \sin \omega + A \cos \omega)}{2\,nae} \left(\frac{a_1}{r_1}\right)^2 + \frac{3K_1 (1 - e^2)^{1/2}}{4n} \left(\frac{a_1}{r_1}\right)^3$$

$$[3A^2 + 3B^2 - 2 + 10AB \sin 2\omega + 5\,(A^2 - B^2) \cos 2\omega] - \dot{\Omega} \cos i$$

$$\dot{M} = n + \frac{3K_2 (1 + e^2)}{2\,nae} \left(\frac{a_1}{r_1}\right)^2 (B \sin \omega + A \cos \omega) - \frac{K_1}{4n} \left(\frac{a_1}{r_1}\right)^3$$

$$[(7 + 3e^2)\,(3A^2 + 3B^2 - 2) + 30\,AB\,(1 + e^2) \sin 2\omega + 15\,(A^2 - B^2)$$

$$(1 + e^2) \cos 2\omega]$$

where $C = \frac{\partial B}{\partial i}$ (See equation 3-9)

$$K_1 = \frac{GM_1}{a_1^3} - \frac{k\,(1 + \beta)\,A}{a_1^3\,M} = \frac{\alpha_1}{a_1^2} \qquad (3\text{-}12)$$

$$K_2 = \frac{k}{a_1^2} \left[(1 + \beta_0) \frac{A_0}{M_0} - (1+\beta) \frac{A}{M} \right] = \frac{\alpha_1 - \alpha_2}{a_1^2}$$

IV. SECULAR EFFECTS

The set of differential equations presented in the last section contain the long period and secular perturbation terms due to the presence of the sun in the satellite problem. (Recall that the short period terms were averaged out of the equations.) To separate the secular terms from the periodic terms, observe that if any of the terms on the right hand side of equation (3-11) contains a constant term or a periodic term with a non-zero average value, then that term will contribute a secular term to the orbital parameters. Therefore, eliminating the long periodic terms and retaining those which produce secular effects, reduces equation (3-11) to,

$$\dot{a} = \dot{e} = \dot{(i)} = 0$$

$$\dot{\Omega} = -\frac{3K_1 \cos i\,(2 - 3\sin^2 i_1)\,(2 + 3e^2)}{16\,n\,(1 - e^2)^{1/2}} \left(\frac{a_1}{r_1}\right)^3 \qquad (4\text{-}1)$$

$$\dot{\omega} = \frac{3K_1 (1 - e^2)^{1/2} (2 - 3\sin^2 i_1)\,(2 - 3\sin^2 i)}{16\,n} \left(\frac{a_1}{r_1}\right)^3$$

$$\dot{M} = n - \frac{K_1 (7 + 3e^2)\,(2 - 3\sin^2 i_1)(2 - 3\sin^2 i)}{16n} \left(\frac{a_1}{r_1}\right)^3$$

V. DISCUSSIONS

The effects of the differential equations derived in the previous two sections can be most clearly viewed by considering the two constants defined by equation (3-12). For typical satellite parameters, these constants can be approximated by

$$K_1 = \frac{GM_1}{a_1^3} - \frac{k\,(1 + \beta)\,A}{M\,a_1^3} \approx \frac{GM_1}{a_1^3} \qquad (5\text{-}1)$$

$$K_2 = \frac{k}{a_1^2}\left[(1 + \beta_0)\frac{A_0}{M_0} - (1 + \beta)\frac{A}{M}\right] \approx -\frac{k\,(1 + \beta)\,A}{a_1^2\,M} \qquad (5\text{-}2)$$

This implies that all the terms containing K_1 are due almost entirely to solar gravitation attraction and terms containing K_2 are produced by solar radiation pressure. (For comparison purposes, $|K_2|$ is on

the order of magnitude of $|aK_1|$.) With the above considerations, it is seen that all elements except the semimajor axis contain long period terms, but the semimajor axis is affected by short period variations only and is therefore not subject to substantial perturbations. This can be reasoned by the argument that on the average the net gain in energy during one portion of the orbit is equal to a loss occurring during another part of the orbit.

Considering the differential equations governing secular effects (4-1), it is seen that all terms containing K_2 have been eliminated and only K_1 terms remain. Only three of the parameters contain secular effects and these are almost entirely due to solar gravitational attraction. The only role that radiation pressure seems to play is to reduce the magnitude of the secular terms slightly as evidenced by equation (5-1). This is understandable since the two forces are directed out of phase with one force adding energy to the orbit when the other is decreasing the orbital energy.

The actual solution to the solar perturbations is not presented here but the effects can be corporated into an existing solution for the effects of earth's oblateness if that solution is modified to accept forces which result in scalar disturbing functions which are not independent of the argument of the ascending node.

VI. SHADOW EFFECTS

In Section III, both the solar radiation force and the sun's gravitational force were seen to be inverse-square law forces. However, there is an important difference between these forces in that the solar radiation pressure is zero whenever the satellite is in the earth's shadow. This causes the radiation force to be discontinuous and the averaging analysis of that section may not apply. To perform a detailed analysis of the shadow effect, it will be necessary to consider two disturbing functions R_1 and R_2, where

$$R_1 = K_2 r \left(\frac{a_1}{r_1}\right)^2 \cos \phi + \frac{K_1 r^2}{2} \left(\frac{a_1}{r_1}\right)^3 (3 \cos^2 \phi - 1) \qquad (6\text{-}1)$$

and applies when the satellite is in sunlight and

$$R_2 = \frac{G m_1}{a_1^3} \frac{r^2}{2} \left(\frac{a_1}{r_1}\right)^3 (3 \cos^2 \phi - 1) \qquad (6\text{-}2)$$

is to be used when the satellite is in the shadow region. The satellite would be within the earth's shadow whenever the following two conditions are satisfied simultaneously:

$$\cos \phi < 0 \qquad (6\text{-}3)$$

and $$r \sin \phi < r_0$$

where r_0 is the radius of the earth. The shadow entry and exit points would have to be determined for each orbit of the satellite. This means that the short period terms may have to be retained since the discontinuity in the force could possibly cause an accumulative effect over a long interval of time (3). The radiation pressure will have no effect on the period if the orbit is circular. If the orbit is noncircular and partly in shadow, the satellite can enter and leave the shadow region at different distances from the sun, resulting in a net gain or loss of energy from the radiation field. In fact, there has been cases where for a short period of time an artificial satellite has gained more energy from solar radiation than was lost due to air drag (4). However, it is also believed that over the long term shadow effects do not change the nature of the perturbations although it produces a change in their amplitudes (5, 6). At the present, it is uncertain how the shadow effect will be incorporated into the perturbation analysis.

REFERENCES

1. Kevorkian, J., "The Two Variable Expansion Procedure for the Approximate Solution of Certain Nonlinear Differential Equations," in Space Mathematics Part 3, J. B. Rosser, Ed. (American Mathematical Society, Providence, R. I.), p. 206.

2. Chen, C. H., "A Study of Lunisolar Perturbation in the Motion of Artificial Satellites," Inter-Office Memorandum No. 9242-72-71, Northrop Services, Inc., Huntsville, Alabama, March 1972.

3. Kozar, Y., "Effects of Solar Radiation Pressure on the Motion of an Artificial Satellite," Smithsonian Institution Astrophysical Observatory, Special Report No. 56, January 30, 1961.

4. Zadunaisky, P. E., I. I. Shapiro, H. M. Jones, "Experimental and Theoretical Results on the Orbit of Echo I," Smithsonian Institution Astrophysical Observatory, Special Report No. 61, 1961.

5. Levin, E., "Solar Radiation Pressure Perturbations of Earth Satellite Orbits," AIAA Journal, Vol. 6, 1. 120, 1968.

6. Musen, P., "The Influence of the Solar Radiation Pressure on the Motion of An Artificial Satellite," Journal of Geophysical Research, Vol. 65, No. 5, p. 1391, 1960.

Page Intentionally Left Blank

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

A MULTI-RING IONOSPHERIC PLASMA PROBE

Prepared by:	John W. Sheldon
Academic Rank:	Associate Professor
University:	Florida International University
Laboratory:	Space Sciences Laboratory Nuclear & Plasma Physics Division Plasma & Solid State Physics Branch
Research Counterpart:	Nobie H. Stone
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

LIST OF FIGURES

NOMENCLATURE

a	=	probe cylinder outside radius
$\vec{c}$	=	ion velocity relative to the satellite
c_s	=	satellite speed
e	=	electron charge
f	=	radial function
I_g	=	geometric capture ion current
I_i	=	net current to ith collector
J_1	=	first order Bessel function
k	=	Boltzmann's constant
$\vec{k}$	=	unit vector pointing upstream
m	=	ion mass
n	=	ion density
n_{eo}	=	ambient electron density
n_o	=	ambient ion density
r_i	=	outside radius of ith ring
R	=	resistance
T	=	electron temperature
v_e	=	average ambient electron thermal voltage
V	=	electric potential on the end plate
V_s	=	sweep voltage
ΔV_{ij}	=	potential difference between i and j rings
α_ℓ	=	ℓth zero of zero order Bessel function
ε_o	=	dielectric constant of free space
η	=	ratio: (electron density)/(ion density)
λ	=	Debye length
ξ	=	plasma flow parameter
ϕ	=	electric potential
ϕ_p	=	plasma potential

Subscripts

e	=	electron
i,j	=	collector indicies

NOMENCLATURE (continued)

ℓ	=	index for numbering zeros of the zero order Bessel function
p	=	plasma
s	=	satellite
o	=	ambient
1	=	first order perturbation

A Multi-Ring Ionospheric Plasma Probe

By

John W. Sheldon

ABSTRACT

An ionospheric plasma probe has been constructed which consists of a long cylinder with the end facing the flow closed by an end plate made up of multiple annular rings and a center disk. A theoretical argument is given which yields the plasma potential and electron temperature in terms of known plasma parameters and the currents to the various rings of the end plate.

This probe was successfully operated in an ionospheric flow simulation facility and the resulting plasma potential (1.21 $\pm$ .05 volts) is in excellent agreement with the traditional Langmuir analysis (1.22 volts).

INTRODUCTION

Electrostatic probes known as Langmuir probes are commonly flown on artificial earth satellites to monitor the electron density and electron electron temperature in the ionosphere. These probes are actually measuring the electron gas properties in a plasma disturbed by the motion of the satellite and the probe rather than properties of the ambient ionosphere. The relationship between the disturbed plasma properties and the ambient ionosphere has been the topic of considerable theoretical investigation [Al'Pert, et al. (1963); Gurevich, et al.(1969); Liu (1969); Taylor (1967)].

The results of these investigations have been applied to the operation of a Langmuir probe in a flowing plasma. In the present paper, the theoretical procedure of Taylor (1967) is used to design an electrostatic probe for use in an ionospheric plasma, and preliminary experiments in an ionospheric flow simulation device are described. It is shown to offer a convenient alternative to the Langmuir probe for measuring plasma potential in a flowing collisionless plasma.

The probe consists of a long conducting cylinder with its axis oriented parallel to the flow. The end plate facing the flow consists of multiple, electrically isolated conducting concentric rings and a center disk. A theoretical description of the electrical current to this probe is developed in the following sections; the pertinent details of the probe design, instrumentation and operating procedures are described; and prelimianry results are given.

OBJECTIVES

The primary goal of this work has been to develop a diagnostic technique for determining plasma potential that will provide a convenient alternative to the traditional Langmuir analysis. The method should take into account the special nature of the flowing ionospheric plasma. Of secondary, but significant importance, is the successful application of the Taylor heuristic theory to an ionospheric flow problem which can be investigated in the laboratory. Agreement between the theory and experiment will lend confidence to further computations on other body configurations.

THEORY

A typical artificial earth satellite travels through the ionosphere at a speed c_s well above the ambient ion thermal velocity, but well below the ambient electron thermal velocity v_e. The ion mean free path is much larger than typical vehicle dimensions in the satellite altitude regime (above 200 km). The flow of rarefied charged particles and the self-consistent electric field which develops in the vicinity of a conducting body under these conditions is usually treated as follows [Taylor (1967)]. The electron distribution is given by the equilibrium distribution for charged particles

in an electric field. The ion distribution is determined by the Vlasov equation. Both the electron and ion distributions are dependent on the electric potential which must be determined consistent with Poisson's equation. The ambient ion and electron temperatures T are assumed equal. The vehicle surface is considered to be a perfect conductor, and reflection of charged particles from the surface is neglected.

By using the above assumptions, the dimensionless steady state Poisson, energy and continuity equations are given by

$$\nabla^2\phi = \eta e^{\phi} - n \quad , \tag{1}$$

$$(\vec{c} \cdot \vec{\nabla})\,\vec{c} + \xi\vec{\nabla}\phi = 0 \quad , \tag{2}$$

and

$$\vec{\nabla} \cdot (n\vec{c}) = 0 \quad , \tag{3}$$

respectively. The electric potential ϕ relative to the plasma potential, ϕ_p is in units of kT/e where k is Boltzmann's constant and e is the electron charge. Distance is in Debye lengths, $\lambda = (\varepsilon_o kT/e^2 n_o)^{1/2}$ where ε_o is the dielectric constant of free space. The spacially dependent ion velocity relative to the satellite, $\vec{c}$ is in units of c_s, the ion density n is in units of the ambient ion density n_o, $\eta = n_{eo}/n_o$ where n_{eo} is the ambient electron density and $\xi = kT/mc_s^2$, where m is the ion mass. For the ionospheric plasma $\xi << 1$. Following Taylor $n(\vec{r}) \simeq 1 + \xi n_1(\vec{r})$ and $\vec{c}(\vec{r}) \simeq -\vec{k} + \xi\vec{c}_1(\vec{r})$, where $c = n = 1$ for the undisturbed plasma and $\vec{k}$ is a unit vector pointing in the upstream direction. Using these approximations in Equations (1), (2) and (3) one obtains

$$\nabla^2\phi = \eta e^{\phi} - 1 \tag{4}$$

$$\vec{\nabla} \cdot (-n_1\vec{k} + \vec{c}_1) = 0 \tag{5}$$

$$-(\vec{k} \cdot \vec{\nabla})\vec{c}_1 + \vec{\nabla}\phi = 0 \tag{6}$$

The solution of Equations (4), (5), and (6) with $\eta = 1$ and $\phi << 1$ was obtained in cylindrical coordinates in a previous work [Sheldon

(1972)]. The boundary conditions for a semi-infinite cylinder at the plasma potential and the front disk at a potential differing by a small value V from the plasma potential were used to calculate the net current to the front disk.

For the present purpose the front disk is made up of a disk and annular rings with equal areas and potential V relative to the plasma potential. It is a trivial extension of the previous note [Sheldon (1972)] to calculate the current I_i to the ith annular ring between radii r_{i-1} and r_i including the possibility of $\eta \pm 1$. The result can be put in the dimensional form

$$I_i = I_g\{[1 - \eta(v_e/c_s)][(r_i/a)^2 - (r_{i-1}/a)^2] \qquad (7)$$

$$-[(4eV\eta v_e)/(kTc_s)][f(r_i/a) - f(r_{i-1}/a)]\}$$

where a is the cylinder radius, I_g the geometric capture ion current is given by

$$I_g = en_o c_s \pi a^2$$

and

$$f(r/a) = \sum_{\ell} \frac{rJ_1(\alpha_\ell r/a)}{\alpha_\ell^2 \, aJ_1(\alpha_\ell)} \quad .$$

J_1 is the first order Bessel function and α_ℓ is the ℓth zero of the zero order Bessel function. The radial function f(r/a) is given in Figure 1.

In obtaining Equation (7) the condition

$$\xi << \eta(v_e/c_s)$$

which is valid for typical satellite ionospheric flow has been used.

If the r_i are chosen so that the disk and annular rings making up the front plate are of equal areas, then, when V = 0, Equation (7) gives all the I_i equal. Physically this is the condition in which both the cylinder and end plate are at the plasma potential so no sheath will exist and the charged particles arrive without electric field deviations. The use of this effect to determine the plasma potential is described in the following section.

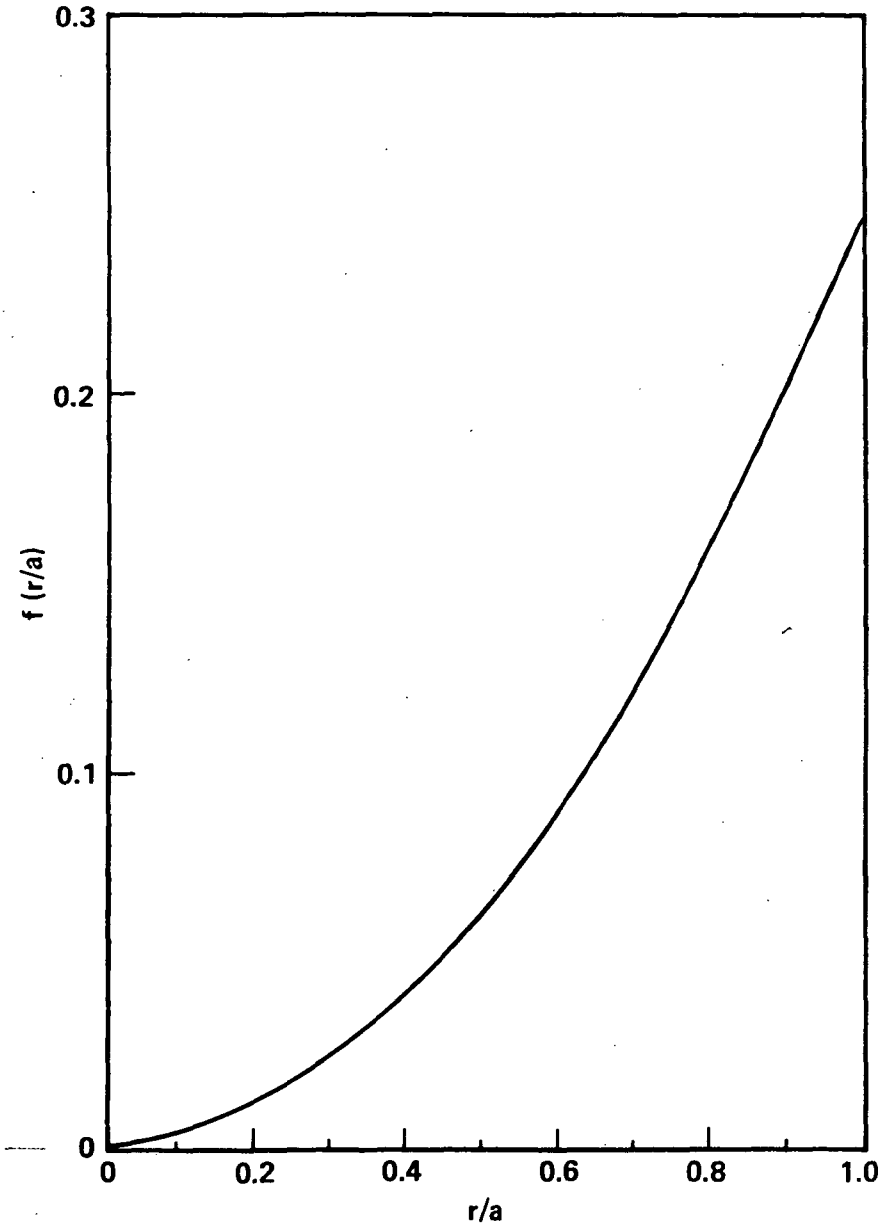


Figure 1. The function f(r/a).

PROBE DESIGN AND INSTRUMENTATION

A sketch of the probe is presented in Figure 2. All of the conducting surfaces are stainless steel. The tube is 23 cm long and 3.68 cm O.R. The disk and rings making up the front plate are attached to a lava insulator with a vacuum grade resin. The ring and disk dimensions were originally chosen to give equal collector areas but a construction error produced an error variation of $\pm$ 10%. Another construction error left a 0.25 cm wide insulator ring exposed between the outer ring (no. 4) and the cylinder. Because of this later error no data are reported for ring no. 4.

A diagram of the probe instrumentation is shown in Figure 3. The voltage difference ΔV_{ij} is related to the current differences since

$$\Delta V_{ij} = R(I_i - I_j)$$

where the positive side of the meter was always on the inner-most collector. ΔV_{ij} was amplified at A_1 and put in the Y channel of X-Y recorder no. 1. The sweep voltage V_s, a ramp function from the signal generator G, was put in the X channel of recorder no. 1. The cylinder voltage was either swept at V_s or fixed by the power supply P according to the position of the switch S_2.

When the sweep voltage, V_s, passes the plasma potential $I_i = I_j$ for all combinations of i and j [V = 0 in Equation (7)] and $\Delta V_{ij} = 0$, that is, all ΔV_{ij} versus V_s curves should go through $\Delta V_{ij} = 0$ when $V_s = \phi_p$. This can be used to determine the plasma potential. R (1 megohm) was chosen so that $RI_i \ll \phi_p$. In an initial run at a given plasma condition ϕ_p is estimated by the curve crossings with the cylinder voltage swept (V_s). Then with this estimated plasma potential placed on the cylinder by the power supply P the front plate voltage is again swept and the curve crossings noted. Consistent results were usually obtained after two sweeps.

The voltage drop across R for the center disk was amplified at A_2 and put into the Y channel of X-Y recorder no. 2 where the same sweep voltage V_s mentioned above served as the X channel. Thus a conventional Langmuir curve could be obtained from this disk for comparison with the ring probe technique.

EXPERIMENTAL RESULTS

The multi-ring probe was placed in an ionospheric flow simulation chamber which utilized a Kaufman ion engine for its plasma source. This facility has been described elsewhere [Stone and Rehmann (1970)]. All data were obtained for an $N_2^+ - \bar{e}$ plasma with

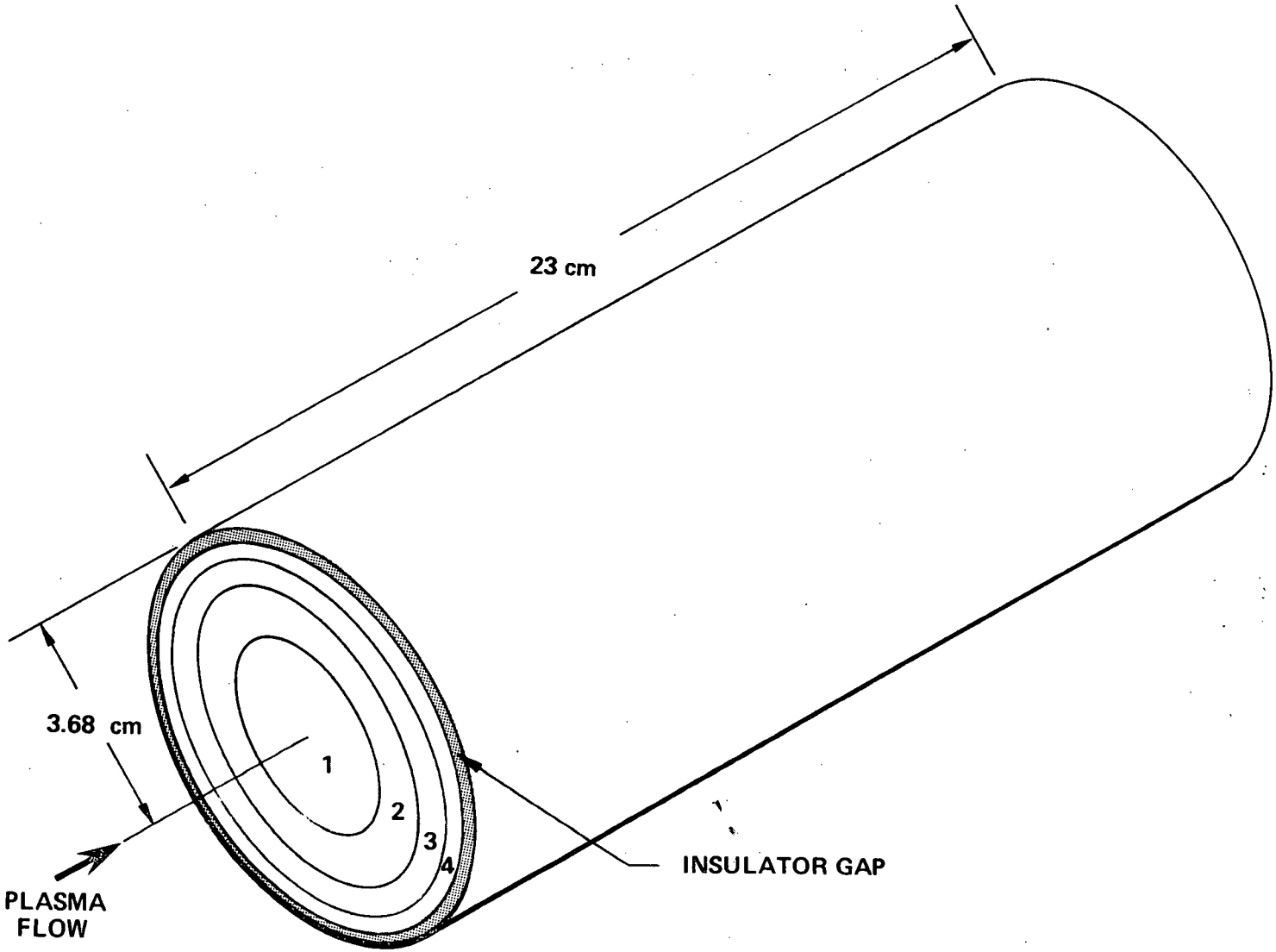


Figure 2. The cylindrical probe with multi-ringed end plate.

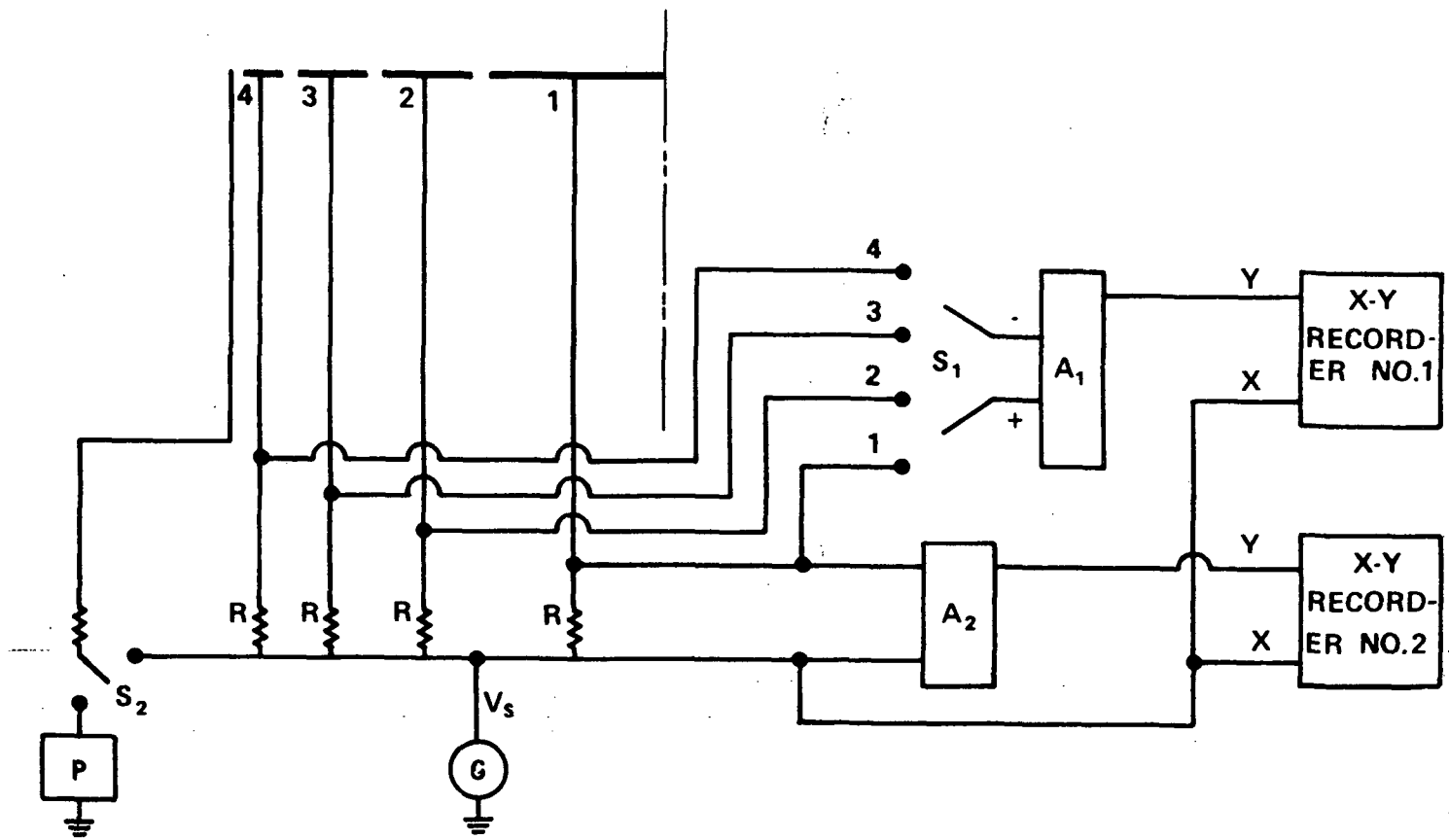


Figure 3. Multi-ring probe instrumentation

the ion energy, E = 20 electron volts and $\eta \simeq 1$ as determined by setting the net engine emission current at zero. The ion density in the test section was typically 10^6 cm^{-3} as determined by a Faraday Cup collector.

Figure 4, a tracing of the X-Y recorder no. 1 output, shows the ΔV_{12}, ΔV_{13} and ΔV_{23} curves that resulted from an initial sweep of V_s with the cylinder potential also swept. Since the curves tend to cross at $V_s \simeq 1.18$ volts the cylinder was biased by the power supply P at this value and the front plate potentials were again swept. The results of this second iteration are presented in Figure 5. The curves now cross at $\Delta V = 0$ giving $\phi_p \simeq 1.21 \pm 0.5$ volts. The output of X-Y recorder no. 2 replotted on a semilog scale for Langmuir-type analysis is shown in Figure 6. The plasma potential determined by this conventional method is 1.22 volts. An electron temperature of 3180°K was obtained from the Langmuir analysis, therefore, the expansion parameter $\xi = (kT)/(2E) \simeq .007$.

CONCLUSIONS AND RECOMMENDATIONS

The utility of the multi-ring probe for determining plasma potential has been demonstrated to a limited extent. Further work is needed in order to have confidence in the method. Certainly the use of Taylor's theory is justified, both on a theoretical basis ($\xi \simeq .007$, where it was assumed that $\xi << 1$), and considering the agreement between the plasma potential as determined by Taylor's theory ($1.21 \pm .05$ volts) with the Langmuir analysis result (1.22 volts).

The multi-ring probe could, in principle, be employed to determine electron temperature if the slope of the $I_i - V$ curves and the current, I_i at the plasma potential were used in Equation (7) with η and c_s known. This would be an interesting and possibly rewarding topic for future work. Laboratory experiments along this line are planned following the correction of inaccuracies in the probe construction.

ACKNOWLEDGEMENTS

It is with pleasure that I recognize the valuable assistance of Mr. Nobie H. Stone on every phase of this project, but especially the laboratory work. The role of our laboratory technician, Mr. H. Scarbrough, is also gratefully acknowledged.

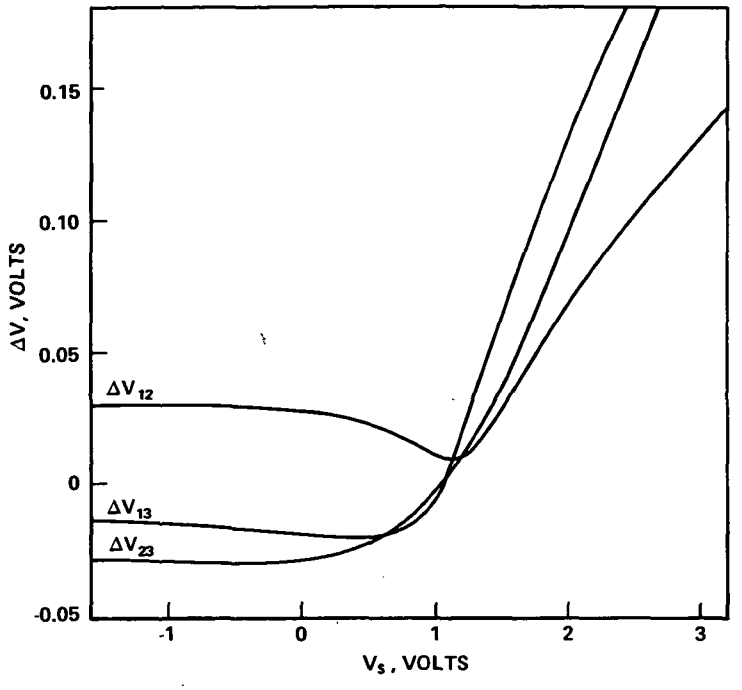


Figure 4. X-Y recorder no. 1 trace, cylinder voltage swept.

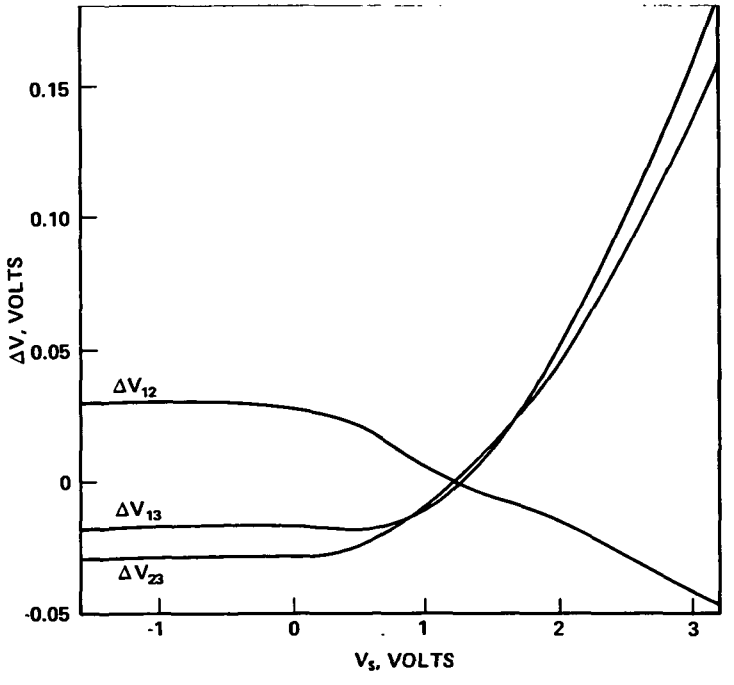


Figure 5. X-Y recorder no. 1 trace, cylinder voltage constant = 1.18 volts.

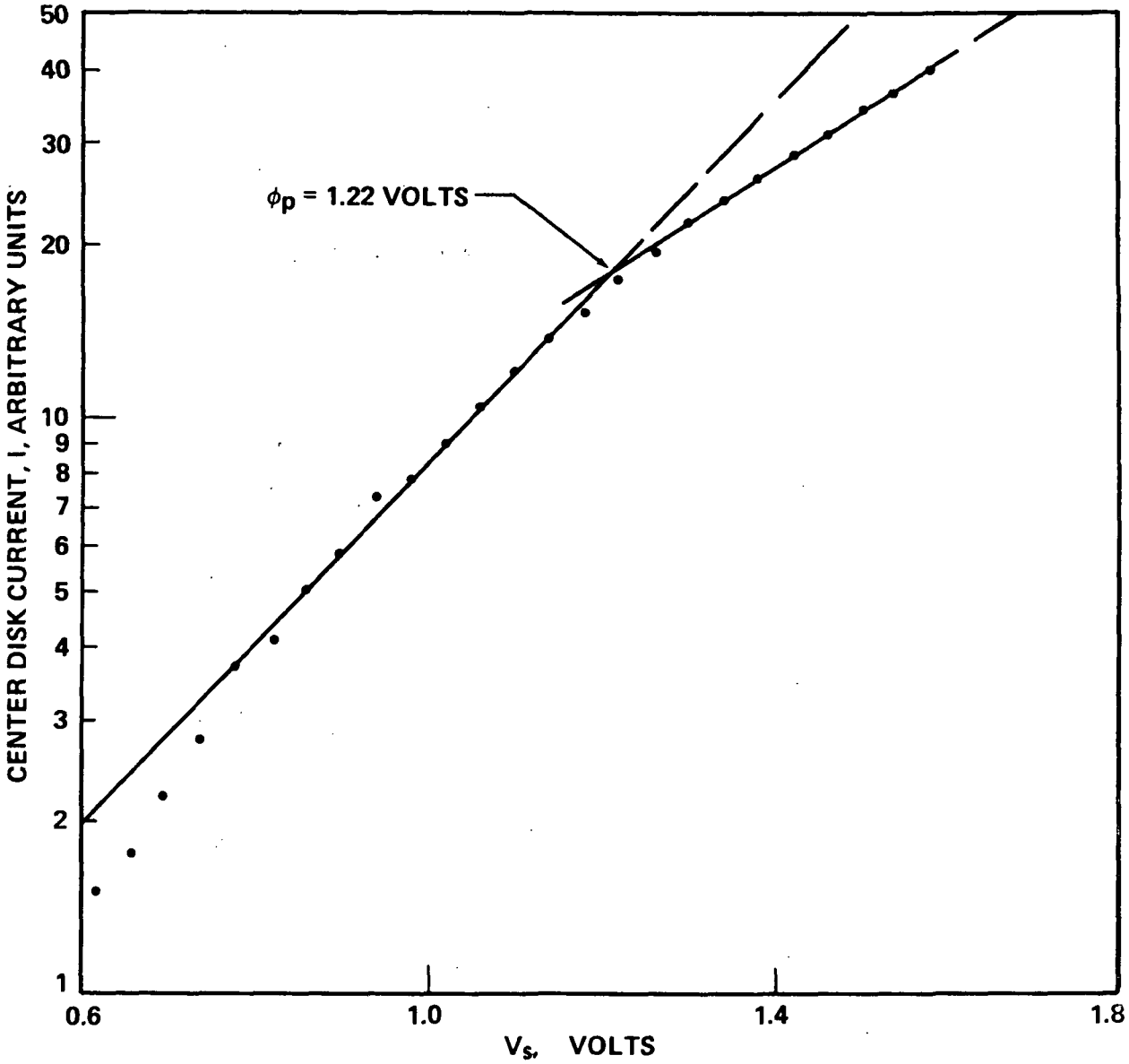


Figure 6. Langmuir plot for collector no. 1, center disk.

REFERENCES

Al'Pert, Ya. L., Gurevitch, A. V., and Pitaevskii, L. P., 1963, Soviet Phys. Usp. 6, 13.

Gurevich, A. V., Pitaevskii, L. P., and Smirnova, V. V., 1969, Space Sci. Rev. 9, 805.

Liu, V. C., 1969, Space Sci. Rev. 9, 423.

Sheldon, J. W., 1972, Planetary Space Sci., to be published.

Stone, N. H. and Rehmann, W. K., 1970, NASA TND-5894.

Taylor, J. C., 1967, Planetary Space Sci. 15, 155.

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

I. CABLE CONNECTED SPINNING SPACECRAFT

II. THE CANONICAL EQUATIONS

III. URBAN MASS TRANSPORTATION

Prepared by:	Amnon Sitchin
Academic Rank:	Associate Professor
University:	University of Alabama
Laboratory: (Division) (Branch)	Aero-Astrodynamics Dynamics and Control Division Orbital Analysis Branch
Research Counterpart:	R. C. Lewis
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

I. CABLE CONNECTED SPINNING SPACECRAFT

II. THE CANONICAL EQUATIONS

III. URBAN MASS TRANSPORTATION

By

Amnon Sitchin

ABSTRACT

Last year's work on the dynamics of cable-connected spinning spacecraft was completed by formulating the equations of motion by both the Canonical equations and Lagrange's equations and programming them for numerical solution on a digital computer. These energy-based formulations will permit future addition of the effect of cable mass. Comparative runs indicate that the Canonical formulation requires less computer time.

Available literature on urban mass transportation was surveyed. Areas of the PRT (Private Rapid Transit) concept of urban transportation to which the expertise of the Dynamics and Control Division could be applied in support of the NASA/DOT program of study and research in automated ground transportation systems are identified.

INTRODUCTION

I. CABLE-CONNECTED SPINNING SPACECRAFT

The work that was begun last year on the cable-connected spinning spacecraft was completed this year.

Two energy-based computer programs that numerically integrate the motion in a marching algorithm are available: one using the Canonic Equation method and one using the Lagrange's Equation method. As is discussed in the next section, the Lagrange's formulation was prepared for comparison only. It is recommended that the Canonic formulation be used.

The Canonic Equations computer program is named CABLES and consists of the following parts:

MAIN - Reads input data, assigns values to the Runge-Kutta coefficients, initializes the momenta values, and cycles the integration routine. Units used are slugs for mass, lbs. for cables, ft. and radians for position, sec. for time.

DERIV - Evaluates the cables lengths, cable forces, generalized cable forces, and the derivatives to be integrated ($\dot{q}$'s and $\dot{p}$'s).

BELMS - Evaluates the direction cosine matrix BA.

PBMAT - Evaluates the elements of the PBA matrix which is the partial derivative of the BA matrix w.r.t. the q's.

AMATAB - Evaluates the elements of the A matrix for the equation $P = AV$.

RKM - Seventh order Runge-Kutta integration routine with stepsize control.

OUTPUT - Outputs the desired information. The current output is: Time, distance of c.g. of body A from origin of coordinates, α, β, γ, ψ, ϕ, θ, a_1, a_2, a_3, the corresponding v's, p's and $\dot{p}$'s. It also calculates and prints the direction cosines of each body's axes relative to the projection of the line connecting the c.g.'s of both bodies on the space z_2 z_3 plane (the plane in which the bodies are designed to spin, preferably undisturbed). See figure 1 for coordinates.

FIGURE 1

The X axes are oriented α, β, γ w.r.t. Z.

The Y axes are oriented Ψ, ϕ, θ w.r.t. Z.

If the Z axes are not inertial (but still not rotating), the linear motion of point Θ can be easily added by modifying the expression for the energy due to the motion of the c.g. If the system is not force-free, the force has to be resolved into generalized components. Please notice that any potential can best be accounted for in terms of the generalized forces derived from it.

The energy formulation presented here permits the treatment of cables with masses and damping. The problem, however, becomes extremely complicated.

II. COMPARISON OF DIFFERENT FORMULATIONS OF THE SYSTEM'S EQUATIONS

The equations of motion were set up in Newtonian, Lagrange's equations, and Canonic formulation, and programmed for numerical solution on the Univac 1108. The Canonic formulation required least computer time. A more detailed study of the different formulation will be published as a TMX.

III. URBAN MASS TRANSPORTATION STUDY

In addition to the material available on file in the Program Development Office, bibliography lists and reference material were obtained from the NASA-DOD computerized index file, the Urban Transportation Literature Collection at the University of Minnesota, the Ford Motor Company and others.

Some observations can be made:

a. The subject of urban mass transportation is of imense proportions.

b. Other than the statement that urban transportation is in a mass, the problem is not defined.

c. Social and political factors will override any technical factors in both the identification of "the problem" and the attempted solutions.

d. Transportation systems will have to be tailor-designed to the location of their application.

Since nothing is, as yet, clearly defined, any work on the subject can be divided into three categories:

a. General studies of available literature and data, leading to a better acquaintance with the problem and proposed systems.

b. Detailed work on a specific system.

c. Establishment of system requirements that are general enough to be regarded as guidelines independent of any specific system or location. Into this category will fall the establishment of safety requirements (similar to current DOT safety requirements for automobiles), establishment of human factors demands such as acceleration and jerk rate for normal and emergency operation, etc.

While NASA could, and probably would, get involved in all of the different modes of automated rapid transit, this study, and the recommendations below, are limited to urban PRT.

The Dynamics and Control Division of the Aero-Astrodynamics Laboratory can participate in the work by:

a. designating at an early date personnel that will work in this area to acquaint themselves with the problem and proposed solutions;

b. establishing human comfort and safety criteria to which all PRT systems will have to conform;

c. reviewing proposed vehicle control modes for their compliance with safety and comfort standards of item b;

d. reviewing proposed vehicle control modes to determine their dependency on assumed hardware (drive, brake, switch) response capabilities. Making a critical examination of the reasonableness of these assumed hardware capabilities;

e. Studying the effect of necessary changes as a result of items c and d;

f. accepting, rejecting, or proposing new operations strategy and control mode resulting from item e;

g. making a sensitivity analysis of all surviving possibilities;

h. doing additional tasks as the need developes.

GLOSSARY OF TERMS

Asynchronous Control - Position and speed related only to car in front or near merge point.

Brake-Light Signal - Information that a car is decelerating is available only to following car.

Brick-Wall Stop - Object generating the need to stop is not moving.

Controlled Collision - Emergency stopping strategy where a car is allowed to slam into the car ahead.

Dual-Mode - Capability to travel on conventional streets as well as on automated guideways.

Functional Trains - Operation with headway less than required for brick-wall no-contact stop.

Headway - The separation between vehicles.

Off-Line Station - Vehicles that stop at stations do not obstruct through traffic.

On-Demand Service - Car made available to passanger at origin for his exclusive use to his selected destination.

Platoon - A string of closely spaced vehicles.

People Movers - Large vehicles; many of the passangers expected to be standing.

Personal Rapid Transit - Small vehicles for 3 to 6 seated passangers.

Quasi-Synchronous Control - The entire network is divided into sectors. Synchronous control is established within each sector.

Scheduled Service - Passangers do not have exclusive use of car, nor control over intermediate stops.

Synchronous Control - The entrie network is composed of position slots, all moving at the same speed, under the control of a single central computer.

ABBREVIATIONS

ACT - Activity Center Transportation

AHS - Automated Highway System

CBD - Central Business District

DMV - Dual Mode Vehicle

O-D Pairs - Origin-Destination Pairs

PRT - Personal Rapid Transit

TACV - Tracked Air Cushion Vehicle

UMTA - Urban Mass Transportation Authority

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

AUTOMATIC CONTROL OF PERSONAL RAPID TRANSIT VEHICLES

Prepared by:	Paul D. Smith
Academic Rank:	Assistant Professor
University:	Auburn University
Laboratory: (Division) (Branch)	Program Development Design Office-Electronics Electronics and Sensors
Research Counterpart:	Charles P. Elms
Date:	August 11, 1972
Contract No.:	NGT-01-003-045

AUTOMATIC CONTROL OF PERSONAL RAPID TRANSIT VEHICLES

by

Paul D. Smith

ABSTRACT

One promising solution to the urban transportation problem is the development of personal rapid transit systems.

The requirements for automatic longitudinal control of a string of closely packed personal vehicles are outlined. Optimal control theory is used to design feedback controllers for strings of vehicles and the previous work in this area is reviewed. An important modification of the usual optimal control scheme is the inclusion of "jerk" in the cost functional. While the inclusion of the jerk term has been considered by other investigators, the effect of its inclusion has not been sufficiently studied. Adding the jerk term will increase passenger comfort.

A proposed outline of investigation is included to solve the optimal control problem and the sensitivity problem associated with developing engineering specifications on the vehicle controls.

INTRODUCTION

The President's State of the Union Message presented to Congress January 20, 1972, committed this nation to a substantial goal for personal rapid transit:

> "... our outstanding capabilities in space technology should be used to help the Department of Transportation develop better mass transportation systems. As has been said so often in the last 2 years, a nation that can send three people across 240,000 miles of space to the moon should be able to send 240,000 people 3 miles across a city to work."

Personal Rapid Transit (PRT) designates the class of fixed-guideway urban transit systems for which the stations are off the main line, the vehicles are automobile size and operate individually under automatic control. Typical PRT trips are non-stop from origin to destination.

Personal Rapid Transit is presently under serious consideration by individuals and organizations throughout the world as a major alternative to present modes of urban transport. Over 50 urban areas in the United States have formally requested the implementation of a PRT system and in both Japan and Germany PRT systems are under construction now, to be in service by 1975.

Personal Rapid Transit has been invented and reinvented in various forms by those who have examined carefully and dispassionately the desirable and undesirable features of both the automobile and conventional forms of public transport. All conventional forms of public transport have been found wanting as all of them involve untenable compromises. Personal Rapid Transit, on the other hand, can be developed to meet the real transportation needs of the urban society without requiring untenable compromises. This statement is illustrated by forty comparisons between conventional transit and PRT listed in Appendix A of Anderson et al (1).

The rational for PRT can be developed as follows: In order to attract a significant fraction of automobile drivers to public transit without coercion, the transit trip must be faster than the auto trip for a very significant number of the trips typically made by urban residents. This requires that the transit vehicles travel on an exclusive guideway either above or below the street level. Accessibility and convenience further require an area-wide network of interconnected guideways.

The major innovation which then leads to PRT is obtained by following the history of highway development. Decades ago, state highways were built around cities in order to speed up interstate travel. This led to the urban freeway. The urban freeway attracts auto drivers in great numbers because, under unsaturated conditions, the trip on a

freeway is non-stop. The driver starts and stops only when these actions have something to do with his own trip. Because he does not have to stop for reasons unconnected with his own trip, his trip is accomplished in minimum time, with minimum frustration and maximum comfort to himself.

A common theme expounds on the many failures of the automobile as a method of urban transportation. Ironically, however, it is the exact reverse, namely the success -- the striking success of the automobile that has created so many of our urban transportation problems. Thus, it would seem that any form of transportation which seeks to replace the form of transportation offered by the automobile by another form is doomed to failure. Indeed, this has been found to be the case even in the most modern of conventional rapid transit systems. Historically, when a modern conventional rapid transit system is opened, one sees a few percentage point rise soon replaced by ever declining numbers of riders.

By adopting the eminently successful characteristics of the urban freeway, one is led to a system of fixed guideways in which the stations are on by-pass tracks off the main line. The by-pass station has the advantage that it permits vehicles to wait for people - not people for vehicles. Maximum capacity is obtained if all acceleration and deceleration maneuvers associated with leaving and entering stations and entering connecting links are accomplished on the by-pass tracks. As an option, these by-pass tracks sometimes can lead not to a station but to the street. This option called "Dual-Mode" has certain advantages and disadvantages when compared with conventional Personal Rapid Transit.

The use of short headways between vehicles is needed to provide adequate capacity with auto-sized vehicles. Small vehicles have two primary advantages. First, the guideways can be light, low cost and have low visual impact. It appears now that the guideways so dominates the overall capital cost that the minumum-size vehicle leads also to the minimum overall system cost even though more vehicles are necessary. Second, the option of truly personal service is available. Truly personal service means that each cab is occupied only by persons travelling together, and that it is unnecessary to travel with strangers. Research has indicated that this is a very important feature. In the personal cabs, each passenger would be provided with a seat. This not only increases comfort and hence patronage, but also permits the acceleration and deceleration to be approximately double that tolerable if passengers are allowed to stand. The result is to cut the ramp lengths approximately in half.

Safe and reliable control of a large number of small vehicles switching in and out of stations and from line to line requires use of advanced techniques of automatic control and reliability engineering developed during the last two decades. When compared with manual control, automatic control has the advantage of considerably reduced labor costs. Service can continue twenty-four hours a day. Automatic control also permits much shorter headways and hence increased capacity.

By interconnecting the network, not only can each trip be non-stop throughout the network, but a system of one-way lines becomes practical. One-way lines will double the percentage of urban land area within walking distance of stations. The mileage of lines and number of stations needed to serve a given percentage of the population is minimized by placing the stations midway between the line junctions. Use of one-way lines also means that the visual impact of the system on a given street is minimized. The provision of vehicle transfer throughout the network has the further advantage that the system can be used for movement of goods and for many kinds of specialized services not possible with conventional transit. All of this means that the system can be amortized over many more vehicle trips than can be attracted to conventional transit, and hence revenues will pay a much larger share of the capital and operating costs than is possible with conventional transit.

The above is adapted from Anderson et al (1) and the reader interested in further details will find this work a good starting point.

OBJECTIVE

The objective of this research is the analysis of a disconnected train of small vehicles operating at relatively high speed and close proximity under automatic control. A further objective is the development of computer programs (analog and/or digital) to study the effects of parameter variations (vehicle mass, sensor error, etc.) and external forces on headway for the various maneuvers (merging, queing, acceleration, deceleration, emergency stopping, etc.). With this program one can determine specifications on the elements of the vehicle and network control system. This effort will provide an engineering tool that can be used in future analyses of these systems.

REQUIREMENTS FOR AUTOMATIC LONGITUDINAL CONTROL OF A PRT VEHICLE

The personal vehicle used in the transportation system discussed above could take several forms depending on the requirements imposed by the enviroment in which it would be used. Regardless of the exact form of the personal-vehicle system, several requirements pertaining to its longitudinal control are fundamental. We shall assume that the vehicles operate on a single lane of roadway and that some form of lateral and steering control is provided. The following set of requirements are similar to those given by Fenton and Bender (7) but have been somewhat enlarged upon here.

1. The system must be capable of maintaining a particular separation between adjacent vehicles. Hereinafter, the term headway will refer to the distance from a particular point on one vehicle to the same point on an adjacent vehicle (say, the distance between the front bumpers). In order to achieve a high traffic flow, the headway should be made as small as possible for some fixed velocity. The requirement to maintain a scheduled headway is closely related to the requirement of maintaining zero relative velocity between adjacent vehicles.

2. Each vehicle in a string of vehicles must be stable with respect to the motion of the vehicle directly ahead. That is, bounded perturbations in the motion of one vehicle must result in bounded perturbations in the motion of the vehicle directly behind. This requirement is termed local stability.
3. Perturbations in the motion of any vehicle in a string must be attenuated as they propagate down the string. This requirement is termed asymptotic stability.
4. The application of large accelerating and decelerating forces to the vehicles must be limited in the interest of passenger comfort. In any case the magnitude of these forces must not exceed the capabilities of the vehicle.
5. Also in the interest of passenger comfort, the time rate of change of acceleration, or jerk, should be limited.
6. The above requirements do not guarantee that the vehicles in a string will keep to the desired position schedule after some perturbation. Hence, for overall system control, position and/or velocity control is required.

The above descriptive statements of requirements will be put into mathematical form when the design of an automatic vehicle control system is formulated in the sequel.

BACKGROUND (A GENERAL SURVEY OF RELATED RESEARCH)

Before going into the detailed analysis of the problem, it will be useful to review in a general way some of the work done so far in the area of High Speed Ground Transportation (HSGT) systems.

The use of modern control theory to design longitudinal HSGT control systems was initiated in the United States in 1964 at the Massachusetts Institute of Technology (MIT) under a directive from and supported by the United States Department of Commerce, Hansen (11). These studies included many aspects of HSGT systems such as vehicle aerodynamics, propulsion, vehicle-guideway interactions, human factors, scheduling, network configuration, computer control, communications, and dynamics and control of vehicle groups. It is the last-mentioned category which is of interest herein.

An early result reported by Levine (13) and Levine and Athans (14) at MIT was concerned with the design of an optimal error regulator for a string of high-speed trains. The requirements of a control system for a string of trains travelling on a single, straight, level guideway were put in the form of a quadratic cost functional and the feedback gains for the optimal error regulator solved for by use of optimal linear regulator theory. The resulting system when simulated on an analog computer appeared to perform in the desired manner. A clever scheme for the optimal merging of two trains (two strings of unconnected vehicles) was investigated by Athans (3). This scheme made use of the optimal feedback system developed earlier for single guideway use. These error-regulating systems required the continuous transmission of the

velocity and position of each vehicle to every other vehicle in the system. To simplify the communications system somewhat, suboptimal forms of the original regulator were studied by Athans et al (5) and Garrard et al (8). This resulted in a reduction in the number of feedback loops required. Simulation studies indicated that the performance of the suboptimal system was only slightly inferior to that of the optimal system. It should be noted that there are no communication advantages to the suboptimal control strategies when central computer control is used. With the idea of further reducing the communication system requirements, Levis and Athans (15) developed an optimal sampled-data control system for high-speed trains. In this system, the position and velocity of each vehicle is measured once every T seconds. The forces applied to each vehicle are held constant during the sampling period. In general, there is a trade-off between the size of the sampling period (which governs channel capacity, band-width, etc., of the communication system) and system performance.

Powner et al (21) and Anderson and Powner (2) have investigated the optimal sampled-data control of high-speed trains in the presence of state measurement noise and random disturbances. Use was made of the discrete form of the Kalman estimator together with an optimal sampled-data feedback system similar to that used by Athans and Levis (15).

Peppard and Gourishankar (20), Wilkie (22), Wilkie and van Schieveen (23),Hajdu et al (10) and Garrard and Kornhauser (9) have considered the case where jerk is included in the cost functional. The resulting system was shown to exhibit a smoother change in acceleration; this is desirable for passenger comfort. All other factors being the same, one would expect that the inclusion of jerk in the cost functional will result in the control taking longer to drive the error to zero. This is obscured in Peppard and Gourishankar (20) because they modified two terms in the cost functional (essentially they went from a car-following to a slot following strategy) and in Wilkie and van Schieveen (23) because of an error in figure numbers.

Peppard (19) and Peppard and Gourishankar (20) give results for a variety of weightings of the terms in the cost functional while Wilkie and van Schieveen (23) discuss the general problem of sensitivity of the results to coefficient weights.

Melzer (16) and Melzer and Kuo (17) have developed a general theory for the optimal control of linear systems comprised of an infinite number of identical objects. The method has been applied to the case of an infinite number of vehicles and the structure for the control system obtained.

The research described up to this point has centered around the concept of using a linear optimal control system to regulate the position and velocity of each vehicle in the system. This approach requires the use of a quadratic cost functional since other forms of performance index would result in a nonlinear controller.

A wider range of problems associated with HSGT systems is considered by Hajdu et al (10), Boyd and Lukas (6) and Anderson et al (1).

MATHEMATICAL MODEL OF A STRING OF VEHICLES

In this section a model similar to that used by many investigators will be developed.

Assume that the system is a string of n vehicles. Let k (k=1,...,n) be an index describing the k-th vehicle in the string. Let z denote the position variable.

The following variables are defined:

$z^k(t)$ = the position of the k-th vehicle at time t

$z_t^k(t)$ = the velocity of the k-th vehicle at time t

$z_{tt}^k(t)$ = the acceleration of the k-th vehicle at time t

$f^k(t)$ = the force applied by the propulsion system to the k-th vehicle at time t

m^k = the mass of the k-th vehicle.

Every vehicle in the string will be subject to drag forces. In general, the drag force will be a function of the vehicle type (on wheels, air or magnetically suspended, etc.), shape, the nature of the guideway (open air, tunnel, etc.), and in particular, it will depend on the velocity $z_t^k(t)$ of the k-th vehicle. Therefore, we let

$g(\, z_t^k(t)\,)$ = the drag force acting on the k-th vehicle at time t.

The general equations of motion of the k-th vehicle can be found from Newton's law. If the guideway is flat, then the equations are

$$(1) \qquad m^k \frac{d^2}{dt^2} z^k(t) = -g(\, z_t^k(t)\,) + f^k(t)$$

for k = 1,2,...,n. If the guideway is not flat, then the only modification to the above is to include in $f^k(t)$ the longitudinal component of force due to the slope.

Because the drag term, g(.), is some nonlinear function of the velocity, $z_t^k(t)$, equation (1) is a nonlinear one. The overall system will be designed in such a way that each vehicle will move with the same desired velocity v_o, under normal operating conditions. The velocity v_o will be called the string velocity. Because of this requirement, we can obtain a linearized differential equation for each vehicle whenever the actual velocity $z_t^k(t)$ is near the string velocity v_o. To

do this we use a Taylor series to expand the drag function about the string velocity v_o

$$(2) \qquad g(\, z_t^k(t)\,) = g(v_o) + Dg(v_o)(z_t^k(t) - v_o) + \text{higher order terms.}$$

The term $Dg(v_o)$ is simply the derivative of the function $g(.)$ evaluated at v_o, i.e., the slope of the drag force curve at v_o. If we let $a_o = Dg(v_o)$, then equation (1) becomes

$$(3) \qquad m^k \frac{d^2}{dt^2} z^k(t) = -g(v_o) - a_o(z_t^k(t) - v_o) + f^k(t) .$$

In the usual way equations (3) can be recast as a system of two first-order ordinary differential equations,

$$(4) \qquad \begin{aligned} \frac{d}{dt} z^k(t) &= z_t^k(t) \\ m^k \frac{d}{dt} z_t^k(t) &= -g(v_o) - a_o(z_t^k(t) - v_o) + f^k(t) \end{aligned}$$

or using matrix notation

$$(5) \qquad \frac{d}{dt} w^k(t) = A\, w^k(t) + h^k(t)$$

where

$$w^k(t) = \begin{vmatrix} z^k(t) \\ z_t^k(t) \end{vmatrix} , \quad A = \begin{vmatrix} 0 & 1 \\ 0 & -a_o/m^k \end{vmatrix} , \quad h^k(t) = \begin{vmatrix} 0 \\ -g(v_o) + a_o v_o + f^k(t) \\ \hline m^k \end{vmatrix} .$$

We shall also want to consider a similar form of these equations in order to include the jerk requirement. For this we let

$$
\begin{aligned}
&\frac{d}{dt} z^k(t) = z_t^k(t) \\
(6)\qquad &\frac{d}{dt} z_t^k(t) = z_{tt}^k(t) \\
&m^k \frac{d}{dt} z_{tt}^k(t) = -a_o z_{tt}^k(t) + f_t^k(t) \quad .
\end{aligned}
$$

The third of equations (6) is obtained simply by differentiating the second of equations (4). Using matrix notation equations (6) become

$$
(7)\qquad \frac{d}{dt} w^k(t) = A\, w^k(t) + h^k(t)
$$

where

$$
w^k(t) = \begin{vmatrix} z^k(t) \\ z_t^k(t) \\ z_{tt}^k(t) \end{vmatrix} \qquad A = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -a_o/m^k \end{vmatrix} \qquad h^k(t) = \begin{vmatrix} 0 \\ 0 \\ \dfrac{-a_o + f_t^k(t)}{m^k} \end{vmatrix} .
$$

We procede now to define the "error" states of the k-th vehicle. Let

$$
(8)\qquad x^k(t) = z^k(t) - v_o t
$$

$$
(9)\qquad x_t^k(t) = z_t^k(t) - v_o
$$

$$
(10)\qquad x_{tt}^k(t) = z_{tt}^k(t) \quad .
$$

Substituting these relations into equations (4) we obtain

$$
\begin{aligned}
&\frac{d}{dt} (x^k(t) + v_o t) = x_t^k(t) + v_o \\
&m^k \frac{d}{dt} (x_t^k(t) + v_o) = -g(v_o) - a_o(x_t^k(t)) + f^k(t)
\end{aligned}
$$

or

$$
\begin{aligned}
&\frac{d}{dt} x^k(t) = x_t^k(t) \\
(11)\qquad &m^k \frac{d}{dt} x_t^k(t) = -g(v_o) - a_o(x_t^k(t)) + f^k(t) = u^k(t) - a_o(x_t^k)
\end{aligned}
$$

defining $u^k(t) = f^k(t) - g(v_o)$.

We can term, $u^k(t)$, the difference between the applied force and the drag force at velocity v_o, as the "control" applied to the k-th vehicle. The term, $u^k(t)$, represents the excess force that is actually applied to the k-th vehicle. In practice, however, it must be remembered that a propulsion system is present on each vehicle and $u^k(t)$ is the desired output of the propulsion system. The dynamics of the k-th vehicle as presented here do not include the propulsion system. The input received by the propulsion system is in practice, say, $\underline{u}^k(t)$. For the purposes of this discussion, we will assume that either the input-output transfer function of the propulsion system is unity (i.e., $u^k = \underline{u}^k$) or equal to a constant r so that $u^k = r\underline{u}^k$. In either event, in the sequel we shall be concerned with u^k only. Garrard and Kornhauser (9) model the propulsion system as a first-order lag, which enables them to include jerk without having to resort to third-order equations.

In a similar manner substituting (8), (9) and (10) into (6) we obtain

$$
\begin{aligned}
\frac{d}{dt} x^k(t) &= x_t^k(t) \\
\frac{d}{dt} x_t^k(t) &= x_{tt}^k(t) \qquad (12) \\
m^k \frac{d}{dt} x_{tt}^k(t) &= -a_o x_{tt}^k(t) + f_t^k(t) \ .
\end{aligned}
$$

Equations similar to (11) can also be written for the (k+1)-th vehicle. The state equations for the two-vehicle unit can then be written as

$$
\frac{d}{dt} x(t) = A\,x(t) + B\,u(t) \qquad (13)
$$

where

$$
x(t) = \begin{vmatrix} x^k(t) \\ x_t^k(t) \\ x^{k+1}(t) \\ x_t^{k+1}(t) \end{vmatrix} \qquad u(t) = \begin{vmatrix} u^k(t) \\ u^{k+1}(t) \end{vmatrix} \qquad A = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & -a_o/m^k & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -a_o/m^{k+1} \end{vmatrix}
$$

$$
B = \begin{vmatrix} 0 & 0 \\ 1/m^k & 0 \\ 0 & 0 \\ 0 & 1/m^{k+1} \end{vmatrix}
$$

Similarly, the state equations for a two-vehicle unit using equation (12) as a basis can be written as

$$\frac{d}{dt}\,x(t) = A\,x(t) + B\,u(t) \tag{13}$$

where

$$x(t) = \begin{vmatrix} x^{k}(t) \\ x_{t}^{k}(t) \\ x_{tt}^{k}(t) \\ x^{k+1}(t) \\ x_{t}^{k+1}(t) \\ x_{tt}^{k+1}(t) \end{vmatrix} \qquad A = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -a_{o}/m^{k} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -a_{o}/m^{k+1} \end{vmatrix}$$

$$u(t) = \begin{vmatrix} f_{t}^{k}(t) \\ f_{t}^{k+1}(t) \end{vmatrix} \qquad B = \begin{vmatrix} 0 & 0 \\ 0 & 0 \\ 1/m^{k} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1/m^{k+1} \end{vmatrix}$$

In general, the state equations for an n-vehicle unit are quite similar, the form should be clear.

CONTROL OF A STRING OF VEHICLES

The objectives of an automatic control system, as outlined above, can now be put in the form of a quadratic cost functional of the form (for two vehicles and equations (12))

$$(14)\qquad J = \tfrac{1}{2}\int_0^\infty \Big(C_1(x^k(t) - x^{k+1}(t))^2 + C_2(x_t^k(t) - x_t^{k+1}(t))^2 + C_3(x^k(t))^2 + C_4(x^{k+1}(t))^2 + C_5(x_t^k(t))^2 + C_6(x_t^{k+1}(t))^2 + C_7(x_{tt}^k(t))^2 + C_8(x_{tt}^{k+1}(t))^2 + C_9(f_t^k(t))^2 + C_{10}(f_t^{k+1}(t))^2 \Big)\, dt$$

where $C_1,\ldots,C_{10}$ are non-negative weighting coefficients.

The cost functional can be interpreted as follows:

1. Since the term

$$(x^k(t) - x^{k+1}(t)) = (z^k(t) - \underline{z}^k(t)) - (z^{k+1}(t) - \underline{z}^{k+1}(t))$$

$$= (z^k(t) - z^{k+1}(t)) - (\underline{z}^k(t) - \underline{z}^{k+1}(t))$$

$$= h^k(t) - \underline{h}^k(t)$$

where $\underline{z}$ is the prescribed location and h is the actual and $\underline{h}$ the prescribed headway, the inclusion of the first term in (14) penalizes the system for deviations of the actual headway from the desired headway.

2. The term

$$(x_t^k(t) - x_t^{k+1}(t)) = (z_t^k(t) - v_o) - (z_t^{k+1}(t) - v_o)$$

$$= z_t^k(t) - z_t^{k+1}(t)$$

which is the relative velocity between the two vehicles. Thus the inclusion of the second term in (14) penalizes the system for deviations from zero of the relative velocity.

3. The next four terms in (14) serve to penalize the system for deviations of the position and velocity from their desired values.

4. The next two terms in (14) are the square of the accelerations. Thus the inclusion of these two terms serves to penalize the system for large accelerations and decelerations which could result in passenger discomfort.

5. the last two terms in (14) are proportional to the square of the jerk (when the system is moving at uniform velocity). Thus the inclusion of these terms also serves to penalize the system for the inclusion of effects which could result in passenger discomfort.

Of course, the simpler cost functional (of an analogous nature) could be written for two vehicles using equations (11) and the analogous more complex functionals written for n vehicles using either equations (11) or(12).

The philosophy underlying the use of a quadratic cost functional stems not only from the fact that a solution can easily be obtained, but also because large deviations are penalized much more severely than small ones.

The cost functional (14) can be written as

$$J = \tfrac{1}{2} \int_0^\infty (x^T(t)\, Q\, x(t) + u^T(t)\, R\, u(t))\, dt \tag{15}$$

where T denotes matrix transpose and

$$Q = \begin{vmatrix} C_1 + C_3 & 0 & 0 & -C_1 & 0 & 0 \\ 0 & C_2 + C_5 & 0 & 0 & -C_2 & 0 \\ 0 & 0 & C_7 & 0 & 0 & 0 \\ -C_1 & 0 & 0 & C_1 + C_6 & 0 & 0 \\ 0 & -C_2 & 0 & 0 & C_2 + C_8 & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{10} \end{vmatrix}$$

$$R = \begin{vmatrix} C_9 & 0 \\ 0 & C_{10} \end{vmatrix}$$

Q will be chosen to be positive semi-definite and R positive definite matrices. The objective is to find the optimal control $u^*(t)$ which minimizes (15) subject to (13). Here the initial value x(0) is assumed to be known (it will be the assumed initial error).

THE SOLUTION OF THE PROBLEM

In this section, the applicable theory is summarized for the convenience of the reader.

Available results, Athans and Falb (4).

Given a completely controllable system

$$\frac{d}{dt} x(t) = A\,x(t) + B\,u(t) \tag{16}$$

a cost functional

$$J = \tfrac{1}{2} \int_0^\infty (x^T(t)\, Q\, x(t) + u^T(t)\, R\, u(t))\, dt \tag{17}$$

where R is a positive definite and Q is a positive semi-definite matrix, then the control which minimizes J for any set of initial conditions x(0) is given by

$$u(t) = -R^{-1} B^T K\, x(t) \tag{18}$$

where K is the real symmetric positive definite constant matrix which satisfies the nonlinear matrix algebraic equation

$$-K A - A^T K + K B R^{-1} B^T K - Q = 0 . \tag{19}$$

Since (13) is of the form of (16) and since the cost functional (14) is of the form of (17) it follows that the results above can be used immediately to construct (via (18) and (19)) the optimal corrections $f_t^k(t)$ as functions of the deviations x(t).

Equation (18) is often used in conjunction with an analog computer.

It should be noted that once a particular physical system is specified, the coefficients of the matrix multiplying x(t) in (18) are fixed.

There are a number of ways to obtain the coefficients k_{ij} of the matrix K in (19). The coefficients can be obtained by solving a matrix differential equation (the Riccati equation), see Athans and Falb (4),

by matrix iteration, Kleinman (12), by so-called direct methods using the eigenvectors, by iteration on the individual terms of K, Meyer and Payne (18), and by using the special structure of the transportation problem, Melzer and Kuo (17).

Local stability is guaranteed when the control law is obtained as described above. Asymptotic stability is usually determined by numerical experiment. Analytical results are available, see Peppard (19).

FURTHER ANALYSIS OF THE COST FUNCTIONAL (14)

The cost functional (14) is quite general and can be used to obtain optimal controls for quite varied conditions. For example, Peppard (19), by weighting the coefficients C_7 and C_8 so that $C_7/C_8 = 1000$ assures the resulting two-car system behaves essentially in a car-following manner, i.e., in response to a given error x(t) only the following car receives commands to the propulsion system. Similarly, if the coefficients for both cars are the same then a slot-following strategy would be obtained, i.e., both cars would respond to errors x(t). In fact, if a strict slot-following strategy is to be used then the coefficients C_1 and C_2 in (14) could be set to zero.

Of course, the strategy to be employed will be known in advance in a given system; equation (14) was written as it is to be general and save space.

The problem is to obtain an appropriate weighting of the coefficients in (14) so that the resulting system behaves in a satisfactory manner. Most of the writers approach this problem in an ad hoc manner, trying various combinations of weightings until the system behaves satisfactorily. Wilkie and van Schieveen (23) approach this sensitivity in a systematic manner. From the published results, it seems the ad hoc method is not at all unsatisfactory, particularly in view of the fact that the available results can be used as a guide.

For the ultimate purposes here, i.e., to obtain engineering specifications on the control system, the following (at least) design inputs seem necessary:

1. Vehicle mass
2. Operating velocity
3. Operating headway (2 and 3 are quite critically interrelared)
4. Wind gust velocity
5. Maximum grades
6. Operating strategy
7. Maximum permissible variations in headway and velocity to be permitted under normal operating conditions, 20%? 50%? (Headway is more critical than velocity.)
8. Maximum acceleration ($\frac{1}{4}$g has been suggested in Anderson (1))
9. Maximum jerk ($\frac{1}{4}$g/sec has been suggested in Anderson (1))
10. Kinematic measurements to be made on each vehicle.

10. (continued)
    Position?
    Position and velocity?
    Position, velocity and acceleration?
11. Drag force function.
12. Number of vehicles and which are controlled. (Second vehicle in a two-car group, middle vehicle in a three-car group?)

With the above inputs (and possibly others!), the coefficients in the cost functional can be adjusted to obtain a system which will perform satisfactorily. At this stage a critical review will have to be made of how long it takes the control to drive errors out. If this time is felt to be excessive (in view of overall system performance), then a review of the assumptions in 7 would be in order (those in 8 and 9 are practically sacred once firm numbers have been generated).

Once a satisfactory control system has been found, then the variations described in OBJECTIVES can be implemented.

CONCLUSIONS AND RECOMMENDATIONS

Optimal control theory has been judged to be a satisfactory means of analyzing the dynamics of strings of unconnected vehicles. It is recommended that computer programs be written to further develop PRT technology. An area which merits investigation as the most rational means of assesing the control requirements is the use of stochastic methods.

REFERENCES

1. J.E. Anderson, J.L. Dais, W.L. Garrard, A.L. Kornhauser: Personal rapid transit: a collection of papers on a new type of urban transportation. Institute of Technology, University of Minnesota, April, 1972.

2. J.H. Anderson, E.T. Powner: Optimal digital computer control of cascaded vehicles in high-speed transportation systems in the presence of measurement noise and stochastic input disturbances. Transportation Research, 4, 185-198, 1970.

3. M. Athans: Applications of optimal control to high-speed ground transportation problems. Sixth Allerton Conference on Circuit and System Theory, Monticello, Ill., Oct. 1968.

4. ---------, P.L. Falb: Optimal control: an introduction to the theory and its applications, New York, McGraw-Hill, 1966.

5. ---------, W.S. Levine, A.H. Levis: On the optimal and suboptimal position and velocity control of a string of high-speed moving trains. ESL-R-291, Electronic Systems Laboratory, M.I.T., Cambridge, Nov. 1966.

6. R.K. Boyd, M.P. Lukas: How to run an automated transportation system. IEEE Trans. Sys. Man Cyber., SMC-2, 3, July 1972.

7. R.E. Fenton, J.G. Bender: A study of automatic car following. IEEE Trans. Veh. Tech., VT-18, 134-140, Nov. 1969.

8. W.L. Garrard, G.R. Hand, R. Raemer: Suboptimal feedback control of a string of vehicles moving in a single guideway. To appear in Transportation Research.

9. ----------, A.L. Kornhauser: Optimal control of automated transit vehicles. (Unpublished)

10. L.P. Hajdu, K.W. Gardiner, H. Tamura, G.L. Pressman: Design and control considerations for automated transportation systems. IEEE Proc., 56, 4, 493-513, April 1968.

11. R.J. Hansen: Planning for high-speed ground transportation. IEEE Proc., 56, 4, 472-486, April 1968.

12. D.L. Kleinman: On an iterative technique for Riccati equation computations. IEEE Trans. Automat. Contr. (Corresp.), AC-13, 114-115, Feb. 1968.

13. W.S. Levine: On the linear optimal control of a string of moving vehicles. M. Sc. Thesis, M.I.T., May 1965.

14. W.S. Levine, M. Athans: On the optimal error regulation of a string of moving vehicles. IEEE Trans. Automat. Contr., AC-11, 355-361, July 1966.

15. A.H. Levis, M. Athans: On the optimal sampled data control of strings of vehicles. Transportation Science, 2, 4, Nov. 1968.

16. S.M. Melzer: Linear regulator theories related to the optimal control of a string of moving vehicles. Ph.D. Thesis, University of Illinois, 1970.

17. ----------, B.C. Kuo: A closed-form solution for the optimal error regulation of a string of moving vehicles. IEEE Trans. Automat. Contr. 50-52, Feb. 1971.

18. G.G.L. Meyer, H.J. Payne: An iterative method of solution of the algebraic Riccati equation. IEEE Trans. Automat. Contr., 550-551, Aug. 1972.

19. L.E. Peppard: Automatic and manual control of personal-vehicle systems. Ph.D. Thesis, University of Alberta, Canada, 1971.

20. ----------, V. Gourishankar: Optimal control of a string of moving vehicles. IEEE Trans. Automat. Contr. AC-15, 386-387, June 1970.

21. E.T. Powner, J.H. Anderson, G.H. Qualtrough: Optimal digital computer control of cascaded vehicles in high-speed transportation systems. Transportation Research, 3, 101-113, 1969.

22. D.F. Wilkie: A moving cell control scheme for automated transportation systems. Transportation Science, 4,4, 347-364, Nov. 1970.

23. ----------, H.M. van Schieveen: On the sensitivity of the linear state regulator. To appear in International J. of Control.

Page Intentionally Left Blank

1972

ASEE - NASA SUMMER FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER

(AUBURN UNIVERSITY - UNIVERSITY OF ALABAMA)

THE VALUATION OF SCIENTIFIC AND TECHNICAL EXPERIMENTS

Prepared By:	Fred E. Williams
Academic Rank:	Associate Professor
University:	Georgia Institute of Technology
Laboratory: (Division)	Central Systems Engineering Space Station Division
Research Counterpart:	Stanley A. Johns
Date:	August 11, 1972
Contract No. :	NGT-01-003-045

Page Intentionally Left Blank

ACKNOWLEDGMENT

The author would like to express his appreciation to the ASEE/NASA program administrators, J. Fred O'Brien (Director) and Donald C. Raney (Associate Director), for a very stimulating and enjoyable summer program.

Thanks are also due to Roy Currie and Stan Johns of Central Systems Engineering for initially suggesting this research topic, and to Mrs. Mildred Wilkerson who patiently and competently typed two versions of this report.

THE VALUATION OF SCIENTIFIC AND TECHNICAL EXPERIMENTS

By

Fred E. Williams

ABSTRACT

This study broadly concerns rational selection of scientific and technical experiments for space missions. Particular emphasis is placed on the assessment of "value" or "worth" of an experiment.

It is argued that the assessment of such a "value" necessarily entails judgmental (subjective) inputs; hence "assessment of value" is taken to mean the construction of a numerical index which (1) represents the preferences of an appropriate decision maker (possibly a group) and (2) possesses certain desirable properties. The requirements are satisfied by a Bernoulli Utility Indicator (BUI) representing the decision maker's preferences, hence in operational terms "assessment of value" means specification of such a BUI.

A specification procedure is outlined and discussed for the case of one decision maker. Experiments are viewed as multi-attributed entities and a relevant set of attributes is proposed. Alternative methods of describing levels of the attributes are proposed and discussed. The reasonableness of certain simplifying assumptions such as preferential and utility independence is explored, and it is tentatively concluded that preferential independence almost certainly applies and utility independence appears to be quite appropriate.

The general specification procedure is then extended to include the possibility of more than one decision maker. The feasibility and potential difficulties of this extension are explored, and the suggested procedure is judged to be quite practicable.

I. INTRODUCTION

The selection of scientific and technical experiments for inclusion in space missions is a problem NASA has faced since its inception. Although historically important, at least two factors suggest that this problem will attain even greater importance in the future. First, as the space program moves into the Shuttle era, projected increases in flight activity mean more experiments; hence increased demands on experiment definition, development, and selection activities. Second, as the glamour of space flight recedes and other activities vie for national attention and budget allocations, there are increasing demands for NASA to justify budget requests and to assume accountability for results. Thus considerable attention must be given to assuring the continuation of high quality scientific and technical investigations in the space program.

Most observers agree that any experiment selection procedure should be based primarily on the concept of the "scientific and/or technical value" of an experiment, i.e., the scientific and/or technical benefits to be derived from the experiment. The obvious difficulty is that this value is seemingly clear as an abstract concept but tends to become quite nebulous when subjected to further scrutiny.

The primary objective of this study is to demonstrate one manner in which the concept of the value of an experiment can be made more precise, and to suggest procedures for assessing a numerical measurement of the value of an experiment. The orientation is primarily (and intentionally) pragmatic -- almost none of the theoretical underpinnings is new, although a few recently

developed results are discussed. The emphasis is on demonstrating, in rather concrete detail, how existing theory can be applied to this problem. A secondary objective is to suggest how these numerical measurements might be used in the formulation and solution of standard mathematical models which could be used to model various phases of the selection process.

2. SOME PHASES OF THE EXPERIMENT SELECTION PROCESS

Actually, without further specification "selection of experiments" is an ambiguous phrase. Any experiment passes through several stages of maturity as it moves from the initial gleam in a principal investigator's eye to the final analysis and dissemination of the results. Thus "selection of experiments" really entails selection at several stages of maturity. Although not necessarily definitive, the flow chart in Figure 2.1 gives some indication of this maturation under the current selection process and under a two-phase process recommended by an experiments management working group chaired by Mr. Lee Scherer [24].

At least three key decision points, labelled A, B, and C, can be identified in either process. Point A represents an initial screening where a set of proposals is reviewed and a subset is selected for further study (usually more precise definition of experimental objectives and procedures, design of required hardware, and integration studies). Decision situations at this point are typically characterized by a high degree of uncertainty, particularly regarding feasibility, costs, and design and development time. The value of a proposed experiment would be somewhat more certain, but still rather fuzzy because of the preliminary nature of proposals at this stage.

EXPERIMENT SELECTION/IMPLEMENTATION PROCESS FOR MSF PROGRAMS

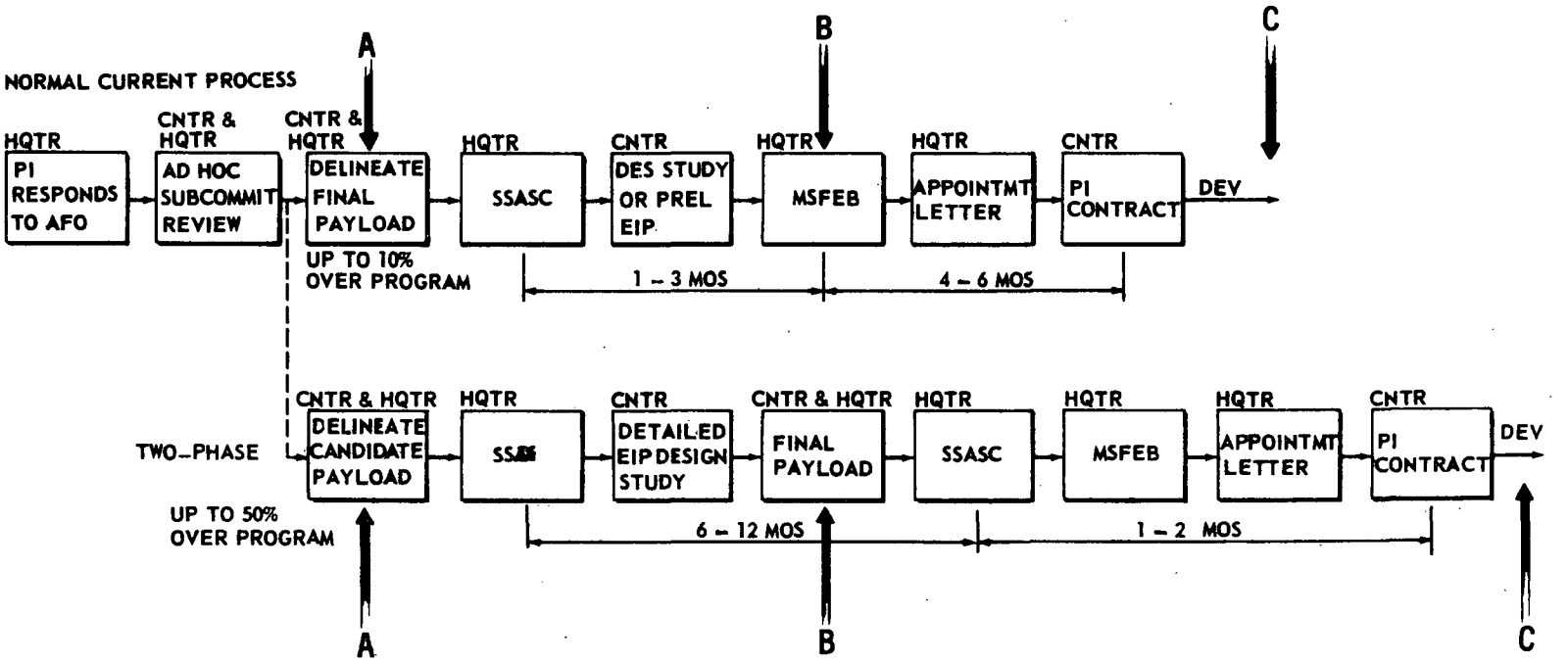


Figure 2.1

Point B represents a second screening where a set of proposals is reviewed and a subset selected for development and implementation. Decision situations at this point are typically characterized by uncertainty regarding development costs and time and, to a lesser extent, feasibility. The value of a proposed experiment should be more clearly defined than at A as a result of the intervening studies.

Point C represents a final selection where a set of experiments is chosen for assignment to a particular mission. Decision situations at this point are typically characterized by a relatively low amount of uncertainty. Primary considerations would appear to be technical, e.g., performance capabilities of the vehicle (capacities, power supply, etc.) and the crew. There is no reason to assume that the value of a candidate experiment would be more clearly defined here than at point B.

The preceeding characterization is intended to be neither a precise description of existing procedures nor a prescription for future procedures. It is, however, intended to convey the essential features of the experiment selection process and to illustrate the major types of decision situations arising in that process.

3. A RECURRING THEME

Although perhaps obvious, it should be noted that the aforementioned situations suggest decision problems having the same abstract structure. In each case a set of candidate experiments is given and a subset is to be chosen that (a) is feasible with respect to available resources and other constraints and (b) is at least as desirable (in some sense) as any other feasible subset.

"Feasibility with respect to available resources and other constraints" is probably characterized in similar manners at A and B, but in a vastly different way at C. At A, for instance, the relevant resources might be funds, manpower, and facilities available for definition and design studies. Other constraints might include such factors as design and development time limitations, or other special factors unique to the specific instance at hand. An illustration of these other constraints is provided by the recent "add-on" experiments for the Skylab. The Announcements of Flight Opportunity (Memorandum Changes 37 and 38 to NHB 8030.1A) were issued late in the program when "time till launch" was relatively short and a rather well established system was designed and being developed. Thus natural constraints were imposed on (a) acceptable design and development times for proposed experiments and (b) acceptable changes to the existing system design required by the proposed experiments.

The relevant resources at B would probably be funds, manpower, and facilities available for development, as well as time constraints and other special factors. On the other hand, at C the constraints should be primarily technical in nature, imposed by the performance capabilities of the vehicle and crew. Power supply, heat transfer capabilities, volume, payload mass capabilities (up and down), spacecraft environment, crew skills, crew time availability, etc., are examples of the factors imposing constraints in decision situations arising at C.

To assert that a particular feasible subset of experiments is "at least as desirable" as any other feasible subset requires that the relative

desirability of two experiments can be assessed and furthermore, that the relative desirability of two subsets of experiments can be assessed. Suppose for the moment that this can be done and that in addition, we can specify a real valued function, v, such that if E_1 and E_2 are any sets of experiments under consideration

(3.1) $v(E_1) \geq v(E_2)$ if and only if E_1 is at least as desirable as E_2

Then any of our decision situations can be abstractly modelled as follows:

(3.2) (a) EXPS is the set of experiments under consideration

(b) P(EXPS) is the set of all subsets of EXPS

(c) $F \subseteq P(EXPS)$ is the set of all feasible sets of experiments (feasibility determined by the factors discussed above)

(d) Find $E^* \in F$ such that

$$v(E^*) = \underset{E \in F}{\text{maximum}}\ v(E)$$

Verbally, 3.2(d) says that the decision problem can be modelled as a mathematical problem involving selecting a feasible subset E* of such that v(E*) is at least as large as v(E) for any other feasible subset E; hence by 3.1, E* is at least as desirable as any feasible subset E.

The recurring theme, then, is that the decision situations arising at points A, B, and C suggest decision problems having the same general structure which can be represented (abstractly, at least) by the same mathematical problem. The detailed form of 3.2(d) might differ, of course, for type A, B, and C decision situations. At least two investigators,

Bartee [2] and Brown [4], have suggested deterministic models for type B and C situations. Bartee suggests subjective programming as a solution technique and Brown suggests integer linear programming with binary variables. Perhaps nondeterministic models are more appropriate for type A situations in recognition of the inherent uncertainty; this remains to be investigated.

At any rate a key to this type of approach is specification of the function v. (Bartee [2] recognizes this problem, but (prematurely, in my opinion) declares the "impossibility" of specifying v and proceeds with his analysis of a type 3.2(d) problem.) Given v, for any experiment e we could properly refer to $v(\{e\})$ as the value of the experiment e. (More precisely, $v(\{e\})$ is a value of e since $v(\{e\})$ depends on the scale chosen -- much like the problem encountered in temperatures where "the" freezing temperature of water is 32 degrees and 0 degrees, depending on whether Farenheit or Centigrade scales are chosen.) Assuming a fixed scale so that we can refer to the value, note that the value of an experiment is simply a real number which is useful for decision purposes, i.e., in the solution of problems such as 3.2(d). It has no mystical qualities -- only those which we, as analysts, bestow upon it. Most importantly, it does not purport to measure the "true, underlying worth" of the experiment (whatever that means). When specified, however, it must represent the preferences of the decision maker faced with the resolution of the aforementioned decision problems, and it must possess certain other properties (discussed later) if it is to be useful in the analysis of problems such as 3.2(d). It is in this sense, then, that "the value of an experiment" is interpreted in this study.

4. THE VALUE OF AN EXPERIMENT -- A PROBLEM IN MEASUREMENT THEORY

4.1 Overview of Measurement Theory

Since the specification of the function v and the investigation of its properties is a problem in the theory of measurement, this section is devoted to a brief introductory overview of that theory. More comprehensive discussions can be found in Suppes and Zinnes [25] and Krantz [19]. Broadly stated, measurement theory is the study of correspondences between systems of empirical objects and systems of mathematical objects (usually real numbers). "A system of empirical objects" means a set of empirical objects, together with certain operations and relationships inherent in the empirical situation. For example the empirical system for the measurement of mass would conceptually be based on the set of all empirical objects. Relevant operations and relations would include the operation of combining objects in the sense illustrated by placing two pieces of metal on the pan of a balance, and the relation of one object being heavier than another. Similarly, the empirical system for temperature measurement would be a subset of the set of empirical objects, but in this case there is no natural operation of combining. The relevant relation is that of one object being hotter than another.

The basic objective of measurement theory is to characterize the observable (or desired) formal properties of the empirical system and then to find a mathematical system having the same structure as the empirical system. (Technically speaking we seek a mathematical system that is isomorphic or at least homomorphic to the empirical system.) Thus in the measurement of mass real numbers are assigned to empirical objects so that the empirical

operations and relations are reflected in the mathematical system. Each empirical object, a, is associated with a positive real number, m(a), so that (1) when objects a and b are combined yielding $a \oplus b$, we have $m(a \oplus b) = m(a) + m(b)$, and (2) $m(a) > m(b)$ if and only if a is heavier than b. In temperature measurement each empirical object, a, is associated with a real number, t(a), so that (1) $t(a) > t(b)$ if and only if a is hotter than b.

Now the functions m and t are not unique, since mass and temperature can be measured on several scales, e.g., pounds and kilograms; degrees Centigrade and degrees Farenheit. It turns out, however, that if m and m^1 are two functions giving measurement values for mass, then m is a positive constant multiple of m^1 (i.e., there is a positive real number, α, so that for any empirical object, a, $m(a) = \alpha m^1(a)$.) Similarly, if t and t^1 are functions yielding temperature measurement values t is a positive affine transformation of t^1 (i.e., there is a positive real number, β, and a real number δ, so that $t(a) = \beta t^1(a) + \delta$.) These facts are summarized by saying that mass measurement leads to a ratio scale and temperature measurement leads to an interval scale.

The two basic problems of measurement theory are illustrated by the preceeding discussion. The first problem is existence -- is there at least one mathematical system having the same structure as a given empirical system? If the answer is yes, the second problem, uniqueness, arises -- is there more than one such mathematical system, and if so, how are they related? For both mass and temperature measurement an infinite number of mathematical systems exist and the relationships between any two (both for

mass or both for temperature) are summarized by the relationships between m and m^1 or t and t^1.

How is this relevant to assessing the value of an experiment? Without going into great detail, consider the simple case of one decision maker (d. m.) facing a type A, B, or C decision situation where EXPS is the set of experiments under consideration. At the very minimum we require that d. m. be able to articulate his preference or indifference -- based solely on scientific and/or technical merit -- between any two experiments in EXPS. That is, assuming that d. m. has a personal conception of scientific and/or technical merit we ask that he ignore (for the purposes of assessing value) such considerations as cost, facilities, time constraints, etc., and compare, say, experiments e_1 and e_2 solely on the basis of scientific and/or technical merit. (This in itself might be a difficult task, and we shall explore it in detail later - assume for now that it can be done.) Let's define some notation for this relationship:

(4. 1. 1) For e_1, $e_2 \in$ EXPS

$e_1 \succcurlyeq e_2$ if and only if { Based on scientific and/or technical merit, d. m. strictly prefers e_1 to e_2 or is indifferent between e_1 and e_2 }

It is sometimes convenient to use the following notation:

Indifference

$e_1 \sim e_2$ if and only if $e_1 \succcurlyeq e_2$ and $e_2 \succcurlyeq e_1$

Strict Preference

$e_1 \succ e_2$ if and only if $e_1 \succcurlyeq e_2$ and not $e_2 \succcurlyeq e_1$

If $e_1 \succcurlyeq e_2$ we shall frequently say that d. m. weakly prefers e_1 to e_2 or simply e_1 is weakly preferred to e_2. We impose certain consistency requirements on d. m., namely:

(4.1.2) (a) For $e_1, e_2 \in$ EXPS

$$e_1 \succcurlyeq e_2 \quad \text{or} \quad e_2 \succcurlyeq e_1 \quad \text{(possibly both)}$$

(b) $(e_1 \succcurlyeq e_2 \quad \text{and} \quad e_2 \succcurlyeq e_3) \Rightarrow e_1 \succcurlyeq e_3$

Thus we require that d. m. be able to compare any two experiments in the sense that he weakly prefers e_1 to e_2 or weakly prefers e_2 to e_1; or possibly both, in which case we'd say he is indifferent between e_1 and e_2. Also we require that his preferences be transitive in that if he weakly prefers e_1 to e_2 and e_2 to e_3 then he must weakly prefer e_1 to e_3.

Thus we have an empirical relationship system (EXPS, $\succcurlyeq$) and 4.1.2 stipulates the formal observable properties we require of this relational system. The measurement theory problem is to find a mathematical system with the same structure as (EXPS, $\succcurlyeq$) and if EXPS has a finite number of elements (and is not empty) this is a trivial task. We need only specify a finite set of real numbers v(e) with the property.

(4.1.3) $v(e_1) \geq v(e_2)$ if and only if $e_1 \succcurlyeq e_2$

Then $(\{v(e) | e \in \text{EXPS}\}, \geq)$ is one satisfactory mathematical system. There are, of course, infinitely many others. We need only transform the numbers

v(e) so that their order is preserved; i.e., if T is a real valued function such that

$$(4.1.4) \quad T(v(e_1)) \geq T(v(e_2)) \quad \text{if and only if} \quad v(e_1) \geq v(e_2)$$

then $(\{T(v(e)) | e \in EXPS\}, \geq)$ is a satisfactory mathematical system. This type of measurement is called ordinal, since preservation of order is the only essential feature of the mathematical system.

Although this type of measurement will almost certainly not be sufficient for most problems characterized by 3.2(d), it does illustrate the measurement theoretic features of the experiment valuation problem. To obtain values useful in analyzing 3.2(d) type problems, we shall find it necessary to add certain features and impose further requirements to obtain the relevant empirical relational systems.

4.2 The Necessity of Judgemental Inputs

Mass and temperature measurement both possess desirable characteristics that might be conveniently summarized by calling these measurements objective. They are objective in the sense that two competent, independent observers could examine one empirical object and -- within instrumentation limits -- arrive at the same mass and temperature for the object (or equivalent values if the observers use different scales). Such objectivity is very desirable, and it would be pleasant if a similar situation were to exist in the experiment valuation problem. Unfortunately it appears that such requirements are too stringent; and in this section I shall try to

convince the reader that it is generally impossible to attain this type of objectivity in experiment valuation.

A principal reason for the objectivity of mass and temperature measurements is that in each case there is precisely one empirical relational system. To illustrate, if a and b are two empirical objects then precisely one of the following three statements is true: (1) a is heavier than b, (2) b is heavier than a, or (3) neither (1) nor (2) is true. Furthermore there are well accepted empirical procedures for determining which of (1), (2), or (3) is true (e.g., placing a and b on opposite pans of a balance and observing the height of the pans). Similarly, at a particular time instant either (1) a is hotter than b, (2) b is hotter than a, or (3) neither (1) nor (2) is true (assuming, of course, that it is meaningful to talk about a temperature of a or b, i.e., that all points of a are "equally hot" and similarly for b). Like the case of mass, there are widely accepted empirical procedures for determining which of these last three statements is true.

Unfortunately this feature of a unique empirical relational system is generally absent in the experiment valuation problem. At the risk of belaboring the obvious this can be clearly illustrated by considering a very small set of experiments, EXPS = $\{e_1, e_2\}$, where e_1 is an astronomy experiment designed to study x-ray emissions of solar flares, and e_2 is a medical experiment designed to evaluate the quantity and quality of sleep during prolonged space flight. As illustrated in section 4.1 possibly the simplest empirical relational system of interest would be (EXPS, $\succcurlyeq$) where, as in 4.1, $\succcurlyeq$ symbolizes the relationship of weak preference. The difficulty,

of course, lies in whose preferences should be used. Using a gross over-simplification to illustrate the point, we would expect an astronomer (A) to strictly prefer e_1 to e_2 ($e_1 \succ_A e_2$). On the other hand we'd expect a physician (M) to exhibit diametrically opposed preferences ($e_2 \succ_M e_1$). Thus two distinct empirical relational systems, (EXPS, $\succcurlyeq_A$) and (EXPS, $\succcurlyeq_M$), vie as candidates for our measurement problem. A little reflection reveals the impossibility of simultaneously representing these two systems with one mathematical system similar to that illustrated in 4.1; hence objectivity in the above sense is impossible -- provided that we're going to consider both the astronomer's and the physician's preferences.

One might object that the example is inappropriate since e_1 and e_2 are such vastly different experiments. There are two responses to this objection. First, comparisons of this type must (perhaps implicitly) be made in practice -- these two experiments are approved for the Skylab, so at some point, implicitly or explicitly, their values were compared. Second, considering only similar experiments, e.g., those relevant to one discipline, might alleviate the difficulty somewhat, but does not eliminate the basic problem. It would be surprising indeed if two astronomers were to exhibit identical preference orderings with respect to a set of, say, ten astronomy experiments; and two different orderings yield precisely the basic difficulties discussed above.

Conceptually one might try to avoid these problems in the following manner. Rather than asking, say, an astronomer and a physician to state their preferences directly, why not try to describe each experiment along a few

attributes and choose these attributes in such a way that one preference ordering would obviously be "best" and accepted by all reasonable observers? (This seems to be the basic approach taken in the MDAC Worth Analysis Study [13].) As an extreme example suppose we could express the total scientific, technical, economic, and social benefits of an experiment in one monetary value. Then these monetary values impose a natural ordering on any set of experiments, resolving our problem. The obvious difficulty lies in obtaining the monetary values. To illustrate, consider how these values might be obtained in a simple example. Suppose there are only two relevant attributes: (1) the scientific significance of the experimental results and (2) the estimated economic benefits resulting from the experiment, say, over a 20 year period. Moreover, assume that two experiments, e_i and e_j, are under consideration and that (a) e_i ranks higher on scientific significance than e_j and (b) the estimated economic benefits are \$1M for e_j and \$.5M for e_i. First note that (a) and (b) are necessarily estimates; thus different observers will hold different degrees of belief in the accuracy of these estimates. But suppose we dispense with that difficulty and assume that the estimates could be made precisely enough to satisfy all observers. To obtain monetary values for e_i and e_j, then, natural starting points are \$.5M and \$1M respectively. But specifying total monetary values for e_i and e_j necessarily requires assessment of tradeoffs between attributes (1) and (2) and observers will typically disagree on the appropriate tradeoffs; hence we're back to the original difficulty.

No matter which way we turn the possibility of obtaining objectivity (in the sense illustrated above) seems nonexistent. It is apparent, then, that the value of an experiment must be largely determined by judgmental (or subjective) inputs, and the task becomes one of specifying how judgmental inputs can be consistently and effectively utilized. Whose judgement should be used is not an easily resolved problem, but this question will be deferred until after we discuss methods for utilizing the judgmental inputs from one decision maker.

4.3 Desirable Properties for a Value of an Experiment

Before turning to the task of specifying experiment values for the case of one d.m., it will be convenient to have a more detailed description of the desirable properties expected of these values.

It appears that three basic properties are required, and we shall discuss these properties rather informally without going into great detail on the empirical and mathematical relational systems. As before the basic empirical relational system is (EXPS, $\succcurlyeq$) where EXPS is a relevant set of experiments and $\succcurlyeq$ denotes weak preference. We shall denote the value of an experiment $e \in$ EXPS by u(e). The first property is order preservation, already discussed in sections 3 and 4.1, but recorded here for completeness.

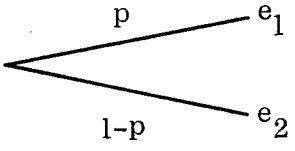
(4.3.1) Property One

For $e_1, e_2 \in$ EXPS

$u(e_1) \geq u(e_2)$ if and only if $e_1 \succcurlyeq e_2$

This property is necessary for problems involving choice under certainty. As noted in section 3 nondeterministic models might be of interest, so we need to extend this property to cover choices with random outcomes, and for our purposes this can be done by introducing the concept of a lottery. Suppose that e_1, $e_2 \in$ EXPS and for some reason either e_1 or e_2 will successfully be performed, but not both. Furthermore suppose the probability that e_1 will successfully be performed is known to be p; hence the probability of successfully performing e_2 is 1-p. (It is not easy to visualize an actual realization of this situation, but that is unimportant for our immediate purpose.) Then this situation can be described as a lottery on the set EXPS and graphically represented as follows:

(4.3.2) ℓ :



This figure is simply a graphical way of depicting a simple random process that with probability p results in the successful performance of e_1 and with probability 1-p leads to the successful performance of e_2. (It is interesting to note that, technically, a lottery on the set EXPS is simply a probability distribution defined on EXPS.)

To compare choices among lotteries we require that u have the following property:

(4.3.3) Property Two

The value associated with ℓ is

$$u(\ell) = pu(e_1) + (1-p)u(e_2)$$

Further in comparing two lotteries ℓ and ℓ_1,

$$u(\ell) \geq u(\ell_1) \quad \text{if and only if} \quad \ell \succcurlyeq \ell_1$$

Note that property two implies property one, but not vice versa. We won't pause here to discuss property two in detail since this is done in section 5.1.

Finally, the last property we require is, roughly speaking, that the value of any two experiments, e_1 and e_2, taken together is the sum of the values of the individual experiments. This property, which is desirable if problems of the form 3.2(d) are to be considered, can be expressed as follows:

(4.3.4) Property Three

If e_1, $e_2 \in$ EXPS, the value associated with $\{e_1, e_2\}$ is

$$u(\{e_1, e_2\}) = u(e_1) + u(e_2)$$

Actually 4.3.4 is a simplified version of the property required in some specific versions of type 3.2(d) problems, but we defer detailed discussion of these matters until section 9.

5. THE GENERAL APPROACH FOR ONE DECISION MAKER

5.1 Bernoulli Utility Indicators and Their Specification

In this section we assume there is one d.m. and sketch the general approach for obtaining a function u satisfying properties one and two.

Such functions have been widely studied and are referenced under at least three synonomous names: Bernoulli Utility Indicators (BUI's), measurable utility indicators, and von Neumann-Morgenstern Utility Indicators. The technical properties of BUI's are well documented and I shall only summarize them here (for more details see, e.g., von Neumann and Morgenstern [27], Blackwell and Girshick [3], Luce and Raiffa [20], Fishburn [7], and Hadley [14]).

For this discussion let's free ourselves of any specific context and consider a d.m. faced with choosing among lotteries on a set X, which we call the set of outcomes or consequences. Outcomes can be simple or very complex, for purposes of this discussion it's unnecessary to specify them in more detail. (If the reader wants to think of a concrete example, the simple case of monetary payoffs will suffice. In this case an element of X is simply a monetary amount received by d.m.) For ease of exposition assume that X has n elements, where n is a known positive integer. Thus

$$(5.1.1) \quad X = \{x_1, x_2, \ldots, x_n\}$$

Weak preference regarding outcomes is symbolized as follows:

(5.1.2) For x_i, $x_j \in X$

$x_i G x_j$ if and only if d.m. weakly prefers x_i to x_j

(G is shorthand for at least as good as)

(X, G) is an empirical relational system relevant for choices under certainty, but for comparing lotteries we need more. Since a lottery is technically a probability distribution on X, define PR(X) as follows:

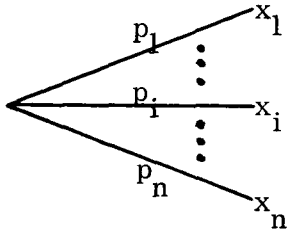
(5.1.3) $PR(X) = \{\underline{p} \mid \underline{p}$ is a probability distribution on $X\}$

We can think of elements of PR(X) in at least two equivalent ways. First, as a vector of nonnegative numbers that sum to one, i.e.

(5.1.4) $\underline{p} = (p_1, p_2, \ldots, p_n)$ where $p_i \geq 0$ and $\sum_{i=1}^{n} p = 1$

and p_i is interpreted as the probability that outcome x_i will materialize. Second, we can use the graphical depiction introducted in section 4.3.

(5.1.5)



Where $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$

Weak preference between elements of PR(X) is denoted as follows:

(5.1.6) For $\underline{p}, \underline{q} \in PR(X)$

$\underline{p} \succcurlyeq \underline{q}$ if and only if d.m. weakly prefers $\underline{p}$ to $\underline{q}$

(PR(X), $\succcurlyeq$), then, is an empirical relational system relevant to choices among lotteries, and this is the appropriate system for our measurement problem. Technically a BUI, U, is a real valued function defined on PR(X) with the properties

(5.1.7) For $\underline{p}, \underline{q} \in PR(X)$ and $0 \leq \lambda \leq 1$

(a) $U(\underline{p}) \geq U(\underline{q})$ if and only if $\underline{p} \succcurlyeq \underline{q}$

(b) $U(\lambda\underline{p}+(1-\lambda)\underline{q}) = \lambda U(\underline{p}) + (1-\lambda)\ U(\underline{q})$

Properties (a) and (b) are simply technically versions of properties one and two. Note that $\underline{p}+(1-\lambda)\underline{q}$ is just another probability distribution on X -- a weighted mixture of $\underline{p}$ and $\underline{q}$ with weights λ and $1-\lambda$, respectively. To illustrate such a mixture assume that d. m. can play one of two roulette wheels, the first described by $\underline{p}$ and the second described by $\underline{q}$. Next suppose that d. m. can no longer choose which wheel to play, but that the wheel will be chosen by a random drawing that yields the first wheel with probability λ and the second with probability $(1-\lambda)$. A ticket to the drawing can be described as $\lambda\underline{p}+(1-\lambda)\underline{q}$.

Now at first glance working with $(PR(X), \succcurlyeq)$ seems hopelessly complicated since PR(X) has an uncountably infinite number of elements. Fortunately the situation isn't that bleak; in fact it turns out that at most n-2 elements of PR(X) need be considered. For each $x \in X$ there is a corresponding degenerate distribution, $\underline{q}_x \in PR(X)$, that gives probability one to outcome x and probability zero to all other outcomes, for example:

(5.1.8) The degenerate distribution for $x_i \in X$:

In vector form: $\underline{q}_{x_i} = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$

i th component

In graphic form: $\underline{\quad 1 \quad}\ x_i$

Now for practical purposes $\underline{q}_{x_i}$ and x_i are identical, so it's natural to require that (X, G) and (PR(X), $\succcurlyeq$) agree in the following sense:

(5.1.9) For $x_i, x_j \in X$

$x_i G x_j$ if and only if $\underline{q}_{x_i} \succcurlyeq \underline{q}_{x_j}$

Further, no generality is lost if the outcomes are numbered from most to least preferred:

(5.1.10) For $x_i, x_j \in X$

$x_i G x_j$ if and only if $i \leq j$

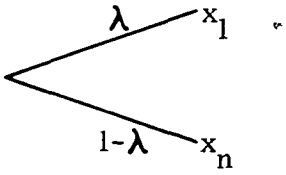
Under 5.1.9 and 5.1.10 specification of U can be completed by considering only the n-2 degenerate distributions $\underline{q}_{x_i}$, i = 2, 3, ..., n-1. The procedure is essentially as follows:

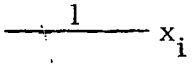
(5.1.11) Procedure for specifying U

(a) Set $U(\underline{q}_{x_1}) = 1$, $U(\underline{q}_{x_n}) = 0$

(b) For i = 2, 3, ..., n-1 face d.m.

with two alternatives (b.1) and (b.2)

(b.1) 

(b.2) 

(c) Determine the unique value, λ_i, for which d. m. is indifferent between (b. 1) and (b. 2) and set

$$U(\underline{q}_{x_i}) = \lambda_i$$

That the procedure in 5.1.11 effectively specifies U follows from two easily demonstrated facts. First 5.1.7(b) can be extended to:

(5.1.12) If $\underline{p}^1, \underline{p}^2, \ldots, \underline{p}^m \in PR(X)$ and $\lambda_1, \lambda_2, \ldots, \lambda_m$ satisfy $\lambda_i \geq 0$ and $\sum_{i=1}^{m} \lambda_i = 1$,

$$U\left(\sum_{i=1}^{m} \lambda_i \underline{p}^i\right) = \sum_{i=1}^{m} \lambda_i U(\underline{p}^i)$$

Second any distribution $\underline{p} \in PR(X)$ can be written as a weighted mixture of the degenerate distributions using 5.1.12:

(5.1.13) If $\underline{p} = (p_1, p_2, \ldots, p_n) \in PR(X)$ then

$$\underline{p} = \sum_{i=1}^{n} p_i \underline{q}_{x_i}$$

Since the expression for $\underline{p}$ in 5.1.13 satisfies the provisions of (5.1.12) it follows that

(5.1.14) For $\underline{p} = (p_1, p_2, \ldots, p_n) \in PR(X)$

$$U(\underline{p}) = U\left(\sum_{i=1}^{n} p_i \underline{q}_{x_i}\right) = \sum_{i=1}^{n} p_i U(\underline{q}_{x_i})$$

and the procedure in 5.1.11 completely specifies U.

U is specified, then, by ranking the outcomes, arbitrarily assigning values to the most and least preferred, and determining the remaining values by comparing alternatives such as (b.1) and (b.2) in 5.1.11.

Three issues remain. First, what is the relationship of U to the function u discussed in section 4.3? Second, under what conditions does U exist? Third, what are the uniqueness properties of U? The relationship between u and U is shown in 5.1.15.

(5.1.15) Given (X, G), (PR(X), $\succcurlyeq$) satisfying (5.1.9) and U a BUI on PR(X). Define u as follows:

$$u(x) = U(q_x) \quad \text{for } x \in X.$$

Then u has properties one and two of section 4.3.

Alternatively, of course, if we had started with u satisfying properties one and two of section 4.3, we could have reversed the above procedure and defined U in terms of u. Thus the distinctions between u and U are essentially technical, and in fact are ignored in most sources -- except for axiomatic studies. Following this practice we call both u and U BUI's.

Existence and uniqueness properties depend, of course, on the precise assumptions imposed on the system (PR(X), $\succcurlyeq$). Several authors have given axiom systems for (PR(X), $\succcurlyeq$) which are both necessary and sufficient for the existence of a BUI. (See, e.g., von Neumann and Morgenstern [27], Blackwell and Girshick [3], Herstein and Milnor [15], Chernoff and Moses [5], DeGroot [6], and Luce and Raiffa [20].) The exact statement of these axioms depends on the system being studied (particularly

regarding the number of elements in X), but for the situation discussed here the axioms can be loosely characterized as follows:

(5.1.16) Characterization of requirements on $(PR(X), \succcurlyeq)$ for the existence of U

(a) For $\underline{p}, \underline{q} \in PR(X)$: $\underline{p} \succcurlyeq \underline{q}$ or $\underline{q} \succcurlyeq \underline{p}$ (or both)

(b) $(\underline{p} \succcurlyeq \underline{q}$ and $\underline{q} \succcurlyeq \underline{r}) \Rightarrow \underline{p} \succcurlyeq \underline{r}$

(c) If $\underline{p} \succ \underline{q} \succ \underline{r}$, there exists a unique λ, $0 \leq \lambda \leq 1$ such that

$$\underline{q} \sim \lambda \underline{p} + (1-\lambda)\, \underline{r}$$

(recall $\underline{s} \sim \underline{t}$, means $\underline{s} \succcurlyeq \underline{t}$ and $\underline{t} \succcurlyeq \underline{s}$)

In most studies (a) and (b) are direct requirements, while (c) is usually a theorem derived from two or more less restrictive assumptions.

If we choose to work with u these requirements can be summarized as follows:

(5.1.17) Characterization of requirements for the existence of u

(a) For x, $y \in X$, xGy or yGx (or both)

(b) (xGy and yGz) $\Rightarrow$ xGz

(c) If x is strictly preferred to y and y is strictly preferred to z, then there is precisely one value λ, $0 \leq \lambda \leq 1$ such that

(c.1) λ to x, $1-\lambda$ to z $\sim$ (c.2) 1 to y

(i.e., such that d.m. is indifferent between c.1 and c.2)

(d) Equivalent lotteries may be substituted for outcomes.

We've seen that there are two distinct, but related, preference orderings, G and $\succcurlyeq$. G denotes weak preference regarding elements of X (outcomes) while $\succcurlyeq$ denotes weak preference regarding elements of PR(X) (lotteries over outcomes). In most sources this distinction is not retained and $\succcurlyeq$ is typically used to denote weak preference regarding elements of X as well as PR(X). In view of 5.1.9 this is reasonable since $x \succcurlyeq y$ can be interpreted as $\underline{q}_x \succcurlyeq \underline{q}_y$, hence xGy. We shall adopt this convention and henceforth use only $\succcurlyeq$ to denote weak preference.

In all cases BUI's are unique up to positive affine transformations. Like temperatures, BUI values are measurements on interval scales. Given a particular situation u (equivalently U) is specified with respect to an arbitrarily chosen unit and origin. Thus some care must be exercised in using and interpreting these values.

Given this background a general procedure for specifying a function u with properties one and two should be clear. Assume the set, EXPS, of relevant experiments has n elements indexed so that the lower the index, the more preferred the experiment. That is,

(5.1.18) (a) $\text{EXPS} = \{e_1, e_2, \ldots, e_n\}$

(b) For e_i, $e_j \in \text{EXPS}$

$e_i \succcurlyeq e_j$ if and only if $i \leq j$

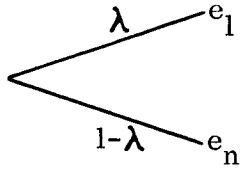
(As before, $e_i \succcurlyeq e_j$ means d. m. weakly prefers e_i to e_j)

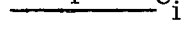
Assuming that d. m. 's preferences satisfy sufficient requirements for the existence of a BUI, u can be specified as follows:

(5.1.19) Procedure for specifying u

(a) Set $u(e_1) = 1$ and $u(e_n) = 0$

(b) For i = 2, 3, ..., n-1 face d. m. with two alternatives (b.1) and (b.2)

(b.1) 

(b.2) 

(c) Determine the unique value, λ_i, for which d. m. is indifferent between (b.1) and (b.2) and set

$$u(e_i) = \lambda_i$$

The procedure in 5.1.19 solves our problem in principle, but as a practical matter it might be infeasible as stated. The source of this infeasibility is step (c), specification of the λ_i for a particular e_i. Experiments are complicated entities with many attributes, and step (c) requires simultaneous consideration of three experiments (e_1, e_n, and e_i) plus the probability λ_i. It might simply be impossible for d. m. to consider and weigh all the relevant factors at one time. As a practical matter we need a more piecemeal attack on this process and probably the best such approach is one that will effectively reduce the complexity of comparing two experiments.

5.2 Multiattributed Consequences

One basic approach to reducing the complexity of comparing two experiments is to consider an experiment as a multiattributed entity. Thus

an experiment is viewed as being described along several relevant attributes (dimensions, criteria) such as the scientific significance of the results, the relevant discipline(s), the technical significance of the results, the timeliness of the experiment, etc., each attribute contributing to the experiment's aggregate worth. Starting with such a description, d. m. can systematically evaluate tradeoffs between the attributes, and assessments such as 5.1.19(c) can be performed in a series of small steps, none of which imposes unreasonable requirements on d. m.

Before turning to the details of such a procedure I shall quickly review and summarize the basic ideas in a more general setting. For more systematic and detailed discussions of these topics see Raiffa [23], Keeney [17], or Ting [26]. As a concrete example for this discussion assume that d. m. is a student preparing to select his first full time job. Further suppose that after considerable reflection on the matter he has concluded that the following six attributes are sufficient for decision purposes:

(5.2.1) Attributes

(1) Annual Salary

(2) Fringe Benefits

(3) Geographic Location

(4) Professional Mobility

(5) Opportunity for Advancement

(6) Type of Work

Under 5.2.1 a position could be described as a six component vector, e.g.

(5.2.2) Description of a position

$$\underline{x} = (x_1, x_2, x_3, x_4, x_5, x_6)$$

Where x_1 = annual salary for the position

x_2 = fringe benefits for the position

x_3 = geographic location of the position

Etc.

and we refer to x_i as the level of the i^{th} attribute. A more specific example of a position's description is:

(5.2.3) Specific example of a description

$\underline{x}$ = ($9,000, Benefit Package A, Atlanta, High, Medium, Adm.)

Interpretation: A job described by $\underline{x}$:

(a) has an annual salary of $9,000

(b) has fringe benefit package A

(c) is located in Atlanta, Georgia

(d) has high professional mobility

(e) has medium opportunity for advancement (within the organization)

(f) is administrative in nature

In this example the level of attribute 1 is a real number, while the levels of the other attributes are described verbally. If we denote the set of all possible levels of the i^{th} attribute by X_i, then every possible description of a position is an element of the cartesian product of the X_i's. If X is defined as follows,

(5.2.4) For i = 1, 2, . . . , 6 let

X_i = set of all possible levels of attribute i

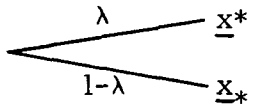
Define X as:

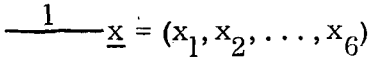
$$X = \prod_{i=1}^{6} X_i = \left\{(x_1, x_2, x_3, x_4, x_5, x_6) \middle| x_i \in X_i;\ i=1, 2, \ldots, 6\right\}$$

then X contains all possible descriptions of positions. X might also contain many elements that don't correspond to actual positions, but that's irrelevant.

Now for decision purposes X is the set of outcomes or consequences we discussed in section 5.1, except, of course that in 5.2.4 each element of X is described along many attributes. (Unfortunately there is a slight possibility of confusion due to notation -- in 5.1 x_i was a typical outcome, an element of X, while in 5.2.4 x_i is the level of the i^{th} attribute of a typical outcome $\underline{x} \in X$.) For obvious reasons situations with outcome sets X as in 5.2.4 are called decision situations with multiattributed outcomes (consequences) or more succinctly, multiple criteria decision situations.

Now suppose d. m. 's preferences admit the existence of a BUI, u, on X, and for simplicity assure that $\underline{x}^*$ and $\underline{x}_*$ are, respectively, most preferred and least preferred elements of X. (The existence of such elements is not a restrictive assumption in practice.) As before, for $\underline{x} \in X$, $u(\underline{x}) = u(x_1, x_2, \ldots, x_6)$ is specified to be that value of λ such that d. m. is indifferent between the alternatives in 5.2.5.

(5.2.5) (a) 

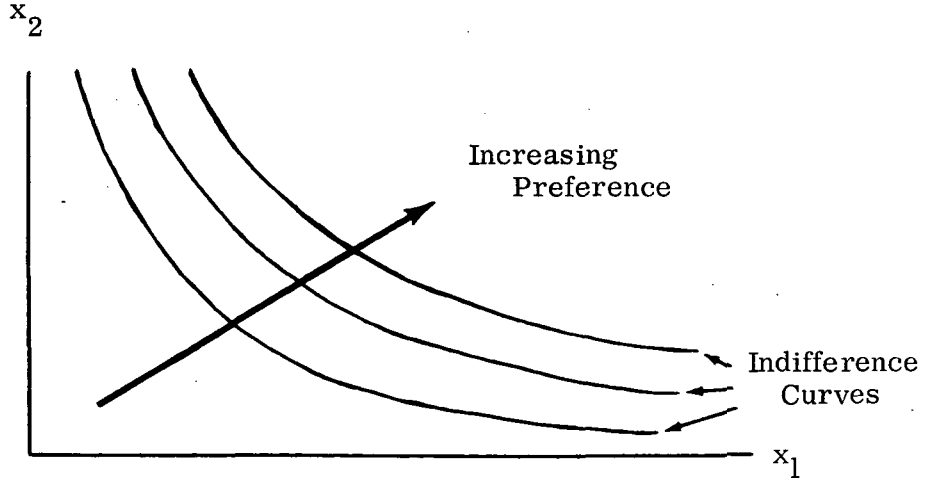
(b) 

But the goal was to avoid this type of direct comparison, and so far we've only succeeded proliferating notation. Two basic approaches are available. First, tradeoffs between the attributes can be assessed, thus effectively reducing the dimensionality of an outcome (i.e., the number of attributes d.m. must simultaneously consider). Second, if it is reasonable to assume a specific functional form for u, this information can frequently be used to aid in specification of u.

To illustrate the use of tradeoffs consider a simple case with two attributes.

$$(5.2.6) \quad X = X_1 \times X_2$$

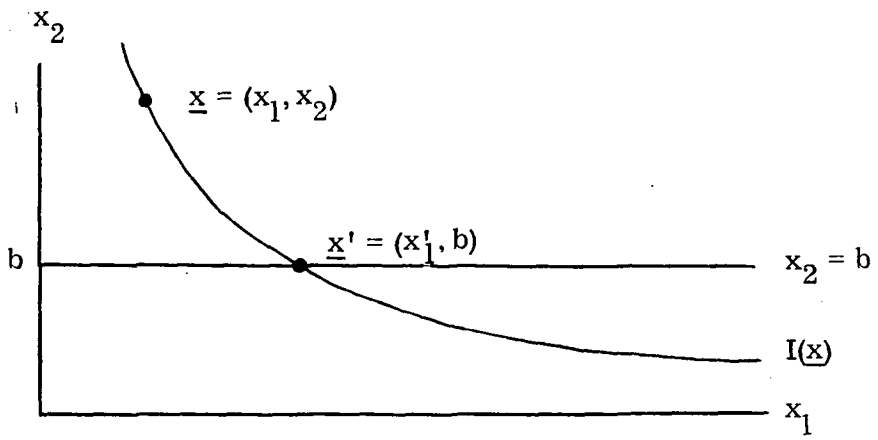
For convenience suppose that the levels of the attributes are expressed as non-negative real numbers, higher numbers reflecting higher preference. Then the situation can be graphically depicted as in Figure 5.1. The curves are indifference curves -- d.m. is indifferent between any pair of points lying on the same indifference curve.



Graphic Depiction of X for Two Attributes

Figure 5.1

Now points $\underline{x} = (x_1, x_2)$ in X can be effectively reduced to one dimension by choosing a convenient level for, say x_2. Take $x_2 = b$, and consider any point $\underline{x} = (x_1, x_2)$. As depicted in figure 5.2 $\underline{x}$ lies on an indifference curve, $I(\underline{x})$, which intersects the line defined by $x_2 = b$, at $x' = (x_1', b)$.

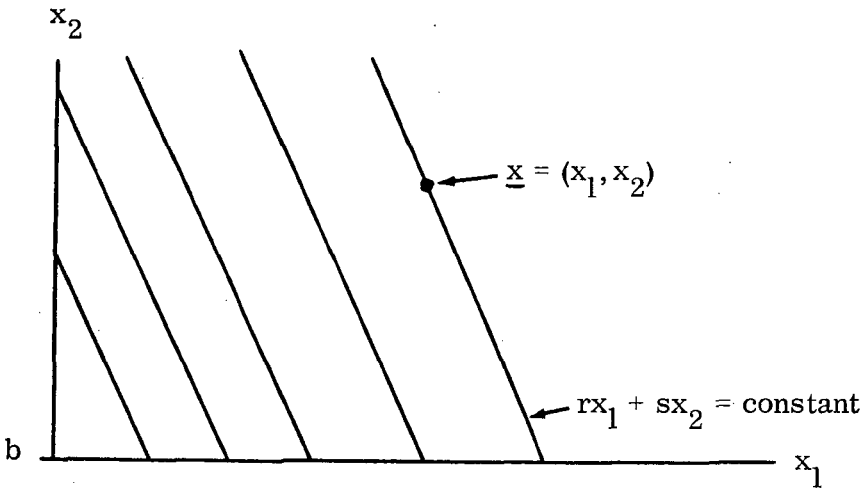


Reduction of $\underline{x}$ to an Equivalent Point (x_1', b)

Figure 5.2

Since $\underline{x}$ and $\underline{x}'$ both lie on $I(\underline{x})$ they are indifferent, hence for decision purposes $\underline{x}$ can be replaced by $\underline{x}'$. Thus any point $\underline{x} = (x_1, x_2)$ can be reduced to an equivalent point $\underline{x}' = (x_1', b)$ and since all points $\underline{x}'$ have identical second coordinates, comparison of such points requires comparison on the first attributes only.

It would indeed be fortunate if the curves $I(\underline{x})$ could be expressed in closed analytical form, since this aids the determination of x_1'. For example, linear $I(\underline{x})$ curves are shown in Figure 5.3.



Linear Indifference Curves

Figure 5.3

Choosing b=0 a point $\underline{x} = (x_1, x_2)$ can be reduced to an equivalent point $\underline{x}' = (x_1 + sx_2/r, 0)$ where $(-s/r)$ can be interpreted as the rate of substitution of x_1 for x_2 (the tradeoff rate of x_1 for x_2). Obviously one doesn't expect life to be this simple, and in most cases no such formula will exist. The basic approach, though, is independent of such neat, closed-form expressions.

It should be obvious that the decision to use x_1 as the numeraire was arbitrary. It would have been just as easy to illustrate the ideas by choosing a constant value of x_1, say $x_1 = a$, and then reducing points to the form $\underline{x} = (a, x_2')$. In applications the numeraire choice can usually be made on the basis of convenience, ease of interpretation, and other such considerations.

Having successfully reduced the outcome descriptions to one dimension, say x_1, it suffices to specify a BUI, u, on X_1. This is done via

comparisons of alternatives such as those in 5.2.5, but the important feature is that with only one dimension, comparisons of this type are practical.

Handling situations with more than two attributes involves successive applications of these procedures. Thus in the six attribute job example, $\underline{x} = (x_1, \ldots, x_6)$, and d. m. might first effectively eliminate x_6 by reducing $\underline{x}$ to an equivalent $\underline{x}' = (x_1, x_2, x_3, x_4, x_5', f)$ then reducing $\underline{x}'$ to an equivalent $\underline{x}'' = (x_1, x_2, x_3, x_4', e, f)$, and continuing in this fashion, finally arrive at the point (x_1', b, c, d, e, f). Then u could be specified on X_1. For more detailed discussions of these ideas the reader should consult Raiffa [23] or Ting [26].

To illustrate the use of a specific functional form consider again a case with two attributes. Under some conditions it turns out that u assumes a rather specific functional form, one of which is the so called additive form.

(5.2.7) Let $X = X_1 \times X_2$ and u be a BUI on X.
Then u is an additive BUI if and only if
$u(x_1, x_2) = w_1(x_1) + w_2(x_2)$
Where w_1 and w_2 are BUI's on X_1 and X_2, respectively

Furthermore, if there are most preferred, $\underline{x}^*$, and least preferred, $\underline{x}_*$, elements of X, u can be written as follows:

(5.2.8) $u(x_1, x_2) = k_1 u_1(x_1) + k_2 u_2(x_2)$
Where u_1 and u_2 are BUI's on X_1 and X_2,
and $k_1 \geq 0$, $k_2 \geq 0$, $k_1 + k_2 = 1$

meaning that u is the weighted sum of two BUI's. In this case specification of u requires only the specification of u_1, u_2, and either k_1 or k_2.

Suppose additivity were appropriate in our job example. Then u would be of the form:

$$(5.2.9) \quad u(x_1, \ldots, x_6) = \sum_{i=1}^{6} k_i u_i(x_i)$$

$$\text{Where } k_i \geq 0 \text{ and } \sum_{i=1}^{6} k_i = 1$$

In practical terms this means that d. m. need only specify six BUI's and the weights k_i. u_1 is a BUI on X_1, and to specify u_1 d. m. need only consider one attribute -- salary. Similarly to specify u_2 d. m. need only consider fringe benefits, etc. Thus the additivity of u reduces the almost impossible specification task in 5.2.5 to a sequence of manageable tasks.

It might be too optimistic to expect additivity of all the attributes. Thus d. m. 's preferences might make the following form appropriate:

$$(5.2.10) \quad u(x_1, \ldots, x_6) = k_1 u_1(x_1) + k_2 u_2(x_2) + k_3 u_3(x_3, x_4, x_5, x_6)$$

$$\text{Where } k_i \geq 0 \text{ and } k_1 + k_2 + k_3 = 1$$

Specification of u requires d. m. to specify u_1, u_2, and u_3 and the weights k_i. Tradeoff procedures could be used to specify u_3, so this example illustrates a case in which both approaches can be applied.

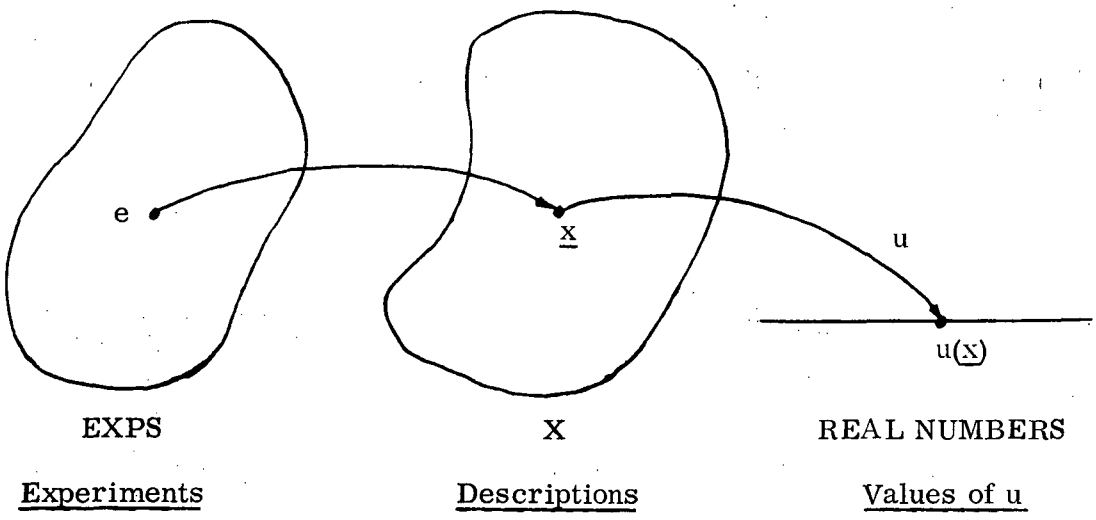
In addition to additivity two other basic forms of u are quasi-additive and multiplicative. Each of these three forms is appropriate under

particular preference conditions that we shall not discuss in detail now. Suffice it to say that in the past few years several authors have studied these matters and derived conditions on preferences that lead to particular forms of u. (See e.g., Fishburn [8] [9] [10] [11] [12], Keeney [16] [18], Raiffa [23], and Luce and Tukey [21].) In each case a specific functional form for u leads to an assessment procedure requiring much simpler comparisons than those in 5.2.5.

We're now in a position to give a more detailed description of the specification of u. Given the set, EXPS, of relevant experiments, d.m. must first choose the set of relevant attributes and the manner in which levels of each attribute are to be expressed (e.g., as real numbers, verbal expressions, etc.). Suppose d.m. chooses m attributes, and the set of possible levels of the i^{th} attribute is X_i. Then define X as follows:

$$(5.2.11) \quad X = \prod_{i=1}^{m} X_i$$

and X contains all possible descriptions of experiments in EXPS. The next step is to describe each experiment, $e \in$ EXPS, in terms of the relevant attributes, i.e., to associate e with precisely one $\underline{x} \in X$. This being done, the possibility of simplifications such as special functional forms of u should be explored. Finally, all possible simplifications having been made, u is specified. Given u and any $\underline{x} \in X$, we can find $u(\underline{x})$. Thus since any experiment $e \in$ EXPS is described by a vector $\underline{x} \in X$, we naturally say that $u(e) = u(\underline{x})$. The process might be graphically depicted as in figure 5.4.



Graphical Depiction of Procedure for Specifying u

Figure 5.4

Aside from details two issues remain to be discussed. First we've assumed the existence of EXPS, a set of relevant experiments and u is specified with respect to EXPS. In practice, of course EXPS changes over time and we've not explicitly accounted for that fact. This is easily done in the following manner. Let $EXPS_t$ be the set of experiments at time t, where for convenience time is indexed in discrete units (months, years, etc.). Now suppose u is specified for $EXPS_t$ and we now wish to specify u for $EXPS_{t+1}$. If d. m. 's preferences regarding the elements of $EXPS_t$ haven't changed (they clearly could change over time, a problem discussed in section 8. 3), then all that needs be done is to specify u for the "new" experiments, i. e., for experiments e such that $e \in EXPS_{t+1}$ and $e \notin EXPS_t$. Some care must be taken to insure that the resulting values are comparable (on the same scale) with the values u(e) for $e \in EXPS_t$, but this is a simple matter. It requires no assumptions or procedures different from those required to specify u on $EXPS_t$.

Of course if d. m. 's preferences for elements of $EXPS_t$ change, then u must be specified on the set $EXPS_t \cup EXPS_{t+1}$. In either case there are two interesting options regarding the actual numerical values, u(e). The first option is to retain a fixed range of possible values, say the interval 0 to 1000. Using this option it would be possible for $e \in EXPS_t$ and $e \in EXPS_{t+1}$ and the actual numerical value u(e) could change from t to t+1, even if d. m. 's preferences regarding elements of $EXPS_t$ do not change from t to t+1. If this happens, however, all numerical values u(e), for $e \in EXPS_t$, are transformed in precisely the same manner. The second option avoids this feature by permitting u(e) to be any real number, negative or positive, although in practice negative numbers could probably be avoided if this seemed desirable. Of course under the second option the actual numerical values u(e) depend on the unit and origin chosen for specifying u in the initial set of experiments at time t=0. This should not be bothersome, however; since BUI's are measurements on interval scales we can hope for no more than the three properties discussed in section 4. 3.

The second issue concerns explicit recognition that an experiment can fail. Thus far the discussion has deftly avoided mentioning the possibility of failure. Describing an experiment, e, as a vector, $\underline{x}$, of several attribute levels implicitly assumes complete success of e. This assumption is unrealistic for obvious reasons -- apparatus does fail occasionally and unforeseen difficulties do arise -- and it seems imperative that the value of an experiment should account for such possibilities. Fortunately, the ability to handle such situations is already built into our

procedures, as illustrated by the following example. Suppose the relevant set of experiments is EXPS, each experiment is described as a vector $\underline{x} = (x_1, x_2, x_3, x_4) \in X$, and u is a BUI on X. Now suppose $e_1 \in$ EXPS has three possible "levels" of success, each with a known (estimated) probability:

(5.2.12)	Level of Success	Notation	Description in X	Probability
	Complete Success	e_1	$\underline{x}=(x_1, x_2, x_3, x_4)$	p
	Partial Success	e_1'	$\underline{x}'=(x_1', x_2', x_3', x_4')$	p'
	Complete Failure	e_1''	$\underline{x}''=(x_1'', x_2'', x_3'', x_4'')$	p''

Then since $\underline{x}$, $\underline{x}'$, $\underline{x}'' \in X$ and u is a BUI on X it follows that the value of e_1 is:

$$(5.2.13) \quad u(e_1) = pu(\underline{x}) + p'u(\underline{x}') + p''u(\underline{x}'')$$

In other words the assumption of complete success means that the description $\underline{x}$ is known with certainty. Relaxing that assumption to include three possible levels of success, each with a known probability, means that one of the descriptions $\underline{x}$, $\underline{x}'$, or $\underline{x}''$ will materialize. Loosely speaking, explicitly including the possibility of failure changes our representation of e_1 from the certain description $\underline{x}$ to a lottery over descriptions, $\underline{x}$, $\underline{x}'$, and $\underline{x}''$. Since u is a BUI on X, the value of this lottery is precisely that given in 5.2.13.

Even though our procedure handles different possible levels of success in a very natural manner, it does so at some cost. It's obvious that including several levels of success for each experiment could greatly increase the complexity of assessing u. Further, each level of success must be

described by a vector $\underline{x}$, and it will become apparent in section 6 that these descriptions don't fall as freely as rain drops. For more practical considerations, though, a very few levels of success should suffice. Indeed, in many cases two levels -- total success and total failure -- should suffice.

This completes our discussion of the general procedure for specifying u in the case of one d. m. In sections 6 and 7 we turn to some of the details of this procedure, suggesting some specific attributes and their measurement and then exploring the reasonableness of some possible simplifying assumptions. In section 8 we relax the assumption of one d. m. and discuss a suggested assessment procedure in this more realistic setting. In section 9 we discuss the conditions under which u has property 3 of section 4. 3.

6. RELEVANT ATTRIBUTES AND THEIR MEASUREMENT

In this section we turn to a detailed discussion of some relevant attributes and the manner in which the levels of those attributes might be expressed. At this stage it's impossible to provide any definitive statements since in many areas there are various options concerning details and good options would need to be determined in practice. We do, however, attempt to provide a rather concrete illustration of some principal aspects of the procedure discussed in section 5. 2.

For illustrative purposes we have chosen to consider six attributes as relevant for experiment valuation, and for reasons discussed below we shall explicitly consider only five of them. These attributes are not suggested as

definitive and there is obvious possibility of considering a different set of attributes. On the other hand the attributes suggested here were not haphazardly chosen -- indeed, they were carefully chosen for two reasons. First they have been suggested, in one form or another, in various NASA documents [28] [29], and second, they incorporate all other criteria either mentioned in those documents or in my discussions with various NASA personnel. The six attributes are as follows:

(6.1) Suggested List of Relevant Attributes

(a) Discipline(s): The relevant discipline or disciplines for the experiment. For example, one experiment might pertain to biology, another to astronomy, another to two or more disciplines, etc.

Attributes (b) and (f) should be understood under the assumption that the experiment will be completely successful as designed.

(b) Potential Scientific Significance of the Results:

The scientific importance of the projected results, per se, to the relevant disciplines. That is, are the projected results basic, fundamental knowledge in the relevant disciplines and likely to generate additional questions and areas for further investigation? Or, on the other hand, are the projected results of a rather narrow, specialized nature, likely to be of interest to only the principal investigator and a few others?

(c) Scientific Timeliness: The degree to which the investigation addresses questions and areas on a "critical path" of progress in the relevant disciplines. Questions and areas on a "critical path" in a discipline would be those that must be investigated before other questions and areas can be explored, i.e., those that logically, or for practical purposes, must precede other experiments.

(d) Potential Scientific and Technical Significance of Any Required Technological Advances:

The degree to which any required technological advances (if any are required) can be transferred and utilized in other areas of scientific and technical investigation. This attribute is intended to be completely distinct from (b) which concerns the experiment's projected results. To illustrate the present attribute, suppose that experiment e requires the design and fabrication of a sophisticated piece of new apparatus. Then regardless of the scientific results of e the new apparatus might have great potential for applications in other areas, and this potential is an important determinant of the desirability of e.

(e) Technical Timeliness: The degree to which the experiment addresses questions and areas on a "critical path" of progress in the relevant technical area(s).

(f) Potential Economic Benefits: The short to intermediate term (0-10 years) market potential (addition to GNP) that might reasonably be attributed to the experiment (both the experiment's projected results and any required technological advances).

Another attribute considered but not included is some notion of the "newness" of an experiment, i.e., the extent to which the experimental concept breaks new paths in the scientific discipline or technical area. After some reflection, however, it was decided that this aspect would be incorporated in attributes (b) and (d).

Having introduced these six attributes we immediately reduce the list -- at least for the present -- to attributes (b) and (e). If we persist in excluding (a), it will be necessary to consider different groups of experiments corresponding to the various disciplines, and to apply our procedures to each group. This leads to different BUI's for each group, and if the ultimate goal is comparable experiment values there is no escaping explicit consideration

of tradeoffs between disciplines. That is, at some point it becomes necessary to compare the relative desirabilities of a particular experiment in discipline a and another experiment in discipline b.

This kind of comparison, of course, is extremely difficult and it might be best to avoid the difficulty in the following, circuitous manner. At a particular phase of the selection process, rather than considering one problem of the form 3.2(d) for all experiments, consider one 3.2(d) type problem for experiments in each discipline group. Of course this process doesn't dispense with interdisciplinary comparisons, these comparisons are simply made at another level -- when the available resources are allocated to the various disciplines.

At any rate whether we aspire to specify experiment values that are comparable among disciplines or only within a particular discipline, attributes (b) through (e) need to be considered. Thus it seems reasonable to postpone explicit consideration of (a), and we shall do so.

Having focused on this set we must now select a method for describing the various levels of the attributes. This is not a trivial task, and the main difficulty centers on the scale properties required from the level descriptions. Loosely speaking, more structure is better than less but obtaining structure can be difficult. To illustrate consider attribute (b), the scientific significance of the results. If x_1 is a level of scientific significance for experiment e, what is the nature of x_1? At one extreme we might impose what is probably the maximum possible structure and express x_1 as a percentile to be interpreted in the following manner:

(6.2) $x_1 = 95\%$ if and only if $\left\{\begin{array}{l}\text{the scientific significance of e is at least as great as the scientific significance of 95\% of all conceivable experiments in this discipline and 5\% of these experiments have scientific significance at least as great as e}\end{array}\right\}$

At the other extreme we might express x_1 as a value on a rather unstructured scale such as:

(6.3) Possible values of x_1:

(1) Excellent

(2) Good

(3) Fair

(4) Poor

Obviously expressing x_1 in terms such as 6.2 conveys, loosely speaking, more information than expressing x_1 in terms such as 6.3. Both magnitudes and differences of percentiles have rather precise interpretations while the levels in 6.3 can, at most, be interpreted only as indicative of rankings. Of course it's also much more difficult to specify percentile levels than to specify levels such as those in 6.3. Between these two extremes various compromises can be attained by relaxing the precision required in 6.2. Thus rather than requiring a percentile level expressed in two significant digits, we might relax the precision to larger intervals such as the top ten percent, the next ten percent, etc.; or the top twenty-five percent, the next twenty-five percent, etc. In these cases the levels of x_1 are illustrated as follows:

(6.4) Decile Intervals

$x_1 = 10$ if and only if {the scientific significance of e is in the top 10% of all conceivable experiments in this discipline}

$x_1 = 9$ if and only if top 20% but not top 10%

$x_1 = 8$ if and only if top 30% but not top 20%

Etc.

Quartile Intervals

$x_1 = 4$ if and only if top 25%

$x_1 = 3$ if and only if top 50% but not top 25%

$x_1 = 2$ if and only if top 75% but not top 50%

$x_1 = 1$ if and only if not top 75%

Expressing x_1 levels in this manner preserves some of the desirable properties of the 6.2 type expression without imposing unreasonable requirements for precision.

Some possible methods of describing the levels of attributes (b) through (f) are summarized in Table 6.1.

ATTRIBUTE	LEVEL	MORE STRUCTURE ←			→ LESS STRUCTURE
Sci. Sign.	x_1	Percentile	Decile	Quartile	Arbitrary Ordinal
Sci. Time	x_2	Percentile	Decile	Quartile	Arbitrary Ordinal
Tech. Sign.	x_3	Percentile	Decile	Quartile	Arbitrary Ordinal
Tech. Time	x_4	Percentile	Decile	Quartile	Arbitrary Ordinal
Econ. Bnft.	x_5	$	Intervals of $		Arbitrary Ordinal

Possible Scales for Attribute Levels

Table 6.1

At this point it's impossible to select one of these possibilities as definitely more desirable than the others. Like the selection of a best set of

attributes, best methods of describing the levels would evolve in practice. For definiteness let's assume that x_1, x_2, x_3, and x_4 are expressed in deciles as in 6.4, and that x_5 is expressed in integer values, 0, 1, 2, ..., 20, where $x_5 = 0$ means the potential economic benefits are \$0, $x_5 = 1$ means potential economic benefits greater than \$0 and less than or equal to \$1M, $x_5 = 2$ means greater than \$1M and less than or equal to \$2M, etc.

An illustration of a description of $\underline{x}$ and its interpretation is:

(6.5) $\underline{x} = (8, 10, 2, 1, 1)$

Interpretation: An experiment e described by $\underline{x}$:

(a) is between the 70th and 80th percentile on scientific significance.

(b) is in the top 10 percent on scientific timeliness.

(c) is between the 10th and 20th percentile on technical significance.

(d) is in the lowest 10 percent on technical timeliness.

(e) has projected economic benefits greater than \$0 and less than or equal to \$1M.

Now the symbols chosen to denote attribute levels are arbitrary -- e.g., any set of 10 distinguishable symbols can be used to denote the levels of x_1, x_2, x_3, or x_4. The numerical symbols we've used are convenient since they reflect natural ordering of the levels of the attributes, but they require careful interpretation. Perhaps the most important point is that identical levels for different attributes can be interpreted in completely different ways -- e.g., $x_4 = 1$ and $x_5 = 1$ and the numeral '1' is interpreted as a decile for x_4 and an interval of dollar values for x_5.

We defer discussion of obtaining the descriptions $\underline{x}$ until section 8.1. For the present we assume that each experiment is described by a vector $\underline{x}$ and turn now to some of the details of specifying a BUI on X. Specifically, the next section concerns some potential simplifications and explores the likelihood that any of these simplifications can be used.

7. SIMPLIFICATIONS -- UTILITY AND PREFERENTIAL INDEPENDENCE

Under the assumptions of section 6 the set containing all possible descriptions of experiments is:

$$(7.1) \quad X = \prod_{i=1}^{5} X_i$$

$$\text{Where } X_i = \{1, 2, \ldots, 10\} \qquad i = 1, 2, 3, 4$$

$$X_5 = \{0, 1, 2, \ldots, 20\}$$

We turn now to some details concerning the specification of a BUI, u, on X. Now X has natural most preferred and least preferred elements, $\underline{x}^* = (10, 10, 10, 10, 20)$ and $\underline{x}_* = (1, 1, 1, 1, 0)$, respectively. But this isn't too helpful since comparing alternatives of the type illustrated in 5.2.5 is a practical impossibility. Some simplifications are necessary and we shall explore the possibility that some particular simplifications can be applied. The basic ideas and results used here are rather recent developments due to Raiffa [23], Keeney [18], and Ting [26].

To introduce the required concepts we temporarily consider an outcome set with two components $\underline{y}$ and $\underline{z}$, each of which might be vectors. Let Y and Z be the sets of possible values of $\underline{y}$ and $\underline{z}$ respectively.

$$(7.2)\quad X = Y \times Z = \left\{ (\underline{y}, \underline{z}) \,\middle|\, \underline{y} \in Y \text{ and } \underline{z} \in Z \right\}$$

To illustrate the intended interpretation of $\underline{y}$ and $\underline{z}$, we might let $\underline{y} = (x_1, x_2)$ and $\underline{z} = (x_3, x_4, x_5)$. Then if $Y = X_1 \times X_2$ and $Z = X_3 \times X_4 \times X_5$, the sets in 7.1 and 7.2 are equivalent.

Now if $\succcurlyeq$ denotes weak preference with respect to elements of X we can introduce the idea of preferential independence.

(7.3) Y is preferentially independent of Z

if and only if

$$\left[(\underline{y}, \underline{z}) \succcurlyeq (\underline{y}', \underline{z}) \text{ for some } \underline{z} \in Z \right] \Rightarrow \left[(\underline{y}, \underline{z}') \succcurlyeq (\underline{y}', \underline{z}') \text{ for all } \underline{z}' \in Z \right]$$

Thus Y is preferentially independent of Z if and only if the conditional preferences between $\underline{y}$ and $\underline{y}'$ given a fixed $\underline{z}$ do not depend on the particular $\underline{z}$ chosen. In our case with Y and Z as illustrated above this would mean that, for example,

$$(7.4)\quad (8, 7, 4, 1, 18) \succcurlyeq (5, 9, 4, 1, 18) \Rightarrow$$

$$\Rightarrow (8, 7, x_3, x_4, x_5) \succcurlyeq (5, 9, x_3, x_4, x_5)$$

for all values of x_3, x_4, x_5

Thus if the levels (8, 7) on attributes 1 and 2 were weakly preferred to the levels (5, 9) when the remaining attribute levels are 4, 1, and 18, preferential independence implies that $(8, 7, x_3, x_4, x_5)$ must be weakly preferred to $(5, 9, x_3, x_4, x_5)$ for any values of x_3, x_4, and x_5. One might conjecture that Y preferentially independent of Z implies Z preferentially independent of Y,

but this is not true in general. (For examples and discussion see Raiffa [23], pp. 26 ff. The property he calls weak conditional utility independence is what we've referred to as preferential independence.)

Another useful way of illustrating preferential independence is in terms of the indifference curves introduced in section 5.2. In our example, fixing x_3, x_4, and x_5 at 4, 1, and 18 yields a set of indifference curves in the $X_1 \times X_2$ plane. Preferential independence says these indifferences curves remain unchanged regardless of the specific values chosen for x_3, x_4, and x_5.

Before turning to the idea of utility independence we pause to introduce some notational conventions that will facilitate later discussions.

(7.5) Let $X = \prod_{i=1}^{m} X_i$ be a set of outcomes.

For $k = 1, 2, \ldots, m$

$$X_{\bar{k}} = \prod_{\substack{i=1 \\ i \neq k}}^{m} X_i$$

Thus in our example:

$$X_{\bar{1}} = X_2 \times X_3 \times X_4 \times X_5$$

$$X_{\bar{2}} = X_1 \times X_3 \times X_4 \times X_5$$

Etc.

Elements of $X_{\bar{k}}$ are denoted as $x_{\bar{k}}$. This notation can be generalized to more than one excluded subscript as illustrated for our example as follows:

(7.6) $X_{\overline{12}} = X_3 \times X_4 \times X_5$

$X_{\overline{24}} = X_1 \times X_3 \times X_5$

$X_{\overline{135}} = X_2 \times X_4$

Etc.

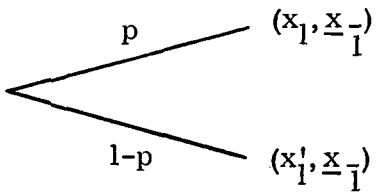
Using this notation the sets Y and Z illustrated above can be described as

(7.7) $Y = X_1 \times X_2 = X_{\overline{345}}$

$Z = X_3 \times X_4 \times X_5 = X_{\overline{12}}$

We shall need to discuss lotteries on X in which some of the components of outcomes $\underline{x}$ vary while others remain fixed. For example the lottery might be

(7.8)



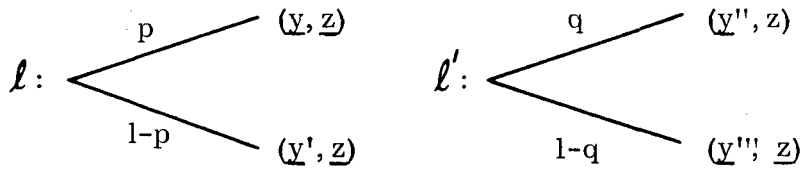
In this lottery only the values in X_1 vary -- $\underline{x}_{\bar{1}}$ is fixed. We describe such a lottery as a lottery over outcomes $(\tilde{x}_1, \underline{x}_{\bar{1}})$ where the tilde over x_1 denotes that the x_1 components vary (are random variables) in the lottery. Thus a lottery over $(x_1, \tilde{\underline{x}}_{\bar{1}})$ is one in which x_1 is fixed and the components $\underline{x}_{\bar{1}}$ vary.

With this notation we can now introduce the idea of utility independence.

(7.9) Let $X = Y \times Z$. Then Y is utility independent of Z if and only if d. m. 's preferences between lotteries over $(\tilde{y}, z)$ is independent of the particular $\underline{z}$ chosen.

To illustrate suppose d. m. is indifferent between two lotteries ℓ and ℓ':

(7.10)



Then Y utility independent of Z means that d. m. would be indifferent between the two lotteries resulting from substituting $\underline{z}'$ for every $\underline{z}$ in ℓ and ℓ'. And this must hold regardless of the values of $\underline{z}$ and $\underline{z}'$ chosen.

Like preferential independence, utility independence is not necessarily reflexive. Thus Y utility independent of Z doesn't necessarily imply Z utility independent of Y. (Again consult Raiffa [23] for discussion and examples.)

The possible relevance of all of this to our immediate problem is due to a recent result obtained by Keeney. Paraphrasing, his theorem is as follows:

(7.11) (Keeney [18]) Let $X = \prod_{i=1}^{m} X_i$, $m \geq 3$.

If for some X_i,

(a) $X_i \times X_j$ is preferentially independent of X_{ij} for all $j \neq i$

(b) X_i is utility independent of $X_{\bar{i}}$

then either

(c) $u(\underline{x}) = \sum_{r=1}^{m} k_r \, u_r(x_r)$

or

(d) $1 + ku(\underline{x}) = \prod_{r=1}^{m} \left[1 + k \, k_r(x_r)\right]$

Where u and u_r are BUI's scaled from 0 to 1, the k_r are scaling constants with $0 \leq k_r \leq 1$ and $k > -1$ is a scaling constant.

The form of u in (c) is the additive form discussed in section 5.2 and (d) is the multiplicative form mentioned there. The names and details are not so important, but the key consideration is that if the conditions of this theorem are met, the assessment of u can be segmented into small managable tasks -- specification of five functions u_r and the constants.

Of course it's impossible to state definitely that the theorem applies to our problem but at present it seems reasonable to expect that it would. To illustrate, take i = 1. Then the first antecedent condition is

(7.12) $X_1 \times X_j$ preferentially independent of $X_{\overline{1j}}$ for j = 2, 3, 4, 5

e.g. j = 3:

$$\left[(x_1, x_3, \underline{x}_{\overline{13}}) \succcurlyeq (x_1', x_3', \underline{x}_{\overline{13}})\right] \Rightarrow$$

$$\Rightarrow \left[(x_1, x_3, \underline{x}'_{\overline{13}}) \succcurlyeq (x_1', x_3', \underline{x}'_{\overline{13}})\right]$$

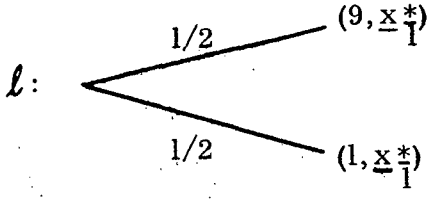
for all values of $\underline{x}'_{\overline{13}}$.

This seems to be a reasonable assumption -- that conditional preferences for paired levels of scientific significance and technical significance are independent of the levels of the remaining three attributes. Of course similar independence must hold for j = 2, 4, 5, but it seems reasonable to make those assumptions.

The second antecedent condition requires (if i=1) that X_1 is utility independent of $X_{\overline{1}}$. There are alternative methods of verifying (or refuting) this assumption and we can only illustrate some support for its plausibility. Utility independence means that preferences for lotteries over $(\tilde{x}_1, \underline{x}_{\overline{1}})$ are independent of the particular $\underline{x}_{\overline{1}}$ value. To illustrate

(7.13) Let $\underline{x}^*_{\overline{1}} = (10, 10, 10, 20)$

Consider the lottery:



Now suppose we ask d. m. to specify the value $\hat{x}_1$ so that he's indifferent between $(\hat{x}_1, \underline{x}_{\bar{1}}^*)$ and the lottery ℓ. The outcome $(\hat{x}_1, \underline{x}_{\bar{1}}^*)$ is called the certainty equivalent for ℓ. Now suppose we substitute $\underline{x}_{\bar{1}*} = (1, 1, 1, 0)$ for $\underline{x}_{\bar{1}}^*$ in ℓ, obtaining ℓ'. Let the certainty equivalent for ℓ' be $(\hat{\hat{x}}_1, \underline{x}_{\bar{1}*})$. Utility independence requires that $\hat{x}_1 = \hat{\hat{x}}_1$. Of course utility independence also requires that substitution of any $\underline{x}_{\bar{1}}$ in ℓ yields a lottery with certainty equivalent $(\hat{x}_1, \underline{x}_{\bar{1}})$. So consideration of specific cases such as ℓ can only lend supporting evidence for utility independence or refute its appropriateness.

At any rate the equality of $\hat{x}_1$ and $\hat{\hat{x}}_1$ seems reasonable and after a fair amount of reflection on the matter I've been unable to construct a realistic example refuting the assumption that X_1 is utility independent of $X_{\bar{1}}$. Obviously this is not sufficient assurance that the assumption is reasonable -- its appropriateness must be judged in practice. At this point we can only say that it should be entertained in any practical situation.

In sum we've seen that it's not unreasonable to expect the assessment of u to be greatly facilitated by some simplifying assumptions. On the other hand what if these simplifying assumptions are deemed unacceptable in practice? In particular what happens if the utility independence assumptions of 7.11 aren't acceptable? Then we're left with no alternative but the assessment of tradeoffs between the attributes. Can this always be done? Ting [26] has shown that, roughly speaking, preferential independence and the applicability of tradeoff procedures are equivalent. Since it seems safe to assume preferential independence between any X_i and X_j $(i \neq j)$, tradeoff procedures would apply. (Actually Ting's assumptions require a bit more than we've

assumed -- in particular the use of discrete levels for the attributes might introduce some minor difficulties -- but let's worry about such detailed considerations when (and if) they arise in an actual application.) Since we've only got five attributes, the assessment of tradeoffs between the attributes is not an unreasonable task.

This concludes our discussion of a general procedure for assessing u in the case of one d. m. A few detailed aspects of the procedure have been deliberately omitted. For example if the additive form were appropriate, what are the precise methods for assessing the functions u_r and the weights k_r? Also, if different degrees of success are permitted, how are the relevant probabilities assessed? These aren't unimportant issues, but they are relatively minor. Various methods exist and the selection of one or more is a matter best left until an actual problem is being analyzed.

8. SUGGESTED ASSESSMENT PROCEDURE

8.1 Procedure and Participants

The last three sections have outlined an assessment procedure for one decision maker, but it's obvious that many individuals are and should be involved in any realistic assessment process. We turn now to the general features of a suggested assessment procedure in this more realistic setting.

The main ideas of this process are identical to the case of a single d. m. Basically two types of groups are involved. One group of individuals essentially replaces d. m. in the previous process. It is this group's preferences that the BUI u represents. There is only one such group, and its members are primarily policy-level NASA administrators. (It would probably

be desirable to include some members from the scientific community.) This group's primary function is to determine the relevant attributes, the manner in which the levels are to be described, and the necessary judgmental inputs for assessing u. Of existing NASA groups the Manned Space Flight Experiment Board is most similar in purpose to the suggested group.

Groups of the second type have the essential function of describing the levels of selected attributes for actual experiments. There are many such groups comprised of scientists and technical specialists from NASA and the scientific community. For example the levels of scientific significance and scientific timeliness for astronomy experiments are determined by a group comprised mainly of astronomers. On the other hand the levels of these attributes for life sciences experiments are determined by a group comprised mainly of life scientists. The technical significance and timeliness levels are determined by a group of competent technical personnel, while the potential market value might be determined by a group of applied economists and/or statisticians. These groups are also responsible for providing probability estimates when varying levels of success are considered. Actual NASA groups most similar in purpose to the suggested type two groups are the scientific and technical rating panels which have been used in evaluating proposed experiments.

The suggested general approach, then, can be summarized as follows. Given a set of experiments type two groups are responsible for describing these experiments along the relevant attributes. Using these descriptions group one "puts it all together," i.e., they provide the necessary judgmental inputs to specify u. Also, if comparable values among disciplines are desired, group one would be responsible for supplying the required judgments.

8.2 Feasibility

This approach seems entirely feasible since it is very similar to existing, well accepted NASA procedures such as those used in evaluating the Skylab add-on experiment proposals [29] and those used by many source evaluation boards (SEB's). The details, of course, differ in that our procedure requires, perhaps, more explicit consideration of details such as comparisons and tradeoffs. On the other hand our procedure compensates by segmenting the task and requiring relatively few judgmental inputs from each group. To illustrate, let's roughly compare the suggested procedure's requirements with those imposed by the Skylab add-on evaluation process.

Two types of rating panels, scientific and technical, were used in the Skylab add-on evaluation process. Scientific panels were asked to place each proposal reviewed into one of four categories: Category I -- Recommended for Flight on Skylab A; Category II -- Recommended for Flight at First Available Opportunities After Category I Experiments; Category III -- Recommended for Further Definition and/or Development Before Consideration for Flight; and Category IV -- Not Recommended for Flight. Criteria to be considered included (1) the experiment's potential for scientific accomplishment, if successful, (2) scientific timeliness, (3) probability of positive results, (4) adequacy of data processing and analysis methods, (5) investigator's prior experience and competence, and (6) interest of the investigator's institution, especially with respect to providing continuing support.

Technical panels were asked to place each proposal in one of four similar (not identical) categories on the basis of at least six criteria: (1) the experiment's potential contribution and timeliness in support of the materials science and manufacturing in space program objectives, (2) the experiment's relevance to potential processes for manufacturing products in space for use on earth, (3) market potential or economic benefit, (4) the ability of the investigator to satisfy flight program requirements and meet flight program schedules, (5) level and validity of funding requests, and (6) requirements for support from NASA manpower and facilities.

It's rather difficult to compare the ultimate goals of these panels with specific aspects of our suggested process. These panels are probably most similar to our suggested type two groups, but note that these panels were being asked, in effect, to suggest tentative solutions to type 3.2(d) problems in one fell swoop. In each case the relevant criteria included not only those relevant to the experiment's value, but some that were relevant to probability of success, and some that were really relevant to resource limitations. Furthermore, the panels were given no explicit instructions regarding relative importance of the criteria and/or procedural details. Obviously the simultaneous consideration of such diverse factors without an overall conceptual framework is a rather demanding task.

On the other hand our procedure would require the scientific panel to describe each proposal along the relevant attributes, say attributes (b) and (c), while the technical panel would describe each proposal along, say, attributes (d), (e) and (f). Further, the scientific panel would be responsible

for describing possible levels of success and estimating the requisite probabilities. Thus while our procedure requires each type two group to make more explicit (and perhaps more detailed) judgments, it also requires each group to consider fewer items simultaneously. Furthermore, it does not require these groups to simultaneously consider probabilities, resouce availabilities, and attributes relevant to value, and to combine these factors in an ill-defined manner to arrive at a final rating.

Of course our procedure also requires group one to provide the necessary inputs for the assessment of u. Hence it requires explicit consideration of detailed comparisons and tradeoffs not required by the Skylab add-on evaluation process. But to a great extent this is a "one shot" investment. Initial specification of relevant attributes, methods of describing attribute levels, tradeoffs, etc., would certainly be time consuming and demanding. But once accomplished they need not be respecified each time experiment values are to be determined. As a specific example suppose the initial specification yielded the attributes and levels discussed in section 7. Further suppose group one's preferences led to an additive form of u over all six attributes. Then initial specification of the functions u_r and the weights k_r ($r = 1, 2, \ldots, 6$) would be demanding, but the process would hardly begin anew each time a new set of experiments was processed. Periodic reviews and adjustments would obviously be required and although changes might occur over time, it seems unlikely that drastic shifts in preference would occur. At any rate complete respecification of u would almost certainly impose far fewer demands than the initial specification.

8.3 Potential Difficulties

Our procedure is not without some potential rough spots, and undoubtedly the most serious potential problem concerns aggregation of various individuals' judgments to form acceptable group judgments. In group one and any of the type two groups differences of opinion are almost certain to exist among members. In general terms the issue is that it might be difficult to arrive at group judgments which simultaneously reflect each individual member's judgment.

To be more precise suppose we have a group comprised of n individuals, individual i having a preference ordering, $\succcurlyeq_i$, with respect to a set of outcomes, X. Then by group preferences we mean a preference ordering, $\succcurlyeq$, with respect to outcomes in X. To assert that $\succcurlyeq$ adequately represents the individuals' preferences, we must stipulate the relationships that necessarily exist between the various $\succcurlyeq_i$'s and $\succcurlyeq$, and under some very reasonable-looking relationships, it has been shown that no ordering, $\succcurlyeq$, exists (see Arrow [1]). Of course there are conditions under which an ordering, $\succcurlyeq$, exists, and a fairly extensive literature exists on this general topic (see the references in Arrow[1]). Detailed consideration of these issues is beyond the scope of the present study, but we do want to point out that there are pragmatic -- if not totally satisfactory -- methods of resolving these difficulties. As a simple example the chairman of each group could be vested with the ultimate authority to impose his personal preferences should his group arrive at an impasse. (For further discussion of these matters see Raiffa [22].)

There are three minor potential difficulties. First, the selection of group members (particularly for group one) is not a trivial task. Since the

judgments of these individuals would have great impact on the space program (at least in the experiments area), criteria for selection would need to be carefully considered. While important, this is certainly not insurmountable. Second, our procedure would certainly require the use of one or more analysts to help in the detailed, technical facets of selecting attributes, methods of describing levels, and methods of assessing u. This would be particularly important and necessary in the formative stages of the procedure. Again, this seems minor. Third, as mentioned earlier, judgments can change over time and some provisions should certainly be made for the systematic and regular review and updating of the judgmental inputs.

In sum, then, there are some potential difficulties associated with the proposed procedure, but individually or in the aggregate, it appears that they can be coped with.

9. THE QUESTION OF ADDITIVITY

We turn now to Property Three of section 4.3 Suppose there are n relevant experiments and we've specified a BUI, u. Thus the situation is

$$\text{(9.1)} \quad \text{EXPS} = \{e_1, e_2, \ldots, e_n\}$$

u is a BUI on EXPS

Consideration of type 3.2(d) problems requires the ability to compare subsets of EXPS as well as individual pairs of experiments. For example we must be able to compare relative desirabilities of, say, $\{e_1, e_3, e_n\}$ and $\{e_2, e_6, e_7\}$.

A naive approach might be to simply compare $u(e_1) + u(e_3) + u(e_n)$ and $u(e_2) + u(e_6) + u(e_7)$, then assert that the largest sum is associated with the most preferred set. This approach is quick and easy but the obvious question is: "Does it make any sense?" The basic assumption of this procedure is that the value of a set of experiments, say $\{e_1, e_3, e_n\}$, is the sum of the values of the individual experiments, e_1, e_3, and e_n. In this section we explore this assumption and conclude that it's unlikely to be generally valid in the form just stated. We shall see, however, that under some reasonable assumptions we can construct a BUI useful for comparing sets of experiments.

Surprisingly, I have been unable to find any discussions of this problem in the literature. Perhaps the most relevant paper is Fishburn's study [12] although he doesn't deal directly with our problem. We shall, however, be able to use the basic properties of BUI's along with the idea of utility independence to explore the conditions under which the above assumption is valid. We need the following notation:

(9.2) For $e_i \in$ EXPS: $u_i = u(e_i)$

For convenience assume the elements of EXPS are subscripted so that

$u_i \geq u_j$ if and only if $i \leq j$

Now our naive approach might be characterized a bit more formally as follows:

(9.3) (a) For $E \subseteq EXPS$ define W as

$$W(E) = \sum_{e \in W} u(e)$$

(b) For $E_1, E_2 \subseteq EXPS$

$$E_1 \succcurlyeq E_2 \quad \text{if and only if} \quad W(E_1) \geq W(E_2)$$

As usual $\succcurlyeq$ denotes weak preference, but note that $E_1 \succcurlyeq E_2$ denotes weak preference with respect to sets of experiments, and not simply weak preference between two individual experiments, e.g., $e_1 \succcurlyeq e_2$. There is an obvious distinction between these preference orderings, and where any possibility of confusion exists, we encircle the "$\succcurlyeq$" to denote weak preference between sets of experiments, e.g., $E_1 \circledcirc E_2$. Let's examine some results of 9.3. Suppose we first consider only one-element subsets of EXPS.

(9.4) For $e_i, e_j \in EXPS$

$$\{e_i\} \circledcirc \{e_j\} \Leftrightarrow W(\{e_i\}) \geq W(\{e_j\}) \Leftrightarrow$$

$$\Leftrightarrow u_i \geq u_j \Leftrightarrow e_i \succcurlyeq e_j$$

Thus 9.3 orders the one element subsets of EXPS in precisely the same manner $\succcurlyeq$ orders EXPS, and this is, of course, a very desirable feature. Also, since there's only a technical distinction between $\{e_i\}$ and e_i, it's fairly obvious that W will serve as a BUI for lotteries on the one-element subsets of EXPS.

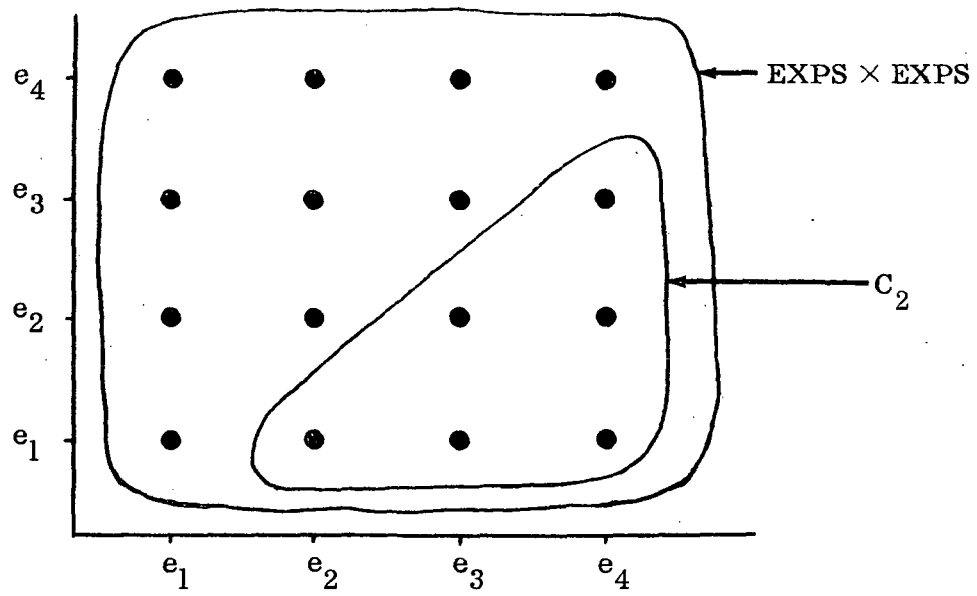
Now let's consider two-element subsets of EXPS.

(9.5) For $e_i, e_j, e_k, e_l \in$ EXPS

$$\{e_i, e_j\} \otimes \{e_k, e_l\} \Leftrightarrow u_i + u_j \geq u_k + u_l$$

Now there's no natural way to check the appropriateness of (9.5) because $\geq$ is only relevant for comparing individual experiments. We shall be able to show, however, that under some independence assumptions 9.5 does make sense. To be more precise, letting C_2 denote the class of all two element subsets, we show that W is a BUI on C_2 under some independence assumptions.

To state the assumptions we note that C_2 can be identified in a natural way with a proper subset of EXPS × EXPS. Figure 9.1 illustrates one such identification for n = 4.



Identification of C_2 with a Proper Subset of EXPS × EXPS

Figure 9.1

Of course there are many other ways to identify C_2 with a subset of EXPS$\times$ EXPS, but that's irrelevant for our purposes.

Suppose, for the moment, that we could specify a BUI, v, on EXPS$\times$ EXPS. Further suppose that u were additive, i.e., that v could be written in the form 5.2.7 or 5.2.8, i.e.

(9.6) For e_i, $e_j \in$ EXPS

$$v(e_i, e_j) = k_1 w_1(e_i) + k_2 w_2(e_j)$$

Where k_1, $k_2 \geq 0$; $k_1 + k_2 = 1$,

and w_1 and w_2 are BUI's on EXPS.

Obviously we'd use u for both of the w_i's, so the form of v would be:

(9.7) $$v(e_i, e_j) = k_1 u(e_i) + k_2 u(e_j)$$

$$k_1, k_2 \geq 0; k_1 + k_2 = 1$$

Now it's reasonable to require $v(e_i, e_j) = v(e_j, e_i)$ since the pairs (e_i, e_j) and (e_j, e_i) both correspond to the same two-element subset $\{e_i, e_j\}$. Under this assumption we have

(9.8) $$v(e_i, e_j) = k_1 u_i + k_2 u_j$$

$$v(e_j, e_i) = k_1 u_j + k_2 u_i$$

and if $v(e_i, e_j) = v(e_j, e_i)$ for all i, j, clearly $k_1 = k_2 = 1/2$. We can summarize as follows:

(9.9) If v is of the form in 9.7 and

$$v(e_i, e_j) = v(e_j, e_i) \text{ for all } i, j, \text{ then}$$

$$v(e_i, e_j) = 1/2\ u(e_i) + 1/2\ u(e_j)$$

Now note that v is closely related to W.

$$(9.10) \quad v(e_i, e_j) = 1/2\ (u_i + u_j) = 1/2\ W(\{e_i, e_j\})$$

Indeed W is a positive affine transformation of v, hence W is a BUI representing the same preference ordering as v. To summarize

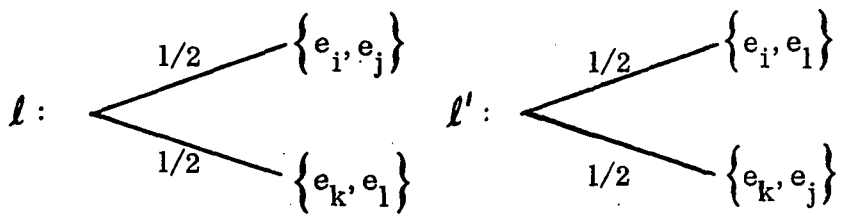
(9.11) Under the assumptions 9.6, W is a BUI on C_2.

Are the assumptions 9.6 reasonable? Let's back up a step and explore some conditions under which these assumptions hold. Keeney ([17] Theorem 3, p. 282) has derived some results that, when applied to our problem essentially say that the assumptions in 9.6 hold if EXPS is utility independent of EXPS and there exist e_i, e_j, e_k, e_l so that 9.12 is satisfied.

(9.12) (a) $\{e_i, e_j\}$ and $\{e_i, e_l\}$ are not indifferent

(b) $\{e_i, e_j\}$ and $\{e_k, e_j\}$ are not indifferent

(c) The lotteries ℓ and ℓ' are indifferent



Essentially 9.12 says that there are four experiments in EXPS for which d. m. is insensitive to the manner in which these experiments are paired in the fifty-fifty lotteries ℓ and ℓ'. This seems to be a reasonable assumption. Recall that utility independence of EXPS and EXPS means d. m. preferences for lotteries $(\tilde{e}, e_i)$ are independent of the specific e_i chosen. Again, this seems to be a very reasonable assumption.

Thus we've seen that it's quite reasonable to expect our naive procedure in 9.3 to yield a function W which is simultaneously a BUI on C_2 and C_1 (the classes of all two-element and one-element subsets of EXPS, respectively). However, we're not home free yet. What about comparisons involving one and two-element subsets, e.g., $\{e_3\}$ and $\{e_6, e_7\}$? The difficulty is that W induces a reasonable ordering of C_1 and a reasonable ordering of C_2, but there's no link yet between the two orderings. To illustrate consider the case of n = 4. Under our assumptions we can definitely specify the ordering W induces on C_1. On C_2, however, we can only illustrate since the exact ordering depends on the numerical values u_i.

(9.13) Orderings induced W for n = 4:

(a) C_1 (definite):

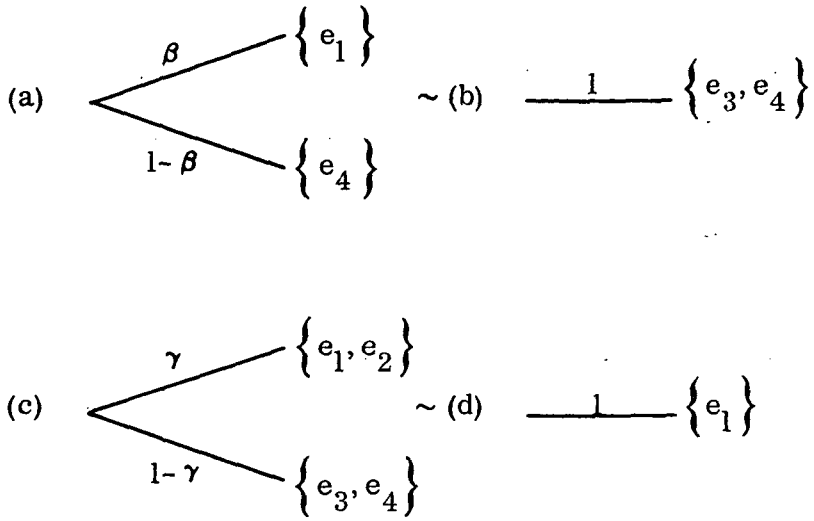
$$\{e_1\} \succcurlyeq \{e_2\} \succcurlyeq \{e_3\} \succcurlyeq \{e_4\}$$

(b) C_2 (illustration):

$$\{e_1, e_2\} \succcurlyeq \{e_1, e_3\} \succcurlyeq \{e_1, e_4\} \succcurlyeq \{e_2, e_3\} \succcurlyeq$$
$$\succcurlyeq \{e_2, e_4\} \succcurlyeq \{e_3, e_4\}$$

The problem now is to establish the linkages between the two preference orderings. This is easily done, but it requires two judgments from d. m. Suppose, for example, that d. m. weakly prefers $\{e_1\}$ to $\{e_3, e_4\}$. Then he must specify, for example, β and γ so that the following indifferences hold.

(9.14)



Using W and the empirically determined β and γ, it is an easy matter to obtain one BUI on $C_1 \cup C_2$, and we omit the details of the procedure. Suffice it to say that we can obtain a BUI, w, on $C_1 \cup C_2$ of the form shown in 9.15.

$$(9.15) \quad w(E) = \begin{cases} \sum_{e \in E} u(e) & \text{if } E \in C_1 \\ a_2\left(\sum_{e \in E} u(e)\right) + b_2 & \text{if } E \in C_2 \end{cases}$$

Where $a_2 > 0$ and b_2 is a real number, both depending on β and γ.

Now the BUI w sufficies for analysis of type 3.2(d) problems involving sets in $C_1 \cup C_2$. Further, we can use w to analyze deterministic or non-deterministic versions of these problems.

However, it's not clear that w orders the sets in $C_1 \cup C_2$ precisely in the same manner as our naive procedure 9.3 orders these sets. In fact we can illustrate some conditions under which the two orderings are different in the case n = 4. Maintaining assumption 9.6, suppose d.m. weakly prefers $\{e_3, e_4\}$ to $\{e_1\}$ and the following numerical values are determined:

(9.16)
$$u_1 = 10$$
$$u_2 = 8$$
$$u_3 = 5$$
$$u_4 = 3$$
$$a = 1$$
$$b = 3$$

Then the numerical values of W (9.3) and w (9.15) are given in tabular form in 9.17.

(9.17)

S	$\{e_4\}$	$\{e_3\}$	$\{e_2\}$	$\{e_1\}$	$\{e_3, e_4\}$
W(S)	3	5	8	10	8
w(S)	3	5	8	10	11

S	$\{e_2, e_4\}$	$\{e_2, e_3\}$	$\{e_1, e_4\}$	$\{e_1, e_3\}$	$\{e_1, e_2\}$
W(S)	11	13	13	15	18
w(S)	14	16	16	18	21

This example shows clearly, then, that our naive procedure and 9.15 don't necessarily order the sets in $C_1 \cup C_2$ in the same manner. Indeed, in this case w and W produce the same orderings if and only if

$$(9.18) \quad u_3 + u_4 \geq u_1$$

and there is no reason to expect that 9.18 will hold in general.

We've only discussed one and two-element subsets, but these ideas are easily extended. For $i = 1, 2, \ldots, n$ let C_i denote the class of i-element subsets of EXPS. Then under independence assumptions analogous to 9.12, W is a BUI on each C_i. By providing judgmental inputs similar to those illustrated in 9.14, these BUI's can be "linked" to form one BUI, w, on $C_1 \cup C_2 \cup \cdots \cup C_n$, where w has the following form:

(9.19) If $E \in C_i$

$$w(E) = a_i \left(\sum_{e \in E} u(e) \right) + b_i$$

Where $a_i > 0$ and b_i are empirically determined constants ($a_1 = 1$, $b_1 = 0$).

Thus if rather reasonable assumptions analogous to 9.12 apply, it's possible to obtain a BUI, w, on $C_1 \cup C_2 \cup \cdots \cup C_n$. In general, though, our naive procedure produces a different ordering of the subsets of EXPS than w produces.

Now w suffices for deterministic and nondeterministic versions of type 3.2(d) problems where EXPS is the relevant set of experiments. Unfortunately, though, w is not generally equivalent to the "pure" additive form of W in 9.3.

This possibility of nonequivalence is unfortunate because on theoretical grounds it eliminates the use of certain formulations of type 3.2(d) problems, particularly integer linear programming formulations. This is not to say that such formulations are always inappropriate, but it does mean that we cannot always justify direct use of W. Indeed there might be cases where direct use of W is either theoretically appropriate or close enough for practical purposes. Also, there might be cases where w can be indirectly used in an integer programming version of a type 3.2(d) problem -- provided the problem is "shrewdly" formulated. We cannot, however, expect these fortunate circumstances to generally exist. At any rate the alternative of direct search methods is always available and such methods will suffice for practical purposes provided the number of elements in EXPS is not unduly large.

10. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This study has sketched the broad features of an approach to the problem of rational selection of scientific and technical experiments for space missions. A central feature of that approach is the concept of the value of an experiment, and this study has suggested and developed one method of making that concept more precise and operational. The suggested method requires judgmental inputs of informed and concerned individuals and yields numerical values with enough properties to be useful for decision purposes at various stages in the selection process. The major facets of this approach have been explored in some detail, and it appears that the approach is both theoretically sound and practicable.

Some detailed work would be necessary, of course, before this methodology could be implemented. Specifically, some of the detailed issues mentioned in sections 6, 7, and 8 would need to be resolved. While these are not trivial matters, they certainly can be handled without too much difficulty.

Also some attention needs to be focused on more detailed and specific formulations of type 3.2(d) problems for various stages of the experiment selection process.

Now if we have no clear, specific, and operational concept of an experiment's value, formulation and analysis of type 3.2(d) problems seems rather superfluous. For this reason I recommend that attention initially be focused on the value assessment process. When it becomes clear that this process can be implemented, some attention should be turned to the specific details of appropriate formulations of type 3.2(d) problems.

REFERENCES

1. Arrow, Kenneth J., Social Choice and Individual Values, Monograph 12 in the Cowles Commission Series, Wiley, New York, 1963 (2nd edition).

2. Bartee, Edwin M., "Optimization of Experimental Lunar Payloads," Management Science, Vol. 14 No. 2, October, 1967, pp. B-28 - B-40.

3. Blackwell, D. and M. A. Girshick, Theory of Games and Statistical Decisions, Wiley, New York, 1954.

4. Brown, Larry G., "Use of Zero - One Linear Programming for Scheduling Experiments for the Space Station," Research Reports, 1971 NASA - ASEE Summer Faculty Fellowship Research Program, (MSFC - Alabama/Auburn Program) pp. 23-31.

5. Chernoff, H. and L. E. Moses, Elementary Decision Theory, Wiley, New York, 1959.

6. De Groot, M., Optimal Statistical Decisions, McGraw-Hill, New York, 1970.

7. Fishburn, Peter C., Decision and Value Theory, Wiley, New York, 1964.

8. Fishburn, Peter C., "Independence in Utility Theory with Whole Product Sets," Operations Research, Vol. 13, 1965, pp. 28-45.

9. Fishburn, Peter C., "Additivity in Utility Theory With Denumerable Product Sets," Econometrica, Vol. 34, 1966, pp. 500-503.

10. Fishburn, Peter C., "A Note on Recent Developments in Additive Utility Theories for Multiple Factor Situations," Operations Research, Vol. 14, 1966, pp. 1143-1148.

11. Fishburn, Peter C., "Additive Utilities With Incomplete Product Sets: Applications to Priorities and Assignments," Operations Research, Vol. 15, 1967, pp. 537-542.

12. Fishburn, Peter C., "Conjoint Measurement in Utility Theory With Incomplete Product Sets," Journal of Mathematical Psychology, Vol. 4, 1967, pp. 104-119.

13. Gradecak, V., "Worth Analysis Study," (Oral Presentation) NASA Contract NAS8-25140, 1971.

14. Hadley, George, Introduction to Probability and Statistical Decision Theory, Holden-Day, San Francisco, 1967.

15. Herstein, I. N. and J. Milnor, "An Axiomatic Approach to Measurable Utility," Econometrica, Vol. 23, April, 1953, pp. 291-297.

16. Keeney, Ralph L., "Quasi - Separable Utility Functions," Naval Research Logistics Quarterly, Vol. 15, 1968, pp. 551-565.

17. Keeney, Ralph L., "Utility Functions for Multiattributed Consequences," Management Science, Vol. 18, No. 5, Part 1 (Jan., 1972).

18. Keeney, Ralph L., "Multiplicative Utility Functions," Technical Report 70, Operations Research Center, Massachusetts Institute of Technology, March, 1972.

19. Krantz, David, "A Survey of Measurement Theory," in G. B. Dantzig and A. F. Veinott (eds.), Mathematics of the Decision Sciences (Part 2), American Mathematical Society, Providence, Rhode Island, 1968, pp. 314-350.

20. Luce, R. D. and H. Raiffa, Games and Decisions, Wiley, New York, 1959.

21. Luce, R. D. and J. W. Tukey, "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement," Journal of Mathematical Psychology, Vol. 1, 1964, pp. 1-27.

22. Raiffa, Howard, Decision Analysis: Introductory Lectures on Choices Under Uncertainty, Addison-Wesley, 1968.

23. Raiffa, Howard, "Preferences for Multi-Attributed Alternatives," RM - 5868 - DOT/RC, The RAND Corporation, April, 1969.

24. Scherer, Lee and Committee, "Report of Space Investigations Management Committee," Internal NASA Report APO-1538, July 6, 1971.

25. Suppes, P. and J. Zinnes, "Basic Measurement Theory," in R. D. Luce, R. R. Bush, E. Galanter (eds.), Handbook of Mathematical Psychology, I, Wiley, New York, 1963, pp. 1-76.

26. Ting, H. M., "Aggregation of Attributes for Multi-Attributed Utility Assessment," Technical Report No. 66, Operations Research Centers, Massachusetts Institute of Technology, August, 1971.

27. von Neumann, John and Oscar Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, N. J., 1947 (2nd edition).

28. _____, "Management Instruction on Submission and Selection of Advanced Research and Technology Space Flight Experiments," Internal NASA Document NMI 8030. 5, April 12, 1968.

29. _____, "Plan for Proposal Evaluation of Materials Science and Manufacturing in Space Experiments for Skylab and CSM Missions," Internal NASA Document, Office of Manned Space Flight, Advanced Missions Program, March 15, 1972.

NAMES AND ADDRESSES OF FELLOWS AND STAFF

DeHoff, Paul H.
Bucknell University
Lewisburg, PA 17837
524-1113

DeSmet, Donald J.
University of Alabama
Physics Dept.
Box 1921
University, AL 35486
348-5050

Dodd, Jimmy L.
Mississippi State Univ.
Dept. of Electrical Engr.
Drawer EE
State College, MS 39762
601/325-3912

Dudgeon, James E.
University of Alabama
Electrical Engr. Dept.
Box 6167
University, AL 35486
348-6351

Elliott, J. Mark
Mechanical Engr. Dept.
Univ. of South Ala.
Mobile, AL 36688
460-6421

Glaser, Jerome C.
Aerospace Engr. Dept.
Iowa State Univ.
304 Engr. Bldg. No. 2
Ames, Iowa 50010
515/294-6796

Griggs, Edwin I.
Tenn Tech. Univ.
Dept. of Mechanical Engr.
P. O. Box 5213
Cookeville, TN 38501
615/528-3357

Hawks, Keith H.
Mechanical Engr. Dept.
Purdue University
Lafayette, IN 47907
493-3940

Hendricks, William L.
Aerospace Engr. Dept.
Calif. St. Polytechnic College
San Luis Obispo, CA
546-2755

Honea, Robert B.
1920 N. W. 7th Place
Gainesville, FL 32601

Jeng, Duen-Ren
Mechanical Engr. Dept.
Univ. of Toledo
2801 West Bancroft
Toledo, OH 43606
491/531-5711 EXT 2664

Karr, Gerald R.
Fluid & Thermal Engr.
Univ. of Ala. in Huntsville
Huntsville, AL

Kusmiss, John H.
Western Michigan Univ.
Physics Dept.
Kalamazoo, MI 49001
383-0995

Martin, Frank E.
Physics Dept.
Central Missouri State University
Warrensburg, MO 64093
816/747-7967

Moore, Jerry D.
La. Tech. Univ.
Electrical Engr. Dept.
Ruston, LA 71270
318/257-4801

Morrison, Richard A.
Dept. of Physics
Talladega College
Talladega, AL
205/362-8800

Olson, Judy M.
University of Ga.
Dept. of Geography
Athens, GA 30601
404/542-4141

Parker, Murl Wayne
Miss. State Univ.
Dept. of Industrial Engr.
P. O. Drawer U
State College, MS 39762
601/325-3828

Rogers, Jerry W.
Electrical Engr. Dept.
Mississippi State Univ.
Drawer EE
State College, MS 39762
601/325-3912

Rudland, Robert S.
1411 Peachtree Battle Ave.
Atlanta, GA 30327

Sawyer, Gary A.
Electrical Engr. Dept.
South Dakota State Univ.
Brookings, SD 57006
605/688-5924

Sheldon, John W.
Dept. of Physical Sciences
Florida International Univ.
Tamiami Trail
Miami, FL 33144

Sitchin, Amnon
c/o Ford Motor Company
Lt. Truck Advanced Engr.
20,000 Rotunda Dr. Bldg. #1
Dearborn, MI 48121

Smith, Paul D.
Civil Engr. Dept.
Auburn Univ.
Auburn, AL 36830
205/826-4320

Williams, Fred E.
Ga. Tech.
College of Industrial Management
225 North Ave. N.W.
Atlanta, GA 30332
404/894-2626

O'Brien, J. Fred, Jr. (Dir.)
Engineering Extension Service
Auburn Univ.
Auburn, AL 36830

Raney, Donald C. (Assoc. Dir.)
Box 6307
University, AL 35486

Kent, Marion I.
AD-S
Marshall Space Flight Center, AL 35812

☆ U. S. GOVERNMENT PRINTING OFFICE: 1972 — 746-823 / 4748