NHSACR-164 $3 / 1$
N81:11977

# NASA CONTRACTOR REPORT 

NASA-CR-161511

$$
19810003469
$$

NASA CR-161511

# RESEARCH REPORTS - 1980 NASA /ASEE SUMMER FACULTY FELLOWSHIP PROGRAM 

University of Alabama
University, Alabama
and
University of Alabama in Huntsville Huntsville, Alabama

October 1980

SEP 11981



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

15. SUPPLEMENTARY NOTES

## 16. ABSTRACT

This document is a collection of reports on the research conducted by the participants in the 1980 NASA/ASEE Summer Faculty Fellowship Program which was held at the George C. Marshall Space Flight Center (MSFC) for the sixteenth consecutive year. The 1980 program ran from May 26, 1980 through August 1, 1980 and was conducted by the University of Alabama in cooperation with MSFC and the University of Alabama in Huntsville. As in the past, the program was operated under the auspices of the American Society for Engineering Education (ASEE) with sponsorship and funding by the University Affairs Office, NASA Headquarters, Washington, D.C. The University of Alabama and Marshall Space Flight Center conducted one of the seven Aeronautical and Space Research Programs which were funded nationally during the summer of 1980. The other programs were conducted at six other NASA centers throughout the country in cooperation with other universities.

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

| 17. KEY WORDS |  | 18. DISTRIBUTION STATEMENT <br> Unclassified-Unlimited <br> J. A. BETHAY <br> Dfrector, Executive Staff, MSFC |  |
| :---: | :---: | :---: | :---: |
| 19. SECURITY CLASSIF. (of thil report) Unclassified | 20. SECURITY CLASSIF. (of this pag <br> Unclassified | 21. NO. OF PAGES <br> 731 | 22. PRICE <br> NTIS |

## RESEARCH REPORTS

1980 NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAMGeorge C. Marshall Space Flight CenterThe University of Alabama
and
The University of Alabama in Huntsville
EDITORS:
Dr. B. F. Barfield, Professor of Mechanical Engineering
The University of Alabama
Mr. Marion I. Kent, Office of University Affairs
Marshall Space Flight Center
Dr. James Dozier, Director, Research and Technology Office
Marshall Space Flight Center
Dr. Gerald Karr, Professor of Mechanical Engineering
University of Alabama in Huntsville

$$
\text { NASA } \quad C R-161511
$$

## PREFACE

For the sixteenth consecutive year, a NASA/ASEE Summer Faculty Fellowship Research Program was conducted at the Marshall Space Flight Center (MSFC). The program was conducted by The University of Alabama and MSFC during the period May 26, 1980 through August 1, 1980. The program was operated under the auspices of the American Society for Engineering Education (ASEE). The program at MSFC, as well as those at five other NASA Centers, was sponsored and funded by the Office of University Affairs, NASA Headquarters, Washington, D.C. The basic objectives of the programs, which are in the seventeenth year of operation nationally, are:
a. To further the professional knowledge of qualified engineering and science faculty members;
b. To stimulate an exchange of ideas between participants and NASA:
c. To enrich and refresh the research and teaching activities of participants; instutions; and,
d. To contribute to the research objectives of the NASA Centers.

The Faculty Fellows spent ten weeks at MSFC engaged in a research project commensurate with their interests and background and worked in collaboration with a NASA/MSFC Colleague. This document is a compilation of Fellow's reports on their research during the Summer of 1980. University of Alabama Report No. BER-259-94 presents the Co-Directors' report on the administrative operations of the Research Fellowship Program. Further information can be obtained by contacting any of the editors.

## First-Year Fellows

I. Abrams, D.P.: "Development of a Facility for Large-Scale Testing of Multistory Buildings"
II. Brindley, Tom: "MSFC personnel Management Tasks: Recruit-ment and Orientation of New Employees"III. Brown, R.C.: "Analysis and Improvement of Contact DynamicsSimulation Using the 6 DOF Motion System"
IV. Chan, C.H.: "Pilot Signals for Large Active Retrodirective Arrays"V. Davis, Jack: "Ion Plating Studies for High TemperatureApplications"
VI. Doss, Deva: "Statistical Wind Profile Gust Model"
VII. Eastman, Robert: "Upgrading the Free Flying Rendezvous and Docking Simulator and the Orbital Servicer System:
VIII. Foster, Larry M.: "Material and Process Constants for aFlat Interface in the Bridgman-Stock-barger Technique"IX. Grone, Robert: "A Ranking Algorithm for Spacelab Crew andExperiment Scheduling"
X. Jandebeur, Thomas: "Continuous Flow Electrophoresis: An Upward Flowing System"
XI. Kleiner, Alex: "Modeling a Maintenance Simulation of the Geosynthronous Platform"
XII. Martin, Benjamin: "Axially Symmetric Motion of a RotatingStratified Fluid"
XIII. Mathis, Frank: See Grone above
XIV. Self, Michael: "Payload Crew Training Complex (PCTC) Utilization and Training Plan"
XV. Shand, Julian: "Some Methods in High Energy Cosmis Ray Measurement"
XVI. Sheridan, Peter S.: "Photochemistry of Rhodium(III)-Carbonato Complexes"
XVII. Springer, John: "The Growth and Characterization of Triglycine Sulfate"
XVIII. Voss, James: "Fluid Dynamic Analysis of the Space Shuttle Main Engine High Pressure Oxidizer Turbopump Slinger Seal"
XIX. Wallace, Donald: "Validation of the Solar Heating and Cooling High Speed Performance (HISPER) Computer Code"
XX. Wang, David: "Incipient Failure Detection (IFD) of Ball Bearings"
XXI. Workman, Gary: "Pulsed Eddy Current Testing:
Second-Year Fellows
XXII. Antar, Basil: "Stability of the Stratified Cylindrical Annulus Flow"
XXIII, Barnes, Ronnie: "On Rosen's Theory of Gravity and Cosmology"
XXIV. Bond, Robert: "Measurements of the Sampling Volume of an Infrared Laser Doppler Velocimeter for the Study of Feasibility of Simultaneous ThreeDimensional Velocity Profile Mapping"
XXV. Davenport, Harold: "A Monte Carlo Approach to Tolerance Analysis of a Near-Diffraction Limited Optical Imaging System"
XXVI. Edwards, Eugene: "Power Distribution for Electron Beam Welding"
XXVII. Fitzjarrald, D.: "Measurement of Velocity Fields in Fluid Dynamics Experiments"
XXVIII. Greene, Michael: "On the Control and Stability of the PinholeOcculter Flexible Boom Facility"
XXIX. Kajfez, Darko: "FFT Methods in Signal Processing of the Coal Interface Detector Radar"
XXX. Korbly, Letitia: "The Motion of a Stratified Column of Fluid"
XXXI. Perel, David: "Improvements to the NASAP Code"
XXXII. Stoll, Henry: "Deployable Structure Design for the Science and Applications Space Platform"
XXXIII. White, William: "Payload Maintenance Cost Mode1 for the Space Telescope"

# MARSHALL SPACE FLIGHT CENTER 

 THE UNIVERSITY OF ALABAMADEVELOPMENT OF A FACILITY FOR LARGE-SCALE TESTING OF MULTISTORY BUILDINGS

| Prepared By: | Daniel P. Abrams, Ph. D. |
| :---: | :---: |
| Academic Rank: | Assistant Professor |
| University and Department: | University of Colorado at Boulder <br> Civil,Environmental and Architectural Engineering |
| NASA/MSFC: (Laboratory) (Division) | Space Sciences A + mospheric Sciences |
| MSFC Counterpart: | Dr. Nicholas Co Costes |
| Date: | August 1, 1980 |
| Contract No. : | NGT-01-002-099 <br> (University of Alabama) |

## By

Daniel P。Abrams, Ph. D. Assistant Professor of Civil Engineering University of Colorado Boulder, Colorado

Abstract

Current experimental research pertaining to response of structures subjected to lateral loads such as strong winds or earthquake motions is limited to either tests of single structural components or small-scale (approximately one-tenth) multistory structures. The feasibility of developing a facility where large-scale multistory structures could be loaded to failure is discussed.

The test facility would consist of a series of hydraulic actuators mounted on reaction frames currently in use at the Marshall Space Flight Center for structural testing of spacecraft components. The actuators would be controlled from signals computed by an on-line analysis of measured data. This method of loading could be used to simulate inertial forces resisted by a structure behaving in the nonlinear range of response subjected to motion at the base or to impulses along the height as would occur during a strong earthquake or wind.

Development and utilization of the test facility would occur in two phases. In the first phase, the testing system would be developed and verified for loading a structure within a single plane. The second phase would utilize the loading system by testing a half-scale ten-story reinforced concrete building 45 feet tall. Further development of the test facility may include the capability of loading individual floor levels to simulate translational and rotational inertia in three directions.

The writer expresses his appreciation to the National Aeronautics and Space Administration, and the University of Alabama and the American Society of Engineering Education for funding and conducting the Summer Faculty Fellowship Program at the Marshall Space Flight Center under the direction of Marion Kent of NASA and Dr. Robert Barfield of the University of Alabama.

Special gratitude is extended to Dr. Nicholas Costes who provided many creative ideas and stimulating conversations, and was responsible for introducing the writer to numerous helpful individuals at Marshall.

The writer wishes to thank Dr. George McDonough for his support of the project and George Shofner, Charles Watson, Frank Vinz and Gerry Waggoner for their technical advice and constructive comments.

Credit is given to Willa Eslick for her patience in typing the paper.

## INTRODUCTION

"Earthquakes have caused, and can cause in the future, enormous loss of life, injury, destruction of property, and economic and social disruption."

This statement quoted from the Earthquake Hazards Reduction Act of 1977 [1] demonstrates the importance for safer structures in regions of high seismicity. More specifically, the primary objective of the Act. passed by Congress is
"The development of technologically and economically feasible design and construction methods and procedures to make new and existing structures, in areas of seismic risk, earthquake resistant..."

In accordance with this objective, several investigators have examined both analytically and experimentally the response of structures subjected to strong ground motions. Of the experimental studies, investigations have been limited to response of one two general specimen types: large-scale single components of more complex structures, or small-scale (on the order of one-twelfth) multistory structures. Large-scale multistory structures have not been tested to destruction because costs of developing a loading system and constructing a test structure were thought previously to be prohibitive.

This paper presents the feasibility of developing a facility where largescale multistory structures could be loaded to failure. The proposed test. facility would consist of a series of hydraulic actuators mounted on reaction frames currently in use at the Marshall Space Flight Center (MSFC) for structural testing of spacecraft components. Jevelopment and utilization of the test facility is discussed in terms of the cost-value relationship of using large-scale test structures for earthquake engineering research.

## OBJECTIVES

Results of an experimental study using a large-scale test structure could be the necessary evidence required to transfer existing research findings to standards for building earthquake-resistant structures. Moreover, tests of large-scale structures may help discover and develop new methods of designing and constructing safer structures. Measured response of a largescale multistory structure would:
(1) verify the capability of the testing system to simulate inertial loads resulting from a programmed impulse;
(2) for the first time, provide comprehensive data for an indepth study of both the overall and local response of a multistory structure subjected to a strong ground motion;
(3) verify currently used small-scale modeling procedures thus extend experimental research of multistory structures using shaking tables;
(4) provide a data base for testing newly developed analytical models used for calculating nonlinear dynamic response of multistory structures;
(5) serve as a "proof test" of recent improvements in earthquakeresistant design procedures developed with mathematical and small-scale physical models;
(6) provide data to evaluate methods of rehabilitation of existing and damaged structures such as epoxy grout injections or masonry in-fill panels.

The objective of the study presented in this paper is to examine the feasibility of developing a large-scale testing facility and to outline a test program utilizing the facility to illustrate these attributes of large-scale testing.

A hypothetical structure (Fig. 1) has been chosen as an example of a typical test article. The structure is a one-half scale reinforced concrete building consisting of two three-bay frames coupled to a slender shear wall at ten levels with floor slabs poured monolithically. The configuration was chosen to replicate a similar configuration previously used with small-scale models [2,3,4]. The scale was chosen so that sizes of members and reinforcement would be representative of actual construction. It is planned to develop a testing system to load this structure laterally at all ten levels from a single direction.

## OUTLINE OF TEST PROGRAM

An outline of the test program is presented in Fig. 2. The program would consist of three stages. In the first stage the testing system would be developed using existing loading equipment at MSFC and a computer analysis that would be developed. The second stage would consist of verifying operation of the testing system by constructing and loading a replica of a smallscale (approximately $8-\mathrm{ft}$. tall) ten-story structure tested previously on a shaking table [2]. The loading procedure which would consist of a series of slowly applied loading reversals simulating inertial forces could be validated
by comparing displacement response of the structure tested on a shaking table with that of the structure loaded laterally with actuators. If agreement can be made, then the loading system would be developed further to test the $45-\mathrm{ft}$ tall structure shown in Fig. 1 using the same load control system and on-line computer analysis with larger size actuators and reaction structure.

It is planned to test the $45-\mathrm{ft}$ tall half-scale structure using the same input base - acceleration history as used for the 8 -ft tall structure. In this way, scaling relationships may be examined for modeling reinforced concrete at small scales. After the structure has been tested, it can be repaired and loaded once again to investigate methods of rehabilitation. The test structure then can be used as a loading system for investigating the behavior of masonry sub-panels by repairing the structure once again, in filling each frame panel with masonry and retesting. In this manner, each of sixty frame panels, or masonry elements, will be subjected to a different ratio of shear to normal stress thus providing a larger population of data than currently exists in all of the past masonry research.

## LOADING SYSTEM

The proposed loading system consists of a series of hydraulic actuators mounted to an existing reactor frame used presently for testing portions of the liquid oxygen tanks of the Space Shuttle. The reactor structure (Fig。3) which is located within the structural test tower (Fig. 4) of the load test annex of building 4619 has sufficient lateral-load capacity to support the 20 actuators (two per level, Table l) required to load the 45 -ft tall test structure. Minor modifications of the reaction structure would be necessary to support individual actuators. In addition, the reaction structure could be used to support actuators mounted in transverse and vertical directions for further development of the loading facility.

A contingency plan could also be developed to use the existing test tower (Fig.4) if removal of the reaction frame were necessary for other reasons. The test area measures 5l x $49 \times 47$ meters high. A double trolley crane with a total capacity of 54 metric tons is accessible for erection of the test structure. Further detailed information about the facility is given in Reference 5.

The existing load control system located in building 4619 has been cited [6] by the Japanese as being a unique system for controlling several actuators simultaneously. A closed-loop control system using a PDP-ll computer monitors response of a specimen at intervals of $1 / 60$ th of a second and adjusts the load to attain a preselected specimen response. This method of control is particularly useful if it is desired to displace a structure through
a specified deflected shape. Actuators loading a structure at several levels can be programmed to displace prescribed amounts despite interaction among actuators which has troubled investigators previously. The capability of the load-control system to coordinate simultaneous actuator movement is proven by numerous experiments including a recent one where several actuators were used to load in three directions an ortho-grid assembly used for part of the floor system in Space lab. The proposed loading facility would utilize ten of fifty-eight channels of the existing load-control system.

The actuators would be controlled from signals generated by an on-line analysis of measured data to simulate inertial loads resulting from a prescribed impulse. For a standard time increment of the impulse duration, each actuator would displace an amount calculated from previous force and displacement measurements and the prescribed forcing function. The differential equation to be satisfied for each floor level is:

$$
\begin{equation*}
m \ddot{x}+F_{T}=p(t) \tag{1}
\end{equation*}
$$

where $m$ is the mass of the floor and lumped portions of the columns and wall, $\ddot{x}$ is the acceleration of the mass, $F_{r}$ is the restoring force of the structure, and $p(t)$ is the prescribed forcing function. Using finite differences, the acceleration may be related to the displacement over a time increment by

$$
\begin{equation*}
\ddot{x}_{i}=\left(x_{i+1}-2 x_{i}+x_{i-1}\right) /(\Delta t)^{2} \tag{2}
\end{equation*}
$$

where $\Delta t$ denotes the time interval and i refers to the particular time increment. The new displacement, $x_{i+1}$, that controls the actuator movement, may be calculated from $\dddot{x}_{i}$ which is determined from equation (1)

$$
\begin{equation*}
\ddot{n}_{i}=p(t)_{i}-F_{r i} \tag{3}
\end{equation*}
$$

For an impulse at the base of a multistory structure, $p(t)$ becomes m $\ddot{x}_{0}$ where $\ddot{x}_{o}$ is the input acceleration at base. Any prescribed ground motion including translational, rocking or torsional motions may be input to the online analysis to load the fixed-base structure "pseudo dynamically."

The restoring force, $\mathrm{F}_{\mathrm{r}}$, is measured continuously during the loading which is particularly important for a nonlinearly behaving structure because the restoring force or stiffness of the structure is constantly changing and sometimes unpredictable. For a nonlinear dynamics problem the fact that the inertial forces are dependent on the stiffness, and the stiffness is dependent on the loading history, justifies the on-line analysis as an irreplaceable technique for simulating inertial loads.

Because the pseudo-dynamic test method is not velocity dependent, the actuators may be operated as slow as the hydraulic power supply requires. Some investigators have even advocated that slowing the loading process down to a nearly static rate is good for observing propagation of cracks and other types of progressive damage. It should be noted that the loading method does not account for effects of viscous, or velocity dependent damping, and strain-rate effects. However, it is reasonable to assume that most of the energy dissipation in reinforced concrete structures behaving within the nonlinear range of response is attributable to hysteretic damping which is not velocity dependent. Furthermore, strain-rate effects for velocities corresponding to the fundamantal mode and possibly the second mode of vibration of prototype structures have been shown to be negligible [7].

## EXPERIMENTAL PROCEDURES

To economize on formwork the test structure (Fig. 1) would be erected from planar wall and frame elements cast in the horizontal plane and lifted to the vertical position. Reinforcement at base of columns and wall would be anchored in concrete blocks prestressed to the floor of the test area to prevent welding of reinforcing bars. Floor slabs could then be cast to complete erection of the structure. A minor amount of disassembly of the reaction structure may be necessary for lifting and placement of the structure.

Instrumentation would consist of two displacement and two load transducers at each level of the structure. Additionally, strain gages would be placed on reinforcing bars to detect internal forces in members and to infer bond characteristics between concrete and reinforcing bars. An estimated maximum of 200 channels of data would be a small percent of the 6000 channel Structural Test Data Acquisition System (STDAS) available at MSFC.

The test structure would be subjected to a series of lateral loads simulating response to the ground motion measured at El Centro, California, in the 1940 Imperial Valley earthquake (NS component). The intensity of the input motion would be scaled according to an analysis used for design of the structure so that similar levels of damage would be attained as for the smallscale structures. After the first test run, or loading, the structure would be loaded corresponding to an input motion of twice the intensity to further investigate strength and stiffness deterioration and energy dissipation characteristics. This pattern may be repeated until destruction occurs. At that time, the structure may be repaired using epoxy grout injections in the cracked concrete and tested once again. If the rehabilitation process can regain a sufficient amount of the original stiffness, the structure can be repaired once again and then strengthened with concrete-block masonry
placed in each frame panel. Instrumentation can be installed to measure normal and shear stresses on each masonry panel when the structure is loaded once again.

In addition to the pseudo-dynamic testing, a series of low-amplitude free and forced vibration tests can be run to determine dynamic characteristics such as modal frequencies, shapes, and damping factors.

## COST ESTIMATE

Approximate costs of developing a large-scale testing facility to load a $45-\mathrm{ft}$ tall ten-story structure are itemized below. Costs include verification of the pseudo-dynamic loading system, and construction of the reinforced concrete test structure. Cost estimates have been compiled by engineers at MSFC who would be directly involved with the proposed work.
(1) Verify pseudo-dynamic testing method. Costs include testing an 8 - ft tall model structure with the newly developed test method.

| Reaction structure | $\$ 5,000$ |
| :--- | :---: |
| Loading system | $\$ 20,000$ |
| Instrumentation | - |
| Construction of $8-\mathrm{ft}$ model | $\$ 8,000$ |
|  | $\$ 33,000$ |
| Man months | 30 |

(2) Develop large-scale testing facility. Costs include modifying existing structural test equipment and purchasing needed actuators, hardware, cables, hydraulic lines and miscellaneous items.

| Reaction structure | $\$ 30,000$ |
| :--- | :---: |
| Loading system | $\$ 65,000$ |
|  | $\$ 95,000$ |
| Man months | 35 |

(3) Constructing and testing $45-\mathrm{ft}$ tall structure

Construction of structure $\$ 85,000$
Instrumentation $\$ 7,000$
Disassembly $\$ 10,000$

$$
\$ 102,000
$$

Man months 25
(4) Total

Man months
\$230, 000
90

The overall cost of the project appears to be justifiable when compared with a similar attempt in Japan [6] and a previous feasibility study [8], each costing well over twenty million dollars.

## EXTENDED DEVELOPMENT OF THE TEST FACILITY

Development of the loading system could be extended further in several steps to result ultimately in a loading system that could simulate translational inertia of each floor slab in three directions in addition to rotary inertia about three axes. Base motions input to the on-line analysis could then contain three translational components: two horizontal and one vertical, and three rotational components: torsion and rocking about each horizontal axis.

The existing reaction structure, hydraulic power supply, and loadcontrol system could be utilized. Two additional horizontal actuators. and two additional vertical actuators would be required per floor level.

Operation of the loading system could be verified by constructing several test structures at a small scale in duplicate. One of each pair of structures would be tested on the six-degree-of-freedom shaking table located in building 4663. The other structure of the pair would be tested using the newly developed loading system. Comparing response of each structure would verify the reliability of the pseudo-dynamic test method for multiaxial loading.

Furthermore, the existing structural test tower (Fig. 4) located in the local test annex of building 4619 is sufficiently tall and strong to accommodate a test structure up to l00-ft tall. A twenty-story half-scale, or ten-story full-scale, building could be tested with the additional cost primarily attributable to the purchase and support of additional actuators.

The feasibility of developing a facility for testing large-scale multistory buildings to failure has been presented. The facility would consist of a series of hydraulic actuators mounted to existing structural test equipment at, the Marshall Space Flight Center. Displacement of each actuator would be controlled by an on-line analysis of measured data to simulate motion of a multistory structure subjected to a prescribed series of impulses input at the base as would occur during an earthquake. Construction, instrumentation, and test procedures are outlined for testing a 45-ft tall reinforced concrete structure to illustrate the worth of large-scale testing and the approximate costs.

The role of a large-scale testing facility for earthquake engineering research has been shown to be consistent with the objectives of the Earthquake Hazards Reduction Act of 1977. Development of such a facility at the MSFC using existing structural test equipment appears to be feasible technically and justifiable economically.

## REFERENCES

1．＂Earthquake Hazards Reduction Act of 1977＂，Congressional Record，Vol．123，October 7， 1977.

2．Abrams，D．P．and M．A．Sozen，＂Experimental Study of Frame－ Wall Interaction in Reinforced Concrete Structures Subjected to Strong Earthquake Motions＂，Civil Engineering Studies，Structural Research Series，No．460，University of Illinois，Urbana，May 1979.

3．Healey，T．J．，and M．A．Sozen，＂Experimental Study of the Dynamic Response of a Ten－Story Reinforced Concrete Frame with a Tall First Story＂，Civil Engineering Studies，Structural Research Series， No．450，University of Illinois，Urbana，August 1978.

4．Moehle，J。P。，and M。A．Sozen，＂Earthquake－Simulation Tests of a Ten－Story Reinforced Concrete Frame with a Discontinued First－Level Beam＂，Civil Engineering Studies，Structural Research Series No．451， University of Illinois，Urbana，August 1978.

5．＂Potential Utilization of the NASA／George C．Marshall Space Flight Center in Earthquake Engineering Research＂，Earthquake Engineering Research Institute，Berkeley，California，December 1979．

6．U．S．－Japan Planning Group，＂Recommendations for a U．S．－Japan Cooperative Research Program Utilizing Large－Scale Testing Facilities＇， Earthquake Engineering Research Center，University of California， Berkeley，Report No．EERC－79／26，September 1979，pg。 92.

7．Staffier，S．R．and M．A．Sozen，＂Effect of Strain Rate on Yield Stress of Model Reinforcement，＂Civil Engineering Studies，Structural Research Series No．415，University of Illinois，Urbana，February 1975.

8．Penzien，J．，J．Bouwkamp，R．W。Clough，and D．Rea，＂Feasibility Study Large－Scale Earthquake Simulation Facility，＂Earthquake Engi－ neering Research Center，University of California，Berkeley，Report No．EERC－67／1，September 1967.

## Table 1

Actuator Requirements (2 per level)

| Level | $\frac{\text { Capacity }}{\text { (kips) }}$ | }{} | Stroke <br> (inches) |
| :---: | :---: | :---: | :---: |
| 10 | 35 | 5 |  |
| 9 | 35 | 5 | 48 |
| 8 | 35 | 5 | 48 |
| 7 | 35 | 5 | 38 |
| 6 | 25 | 4 | 38 |
| 5 | 25 | 4 | 29 |
| 4 | 25 | 4 | 29 |
| 3 | 25 | 4 | 19 |
| 2 | 15 | 3 | 19 |
| 1 | 15 | 3 | 10 |
|  |  |  | 10 |



Fig. 1 Test Structure


Fig. 2 Outline of Experimental Program


Fig. 3 Reaction Structure


Fig. 4 Structural Test Tower

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM 

MARSHALL SPACE FLIGHT CENTER 'THE UNIVERSITY OF ALABAMA

MSFC PERSONNEL MANAGEMENT TASKS: RECRUITMENT AND ORIENTATION OF NEW EMPLOYEES

Prepared $13 y$ :

Academic Rank:

University and Department:

NASA/MSFC:
Oflice
Jivision

MSFC; Connterpait:

Date:

Contract No.:
'Thomas A. Brindley, Ph. D.

Associate Professor

The University of Alabama in Huntsville Department of Education

Personnel Office
Personnel Development

James R. Johnson

Migusi 1, 1980

N(;"-0)-002-099
(laiversity of Alabimma)

# MSFC PERSONNEL MANAGEMENT TASKS: RECRUITMENT AND ORIENTATION OF NEW EMPLOYEES 

BY
Thomas A. Brindley, Ph. D. Associate Professor of Education The University of Alabama in Huntsville Huntsville, Alabama

## ABSTRACT

In order to encourage highly motivated young students to learn about NASA and consider it for a career, a formal program will be initiated whereby selected students can work on a voluntary basis at Marshall Space Flight Center. These students will work under the guidance and supervision of expert NASA employees. The first task has been to develop the working plan and procedures for this program, called Student Volunteer Service Program, in the writing of MSFC official guidelines, the Marshall Management Instruction (the MMI) which is a binding document that defines policy and establishes procedures and guidelines. Particular considerations written into the MMI after numerous consultations, interviews, and discussions about a satisfactory policy, include: arrangements to be made between the student, the school authorities, and concerned MSFC employees; management of the work assignments; and procedures for the student's welfare and safety.

The second task has been the development of a Recruitment Brochure for the attraction of new employees, especially scientists and engineers. Since private industry has certain advantages in hiring, particularly in offering higher initial salaries, it is imperative that NASA stress the valid reasons why working at MSFC would be of great benefit to the prospective employee. Thus, the brochure needs to stress aspects of ongoing NASA projects which show how various disciplines of science and engineering apply.

The third task assigned has been to develop a plan called Orientation of New Employees. The main intention is to establish a procedure whereby new employees can best become acquainted in their new work with a minimum of stress involved in coming to a new organization. This plan attempts to schedule various experiences that will enable the new employee to understand not only the particular work situation for which the employee is assigned, but to show the overall mission and ope rations of the Marshall Space Flight Center. The plan is to indicate the types of orientation believed necessary to aid the new employee in adjusting to work at MSFC.

## ACKNOWLEDGEMENTS

I wish to thank Jim Johnson, Norm Hochberger, and Katy Lyle in particular for their constant and diligent review of my work, for their many suggestions, and for their great encouragement. Many others in the Personnel Office of MSFC have been very helpful and supportive not only in acquainting me with aspects of the specific tasks but in familiarizing me with the overall concerns of personnel management in general. I am also indebted to the many supervisors, scientists, engineers, technicians, and other employees in many different organizations of MSFC for taking their time to introduce me to various missions and to the many accomplishments of the dedicated people here at Marshall Space Flight Center. And, truly, I wish to express grateful appreciation to Helene Zazulak, who typed many manuscripts and this final report for me this summer.

## INTRODUCTION

The Personnel Development Division of the Personnel Office requested that three tasks be accomplished by this NASA/ASEE Summer Faculty Fellow. The tasks were to investigate the needs and ramifications of several aspects of recruiting and orientation of employees and to write up a study or report on each of the subjects, taking each project as far as it could be carried in the ten weeks. Specifically, these management tasks were:

1) the orientation of new employees;
2) a recruiting brochure for the hiring of new scientists and engineers; and,
3) a student volunteer service program.

These three tasks have been completed and are reported here in respective sections.

The research undertaken involved the reading of all materials pertinent to past actions on the subjects and numerous interviews, consultations, and discussions with concerned MSFC employees about satisfactory policies related to the se tasks. The tasks also demanded that this writer understand as much as possible about NASA and MSFC in a holistic, comprehensive way. Since the author had little previous acquaintance with MSFC in either a technological sense or in an organizational management sense, it was necessary to learn about NASA/MSFC in a short time so that the tasks could be begun. Thus, in effect this writer experienced his own orientation program.

## VOLUNTARY STUDENT SERVICE PROGRAM

In 1978, the MSFC Personnel Office in response to various university and school requests in the Huntsville area initiated inquiries into the feasibility of forming a plan to describe means whereby exceptional students could work at MSFC on a voluntary basis. The initial exchange of memorandum and letters regarding the subject were described as the "Worksite Experience Program."

Laboratory supervisors were asked if they had needs in their organizations for visiting students. Several responded that they would be interested in allowing students to work in their laboratories. The Associate Director of the Center and the Director of Science and Engineering were especially interested in pursuing this program.

Coincidentally and unrelated to local efforts, guidelines were being discussed at the national level to provide for voluntary student services. The Civil Service Reform Act of 1978, allowed agencies to devise memoranda or official guidelines which would establish workable programs of this nature. MSFC temporarily delayed pursuing this further until the assignment for this NASA/ASEE Fellow.

Until the Civil Service Reform Act of 1978, there were no provisions for agencies of the Federal government to accept voluntary or gratuitous services from interested persons or parties. Thus, visiting individual researchers and interested students could find no means to offer services to a Federal agency because no employment status existed for them. There were no arrangements or agreements to allow students to take part in government programs gratuitously. Thus, a talented high school student could not work at MSFC in a laboratory whe re the youngster could gain in knowledge and experience. Such an opportunity would also encourage the youngster to think about future employment at NASA after completing college work.

The first task assigned and completed was to write a proposed Marshall Management Instruction (the MMI) which is a binding document that defines policy and establishes procedures and guidelines. The MMI sets up
a formal mechanism whereby selected, highly motivated young students can work on a voluntary basis at MSFC. The following study plan, following the format of the Marshall Management Instruction, describes the Voluntary Student Service Program.

## STUDY PLAN: THE MARSHALL MANAGEMENT INSTRUCTION (DRAFT)

SUBJECT: Student Volunteer Service Program

## 1. PURPOSE

a. To provide a means whereby qualified students can volunteer their services to the Center in order to gain experience and knowledge.
b. To establish Center policy and procedures and to identify responsibilities in the management of the Student Volunteers.
c. To establish guidelines for the Student Volunteers while they are working at the Center.
2. SCOPE

This instruction is applicable to all elements of MSFC.
3. REFERENCES
a. FPM Chapter 308, Subchapter 7.
b. FPM Letter 308-16, November 8, 1979.
4. POLICY
a. High School students who show outstanding promise in their school work or school activities as determined by their school teachers and administrators in all areas of school study, especially, but not limited to science subjects, may wish to receive additional knowledge and expertise and be provided advanced opportunities to learn through on-the-job experience. These students will be allowed to work in a nonpay status on a voluntary basis at the Marshall Space Flight Center during a specified time and under the direction of a designated NASA employee.
b. Each student assignment will be established and conducted through written agreement between the Marshall Center and the educational institution or with the organization so
designated by schools or boards of education.*
c. The participating student is not to be considered a Federal employee for any purposes other than injury compensation or laws related to the Federal Tort Claims Act. Service is not creditable for leave accrual or any other employee benefits.
d. Assignments will be tailored individually at the chief discretion of the project officer (Mentor).

## 5. DEFINITIONS

a. The definitions of "volunteer service," "student," "half-time student," and "agreement" are those contained in FPM Chapter 308, Subchapter 7 and FPM Letter 308-16.
b. Mentor -- the MSFC employee responsible for the student's work, welfare, and learning experiences, whois in charge of the student's assignments, scheduling, reporting, and evaluation, and who supervises the student.

## 6. BACKGROUND

The authority for the acceptance of unconditional gifts and services donated to the agency is contained in the National Aeronautics and Space Act, Section 203 (c)(4). Based upon this authority the Center has been developing a program to allow participation in Center activities by students who are willing to donate their time and effort, in response to interest expressed by local educational institutions. The Civil Service Reform Act of 1978 authorized the acceptance of volunteer service and the Office of Personnel Management implemented regulations outlining procedures for the development of such programs government-wide.

## 7. GENERAL PROVISIONS

a. Responsibilities for the student's work at the Center shall be mutually entrusted to the Mentor. Participation in the program is conditioned upon the student's consent by written agreement to certain specified guidelines. Although students are not considered Federal employees, they will be allowed to work alongside and at the direction of Federal employees at the Center.
b. This programis limited to students enrolled not less than half time in high school, trade school, technical or vocational institution, or junior college, college or university.

* Agreements and forms, as attachments, are not included here.
c. Students receive no compensation.
d. The student offers service (comes to work at NASA) with permission of sponsoring school or institution where student is enrolled.
e. Nominations for participation will be made by the sponsoring school or institution where the student is enrolled. Selection shall be within the sole discretion of the Center.
f. The student will not replace or displace any Center employee or staff a vacancy.
g. The student will not be used to replace or substitute for routine and/or clerical duties while on assignment.
h. The student will not be allowed to enroll or partake in MSFC training courses.
i. Students will be at least 16 years old and must be at least in their junior year in High School. Students must be citizens of the United States. Participation shall not be based upon race, color, religion, sex, national origin, age, mental or physical handicap, marital status, or political affiliations, as prohibited by law, rule, or regulation.


## 8. RESPONSIBILITIES AND PROCEDURES

a. The academic institution will be responsible for:
(1) Requesting in writing to the MSFC Center Director placement of nominated students in designated disciplines.
(2) Recommending students who have demonstrated a serious interest in gaining work experience at the Center and who have demonstrated outstanding curricular and/or extracurricular achievements as deemed appropriate by the academic institution.
(3) Executing and complying with the terms of the Memorandum of Understanding with MSFC.
b. The MSFC Personnel Office will be responsible for:
(1). Selecting participants from those nominated by their educational institution.
(2) Identifying work locations.
(3) Appointing a Mentor to supervise the student during the student's assignment at MSFC.
(4) Negotiating appropriate agreements.
(5) Documenting volunteer service on Standard Form 50 (SF-50), "Notification of Personnel Action."
(6) Establishing an Official Personnel Folder (OPF).
(7) Evaluating the program.
(8) Submitting any required reports.
(9) Managing a one week orientation program to acquaint the student with the overall NASA-MSFC mission.
c. The MSFC Security Office will be responsible for:
(1) Providing proper authorization of the student's special status while assigned or visiting the Center.
(2) Acquainting the student with necessary security measures.
(3) Limiting the student only to non-security materials and areas.
d. The MSFC Mentor will be responsible for:
(1) Supervising the student's work and learning experiences.
(2) Supervising the student's assignments and schedules.
(3) Assigning the student a definite worksite and monitoring the student's activities.
(4) Evaluating the student's performance and submitting any necessary reports to the Personnel Office.
(5) Maintaining informal relationships with responsible parties of the student's academic institution.
e. The student will be responsible for:
(1) Complying with rules and regulations governing the internal operations of the Center. Any infractions of responsibilities may be cause for termination of the program.
(2) Being at the work location during the hours specified in the work experience Memorandum of Understanding or in other designated places upon approval of the Mentor.
(3) Following the authority of the assigned Mentor.
(4) Contacting the Mentor if absence from work is necessary.
(5) Caring for and properly maintaining all equipment and facilities used.
(6) Providing transportation to and from the MSFC worksite.
f. Mutual agreements between parties:
(l) A written agreement, called the Memorandum of Understanding, will be signed by the Director of the Personnel Office and the representative of the participating institution. The agreement will explain the general assignments, time of assignment, and any special requirement desired by the parties.
(2) A "Notification of Personnel Action," Form 50, will be completed and signed by Department Head and Applicant.
(3) An "Application For Use of Research Facility, Equipment, or Computation Service" will be completed and signed by designated persons.
(4) A 'Memorandum Agreement Relative to an Unconditional Donation of Services and Acceptance Thereof," will be completed and signed by designated persons.

The second management task has been the development of a recruitment brochure for the attraction of new employees, especially scientists and engineers. In 1968, the first reduction in force (RIF) affected MSFC. From time to time thereafter various reductions of employees occurred through 1979, so that the work force which was once over 7000 in the heyday of the Apollo-Saturn missions fell by Summer, 1980, to 3600 employees. Coupled with this, there are over 1500 employees who have reached twenty years of service and are eligible for retirement. The rate of employee attrition will certainly increase in the next five years unless new people can be brought in. Although there has been some hiring in critical skill areas over the last twelve years, there is abundant need to recruit and hire scientists and engineers at MSFC to maintain a strong and active work force. Even though there is a 3561 end of fiscal year 1980 employment ceiling at MSFC imposed by Congressional budgetary considerations, new hiring is very necessary.

Recruitment teams are sent to selected universities each year. There are numerous kinds of materials, but there is no special handout available for recruiters to utilize which is aimed definitely at attracting employees to this center to work on the kinds of projects in process for which MSFC is renowned, and to praise the advantages of living in the Huntsville, Alabama area. A recruitment brochure is deemed necessary to enhance the search for new employees of critical skills in science and engineering.

There were several important considerations that were discussed in the numerous interviews and inquiries taken by this researcher in order to determine a final plan for the nature of the brochure. It was determined that the particular engineering and scientific skills in which a prospective employee has been educated should be shown in their direct relationship to the kinds of projects and activities being developed at this particular center. A prospective employee should feel that his or her university preparation and past experience will be properly utilized at MSFC. The employee has to feel that there are opportunities to work on significant projects, to acquire a sense of mission and purpose, to see one's work in meaningful relation to the whole, and to work with materials and processes that enhance one's skills and draw upon one's best talents. The new employee has to be shown that MSFC allows for personal and professional growth and development and that MSFC provides abundant opportunities for advancement and recognition. Since private industry has certain advantages in hiring, particularly in offering higher initial salaries, a recruiting brochure must show how the Federal pay and
promotion scale and the MSFC organization allow an employee to gain in pay and status rather rapidly over the years. New employees are concerned about the area in which they live. Thus, the many advantages of the good life in the Tennessee Valley should be described.

The actual development of the brochure itself included a number of steps. The researcher had to ascertain what the needs are at MSFC and what kind of employees are being sought. The researcher had to learn what MSFC offers that would be of concern and interest to possible new employees. What especially attracts recent college graduates and/or experienced hands? Can NASA/MSFC attract candidates of the highest quality in the face of considerable competition from other governmental organizations and from private industry?

Various federal agencies, other NASA centers, and some aerospace companies were contacted about their own recruiting efforts. After this information was gathered, the researcher had to discuss the plan of the brochure with various supervisors, chiefly the head of recruitment.

Then the researcher made several plans, called thumbnail sketches. Priorities of arrangement and order were determined. The final sketch was presented to the graphics department to make a draft mockup. This is the stage reached by the end of the seventh week. Thereafter, each draft must then be developed further and refined. Appropriate pictures must be acquired and adequate text must be written. Then a final mockup is made by the graphics department. This is presented to various directors for approval and then submitted to the printers for the final edition.

The organizational pattern of the draft mockup recruitment brochure as completed follows in this order. First, an attractive cover is planned to show the chief ongoing project for which MSFC is mainly responsible -- the Space Shuttle itself. At the beginning of the booklet the Director of the Center introduces MSFC to prospective employees. The Center itself is shown and the geographic relationship of MSFC to other field centers is represented. Chief accomplishments of NASA in the past, in the present, and planned for the future are pictorially shown. Then the main missions and projects of the Marshall Center are depicted. These include the Space Shuttle, the Spacelab Payloads, various important scientific laboratories, and some of the spin-offs, materials processing in space, and special assigned tasks at MSFC, such as solar heating and cooling. Integrated into the portrayal of these various missions are references indicating how the education and experience of qualified engineers and scientists work on each of the projects in particular. Then some pages describe briefly how future employees will benefit by
working at MSFC. Several pages show how employees have special advantages and opportunities at MSFC in advancement, career development, pay and status, training programs, and professional enhancement. Finally, the attractive surroundings of the Tennessee Valley and the numerous educational, cultural, and social advantages of this area of America are portrayed. The sense of living in a multicultural, pleasant, and attractive community is stressed.

The various thumbnail drafts and the first mockup from graphics are already completed at the time of this writing. However, the bulk of the draft brochure precludes it being inserted as a part of this report. The first draft mockup brochure serves as a springboard for its final development and incorporates and organizes the many conceptual ideas into a meaningful plan for further action. The ongoing continuation of the recruitment brochure lies in the hands of the Personnel Director.

## ORIENTATION OF NEW EMPLOYEES

The third task completed has been to develop a study plan called the Orientation of New Employees. The main purpose is to establish a procedure indicating guidelines whereby new employees can best become acquainted in their new work with a minimum of stress involved in working in a new organization.

Presently very little is being done at MSFC in a formal manner to orient new employees about the ways of the Center. There is a one time meeting given by a representative of the Personnel Office, as presented by law in the guidelines of the Office of Personnel Management to describe the necessary work conditions and benefits of employees. This is not a true orientation since it is limited to a description of procedural matters relative to all NASA employees.

The Director and supervisors of the Personnel Office recognize the need for a completely developed orientation plan. But the re are differing opinions among them and among various supervisors and employees throughout MSFC about the merits and the type of orientation that would be necessary. Informal inquiries taken among several new employees also show a division of opinion about the nature and need of an orientation program for the recently hired.

Some feel that orientation efforts are a waste of time, others feel they are beneficial. Some believe orientation meetings should be brief and narrowly construed. Others feel that meetings should be very
comprehensive of the entire MSFC operations beyond one's own workplace. Others feel that there is no need to learn about areas unrelated to one's work. Some feel that an orientation schedule should occur within the first few days of a new person's hire or, at most, but a week long. Others feel that a proper orientation cannot be hurried or overloaded in a short time frame, but should be extended over several months. There are many concerns about the nature of an orientation program. Does it encourage better morale in the work force? Does it enhance the sense of mission? Does it build cooperative efforts?

The main argument seems to be whether a new employee needs to become acquainted with the general workings of the new organization, or whether the new employee needs to learn only about one's own particular job and only needs to learn about the wider ramifications of the organization as the person's contacts develop from the job site. Although this might be an individual matter, there seems to be a broad difference in approach between generalists and particularists based upon one's own research training and education. Some scientists and engineers want to work alone and without much to do with others. Some scientists and engineers wish to see how their own discipline relates with others and wish to ascertain how their own particular skills might be relevant to other parts of the whole center.

Management, on the other hand, has to determine what is best for the new employee and what kind of orientation would lead to the best satisfaction of the missions and purposes of the center as a whole. There are conflicting views here, too. Some personnel supervisors feel that the integrated nature of much of the work at NASA demands a general understanding of MSFC as a whole and of different projects and laboratories other than one's own specialty or assigned area. Other supervisors feel that an extensive orientation program would be too time consuming on the parts of both the participants and the presenters and would be too costly.

The orientation plan devised by this researcher incorporates many approaches to meet expected needs of new employees to help them adjust to work at MSFC. This plan attempts to schedule various experiences that will enable the new employee to understand not only the particular work situation in which the employee is assigned, but to show the overall mission and operations of the Marshall Space Flight Center. The study plan which follows here is devised to allow for maximum opportunities, but it may be pared down or condensed if so desired, as the plan moves up the ladder in the approval process.

## STUDY PLAN: ORIENTATION OF NEW EMPLOYEES

## PROBLEM:

Since 1968, when MSFC's first reduction in force took place, NASA-MSFC has been in a reduced hiring mode. As a result, new employee orientation during this time has been based primarily upon satisfying general information needs. Much of a new person's orientation is conducted informally at the work site by the employee's immediate supervisor. The Personnel Office, on the other hand, gives the new employee basic procedural information concerning OPM regulatory requirements and benefits. This is usually done on a one-to-one basis and later in a short general meeting for those brought on-board that week.

In the last year, NASA-MSFC hiring of new employees has increased. Yet, no plan has been developed to acquaint these new employees with the overall mission and purposes of NASA or with a proper perspective in seeing the relationship of one's own work to the whole. This has generated a need for a more comprehensive, in-depth orientation of new employees.

## OBJECTIVES:

The objectives of a plan for the orientation of new employees are:

1. To inform each employee about the basic requirements of adjusting to the new organization. This includes acquiring a knowledge of benefits, hospitalization, security, NASA-MSFC regulations and various other procedural and administrative matters.
2. To provide a method whereby the employee is thoroughly oriented to the work situation. Presently, the employee is acquainted with the work situation in an informal manner.
3. To provide a general understanding of the overall NASA mission and to show how the various organizations work together at MSFC to achieve national goals. Not only will the employee begin to sense the importance of NASA-MSFC as a whole, but also will learn how the new employee's particular work fits into a wider framework.

This study deals with these objectives and outlines the structure and content of a proposed orientation plan that will satisfy the needs
of the new employees and be beneficial to NASA.
PROPOSAL:
The "Orientation Plan for New Employees" will consist of three phases as follows:

1. Administrative Orientation -- procedures, responsibilities and rights, guidelines and benefits.
2. Work Site Orientation -- the particular tasks of the employee at the immediate place of work.
3. NASA-MSFC Orientation -- the overall perspective and acquaintanceship of the mission of MSFC as a whole.

PHASE ONE: ADMINISTRATIVE ORIENTATION

1. Personnel Staffing Specialist Meeting

The very first acquaintanceship of the new employee upon being hired will be a scheduled visit with an assigned Personnel Staffing Specialist (PSS) on an informal, one-to-one, personal basis. The most immediate concerns of the new employee, such as housing, locating main facilities of the Center, work hours, health plans, and pay procedures, will be discussed. The main attempt here is to give the new employee a feeling that he or she is welcome and has a place to find information immediately. A designated number of new employees will be assigned to each of the Personnel Staffing Specialists. Even new employees who have worked with other federal agencies would be expected to participate in this orientation. The PSS will "follow through" with each of the new employees over the course of the entire three phase orientation period. From time to time the PSS will contact the newcomer to inquire about the person's welfare at MSFC. The PSS should give the newcomer a contact phone number on the first visit. Also, the new employee will be given a packet, which includes necessary information about MSFC and Huntsville, employee unions, equal opportunity, and regulations, to keep and study.
2. Personnel Administrative General Meeting

At a designated time, preferably at the beginning of every two week pay-period in order to keep in phase with payroll records, the new employees hired up to that date will meet as a group for a general
orientation conducted, as presently done, by the Personnel Office. The necessary items scheduled for discussion will follow the "New Employee Orientation Checklist," MSFC Form 3688 (April, 1968), which is signed upon completion by both the employee and Staffing Officer and kept in the employee's 201 File. Discussion of any particular points will be encouraged and details will be provided as needed by the new employees.

## PHASE TWO: WORK SITE ORIENTATION

It is of utmost importance that the new employee learn to feel at home in his/her own work environment as soon as possible. The adjustment to a new situation will produce some frustrations, but it need not be stressful if proper considerations are given. The new employee will be under the authority of the immediate supervisor. This person should pay great attention to the psychological welfare of the newcomer. Introduction to all fellow employees, an opportunity to see others at work and to see who is responsible to whom, a review of work procedures, and an acquaintance with the location of key facilities and equipment are some of the concerns necessary to the well being of the new employee. This acquaintanceship period takes time, cannot be rushed, and should include some. free time to digest the new information. In essence, one is learning some of the informal ways of the organization.

During this time the Personnel Staffing Specialist will be in contact to review progress and handle problems. This is in line with regular personnel procedures which require a ninety day follow-up. Sometimes, several other orientation activities (of Phase Three, below) will be provided during the time the new employee is on the job in this early acquaintaceship period. Such activities might include the bus tour of MSFC or selected lectures and/or interviews by NASA specialists on particular topics. No additional paperwork is required by the supervisor other than the checklist now used routinely by the supervisor in reviewing the completion of special tasks in this part of the orientation. The intent of work site orientation is to keep procedures as natural and informal as possible.

## PHASE THREE: NASA-MSFC ORIENTATION

1. The major purpose of this phase is to assure that new employees are given opportunities to see their own particular work in relation to the whole of MSFC and of the entire NASA agency. It is important to the overall mission of NASA that integration and cooperation of efforts be understood by all employees not only better to achieve the goals of NASA, but to enhance esprit de corps. Also, an
understanding of the interrelationships of NASA's many projects often leads to a better feeling about one's own work and aids an individual in one's own personal and professional advancement.
2. Since the new employee will now be at work on the job site, this phase of the orientation must be scheduled so as to cause minimum disruption to the person's work schedule. It is proposed that this general orientation be staggered over a number of weeks to allow the new employee sufficient time to digest the new information and to prevent overburdening the offices of those who present materials by taking time out of their regular workload.
3. Administration of all aspects of Phase Three, including scheduling and monitoring, will be under the management of the Personnel Office.
4. All new employees should be involved together as a group in a one day common orientation. In the morning a bus tour should take the newly hired to various locations of MSFC. They should visit the main projects, test stands, certain laboratory facilities, museum mockups, and special projects, such as the Neutral Buoyancy Facility.

In the afternoon a general meeting should be arranged in the Tenth Floor Conference Room of Building 4200. At this time various selected Center management officials will introduce themselves and conduct a presentation on various important general MSFC topics. Such topics might include a general discussion of the Space Shuttle or of the importance of coordinating payloads.

A representative of the Office of the Director or one of the Executive Staff Offices should give an introductory address called the Center Overview, followed by other senior management speakers, who can lend an aura of importance to the new hiree's appointment into the MSFC workforce. This Conference Room meeting should include as many new hirees as possible at the same time.
5. Up to this time the orientation is the same for all types and categories of employees: scientists and engineers, administrative and professional, technicians, clerical and operational. At this point the orientation is divided into two modules. Each module applies to different categories of new personnel. The difference in orientation lies in the degree of depth and scope. The two modules are:

1) Scientists and Engineers, Business and Professional, and
2) Clerical (NCC 500).
6. New employees of both categories (modules) should visit certain basic projects, labs, or offices in common and as a large group. These include:
1) selected science and selected engineering laboratories;
2) key projects, such as Shuttle propulsion.
7. The module plan could be modified or compressed if the plan involves too much time of the employees and of the supervisors by cutting down on the staggered sessions and by having the new employees meet together in larger groups.

> Module One: Scientists, Engineering, Technicians, Business and Professionals

1. On a bi-weekly basis a certain number of new hirees will meet as a group to interface with each other, to share experiences, and occasionally to listen to guest speakers about special projects. The purpose of this meeting is to provide for some identifiable group loyalty and cohesion to disparate members of the MSFC work force in order to build up a sense of common solidarity across NASA as a whole, to build a sense of community and shared associations, and to provide for cross-fertilization of ideas. (Precedence for this has already been established by the NASA/ASEE Summer Faculty weekly "fellowship" meetings).
2. An opportunity for new employees to meet key personnel can be developed by inviting such people to these bi-weekly meetings to converse with the group, or to schedule specially arranged times when small numbers of hirees can visit in selected offices.
3. Rather fundamental kinds of programs, projects, and interfaces ought to be included in the itinerary. These might include:
a. Advanced Systems, planning layout

Examples: (1) Description of all future projects;
(2) Coal Gasification;
(3) Nuclear Waste Disposal;
(4) Solar Energy Use.
b. Space Shuttle operations

Examples: (1) Configuration Change Board of Solid Rocket

Boosters in Building 4610;
(2) Telecon Level II PRCB in Building 4202, Room 411 .
c. Payload and Spacelab Projects
(1) Overview.
d. Science and Engineering Laboratories

Examples: (1) Metallurgical analysis and use of electron microscopes;
(2) Cell separation in zero gravity
e. MSFC Administrative Support Examples: (1) Personnel Training;
(2) Procurement;
(3) Plans and Analysis Office;
(4) Overall administration.

## Module Two: Clerical

1. Since the new employees of this group will not have had extensive, formal scientific and engineering education, the presentations should be given in less complicated technical detail. However, the general nature of the various projects and tasks indicated in Module One should be of interest to this group. Thus, a schedule should be developed in the Personnel Office for this module following the same pattern.
2. Scheduling for both modules will be done in accordance with new hiring activity and with the concern to distribute presentations among many different people and offices.
3. A special session for Module Two would be a general, overall tour and explanation of a selected number of key installations. This would be different from the usual NASA tour. This would encompass such office functions as procurement, personnel training, computer services, and the public affairs office. Selected scientific laboratories would be included on the itinerary. This tour would be planned to encompass an entire morning.

## ORIENTATION PLAN RECOMMENDATIONS:

1. That a comprehensive, in-depth new employee orientation be implemented at MSFC.
2. That the administration and coordination of the Orientation of New Employees be the responsibility of the Personnel Office.
3. That an MMI be published identifying Center policy and assigning duties and responsibilities.

The Personnel Office of Marshall Space Flight Center assigned three management tasks to this researcher. All three were completed and the results were presented to the Personnel Director. A meeting of selected supervisors was called so that the researcher could orally present a status report on these topics. They expressed satisfaction with the projects and the work done. Final suggestions were made, although constant surveillance and assistance by the se same people had been given continuously during the entire time of investigation and writing.

This report has described the nature of the three tasks and the substance of their results. Since each of the tasks is a topic that is ongiong in nature, the final products will not be ready until all reviewers and authorities give approval. Undoubtedly some changes and revisions will be made. Nevertheless, the basic structure, the main outline and format, and the conceptual plans have been determined at this point.

There was no shortcut in the undertaking of these tasks. First, a thorough orientation and grounding into the nature of the scientific and engineering projects which make up the mission of NASA-MSFC and the nature of the administration and organization of MSFC had to be obtained by this researcher. Then, many different viewpoints and opinions had to be gathered and discussed about the specific nature of the tasks. Materials relevant to the tasks were read and analyzed, but these usually consisted of memos, letters, and similar works previously done. Since most of these materials are on file in various offices at MSFC, they do not form a descriptive bibliography or a list of references, and, consequently, are not stated in this report. Finally, each draft of each task had to be reviewed by an authority on the subject before the next revision could be written. Most often the suggestions of several people were incorporated into the next stage of the writings.

It is recommended that these three projects be continued immediately until their final approval is given by the Director of the Center. A Marshall Management Instruction (MMI) on the Student Volunteer Service Program should be completed in final form following the model presented here. The Recruitment Brochure should follow the general the me as worked upon to this point. Pictures should be selected and minimal text written. A comprehensive Orientation Plan for New Employees should be implemented at MSFC and an MMI should be written to identify Center policy and guidelines. All three of the se tasks should be administered by the Personnel Office of MSFC.

| Prepared By: | Richard Brown, Ph.D. |
| :---: | :---: |
| Ncademic Rank: | Associate Professor |
| University and Department | University of Alabama University, al Department of Mathematics |
| NASA/MSFC: | Data Systems Laboratory <br> Engineering Computers Division <br> Simulation Systems Branch |
| MSFC Counterpart: | Fiank L. Vinz |
| Date: | Auquest 1, 1980 |
| Contract No: | $\begin{aligned} & \text { NC'I'0)1-002-099 } \\ & \text { (Inivarsity of Alabama) } \end{aligned}$ |

ANALYSIS AND IMPROVEMENT OF CONTACT DYNAMICS SIMULATION USING THE SIX DEGREE OF FREEDOM MOTION SYSTEM

BY
Richard C. Brown, Ph.D.
Associate Professor of Mathematics University of Alabama, Tuscaloosa, Alabama


#### Abstract

The 6DOF motion system has been employed for closed loop simulation of the contact dynamics involved in the docking of two satellites (for example TRS and Skylab). Under most initial conditions however the system gives inaccurate results. Typically in the simulation if docking fails to occur the rebound velocity of the TRS exceeds its approach velocity, thus generating an artificial and unacceptable gain of momentum in the system.

It has long been well known that the presence of a time lag can generate instability in closed loop systems. To see if the time lag (about $200 \mathrm{ms}$. ) in the 6 DOF system can satisfactorily account for its errors, a simple computer mode of the system incorporating this lag and other revelant parameters has been implemented. The model in fact does predict all the observed rebound errors with reasonable precision.

Also, since it is probable that the time delay can only be reduced but not eliminated, various compensation schemes have been devised. Two give adequate correction for delays under 100 ms for a non-preloaded probe. More powerful schemes may be adequate for the present 200 ms . environment and the presence of a preload.

Our study relies heavily on the mathematical theory of second order differential equations with retarded arguments and standard tools of numerical analysis.


## ACKNOWLEDGEMENT

I want to express my thanks and appreciation to the members of the Simulation Systems Branch - especially F. L. Vinz, David Allen, J. S. Spear and Bert Gover for many helpful discussions concerning the 6 DOF system.

I am also grateful to Marion Kent and Professor B. F. Barfield for their effective and dedicated efforts in the administration of the NASA/ASEE Faculty Fellowship program.

NOMENCLATURE

| r |  | Time lag of the 6 DOF system |
| :---: | :---: | :---: |
| $\Delta t$ | - | Sample frame time of the $A D$ converter between the sensor and the EAI 8900 |
| $F(t)$ | - | Sensor force signal |
| $\Delta \mathrm{F}$ | - | Change in force signal in time step |
| 7 | - | Probe compression |
| $\bar{\eta}$ | - | Maximum probe compression |
| K | - | Hooke's constant of probe |
| M | - | Mass of probe |
| e | - | Base of the natural logarithmns |
| $v_{*}$ | - | Initial velocity of probe |
| $v_{f}$ | - | Observed rebound velocity of probe |
| $v(t)$ | - | Velocity of probe |
| $a(t)$ | - | Acceleration of probe |
| IVP | - | Initial value problem |
| $\omega_{\text {c }}{ }^{2}$ | - | K/M |
| $\mathrm{F}_{\mathrm{p}}$ | - | Preload of probe |
| $x$. | - | $\eta+F_{p} / k$ |
| $x_{T}$ | - | $x(t-\tau)$ <br> Convolution, $f=g:=\int_{0}^{+} f(t-1) g(s) d s=\int_{0}^{t} f(s) g(t-s) d s$ |
| $\mathrm{F}_{t}$ | - | Force threshold |
| $\lambda(t)[a, t]$ | - | Characteristic function based on the interval [a,b] |
| h | - | Step in a digital simulation |

$$
\begin{array}{ll}
M & \text { Step parameter ; } T=n, \mu \\
O\left(h^{*}\right) & -\quad \text { Error proportional to } h^{k} \\
v_{n} & -v(n h) \\
q_{i n} & -a(n h) \\
v_{0} & -\quad \text { Compensated version of } v \\
x_{c} & -\quad \text { Compensated version of } x
\end{array}
$$

## I INTRODUCTION

The LINK Six-Degree-of-Freedom (6 DOF) Motion System has been employed by the Marshall Space Flight Center for the physical simulation of the contact forces and dynamics in the docking of two satellites - for example the Teleoperator Retrieval System (TRS) and Skylab. A similar device was developed at the Johnson Space Center in Houston for contact dynamics simulation prior to the Apollo-Soyuz mission of 1975.

In this section we give an informal description of the system and its performance. Some illustrations, photographs, and values of some key parameters may be found in the Appendix. See in particular the block diagram, fig. 10.

A docking probe assembly is placed on the platform of the LINK 6 DOF Motion System. The platform is free to move with any general motion within its limits of travel. Above the probe is a docking collar or other target which is fixed. The software of the EAI 8900 hybrid computer is capable of simulating a set of initial conditions which might include velocities, accelerations, and rotations preliminary to docking.

Since, as mentioned above, the target is stationary, this information is calculated relative to a coordinate system fixed in the target. The initial conditions are translated by the CAT digital computer into commands to change lengths of the six legs supporting the 6 DOF platform. These commands are executed by the hydraulic actuators in the legs causing the 6 DOF system and the probe to perform motions appropriate to the initial conditions.

When the probe is in contact with the target, sensors, also placed on the table, record the forces and torques encountered by the probe. In the open loop mode this information is recorded on strip charts for later use. If desired software can be developed so that the probe simulates dynamic interaction with the target. This option has so far been rejected since the resulting motion would not be generated by the forces actually encountered in the contact phase. A more useful choice is closed loop simulation where the forces are used in a feedback loop through the EAI 8900 and the CAI computers to control the motion of the probe itself. Basically two integrations of the hybrid subsystem give velocity and position functions of the TRS relative to the target. As in the case of the initial conditions the position function is converted into leg length commands. The result is that the probe is supposed to move as if it was freely responding to its own contact dynamics. This mode when perfected should be a valuable tool in the testing and design of docking hardware.

In preliminary testing of the system, the closed loop simulation has been radically simplified so that orobe-target interaction is "head on;" only (vertical) translational motion is accepted. Also no thrusts firings of the chase vehicle are allowed. Table l presents results of some of these tests.

Obviously apart from the unacceptable level of noise, there is an artificial and unwarranted gain of momentum in the system. The analysis and attempted resolution of this problem is the central concern of the present study.

We close this section with a brief survey of the rest of the paper. In Section II a simple mathematical model based on the time delay inherent in the system is develoned. It is also shown how the model can be modified so that it is "noisy." Section III developes several methods to compensate for the delay. The last two sections contain respectively data and graphs illustrating our results and recommendations for further action.

It has long been known that the presence of time delay can affect the stability of a closed loop feedback system - see for example.

Minorsky [10], [11] or Callendar, Hartree and Porter [4]. It has been verified that a time delay defined as the time it takes for the legs to complete a response to a signal from the force sensor, exists in the 6 DOF system.

This delay (let us call it "r") is approximately 200 ms ; the various contributions to it by hardware subsystems are shown in the block diagram Fig. 10.

A simple physical argument shows how $r$ might explain the high rebound velocities of table l. Suppose at time $t$ the probe has reached maximum compression $\bar{\eta}$ and should turn around. The command to do this is executedt+r later; in the meantime, the probe has continued to move and overshoth . Therefore during a time period of length $\boldsymbol{\tau}$ additional force has been communicated to the software resulting in a gain in energy. If the probe did not detach, $\boldsymbol{\tau}$ sec. after reaching its rest position the overshot would be repeated and one might expect a sequence of oscillations of increasing amplitude.
2.1 A Mathematical Model - Let us now try to make the above argument mathematically precise. Suppose in the interval $[t, t+\Delta t]$ a change in the force signal $\Delta F$ is sampled. When the EAI 8900 computes a change of compression $\Delta \eta$ corresponding to $\Delta F$, it does this by digitally solving the equation(l)and computing the solution at $t+\Delta t$

$$
\begin{equation*}
M \ddot{\eta}=F(t) \tag{2.1}
\end{equation*}
$$

Now $\eta(t+\Delta t)$ should equal $n(t)+\Delta k$. However $\Delta h$ is not executed until the actuators have finished their response at t. $T^{\circ}$. Thus $n(t+r)$ corresponds to $F(t)$ - $\eta(t)$ equivalently corresponds to $F(t-r)$ It follows - ( $\tau$ includes the sample interval or "frame time" $\Delta t$ )that the system behave as if it were solving the equation

$$
\begin{equation*}
M \ddot{u}=F(t-r) \tag{2.2}
\end{equation*}
$$

Thus we could model the effect of delay by using (2.2) instead of (2.1) in a computer simulation.

In case the probe is an ideal spring $F=-k \eta$ and (2.2) becomes
(2.2) and (2.3) are examples of retarded differential equations. Because of their importance to control theory applications, such equations have been intensively studied [3 ], [6], [7] in recent years. The resulting theory has many similarities with the classical theory of differential equations but also important differences. In particular it can be shown that if

$$
T<2 / \omega_{0} e
$$

then "practically all" solutions oscillate with exponentially increasing amplitude (R. D. Driver, private communication, also cf. Driver [6]).

Also unlike differential equations, a correctly formulated initial value problem (IVP) for (2.2) - (2.3) requires initial conditions on an initial interval rather than just at a point. For example, consider the first $T$ seconds after contact: The probe continues to compress oblivious of contact since the legs have not yet moved in response to $F(t)$. Thus

$$
\begin{equation*}
\eta(t)=N_{0} t \tag{2.4}
\end{equation*}
$$

$$
\dot{n}(t)=N_{0}, \quad 0 \leq t \leq T .
$$

These replace the point conditions of a differential equation IV.
2.2 A Few Refinements - Essentially equations (2.1) and (2.2) Will constitute the mathematical model of the 6 DOF system in our report. Before solving it we discuss some possible improvements (only the most significant of which have been implemented).
2.2.1 Probe Nonlinearity - As is shown in Fig. 2, the probe is not an ideal spring. It is both nonlinear and preloaded. The nonlinearity can be modeled by an equation of the form

$$
M \ddot{n}+K(n(t-T))=0
$$

For the graph of $k(u)$ see Fig. . To approximate $K(n)$ one might choose a piecewise linear function, a least squares parabolic fit, or a spline. None of these choices is difficult to implement on the computer. However, since $K(n)$ is nearly
constant in the neighborhood of $50-60 \mathrm{lbs} / \mathrm{ft}$, its variation is probably not a source of major error, we will omit this feature of the spring from our model.

Of much greater significance is the existence of a preload ( $F_{p}$ ) of approximately 69 lbs. By energy considerations a preloaded spring is equivalent to a compressed longer spring having the same Hook's constant; the fictitious compression is $F_{p} / K$. This suggests that we define $x(t):=n(t)+F_{p} / k$ and solve

$$
\begin{align*}
& \dot{x}+\omega_{0}^{2} x_{T}=0 \\
& x(t)=x_{0}+v_{0} t \\
& \dot{x}(t)=N_{0} \quad 0 \leq t \leq T  \tag{2.5}\\
& x_{0}:=F_{P} / K \\
& \omega_{0}^{2}:=K / M \\
& x_{T}:=x(t-r)
\end{align*}
$$

It will turn out that the preload magnifies the effect of a pure delay.

In this report we will also consider an ideal probe with no preload and $K=1200$ lbs/ft., since this value has been used in a recent analog simulation [1] of the effect of delays on spring motion in the 6-DOF system.
-2.2.2 Filters - As the block diagram Fig. 10 shows there are several filters in the system. They contribute to the total delay and probably somewhat distort the signal. In [!] it has been shown that the addition of a low pass filter with transfer function $6 / \$+6$ gives an additional destabilizing effect.

In a digital simulation the presence of such a filter changes (2.3) to the equation

$$
M \ddot{\eta}+k e(t) * n(t-T)=0
$$

$$
\begin{equation*}
e(t):=1-e^{c t} . \tag{2.6}
\end{equation*}
$$

In general the presence of several filters with constants $c_{1} \ldots c_{n}$ gives the equation
(2.5) $\quad M \ddot{n}+K l_{1} \cdot e_{2} * \cdots e_{6} * n(t-T)$
where

$$
e_{i}:=1-e^{-c_{i} t}
$$

The presence of convolutions can be handled by nested integration loops in a digital simulation. Since the parameters of the filters actually in the system are unknown we ignore their effect except in sofar as they contribute to the delay. Our results compared with tests suggest that this is not a crucial omission.
2.2.3 Force Threshold - In addition to the filters a force threshold ( $\mathrm{F}_{\mathbf{t}}$ ) is used to reduce noise. This means that $\mathrm{F}(\mathrm{t})$ is accepted if and only if its absolute value exceeds $F_{\neq}$i.e.,

$$
F(t)= \begin{cases}F(t) & \text { if }|F(t)| \geqslant F_{t} \\ 0 & \text { otherwise. }\end{cases}
$$

Thus the IVP (2.5) becomes

$$
M \ddot{x}+k \lambda(x)\left[E_{ \pm} / k, \infty\right) x=0
$$

$$
\begin{align*}
& x(t)=w_{0} t+F_{p} / k  \tag{2.6}\\
& \dot{x}(t)=v_{0}, 0 \leq t \leq r
\end{align*}
$$

Fortunately (2.6) is equivalent to a simpler IV P where $\mathrm{F}_{\mathrm{i}}$ plays a role like the preload. The probe must travel $\mathrm{F}_{\mathrm{t}} / \mathrm{K}$ before force due to compression is "accepted." The elapsed time is $\bar{E}=\mathrm{F}_{\mathrm{t}} / K v_{0}$ : If we regard $\bar{E}$ as time zero, (2.6) becomes

$$
\begin{aligned}
& \ddot{x}+w_{0}^{2} x_{T}=0 \\
& x(t)=x_{0}+w_{0} t \\
& \dot{x}(t)=v_{0} \\
& x_{0}=\left(F_{p}+F_{t}\right) / k \\
& n(t)=x(t)-F_{0} / k
\end{aligned}
$$

(2.7)

Note that the system should be very sensitive to $F_{t}$. If $F_{t}$ is large and $v_{0}$ is small the only force affecting the probe might be that of the bottoming spring. For example if $K=60, F_{t}=50$ $x_{0}$ • 14.4 in which is greater than the maximum compression of the probe. In the absence of delay, even threshold of 30 lbs. would produce a severe distortion. When delay is present its destablizing effects are greatly magnified. In the case of a spring with $K=1200$ and no preload the effects thought not so severe are still noticable.

The presence of a force threshold also causes some ambiguity in time measuresments during the simulation. Suppose $t$ " is the time during rebound when $K \eta=F_{t}$. Since $F(t)$ is measured directly on the strip chart the simulation period is $E-F_{t} / K v_{0}^{\prime}, t_{t}+F_{t} / K v_{0} \mathcal{I}$. The period in which the forces are producing a response in the legs is [ $T$, $t_{F}+r$ ]. Currently the exit velocity is measured at $\boldsymbol{t}_{\mathbf{F}}+\boldsymbol{r}$, time $\mathbf{t}_{\mathbf{F}}+\boldsymbol{T}+\mathrm{F}_{\mathbf{t}} / / K V_{\boldsymbol{o}}$ from contact. To avoid confusion one needs to keep these various times clearly in mind especially when $\mathrm{F}_{\mathrm{t}}$ is high and/or $\mathrm{K}, \mathrm{V}_{\mathrm{o}}$ is low.
2.2.4 Variable Delay - The IVP (2.6) is a continuous realization $\overline{O f}$ a discmetry sampled system. The AD and DA converters (1), (2) (see Fig. 10 ) sample the signal at discrete intervals $\Delta t$ approximately $31-35 \mathrm{~ms}$. If all the converters where synchronous little would be lost since we in fact solve the I V P in steps on the computer. However, the converters (3), (4) sample at intervals of 25 ms . This means that periodically (3) samples the same signal twice, producing an additional 25 ms delay. Since they do not start in synchrony (they start where they stopped) the effect varies from run to run. We have put a "noisy" delay into the model. The results suggest that the asynchronization explains some of the noise in the data.
2.2.5 Noise from the table - Since the force sensors are in the table, they may feel some of the inertial forces in the table itself. Again forces due to this source are executed seconds later. This suggest that we consider the neutral to second order equation.

$$
\begin{equation*}
\ddot{x}-E \frac{M_{1}}{m} \ddot{x}_{r}+\omega_{0}^{2} x=0 \tag{2.8}
\end{equation*}
$$

with the same initial conditions of (2.7) where $E$ is a noise parameter given by a random number generator and $M_{1}$ is the mass of the table + probe $=323.5$. Preliminary tests (see Apps.ulx ) show that this complication is indeed a potent sour ce of noise.
2. 3 Solving the Model
2.3.1 Formulas in closed form - Suppose we linearise (2.7) by taking the first two terms of the Taylor expansion, we obtain

$$
\ddot{x}+\omega_{0}^{2} x-\omega_{0}^{2} \tau \dot{x}=0
$$

$$
\begin{equation*}
x(0)=x_{0} \tag{2.9}
\end{equation*}
$$

$$
\dot{x}(0)=v_{0}
$$

If $\omega_{0}^{2} r \ll 1$ and $x_{0}=0$ (no preload, no threshold) an approximate solution to (2.9) is

$$
x(t)=e^{\pi / 2 \omega^{r} t} \sin \omega_{0} t
$$

Hence

$$
\begin{equation*}
v_{F}=v_{0} e^{H / 2 \omega_{v} T} \tag{2.10}
\end{equation*}
$$

Neither of these formulas are accurate for much beyond a few seconds. However, (2.9) gives an error of under $3 \%$ for $7 \leq .2$. If there is a force threshold the formula

$$
\begin{equation*}
v_{F}=V_{0} e^{\pi / 2 \omega_{0} T}\left(V_{0}+F_{t} T / 2 m\right) \tag{2.11}
\end{equation*}
$$

yields an error < 15\% provided $F_{t}<120$ lbs.
An exact closed form solution to (2.7) or 2.8 ) can be given by an inductive procedure known as the "Method of Steps." Suppose of and $x$ are known on the interval $[(k-1) r, k r]$. Then setting

$$
t=k \tau+s
$$

we have for $k T \leq t \leq(k+1)$ ?

$$
\begin{align*}
& v(t)=v(k r)-w_{0}^{2} \int_{k T}^{t} x(s) d s  \tag{2.12}\\
& x(t)=x(k r)+\int_{k T}^{t} v(u) d u .
\end{align*}
$$

Since the initial functions are simple the solution may be calculated in closed form, step by step. Thus for $\boldsymbol{T} \leqslant t \leq 2 T$

$$
\begin{aligned}
& v(t)=v_{0}+1 / 2 w_{0}^{2} v_{0}(t-T)^{2} \\
& x(t)=x_{0}+v_{0} r+1 / 3!w_{0}^{2} v_{0}(t-T)^{3}
\end{aligned}
$$

etc.
2.3.2 Numerical Integration- The Method of Steps gives a solution in the form of an infinite series. Since the result tells us nothing about the properties of the solution, nothing is lost if we numercially integrate (2.7). The following algorithm which uses the corrected trapezoid rule ( L ]. 74 ) has $\theta\left(H^{\circ}\right)$ discretization error and has been implemented on the sis sigma 5 computer and the writers TI 59.

Input $\quad K, V_{0}, X_{0}, F_{t}, F_{p}, M, K, T, N$

$$
\begin{aligned}
& H:=\tau / N \\
& x_{0} \leftarrow x_{0}+\left(F_{+}+F_{p}\right) / k
\end{aligned}
$$

For $\quad i=1$ to $N+1$

$$
\begin{aligned}
& x_{i}=x_{i-1}+v_{0} H \\
& v_{i}=v_{0}
\end{aligned}
$$

continue

## For $i=k+2$ until aatesfied:

$v_{i+1}=v_{i}-\omega_{0}^{2}\left(H / 2\left(x_{i+1}-n+x_{i-n}\right)+H^{2} / 12\left(V_{i-n}-V_{i+1-n}\right)\right.$
Write $V_{i r}$

$$
\begin{aligned}
& x_{i+1}=x_{i}+H / 2\left(v_{i}+v_{i+1}\right)-\omega_{\frac{0}{12}}^{2} H^{2}\left(x_{i+1-n}-x_{i-4}\right) \\
& n_{i+1}=x_{i+1}-F_{i} / k
\end{aligned}
$$

Write $n_{i+1}$

## Continue

2.4 Results - Figures (2)-(7) give the results of several runs with $\overline{d i f f e r e n t ~ i n i t i a l ~ c o n d i t i o n s ~ a n d ~ p a r a m e t e r s . ~ N o t e ~ t h a t ~ t h e ~}$ results are consistent with tests of the 6 DOF system and also agree almost perfectly with the results of the analog simulation [1].

It is hoped that most of the delay can be eliminated by suitable modifications in the hardware. Nervertheless even delays as low as 50 ms can cause noticeable instability. Therefore in this section we develop several algorithms that in some cases can produce near normal performance in spite of delay. Test results for some of the methods described below are given in section IV.
3.1 Taylor Approximations

Since

$$
x_{T}=x-T \dot{x}+\frac{T^{2}}{2} \ddot{x}+\cdots+(-1)^{n} \frac{T^{n}}{n!} x^{(n)}(1), t-\tau<1<t
$$

we can truncate the series and solve for an approximation $x_{c}$ to $x-i . e$,

$$
x_{c} \approx x_{T}+T \dot{x}-T^{2} \ddot{x}+\cdots(-1)^{n-1} \frac{T^{n-1}}{(n-1)!} x^{(n-1)}
$$

with error $\theta\left(r^{\prime \prime}, x^{n}\right)$. We then solve.

$$
\ddot{x}+w_{0}^{2} x_{c}=0
$$

$$
\begin{equation*}
x(6)=x_{0} \tag{3.1}
\end{equation*}
$$

$$
\eta=x-F_{p} / k
$$

instead of (2.1).
3.1.1 Range Gutta - In theory (3.1) can be solved for arbitrarily high n by the Runge Kutta algorithm without knowledge of high derivatives. For example if $n=4$ (3.1) can be converted to the

$$
\begin{align*}
& \left(\begin{array}{l}
x_{1} \\
x_{2} \\
x_{3}
\end{array}\right)^{\prime}=\left(\begin{array}{lll}
0 & 1 & 0 \\
0 & 0 & 1 \\
E_{5} & c_{1} & c_{2}
\end{array}\right)\left(\begin{array}{l}
x_{1} \\
x_{2} \\
x_{3}
\end{array}\right)  \tag{3.2}\\
& \left(\begin{array}{l}
x_{1} \\
x_{2} \\
x_{3}
\end{array}\right)(0)=\left(\begin{array}{l}
x_{0} \\
v_{0} \\
0
\end{array}\right)
\end{align*}
$$

where
$E_{T}:=$ backwards shift

$$
\begin{aligned}
& c_{1}:=-3!/ T^{2} \\
& c_{2}:=\frac{3}{T}\left(\frac{2}{T^{2}} \omega_{0}^{2}+1\right) \\
& x_{i}:=x^{(i-1)} \quad i=1,2,3
\end{aligned}
$$

Writing (3.2) in the abbreviated form

$$
x^{\prime}=A x
$$

$$
x(0)=x_{0}
$$

and applying the Range Gutta algorithm.
we obtain

$$
x_{i+1}=x_{i}+\mu\left(\frac{k_{1} i}{6}+\frac{k_{2} i}{3}+\frac{k_{2 i}}{3}+\frac{k_{4} i}{6}\right)
$$

where

$$
\begin{aligned}
& k_{1 i}=X_{i}+H / 2 A Y_{i} \\
& k_{2 i}=X_{i}+H / 2 A k_{1 i} \\
& K_{3 i}=X_{i}+H / 2 A k_{2 i} \\
& k_{4 i}=X_{i}+H A k_{3 i}
\end{aligned}
$$

The difference between the above version of Range Kutta and the conventional one for differential equations is that (because of the shift operation) $k_{1 i}$ requires $y_{i-\ldots}$. $k_{\mathbf{z}_{1}}$ requires $k_{1,-m}$ etc. However, the effect this deviation has on the stability and accuracy of the algorithm is unknown and will require additional investigation. Higher order versions of Range Gutta are also possible. They can be used alone or combined with equivalent Adams predictor corrector methods. If they work quite high order Taylor approximations can be easily generated.
3.1.2 The Algorithm "FIX*T4" - The following direct approach to (3.1) has been implemented and found to work well. We write.

$$
\begin{aligned}
& x_{T}=x-T \dot{x}+\frac{T^{2}}{2} \ddot{x}-\frac{T^{3}}{6} \dddot{x}-\frac{T^{4} x^{10}}{24} \\
& \dot{x}_{T}=\dot{x}-T \ddot{x}+T^{2} / 2 \ddot{x}-T^{3} / 6 x^{(V V)} \\
& \ddot{x}_{T}=\ddot{x}-T \dddot{x}+\frac{T^{2}}{2} x^{1 v} \quad I I T-12
\end{aligned}
$$

$$
\begin{aligned}
& \dddot{x}_{T}=\ddot{x}-T x^{(i v)} \\
& x_{4}^{i v}=x^{i v}
\end{aligned}
$$

By back substitution the present values of the derivatives can be eliminated. We finally obtain

$$
x_{c}=x_{T}+r \dot{x}_{T}+T^{2} \ddot{x}_{T}+\frac{I^{3}}{6} \ddot{x}_{T}+\frac{T^{9}}{24} x_{r}^{(v)}+\theta\left(r^{5}, x^{(5)}\right)
$$

We can truncate $t h$ series obtaining second, third on fourth archer connections

The 4 th order version, however, is as easy to implement as the others and more accurate
"FIX*T4" For $i=1$ to mn

$$
\begin{aligned}
v_{i} & =v_{0} \\
x_{i} & =i+v_{0}+x_{0} \\
a_{i} & =0
\end{aligned}
$$

continue
For. $=m+1$ until satisfied

$$
\begin{align*}
v_{i+1}= & v_{i}-\omega_{0}^{2}\left\{\frac{H}{2}\left(x_{i-m}+x_{i+1-m}\right)+H^{2} / 12\left(v_{i-m}-v_{i+1-m}\right)\right. \\
& +\left(x_{i+1-m}-x_{i-m}\right)+r^{2}\left(v_{i+1-m}-v_{i-m}\right) \\
& \left.+T^{3} / 6\left(a_{i+1-m}-a_{i-m}\right)\right\}  \tag{3.3}\\
a_{i+1}= & -\omega_{0}^{2} x_{i+1-m} \\
a_{i+1}^{c}= & -w_{0}^{2}\left(x_{i+1-m}+v_{i+1-m}\right) \\
x_{i+1}= & x_{i}+H / 2\left(v_{i}+v_{i+1}\right)+H_{12}^{2}\left(a_{i}^{c}-a_{i+1}^{c}\right)
\end{align*}
$$

Continue.
The algorithm is $\theta\left(h^{9} ; x^{i v}\right)$.
Remark - The corrected trapezoid rule and the two estimates of $\ddot{x}_{i} .-q_{i}, q_{i}$ are necessary to maintain accuracy.

Warning - In a 6 DOF simulation

$$
a_{i+1}=a_{i+1}^{c}=F_{i+1}=F(i+1 \Delta t)
$$

(The frame time $\Delta t$ is the step). Also all delayed terms $x_{r},{ }^{\circ} x_{r}$ etc. are the present values of $x, \dot{x}$ etc., generated in the hybrid (recall that the backward shift introduced into the algorithm simulates the loop circuit time in the 6 DOF system) with this in mind, we write the algorithm in a form suitable for implementation on the 6 DOF. Let $v_{i+1}, x_{i+1}$ be the velocity and position produced by the hybrid from $F$ by integration. Let $V_{i+1}^{c} X_{i+1}^{c}$ be the compensated versions. Then the main loop in (3.3) looks like

For $n=m+1$ until satisfied:

$$
\begin{align*}
V_{i+1}= & \text { Integration of } \mathrm{F} \\
x_{i+1}= & \text { Integration of } \mathrm{v} \\
V_{i+1}^{c}= & V_{i}-\omega_{0}^{2}\left\{H / 2\left(x_{i}+x_{i+1}\right)+H^{2} / 12\left(V_{i}-V_{i+1}\right)+T\left(x_{i+1}-x_{i}\right)\right. \\
& \left.+T^{2} / 2\left(V_{i+1}-V_{i}\right)+T^{3} / 6\left(F_{i+1}-F_{i}\right)\right\} \tag{3.4}
\end{align*}
$$

$$
\begin{aligned}
& x_{i+1}^{c}=x_{i}+H / 2\left(V_{i}+V_{i+1}^{c}\right)+H / 12\left(F_{i}-F_{i+1}\right) \\
& V_{i} \leftarrow V_{i+1}^{c} \quad x_{i} \leftarrow x_{i+1}^{c}
\end{aligned}
$$

(i.e., these values replace $v_{i}, x_{i}$ in the next iteration of the loop)

## Continue

In the remainder of the paper all algorithms will be written the form (3.3) instead of (3.4) and it will be understood that terms $x_{r}$ and their digital form $x_{i-m}$ etc. refer to present not past values of variables calculated in the hybrid.
3.2 Polynomial Extrapolation - Porter and Stoneman [12] (also see [13] have proposed the use of Newton-Gregory interpolating polynomials to compute $x_{c}$. Specifically they fit a polynomial to $x_{n+1-m}, x_{n+1}-(m+1), \ldots x_{n+1}-(m+4)$ and extrapolate to compute $x_{c}=$ $x_{m}$. Thus by the Newton-Gregory backwards interpolation formula

$$
\begin{align*}
x_{n+1}^{c} ः & \left.x_{n+1-m}+T \nabla x_{n+1-m} T \frac{T( }{2!} H^{2}+T\right)  \tag{3.5}\\
& \nabla^{2} x_{n+1-m} \\
& +\frac{\nabla^{3} x_{n+1-m}}{3!H^{3}} r(T+H)(2 H+T) \cdots \quad \theta\left(H^{6+1} x^{6+1}\right)
\end{align*}
$$

where

$$
\begin{aligned}
\nabla x_{n+1}-m & =x_{n+1-m}-x_{n+1-(m+1)} \\
\nabla^{2} x_{n+1}-m & =\nabla x_{n+1-m}-\nabla_{n+1-(m+1)} \text { etc }
\end{aligned}
$$

No higher than a cubic should be taken. The procedure might begin with a 4 th order Taylor method to compute enough $x$ values beyond $[0, r]$ to fit the cubic. For example, the first cubic could be passed through $x_{m+1} \ldots x_{m+4}$, or through $x_{m} \ldots x_{m+1}$ etc. The best choice would be found experimentally. Since interpolating polynomials are extremely sensitive to noise, we suggest substituting a least squares polynomial of low degree based on as many previous points as possible. If we write the polynomial in the form.

$$
P_{v}(x)=a_{0}+a_{1} x+\cdots a_{v} x^{v}
$$

and set $\bar{a}=\left(q_{0}, a_{1}, \ldots a_{v}\right), \bar{a}$ is the solution of the system.
(3.6) $C \bar{a}=\bar{r}$
where $C$ is the $v+l x v+1$ matrix with general component

$$
c_{i j}=\sum_{\rho-m}^{e:} t_{e}^{i+j-2}
$$

where

$$
\bar{r}=\left(\begin{array}{l}
\sum_{p=-1}^{p} x_{l} \\
\sum_{p=1}^{p} t_{l} x_{l} \\
\vdots \\
\sum_{p=1}^{p} t_{l} x_{p}
\end{array}\right) \quad 1 \quad t_{l}:=l H .
$$

and p is a preassigned integer.
Remark - matrix system (3.6) becomes poorly conditioned for large $V$; most machines will give meaningless results at $V=7$ or 8. We recommend double precision arithmetic and $v \leq 4$.

Because velocities are available at the same time as the position function, anther choice would be to use this information and require that the polynomial pass through $x_{n} \ldots, v_{n-m} . .$. etc. For formulas for this type of interpolating polynomial see [5] p62ff. The error is $\theta\left(H^{+1}\right)$ where $m$ is the degree of the polynomial (egg., 5 if three points are used.)

## Remarks

(1) Although error terms for polynomial extrapolation schemes are similar to Taylor methods, their performance may be superior (particularly for the least squares version) since more information is used and sensitivity to noise is reduced.
(2) Warning - Neither method (or any other known to the author) will make the system stable for more than a short time. The high derivatives present in the error terms ruin the approximation as
3. 3 Transfer Function Approximation - The Laplace transform of $X_{T}$ is $\boldsymbol{e}^{T 1} \mathrm{IX}$. Therefore, it is tempting to find an approximation $P(s)$ to $e^{\text {rs }}$ and define

$$
\begin{aligned}
x_{c} \geqslant x & =\mathcal{L}^{-1}\left(P(1) e^{-T s} \mathcal{L} x\right) \\
& =\mathcal{L}^{-1}(P(1)) * x_{T}
\end{aligned}
$$

Thus we solve.
(3.7)

$$
\ddot{x}=-w_{0}^{2} f^{-1}(P(Q)) \star x_{T}
$$

The advantage of this method is that if $P(s)$ is properly chosen, (3.7) can be solved on an analog computer much more easily than it can be on a digital machine. This technique provided $P(s)$ is "close" to $e^{\top s}$ can be shown to be as good as polynomial extrapolation (see [13] p32ff).
Some Approximations for $e^{r_{1}}$
3.3.1 Chebyshev polynomials lead to approximations which are much more accurate than Taylor polynomial approximations of the same degree.

Thus
(3.8) $\quad e^{\text {TA }} \approx .994792+.997396 T 1+.541687(T 1)^{2}+.177083(T 9)^{3}$

$$
\begin{equation*}
=\quad .994571 \rightarrow .947308 \mathrm{Ts}+.542991(71)^{2}+.177349(\mathrm{T1})^{3} \tag{3.9}
\end{equation*}
$$

$$
\approx .994571+.99730872+.542991(71)^{2}+.177347(\mathrm{r})
$$

See Atkinson [2] Ch 4
The error of each of these approximations is less than $6 \times 10^{-1}$ if $\mid$ 个s/<1

Tests based on (3.10) have shown that it is generally superior to a fourth order Taylor approximation when there is no threshold and no preload. For a description of the resulting algorithm "FIX*C" see p below.
3.2.2 Padé Approximations - It is possible to approximate accurately by rational functions of equal degree in the numerator and denominator. The functions are called Fade approximations. Inverting them offers the advantage that numerical or analog differentation is avoided. Also fade approximations are much more accurate than Taylor series approximations of equal degree. For example the end degree Fade

coincides with the first five terms of the Taylor series of ${ }^{75}$ the sixth term is $(T s)^{s} / 144$ which is close to $(r s)^{s} / 120$. Hence this approximation is $\theta(r s)^{s}$.

The third degree

$$
\begin{equation*}
\frac{\left.1+r / 2+\frac{(r 5}{10}\right)^{2}+\frac{(r 5)^{3}}{120}}{1-r 5 / 2+\frac{(r 5)^{2}}{10}-\frac{(r 5)^{3}}{120}} \text { of }[9] \text { e145 st } \tag{3.12}
\end{equation*}
$$

is even better, it concides with the first seven terms of the Taylor seri.es. Hence, if our previous discussion is correct these approximations should be similar to cubic and six degree polynomial extrapolation. Pade approximation can be further improved using Chebychev polynomials (See Ralson [14] p 296) Thus

$$
\begin{equation*}
e^{T 3}=\frac{2353+1176 T s+192(T S)^{2}}{2353-1175 T S+192(T S)^{2}} \tag{3.13}
\end{equation*}
$$

The error of the approximation for $\mid$ isl <l is $\theta 10^{-4}$
Inversion of fade approximations and appeal to (3.7) result in complicated integro-differential equations. For example the second degree Fade yields

$$
y^{\prime \prime}+\omega_{0}^{2} y_{T}+\frac{12}{T}\left(e^{\frac{3 t}{T}} \sin \frac{3}{5} t+e^{\frac{1+}{r}} \cos \frac{3 t}{T}\right)-y_{T}=0
$$

Unfortunately all fade approximations for have poles in the half plane. We conjecture that because of the short contact time this source of instability will not be of major imporance.
3.4 Richardson Extrapolation - We finally mention a technique which can "boost" the degree of approximation when used in conjunction with the previous techniques; that is if the error term of a given method is $O_{\left(H^{\circ}\right)}$ we will obtain by modifying the method an error term $\sigma\left(H^{n+1}\right)$

Suppose

$$
\begin{equation*}
x_{c}(H)=x+C\left(H^{*}\right)+O\left(H^{(H+1)}\right) \tag{3.14}
\end{equation*}
$$

Having the step gives
(3.15) $\quad X_{C}\left(\frac{H}{2}\right)=x+\frac{C}{2 m}\left(H^{n}\right)+O\left(H^{n+1}\right)$
(The notation " $X_{c}$ " indicates the dependence of the compensated position on the step h). Multiplying (3.14) by $2^{n}$ and subtracting 3.13 gives

$$
2^{n} x_{c}(H / 2)-x_{c}(H) \approx\left(2^{n}-1\right) x+\theta\left(H^{n+1}\right)
$$

Therefore

$$
x=\frac{2^{n} x_{c}(H / 2)-x_{c}(H)}{2^{n}-1}+\theta\left(H^{n+1}\right)
$$

## IV GRAPHS AND TABLES

## KEY TO GRAPHS

In figures 2 and 3 the vertical scale is $1 \mathrm{~cm}=.01 \mathrm{ft}$. while in figures 4 and 5 it is $1 \mathrm{~cm}=.001 \mathrm{ft}$. In fig. 7 the vertical scale is $1 \mathrm{~cm}=.01 \mathrm{ft} / \mathrm{sec}$.

In all position vs time or velocity vs time graphs the horizontal scale is $1 \mathrm{~cm}=.1 \mathrm{sec}$.
"l" signifies theoretical performance with the given parameters.
"2" signifies uncompensated performance.
" 3 " is FIX*C correction.
The dashed line in Figs. 6,7 is FIX*T4 correction.
"4" in Fig. 2 is performance with a thirty pound force threshold.
Fig. 2 shows the distorting effect of the force threshold (at two levels) with attempted correction.

Fig. 3 represents theoretical, uncompensated and compensated performance of an Apollo type probe with typical parameters.

Fig. 4 contrasts the performance and compensation for a preloaded vs non-preloaded probe.

Fig. 5 is similar to Fig. 3 but the probe has no preload.

## 6 DOF CONTACT DYNAMICS TEST RESULT (8727-28779)

Test involved vertical motion only and a .75" plywood target simulating the docking collar.
I.

Vo
.05
.05
.05
.05
II.

## $V_{0}$

.10
.10
.10
.10
III.

| $V_{0}$ | $V_{\mathbf{4}}$ |
| :--- | :--- |
| .1 | .2216 |
| .2 | .3826 |
| .2 | .4101 |

.2216
.4101

$$
\begin{aligned}
& \quad \mathrm{V}_{f} \\
& \\
& .1965 \\
& \\
& \\
& \\
& \\
& \hline
\end{aligned} .2122245
$$

. 2

$$
\begin{array}{ll}
\mathrm{V}_{4} & 8 / 29 / 79 \\
.1927 \\
.1907 \\
.1971
\end{array} \quad \mathrm{~F}_{\mathrm{t}}=30 \mathrm{lbs} .
$$

## 8/17/79

$\mathrm{F}_{\mathrm{t}}=30 \mathrm{lbs}$

8/28/79
$F_{t}=35$

THE EFFECT OF FORCE THRESHOLD ON REBOUND PARAMETERS

| $\mathrm{F}_{t}$ | $\bar{n}$ - model | $V_{+}$- formula | $\mathrm{V}_{f}$ - Analog Simulation | $\mathrm{V}_{f}$ - Model |
| :---: | :---: | :---: | :---: | :---: |
| 0 | . 0637 | . 159 | . 157 | . 157 |
| 20 | . 0668 | . 1666 | - | . 162 |
| 30 | . 0701 | . 1704 | . 165 | . 170 |
| 40 | . 0746 | . 1742 | - | . 171 |
| 60 | . 855 | . 1818 | . 182 | . 182 |
| 70 | . 978 | . 1850 |  | . 188 |
| 80 | . 984 | . 1896 | - | . 1968 |
| 90 | . 102 | . 1934 | - | . 1974 |
| 100 | . 112 | . 1973 | - | . 198 |
| 120 | . 127 | . 2048 | . 22 | . 223 |
| 240 | . 22 | . 2510 | . 312 | . 332 |

## PROBE DEFLECTION PROFILE

| $\eta(f t)$ | $F(l b s)$ |
| :--- | :---: |
| 0 | 69 |
| .125 | 74.8 |
| .25 | 91 |
| .417 | 89.3 |
| .5 | 93.75 |
| .625 | 101 |
| .708 | 105.2 |
| .792 | 118.5 |
| .88 | 121 |
| .9 | 125 |
| .911 | 130 |

For $\eta \geqslant .912$ there is a bottoming spring with $k=480,0001 b s / f t$

TABLE 3


I $\quad F_{p}=0 \quad K=1200$ LBS $/ F T \quad M=311$ SLUGS



## SIMULATION RESULTS CORRESPONDING TO TABLE I

$$
\begin{aligned}
& \mathrm{F}_{\mathrm{p}}=69 \mathrm{LBS} \mathrm{~K}=60 \mathrm{LBS} / \mathrm{FT} \quad \mathrm{M}=311 \mathrm{SLUGS} \\
& \mathrm{~F}_{\downarrow}=30 \mathrm{LBS} \text {. } \\
& \text { I } \\
& V_{0}=.05 \\
& V_{f}=.163 \\
& \text { II } \\
& V_{0}=.1 \\
& V_{t}=.227 \\
& \text { III } \\
& V_{0}=.2 \\
& V_{f}=.33 . \\
& \text { IV } \quad F_{t}=35665 \\
& V_{0}=\cdot 2 \\
& V_{f}=.343
\end{aligned}
$$

TABLE 6

## APPROXIMATE FORMULAS FOR A DELAYED SPRING

Assuming $w_{0} T / 2 \ll 1$ the following approximate formulas may be derived for the IVP:

$$
\begin{aligned}
& \ddot{\eta}+w_{0}^{2} n_{T}=0 \\
& \eta(t)=v_{0} t \quad 0 \leq t \leq T \\
& \dot{\eta}(t)=v_{0}
\end{aligned}
$$

$$
n(t)=v_{0} / \omega_{0} \exp \left(\omega_{0}^{2} r t / 2\right) \sin \omega_{0} t
$$

$$
\bar{n}=v_{0} / \omega_{0} \quad \exp \left(\pi+\omega_{0} / 4\right)
$$

$$
V_{F} \geqslant \quad V_{0} \exp \left(\pi+w_{0} / 2\right)
$$

These formulas give relative errors of under $3 \%$.
If a force threshold is present then

$$
V_{F} \approx \exp \left(\frac{\pi}{2} \omega_{0} r\right)\left(V_{0}+F_{T} r / 2 m\right)
$$

TABLE 7





The second order delay equation model developed in this paper agrees with a recent analog simulation [1] of the effect of delay. It is also reasonably consistent (See Table 6 with the observed behavior (up to noise) of the 6 DOF system when $\mathrm{V}_{0}=.1$ or .2. The most disagreement happens at $\mathrm{V}_{6}=.05$ when our simulation gives a rebound velocity about $13 \%$ below the mean of the measurements in Table 1.

The model predicts that for a pure spring with no preload or force threshold $V_{f} / V_{0}$ depends only onr and and not on $V_{0} . V_{0}$ can also be computed quite accurately by a simple formula without numerically integrating the equation in the model. The use of a threshold increase the effect of delay. The results are relatively most severe at high threshold levels and low initial velocities. For thresholds under 100 lbs and delays less than .2 sec $V_{f}$ can be predicted with reasonable accuracy by an approximate formula. Verification of these aspects of the model in hardware tests would strengthen the hypothesis that delay is the predominate factor in the observed errors in the system.

These tests and the isolation of other factors - such as phase shifts due to the filters-should be possible when the noise of the system is reduced. Some of this noise is probably caused by the sensitivity of the force sensors on the table to inertial effects, and by variable delay due to the aysynchronization of the AD/DA converters (see 2.2.4-2.2.5 above). There may also be another source of apparent noise: for the parameters chosen to model the probe and a force threshold the rebound velocity changes rapidly towards the end of the simulation. If the rebound velocity is measured at $t_{f}+r$ (cf. 2.2.3) an error of $\pm$ $\Delta t$ in the time of measurement can produce a large relative change in $\mathrm{V}_{\mathrm{f}}$. For example when $\mathrm{V}=.2 \mathrm{~V}_{\boldsymbol{f}}$ is .343 ; 033 sec later it is 354; 066 sec later it is. $38,066 \mathrm{sec}$ before it is .32 . Hence, slight differences in time of final measurement in repeated runs can cause variations close to the order of magnitude of those observed.

Every effort ought to be made to reduce noise and delay. Several means to this end are suggested in[1 ]. We agree with these recommendations. Since a delay as low as 50 ms still causes considerable errors and this may be close to an achievable lower limit, means of compensating for delay need to be implemented.

In this study both a fourth order Taylor method and a Chebychev polynomial approximation to the transfer function $e^{\text {s }}$ gave good results for for an unpreloaded spring when no force threshold is used. The velocity profile correction in this case was excellent; penetration errors were reduced to under $10 \%$ (most of
the time less than 5\%). The Chebychev algorithm "FIX*C" was the most accurate.

We regret however that neither algorithm works in the case of a preloaded spring, when the parameters model the probe tested in Table l. Indeed its is remarkable that they hardly perturb the uncompensated position time profiles even when $\tau<.05 \mathrm{sec}$. Their effects on an unpreloaded spring with a force threshold are more complicated but also unsatisfactory.

If these options are not satisfactory, a deeper mathematical analysis needs to be made on whether compensation of a preloaded system with delay is possible. On this topic the following rough observation seems appropriate. Since the difference between the compensated and the theoretically correct simple harmonic system is $\sigma\left(\tau^{n}, X^{(9)}\right.$ where $x$ is the solution of the compensated system an $n$ is the order of the method, the derivative $x^{(n)}$ may grow faster beyond $x=x_{c}$ - than $\tau^{n}$ goes to zero. If this phenomenon occurs for all realistic compensation schemes it is doubtful that the 6 DOF system can simulate contact dynamics unless the delay is reduced far below 50 ms .
l. Allen, D. W., W. M. Gover, L. G. Thomas, and F. L. Vinz,
2. Atkinson, K.E.,
3. Bellman, R. and
K. L. Cooke,
4. Callender, A., D. R. Hartree, A. Porter,
5. Conte, S. D. and C. de Boor
6. Driver R. D.,
7. El'sgol'ts, L.E. and S. B. Norkin,
8. Hale, J.,
9. Levine, L.,
10. Minorsky, N.,
11. $\qquad$
12. Porter, A. and $F$. Stoneman,
13. Ragazzini, J.R. and G. F. Fianklin
"An Investigation of the Performance of the Simulation Equipment used in Contact Dynamics," manuscript.

An Introduction to Numerical Analysis, John Wiley, New York, 1978

Differential - Difference Equations, Academic Press, 1963
"Time-lag in a Control System," Phil. Trans. Roy Soc. London, A235 (1936), 415-444.

Elementary Numerical Analysis, $\overline{3 r d} \overline{\text { Ed. }}$ Mc $\overline{G r a w-H i l l}$, New York, 1980.
"Linear Differential Systems with Small Delays," Journal of Differential Equations, 21 (1976), 148-165.

Introduction to the Theory and Application of Differential Equations with Deviating Arguments., Academic Press, New York, 1973.

Theory of Functional Differential Equations, Spi inger Verlag, New York, 1977.

Methods for Solving Engineering $\frac{\text { Problems }}{\text { McGraw }} \frac{\text { Using }}{} \frac{\text { An alog }}{\text { Computers }}$, McGraw Hill, New York, 1964
"Self-excited Oscillations in Dynamical Systems Possessing Retarded Actions," J. Appl. Mech. 9 (1942), GS-71.

Nonlineal Oscillations, Chap. 21, Van Nostrand, Princeton, 1962.
"A New Approach to the Design of Pulse-monitored Servo Systems," Proco IEE, 97 pt II (1950)m 597.

Sampled-Data Control Systems, McGraw-Hill, New York, 1958.
14. Ralston, A.,
15. Wolfe, M.A.,

A First Course in Numerical Analysis, $\overline{M c G r a w-H i l l, ~ N e w ~ Y o r k, ~} 1965$.

A First Course in Numerical Analysis, Van Nostrand, London, 1972.

The following is an attempted simulation of the effects of noise caused by the inertial forces of the table affecting the sensor. We integrate equation (2.8) (see $\xi$ 2.2.5), thus representing a "shaking" of the sensor at each step in the movement of the table. Here the "noise parameter" $\epsilon$ is a uniformly distributed random variable taking values between 0 and either .5 or 1 . The mass of the table is about 12 slugs so that the mass of the table and probe is 323 slugs. Note that the rebound velocities with the model exceed those observed. parameters: $F=69 \mathrm{lbs} \quad K=60 \mathrm{lbs} / \mathrm{ft} \quad \tau=.2 \mathrm{sec}$.
I. $f=.5$
A. $V_{0}=.05$
C. $V_{0}=.2$
$V_{f}$
.1778
$V_{+}$
.204
.395
.2001
.4012
.1967
.383
.398

B $V_{0}=.1$
$V_{\downarrow}$
.2773
.2525
.269
.2649
II. $=1$
A. $V_{0}=.05$
C. $\quad V_{0}=.2$
Vf
.2248
.2186 . 46
.2356
.4537
.2433 . 4376
B. $V_{0}=.1$
$V_{f}$
.3100
.3429
. 3828
. 3298

## APPENDIX II

## LIST OF PROGRAMS

All programs are in BASIC completely interactive and implemented on MFC's Xerox Sigma 5 system.

## Name

LAG*R
PSP*I

FIX*C
FIX*T4

LAG*N

## Purpose

Simulates the basic delay model.
Theotetical spring with arbitrary initial conditions

Chebychev polynomial correction
Fourth order Taylor method correction

Simulates the delay model with inertial noise

## APPENDIX III

## THE ALGORITHM "FIX*C"

Let $C_{1}:=.994526$

$$
\begin{aligned}
& C_{2}:=.995682 \\
& C_{2}:=.543081 \\
& C_{4}:+.179519
\end{aligned}
$$

Set

$$
x_{c}:=c_{1} x_{T}+c_{2} x_{T}^{\prime}+c_{3} \ddot{x}_{T}+c_{1} \ddot{x}_{T}
$$

The following algorithm solves the IVP

$$
\begin{aligned}
& x^{\prime \prime}+w_{0}^{2} x_{c}=0 \\
& x(t)=x_{0}, \quad \dot{x}(t)=v_{0}, \quad 0 \leq t \leq r
\end{aligned}
$$

For $i=1$ to $m$

$$
\begin{aligned}
v_{i}: & =v_{0} \\
x_{i}: & =i+v_{0}+x_{0}
\end{aligned}
$$

Continue
For $i=m+1$ until satisfied

$$
\begin{aligned}
& v_{i+1}= v_{i}-\omega_{0}^{2}\left\{H / 2 C_{1}\left(\left(x_{i-m}+x_{i+1-m}\right)+H / c\left(v_{i-m}-v_{i+1-m}\right.\right.\right. \\
&+c_{2}\left(x_{i+1-m}-x_{i-m}\right)+c_{3}\left(v_{i+1-m}-v_{i-m}\right) \\
&+c_{4}\left(a_{i+1-m}-a_{i-m}\right) \\
& \text { IIT-40 }
\end{aligned}
$$

$$
\begin{aligned}
& a_{i+1}=-w_{0}^{2} x_{i+1-m} \\
& a_{i+1}^{c}=-w_{0}^{2}\left(x_{i+1-w}+v_{i+1-m}\right) \\
& x_{i+1}=x_{i}+H / 2\left(v_{i}+v_{i+1}\right)+H^{2} / 12 a_{i}^{c}-\left(a_{i+1}^{c}\right)
\end{aligned}
$$

## APPENDIX IV

ILLUSTRATIONS OF THE 6 DOF MOTION SYSTEM


SIX-DEGREE-OF-FREEDOM MOTION SYSTEM SKETCH


Fig 10

|  |  | POSITION | RATE | ACCELE | RATION |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | NO LOAD | 23,003 LB. LOAD |
|  | PITCH | $+30^{\circ},-20^{\circ}$ | $\pm 15^{\circ} / \mathrm{SEC}$. | $\pm 6.5$ RAD/SEC ${ }^{2}$ | $\pm 2$ RAD/SEC ${ }^{2}$ |
|  | ROLL | $\pm 22^{\circ}$ | $\pm 15^{\circ} / \mathrm{SEC}$. | $\pm 7.0$ RAD/SEC ${ }^{2}$ | $\left.\begin{array}{l} +1.6 \\ -2.0 \end{array}\right] \text { RAD/SEC }{ }^{2}$ |
| 日 | YAW | $\pm 32^{\circ}$ | $\pm 15^{\circ} \mathrm{ISEC}$ | $\pm 6.0$ RAD/SEC ${ }^{2}$ | $\pm 2.0$ RADISEC ${ }^{2}$ |
| $\stackrel{1}{*}$ | VERTICAL | 39 IN. UP, <br> 30 IN. DOWN | $\pm 24$ IN. /SEC. | $\pm 1.6 \mathrm{~g}$ | $\pm 1.0 \mathrm{~g}$ |
|  | LATERAL | $\pm 48$ INCHES | $\pm 24$ IN. /SEC. | $\pm 2.4 \mathrm{~g}$ | $\pm 0.6 \mathrm{~g}$ |
|  | LONGITUDINAL | $\pm 48$ INCHES | $\pm 24$ IN. /SEC. | $\pm 2.0 \mathrm{~g}$ | $\pm 0.6 \mathrm{~g}$ |

..............
NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAMMARSHALL SPACE FLIGHT CENTERTHE UNIVERSITY OF ALABAMA
PILOT SIGNALS FOR LARGE ACTIVE RETRODIRECTIVE ARRAYS
Prepared By: C. H. Chan, Ph.D.
Academic Rank: Professor of Physics
Unviersity and Department: University of Alabama at Huntsville
Department of Physics
NASA/MSFC
(Laboratory)(Division)(Branch)
MSFC Counterpart: Woolsey Finnell III
Preliminary Design OfficeSubsystems Design DivisionElectrical System Branch
Date: August 1, 1980
Contract No.: NGT-01-002-099(University of Alabama)

## PILOT SIGNALS FOR LARGE ACTIVE RETRODIRECTIVE ARRAYS

BY

C. H. Chan, Ph.D.<br>Professor of Physics University of Alabama at Huntsville Huntsville, Alabama

ABSTRACT

It has been suggested that for large active retrodirective arrays, as in the solar power system, a two-tone uplink pilot signal with frequencies symmetrically situated around the downlink frequency be used in order to reduce ionospheric biases and to lower the cost since a two-tone receiver is economically much cheaper than a single-tone phase-locked receiver. Unfortunately such a system now faces the following well-known difficulties: (i) the $\pi$-ambiguity, (ii) a large phase difference between the downlink and uplink signa1s.

We show in this report how the $\pi$-ambiguity can be easily removed by using a two-tone uplink signal with both frequencies situated at one side of the downlink frequency, and the phase difference can be greatly reduced with a three-tone or a four-tone uplink pilot signal.
I wish to thank the NASA/ASEE Summer Faculty Fellowship Program and its directors for providing the opportunity to participate in the Summer (1980) research project. Appreciation is extended to Mr. W. Finne11. III for his support of my work this summer. Special thanks are to Mr. R. H. Durrett and Mr. R. A. Inman of EC-33 for their many enlightening discussions.

## I. Introduction

It has been suggested ${ }^{1}$ that active retrodirective arrays ${ }^{2,3}$ would be particularly suitable as solar power satellite's antennas ${ }^{4}$ because they are inherently failsafe. The active retrodirective array works on the so-called phase-conjugation principle. It electronically points a microwave beam back at the apparent source of an incident pilot signal. Retrodirectivity is achieved by retransmitting from each element of the array a signal whose phase is the conjugation of that received by the element. In the satellite power system, the pilot source on ground may be situated at the center of a large rectenna and the retrodirective array is the space antenna in geosynchronous orbit.

Retrodirectivity can be most easily achieved if the uplink signal and the downlink beam have the same frequency. But due to input-output isolation problems. the uplink frequency is either upshifted or downshifted from the downlink frequency, a phase-locked receiver is used to achieve phase conjugation. When the uplink and downlink frequencies are different and because the ionosphere and transmission lines are dispersive, the conjugated uplink phase is no longer exactly equal to the downlink phase and the beam coherence at the rectenna can be lost. The downlink beam then points to wrong directions and this is known as beam squint.

A two-tone pilot uplink signal with frequencies symmetrically situated around the downlink frequency has subsequently been suggested. The two-tone uplink signal circumvents the beam squint problem. It reduces ionospheric biases and biases due to the dispersion of the transmission line. It also lowers the cost since a two-tone receiver is much

$$
I V-4
$$

cheaper than a single-tone phase-locked recelver. But, it introduces a new problem, known as the $\pi$-ambiguity, and the old problem that the conjugated uplink phase could still differ from the downlink phase by an intolerable amount remains. Raytheon, ${ }^{5}$ Boeing ${ }^{6}$, and Rockwell ${ }^{7}$ have all made further suggestions to remedy the problems. Their solutions not only are complicated and require much extra hardware in the already very complicated phase conjugation circuitry, but certain problems still remain.

In this report new designs of two-tone and multi-tone uplink signals with frequencies situated at one side of the downlink frequency are suggested. This method removes the above mentioned difficulties and does not require extra components in the phase conjugation circuit. We shall review in our next section the basic principle of phase conjugation and where the problems are in a symmetrically situated two-tone uplink signal. In section III we show how these problems can be circumvented with our new designs.

## II. Difficulties With Symmetrically Situated Two-Tone Uplink Signal

A retrodirective array electronically transmits a microwave beam back to the apparent source of a coherent pilot signal. The beam radiated by self-phasing antenna may or may not be coherent across the aperture but it is coherent when it arrives back at the source. Retrodirectivity is the result of phase conjugation of the pilot signal received by each element of the array. Let the phase of the pilot signal of angular frequency $\omega$ received by the $k$ th element of the array at time $t$ be $\phi_{k}(t)=\omega\left(t-r_{k} / c\right)$ where $r_{k}$ is the distance from the $k t h$ element to the source. We define the conjugate of $\phi_{k}$ to be

$$
\phi_{k}^{*}(t)=w\left(t+r_{k} / c\right)+\phi_{o}
$$

where $\phi_{0}$ is an arbitrary phase offset but is constant over the entire array. The phase of the beam received from the kth element by a receiver located at the pilot source $(r=0)$ is, at time $t$,

$$
\phi_{k}(t, 0)=\omega\left(t+\frac{r_{k}}{c}-\frac{r_{k}}{c}\right)+\phi_{o}=\omega t+\phi_{0}
$$

Thus the contributions to the field at $r=0$ from various elements of the array are all in phase at that point.

In the above simple example the uplink frequency was chosen to be the same as the downlink frequency. This restriction is neither necessary nor desirable and is usually avoided because of input-output isolation problems. When these two frequencies are different, a phase-locked receiver Is used. Retrodirectivity can still be achieved - provided that the propagating medium is nondispersive.

Due to the fact that single-tone phase-locked receivers are expensive and the ionosphere and transmission lines are dispersive, a two-tone uplink signal with frequencies symmetrically situated around the down-link
frequency was suggested and the average of the phases of the uplink tones be taken as a good estimate of the phase at the downlink frequency. Such a system lowers the cost and removes partially the difference between the uplink and downlink phases but it also introduces a new problem known as the $\pi$-ambiguity. We shall review here where these problems are. We shall use the ionosphere as an example to study the effect on the phase conjugation due to the dispersive property of the propagating medium.

The dispersion relation for ionosphere with $\omega \gg \omega_{p}$ is

$$
\begin{equation*}
k=\frac{\omega}{c}\left(1-\frac{\omega}{\frac{p^{2}}{2 \omega^{2}}}\right) \tag{1}
\end{equation*}
$$

where $c$ is the speed of light in vacuum, $k$ the wave number, $\omega$ the angular frequency and $\omega_{p}$ is the plasma frequency. The plasma frequency $\omega_{p}$ is related to the electron density as

$$
\begin{equation*}
\omega_{p}^{2}=\frac{N e^{2}}{\varepsilon_{\mathrm{o}} \mathrm{~m}} \tag{2}
\end{equation*}
$$

where $\varepsilon_{0}$ is the vacuum permittivity, $e$ the charge of the electron, $m$ the mass of the electron and $N$ is the volume electron density. From Eqs. (1) and (2)

$$
k=\frac{2 \pi f}{c}-\frac{N e^{2}}{4 \pi \varepsilon_{0} f c m}
$$

where $\mathrm{f}=\omega / 2 \pi$ is the frequency. For a beam to transverse a path length $L$, the total phase change is

$$
\begin{align*}
& \phi=\int k d s \\
& =\frac{2 \pi f L}{c}-\frac{e^{2}}{4 \pi \varepsilon_{0}} \frac{1}{f c m} \int N d s \\
& =A f-\frac{B}{f} \tag{3}
\end{align*}
$$

where $A=\frac{2 \pi L}{c}, \quad B=\frac{e^{2}}{4 \pi \varepsilon_{0}} \frac{1}{c m} \int N d s$, and $\int N d s$ is the columal electron density.

With uplink frequencies situated around the downlink frequencies, we have $f_{1}=f_{D}-\Delta f$ and $f_{2}=f_{D}+\Delta f$ where $f_{D}=2.45 \mathrm{GHz}$ is the downlink frequency. Using the notation of Eq. (3)
and

$$
\begin{aligned}
& \phi\left(\mathrm{f}_{1}\right)=\mathrm{Af}_{1}-\frac{\mathrm{B}}{\mathrm{f}_{1}}=\phi_{1} \\
& \phi\left(\mathrm{f}_{2}\right)=\mathrm{Af}_{2}-\frac{\mathrm{B}}{\mathrm{f}_{2}}=\phi_{2}
\end{aligned}
$$

then the average of the two phases is

$$
\begin{align*}
& \bar{\phi}=\frac{1}{2}\left(\phi_{1}+\phi_{2}\right)  \tag{4}\\
& =A f_{D}-\frac{B}{2}\left(\frac{1}{f_{D}-\Delta f}+\frac{1}{f_{D}+\Delta f}\right) \\
& =A f_{D}-\frac{B}{f_{D}}\left(1+\varepsilon^{2}+\varepsilon^{4} \ldots .\right)
\end{align*}
$$

where $\varepsilon=\Delta f / f_{D}$ is a small number. We may compare this phase $\bar{\phi}$ with the downlink phase

$$
\phi_{\mathrm{D}}=\mathrm{Af} \mathrm{f}_{\mathrm{D}}-\frac{\mathrm{B}}{\mathrm{f}_{\mathrm{D}}}
$$

e.g. their difference is of the order

$$
\Delta \phi=\left|\phi_{\mathrm{D}}-\bar{\phi}\right| \approx \frac{\mathrm{B} \varepsilon^{2}}{\mathrm{f}_{\mathrm{D}}}
$$

To estimate this difference, we sha11 assume

$$
\int \mathrm{Nds}=5 \times 10^{17} \text { electrons } / \mathrm{m}^{2}
$$

a rather large value for $\int N$ ds but taking into account for the possible worst condition. With $\Delta f=50 \mathrm{MHz}, \Delta \phi$ is estimated to be about $4^{\circ}$, which is not too small.

When we took the average of phases $\phi_{1}$ and $\phi_{2}$ in Eq. (4), there could introduce an ambiguity known as the $\pi$-ambiguity. Let

$$
\phi_{1}+\phi_{2}=\mathrm{K}(2 \pi)+\Delta
$$

where $0 \leq \Delta<2 \pi$ and $K$ is a positive, zero or negative integer. Hence

$$
\bar{\phi}=\mathrm{K} \pi+\frac{\Delta}{2}
$$

In performing the phase average, the $K \pi$ term could get lost. For $K$ even, no damage is done. For $K$ odd a $\pi$ error is introduced and one would conjugate the wrong phase.

In order to remove these two difficulties, especially the $\pi$-ambiguity, Raytheon ${ }^{5}$, Boeing ${ }^{6}$, and Rockwel1 ${ }^{7}$ have all made suggestions. Their solutions are very complicated and usually require a lot of hardware in the receiving and phase conjugation circuitries with much added cost. Furthermore, their soultions do not solve the problem completely. We shall show in our next section several simple solutions which circumvent the above mentioned difficulties and do not add extra costs.

In this section, several uplink designs are proposed. The first simple design avoids the $\pi$-ambiguity and the rest are improved versions of the first one. They are all free from the $\pi$-ambiguity but reduce the phase difference $\Delta \phi$ to various orders.
(i) This is also a two-tone uplink, but the two frequencies are both on one side of the downlink frequency with

$$
\begin{equation*}
f_{1}=f_{D}-\Delta f \tag{5}
\end{equation*}
$$

and $\quad f_{2}=f_{D}-2 \Delta f$
where $f_{D}$ is the downlink frequency. We now let $\bar{\phi}=2 \phi\left(f_{1}\right)-\phi\left(f_{2}\right)$ to be the estimation of the downlink phase. With the notation of Eq. (3)

$$
\begin{aligned}
\bar{\phi} & =A f_{D}-B\left(\frac{2}{f_{D}-\Delta f}-\frac{1}{f_{D}-2 \Delta f}\right) \\
& =A f_{D}-\frac{B}{f_{D}}\left(1-2 \varepsilon^{2}+. . . .\right)
\end{aligned}
$$

where again $\varepsilon=\Delta f / f_{D}$ is a small number. In this simple design the $\pi$-ambiguity is removed since no division of the phase is used anywhere. The combined phase $\bar{\phi}$ is also a good estimate of the downlink phase $\phi_{D}$. Their difference $\Delta \phi=\left|\phi_{D}-\bar{\phi}\right|$ is

$$
\begin{equation*}
\Delta \phi=\frac{2 \mathrm{~B}}{\mathrm{f}_{\mathrm{D}}} \varepsilon^{2} \tag{6}
\end{equation*}
$$

which is of the same order as the one with two uplink frequencies situated symmetrically around the downlink frequency and is about $8^{\circ}$ when the same values for $\int N \mathrm{ds}$ and $\Delta \mathrm{f}$ are used as in the last
section. We also note that in this simple design no extra components are required in the receiving circuitry.
(ii) This is an improved version of the first design. It requires three uplink tones but it also greatly improves the accuracy in estimating the downlink phase $\phi_{D}$. The three frequencies are

$$
\begin{aligned}
f_{1} & =f_{D}-\Delta f \\
f_{2} & =f_{D}-2 \Delta f \\
\text { and } \quad f_{3} & =f_{D}-3 \Delta f
\end{aligned}
$$

We now let

$$
\begin{equation*}
\bar{\phi}=3 \phi\left(f_{1}\right)-3 \phi\left(f_{2}\right)+\phi\left(f_{3}\right) \tag{7}
\end{equation*}
$$

to be the estimation of the downlink phase. With the notation of Eq. (3)

$$
\begin{aligned}
& \bar{\phi}=A f_{D}-B\left(\frac{3}{f_{D}-\Delta f}-\frac{3}{f_{D}-2 \Delta f}+\frac{1}{f_{D}-3 \Delta f}\right) \\
& =A f_{D}-\frac{B}{f_{D}}\left(1+6 \varepsilon^{3}+\ldots\right)
\end{aligned}
$$

where again $\varepsilon=\Delta f / f_{D}$. In this design, there is no $\pi$-ambiguity as before and the difference between $\phi_{D}$ and $\bar{\phi}$ is reduced by an extra factor $\varepsilon$, e.g. the difference $\Delta \phi$ is now

$$
\begin{equation*}
\Delta \phi=\frac{6 B}{f_{D}} \varepsilon^{3} \tag{8}
\end{equation*}
$$

With $\Delta f=50 \mathrm{MHz}$ and $\int \mathrm{N} d s=5 \times 10^{17}$ electrons $/ \mathrm{m}^{2}$, this difference $\Delta \phi$ is only about $0.5^{\circ}$, which is samll.
(iii) This version can be used in the event we would like to have an even smaller $\Delta \phi$ or we would like to use a larger $\Delta f$, which
would otherwise result in a too large $\Delta \phi$ even with a three-tone uplink design. These new specifications can be achieved at the expense of adding a fourth tone. The four frequencies are

$$
\begin{aligned}
& f_{1}=f_{D}-\Delta f \\
& f_{2}=f_{D}-2 \Delta f \\
& f_{3}=f_{D}-3 \Delta f \\
& f_{4}=f_{D}-4 \Delta f
\end{aligned}
$$

and we now let

$$
\begin{align*}
\bar{\phi} & =4 \phi\left(f_{1}\right)-6 \phi\left(f_{2}\right)+4 \phi\left(f_{3}\right)-\phi\left(f_{4}\right)  \tag{9}\\
& =\Lambda f_{D}-B\left(\frac{4}{f_{D}-\Lambda f}-\frac{6}{f_{D}-2 \Delta f}+\frac{4}{f_{D}-3 \Delta f}-\frac{1}{f_{D}-4 \Delta f}\right) \\
& =A f_{D}-\frac{B}{f_{D}}\left(1-24 \varepsilon^{4}+\ldots\right)
\end{align*}
$$

In this design, the phase difference is reduced further by a factor $\varepsilon$, e.g. $\Delta \phi$ is only

$$
\begin{equation*}
\Delta \phi=\frac{24 \mathrm{~B}}{\mathrm{f}_{\mathrm{D}}} \varepsilon^{4} \tag{10}
\end{equation*}
$$

With $\Delta f=50 \mathrm{MHz}$ and $\int N \mathrm{ds}=5 \times 10^{17}$ electrons $/ \mathrm{m}^{2}$, this difference Is only about $0.04^{\circ}$. Even with $\Delta f=100 \mathrm{MHz}$, this difference is only $0.66^{\circ}$ which is still very small.

In all these three designs the uplink frequencies are all on the lower side of the downlink frequency and they are all equally spaced. These are not the only choices. One can use all frequencies on the upper side of the downlink frequency and they need not be equally spaced
either. As a simple example one may very well have a two-tone uplink with

$$
\begin{align*}
& f^{\prime}=f_{D}+2 \Delta f  \tag{11}\\
& f^{\prime \prime}=f_{D}+3 \Delta f
\end{align*}
$$

and $\bar{\phi}=3 \phi\left(f^{\prime}\right)-2 \phi\left(f^{\prime \prime}\right)$. It will work just as well.
So far we have used the ionosphere as an example to show how the dispersion of the transmission medium could introduce sizable biases between the uplink and downlink phases. For a two-tone uplink signal, this phase difference is about $8^{\circ}$. We suggested one way to suppress it is to use a three-tone uplink. However, it turns out that if the biases were purely due to the ionosphere, it is really not necessary to use a three-tone uplink. This is because even though the phase difference between $\bar{\phi}$ and $\phi_{D}$ in our two-tone uplink design is large, its variation from subarray to subarray will be decimal if not infinitesimal. To corroborate more on this statement, from Eq. (6) and the definition of $B$, the difference between the uplink and downlink phases from the pilot to the $k$ th subarray for a two-tone uplink is

$$
\begin{equation*}
\Delta \phi_{k}=\frac{2(\Delta f)^{2}}{f_{D}^{3}} \frac{\varepsilon^{2}}{4 \pi \varepsilon_{o}} \frac{1}{c m} \int N \quad d s_{k} \tag{12}
\end{equation*}
$$

where $\int \mathrm{N} \mathrm{ds}{ }_{k}$ is the columnal electron density along the path from the pilot to the kth subarray. Since the horizontal dimension of the region of the ionosphere that would be transversed by the pilot signal to any subarray is very small, typically of the order less than 100 meters, the transverse variation of the total electron along a path of several hundred kilometers within a tube of diameter less than 100 meters would be very small. Hence even though $\Delta \phi_{k}$ is estimated to be about $8^{\circ}$, its variation is at least several orders less. Then with
the use of the phase from one of the subarrays as the reference phase, this difference $\Delta \phi_{k}$ can be subtracted out and the remainder be treated as a constant phase offset which has no effect on the retrodirective beam. In this sense, though the ionosphere is dispersive, it does not cause any problem, and a two-tone uplink is sufficient.

On the other hand, with the use of "central phasing" we cannot avoid the extra path length of transmission lines for some subarrays. Since these lines are not dispersionless, they will introduce a sizable phase difference. This phase difference can also be estimated. If we assume that this transmission line is a wave guide, its dispersion relation is well known

$$
\begin{equation*}
\mathrm{k}=\frac{2 \pi}{\mathrm{c}} \sqrt{\mathrm{f}^{2}-\mathrm{f} \lambda^{2}} \tag{13}
\end{equation*}
$$

where $f_{\lambda}$ is the cut-off frequency. Hence the phase for any link is

$$
\begin{equation*}
\phi(\mathrm{f})=\frac{2 \pi \ell}{\mathrm{c}} \sqrt{\mathrm{f}^{2}-\mathrm{f} \lambda^{2}} \tag{14}
\end{equation*}
$$

where $\ell$ is length of the transmission line. Just as before the phase differences for various uplinks can be calculated. For a two-tone. uplink with $\bar{\phi}=2 \phi\left(f_{1}\right)-\phi\left(f_{2}\right), \ell=500$ meters and $f_{\lambda}=\frac{1}{2} f_{D}$, where $f_{1}$ and $f_{2}$ are given as in Eq. (5), $\Delta \phi$ is

$$
\Delta \phi=\left|\phi_{D}-\bar{\phi}\right|=256^{\circ}
$$

Similarly for a three-tone uplink with $\bar{\phi}=3 \phi\left(f_{1}\right)-3 \phi\left(f_{2}\right)+\phi\left(f_{3}\right)$, one obtains

$$
\Delta \phi=23^{\circ}
$$

For a four-tone uplink with $\bar{\phi}=4 \phi\left(f_{1}\right)-6 \phi\left(f_{2}\right)+4 \phi\left(f_{3}\right)-\phi\left(f_{4}\right)$,
$\Delta \phi$ is further reduced to

$$
\Delta \phi=3^{\circ}
$$

As we see in this example, the phase difference for a two-tone uplink
Is very large but this number is greatly reduced in a three-tone or four-tone uplink. These multi-tone uplink signals can be used as a useful alternative method to suppress biases due to the dispersion of the transmission line and the medium.

In our last section we illustrated how simple designs can be used to eliminate the $\pi$-ambiguity and reduce the ionospheric biases and biases from dispersive transmission lines. We also note that none of our designs require extra components in the receiving circuitry. A11 one is required to do is to obtain $\bar{\phi}$, which can be achieved rather easily, and simply conjugate it and use it as the phase of the downlink signal leaving the space antenna.

It is also important to remember that we are here to design a pilot beam system as simple as possible with the phase received by the array as close to the downlink phase as possible. We are not asked to and it is not necessary to determine the ionospheric electron density as required in some other designs.

Lastly our designs of pilot beam can be implemented easily in any large retrodirective arrays. Their advantages are (i) avoiding using phase-locked receiver (ii) free from phase ambiguity (iii) greatly reducing biases due to dispersion of the transmission line and medium (iv) very simple to be constructed.

It will be extremely interesting to have such a system built and tested in the very near future.

## References

1. R. C. Chernoff, "Large Active Retrodirective Array for Space Application", IEEE Transaction on Antennas and Propagation, Vol. 27, PP 489-496, July, 1979.
2. M. I. Skolnik and D. D. King, "Self-Phasing Array Antennas", IEEE Transaction on Antennas and Propagation, Vol. 12, pp 142-149, March, 1964.
3. B. A. Sichelstie1, W. M. Waters and T. A. Wild, "Self-focusing Array Research Model', IEEE Transaction on Antennas and Propagation, Vol. 12, pp 150-154, March, 1964.
4. P. E. Glaser, "Power From the Sun: Its Future", Science, Vo1. 162, pp 857-861, Nov. 22, 1968.
5. Raytheon Report No. ER79-4032, "Solar Power Satellite (SPS) Pilot Beam and Communication Link Subsystem Investigation Study - Phase $\mathbf{1}^{\prime \prime}$, Jan 31, 1979, Contract No. NAS8-33157.
6. Boeing Report No. D180-24635-1, "Microwave Power Transmission System", 1978.
7. A. K. Nandi and C. Y. Tomita, "Ionospheric Effects in Active Retrodirective Array and Mitigating System Design', NASA, Solar Power Satellite, Workshop on Microwave Power Transmission and Reception, Jan. $15-18,1980$.

$$
\text { IV - } 17
$$

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM 

## MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA

ION PLATING STUDIES FOR HIGH TEMPERATURE APPLICATIONS

Prepared by:
Academic Rank:
University and Department:

NASA/MSFC:
(Laboratory)
(Division)
(Branch)
MSFC Counterpart:
Date:
Contract No.

Jack H. Davis, Ph.D.
Associate Professor
The University of Alabama in Huntsville Department of Physics

Materials and Processes
Engineering Physics
Physical Sciences
Ann F. Whitaker
August ll, 1980
NGT-01-002-099
University of Alabama

## A. Abstract

An experimental project was undertaken to ion plate by electron beam evaporation Al films onto 4340 steel substrates using (and at the same time troubleshooting) the custom built V.T.A. 7375 electron beam ion plating system. A careful recent literature and commercial vendor survey indicates possible means (reported herein) of improving the trouble plagued V.T.A. system.

I wish to express appreciation to the National Aeronautics and Space Administration, the American Society of Engineering Education, and the University of Alabama for operating and supporting this summer program. Special thanks are due the Marshall Space Flight Center and personnel of the Materials and Processes Laboratory. I have enjoyed meeting many new people - some with a fascinating special expertise with whom further collaboration is expected. Everyone has readily offered their help and assistance. I have relied heavily on the clear, lucid lab logs and final reports of my predecessor, R. J. Holliday. Special thanks to Hobert Gregory and John McClure.

## C, Introduction

Holliday ${ }^{12}$ gave the following concise introduction to ion plating.

Since first being reported in the literature in 1963 (ref. 1), ion plating has progressed to the point that today it is used in several commercial processes including the aluminum coating of fasteners used in the manufacture of aircraft (ref. 2). Ion plating has several positive qualities. Probably the most important of which are outstanding film adhesion and deposition on all sides of the substrate (including coverage into cavities) whereas normal vacuum deposition gives the usual line-of-sight coverage. Other qualities obtainable, depending on the material deposited, include exceptional corrosion resistance, high film purity, fine grain structure, very low-coefficient of friction, and improved mechanical properties of metals (ref. 3-5).

Several articles have been published describing the ion-plating process (e.g., ref. 6 and 7) Basically the process consists of two phases. In the initial cleaning phase the object to be plated (substrate) is made the cathode of a de inert gas discharge or plasma. Some of the inert gas atoms are ionized and accelerated towards the substrate. The bombardment of the substrate by these ions having high kinetic energy produces a clean surface preparatory to the actual plating process. In the plating phase the material to be plated is evaporated from the anode while maintaining the inert gas discharge. Some of the coating material atoms are also ionized and accelerated towards the substrate. These ions follow the electric field lines which terminate on all sides of the biased substrate and thereby help coat not only the front but all sides of the substrate. However, since the degree of ionization in many plasmas is very low (0.1-2\%), it has been suggested that this mechanism is probably secondary to gas scattering of neutral atoms in contributing to the high throwing power of ion plating (ref. 8). Thus the energies of the impinging particles range from that obtained from ions accelerated by the potential difference between the anode and cathode to the thermal energy of unionized atoms. These factors, along with the heating of the substrate surface due to its continued bombardment, produce the graded-fused interface, which provides the superior adherence and improved mechanical properties characteristic of ion plating.

## D. Objective

A Materials and Processes Laboratory objective is to have the capability to coat a wide variety of mechanical parts with a host of good-bonding protective thick (up to 40u) films. Ion plating is well suited for providing good film bonding.

Electron beam evaporation of source materials provides a technique applicable to the widest range of materials, even the refractory metals which are vital to both high temperature rocket engines and to high temperature superconductors. However, most electron beam evaporation ion plating systems were not off the shelf but were custom built and debugged.

The immediate objective was to debug and put into smooth operation the trouble plagued (for its five year 1ife) electron beam evaporation ion plating system Model 7375 custom built by Vacuum Technology Associates (now out of business).

## E. Experimental Hardware

The ion plating apparatus used, V.T.A. Model 7375, was custom built by Vacuum Technology Associates and consisted of three main components: (1) The vacuum system, a Varian NRC 3117, attained an ultimate pressure of $10^{-6}$ torr using a 6" diffusion pump. (2) The electron beam gun system was a Sloan Multihearth $270^{\circ}$ gun powered by a Sloan Model Five/Ten Power Supply. (3) The high voltage supply for the substrate bias was a $10 \mathrm{KV} / 500$ ma unit made by Vacuum Technology Associates.

A typical ion plating apparatus is shown in Fig. 1.


Figure 1. Ion Plating Apparatus with an Electron Beam Gun
(D. G, Teer 1977)

## F.1. Experimental Improvements Made

a. A brass cathode sample holder was removed from the system since brass was sputtercoating the substrate during precleaning and during ion plating. The chance of Zn contamination is thus reduced.
b. A cylindrical ( $5 \frac{1}{2}$ in. diameter by 3 in. long) grounded skirt extension was constructed which provides two inches of extra shielding to the top of the sample's support. This could reduce some of the above brass sputtering and efficiently concentrate the discharge near the sample substrate.
c. Due to modifications, the position of the sample holding rectangle may now be continuously located by a screw adjustment from direct contact with the cooled cathode plate to a position $5^{\prime \prime}$ lower.
d. Greater plating rates are anticipated as minimum sample substrate to hearth distances from about $4^{\prime \prime}$ to about $2^{\prime \prime}$. Some concern existed about reduced throwing power with the sample close; however, Hollidayll found no reduction in throwing power as hearth-substrate distances are reduced to from 5" to $3^{\prime \prime}$. The mean-free-path is in the several mm range so the transport is diffusion controlled.

## F.2. Discussion of Recent Literature

a. In a recenfo(Oct. '79) review article on ion plating sources, T. Spalvins ${ }^{16}$ mentions the following advantages of Wan's ${ }^{19}$ hollow cylinder electron beam source:
(1) It operates in the plasma.
(2) It does not have the hot filament or the filament related problems.
(3) It is easy to design and construct.

The main disadvantage is some contamination due to the cylinder sputtering into the substrate.
b. Ion plating by radio frequency induction evaporation is mentioned by Spalvins ${ }^{18}$. The source metal is placed in a ceramic crucible which is centered in the inducation coil. Frequency reduction by a factor of $1 / 6$ to 75 MHz prevented coil arcing in the plasma.
c. Rules for ion plating: ${ }^{16}$
(1) Ideal conditions for ion plating include:
(a) Edge of cathode dark space is as close as possible to the boat.
(b) Dark space is as wide as possible.
(2) A cathode to anode distance of 15 cm is adequate to establish a glow discharge.
(3) A metal screen is used to coat plastics and insulators.
(4) Precleaning (surface oxide removal) is usually complete when the discharge current flattens out with time. It may take one hour for stainless steel.
(5) Bare wire should not be used to make high voltage connections according to Mattox. 6
d. References 19 and 20 are listed because they are recent.
e. Discussion

The only reason which comes to mind for having a wide dark space is that the high electric field is only in the dark space and not in the plasma. The larger the dark space volume the greater the region from which to draw ions into the cathode substrate. If the boat to dark volume boundary distance is so reduced (the diffusion transport distance is small), the plating rate will be greater.

The metal shield stack ( $3^{\prime \prime}$ high x 4. $5^{\prime \prime}$ diamter) is now recognized to have a detrimental effect since it acts as a Faraday (shield) cage around the hearth zeroing the E field needed to pull the positive ions to the substrate.

Even with the above stack shield removed, the top of the crucible is surrounded on 3 sides by a $1 \frac{1^{\prime \prime}}{4}$ high by $3^{\prime \prime}$ wide magnetized rectangle having 3 sides $1 \frac{1}{4} "$ above the top of the crucible (about $2^{\prime \prime}$ above the crucible bottom). This shielding could reduce the local electric field near the melted source enough to dramatically reduce the coating rate. The Sloan $2700^{\circ}$ evaporation hearth and gun were designed for vacuum evaporation and not for ion plating; thus, the irregular top
profile is not unexpected．The shielding rectangle arms are magnetized and could not be ground down or removed without a magnetic compensation．Perhaps a large external Helmholtz coil could hopefully compensate and permit removal．

For both ion plating systems shown in ref． 9 and 10 the crucible cup top is flush with the flat baffle plate． Likewise in Teer＇s 21 ion plating unit，nothing is shown higher than the hearth．The top of the hearth is flush with the flat baffle plate．See Fig．1．The Heinz ${ }^{13}$ hearth is only slightly shielded by the horizontal gun．

The sparce $1 / 2 \mu$ coating on sample $⿰ ⿰ 三 丨 ⿰ 丨 三 11$ could be due to excessive shielding of the source．

## F．3．Vendor Search

On June 3rd letters requesting complete brochures with prices on electron beam evaporation ion plating systems and components were sent to all 16 vendors listed under Vacuum Coating Equipment on the Science Instrumentation Index．Not that a new system purchase is expected but that new components might be acquired for a much－needed upgrading of the V．T．A． system．

Denton Vacuum offered a complete unit with a DEG－801 Gun on the DIP－1 ion plating unit．

Hughes uses extensively，but rarely builds，ion plating units for customers．Mr．Cristy at 213／648－2345，Ext．84369， may however be of help．

Nanotech＊literature and prices just arrived listing a wide range of ion plating units with accessories．In fact they offer an option much like NASA＇s V．T．A．system except the e－beam gun is a Sloan 8 kW or 12 kW （costing about $\$ 31,000$ ）instead of our Sloan 5 kW ．Unfortunately no gun mounting drawings were sent．Sloan no longer lists the 5 kW （Five／Ten）power supply．Since the V．T．A．system hearth is rated 12 kW it may be underpowered．

Inflicon Leybold－Heraeus responded indicating no ion plating systems for sale．

## F．4．Films Plated

a．e－beam generated films．
All samples were 非 4340 hot roll steel disk one inch in diameter by $1 / 4$ inch thick．Samples 非3，非4，and 非11 were
＊Prestwich，Manchester M258WD England
Phone：061－773 8514

| TABLE I（Accuracy $\pm 30 \%$ ） |  |  |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Sample | Coating Date |  | $\begin{aligned} & \frac{\text { Back }}{\text { Front }} \\ & (\mu(\mu) \\ & \text { AI Film } \\ & \text { Thickness } \end{aligned}$ | $\begin{aligned} & (\ln ) \\ & \text { Elect. } \\ & \text { Spacing } \end{aligned}$ | Shield <br> Used <br> Hearth？ <br> Cathode？ | $(\underset{\Delta ⿴ 囗}{(\mathrm{mg})}$ | $\underset{\text { Material }}{\text { Film }}$ | $\begin{aligned} & \text { Front u" } \\ & \text { Surf. Smi } \\ & \hline \text { Subst. } \end{aligned}$ | $\begin{gathered} \mathbf{r} \cdot \mathrm{m} . \boldsymbol{s} \\ \text { othnes } \\ \hline \text { FIIm } \end{gathered}$ | $\begin{aligned} & \text { 會 } \\ & \text { 20 } \end{aligned}$ | $\begin{aligned} & \text { 合 } \\ & \text { 总 } \\ & \text { 見 } \end{aligned}$ |  | Comments |
| 3 | 6 June | 5 | $\overline{8 \pm 2}$ | 3－4 | Not <br> Cathode |  | Al Marz Grade | 100 | 100 | （ 4 | － | － | Numbers <br> listed are <br> estimates |
| 4 | 12 June |  | 2.5 | 5－6 | Hearth \＆ Cathode | －． 1 | $\frac{\mathrm{Brass}}{\mathrm{Al}}$ | 100 | 100 | $\left.{ }^{1} 200\right)$ | － | － |  |
| 11 | 25 June | 24 | $\frac{0}{.5}$ | －8 |  | ． 3 | Marz Al | 5 | 5 | 1.5 | （ $\begin{gathered}17 \\ (10 \\ 35\end{gathered}$ | ${ }_{2}$ to |  |
| 14 | 31 July | 30 | 8 | 2 | None | 20 | Marz 11 | 5 | 40 | （20－5） | 20 |  | Resiatance evap． |
| 12 | 1 Aug | 30 | $\Gamma 2$ | 2 | None | 14 | Marz Al | 5 | 30 | 3.5 （10－ $20)$ | 60 |  | Resiatance evap． |

ion coated with Al during trial runs of the V．T．A． 7375 system．Samples $\# \# 3$ and $⿰ ⿰ 三 丨 ⿰ 丨 三 4$ had surface roughness of $\sim 100 \mu$ while 非11（12， 13 and 14）were polished to a $\sim 5 \mu^{\prime \prime}$ smooth finish on one side．Film thickness on samples 3，4，and 11 were～8， 2.5 and ． $5 \mu \mathrm{~m}$ respectively．Some direct evaporation was possible on $⿰ ⿰ 三 丨 ⿰ 丨 三 一$ 3 where the recorded pressure was in the $5 \mu$ range．

Sample 非3 was mounted on a bare unshielded 1／2＂ rod．Sample 非 4 used the $5^{\prime \prime}$ cathode and grounded shield but was spaced away by about $3^{\prime \prime}$ by the brass interface．Sample非11 was flat mounted against the cooled cathode（to avoid heat induced blackening）and had an extra $2^{\prime \prime}$ of grounding skirt．Samples 非4 and 非11 also had a grounding cup shield around the hearth．Sample 非l had further shielding in the form of a brim ring around the cup top．
b．While the V．T．A．e－beam evaporation system was being repaired，the Denton resistance heated filaments were used to ion plate $8 \mu$ and $12 \mu \mathrm{~A}$ coatings on samples 非14 and 非12， respectively．The unshielded sample cathode was screw mounted to a vertical $1 / 4^{\prime \prime}$ steel rod and positioned $2^{\prime \prime}$ above the filament．
c．Film surface roughness is several times greater than substrate roughness on samples 非12 and 非14．
d．The Al was M．R．C．Marz grade while the crucibles were Carbon EB－9 Union 76 （POCO）．

F．5．Discussion
a．Samples 非3，非4，and 非11 were coated under explora－ tory conditions with the e－beam power being gradually increased．The cathode voltage was adjusted to prevent arcing in the chamber．The above variations plus the fact that only three films were coated suggests that only tenta－ tive conclusions should be drawn．The worst conditions （lack of shielding and close substrate source distance）gave the greatest film thickness．
b．Holliday $11 \& 12$ and this author observed that their ion plated films were rather dull．Jones，Griffith，and Williams ${ }^{23}$ ，ion plating on plastic，also found reflectivities reduced relative to vacuum coating，by about $2.5 \%, 10 \%, 18 \%$ for $\mathrm{Au}, \mathrm{Al}$ ，and Cu ，respectively．Could the trace contami－ nants always present ${ }^{22}$ from e－beam evaporation hearth liners in vacuum plating Al account for the decrease reflectivity？ Perhaps ion impact induced surface roughness makes the sur－ face less specular（smooth）．If one found that plasma cleaning on a $100 \%$ reflecting mirror surface causes loss of reflectivity in a lab experiment then the dullness in our films could be explained．The $\mu$ sized column film structure is common in ion plating and thus could easily account for the depressed reflectivities．
c. The trouble with the V.T.A. system may be inefficiency due to lack of optimum adjustment and that troubles arise as increased high powers are used to compensate for the inefficiency. The Al was easily melted with .6 kW power (only $12 \%$ of the rated full power of 5 kW ). The highest power recorded was 2.2 kW which was held for the last 5 minutes before the system blinked out 25 June. The A1 melt temperature was considerably above the melting point.
G.l. Suggested improvements which may be incorporated into the V.T.A. Model 7375 electron beam evaporation ion plating system are listed:

## a. Baffle plate improvements

During the past two years previous prospective V.T.A. system users ${ }^{12}$ designed a complex aluminum baffle plate using a small orifice which does an excellent job of providing the correct macroscope gauge pressure, both for the plasma above and for the filament chamber below the plate. The following further improvements are recommended for this baffle to lower further the filament pressure.
(1) Remachine this baffle totally from stainless steel instead of aluminum to match more closely the refractory materials (tungsten and molybdenum) in the Sloan e-beam gun. The present Al readily melts and evaporates around the Al orifice casting a direct shadow in the lower high vacuum region onto the gun and onto the high voltage electrodes causing shorts and arcs.
(2) Place stainless steel shim stock over the present Al baffle in the heat sensitive areas. This has been done.
(3) Bore a $1^{\prime \prime}$ diameter hole in the Al plate centered on the orifice hole. Machine a set of 7 stainless steel seats with orifice sizes $1 / 8,3 / 16,1 / 4,5 / 16,3 / 8,7 / 16$, $1 / 2$. The smaller sizes would suffice for material which requires no sweep and would provide a higher vacuum around the filament.
(4) Use only a very small quantity of vacuum grease on the 0-rings. Use only enough to give a wet look but not enough for the grease to be visible.
(5) A rectangular orifice may enhance the electron beam current while maximizing the gas flow impedance.
(6) The above rectangles could be continuously adjustable slits. If made of a bimetallic strip the adjustment would be automatic.
(7) Add extra fins and shields around the e-beam path to shield the exposed insulators from unintentional e-beam induced evaporation from around the orifice.
(8) The orifice could be placed with its axis horizontal at the top of the beam's arch, thus causing the gas to jet $180^{\circ}$ away from the filament and anode, instead of the present $90^{\circ}$ jet. A secondary advantage would be that more of the beam's path would be in the higher vacuum. A disadvantage would be increased coating around the orifice.
(9) Again the orifice is only an $\sim$ inch from the filament while the ion gauge is about $10^{\prime \prime}$ away. A gauge should be placed within an inch of the orifice with its short tube pointed toward the orifice jet to get the best indication of the filament pressure.
b. Improving other parts of the V.T.A. system other than the baffle plate:
(1) Place low melting point low vapor pressure materials on parts of the vacuum system for overheating indicators for sensitive components such as magnetic coils which must operate under $400^{\circ} \mathrm{C}$.
(2) Purchase at least all the parts listed on the suggested spare parts inventory (cost about \$2000).
(3) Add a shutter to shield initial volatile outgassing impurities from the substrate during the high vacuum soak of the crucible and source.
(4) In view of the manpower shortage connect a $4-$ pen recorder to the ion plating system to monitor the cathode voltage and current and the electron beam (emission) current. The fourth pen could monitor such quantities as sample temperature by thermocouple or magnet current.
(5) Dr. Arthur Nunes suggested using a focusing current ring (or an electrostatic ring) around the orifice to contact the beam to a finer thread to minimize orifice material evaporation.
(6) Apply the magnetic field sweep above the orifice so as to keep the orifice size as small as possible.
(7) Use a double hearth liner crucible stack with the inner crucible cut off to provide a lower target which should, because of greater insulation, rise to a higher temperature with less power to provide a higher plating rate.
(8) Entering the Ar leak from the bottom (instead of from the top) would cause some upflow which could enhance the transport of metal to the substrate increasing the deposition rate.
(9) Even with the cathode in its lowest position it is still ~ $6^{\prime \prime}$ from the hearth. Extending the substrate to the $2^{\prime \prime}$ to $3^{\prime \prime}$ range would produce a good enhancement in the deposition rate at no loss of throwing power according to Prof. Holliday's experimental data. 11 Suggested optimum distance for best ion plating is about six inches between substrate and source so some loss of adhesion may be expected.
(10) Switch the two baffle O-rings to Viton to assure better high temperature performance.
(11) The V.T.A. 15 KV 500 ma cathode supply was found inoperative due to both a blown fuse and a blown milliampmeter. An 8 amp fuse was found necessary to withstand the current surges during critical arcing while providing some meter protection. This recently added fuse holder should be upgraded to a 15 KV rating.
G.2. Major additions to the V.T.A. Model 7375 ion plating system.
a. Acquire a "Balzers" e-beam gun and hearth which measures the positive ion current which is "proportional to the ion plating rate." Geometry dependent coating rates with ion plating are difficult to determine compared to vacuum evaporation.
b. Purchase an identical Sloan e-beam evaporation unit so that one unit may be refurbished while the other unit is in operation.
c. Use boronnitride-titanium diboride ribbon evaporator boats with resistance heating of Al as done by Walley and Cross 14 in the article "An Aluminum Evaporation Source for Ion Plating." The electrodes must be water cooled and conducting graphite paper provides better electrical contact to the electrodes. This $\mathrm{BN}^{2}-\mathrm{TB}_{2}$ is cited in the Handbook of Thin Film Technology, page 1-47, as a crucible material sold under the name (HDA Composite Ceramic) by Union Carbide, New York, NY.

## G.3. Conclusions

The V.T.A. system shows promise in that it operated for about a month with no major problems. The V.T.A. low coating rate will probably improve when the excess electrostatic shielding is reduced. However for quick ion plating of Al, Denton resistance evaporation unit is much more reliable.

## REFERENCES

1. Mattox, D. M., "Film Deposition Using Accelerated Ions," Rept. No. SC-DR-281-63, Sandia Corp., (Nov., 1963); Mattox, D. M., Electrochem. Technol., 2295 (1964).
2. "Versatile Aluminum Coating Controls Corrosion," Iron Age, (February 9, 1976).
3. Spalvins, T., "The New Applications of Sputtering and Ion Plating," NASA TMX-73551 (1977).
4. Chambers, D. L. and Carmichael, D. C., Research/Development, 22, 32 (1971).
5. Spalvins, T., J. Am. Soc. Lub. Eng., 27, 40 (1971).
6. Mattox, D. M., Proceedings of the Conference on Sputtering and Ion Plating, NASA SP-5111, NASA Lewis Research Center, Cleveland, Ohio, p. 41 (1972).
7. Mattox, D. M., J. Vac. Sci. Technol., 10, 47 (1973).
8. Aisenburg, S. and Chabot, R. W., J. Vac. Sci. Technol., 10, 194 (1973).
9. Smith, H. R., "Current Developments in Ion Plating," Society of Vacuum Coaters Conference, Detroit, Michigan, March 23, 1972.
10. Harker, H. R. and Hill, R. J., "Deposition of Multicomponent Phases by Ion Plating," Journal of Vacuum Science and Technology, v.9, 非6, p. 1395 (1972).
11. Holliday, R. J., "Ion Plating Investigation," BER Report No. 224-94, The University of Alabama, p. X111 (Sept. 1978).
12. Holliday, R. J., "Ion Plating Studies for High Temperature Applications," University of Alabama in Huntsville, Report No. 227, p. Xlll (Oct. 1979).
13. Heinz, B. and Kithel, G., "Ion Plating with Electron Beam Evaporation," Conference on Ion Plating and Allied Techniques, p. 73, Edinburgh, 1977.
14. Walley, P. P. and Cross, K. B., Ibid., p. 82.
15. Maisse1, L. I. and Glang, R., Handbook of Thin Film Technology, 1970, McGraw Hill, New York.
16. Spalvins, T., Survey of Ion Plating Sources, Amer. Vac. Soc. 26th National Symp., New York, NY, (Oct. 2-5, 1979).
17. Wan, C. J., Chambers, D. L., and Carmichael, D. C., J. Vac. Sci. Technology 8, Vm. 99 - Vm. 104 (1971)
18. Spalvins, T. and Brainard, W. A., NASA TMX-3330 (1976).
19. Masatoshi, Y., et. al., "Studies on Ion Plating Process" for National Symposium and Exhibition, San Francisco, CA, Society for the Advancement of Material and Process Engineering, (May 1979).
20. Cappelli, P. G., "Coating Processes," Proceedings of the Conference on High Temperature Alloys for Gas Turbines, p. 177 of Applied Science Publishers (1978).
21. Teers, D. G., "The Ion Plating Technique," p. 13, op. cit. Ref. 13.
22. Graper, E. B., 'Evaporation (Vacuum) of A1 from an Electron Beam Source," Journal of Vacuum Science and Technology, g 非1, 33 (1971).
23. Jones, K., Griffith, A. J., and Williams, E. W., p. 115, op. cit. Ref. 13.
'NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM
MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA

STATISTICAL WIND PROFILE GUST MODEL

## Prepared by:

Academic Rank:

## University and Department:

## NASA/MSFC:

 (Laboratory)(Division)

MSFC Counterpart:
Date:
Contract No. :

Deva C. Doss, Ph. D.
Professor
The University of Alabama in Huntsville Department of Mathematics

## Space Sciences

 Atmospheric SciencesO. E. Smith

August 1, 1980
NGT-01-002-099
(University of Alabama)

## BY

Deva C. Doss, Ph.D. Professor of Mathematics<br>The University of Alabama in Huntsville Huntsville, Alabama

## ABSTRACT

The purpose of this study is to develop a statistical wind profile gust model for the Space Transportation Operations and Trade Studies by using 1800 Jimsphere wind profile data collected a+ Cape Kennedy during 1965-1972. Wind profiles from the surface to 20 km in component form, i.e., zonal and meridional are processed through the digital filters of different wave length ranges based on the MartinGraham cosine rolloff model. The residuals obtained from the filtering processes form the data base for the statistical analysis.

For each wind component the gust and gust length at a specified reference altitude in a residual profile are defined. A two-parameter gamma probability marginal distribution seems to fit well the component gust amplitude, and the gust length when redefined. The problem of finding an appropriate bivariate joint distribution of the gust amplitude and length remains to be solved.

The probability distribution of the modulus of the gust amplitudes has been derived under the assumption that they are independently distributed as gamma variates. It seems to fit the observed data.

Accurate wind profile measurements from the surface of the earth to 20 km are made possible by the Jimsphere-radar system. This system consists of a 2 m diameter, constant-volume balloon and an AN/FPS'-16 or equivalent high precision radar which tracks the position of the balloon. The position data collected by the system are smoothened over a 25 m interval. The resulting data are differenced over 50 m intervals to produce wind profile data points.

This study utilizes 1800 wind profile data in component form collected at Cape Kennedy during 1965-1972. They are processed through the digital filters of different wave length ranges based on the Martin-Graham cosine rolloff model. The residuals obtained from the filtering processes form the data base for the statistical analysis (see Figure l).

### 1.1 Definitions

Adelfang [1] in a recent report defines gust amplitude and length for a wind component in a residual profile at a specified reference altitude $H_{0}$. For instance in Figure 2 the gust amplitude (gust) for the zonal wind component $u^{\prime}$ is defined as the maximum value of $\left|u^{\prime}\right|$ in the vicinity of altitude $H_{0}$ with like sign to $u^{\prime}$ at $H_{o}$. The gust length $L_{u}$ is defined as the altitude difference of the zero crossings on either side of the gust, i. e.,

$$
\begin{equation*}
\mathrm{L}=\mathrm{H}_{2}-\mathrm{H}_{1} \tag{1}
\end{equation*}
$$

where $\mathrm{H}_{1}=$ altitude of the first zero crossing for the downward scan,

$$
H_{2}=\text { altitude of the first zero crossing for the upward scan. }
$$

Similarly, the gust amplitude and length are defined for the meridional gust component $\mathrm{v}^{\prime}$.

### 1.2 Purpose of the Study

One of the objectives of this study is to develop a theoretical probability gust model for each component as well as a joint model for both the components. Then the probability distribution of the modulus of the gust amplitudes can be derived.

In statistical literature it is not uncommon to fit a gamma probability


Figure 1. Cape Kennedy Residual Profiles

model to a climatological measurement. However, when more than one random variable is involved, it is not yet clear whether or not a particular species of several multivariate gamma probability distributions is adequate to describe the behavior of random variables under consideration. This difficulty stems from the fact that any given univariate marginal distributions do not uniquely determine their joint probability distribution. If one marginal of a component of a random vector happens to be different from that of another component, egg., gamma and beta, then the complexity of reconstructing their joint distribution increases.

### 2.1 Gust Amplitude

Adelfang [1] has shown that the gust amplitude of any wind component at a given reference altitude for filters of different wave length ranges seems to be distributed as a two-parameter gamma distribution and the fit seems to be good. A gamma probability density of a random variable $x$ is here defined as

$$
\begin{equation*}
f(x)=\beta^{\gamma} x^{\gamma-1} \exp (-\beta x) / \Gamma(\gamma), x \geqslant 0, \tag{2}
\end{equation*}
$$

$$
\because 0 \quad \text { otherwise, }
$$

where scale parameter $\beta$ and shape parameter $\gamma$ are positive and estimated by the maximum likelihood procedure as in Whom [3]

$$
\begin{equation*}
\hat{\beta}=\hat{\gamma} / \bar{x} \tag{3}
\end{equation*}
$$

and

$$
\begin{align*}
& \hat{\gamma}=\left\{1+(1+4 A / 3)^{1 / 2}\right\} / 4 A  \tag{4}\\
& \hat{A}=\overline{\ln (\bar{x})-\overline{\ln (x)}} \tag{5}
\end{align*}
$$

### 2.2 Gust Length

Although the gamma and truncated normal probability distributions seem to fit fairly the data on gust length, a pronounced bimodality appears to persist in the observed frequency distribution of the gust length. Therefore, there are two options to pursue:

1. To fit a mixture of two distributions
2. To redefine the gust length and hopefully avoid bimodality.

Since the first option involves computation of more parameters and analytical complexity, a redefinition of the gust length is in order. Instead of defining the gust length as the distance between two successive zero crossings, it can be defined as the length of the interval containing the reference altitude $H_{0}$, whose end points form successive extrema of the wind profile. In Figure 2, the distance from $A$ to $B$ is the new gust length for $u^{\prime}$ and the distance from $C$ to $D$ is the new gust length for $v^{\prime}$.

It is encouraging to find that newly defined gust length seems to fit a gamma probability distribution without exhibiting any bimodality. It is premature at present to accept this observation without further investigation.

## 2. 3 Joint Distribution of Gust Amplitude and Length

Once it is assured that the gust length follows a gamma model a search for an appropriate bivariate gamma distribution should be made for the gust amplitude and length. Three bivariate gamma distributions are given in Mardia [2], p. 88. When the shape parameters of gamma marginals are almost equal, the first bivariate distribution involving a modified Bessel function may be considered. Because of the restriction, namely, $y \geqslant x$ imposed on the third distribution, it cannot be employed in fitting the gust amplitude and length.

## 2. 4 The Modulus of the Gust. Amplitude

Since it is found that the correlation between the gust amplitudes of the components $u^{\prime}$ and $v^{\prime}$ is less than 0.2 , the maximum of modulus

$$
\begin{equation*}
R=\left\{\left(u^{\prime}\right)^{2}+\left(v^{\prime}\right)^{2}\right\}^{1 / 2} \tag{6}
\end{equation*}
$$

can be derived from the gamma marginal of the gust amplitudes under the assumption that they are independently distributed. In fact, the probability density of $R$ is:

$$
\begin{equation*}
\frac{1}{4} \frac{\beta_{1}^{\gamma_{1}} \beta_{2}^{\gamma_{2}} R^{\gamma_{1}+\gamma_{2}}}{\Gamma\left(\gamma_{1}\right) \Gamma\left(\gamma_{2}\right)} \sum_{n=0}^{\infty} \frac{\left(-\beta_{1} R\right)^{n}}{n!} \sum_{m=0}^{\infty} \frac{\left(-\beta_{2} R\right)^{m} \Gamma\left(\frac{\gamma_{1}+n}{2}\right) \Gamma\left(\frac{\gamma_{2}+n}{2}\right)}{m!\Gamma\left(\frac{Y_{1}+\gamma_{2}+n+m+2}{2}\right)} \tag{7}
\end{equation*}
$$

where all parameters $\beta_{1}, \beta_{2}, \gamma_{1}, \gamma_{2}$ are positive. This has to be validated by the data.

### 2.5 Sum of Two Adjacent Amplitudes

Denote the amplitude of the zonal component at the reference altitude $\mathrm{H}_{\mathrm{o}}$ by $\left|u^{\prime}\right|$ and the amplitude just above by $\left|u^{\prime}{ }_{2}\right|$. Then it has been found for a set of data that the correlation between them is in the vicinity of 0.5. However,
the observed data on their sum $\left|u_{1}^{\prime}\right|+\left|u_{2}^{\prime}\right|$ seems to fit a gamma distribution. (Similar observation may be true also for the other component.) It raises a theoretical question: When is the sum of two correlated gamma random variates distributed as a gamma?

## CONCLUSIONS AND RECOMMENDATIONS

This report presents several unanswered questions to which answers have to be sought. It is, with reasonable assurance, seen that the gust. amplitude follows a two-parameter gamma distribution the probability model of the newly defined gust length seems promisingly to be a gamma distribution.

The joint distribution of the gust amplitude and length, and the distribution of the modulus of gust amplitudes need to be explored. In fact finding the joint distribution of the gust amplitudes and length, of both the components should be the ultimate goal. One realizes that there are many loopholes in this investigation, but it is hoped that by this time next year the statistical model for the gust is resolved.

## REFERENCES

1. Adelfang, S. $I_{n}$, "Vector Wind Profile Gust Model." Computer Sciences Corporation Final Report for NASA/MSFC under Contract NAS8-33433, April 9, 1980.
2. Mardia, K. V., "Families of Bivariate Distributions." Griffin's Statistical Monographs and Courses. Hafner Publishing Co., Darrien, Conn. 1970
3. Thom, H. C. S., "Some Methods of Climatological Analysis." WMO Technical Note 81, WMO-MO 199. TP. 103, 1966

Prepared By:
Academic Rank:
University and Department

NASA/MSFC:
(Laboratory)
(Division)
(Branch)
MSEC Countorpart:
Datr:
Contract Nu .

Robert M. Eastman, Ph.D., P.E
Professor
University of Missouri - Columbia
Department of Industrial Engineering

Electronics and Control
Guidance, Control \& Instrumentation Control Mechanisms

Kenneth R. Fernandez
July 25, 1980
NGT-01-002-099
(University of Nlabama)

# UPGRADING THE FREE FLYING RENDEZVOUS AND DOCKING SIMULATOR AND THE ORBITAL SERVICER SYSTEM 

BY<br>Robert M. Eastman, PH.D., P.E. Professor of Industrial Engineering University of Missouri Columbia, Missouri

## ABSTRACT

The objective of the research was to make recommendations for upgrading two teleoperator/robotics test and simulation systems based upon a review of latest technology advances in the involved disciplines. The Free Flying Rendezvous and Docking Simulator has two vehicles, the Free Flying Mobility Unit and the Free Flying Target Assembly, operating on air pads on the air bearing epoxy flat floor. The Orbital Servicer System has a six degrees-of-freedom manipulator arm for transferring modules between the Orbital Servicer and a prepared satellite.

Based on the research, a new second-generation Free Flying Mobility Unit is recommended. It will add a sixth degree-of-freedom and incorporate other improvements which will greatly expand the Center's capability to perform evaluation tests and demonstrations of advanced systems concepts for rendezvous and docking in support of the Teleoperator Maneuvering System (TMS) Program. Current plans for the TMS call for a technology readiness in 1985 for placement and retrieval of a satellite/payload in a predetermined orbit with the TMS moving from and returning to the Shuttle Orbiter. The second generation Free Flying Mobility Unit will incorporate TMS design features which will enable it to perform realistic simulations for evaluating rendezvous concepts, ranging sensors, lighting and video systems, and docking mechanisms. At a later time it will incorporate advanced technology features for future TMS missions and will be invaluable in defining the requirements and evaluating promising concepts for automated rendezvous and docking.

The Orbital Servicer System provides the capability for testing and demonstrating concepts for on - orbit servicing of compatibly designed satellites/payloads. The TMS will be the transporting vehicle for the servicer. The manipulator arm of the Orbital Servicer System is presently computer controlled in the trajectory portion of the module transfer operation. The ultimate objective is to fully automate its operation having the manipulator arm move from its stowed position, latch on to a mock-up faulty module, remove the module and stow it in the base of the servicer, latch on to a replacement module, place it in the satellite, and having the manipulator arm return to its stowed position. This will require additional capabilities in several fields basic to robotics research. These include sensors, artificial intelligence; image analysis, communications, computer programming, pattern recognition, kinematics and manipulator design. It is recommended that the Electronics and Control Laboratory move to acquire the basic competencies in robotics necessary to achieve the major objective of full automation.

## ACKNOWLEDGMENTS

My gratitude to the many NASA people who helped with the research and contributed to our pleasant summer in Huntsville.

Kenneth R. Fernandez as my NASA MSFC counterpart helped with explanation and demonstration of the robotics facilities and test programs. John B. E. Chase was generous with his cditorial help, arrangements, and explanation of the factors and issues involved. Ed Guerin and Dave Henderson (Essex Corporation) explained the Free Flying Rendezvous and Docking Simulator and arranged for me to observe the tests. Don Scott and Keith Clark demonstrated the Orbital Servicer System. Charles Cornelius provided logistic support and organizational enlightenment. F. Brooks Moore and Eugenc Smith furnished the management support without which no research can prosper. Other members of the Electronics and Control Laboratory helped with information and support.

Mary F. Smith typed and retyped the manuscript and helped in numerous other ways. Eloise Meyer, Bettye Morey and Gayle Dodson were gracious and helpful. Marion Kent's advice was most useful and appreciated.

Dr. Robert Barfield, Project Director, handled the contract substance assisted by Barbara Skaggs. The University of Alabama in Huntsville faculty were gracious hosts socially and professionally.

## LIST OF FIGURES

Figure No.TitlePage
1 Free Flying Rendezvous and ..... VII - 2Docking Simulator
2 Free Flying Mobility Unit ..... VII. - 3
3 Free Flying Target Assembly ..... VII - 4
4 Orbital Servicer System ..... VII - 5
5 Conceptual Sketch of Overhead ..... VII - 16 Mounted Secind-Generation Free Flying Mobility Unit
6 Conceptual Sketch Floor ..... VII - 18
Mounted Second-Generation Free Flying Mobility Unit

## INTR ODUCTION

The Free Flying Rendezvous and Docking Simulator consists of the Free Flying Mobility Unit and the Free Flying Target Assembly both of which operate on the air bearing epoxy flat floor. The two vehicles are mounted on air pads which provide frictionless motion in one plane for simulating translational movement in space. Simulated propulsion is provided by compressed air jets under the control of the remotely located operator. The simulator is illustrated in Figures 1, 2, and 3 and described in reference (1).

The simulator is an engineering tool used to define, test, evaluate and demonstrate the rendezvous and docking techniques and procedures, and the components and systems designs needed for rendezvous and docking a remotely controlled teleoperator vehicle with a satellite payload. The data recorded include air consumed (measure of fuel consumption), time to dock, docking success, lighting patterns and camera positions. The evaluation and test data influence the selection of the most promising concepts and designs and the refinement of them for potential use in future projects.

The Orbital Servicer System is an engineering model of an on-orbit maintenance system. It contains full-scale typical satellite replacement modules which are exchanged by the six-degree-of-freedom manipulator arm. It also includes the interfacing control system for the manipulator arm. The Servicer System model is used to test, evaluate, and demonstrate the on-orbit maintenance concept and to define design features that would enhance its performance. At present (1980), the operation is partially automated. A computer program controls the teleoperator arm trajectory movement until the module is close to its destination. Then control is transferred to the human operator who completes the module transfer. He operates from a remote location using video cameras and manual controls. The servicer is illustrated in Figure 4 and described in references (2) and (3).

One of the key questions concerning the Orbital Servicer is whether the teleoperator arm can be fully automated to eliminate the human operator (man-in-the-loop). There are advantages and disadvantages to each alternative and strong advocates on each side. A major purpose of the Orbital Servicer System is testing the feasibility of the fully-automated transfer.


FIGURE 1. FREE FLYING RENDEZVOUS AND DOCKING SIMULATOR

FREE FLYING MOBILITY UNIT
FIGURE 2.


FIGURE 3. FREE FLYING TARGET ASSEMBLY


FIGURE 4. ORBITAL SERVICER SYSTEM

## OBJECTIVES

The first research objective is to recommend design improvement for the Free Flying Rendezvous and Docking Simulator based upon a review of latest developments in associated disciplines. The subsidiary objectives are:

1. A 6 DOF Free Flying Mobility Unit which could be either a modification of the existing 5 DOF vehicle or a new vehicle.
2. Automatic docking capability.
3. Alleviation or elimination of some current limitations including:
(a) Reliability and maintenance problems with existing equipment.
(b) Physical balance of the current equipment, i.e., stability in pitch and roll.

The second research objective is to recommend design improvements in Orbital Servicer System based on review of latest developments in appropriate disciplines. Subsidiary objectives are:

1. Automated fast servicing of a class of satellites by robotic arm from the Orbital Servicer vehicle.
2. To review the proposed new electrical self-aligning connector design.
3. To suggest improved methods of rack positioning.

METHOD OF INVESTIGATION

The following methods of investigation were used:

1. Literature search.
2. Material and information from the NSF Robotics Research Workshop held April 15-17, 1980, at the University of Rhode Is land and from personnel at the McDonnell Aircraft Company (part of McDonnell Douglas Corporation in St. Louis, MO).
3. Observation of the Free Flying Rendezvous and Docking Simulator and the Orbital Servicer System.
4. Discussions with MSFC personnel.
5. Sketches and calculations.
6. Review of previous and associated work.

> SYSTEMS INTEGRATION

NASA Headquarters has assigned MSFC the mission of "Systems Integration" or the co-ordination and interfacing of teleoperator/robotics sub-systems and components into an effective total system to support future NASA programs. The term "System Integration" also connotes optimizing the effectiveness of the system while using scarce resources efficiently.

The two lab facilities, the Free Flying Rendezvous and Docking Simulator and the Orbital Servicer System, are essential for evaluating and demonstrating the integration of the functions and the system performance in a simulated space mission. The full-scale dimension provides the experimental setting necessary for accurate test and evaluation. Scale models can be helpful but there are too many possibilities for discrepancies between actual mission systems and part-scale simulators to permit complete reliance on the part-scale model. Computer modeling and simulation can be efficient and time-saving; however, their results require full-scale physical verification because of the possibilities of conceptual errors and data omissions and inadequacies.

## BASIC ROBOTICS RESEARCH OVERVIEW

This section will be devoted to a discussion, necessarily brief, of several fields of basic robotics research pertinent to the main objectives of full automation of the Orbital Servicer Teleoperator System and to upgrading the Free Flying Rendezvous and Docking Simulator.

The following research areas are fundamental to advanced teleoperator and robotics development:

1. Artificial intelligence
2. Sensors
3. Kinematics
4. Control
5. Communications
6. Manipulator design
7. Image analysis
8. Locomotion

## ARTIFICIAL INTELLIGENCE

Artificial intelligence is the branch of computer science which studies the way machines can replicate human actions requiring intelligence. It involves receiving and evaluating environmental data and making a decision based upon the data and a set of decision criteria. At an advanced level, the machine can learn from experience and adjust its behavior accordingly. Artificial intelligence is necessary for any teleoperator or robot which has an operation which is not completely and undeviatingly predetermined.

Systems have been developed which will perform at least one but never all of the following: (4)
(1) Plans can be generated at multiple levels of detail.
(2) Plans can be viewed as partially ordered sequences of actions with respect to time.
(3) Each action is expected to produce a single state change characterized by a single primary effect.
(4) A plan is not generated at all unless the planner determines that it will be totally successful in meeting all specified goals.
(5) Simple plans requiring information gathering can be generated.
(6) Unsophisticated techniques for dynamic repair of unsuccessful plans during execution have been developed.
(7) Plans can be used to control robot devices with simple use of sensory feedback and systematic replanning.

Major research problem areas in intelligent robots are:

1. Planning for parallel execution
2. Planning for information gathering
3. Planning for planning
4. Learning: Using a plan data base
5. Interactive planning
6. Dynamic plan repair
7. Distributed robotics (i.c., disbots)
a. Hierarchal planning

## b. Control

c. Basis
d. Communication
e. Plan execution

The central problem of robotic intelligence in the real world is that a robot may have a large number of alternative actions, but the user does not want to be burdened with specifying them and the computer is too busy to compute them anyhow (5). ------"Intelligence and decision making" usually mean the ability of a robot to sense some aspect of its environment and take alternative courses of action.

## SENSORS

A sensor is a device which collects data from the environment for transmission to the robotics command system. An important research issue in robotics is sensory control in which the sensor is the first link in the chain which finishes with the completed robot action. Fast and accurate sensing is essential to reaching the goal of fully automated space robotics (6).

Among the research areas to be addressed are:

1. More capable visual processing
2. Faster visual processing
3. Better dynamic models of arms
4. Tactile and force sensors with desirable characteristics
5. Better understanding of the role of compliance

Visual sensing is important in space applications. Some needed improvements are (7):

1. Higher resolution, dimensional density and precision
2. Improved pixel (i.e., picture element)quality
3. Color discrimination
4. Better lenses
5. Better cameras
6. Increased robustness
7. Controlled illumination
8. Software development
9. Range and proximity sensing

$$
\text { VII - } 9
$$

Sensor types which have negligible usefulness for space functions are tactile, force, acoustic and temperature.

Radar as a sensor has advantages and disadvantages. The distances involved may be too short to permit radar to function properly. Even more difficult is the problem of interfering reflections from structural components of the system. These reflections will prevent the system from receiving clear and accurate readings on the position of any system component.

However, some current and research in radar may facilitate using radar in fully automated orbital servicer. It has the capability of rapid location of an object and determination of its velocity and direction of movement provided the reflection problem can be solved.

Video cameras produce a good picture for the human operator. The major problem is automated image analysis of the picture. (Image analysis is the conversion of visual pictures such as video to a set of impulses which can be processed and interpreted by a digital computer). The human operator is much better at analysing the video picture and taking appropriate action than is an automated system.

Lasers may be the best sensors for space robotics applications. Much additional research is needed. Light-beam sensors. (photoelectric cells) can not produce the detailed image needed. Acoustic sensors lack the medium to transmit sound wave. Contact sensors would require delicate, easily-damaged antennas and would probably be unable to gather enough environmental data to be useful.

Sensors present several robotic control problems such as data filtering and smoothing, multidirectional sensors and data acquisition. Control structures, particularly hierarchical, and control logic need further work. The man-machine interface requires additional development (12).

## KINEMATICS

Kinematics and mechanical design refer to the static and dynamic relationships of the physical components of the teleoperator manipulator system. Examples are mass, velocity, momentum, strength and deflection. Correct handling these factors is vital to successful teleoperator and robot operation. Several important research areas remain.

Kinematic preformance can be either time-based (velocity and acceleration) or geometric-based (space and orientation). Further work is needed in both areas (8). This effort should get away from position-byposition analysis to a broader knowledge which eliminates the need for the detailed position approach. A general solution to the optimal number of degrees-of-freedom has yet to be resolved. More DOF produce greater flexibility and versatility; fewer DOF produce a system easier to design and control, cheaper to build and generally more reliable.

Some robotic configurations are more efficient than others. However, the knowledge about these configurations is empirical rather than theoretical. Among the research questions to be resolved are:

1. The optimal manipulator configuration for flexibility, dexterity and reliability.
2. Actuator placement and control.
3. Basic design theory for fluid and electrical actuators.
4. Increasing the number of design parameters which can be rationally chosen given the present state of the art.
5. Elimination or minimization of elastic vibration due to system resonance.
6. Coping with large inertial forces at high speeds.

Another practical design problem is that robot and teleoperator performance fall short of theoretical optimum. Some causes are mechanical interference by structural members, limited rotational range and the mass of components particularly the motors (8).

Speed of operation of robots is vitally important to industrial applications, but less so to space applications. High-speed dynamic effects become critical yet little is known definitively about these effects. Present manipulator configuration design and trajectory are not the best for high speed operation. More needs to be done (9). Three difficulties with high-speed operation are:

1. Non-linear dynamic effects
2. Structural flexibility
3. Discrete time increments resulting from micro-computer controls produce undesirable effects.

Some modifications of manipulator design may be necessary to make the design compatible with automated operation. Present plans are based on the assumption that full automation can be achieved by rewriting the control computer program. This may or may not be a valid assumption. At the very least, the manipulator design should be reviewed for compatibility with full automated control.

## CONTROL

Control theory is the branch of engineering and science which deals with monitoring and direction of the teleoperator to achieve the desired result. Control theory has two facets:

1. Systems study and modeling and (2) measuring (sensing) the states of the system. System and task modeling are essential to advanced robot control development. Modeling can be defined as quantitative descriptions of the system and tasks in measurable, calculable and controllable terms.

Among the research issues are sensitivity studies of the model, robot control in feedback schemes and multi-unit robotic systems (10).

Four assumptions of robotics control theory are:

1. Robots are complex articulated mechanical devices designed to perform operations normally requiring human skills.
2. Robots are equipped with sensors capable of measuring internal and external states.
3. Robots are equipped with computers.
4. Humans can communicate with robots.

Paul discusses control applied to robotic operations in space. He postulates 6 DOF with either velocity or force but not both specified along each DOF. He also discusses kinematic control and Jacobian control. Each method has advantages (ll).

Graupe and Saridis hold that technological advances of the last fifty years depended upon development of system theoretic methodologies (14). Intelligent control uses the digital computer along with advanced techniques of system theory to produce a unified engineering approach. Based on this, control systems for robotic tasks and systems have been developed for prosthesis and orthotic devices and for general purpose robots. They propose a hierarchally intelligent control approach distributed according to "decreasing precision with increasing intelligence." The levels would be:

1. Organization
2. Co-ordination
3. Hardware Control

Present computer-based systems for robot control require the program to direct all movements and actions of the robot. A specially-designed computer language may be used although some programs are in a standard language such as Fortran. The program can include provision for receiving external stimuli and directing action based on the stimuli received. However, there is no provision for stimuli not in the program or for unprogrammed reactions no matter how desirable (15). Libraries of sub-routines can be developed since the programs can be independent os specific locations. The number and usefulness of sub-routines handling contingencies, errors and emergencies has increased.

The major deficiencies of the second-generation programmed robotic systems are (15):

1. The programs are difficult to debug and modify.
2. Achieving the final position with maximum accuracy is inefficient.
3. It is difficult to foresce all possible emergencies and contingencies. This is particularly important in the Space Shuttle - Orbiter Servicer system.

In the next (third) generation of computer robotics control systems, the robot will formulate an appropriate plan of action based on goal directions received from the operator. The robot will have greater artificial intelligence capabilities and much more data than it needs with the present computer program motion direction system. Research on this system is now underway (1980).

The elementary "Pick-and-place" robotic systems are relatively easy to design, program and implement. Many industrial robots are in this category. However, they are unable to handle effectively and reliably events and environments which deviate even slightly from the environment for which the program was designed. However, as the program is modified to handle environmental variations, the complexity of the problem domain and the resulting' program increase many-fold (13).

## COMMUNICATIONS

Communications will be a major robotics design problem. Data impulses from sensors must be transmitted to the control computer and instructions sent from the control center to the teleoperator components. Another major R\&T issue is the computer languages and control algorithms needed for full automation. This is inextricably interconnected to the determination and optimization of the desired operating characteristics.

## IMAGE ANALYSIS

The role of image analysis will be the identification and location of the components of the shuttle and the servicer. These components include:

1. Structural components of the Shuttle.
2. Structural components of the Servicer.
3. Alignment of the space vehicle (Servicer and satellite) during docking.
4. The modules at the beginning, during, and of the end of the transfer process.
5. Alignment of the rack which will receive the module.
6. Any component out of its regular position to the extent it might cause damage.
7. Any stray object in a position to cause a collision.

Image analysis is a particularly difficult field in robotics research and application. The processes by which a human operator's mind receives signals from his sensors, and evaluates, interprets and identifies the object seen, its location and its orientation are incredibly complex. Performing these processes automatically has been done only in simple situations with great difficulty.

Basic research in image analysis is going forward. Further progress is essential to the full automation of the Crbital Servicer Teleoperator.

## LOCOMOTION

Locomotion is defined as movement of the robot over an irregular terrain surface. This contrasts with space travel in which the robot has no terrain contact. At a cursory glance, locomotion has little to do with the NASA Space Shuttle and Orbital Servicer missions. However, some locomotion research results have application to the simulation of rendezvous and docking operations. Furthermore, locomotion could become important to the MSFC robotics efforts if the mission is changed to include mobile robots for space applications.

## THE FREE FLYING RENDEZVOUS AND DOCKING SIMULATOR

The present Free Flying Rendezvous and Docking Simulator uses aging equipment of limited capabilities. The two vehicles were assembled partly from surplus and scrounged components. They require considerable maintenance with increased down time and decreased capability. Further, the data output is limited in comparison to the data potential of a simulator with better equipment and instrumentation.

## Second - Generation Free Flying Mobility Unit

A major research and policy question is "Should NASA build a new, second-generation Free Flying Mobility Unit or should it rebuild and improve the present vehicle?" In either case, the vehicle should incorporate a sixth degrees-of-freedom ( 6 DOF ), balance stability in the pitch, roll and $Z$ axes, increased reliability and more data acquisition instrumentation.

The general specifications for a second-generation Free Flying Mobility Unit are:

1. Incorporate six degrees- of-freedom.
2. Maintaining the vehicle balance during docking maneuvers.
3. Minimize weight without interfering with operating characteristics.
4. Be compatible with existing equipment and existing data.
5. Provide for installing alternative docking mechanisms.
6. Have on-board computer capabilities.
7. Provides for lighting and camera alternatives.
8. Have good maintainability and reliability characteristics.
9. Have the necessary communication and control capabilities.
10. Permit modification for automated docking maneuvers.

## Advantages of a new second-generation unit are:

1. The desired 6th DOF can be incorporated into the design of the new vehicle more easily than into the existing one.
2. Some faults of the present design can be eliminated in the new vehicle design.
3. The reliability of a new vehicle would be much better.
4. Maintenance costs should be lower at least in the early years of its service life.
5. Vehicle balance can be incorporated into the design and operation.
6. Provision for alternative docking mechanisms can be installed.
7. Additional features such as improved lighting and vision can be incorporated.

Disadvantages of a new second-generation unit are:

1. The cost will be considerable. It may be more difficult to obtain the fund appropriation for a new vehicle than to rebuild the present unit with NASA personnel.
2. The procurement lead-time will probably be a minimum of three years between the date of the approval and the date the new vehicle goes into full service.
3. Unforeseen bugs and normal new model difficulties will arise.

Alternative Design Concepts for a Second-Generation Free Flying Mobility Unit
The following alternative design concepts were considered:

1. A vehicle suspended from a mobile overhead track-mounted vehicle.
2. A new floor mounted vehicle with its mass concentrated towards the center of gravity.
3. A floor-mounted vehicle with automated computer - controlled balancing.
4. A vehicle mounted on top of a two-way track system above the floor level. (Dr. Campell's concept) This would have some advantages in stability and simpler motion design. However, the mounting might cause interference problems and possible damage (16).

## Overhead Track-Mounted Vehicle

The concept of an overhead track vehicle with the docking mechanism suspended underneath is illustrated in Figure 5. The track and carrier will permit locating the mechanism in the $X$ and $Y$ directions. An actuator operated

column or beam would raise or lower the mechanism on the $Z$ axis. Pitch yaw and roll would be provided by movement about the lower column end. Communications could be by direct electrical connection rather than by radio.

The advantages of suspending the docking vehicle from an overhead track are:

1. The sixth degree-of-freedom will be easier to incorporate.
2. Some of the technology is known and well-established.
3. Some components may be completely or partially "off-the-shelf." This will reduce the lead-time and possibilities of trouble with new technology.
4. There will be fewer interference problems such as those with the bases of the two present vehicles.

A major disadvantage is:

1. Maintaining vehicle balance through all DOF movements. This will require either a complex mechanical system or a sophisticated computer control program.
The research and design problems include:
2. Adequate building structure which may be difficult to fund.
3. Kinematics and deflection.
4. Mass and momentum.
5. The control system.

## Floor-Mounted Free Flying Mobility Unit

A floor-mounted Free Flying Mobility Unit with computer-controlled balancing is a possibility. A sixth degree-of-freedom could be added. This would have the advantages of floor mounting while providing solutions to some major problems with the present model.

The major difficulty with this concept is the design of the control system and the computer programming. Both would be complex and expensive. Further, modifications to the design or latching mechanisms would require extensive redesign, reprogramming and modification.

A second-generation floor-mounted Free Flying Mobility Unit would be similar in some respects to the present model. The design concept is illustrated in Figure 6. A major change is the use of the three actuators to control the plane of the mechanism. This will provide more possibilities of movement and improve the units balance. Some of the other components can be relocated nearer the center of gravity to improve the balance. Other improvements can be incorporated.

CONCEPTUAL SKETCH FLOOR MOUNTED SECOND-GENERATION FREE FLYING MOBILITY UNIT


FIGURE 6.

## Floor-Mounted Unit with Automated Computer-Controlled Balance

Computer balance control will have the following advantages:

1. Proper design of the balancing system will hold down the mass of the vehicle.
2. The computers can be located remotely from the vehicle. This will hold down the complexity and mass of the vehicle itself.
3. Although trained computer personnel are presently scarce, the number in training who will become available in the near future is increasing substantially.

## Major difficulties with Mechanical Balance Control are:

1. Increase in the vehicle mass. This may cause difficulties in operation. The higher mass will increase the impact forces and may also aggrevate some imbalance situations.
2. Personnel capable of designing, building and maintaining mechanical devices are scarce and will become almost unavailable within a few years.

## Display System Definition

The present display system for the Free Flying Rendezvous and Docking Simulator needs a thorough review to determine whether it meets current requirements and is optimum for the test facility mission.

## FULLY-AUTOMATED ORBITAL SERVICER SYSTEM

The objective of full automation of the Orbital Servicer will require a closely coordinated research and test effort involving several advanced robotics disciplines. The result of successful effort will be the elimination of the man-in-the-loop requirement of the present system.

Full automation of the Orbital Servicer Module transfer operation involves complex difficult research issues. It is not enough to extend the present computer controlled trajectory to cover the final movement of the module into the rack and to complete the latching operation. Functions performed by the man-in-loop must be incorporated into the automation. These functions include collision avoidance, path adjustment, coordination of the teleoperator movement and others.

## In this discussion, the following assumptions are made:

1. The precision and accuracy, capabilities of the present Orbital Servicer robotic system are inadequate for efficient operation and collision avoidance in the full automation mode.
2. The major configuration features of the present Space Shuttle, Orbital Servicer and TMS will remain substantially unchanged.
3. Mass of the components and the assembly.
4. Oscillation and deflection.
5. Placement of sensors, controls, motors and similar parts.
6. Collision and damage avoidance features.

Application of Basic Robotics Research to the Orbital Servicer System
In the section "Basic Robotics Research Overview," several basic knowledge areas are discussed. In this section, application to the Orbital Servicer System will be covered.

The first major difference between the present partially automated, partially manual mode and the fully automated mode is the need for substantially increased precision and accuracy in the final approach, contact and latching. Extending the present computer control program will not provide enough accuracy to complete the final mating and latching successfully. In order to attain this accuracy, knowledge from several basic fields must be added.

The first addition will be sensors to determine accurately the position of the module with respect to the rack. The predicted accuracy with the present system is about $\pm 1 / 2$ inch. This is insufficient to insure proper completion of the maneuver.

A necessary function of the sensors will be to detect deviations from the planned or expected movements and locations. The human operator can recognize situations which deviate from plan and take action to abort or adjust the teleoperator arm action. If there is no man-in-the-loop, the teleoperator must be designed to recognize the situation and to take appropriate action.

The first situation the sensors must recognize is misplacement, deviation or absence of the rack, module or other system component. The other is the presence of an obstruction, accidental or not with which the teleoperator would collide if the programmed paths are continued.

At the time of this report (Summer 1980), it is infeasible to specify the type and technical details of the needed sensors. Several types are mentioned in the previous section.

After the sensors have collected data from the environment, image analysis is necessary to interpret the data for action by the control system. Data impulses received by the sensors must be converted into an information format upon which a decision can be based. This has turned out to be one of the most difficult fields in robotics research and application.

The human's ability to analyze and interpret data collected by his sensory organs is unbelievably complex and difficult to replicate. Yet the teleoperator system must perform this function if it is to carry out its mission.

After the teleoperator system has received and analyzed the environmental data decisions on the teleoperator action must be made by the system without human intervention. This is the domain of artificial intelligence, another basic field in robotics research.

Communications are essential to the teleoperator system operation. Data on the environment and the system must be transmitted to the analysis and control sub-system. Commands resulting from decisions must be transmitted to the functioning sub-systems and components. Accuracy, reliability, speed and other characteristics are important to mission success and efficient operation.

Control function provides the machine direction and monitors of the of the teleoperator system. Again replicating the process of human control in the fully-automated teleoperator system is an exceedingly complex and difficult task. However, modern control theory research has made considerable progress which will be helpful in this effort.

The kinematics and mechanical design of the teleoperator system and its components will require further research and technology. For example, each additional joint adds degrees of freedom. The number of degrees-of-freedom in the teleoperator exceeds the number which can be analyzed kinematically at the present level of knowledge. Further basic research is needed. Other problems concern the mass, momentum and impact characteristics of the teleoperator system and its components.

End effector design is an field of robotics research important to the Orbital Servicer Teleoperator system. Improvements are needed in the latching mechanism and in the latching and unlatching operations. One such improvement is computer control of the latching and unlatching as the module moves into and out of the rack.

Locomotion, a basic field of robotics research at some institutions, has little relevance to the current mission of neither the Orbital Servicer System nor the Free Flying Rendezvous and Docking Simulator. However, should the missions change, locomotion might become essential to mission success.

## Resolver Feedback

The use of resolver feedback for manipulator joint position sensing should be investigated. The current method is subject to inaccuracies due to nonlinearities in the potentiometers, grounding and noise, and variation in the reference voltages. The very accurate resolver feedback is especially needed for the successful implementation of the current dead-reckoning control system.

## ADDITIONAL ORBITAL SERVICER SYSTEM PROBLEMS

In addition to full automation of the Orbital Servicer TMS, there are three subsidiary problems in the assignment.

1. Latching mechanism operation.
2. Rack attachment.
3. Electrical self-aligning connector.

## Latching Mechanism Operation

The mechanism for latching and unlatching the modules needs improvement in operation and in coordination with the full automation trajectory mode.

The following constraints must be met:

1. Automated operation.
2. Positive latching.
3. No damage to the module or the mechanism of the manipulator arm.
4. No major modification of the present latching mechanism, teleoperator end effectors, components or the existing control program.

One alternative solution is a mechanical device actuated by contact or visual sensors. As the device takes over motion control, the computer control would be released. The mechanism would complete the latching operation. There are a number of difficult problems with this alternative such as the sensing device, powering the latching device and operational difficulties with . the mechanism and teleoperator.

A better solution appears to be computer operation and control of the latching operation from start to completion. The coordination problems will be less than with mixed mode control. Full computer control will produce weight savings over mechanical tatching devices. In the future, it will be easier to acquire personnel with the necessary computer and electronic skills than personnel with mechanical design, manufacturing and maintenance skills.

## Rack Attachment

The present method of attaching the racks to the satellite and to the Orbital Servicer allows some movement and variation in the rack position. This condition creates little difficulty with the manual mode of completing the trajectory. However, it produces major problems for full automation. Variation in the rack position makes completion of the trajectory by dead reckoning infeasible since this operating mode requires that the final rack position be known within close tolerances. If trajectory completion requires complex systems design using image analysis, artificial intelligence, pattern recognition, and other knowledge areas, variable rack position will add
substantially to the difficulty, complexity and expense of full automation.
The constraints on the solution are:

1. No major design change or modification of the module, teleoperator rack or present computer control program.
2. Accuracy of the position must be such that automated teleoperator functioning is feasible.
3. No collision, damage, binding or interference be permitted.

## Four Alternate Solutions are:

1. Attach the racks positively and solidy to the Servicer and the Satellite by welding, metal fasteners or similar method.
2. Spring load from one or both sides to center the rack in the desired position.
3. Use artificial vision and feedback with the computer control program to position the module regardless of the location of the rack.
4. Redesign the rack to provide wide entry space tapered to fit the module closely. The best solution appears to be fixed, accurate location of the racks. Full automation will require complex devices and programs to complete the trajectory successfully. Fixed rack positions will hold down the complexity and difficulties of full automation.

## Electrical Self-Aligning Connector

Avoiding damage to electrical connectors during transfer of a module between Satellite and Orbiter Servicer is a serious problem. If a pin is bent or the connector is out of alignment, the connection may fail or the connector be bent and damaged.

The proposed solution is an electrical self-aligning connector, NASA Invention Disclosure Case No. MFS - 25211. The plug and receptacle are conical with annular conductive rings spring-loaded for positive contact. This connector does not require precise alignment and is not as liable to damage as the pin-type connector.

It is recommended that the proposed electrical self-aligning connector NASA Invention Disclosure Case No. MFS - 25211 be built and tested. If tests are successful, the design should be incorporated into the TMS and other programs as appropriate.

## COMPUTER RESOURCES FOR ROBOTICS

Although MSFC is an acknowledged leader in modern computer facilities, continuous acquisition of new computer technology and upgrading of present computer resources is essential to support the robotics activities.

The robotics and teleoperator efforts should have dedicated computer resources in addition to the general computer resources available at MSFC. The dedicated computers will avoid access delays, provide immediate interface and permit efficient programming and output.

Computers themselves are improving steadily. Storage capacity, operating speed, reliability, flexibility and computing power are steadily increasing while size, weight and cost are going down. There are limits to possible improvements. (The exponential growth curve cannot go on forever.) However, the limits do not appear on the near-time horizon.

Among the dedicated computer facilities that will be needed are:

1. On-board operating and control computers for the Free Flying Mobility Unit in the Free Flying Rendezvous and Docking Simulator. This will provide for maintaining the balance, recording data, and other functions in which the human test subject need not be involved.
2. Graphics software and hardwarc for both robotics simulators. This may interface with the major computers to gain increased capacity and allow verification through simulation prior to hardware implementation.
3. Full automation of the Orbiter Servicer System will require substantial increase in computer capacity and power in order to perform the necessary command control, image analysis, and other activities.

## COMPUTER MODELING

Computer models and simulations can be useful even though they cannot replace full-scale physical mock-ups and simulators. A computer model can be developed, tested and implemented more quickly and less expensively than can a full-scale simulator. The information gained can be useful in design development.

Computer modeling is developing a considerable body of basic knowledge, practical experience, applications and published articles. The topic is included in many current Engineering, Math and Science education programs. The number of professionals with education and training in computer modeling and simulation is steadily increasing. Competence in this area will become increasingly available within NASA and from its contractors.

Computer modeling and simulation does have limits. The validity and reliability of the model output depend upon the extent to which the model represents reality and upon the quality of input data. The right parameters and the correct representations of their behavior must be incorporated in the model. In order to keep the model to manageable size, simplifying assumptions and short cuts are used. These may introduce unacceptable errors. The input data quality is vital to good results. Sometimes there is inadequate communication between the mathematical modeler and the designer-operator of the physical system. The result may be failure to interpret and apply the results properly and to detect anomalies in the model and its output.

A major problem in modeling robotics system is the representation of the physical components and the automated sensory - based interpretation of the data. Another is the present method of representing all trajectories as a set of straight-line elements. This requires a large number of steps and is expensive. A third is incorporation of five or more degrees of freedom; most current modules are limited to three (15).

An important application of robotic programming models is the development of "goal domain" programs. This permits the programmer to specify the goal states of the robot while the robot selects the appropriate actions to achieve the desired goal state (17).

Task specification definition needs to be refined and improved. Too great detail in task definition reduces the scope of adaptive programming to meet variations and contingencies. On the other hand, present state of the art in communications, artificial intelligence, computer languages and artificial vision is inadequate to support goal-oriented task definition in NASA's robotics efforts(13).

## CONCLUSIONS AND RECOMMENDATIONS

The major conclusions from the research on robotics in the Free Flying Rendezvous and Docking Simulator and the Orbital Servicer System are:

1. These facilities perform an important service in support of NASA's Space Missions.
2. These facilities should be upgraded and improved.
3. MSFC should move to acquire professional competence in research areas basic to robotics research and technology.

The main recommendations are summarized below. Detailed discussion is the body of this report.

1. A second-generation Free Flying Mobility Unit be acquired. This would provide six degrees of freedom and other improvements infeasible with the present model. The present unit might be kept if its role is complementary to the new unit.
2. The Orbital Servicer System be fully automated.
3. MSFC move to acquire competence in areas necessary to robotics teleoperator development. These include artificial intellignece, sensors, image analysis, pattern recognition, control, communication and computer science applied to robotics.
4. Enhanced computer facilities be acquired for the robotics and teleoperator activities. Some can be additions to Center-wide computer resources; others must be dedicated to the robotics and teleoperator work.
5. Procurement and test of the new electrical self-aligning connector be implemented.
6. Other recommendations are contained in the body of the report. These include rack attachment, latching mechanisms and others.

## REFERENCES

1. Shields, N. Jr., Slaughter, P. H., Bry, R. G., and Henderson, D. E.: "Earth Orbital Teleoperator Systems Evaluation. 1978 Year End Report." Essex Corporation, Contract NAS8-31848, Feb. 1, 1979.
2. Fernandez, K. R., "Application of a Computer Controlled Robot to Remote Equipment Maintenance'. Accepted for Iascon '80, Cincinnati, OH , Sept. 1980.
3. DeRocher, B. et al., "Integrated Orbital Servicing Study Follow-On", Final Progress Review. Martin-Marietta, TRW, Contract NAS8-30820 SA-5, 9, June 1978.
4. Sacerdoti, E. D., "Plan Generation and Execution for Robotics". Proceedings NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
5. Grossman, D. D.; "Robotic Intelligence and Decision Making: Real World Aspects". Proceedings NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
6. Agin, G. J., "Issues Involving Sensory Control". NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
7. Nitzan, D., "Assessment of Robotic Sensors". Proceedings NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
8. Roth, B., "Kinematic Design for Robotics". Proceedings NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
9. Dubowsky, S., "Dynamics for Manipulation: Areas of Future Research". NSF Workshop for Robotics Research, University of Rhode Island, April 15-17, 1980.
10. Bejczy, A. K., "Demands that Robotics Systems Place on Control Theory". NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
11. Paul, R. P.: Present Research approaches to Control in Robotics Systems. NSF Workshop on Robotics Research, April 15-17, 1980.
12. Graupe, D. and Saridis, G. N., "Principles of Intelligent Control for Robotics Prosthetics \& Orthotics". Proceedings, NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
13. Ruoff, C. F., "Future Languages needed for Cooperating Robots and Operating Systems for Robots". Proceedings, NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
14. Shimano, B., "Levels of Communication and Programming Methods". NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
15. Wesley, M. A., High Level Languages and Modeling. NSF Workshop on Robotics Research, University of Rhode Island, April 15-17, 1980.
16. Campbell, R. A., Unpublished memorandum, MSFC.

| MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA |  |
| :---: | :---: |
| MATERIAL AND PR INTERFACE IN THE NIQUE | FOR A FLAT RGER TECH- |
| Prepared By: | Larry M. Foster, Ph. D. |
| Academic Rank: | Assistant Professor |
| University and Department | University of Alabama in Huntsville <br> Mathematics Department |
| NASA/MSFC: <br> (Laboratory) <br> (Division) <br> (Branch) | Space Sciences <br> Space Processing <br> Semiconductors and Devices |
| MSFC Counterpart: | R. J. Naumann |
| Date: | August 1, 1980 |
| Contract No. : | NGT-01-002-099 <br> (University of Alabama) |

## By

Larry M. Foster, Ph. D.
Assistant Professor of Mathematics University of Alabama in Huntsville Huntsville, Alabama

## ABSTRACT

The functional relationships between various material and process parameters necessary for a flat interface in the Bridgman-Stockbarger technique are given. In particular, two boundary condition cases of
(1) $\nabla^{2} T=P_{l} \frac{\partial T}{\partial x}, \quad x<0$ and $0<r<1$
(2) $\nabla^{2} T=P_{s} \frac{\partial T}{\partial x}, \quad x>0$ and $0<r<1$
(3) $\left.\left.-K_{\ell} \frac{\partial T}{\partial x}\right)_{x=0^{-}}+K_{s} \frac{\partial T}{\partial x}\right)_{x=0^{+}}+L=0$
are considered.
Case 1: $\quad T(x, 1)= \begin{cases}T_{c}, & x>0 \\ T_{h}, & x<0\end{cases}$
Case 2: $\quad \frac{\partial T}{\partial r}(x, 1)=\left\{\begin{array}{l}-B_{S} T, \quad x>0 \\ -B_{\ell}\left(T-T_{1}\right), \quad x<0\end{array}\right.$

Necessary interval bounds on $T_{1}$ and $T_{h}$ are given assuming a flat melt-solid interface at $x=0$. Approximations of $T(x, r)$ and related error estimates are also given.

## ACKNOW LEDGMENTS

The author wishes to thank Dr. R. J. Naumann for suggesting this problem and for his many enlightening remarks. In addition, the computer assistance of Ernestine Cothran was invaluable. Finally, special thanks to the George C. Marshall Space Flight Center and the University of Alabama for their support, to my wife Helen for her patience, and to God for His $2^{\text {No }}$ blessings.

```
\(\mathrm{P}_{\ell}\) - liquid P éclèt number
\(P_{s} \quad-\quad\) solid Peclet number
\(\mathrm{T}_{\mathrm{c}}\) - cooler temperature
\(\mathrm{T}_{\mathrm{H}}\) - heater temperature
Q - material melting point
\(\mathrm{K}_{\ell}\) - liquid thermal conductivity
\(\mathrm{K}_{\mathrm{s}} \quad\) - solid thermal conductivity
\(\mathcal{L} \quad-\quad\) crystal growth rate - liquid density - latent heat
\(\mathrm{B}_{\mathrm{s}}\) - solid region liot number
\({ }^{B} \ell\) - liquid region Biot number
' \(T_{1}\) - scaled ambient temperature for liquid region
T - temperature
```


## INTRODUCTION

The partial differential equations defining cyrstal growth by the Bridgman - Stockbarger technique [2] are

$$
\nabla^{2} T=P_{l} \frac{\partial T}{\partial x}, \quad x<0
$$

and

$$
\nabla^{2} T=P_{S} \frac{\partial T}{\partial x}, \quad x>0
$$

where $P_{\ell}$ and $P_{S}$ are the liquid and solid Péclèt numbers respectively, $T$ is temperature and $x$ is distance from a flat solid - melt interface. To conserve energy at the interface,

$$
\left.-K_{\ell}\left(\frac{\partial T}{\partial x}\right)_{x=0^{-}}+K_{s} \frac{\partial T}{\partial x}\right)_{x=0^{+}}+\mathcal{L}=0
$$

where $K_{\ell}$ and $K_{s}$ are the respective liquid and solid thermal conductivities and $\mathcal{L}$ is the product of the crystal growth rate, the melt density and the latent heat of solidification. Two heating mechanisms are investigated; first considered is conductive heating defined by

$$
\begin{array}{ll}
T(x, 1)=T_{c}, & x>0 \\
T(x, l)=T_{H}, & x<0
\end{array}
$$

followed by Newton heating

$$
\begin{aligned}
& \frac{\partial T}{\partial r}(x, 1)=-B_{S} T, x>0 \\
& \frac{\partial T}{\partial r}(x, 1)=-B_{l}\left(T-T_{l}\right), \quad x<0
\end{aligned}
$$

where $T_{c}$ and $T_{H}$ are the respective cooler and heater temperatures, $B_{s}$ and $B_{l}$ are the solid and melt Biot numbers respectively, and $T_{1}$ is the ambient temperature surrounding melt region. The ambient temperature for the solid region has been scaled to 0 .

The assumption of a flat melt-solid interface forces $K_{s}, K_{S}, T_{c}, T_{H}$, $B_{S}, B_{l}, T_{1}, P_{S}, P_{l}$ and $Q$, the material melting temperature, to be functionally dependent. This dependence, the solutions of the above partial differential equations, and various error estimates are given in the next section followed by several numerical experiments. The final section is devoted to conclusions and recommendations.

The solutions of Problems I-II and their computable approximations are given in this section. Also given are the functional relationships between the material and process parameters $\mathrm{P}_{\ell}, \mathrm{P}_{s}, \mathrm{~K}_{\ell}, \mathrm{K}_{s}, \mathrm{~T}_{c}, \mathrm{~T}_{H}, \mathrm{~B}_{s}, \mathrm{~B}_{\ell}$, and $T$, necessary for a solid-melt interface at $x=0$. We begin with two elementary but fundamental facts.

Fact 1. If (1) $\left\{a_{n}\right\}$ alternates
(2) $\left|a_{n+1}\right| \leq\left|a_{n}\right|$
(3) $\quad a_{m} \rightarrow 0$ as $m \rightarrow \infty$
then (1) $\sum_{1}^{\infty} a_{m}$ converges
(2) $\left|\sum_{i}^{N} a_{m}-\sum_{1}^{\infty} a_{m}\right| \leq\left|a_{N+1}\right|$

Fact 2. For $f(x)$ sufficiently smooth,

$$
f^{\prime}(x)=\frac{f(x+h)-f(x)}{h}+O(h)
$$

Problem I.
We wish to solve, in cylindrical coordinates,
(1) $\quad \nabla^{2} T=P_{l} \frac{\partial T}{\partial x}, x<0,0<r<1$
(2) $\nabla^{2} T=P_{5} \frac{\partial T}{\partial x}, x>0,0<r<1$
subject to
(3) $T(x, 1)=T_{c}, x>0$
(4) $T(x, 1)=T_{H}, x<0$
(5) $\left.\left.-\mathrm{K}_{l} \frac{\partial T}{\partial x}\right)_{x=0^{-}}+\mathrm{K}_{5} \frac{\partial T}{\partial x}\right)_{x=0^{+}}+\mathcal{L}=0$
where $\mathcal{L}=$ crystal growth rate $\cdot$ melt density $\cdot$ latent heat

Condition (5) implies a solid-melt interface at $\mathrm{x}=0$, i.e.,
(6) $T(0, r)=Q$.

Unfortunately, conditions (3) - (6) overpose (1) and (2) and generally, overposed problems are impossible to solve. However, in this case, the overposing of (1) and (2) will give constraints on the material and process parameters $\mathrm{P}_{l}, \mathrm{P}_{S}, \mathrm{~T}_{\mathrm{C}}, \mathrm{Q}, \mathrm{T}_{H}, \mathrm{~K}_{\ell}, \mathrm{K}_{S}$, and $\mathcal{L}$.

By separation of variables, (1) and (2) with conditions (3), (4) and (6) have solutions [1]

$$
\begin{align*}
T(x, r) & =  \tag{7}\\
T_{c} & +\sum_{1}^{\infty} \frac{2\left(Q-T_{c}\right)}{\beta_{m} J_{1}\left(\beta_{m}\right)} J_{0}\left(\beta_{m} r\right) E_{x p}\left[\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}}\right) \frac{x}{2}\right] \\
x & >0 \text { and } \quad 0<r<1
\end{align*}
$$

$$
\begin{align*}
& T(x, r)=  \tag{8}\\
& T_{H}+\sum_{1}^{\infty} \frac{2\left(Q-T_{H}\right)}{\beta_{m} J_{1}\left(\beta_{m}\right)} J_{0}\left(\beta_{m} r\right) \operatorname{Exp}_{\rho}\left[\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{m}^{2}}\right) \frac{x}{2}\right], \\
& x<0 \text { and } \quad 0<r<1
\end{align*}
$$

where $\left\{\beta_{m}\right\}$ is the increasing sequence of all positive roots of $J_{0}$. We next show that $Q, K_{l}, K_{s}, T_{c}, T_{H}, \mathcal{Z}, P_{s}$ and $P_{l}$ cannot be arbitrary if (5) is to hold. In fact, (5) will imply that $Q, K_{s}, K_{l}, T_{c}, T_{H}, \mathcal{L}, P_{s}$ and $P_{\ell}$ satisfy a functional inequality.

Remark 1: An intuitive approach to enlisting (5) is to differentiate (7) and (8) with respect to $x$ and insert the resulting series into (5). But this would require $\sum_{1}^{\infty} J_{1}^{-1}\left(\beta_{m}\right)$ to converge, an impossibility. Hence, we will approximate the partial derivatives in (5).

By Fact 2, (5) and (6) imply

$$
\begin{align*}
& -K_{l} \frac{Q-T(-h, 0)}{h}+K_{s} \frac{T(h, 0)-Q}{h}  \tag{9}\\
& \quad+\mathcal{Z}=O(h)
\end{align*}
$$

Fact 3. Let $\left\{\beta_{m}\right\}$ be the increasing sequence of all nonnegative roots of $J_{0}(x)$. Then

$$
\begin{array}{ll} 
& \begin{cases} & \left.\frac{1}{\beta_{m} J_{1}\left(\beta_{m}\right)}\right\} \\
& \left|\frac{1}{\beta_{m} J_{1}\left(\beta_{m}\right)}\right| \text { alternates in sign, }\left|\frac{1}{\beta_{m+1} J_{1}\left(\beta_{m+1}\right)}\right| \leq\end{cases} \\
\text { as and } \frac{1}{\beta_{m} J_{1}\left(\beta_{m}\right)} \longrightarrow 0
\end{array}
$$

Combining Facts 1 and 3 with (7), (8), and (9),
(10)

$$
\begin{aligned}
& \leq O\left(h^{2}\right)+\left|k_{s} 2\left(Q-\tau_{c}\right) e^{\left(P_{s}-\sqrt{\beta_{s}^{2}+\beta_{n}^{2}}\right)} \frac{\left.\beta_{n+1}\right)}{\sigma_{1}\left(\beta_{n+1}\right)}\right| \text {. } \\
& +\left|K_{l} 2\left(Q-T_{H}\right) \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{M+1}^{2}}\right)-\frac{h}{2}}}{\beta_{M+1} J_{1}\left(\beta_{M+1}\right)}\right|
\end{aligned}
$$

In fact, for computational purposes, if $P_{S} \ll 1$ and $P<l$, then the right hand
side of (10) can be replaced by side of (10) can be replaced by
(11)

$$
\begin{aligned}
& O\left(h^{2}\right)+2 K_{s}\left(Q-T_{c}\right) \frac{e^{\left(P_{s}-2 \beta_{M+1} \frac{h}{2}\right.}}{\left|\beta_{M+1} J_{l}\left(\beta_{M+1}\right)\right|} \\
& \quad+2 K_{l}\left(T_{H}-Q\right) \frac{e^{\left(P_{l}+2 \beta_{M+1}\right)-\frac{h}{2}}}{\left|\beta_{M+1} J_{1}\left(\beta_{M+1}\right)\right|}
\end{aligned}
$$

Remark 2: (10 )forces $\mathcal{L}, K_{l}, K_{s}, Q, T_{c}, T_{H}, P_{s}$ and $P_{l}$ to be functionally dependent. For computional use, we may set
(12)

$$
\begin{aligned}
& \left.2 h-\left(K_{1}+K_{3}\right) Q+K_{s} \left\lvert\, T_{c}+2\left(Q-T_{c}\right) \sum_{1}^{M} \frac{e^{\left(P_{s}-\sqrt{\beta_{3}^{2}+4 P_{m}^{2}}\right) \frac{h_{2}}{2}} \beta_{m} J_{1}\left(\beta_{m}\right)}{}\right.\right\} \\
& +K_{\mu}\left|T_{H}+2\left(Q-T_{H}\right) \sum_{1}^{m} \frac{e^{\left(P_{2}+\sqrt{R_{R}^{2}}+4 \mu_{m_{2}^{2}}^{2}\right)}-\frac{-1}{2}}{\beta_{m} T_{1}\left(\beta_{m}\right)}\right| \\
& =0
\end{aligned}
$$

for $h$ sufficiently small and $M$ sufficiently large. This gives an equation linear in $\mathcal{L}, \mathrm{K}_{5}, \mathrm{~K}_{\ell}, Q, \mathrm{~T}_{c}, \mathrm{~T}_{H}$ and nonlinear in $\mathrm{P}_{S}$ and $\mathrm{P}_{\ell}$. In particular, (12) gives $T_{H}$ in terms of $\mathcal{L}, K_{\ell}, K_{s}, Q, T_{c}, P_{s}$ and $P_{\ell}{ }^{\ell}$. Furthermore, if (10) is violated, (1) and (2) cannot have a solution satisfying (3) - (6).
Remark 3: Since $\frac{f(x+h)-f(x)}{h}=f^{\prime}\left(x+\frac{h}{2}\right)+O\left(h^{2}\right)$, it may appear profitable to replace

$$
\begin{aligned}
& \frac{T(h, 0)-Q}{h} \text { and } \frac{Q-T(-h, 0)}{h} \text { in (9) by } \\
& \frac{\partial T}{\partial x}\left(\frac{h}{2}, 0\right) \text { and } \frac{\partial T}{\partial x}\left(-\frac{h}{2}, 0\right) \text { respectively. However, } \\
& \frac{\partial T}{\partial x}\left(\frac{h}{2}, 0\right)=2\left(Q-T_{c}\right) \sum_{1}^{\infty} \frac{P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}}}{\beta_{m} J_{1}\left(\beta_{m}\right)} e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}}\right) \frac{h}{4}}
\end{aligned}
$$

converges much slower than the estimates in (10).

Remark 4: An interval estimate of linear parameters $\mathcal{L}, \mathrm{K}_{\boldsymbol{l}}, \mathrm{K}_{s}, \mathrm{~T}_{c}$ and $\mathrm{T}_{H}$ can be given. For example, suppose $\mathcal{L}, K_{l}, K_{s}, T_{c}, P_{s}, P_{l}$ and $Q$ are fixed. Then (10) implies
(12)

$$
\begin{aligned}
& O\left(h^{2}\right)-K_{s} 2\left(Q-T_{c}\right) \frac{e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{M+1}^{2}}\right) \frac{h}{2}}}{\left|\beta_{M+1} J_{1}\left(\beta_{M+1}\right)\right|} \\
& -\mathcal{L} h+\left(K_{l}+K_{s}\right) Q+2 Q K_{l} \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{M+1}^{2}}\right)-\frac{h}{2}}}{\left|\beta_{M+1} J_{1}\left(\beta_{M+1}\right)\right|} \\
& -K_{s}\left(T_{c}+2\left(Q-T_{c}\right) \sum_{l}^{M} \frac{e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}} \frac{h}{2}\right.}}{\beta_{m} J_{1}\left(\beta_{m}\right)}\right) \\
& -K_{l} 2 Q \sum_{1}^{M} \frac{e^{\left(P_{l}+\sqrt{\left.P_{l}^{2}+4 \beta_{m}^{2}\right)}-\frac{h}{2}\right.}}{\beta_{m} J_{1}\left(\beta_{m}\right)} \leq \\
& K_{l} T_{H}\left(1-2 \sum_{1}^{M} \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{m}^{2}}\right) \frac{h}{2}}}{\beta_{m} J_{l}\left(\beta_{m}\right)}+2 \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{M+1}^{2}}\right.}}{\left|\beta_{M+1} J_{1}\left(\beta_{M+1}\right)\right|}\right)
\end{aligned}
$$

and
(13)

$$
\begin{aligned}
& K_{l} T_{H}\left(1-2 \sum_{1}^{M} \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{m}^{2}}\right)-\frac{h}{2}}}{\beta_{m} J_{1}\left(\beta_{m}\right)}-2 \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{M+1}^{2}}\right) \frac{h}{2}}}{\left|\beta_{M+1} J_{1}\left(\beta_{M+1}\right)\right|}\right) \\
& \leq O\left(h^{2}\right)+2 K_{s}\left(Q-T_{c}\right) \frac{e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{M+1}^{2}}\right) \frac{h}{2}}}{\left|\beta_{M+1} J_{1}\left(\beta_{M+1}\right)\right|} \\
& -2 K_{l} Q \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{M+1}^{2}}\right) \frac{-h}{2}}}{\left|\beta_{M+1} J_{1}\left(\beta_{M+1}\right)\right|}-\mathcal{L} h+\left(K_{l}+K_{s}\right) Q \\
& -K_{S}\left(T_{C}+2\left(Q-T_{c}\right) \sum_{l}^{M} \frac{e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}}\right) \frac{h}{2}}}{\beta_{m} J_{1}\left(\beta_{m}\right)}\right) \\
& -K_{l} 2 Q \sum_{1}^{M} \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{m}^{2}}\right) \frac{-h}{2}}}{\beta_{m} J_{1}\left(\beta_{m}\right)}
\end{aligned}
$$

Solving (12) and (13) for $T_{H}$,

$$
T_{1}(M, h)+O\left(h^{2}\right) \leq T_{H} \leq T_{2}(M, h)+O\left(h^{2}\right)
$$

For practical purposes, set

$$
\begin{equation*}
T_{1}(M, h) \leq T_{H} \leq T_{2}(M, h) \tag{14}
\end{equation*}
$$

Since we only know bounds on $T_{H}$ (from (14)), for computational use, approximate $\mathrm{T}_{H}$ by $\mathrm{T}_{H, A P P}=\frac{\mathrm{T}_{2}(\mathrm{M}, \mathrm{h})+\mathrm{T}_{1}(\mathrm{M}, \mathrm{h})}{\text {. If } \mathrm{T}_{H} \text { is replaced by } \mathrm{T}_{H, A P P} \text { in (8), then }}$ the resulting error is

$$
\begin{equation*}
\left|T_{H}-T_{H, A P P}-2 \sum_{1}^{\infty} \frac{T_{H}-T_{H, A P P}}{\beta_{n} J_{1}\left(\beta_{n}\right)} J_{0}\left(\beta_{n} r\right) e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{m}^{2}}\right) \frac{x}{2}}\right| \tag{15}
\end{equation*}
$$

$$
\begin{align*}
& \text { which has an upper bound of } \\
& \text { (16) } \quad\left(T_{2}(M, h)-T_{1}(M, h)\right)\left|\frac{1}{2}-\sum_{1}^{\infty} \frac{J_{0}\left(\beta_{m} r\right)}{\beta_{m} J_{1}\left(\beta_{m}\right)} e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \beta_{n}^{2}}\right) \frac{x}{2}}\right| \tag{16}
\end{align*}
$$

Problem II.

We wish to solve, in cylindrical coordinates,
(18) $\quad \Gamma^{2} T=P_{3} \frac{\partial T}{\partial x}, \quad x>0 \quad$ and $\quad 0<r<1$
subject to

$$
\begin{align*}
& \frac{\partial T}{\partial r}(x, 1)=-B_{s} T, x>0  \tag{19}\\
& \frac{\partial T}{\partial r}(x, 1)=-B_{l}\left(T-T_{1}\right), x<0 \\
& \left.\left.-K_{l} \frac{\partial T}{\partial x}\right)_{x=0^{-}}+K_{s} \frac{\partial T}{\partial x}\right)_{x=0^{+}}+\alpha=0
\end{align*}
$$

where $\mathcal{L}=$ crystal rate of growth - liquid density - latent heat, $B_{1} \geqslant 0$ and $B_{s} \geqslant 0$. Condition (21) implies a solid-melt interface at $x=0$, i. e.,

$$
\begin{equation*}
T(0, r)=Q \tag{22}
\end{equation*}
$$

As in Problem I, conditions (19) - (22) over pose (17) and (18). This dilemma can be resolved only if the material and process parameters $P_{s}, P_{l}, Q, T_{1}$, $K_{s}, K_{l}$ and $\mathscr{L}$ are functionally dependent.

Separating variables, (17) and (18) subject to (19), (20) and (22) have solutions [1]

$$
\begin{equation*}
T(x, r)=\sum_{1}^{\infty} \frac{2 Q B_{s} J_{0}\left(\beta_{m} r\right)}{\left(\beta_{m}^{2}+B_{s}^{2}\right) J_{0}\left(\beta_{m}\right)} e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}}\right) \frac{x}{2}}, x>0 \tag{23}
\end{equation*}
$$

(24)

$$
T(s, r)=T_{1}+\sum_{1}^{\infty} \frac{2\left(Q-T_{1}\right) B_{l} J_{0}\left(\beta_{m} r\right)}{\left(\alpha_{m}^{2}+B_{l}^{2}\right) J_{0}\left(\alpha_{m}\right)} e^{\left(P_{l}+\sqrt{\left.P_{l}^{2}+4 \beta_{m}^{2}\right) \frac{x}{2}}\right.}, x<0
$$

where $\left\{\alpha_{m}\right\}$ and $\left\{\beta_{m}\right\}$ are the increasing nonnegative sequences of all roots of
(25) $B_{s} J_{0}\left(\beta_{n}\right)=\beta_{n} J_{1}\left(\beta_{n}\right)$
(26) $\quad B_{l} J_{0}\left(\alpha_{m}\right)=\alpha_{m} J_{1}\left(\alpha_{m}\right)$

Fact 4: $\left\{\left(\beta_{n}^{2}+\beta_{s}^{2}\right) J_{0}\left(\beta_{n}\right)\right\}$ alternates in sign and $\left|\frac{1}{\left(\beta_{n}^{2}+B_{s}^{2}\right) J_{0}\left(\beta_{n}\right)}\right|$
decreases to 0. A similiar result holds for $\left\{\frac{1}{\left(\alpha_{m}^{2}+B_{l}^{2}\right) J_{0}\left(\alpha_{m}\right)}\right\}$
As in Problem I, (21) and Fact 2 imply
(27)

$$
-K_{l} \frac{Q-T(-h, 0)}{h}+K_{s} \frac{T(h, 0)-Q}{h}+\mathcal{L}=O(h)
$$

Combining Facts 1 and 4 with (23), (24) and (27),
(28)

$$
\begin{aligned}
& \mathcal{L} h-Q\left(K_{s}+K_{l}\right)+K_{s} \sum_{1}^{M} \frac{2 B_{s} Q}{\left(\beta_{m}^{2}+B_{s}^{2}\right) J_{0}\left(\beta_{n}\right)} e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}}\right) \frac{h}{2}} \\
& \left.+K_{l}\left\{T_{1}+\sum_{1}^{M} \frac{2\left(Q-T_{1}\right) B_{l}}{\left(\alpha_{m}^{2}+B_{l}^{2}\right) J_{0}\left(\alpha_{m}\right)} e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{m}^{2}}\right) \frac{-h}{2}}\right\} \right\rvert\,
\end{aligned}
$$

$$
+\left|\frac{2\left(Q-T_{1}\right) B_{l} K_{l}}{\left(\alpha_{m+1}^{2}+B_{l}^{2}\right) J_{\sigma}\left(\alpha_{m+1}\right)} e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{m+1}^{2}}\right)-\frac{h}{2}}\right|
$$

As in Problem I, (28) implies various interval constraints on $Q, K_{j}, K_{l}, T_{1}$, $P_{s}, P_{l}, B_{s}$ and $B_{l}$. For illustration, consider $T_{1}$. Because we have transformed (in practice) the Newton cooling condition (19) to imply a $0^{\circ}$ ambient temperature around the rod for $x>0,0<Q<T_{1}$. (28) then implies (29)

$$
\begin{aligned}
& O\left(h^{2}\right)+Q\left(K_{s}+K_{l}\right)-2 Q K_{l} B_{l} \sum_{1}^{M} \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{m}^{2}}\right)-\frac{h}{2}}}{\left(\alpha_{m}^{2}+B_{l}^{2}\right) J_{0}\left(\alpha_{m}\right)} \\
& -2 K_{s} B_{s} Q \sum_{1}^{M} \frac{e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}}\right) \frac{h}{2}}}{\left(\beta_{m}^{2}+B_{s}^{2}\right) J_{0}\left(\beta_{m}\right)}-\mathcal{L} h \\
& +\frac{2 B_{s} Q K_{s} e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m+1}^{2}}\right) \frac{h}{2}}}{\left(P_{m+1}^{2}+B_{s}^{2}\right)\left|J_{0}\left(\beta_{m+1}^{2}\right)\right|}-\frac{2 B_{l} Q K_{l} e^{\left(P_{1}+\sqrt{P_{l}^{2}+4 \alpha_{m+1}}\right)-\frac{h}{2}}}{\left(\alpha_{m+1}^{2}+B_{l}^{2}\right)\left|J_{s}\left(\alpha_{m+1}\right)\right|}
\end{aligned}
$$

$$
\begin{aligned}
\geqslant T_{1} K_{l} & \left\{1-2 B_{l} \sum_{1}^{M} \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{m}^{2}}\right) \frac{-h}{2}}}{\left(\alpha_{m}^{2}+B_{l}^{2}\right) J_{0}\left(\alpha_{m}\right)}\right. \\
& \left.-2 B_{l} \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{M+1}^{2}}\right)} \frac{-h}{2}}{\left(\alpha_{M+1}^{2}+B_{l}^{2}\right)\left|J_{0}\left(\beta_{M+1}\right)\right|}\right\}
\end{aligned}
$$

and
(30)

$$
\begin{aligned}
& O\left(h^{2}\right)-\frac{2 B_{s} K_{s} Q e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{M+1}^{2}}\right) \frac{h}{2}}}{\left(\beta_{M+1}^{2}+B_{s}^{2}\right)\left|J_{0}\left(\beta_{m+1}\right)\right|}+ \\
& \frac{2 Q B_{l} K_{l} e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{M+1}^{2}}\right) \frac{-h}{2}}}{\left(\alpha_{m+1}^{2}+B_{l}^{2}\right)\left|J_{0}\left(\beta_{m+1}\right)\right|}+Q\left(K_{s}+K_{l}\right) \\
& -2 K_{l} B_{l} Q \sum_{1}^{M} \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{m}^{2}}\right)-\frac{h}{2}}}{\left(\alpha_{m}^{2}+B_{l}^{2}\right) J_{0}\left(\alpha_{m}\right)} \\
& -2 K_{s} B_{s} Q \sum_{1}^{M} \frac{e^{\left(P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}}\right) \frac{h}{2}}}{\left(\beta_{m}^{2}+B_{s}^{2}\right) J_{0}\left(\beta_{m}\right)}
\end{aligned}
$$

$$
\begin{aligned}
\leq T_{1} K_{l} & \left\{1-2 B_{l} \sum_{l}^{M} \frac{e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{m}^{2}}\right)-\frac{h}{2}}}{\left(\alpha_{m}^{2}+B_{l}^{2}\right) J_{0}\left(\alpha_{m}\right)}\right. \\
& \left.+2 B_{l} \frac{c^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{m+1}^{2}}\right)-\frac{h}{2}}}{\left(\alpha_{M+1}^{2}+B_{l}^{2}\right)\left|J_{l}\left(\alpha_{m+1}\right)\right|}\right\}
\end{aligned}
$$

Solving (29) and (30) for $T_{1}$,

$$
T_{1}^{\text {Low }}(\mathrm{M}, \mathrm{~h})+0\left(\mathrm{~h}^{2}\right) \leq \mathrm{T}_{1} \leq \mathrm{T}_{1}^{\mathrm{H} / \mathrm{GH}}(\mathrm{M}, \mathrm{~h})+0\left(\mathrm{~h}^{2}\right)
$$

For computational purposes, let
(31)

$$
T_{1}^{\text {LoW }}(M, h) \leqslant T_{1} \leq T_{1}^{H / G H}(M, h)
$$

(32)

$$
T_{1, \text { APPROX }}=\left(\mathrm{T}_{1}^{\text {LOW }}(\mathrm{M}, \mathrm{~h})+\mathrm{T}_{1}^{\text {HIGH }}(\mathrm{M}, \mathrm{~h})\right) / 2
$$

The error committed by replacing $\mathrm{T}_{1}$ in (24) by $\mathrm{T}_{1, \text { APPROX }}$ is no more than
(33)

$$
\left(T_{1}^{\text {HiGH }}(m, h)-T_{1}^{\text {Low }}(m, h)\right)\left|\frac{1}{2}-\sum_{1}^{\infty} \frac{B_{l}^{J_{0}\left(\alpha_{m} r\right)}}{\left(\alpha_{m}^{2}+B_{l}^{2}\right) J_{0}\left(\alpha_{m}\right)} e^{\left(P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{m}^{2}}\right) \frac{x}{2}}\right|
$$

Remark 5: It is possible to substitute (23) and (24) into (21) and solve for $\mathrm{T}_{1}$. However, this would require computing
(34)

$$
\sum_{1}^{\infty} \frac{P_{s}-\sqrt{P_{s}^{2}+4 \beta_{m}^{2}}}{\left(\beta_{n}^{2}+B_{s}^{2}\right) J_{0}\left(\beta_{n}\right)}
$$

and
(35)

$$
\sum_{1}^{\infty} \frac{P_{l}+\sqrt{P_{l}^{2}+4 \alpha_{m}^{2}}}{\left(\alpha_{m}^{2}+B_{l}^{2}\right) J_{0}\left(\alpha_{m}\right)}
$$

Although (34) and (35) conditionally converge, the sums in (28) - (30) and (33) converge faster by several orders of magnitude.

Remark 6: Let $B_{S}=B_{l}=P_{S}=P_{l}=0=\mathcal{L}$ and fix $0<h \ll 1$.

$$
\text { Letting } M=39,(10) \text { implies }
$$

$$
Q=\frac{K_{s} T_{c}+K_{l} T_{H}}{K_{l}+K_{s}}
$$

for Problem 1 and (28) gives

$$
Q=\frac{K_{l} T_{t}}{K_{s}+K_{l}}
$$

for Problem 2.

The results of several numerical experiments are given in this section.

Example 1: In Figure 1 and Table 1 we illustrate the dependence on M of the interval estimate (14). Let $\mathrm{T}_{\mathrm{c}}=400, \mathrm{Q}=700$, $P_{S}=P_{l}=0.001, \mathscr{L}=0.01, h=0.1, K_{S}=0.0032$ and $K_{l}=0.002$.

| $M$ | Interval Estimate of $\mathrm{T}_{H}$ |
| :---: | :---: |
| 9 | $[1391.34,1476.13]$ |
| 14 | $[1202.19,1214.26]$ |
| 19 | $[1180.56,1182.74]$ |
| 24 | $[1176.34,1176.54]$ |
| 29 | $[1176.34,1176.54]$ |
| 34 | $[1175.92,1175.93]$ |

Table 1


Figure 1
Example 2: In Table 2 and Figure 2 we illustrate the sensitivity of the interval estimate (14) to changes in $\mathscr{L}$. Let $K_{l}=K_{s}=P_{s}=P_{l}=0.001$,

$$
h=0.1, T_{c}=200.0, Q=700.0 \text { and } M=39 .
$$

| $\mathcal{L}$ | Interval Estimate of $\mathrm{T}_{H}$ |
| :---: | :---: |
| .001 | $[1192.386,1192.389]$ |
| .011 | $[1116.514,1116.517]$ |
| .021 | $[1040.642,1040.645]$ |
| .031 | $[964.770,964.773]$ |
| .041 | $[888.898,888.901]$ |
| .051 | $[813.026,813.029]$ |
| .061 | $[737.154,737.157]$ |

Table 2


Figure 2

## CONCLUSIONS AND RECOMMENDATIONS

Equations (10) and (28) imply that the material and process parameters cannot be arbitrary if a flat melt-solid interface is to be achieved. The sensitivity of one parameter to changes in another parameter (Example 2) suggests great care should be taken in obtaining the various material parameters.

The relationships between $M$ and $h$ that minimize the interval widths in (14) and (31) remain open questions. Further investigations of Problems I and II could incorporate adiabatic boundary zones near the melt-solid interface and the influences of various container configurations.

## REFERENCES

1. Carslaw, H. S. and Jaeger, J. C., Conduction of Heat in Solids, 2nd Ed., Oxford Press, 1959.
2. Chang, C. E., and Wilcox, W. R., "Control of Interface Shape in the Vertical Bridgman-Stockbarger Technique", Journal of Crystal Growth, Vol. 21, pp. 135-140, 1974.

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM 

A RANKING ALGORITHM FOR SPACELAB
CREW $\Lambda$ ND EXPERIMENT SCHEDULING

\author{

Prepared by: Robert D. Grone and Frank H. Mathis <br> Academic Ranks: Assistant Professor <br> Institution and Department: | Auburn University |
| :--- |
| Department of Mathematics |

}

NASA/MSFC:
Division: Mission Analy:it:
Branch: Mission lmbertalion

MSFC Counterparts: Jerry D. Weiler and William C. Askew

Contract No: NGT-01-002-099 (University of Alabama)

A RANKING ALGORITHM FOR SPACELAB
CREW AND EXPERIMENT SCHEDULING
by

Robert D. Grone Assistant Professor Department of Mathematics Auburn University Auburn, Alabama 36849

Frank H. Mathis Assistant Professor
and Department of Mathematics Auburn University Auburn, Alabama 36849

## ABSTRACT

This report is concerned with the problem of obtaining an optimal or near optimal schedule for scientific experiments to be performed on Spacelab missions. The design criteria of a method for solving this problem is two-fold. First, the method must produce a schedule which satisfies all material and resource constraints while maximizing a function assigning a grade to each schedule. Secondly, the method should reduce the amount of preflight manpower and computer time currently necessary to achieve this goal and be rapid and flexible enough to allow real-time support of rescheduling problems arising due to launch delays, equipment malfunction, or other factors.

In this paper the current capabilities in this regard are examined and a method of ranking experiments in order of difficulty is developed to support the existing software. Experimental data is obtained from applying this method to the sets of experiments corresponding to Spacelab missions 1, 2 and 3. Finally, suggestions are made concerning desirable modifications and features of second generation software currently being developed for this problem.

## ACKNOWLEDGEMENT

The authors of this report wish to thank their counterparts and John Jaap for the careful and complete introduction to the problem which they provided and for the generous support afforded in terms of assistance, supplies, and computer access.

## INTRODUCTION

The problem under consideration is that of scheduling a set of experiments to be performed on a given Spacelab mission. Through a preliminary analysis, which is not discussed here, a set of compatible experiments is compiled which are hopefully to be included in a specific mission. The experiments are then converted into model sheets by the principal investigator, who is responsible for the experiment design, and a NASA project engineer familiar with the format and requirements of the model sheets. The model sheet divides each experiment into a sequence of steps, each of which has certain requirements in terms of crew, equipment, and energy usage; and may in addition require that the Spacelab be in a certain position or configuration, or have certain targets available. Examples of targets are planets, communication satellites, and data receiving facilities on the earth's surface. If the experiment is to be performed more than once the number of performances desired is included on the model sheet. A typical Spacelab mission lasts for seven days and involves six crewmembers.

Extremely sophisticated software is available to assist in scheduling of the experiments. The current program is designated by TLP for timeline program. This program takes the model sheets in some order and "front loads" them in this order. By front load, it is meant that the model is scheduled at the earliest possible time which satisfies all the model constraints and which does not conflict with the requirements of previously scheduled models. If it is impossible to schedule a certain model at any time the program then goes to the next model in the sequence. This program is currently being upgraded to a program denoted by ESP for experiment scheduling program, which has the capability of either front loading or back loading any given model, and which is to supersede TLP sometime in 1981.

Even with the extremely advanced and effective software available, the problem of scheduling the experiments in an optimal way currently requires an enormous amount of pre-flight man-hours. By an optimal schedule, we roughly mean one that includes the maximal number of experiments and performances, and which makes maximal use of the available resources. Since each set of models designated for inclusion in a specific mission are only roughly sorted for compatibility, it is frequently impossible to include every performance of every experiment.

The method previously employed has been to feed TLP various random orderings of the model sheets and have the project engineers examine the characteristics of the resulting schedules. From these examinations other orderings are suggested and fed to TLP. After months of such trial and error, the best resulting schedule is then modified and edited by hand by a group of NASA personnel familiar with the characteristics of the given mission. The current schedule for Spacelab mission one has been under development for five years. It is clear that such a time expenditure is not only economically prohibitive, but is unacceptable in light of the fact that the Spacelab program eventually envisions flying several missions a year.

The main object of this paper is to provide a method of ranking the model sheets in order of difficulty. By scheduling the most difficult experiments first it is thought that an optimal schedule will result. In support of this idea a grading function is used which assigns a higher grade to schedules which are in some sense better. The problem is then to find a ranking function of the models which provides an ordering which results in a schedule which maximizes the grading function. Since there are 218 model sheets corresponding to Spacelab mission one, an exhaustive search is clearly out of the question.

The remainder of this paper is devoted to the development and examination of such a ranking function While the method developed did not result in a schedule superior to the schedule hand produced over a five-year period for Spacelab mission one, it did quickly produce schedule for Spacelab missions two and three, which were superior to the best existing schedules for these missions.

These experimental results are described and suggestions are made on the basis of these for possible modifications or additions to the existing software.

The problem is now to produce a ranking algorithm which produces an ordering of the models which produces an optimal schedule in terms of a grading function which is described in the results section. While this has certain aspects of a linear programming problem which might be treated with either the simplex method or the new Khachian method, it is unclear if it is possible to convert this to a linear problem. The method we exhibit is similar to the one described by Garey, Graham, Johnson, and Yao in J. Comb. Theory (A) 21 (1976), 257-298.

We now proceed to describe the method used to rank the models in decreasing order of difficulty. The steps involved are: the construction of a nonnegative real vector corresponding to a single performance of each model, the normalization of these vectors, and the evaluation of the difficulty function on the normalized vectors. The models (vectors) are then arranged in decreasing order of difficulty and this ordering is fed to TLP to produce a schedule.

The vector includes a component corresponding to crew time required. The length of each step is multiplied by the number of crewmen required for the step, and the sum of these terms over all steps in the model is computed to give a figure in crewman-minutes.

The vector includes components for up to 10 resources. On Spacelab mission 1 a total of 4 resources were included: power, high rate multiplex, computer memory, and data transmission. The components for each of these resources were computed by multiplying the length of each step times the level usage of the resource and then summing over all steps.

The vector includes up to 40 components corresponding to various pieces of equipment. As in the case of crew and resources, the number of equipment-minutes is computed by summing the number of minutes a certain piece of equipment is required over all steps of the model. On Spacelab mission 1 , a total of 38 pieces of equipment are considered as components of the vector.

The vector includes components corresponding to attitude and maneuvers. If a step requires that the Orbiter be performing a certain maneuver the length of the step is multiplied by 9 , and if a step requires that no maneuvers are to be performed during its execution the length of the step is multiplied by 1 . The sum over all steps of the model then produces the component corresponding to maneuvers. The attitude (Orbiter orientation) component is computed in an analogous fashion.

The vector includes up to 72 components corresponding to targets; Spacelab mission 1 having a total of 50 such components. Once again, a figure in target-minutes is produced for each target and model by summing the number of minutes each target is required by a particular model over all steps of that model.

The vector includes two components referred to as $\Delta$-start and start-time. For the ith model, let $M_{i}$ and $m_{i}$ denote the maximum and minimum respective times at which the performance of the first step of the model may be initiated. The $\Delta$-start component is

$$
\Delta_{i}=L-\left(M_{i}-m_{i}\right),
$$

where $L$ is the total length of the mission. This component is higher for models which are less flexible with respect to possible starting times. The start-time component is

$$
S_{i}=L-M_{i},
$$

and gives higher values to models which must be started earlier in the mission. The start-time figure is included in view of the front-loading nature of TLP which causes the time and resources at the earlier part of the mission to be used up first.

The next step in the ranking algorithms is the normalization of the vectors corresponding to the models. Let $\ddot{\mathrm{V}}_{\mathrm{ij}}$ denote the $j$ th component of the vector corresponding to the ith model. Hence, if there are models and $n$ vector components then

$$
v=\left[v_{i j}\right],
$$

is an m-by-n nonnegative real matrix. Spacelab mission 1 corresponds in this fashion to a 218-by-95 matrix. The scaling is obtained by dividing the j th component of each
vector by the sum of the $j$ th components of all the vectors. That is: we normalize the matrix $V$ to a column-stochastic matrix by dividing each column by its column sum. The targets constitute an exception to this rule. The target time component is normalized by dividing by the total length the target is available during the mission rather than the time the target is required by all the models.

The next step in the ranking algorithm is to evaluate the difficulty function at each model (row of V). The programmer has the option of assigning nonnegative weights; $w_{j}, j=1, \ldots, n$, to each of the components of a vector (or alternatively, columns of $V$ ). If the ith model has corresponding vector $v_{i}=\left(v_{i 1}, \ldots, v_{i n}\right)$, the difficulty of the ith model is defined by

$$
d_{i}=\sqrt{p_{i}}\left(v_{i 1} w_{1}+\ldots+v_{i n} w_{n}\right),
$$

where $p_{i}$ is the number of performances of the $i$ th model requested. The capability of adjusting the weights allows a scheduler to emphasize or de-emphasize certain aspects of the mission deemed more critical or less critical to produce various rankings of the models.

The above described procedure of ranking the models in order is referred to as the model-sort method. A slightly different method referred to as performance-sort was also implemented and tested on Spacelab missions 1,2 , and 3. The idea behind the performance-sort is to treat each performance of an experiment as a separate model. With the model sort program, all possible performances of a given model are scheduled consecutively. With the performance sort program, each performance of a model corresponds to a vector. For example: Spacelab 1 lists 935 performances of the 218 models. Hence, when using model-sort, a 218-by-95 matrix is envisioned; and when using performance-sort a 935-by- 95 matrix is envisioned. If $\mathrm{v}_{\mathrm{i}}$ is the vector previously described which corresponds to a single performance of the ith model, and the ith model requires $p_{i}$ performances, the difficulty level of the kth performance of the ith model is

$$
\sqrt{k}\left(v_{i 1} w_{1}+\ldots+v_{i n} w_{n}\right) ; k=1, \ldots, p_{i}
$$

Another feature of the ranking method is the selection of fixed models. This allows the programmer to specify a certain subset of the models or performances which will be scheduled first. For example, on Spacelab mission 1, such items as "crewman sleep" and "crewman lunch" have model sheets corresponding to them. These models are scheduled first and then the remaining models are ranked using either the model-sort or performance-sort method. On Spacelab mission 1 a total of 16 models were fixed.

Another feature of the algorithm is the model weight capability. This is in some sense dual to the vector normalization in that a row is multiplied by a nonnegative scalar in the model weight feature. This allows the programmer to assign weights to any or all of the models which will change the corresponding vector. This option lets the operator increase the difficulty function of a model which is considered particularly important for political or other reasons not reflected in the difficulty function. The authors used this feature to schedule models which were not included in an otherwise excellent schedule. Successive doubling of the weight of a certain model insures that it will appear higher and higher in the ordering until it is the first model scheduled, if necessary.

## RESULTS

We proceed to discuss the results obtained using the previously described methods. The schedule produced due to some choice of sorting technique, weights, fixed models, and model weights is evaluated by means of a grading function which is dependent on four parameters: percentage of crew time, percentage of experiment time, percentage of performances, and percentage of models. The average of these four percentages yields the grade of the corresponding schedule. All four parameters are obtained by comparing amount scheduled to amount requested. In this connection, it should be mentioned that the experiment and crew time requested is the minimum experiment time for experiments whose steps have variable length. Hence, it is possible for the percentage score on experiment and crew time, and the grade, to be over 100. Another fact worth noting is that is possible for a fraction of a model to be performed in some cases. For example, if 3 out of 4 steps in a given model were performed this would be counted as .75 of a model. The percentage of models reflects the extent to which experiment on the mission is performed at least once, and the percentage of performances measures the extent to which all requested experiments (including multiplicities) were scheduled.

For each of Spacelab missions 1, 2, and 3 we include tabulated data comparing the effects of varying sorting technique, weights, fixed models, and model weights. 'For each mission we include for comparison the best schedule which had been previously produced by the personnel at Marshall Space Flight Center. Their method was to experiment with various orderings, feed them to TLP, obtain some reasonably good schedules, and then to invoke software referred to as TLE (timeline editor) to then shift and fit experiments. The TLE affords the capability of removing and rescheduling a single performance from an existing schedule.

Spacelab mission 1 has undergone the longest development time, and we were unable to exceed the best previous score by our methods. Mission 1 is particularly complicated due to the fact that the set of models chosen for inclusion were selected more for political considerations of being included on the first mission than for compatibility of the various models. It is thought that subsequent missions will be designed around a particular theme such as atmospheric research, materials processing, space physics, biological sciences, or astronomy. On Spacelab mission 1 it was
found, somewhat unexpectedly, that the model sort technique worked best. On Spacelab missions 2 and 3, the performance sort worked best and produced schedules superior to the previous best. The overall best results were obtained by using the weights which had produced the schedule with maximal score and then increasing the model weights of those models which failed to schedule. On these missions every model was scheduled for at least one performance. The following three tables summarize the results obtained on these three missions, and compares these to the previous best schedule.

## SPACELAB 1

MODELS - 218 -- PERFORMANCES - 935
MIN. EXP. TIME - 4869.71 HRS. -- MIN. CREW TIME - 800.62 HRS.

| WEIGHIING | TYPE OF SORT | MODETS | PERFORMANCES | EXP. TIME | CREW TIME | GRADE |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ALL EQUAL | MODEL | 183.69 | 858.61 | 4939.35 | 756.25 | 92.99. |
| ALL EQUAL | PERFORMANCE | 184.51 | 851.22 | 4847.83 | 757.56 | 92.46 |
| $\begin{aligned} & \text { TARGETS }=10, \\ & \Delta \text {-START }=1 \end{aligned}$ | MODEL | 193.90 | 884.39 | 4664.38 | 762.87 | 93:64 |
| PREVIOUS BEST |  | 203.03 | 897.94 | 4866,52 | 751.04 | 95.73 |

SPACELAB 2
MODELS - 168 -- PERFORMANCES - 1292
MIN. EXP. TIME - 2887.64 HRS. -- MIN. CREW TIME - 647.44 HRS

| WEIGHITNG | TYPE OF SORT | MODELS | PERFORMANCES | EXP, TIME | CREW TIME | GRADE |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ALL EQUAL EXCEPT $\triangle$-START $=0$ | MODEL | 154.54 | 1931.09 | 3804.08 | 682.90 | 106.54 |
| ALL EQUAL EXCEPT $\triangle-S T A R T=0$ | PERFORMANCE | 159. 50 | 19.71 .09 | 3813.67 | 693.83 | 108.28 |
| CREW, RESOURCES, TARGETS, WINDOW SIZE ONLY. <br> (EQUIPMENT AND START TIME $=0$ ) | PERFORMANCE | 162.15 | 1974.09 | 3820.48 | 700.44 | 109.03 |
| ABOVE WITH MODEL WEIGHTS ADJUSTED | PERFORMANCE | 164.15 | 1976.09 | 3821.98 | 702.48 | 109.44 |
| PREVIOUS BEST |  | 161.82 | 1976.65 | 3808.10 | 658.08 | 107.27 |

SPACELAB 3
MODELS - 61 -- PERFORMANCES - 592
MIN. EXP. TIME - 2507.48 HRS. -- MIN. CREW TIME - 569.02


In this section, we indicate some avenues of approach suggested by our investigations. These suggestions arose largely on an intuitive basis, and are in no particular order.

It is certain that experimentation with weights is necessary to implement this algorithm. It appeared that no one weighting would produce the best results for all Spacelab missions. Perhaps a scheme could be developed to give an approximation to the optimal weighting by considering the resource profiles of the specific mission.

The difficulty function could be modified. The difficulty function discussed in the body of this paper was motivated by the sum norm on nonnegative vectors. That is, if $v=\left(v_{1}, \ldots, v_{n}\right)$ is a nonnegative real vector, the sum norm of $v$ is defined via

$$
v_{1}+\ldots+v_{n} .
$$

We also experimented with a grading function related to the supremum norm

$$
\max _{i} v_{i} \cdot
$$

The results obtained were equal or slightly less than the results described in the previous section. It is rather remarkable that a greater difference was not observed. Further experimentation is possible. Perhaps the Euclidean norm might be investigated.

The grading function could be modified. For example, weights could be assigned to the four parameters involved. Also, the grading function does not take into consideration the evenness of the scheduling distribution. It would be better, for example, if the crew were not completely utilized during the first half of the mission and have a lot of unused time during the second half. Perhaps some mechanism might be devised to give preference to schedules whose resource utilization levels are held fairly constant over the duration of the mission.

An examination of what we shall term a dynamic sort scheme was done. In the method described in this paper no consideration is taken of the fact that a model may become more difficult to schedule if models preceding it deplete the resources it requires. We implemented the following iterative technique to circumvent this phenomena. A re-examination of the normalization procedure employed shows that most components of the vector were normalized by the total amount required by all the models in that component. The target components were normalized by the total available target time rather than the total target time requested. We also experimented with normalizing all components by the amount available corresponding to each component. With this normalization scheme we implemented the following dynamic sort option. In this scheme we would find the most dificult model and then re-normalize the remaining vectors by the amounts remaining of the various resources, etc. This iterative technique has the effect of dynamically increasing the weighting factors on those components which are being depleted fastest. We then assigned various weighting factors and computed the corresponding orderings with and without the dynamic sort option. The ordering produced with the dynamic sort option consistently produced a schedule inferior to the one produced without using the dynamic sort. Even in view of these results, we feel that investigation along these lines might yield superior techniques.

All of the previous discussions have focused on possible ways to improve the method of ordering the models to feed to TLP. Our final suggestion arises from limitations imposed by the front-loading nature of TLP, and involve factors internal to TLP rather than external. With the front-loading scheme, a model is scheduled at the earliest possible time. This results in a schedule which has no flexibility during the first part of the mission. It is fairly certain that this period in the mission is very likely to be plagued with correction and adjustment factors, and perhaps space sickness on the part of the crew. Also, the pieces of equipment are scheduled to begin operation at maximum capacity and then level off. This makes the execution of the mission highly dependent on the Spacelab being able to operate at full capacity immediately. The implementation of a front- or back-loading option on ESP is largely motivated by a desire to avoid this type of scheduling. It would be advantageous to schedule each successive model at the "best" possible time rather than the earliest or latest. "Best" might
be defined by scheduling the model at the time during the mission which keeps the overall resource-available profile the flattest over the duration of the mission. If "best" turns out to be impossible to define, it is possible that even a random choice among possible scheduling times for a model in the sequence might result in a superior scheduling method.

# nasa-Asee summer faculty research fellowship program 

## Marshall Space Flight Center

The University of Alabama

Continuous Flow Electrophoresis:
The Effect of Sample Concentration on Throughput and Resolution in an Upward Flowing System

Prepared By:
Academic Rank:
University and Department:

NASA/MSFC:
(Laboratory)
(Division)
(Branch)
MSFC Counterpart:
Date:
Contract No:

Thomas S. Jandebeur, Ph.D.
Assistant Professor
Athens State College Department of Biology

Space Sciences
Space Processing Separation Processes
R. S. Snyder, Ph.D.

August 1, 1980
NGT-01-002-099
(University of Alabama)

Continuous Flow Electrophoresis:
The Effect of Sample Concentration on Throughput and Resolution in an Upward Flowing System

By'
Thomas S. Jandebeur, Ph.D. Assistant Professor of Biology Athens State College Athens, Alabama


#### Abstract

The principle of continuous flow electrophoresis is as follows: a suspension of sample material is introduced as a fine, steady stream into a vertical, flowing "curtain" of electrolyte. A horizontal d.c. electrical field is applied to the curtain. Where a mixture of particles/cells having different electrophoretic mobilities is analyzed, each type particle/cell will assume a different, collectable band position in the curtain electrolyte.


The fundamental behavior of the continuous electrophoretic process in the Beckman Continuous Particle Electrophoresis (CPE) System, with flow in the downward direction, has recently been investigated. The purpose of this research is to determine the effect of sample concentration on throughput and resolution in the CPE System with flow in an upward direction. To maintain upward flow of the curtain buffer and sample stream, and to collect same, certain modifications to the CPE System were made; these also are described.

## ACKNOWLEDGEMENTS

I wish to thank the several individuals who have contributed in a direct way to this research effort: Dr. Robert S. Snyder, Chief, Separation Processes Branch, Space Sciences Division, for having served as my NASA counterpart, and for stimulating my interest in electrophoresis technology; Ms. Jan McGuire, for sharing her understanding of the Beckman CPE System and for guiding my progress; Ms. Teresa Miller, for her practical insight toward the solution of technical problems encountered while "redesigning" the CPE System; and Ms. Evelyn Keenum, for having typed this manuscript.

Thank you, one and all.

## LISI OF TABIES

| Table No. | Title | Page |
| :---: | :---: | :---: |
| 1. | Recovery of Fixed Cells at Indicated Concentration with (Field) and without (Zero) Voltage Applied for First (1) and Second (2) Data Sheets | X-11 |
| 2. | Percent Relative Abundance of Recovered Cells by Tube Number at Indicated Concentration (Cells/ML.) with (Field) and without (Zero) Voltage Applied: First Data Set | X - 12,13 |
| 3. | Percent Relative Abundance of Recovered Cells by Tube Number at Indicated Concentration (Cells/ML.) with (Field) and without (Zero) Voltage Applied: Second Data Set | X - 14, 15 |

## LIST OF FIGURES

| Figure No. | Title | Page |
| :---: | :---: | :---: |
| 1 | Recovery of Fixed Cells vs. Concentration with (Field) and without (Zero) Voltage Applied: First Data Set. | X-16 |
| 2 | Recovery of Fixed Cells vs. Concentration with (Field) and without (Zero) Voltage Applied: Second Data Set. | X-17. |
| 3 | Percent Relative Abundance of Recovered Fixed Cells vs. Tube Number at Indicated Concentration with (Field) and without (Zero) Voltage Applied: First Data Set. | X - 18-21 |
| 4 | Percent Relative Abundance of Recovered Fixed Cells vs. Tube Number at Indicated Concentration with (Field) and without (Zero) Voltage Applied: Second Data Set. | X - 22-25 |
| 5 | Continuous Particle Electrophoresis (CPE) System: Collection Bench for Upward Flow. | X-26 |
| 6 | Continuous Particle Electrophoresis (CPE) System: Collection Bench/Syringe Pump Support Platform. | X-27 |

## INTRODUCTION

As developed by Strickler ${ }^{5}$, Hannig ${ }^{2}$ and their coworkers, continuous flow electrophoresis is a technique for separating small amounts of biological material. In this process, a suspension of sample is introduced as a fine, steady stream into a vertical, flowing "curtain" of electrolyte. A horizontal d.c. field is then applied to the curtain. Where a mixture of particles/ cells having different electrophoretic mobilities is analyzed, each type particle/cell will assume a different, collectable band position in the curtain electrolyte/buffer.

The fluid containing the particles/cells to be separated moves through a thin chamber. The depth of the chamber is dictated by a need to suppress temperature gradients within the chamber and by a need to reduce chamber wall effects, both of which generate undesirable flow characteristics. As a result of these considerations, so-called "narrow gap" ( $0.5 \mathrm{~mm}-3.0 \mathrm{~mm}$ ) chambers lack sufficient sample capacity and/or resolving power, particularly with respect to separating cell populations.

## OBJECTIVES

Though it is unlikely to improve resolution, increasing particle/ cell concentration in the sample should increase sample throughput. However, higher cell concentrations lead to a density mismatch between sample stream and curtain. buffer. Small amounts of sedimenting sample, diluted by the curtain buffer subsequently are rapidly displaced within the curtain. This results in broadening of the sample stream and ultimately in poorer resolution at higher sample concentrations.

Tests conducted by McDonnel1-Douglas ${ }^{6}$ demonstrated upflow machines to be much more sensitive to the described density mismatch than downflow machines, further limiting sample stream concentration, resolution and throughput of material.

The objective of this study, then, is to determine the limits to throughput and resolution effected by sample concentration in the Beckman Continuous Particle Electrophoresis (CPE) System with flow in the upward direction. The data may then be compared to similar data exacted from the same CPE System, but with flow in the downward direction.

To maintain upward flow of the curtain buffer and sample stream, and to collect the sample, certain modifications to the CPE System were necessary; these will be described.

## MATERIALS AND METHODS

To regulate and maintain upward flow of the curtain buffer and sample stream, a 0-140 volt output powerstat was used in tandem with a variable, speed fluid pump. Dilute R-1 buffer ${ }^{4}$ cooled to approximately $4^{\circ} \mathrm{C}$ was pumped through the curtain flow meter at approximately $28 \mathrm{cc} / \mathrm{min}$. with the powerstat set at 90 volts.

Sample injection was accomplished using a Sage Instruments Model 355 Syringe Pump. The sample, contained in a 3 cc syringe fitted with a 21 gauge needle and connected to the sample inlet tube by a short piece of PVC tubing, was injected into the electrophoretic cell at an approximate rate of $1 \mathrm{ml} / \mathrm{hr}$ (range $1 / 100$ at $24.3 \%$ flow). To prevent sedimentation of the sample contained in the syringe, a small stirring bar was placed within the syringe. A magnetic stirrer supported by the syringe pump housing was then inverted over the syringe.

A prototype of the collection system depicted by Figures 5 and 6 facilitated sample collection. The CPE System collecting tubes were connected to the collection bench by 10 inch lengths of 17 gauge PVC tubing fitted with 18 gauge needles at the tubing's distal end. The needle-containing tubes were seated in the 48 hole array of the collection bench. A 48 testtube fraction collector was placed below the collection bench to receive the eluted sample.

The d.c. field was generated by a Hewlett Packard Model 712B power supply; 65 volts/cm constituted the field for each experimental separation.

Concentrated R-1 buffer was used in the electrode rinse and coolant loops. Dilute R-1 buffer ( 800 ml R-1 concentrate added to 1.2 liters of deionized water) was used as curtain buffer.

To establish the effect of concentration on throughput and resolution in this CPE system, fixed turkey and cow red blood cells were used. These cells are stable, well-characterized standards, having different electrophoretic mobilities; further, they are morphologically distinguishable ${ }^{4}$.

## DATA ANALYSIS

Two data sets were collected; the first data set consisting of seven "runs" at sample concentrations ranging from $5 \times 10^{7}$ cells/ml to $2 \times 10^{9}$ cells/m1 (Table $1 ;$ Figure 3 ); the second data set consisting of eight runs at the aforementioned seven concentrations plus one run at a sample concentration of $1.5 \times 10^{9}$ cells/ ml. A 5-minute stabilizing flow followed by 10 minutes of sample collection without (zero) and with (field) voltage applied constitutes a "run".

Aliquots of the collected sample were examined microscopically; recovered cells were counted using an Improved Neubauer hemocytometer. Percent recoveries for each cell type and for the combined cells were then determined (Table 1). Based on the total number of recovered cells of each type and of the combined cells, percent relative abundances of the recovered cells were determined. This data, plotted for both data sets, is presented in Figures 1 and 2 as percent recovery of fixed cells vs. concentration, and in Figures 3 and 4 as percent relative abundance of recovered fixed cells vs. tube number.

## DISCUSSION OF RESULTS

Examination of Table 1 (1st data set) and Figure 1 reveals that, in general, more cells are recovered at lower concentrations than are recovered at higher concentrations, either with (field) or without (zero) voltage applied. This may be explained as a sedimentation effect; i.e., higher cell concentrations engender a density mismatch between sample stream and curtain buffer. Small amounts of sedimenting sample, diluted by the curtain buffer subsequently are rapidly displaced within the curtain. This results in broadening of the sample stream, "spreading" of cells over a broader range of sample collection test tubes (see higher concentrations Tables 2 and 3 and higher concentrations Figures 3 and 4) and ultimately to poorer sample resolution. This interpretation is supported by the lowered recovery of the larger, heavier turkey cells at higher concentrations (Figure 1) and by the enhanced recovery of the smaller, lighter cow cells at higher concentrations (also Figure 1). In part, the apparent increased recovery of cow cells at higher concentrations is, perhaps, due to examining and counting a greater number of sample collection test tubes. In general, the datd of Table 1 (2nd data set) and Figure 2 also fit this interpretation, but with notable exception: the turkey cells do not sediment at higher concentrations. It is, perhaps, significant that the electrophoretic cell was coated with $1 \%$ bovine serum albumin, the electrode membranes and electrode rinse replaced, and the buffer temperatures lowered to approximately $4^{\circ} \mathrm{C}$ between collection of the first and second data sets. These measures, taken to improve CPE chamber performance, may have inadvertently altered the previously noted turkey cell sedimentation behavior.

Tables 2 and 3 and Figures 3 and 4 reveal "spreading" of the sample over a widening range of sample collection test tubes with increasing concentration. This is especially evident at the higher concentrations ( $1.5 \times 10^{9}$ and $2 \times 10^{9} \mathrm{cells} / \mathrm{ml}$ ) at which the sample stream was observed to "fall back" on the inlet tube; periodically, "clouds" of the sample billowed upward in the buffer flow apparently being transported more rapidly than cells remaining in the sample stream.

## CONCLUSIONS

With particular reference to Table 3 and Figure 4, data collected after treating the electrophoretic cell as described in Discussion of Results, and with voltage applied, it appears that maximum resolution in the Beckman CPE System, with flow in an upward direction, is achieved at concentrations ranging from $2 \times 10^{8} \mathrm{cells} / \mathrm{ml}$ to $8 \times 10^{8}$ cells $/ \mathrm{ml}$. The widest peak separation is at $2 \times 10^{9}$ cells/ml; however, the sharpest peaks and least overlap between cell populations is at $8 \times 10^{\text {B }}$ cells $/ \mathrm{ml}$.

Apparently as a result of improved electrophoresis cell performance due to coating the chamber with bovine serum albumin, changing the electrode membranes and rinse, and lowering buffer temperatures, sedimentation effects attending to higher concentrations are diminished.

As measured by recovery of fixed cells, throughput is greatest at lower sample concentrations (ranging from $5 \times 10^{7}$ to $2 \times 10^{8}$ cells/ml in this study), diminished at moderate concentrations ( $4 \times 10^{8}$ to $1 \times 10^{9}$ celis $/ \mathrm{ml}$ ) and increasing at higher concentrations ( $1.5 \times 10^{9}$ to $2 \times 10^{9}$ cells $/ \mathrm{ml}$ ) although not increasing to the higher recovery levels noted for lower sample concentrations.

Throughput as measured by recovery of fixed cells is, then, diminished at the concentrations judged most likely to yield satisfactory resolution. The "tradeoff" appears to be as follows: improved recovery/throughput at the expense of resolution.

## REFERENCES

1. Continuous Particle Electrophoresis System, Beckman Instruments, Inc., Scientific Instruments Division, Fullerton, California; October 1967.
2. Hannig, K., "Preparative Electrophoresis," in Electrophoresis: Theory, Methods and Applications, Vol. II, Ed., M. Bier, Academic Press (1967).
3. Saville, D.A. and Ostrach, S., "The Fluid Mechanics of Continuous Flow Electrophoresis," Final Report on Contract NASA-31399, Department of Chemical Engineering, Princeton University (1978).
4. Seamon, G.V.F., "Electrophoretic Separator Experiments," Final Report on Contract NASA-32609, University of Oregon Health Sciences Center (1979).
5. Strickler, A., "Continuous Particle Electrophoresis: A New Ânalytical and Preparative Capability," Sep. Sci., 2(3):335(1967).
6. Weiss, R.A., "Investigation of the Free Flow Electrophoretic Process", Final Report on Contract NASA-33200, McDonnell-Douglas Astronautics Co., St. Louis, Missouri (1979).

TABLE 1. RECOVERY OF FIXED CEIIS AT INDICATED CONCENIRATION WITH (FIEID) AND WITHOUI (ZERO) VOLTAGE APPLIED FOR FIRST (1) AND SECOND (2) DATA SETS

|  | Data Set | Field |  |  | Zero |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\begin{aligned} & \text { Concentration } \\ & (\mathrm{Cell} / \mathrm{s} / \mathrm{ML}) \end{aligned}$ |  | \% Recovery |  |  | \% Recovery |  |  |
|  |  | Comb. | Turkey | Cow | Camb. | Iurkey | Cow |
| $5 \times 10^{7}$ | 1 | 87.22 | 77.4 | 98.04 | 92.88 | 108.36 | 77.4 |
|  | 2 | 76.38 | 64.06 | 89.7 | 98.56 | 118.28 | 78.84 |
| $1 \times 10^{8}$ | 1 | 65.79 | 69.66 | 61.92 | 74.82 | 98.04 | 51.6 |
|  | 2 | 68.99 | 81.32 | 56.67 | 70.22 | 86.24 | 54.2 |
| $2 \times 10^{8}$ | 1 | 74.18 | 83.84 | 64.5 | 81.27 | 81.28 | 81.28 |
|  | 2 | 69.61 | 80.08 | 59.14 | 53.59 | 46.82 | 60.36 |
| $4 \times 10^{8}$ | 1 | 55.15 | 57.4 | 52.88 | 53.86 | 50.96 | 56.76 |
|  | 2 | 43.12 | 35.72 | 50.52 | 43.12 | 37.58 | 48.66 |
| $8 \times 10^{8}$ | 1 | 48.86 | 45.8 | 51.92 | 42.73 | 42.24 | 43.22 |
|  | 2 | 48.05 | 53.9 | 42.2 | 30.18 | 32.64 | 27.72 |
| $1 \times 10^{9}$ | 1 | 50.31 | 37.66 | 62.96 | 39.22 | 28.9 | 49.44 |
|  | 2 | 29.81 | 15.28 | 44.36 | 48.17 | 28.08 | 68.26 |
| $1.5 \times 10^{9}$ | 2 | 54.7 | 71.95 | 37.45 | 27.6 | 31.7 | 23.48 |
| $2 \times 10^{9}$ | 1 | 53.66 | 26.44 | 80.88 | 41.34 | 29.02 | 53.66 |
|  | 2 | 58.27 | 60.98 | 55.56 | 54.45 | 53.1 | 55.8 |

## $\because$ 要

TABLE 2. PERCENT REIATIVE ABUNDANCE OF RECOVERED CFLIS BY TUBE NLMBER AT INDICATED CONCENTRATION (CEILS/ML.) WITH (FIED) AND WITHOUT (ZERO) VOLTAGE APPLIED FIRST DATA SET

| Tube非 | $5 \times 10^{7}$ |  |  |  | $1 \times 10^{6}$ |  |  |  | $2 \times 10^{8}$ |  |  |  | $4 \times 10^{8}$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Turkey |  | Cow |  | Turkey |  | Cow |  | Turkey |  | Cow |  | Turkey |  | Cow |  |
|  | Zero | Freld | Zero. | Field | Zero | Field | Zero | Field | Zero | Field | Zero | Field | Zero | Field | Zero | Field |
| 21 |  | 0 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 22 |  | 13.33 |  |  |  | 0 |  |  |  |  |  |  |  |  |  |  |
| 23 |  | 13.33 |  | 0 |  | 7.41 |  |  |  |  |  |  |  |  |  |  |
| 24 |  | 46.67 |  | 10.53 |  | 14.82 |  | 0 |  |  |  | 0 |  |  |  |  |
| 25 |  | 13.33 |  | 10.53 |  | 40.74 |  | 4.17 |  | 0 |  | 2.00 |  | 0 |  | 0 |
| 26 |  | 13.33 |  | 31.58 |  | 29.63 |  | 11.11 |  | 7.69 |  | 3.08 |  | 31.46 |  | 1.22 |
| 27 |  | 0 |  | 42.10 |  | 7.41 |  | 29.63 |  | 32.31 |  | 0 |  | 61.80 |  | 4.88 |
| 28 | 4.76 |  |  | 5.26 |  | 0 |  | 40.74 |  | 53.85 |  | 2.00 |  | 6.74 |  | 51.22 |
| 29 | 0 |  | 0 | 0 |  |  |  | 0 |  | 4.62 |  | 29.23 | 0 | 0 |  | 35.37 |
| 30 | 4.76 |  | 20.00 |  |  |  | 0 | 3.70 |  | 1.54 |  | 24.62 | 1.27 |  | 0 | 7.32 |
| 31 | 57.14 |  | 33.33 |  | 0 |  | 5.00 |  | 0 | 0 |  | 15.38 | 35.44 |  | 35.23 | 0 |
| 32 | 19.05 |  | 40.00 |  | 26.32 |  | 20.00 |  | 3.18 | 1.54 | 0 |  | 58.23 |  | 60.23 |  |
| 33 | 14.29 |  | 0 |  | 57.90 |  | 55.00 |  | 31.75 |  | 31.75 |  | 5.06 |  | 4.54 |  |
| 34 | 0 |  | 6.67 |  | 15.79 |  | 10.00 |  | 49.21 |  | 56.92 |  | 0 |  | 0 |  |
| 35 |  |  | 0 |  | 0 |  | 10.00 |  | 15.87 |  | 9.23 |  |  |  |  |  |
| 36 |  |  |  |  |  |  | . |  | . |  | 0 |  |  |  |  |  |

TABIE 2：PERCENT RELATII空 ABUNDANCE OF RECOVERED CEILS BY TUBE NUMBER AT INDICATED CONCENIRATION（CEULS／M．L．） （Cont＇d）WTTH（FIEID）AND WITHOUT（ZERO）VOLTAGE APPLIED FIRST DATA SET

| Tube | $8 \times 108$ |  |  |  | $1 \times 10^{9}$ |  |  |  | $1.5 \times 10^{9}$ |  | $2 \times 109$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Tukey |  | Cow |  | Turkey |  | Cow |  | $\begin{gathered} \text { Turkey } \\ \hline \text { Zero Field } \\ \hline \end{gathered}$ | Cow | Turkey |  |  |  |
|  | Zero | Field | Zero | Field | Zero | Field | Zero | Field |  | Zero Field | Zero | Field | Zero | Field |
| 14 | $\begin{gathered} 0 \\ 1.53 \\ 12.21 \\ 61.83 \\ 24.43 \\ 0 \end{gathered}$ | $\begin{gathered} 0 \\ 8.45 \\ 42.96 \\ 42.25 \\ 6.33 \\ 0 \end{gathered}$ | $\left\|\begin{array}{c} 0 \\ 8.21 \\ 63.43 \\ 28.36 \\ 0 \end{array}\right\|$ | 00.620.620.626.8355.2836.020 | 00.890.8931.2564.292.680 | $\begin{gathered} 0 \\ 2.74 \\ 13.70 \\ 31.51 \\ 30.82 \\ 14.38 \\ 6.16 \\ 0.68 \\ 0 \end{gathered}$ |  0 <br>  0.82 <br>  14.34 <br>  31.97 <br>  32.79 <br> 0 16.80 <br> 0.52 3.28 <br> 40.62 0 <br> 56.77  <br> 1.56  <br> 0.52  <br> 0  |  |  |  |  | 0 |  |  |
| 15 |  |  |  |  |  |  |  |  |  |  | 0.98 |  |  |
| 16 |  |  |  |  |  |  |  |  |  |  | 0.49 |  |  |
| 17 |  |  |  |  |  |  |  |  |  |  | 0 | 0.49 | 0 |  |
| 18 |  |  |  |  |  |  |  |  |  | ¢ | 0.44 | 0 | 0.24 |  |
| 19 |  |  |  |  |  |  |  |  |  |  | 0 | 0.49 | 0.48 |  |
| 20 |  |  |  |  |  |  |  |  | $\stackrel{9}{7}$ | \％ | 1.33 | 0.49 | 1.20 |  |
| 21 |  |  |  |  |  |  |  |  | \％ | F | 1.33 | 1.46 | 0.48 |  |
| 22 |  |  |  |  |  |  |  |  | H | \％ | 0.89 | 5.85 | 1.20 | 0.32 |
| 23 |  |  |  |  |  |  |  |  | H | \％ | 0.44 | 10.73 | 0.96 | 0.16 |
| 24 |  |  |  |  |  |  |  |  | － | 㿾 | 0.44 | 29.27 | 0.24 | 0.96 |
| 25 |  |  |  |  |  |  |  |  | 5 | 5 | 0 | 20.98 | 0.24 | 7.18 |
| 26 |  |  |  |  |  |  |  |  | ． | － | 0 | 17.56 | 0 | 2．5． 52 |
| 27 |  |  |  |  |  |  |  |  | $\stackrel{H}{0}$ | \％ | 0 | 7.32 | 0.96 | 24.88 |
| 28 |  |  |  |  |  |  |  |  | $\underset{ }{H}$ | ${ }^{H}$ | 0.44 | 2.93 | 0.72 | 24.88 |
| 29 |  |  |  |  |  |  |  |  | 5 | \％ | 1.33 | 0.98 | 0 | 14.20 |
| 30 |  |  |  |  |  |  |  |  | ¢ | ¢ | 0.89 | 0 | 1.44 | 1.91 |
| 31 |  |  |  |  |  |  |  |  | 8 | 0 | 2.67 |  | 3.36 | 0.32 |
| 32 |  |  |  |  |  |  |  |  | 9 | － | 9.78 |  | 6.49 | 0 |
| 33 |  |  |  |  |  |  |  |  | 真 | 虫 | 31.56 |  | 30.05 | 0.16 |
| 34 |  |  |  |  |  |  |  |  | 4 | $\pm$ | 35.11 |  | 34.13 | 0 |
| 35 |  |  |  |  |  |  |  |  | U | \％ | 10.22 |  | 11.54 |  |
| 36 |  |  |  |  |  |  |  |  | $\stackrel{-1}{-1}$ | $\stackrel{1}{7}$ | 2.22 |  | 2.64 |  |
| 37 |  |  |  |  |  |  |  |  | ＇ | ＇80 | 0.44 |  | 2.16 |  |
| 38 |  |  |  |  |  |  |  |  | $\checkmark$ | $\pm$ | 0 |  | 0 |  |
| 39 40 |  |  |  |  |  |  |  |  | － | \％ | 0.44 |  | 0.24 |  |
| 40 |  |  |  |  |  |  |  |  | － | － | 0 |  | 1.20 |  |
|  |  |  |  |  |  |  |  |  | － | － |  |  | － |  |

Table 3. PERCENT RELATIVE ABUNDANCE OF RECOVERED CELLS BY TUBE NUMBER AT INDICATED CONCENTRATYN (CELLS/ML.) WITH (FIELD) AND WITHOUT (ZERO) VOLTAGE APPLIED: SECOND DATA 3 ET.


| $\begin{aligned} & \text { 'ube } \\ & \text { \# } \end{aligned}$ | $8 \times 10^{8}$ |  |  |  | $1 \times 10^{9}$ |  |  |  | $1.5 \times 10^{9}$ |  |  |  | $2 \times 10^{9}$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | TURKEY |  | COW |  | TURKEY |  | COW |  | TURKEY |  | COW |  | TURKEY |  | COW |  |
|  | zero | field | zero | field | zero | field | zero | field | zero | field | zero | field | zero | field | zero | field |
| 18 |  |  |  |  |  |  |  |  |  | 0 |  |  |  |  |  |  |
| 19 |  |  |  |  |  |  |  |  |  | 0.68 |  |  |  |  |  |  |
| 20 |  |  |  |  |  |  |  |  |  | 0.46 |  |  |  |  |  |  |
| 21 |  |  |  |  |  |  |  |  |  | 0 |  |  |  | 0 |  |  |
| 22 |  |  |  |  |  |  |  |  |  | 0 |  | 0 |  | 0.20 |  |  |
| 23 |  |  |  |  |  |  |  | - |  | 0.91 |  | 0.88 |  | 0.81 |  |  |
| 24 |  |  |  |  |  |  |  | 1.12 |  | 2.28 |  | 0.44 |  | 0.40 |  |  |
| 25 |  | 0 |  |  |  |  |  | 0 |  | 5.48 |  | 0 |  | 0.81 |  |  |
| 26 |  | 2.29 |  |  |  | 0 |  | 0 |  | 20.78 |  | 0 | 0 | 1.62 |  |  |
| 27 |  | 14.86 |  |  |  | 11.29 |  | 0 |  | 26.94 |  | 1. 75 | 0.23 | 4.24 | 0 | 0 |
| 28 |  | 62.86 |  | 0 |  | 51.62 |  | 0.56 |  | 18.95 |  | 11.84 | 0.23 | 15.35 | 0.44 | 0.22 |
| 29 |  | 15.43 |  | 5.84 |  | 24.19 |  | 0.25 |  | 17.35 |  | 23.25 | 3.02 | 23.03 | 2.43 | 1. 33 |
| 30 |  | 4.57 |  | 65.69 |  | 2.90 |  | 27.53 | 0 | 4.57 | 0 | 26.75 | 10.90 | 45.66 | 12.58 | 12.42 |
| 31 |  | 0 |  | 22.63 |  | 0 |  | 35.39 | 10.35 | 1.37 | 3.50 | 29.39 | 25.99 | 7.68 | 25.39 | 45.23 |
| 32 | 0 |  | 0 | 5.11 | 0 |  | 0 | 32.58 | 1. 04 | 0 | 27.27 | 3.07 | 34.34 | 0.20 | 30.46 | 29.49 |
| 33 | 4.72 |  | 8.89 | 0 | 1.75 |  | 2.89 | 1.12 | 0.52 | 0.23 | 11. 89 | 2.19 | 19.95 | 0 | 22.52 | 9.76 |
| 34 | 83.02 |  | 68.89 | 0.73 | 22.81 |  | 26.35 | 0.56 | 0 | 0 | 3.50 | 0 | 5.10 |  | 6.18 | 1. 55 |
| 35 | - |  | - | - | 62.28 |  | 57.04 | 0 | 56.48 |  | 37.76 | 0.44 | 0.23 |  | 0 | 0 |
| 36 | 12.26 |  | 20.00 |  | 12.28 |  | 12.64 |  | 26.42 |  | 13.29 | 0 | - |  |  |  |
| 37 | 0 |  | 2.22 |  | 0.88 |  | 1. 08 |  | 1.04 |  | 2.10 |  |  |  |  |  |
| 38 |  |  | 0 |  | 0 |  | 0 |  | 2.07 |  | 0 |  |  |  |  |  |
| 39 |  |  |  |  |  |  |  |  | 2.07 |  | 0.70 |  |  |  |  |  |
| 40 |  |  |  |  |  |  |  |  | 0 |  | 0 |  |  |  |  |  |

Figure 1. Percent Recovery of Fixed Cells vs. Concentration with (Field)
and without (Zero) Voltage Applied: First Data Set.


Figure 2. Percent Recovery of Fixed Cells vs. Concentration with (Field) and without (Zero) Voltage Applied: Second Data Set.


Figure 3. Percent Relative Abundance of Recovered Fixed Cells vs. Tube Number at Indicated Concentration With (Field) and Without (Zero) Voltage Applied: First Data Set.
PERCENT RELATIVE ABUNDANCE


Figure 3. Continued


Figure 3. Continued


Figure 3. Continued


Figure 4. Percent Relative Abundance of Recovered Fixed Cells vs. Tube Number at Indicated Concentration With (Field) and Without (Zero) Voltage Applied: Second Data Set.


0
$\qquad$














1

Figure 4. Continued
——————
$+$


Figure 4. Continued
PERCENT RELATIVE ABUNDANCE
(

Figure 5. Continuous Particle Electrophoresis (CPE) System:
Collectiok Bench for Upward Flow.








## 








+1!


$+1+1+1+1+1$ +1+1+1+5 $\frac{1+1}{1+1}$

## $+7+1$



+ H Fit







$1=5=$
$1=0$

| an |
| :---: |
| $=10$ |品

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM 

MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA

MODELING A MAINTENANCE SIMULATION OF THE GEOSYNCHRONOUS PLATFORM

| Prepared by: | A. F. Kleiner, Jr. |
| :--- | :--- |
| Academic Rank: | Professor |
| University and Department: | Drake University <br> Department of Mathematics <br> and Computer Science |
| NASA/MSFC:(Office) <br> (Division) <br> (Branch) | Preliminary Design <br> Mission Analysis and Integration <br> Operations Analysis |
| MSFC Counterpart: | Dr. J. Steincamp |
| Date: | August 1, 1980 |
| Contract No. : | NGT-01-002-099 <br> (University of Alabama) |

by
A. F. Kleiner, Jr. Professor of Mathematics

Drake University
Des Moines, Iowa

## ABSTRACT

With the advent of the Space Transportation System it will be possible to service equipment in space. In particular the effect of a maintenance capability for the Geosynchronous Platform (GSP) needs to be examined. Because of the size and complexity of the GSP, analytic results in this area are beyond reach. Hence a simulation study comparing various maintenance routines appears to be the most fruitful approach. The purpose of this report is to outline the basic model which will be used to conduct such a study. Information is included on the computer model and maintenance. A sample system is also modeled.

## Acknowledgements

The 1980 NASA/ASEE program has provided me with many benefits both professional and personal. I would like to thank the following people for their assistance this summer: Jim Steincamp, my counterpart, for providing me with an interesting problem and the time to understand it; John Cole, Chief of the Operations Analysis Branch for providing support for my work; Dr. Barfield and Marion Kent for the running of the program; Ronnie Barnes, John Springer and Hank Stoll, fellow participants, for a stimulating carpool and friendship for 8 weeks; Dawn Walton, Carol Byler and Janet Coker for typing this report, and most importantly Mary, Alex and Mimi for making a home away from home for the summer.

## 1. Introduction

With the advent of the space transportation system NASA will be able to implement a new approach to the reliability of space missions. Rather than build for the mission life from the start it will be possible to visit spacecraft for the purpose of servicing the equipment. The Geosynchronous Platform (GSP) provides a particularly good example of the possibilities of this service for two reasons: (1) GSP will have an extended mission life, 10 to 16 years, requiring heavy redundance costs if maintenance is not available and (2) GSP will serve the communications industry with projected increases in demand over the course of the mission, thus the ability to increase capability is a valuable option.

An analysis of service for GSP is to be carried out to determine if maintenance is a viable concept and to aid in the scheduling of such if it proves feasible. One approach to this problem is to seek an analytic solution by techniques in decision analysis, such as in [1]. The size of this problem however precludes such an analysis (see for example the "small" sample system in Chapter 3). A numerical approach to the same problem can be developed by means of simulation.

Since the problem of maintenance will arise in other settings besides GSP it is useful to develop a general simulation method. Such a method should allow for the incorporation of special features which arise in individual problems such as the need to increase capacity that is an important part of the GSP. This method should be able to accept different mission types and different maintenance policies and produce data which indicate which policies are better.

This report is a preliminary one on work done to develop such a general method. The method is based on earlier work of Dr. J. Steincamp of the Marshall Space Flight Center. Specific details of the method may be found in [3].

The immediate objective is to construct a simulation model which can be used for an analysis of maintenance for the GSP. It is required to find out if maintenance is an economic approach and if so how it should be carried out.

A more general objective is to develop a method which can answer the above questions for many different space missions. The method is to accept as input the particular system and be capable of running the system through a mission life with different maintenance policies.

This report contains an overview of such a method. The system model is described and illustrated, the basic concepts of a simulation pass are detailed and sections on failures and maintenance are included.

### 3.1 Introduction

An operating, physical system may be considered from many points of view. For the purpose of a maintenance simulation two such points of view are required. When a piece of equipment fails it must be replaced, so its location and accessibility are important. Hence a geometric or physical model of the system which details the replacement relationships among the parts is necessary. More over, in order to carry out a program of part failures on the system, it is necessary to have a model of the system's functional or operational mode.

The operation of the system requires that a physical part fail, the operational effects of this failure be carried out and a determination of the need for maintenance of the system be checked. Hence the physical and operational models must be interrelated. The interrelation will be carried out at the most elementary level of the system by identifying the individual parts which fail, called items, with the lowest level operations carried out by the system, called basic functions. Hence forth the dual nature of these items/basic functions will be emphasized by the name IBF. Thus an IBF can be considered either as a specific physical part which supports a basic function of the system or as a basic operation which is performed by some physical item.

### 3.2 The Physical Model

Inspace fine tuning of systems will not be available in the foreseeable future. Rather, when a part fails some package which contains the failed part will be replaced. This package may contain other parts which have failed as well as good ones. Such a package is called a module. More generally, a module is a replaceable or stockable unit which may contain both items and other modules (sub modules). Each item belongs to same, smallest module. A module which contains only items is called a simple module, one which contains only other modules is called pure and a module which is neither simple nor pure is called mixed. The smallest module which can be used to replace the failed item is said to contain the item.

In a physical system with planned redundancy there will be many instances of duplication. Identical parts will be used in various places and essentially similar but distinct modules will be used for replacement purposes. Although duplicate items will have the same failure distributions, the randomness of the actual failures requires that each physical part be individually represented in the computer model. As a result replacement must be performed and recorded for each module rather than on the classes of similar types.

On the other hand it is desirable to take advantage of the similarities among the items and modules for efficient input of the system. For instance if only separate types of items and modules must be entered in the computer a good deal of labor can be saved. Also the identification of these duplicates can streamline ordering, packing and inventory procedures and costs. This identification is established by the concept of generic items and modules. The essential features of each type of item and module are entered into the machine and these types are referred to as generic. The specific items are then identified with the basic functions of the ope rational model to form the actual IBF's which form the base of the system. It is the IBF's which actually fail according to the failure rule prescribed for each generic item. The actual physical instances of the modules are created by the computer as copies of the appropriate generic module. Records of maintenance are kept for specific modules.

One further feature regarding the physical model of the system is useful. When an IBF fails there are several possibilities for replacement. The IBF may already be scheduled for replacement because it is contained in some module which is to be replaced as a previous failure or because its containing module is a submodule of same module scheduled for replacement. In this case the failure is noted but no adjustment need be made to the list of modules to be replaced. On the other hand if the module containing the IBF was not previously scheduled for replacement it is necessary to add it to the list of modules to be replaced. Moreover any module previously scheduled for replacement whose replacement will be included with the new module should be removed from this list. This procedure is accomplished by distinguishing between kitted and packed modules. A packed module is any module which is scheduled for replacement, either of its own right or because it is contained in some such module. A module is kitted only if it is to be replaced in its own right. Thus, if a module not on the packed list must be replaced, it is to be added to the kitted list, any module which will be replaced with it is removed from the kitted list and all such modules are put on the packed list.

### 3.3 The Operational Model

Each system has one overall mission or function to perform. This function is accomplished by subfunctions each of which contributes something to the operation of the whole system. Usually, each subfunction can be divided further into functions whose composition is the subfunction. This process of dividing functions into lower level functions until functions which are performed by the basic items of the physical model are reached. These functions, which are called basic functions, are identified with the physical items that perform them and form the IBF's upon which the system rests. Thus the operational mode of the system modeled by a treegraph. The nodes correspond to the various functions with the
root node representing the overall system function and the terminal nodes the IBF's.

Each function has a set of attributes which are used for defining the system state, processing failures and keeping records. The tree like structure of the system is defined by the attributes chief and base. The chief of a function is the immediate superior of that function in the tree while the base is the set of immediate subordinates of the function. The duty cycle of a function is that proportion of its chief's operating time which it must operate while the condition indicates if the function exists, is online, in standby or has failed. For the purpose of implementing failures, several attributes such as figure of merit, reserve and a user defined index are useful (these are illustrated in the section on failures). Attributes such as operating life, on line time, number of times failed and time to first failure are useful for record keeping and statistics gathering. Each basic function also has a set of failure distributions which are used to cause failure events (see the section on failures).

### 3.4 A Sample Model

Some of the preceeding concepts are illustrated by the sample model with function tree as displayed in Figures 1-4 and module definitions as in Table 1. This model is for a simplified GSP type system with 119 IBF's, 5 generic modules and 200 functions. Note that in this model several of the modules contain only one IBF and since there is no nesting of modules the distinction between kitted and packed modules does not arise. Each node of the function tree is identified by name and initial duty cycle modifier. Failure rules are included for the non basic functions. The failure distributions for the IBF's are indicated in Table 2.


PARTIAL FUNCTION TREE - SAMPLE SYSTEM


## FIGURE 2



FIGURE 3
XPDR SET DETAILS - SAMPLE SYSTEM



* duty cycle is $66 \%$ until
first failure then $100 \%$.

FIGURE 4
XPDR GROUP ..... 3 ..... 5.2
RCVR PK 1 ..... 2.1
IRA 1 ..... 3.4
FHST ..... 4 ..... 1.2
RW ..... 3 ..... 5.1

Payload Capacity of tug is 7.0 Kgs .
Minimum Capability of system is to be 15 channels over a 10 year life.
Desired statistics are

1. time average of on line capacity (\# of transponders)
2. time to first maintenance
3. number of times maintenance scheduled due to weight
4. number of times maintenance scheduled due to capability
5. sample mean and variance for the randum variables
a. pass duration
b. tug payload weight

## TABLE 1

SAMPLE SYSTEM DEFINITIONS

NUMBER NEEDED
TDA, AMP \& ATTN 36
TWT, AMP \& ISOL 36

RFSW, MVX, FLTR \& ISOL 18
TDA
6

TDAB
4

LOC OSC 4

DRIRU
4

RFSW
3
FHST 3
RW
3
RCVR, INL, SW \& FLTR 2

FAILURE DISTRIBUTION

Expon,.488, . 0488
Expon, 2. 795, . 2795*
Expon, . 811, . 0811
Expon, .4, . 04
Expon, 1.811, . 1181
Expon, 1.6, . 16
Expon, 15.0
Expon, . 025
Expon, 4.411, . 4411
Expon, 15.0
Expon, .095,

* Also a drop dead time of 26,298 hours

TABLE 2

IBF DEFINITIONS FOR SAMPLE SYSTEM

### 4.1 Introduction

As previously indicated, the complexity of this problem suggests an analysis by simulation. For this purpose, the operation of the system across time is best modeled by a discrete event approach with two basic events - failure and maintenance of the system. Various languages are available for such a simulation. In particular, the language SLAM with its discrete event capability and its built-in use of FORTRAN is appropriate in this context (See [2 ]).

Each overall simulation run will consist of introducing a particular model of the physical system, together with a maintenance policy, demand function and mission lifetime. The system is then run through many passes, each pass corresponding to one mission and the model is re-initialized before each pass. Statistics are compiled at the end of each pass and after the last pass a report is printed. Items of interest typically include the time to first maintenance, total number of maintenance trips for each pass, average capability of the system, etc. The remainder of this section will treat PASS, the subroutine package which completes each simulation pass. The problem of loading the system model is a significant one and has not yet been fully developed - see the section on recommendations and conclusions.

### 4.2 PASS

The PASS routine sets the system in initial condition, schedules failures of IBF's, carries out failures, schedules maintenance, when required, performs maintenance, increases capability in response to demand, terminates the mission and keeps records.

Several assumptions about the computer model need to be detailed. The model is to be capable of increasing the system capability in response to demand. This is done by building the model to its full size at the beginning and introducing a condition DOES NOT EXIST for excess equipment. This equipment is then "brought into existence" during maintenance missions as it is needed. Additionally, some procedure must be chosen for terminating a pass. The standard procedure for this method will be to discontinue maintenance when the mission horizon is reached and then to terminate the mission when the system figure of merit drops to some cutoff point.

Table 3 gives the basic flow ideas of PASS and its major subroutines. These ideas are amplified in the pages following Table 3.

INITIALIZE SAMPLE TIME TO FAIL FAIL MAINT DEMAND MAXTIME STAT

| FAIL | MAINT | DEMAND |
| :--- | :--- | :--- |
|  |  |  |
| IMPACT | STORE | TURN ON |
| STAT | HEAL | UPDATE |
| TIME TO FAIL | MODIFIED FAIL | SDEMAND |
| SMAINT | ADDON | TIME TO FAIL |
|  | SAMPLE | STAT |
|  | TIME TO FAIL |  |
|  | STAT |  |

TABLE 3

| INITIALIZE | Sets the model into its initial condition at the start of each pass, each IBF is set to its proper starting condition, duty cycles are computed. |
| :---: | :---: |
| SAMPLE | Input is a set of IBF's, output is the time to fail in on-line (TTFO) and standby (TTFS) modes for each IBF in the input set. |
| TIME TOFAIL | Input is a set of IBF's, output is the time to next failure for this IBF (see section on Failures for details). |
| FAIL | An event subroutine which carries out the details of a failure (see Table 3 and below). |
| MAINT | An event subroutine which carries the details of maintenance (see Table 3 and below). |
| DEMAND | An event subroutine which carries out the details of "creating" new capacity for the GSP (see Table 3 and below). |
| MAXTIME | When the mission horizon is reached this routine terminates all maintenance and sets a test for mission termination. |
| STAT | A general routine which uses the SLAM statistics collection capability to record desired statistics. |
| IMPACT | Uses actions to carry out the effect of failures on the function trees (see section on Failures). |
| SMAINT | Uses a given maintenance policy to test the need for maintenance and to schedule maintenance when required. |
| STORE | Keeps the current records of those failed IBF's which will not be replaced during the current maintenance. |
| HEAL | Similar to initialize, however, all functions, including those which have been added in response to demand, are set to their initial state. |
| MODIFIED FAIL | A reduced version of FAIL, this routine refails all IBF whose records have been saved in STORE, no records are kept and no new failure times are required. |


| ADD ON | All equipment which is "delivered" on the current <br> maintenance mission is "brought into existence" and <br> statistics gathering is started. |
| :--- | :--- |
| TURN ON | New equipment is brought on -line in response to <br> increasing demand. |
| UPDATE | The system definition which is used to HEAL the <br> system for maintenance is brought up to date. |
| SDEMAND |  |
|  | The known demand curve is used to schedule the <br> next DEMAND. |

## 5. Failures

### 5.1 IBF Failures

In general, each IBF has a failure distribution with known parameters. Sample time to fail are drawn from this distribtuion for both on-line (TTFO) and standby (TTFS) modes. The proportion of on-line time (i.e., the duty cycle), $\alpha$, is then used to project the time to fail in each mode. The Red Life Method of failure is assumed here (see [ 4 , page 16]) so that the minimum of these prorated times is used as the time to next failure (TTF) of the IBF. Thus,

$$
\operatorname{TTF}=\operatorname{MIN}\{\operatorname{TTFO} / \alpha, \operatorname{TTFS} /(1-\alpha)\}
$$

Variations of this rule occur when the IBF will not fail in standby mode, TTFS = when it is on, $\alpha=1$, or off, $\alpha=0$, over an entire interval or when a drop dead time (DDT) requires that the item must fail after a fixed time. These variations are handled by adding or dropping terms in the preceeding formula.

I some other IBF fails first, say after $\Delta T$ time units, then the failure times are adjusted by:

$$
\operatorname{TTFO}=\operatorname{TTFO}-\alpha(\Delta \mathrm{T}), \operatorname{TTFS}=\operatorname{TTFS}-(1-\alpha)(\Delta \mathrm{T}), \mathrm{DDT}=\mathrm{DDT}-(\Delta \mathrm{T})
$$

where $\alpha$ is the duty cycle of the preceding period. The computation then proceeds as before.

### 5.2 Failure Propagation

When an IBF fails, a message announcing this failure must be transmitted to its chief. The chief must then react to this message, perhaps issuing orders to its subordinates or send messages further up the function tree. This process must be repeated at each function which receives a message until all effects of the original failure are carried out.

These messages are delivered by means of actions. An action is an internally generated message which includes an ORIGIN (the originating function), a TARGET (the receiving function) and a CMD/CAUSE (a command to change state if the action is directed to a subordinate, a cause if the action is directed to a chief).

Several assumptions are made concerning actions. First among these is the assumption that no function will issue a command to a subordinate unless the subordinate is capable of carrying out the command. Another is that downward actions are always carried out before an upward action. (Actions are filed in a last in - first out set called ACTS).

### 5.3 A Standard Failure Rule

Let $F$ be a function with a $n /(\ell+m)$ failure rule. That is, $F$ has ( $\ell+n$ ) identical subordinates and requires that $n$ be in an up condition for $F$ to be up. Further, in the initial state $\ell$ of the subordinates are up ( $\ell \geqslant n$ ) and $m$ have condition ready. The first $m$ failures have no effect on the operation of $F$, the next $l-n$ failures each cause a decrease in the figure of merit of $F$ while the ( $\ell+m-n+1$ ) th failure causes $F$ to fail. The attributes INDEX, FOM and RESERVE are used as follows: INDEX is a user-defined index which assigns to each subordinate of $F$ an integer such that in the initial state the first $\ell$ subordinates are up and the last $m$ are ready; $F O M$ is the figure of merit of $F$, if $F$ is up $F O M(F)$ is the number of subordinates which are on, otherwise $F O M(F)=0$; RESERVE is a pointer which indicates the lowest index of a subordinate in a ready state, when no such subordinates remain, RESERVE is set to zero. Thus, if $F O M(F) \geqslant N$, and $S$ is a subordinate of $F$, then $\operatorname{INDEX}(S)$ is less than RESERVE(F) if $\operatorname{COND}(S)=U P$ and greater than or equal to RESERVE(F) if COND(S) $=\operatorname{READY}$. The function $F$ can receive two kinds of notices: a failure notice from a subordinate or a notice to go to a ready state from its chief. The next several tables outline the flow ideas in these two cases.

### 5.4 A Special Failure Rule

In the sample system of section 3, the XPDRSET's have a special failure rule. Each XPDRSET starts with six channels, five on-line and one standby. When a first channel fails, the standby, if still available, is turned on. When a second fails, the XPDRSET forwards an action to its chief asking that a channel in some other set be turned on if any standbys are left in the system. This can be implemented by using the figure of merit for each XPDRSET to be the number of channels on-line and the reserve to be the number in standby mode. The same meanings are attached to the attributes for each XPDR GROUP while the attributes of ME are the sum of the attributes of all these XPDR GROUPS.

IF FOM(F) $=0$ THEN RETURN
ELSE
IF FOM(F)=N THEN GO TO FAIL

## ELSE

        \(\operatorname{IF} \operatorname{RESERVE}(F)=0\) THEN \(\operatorname{FOM}(F)=F O M(F)-1\)
        ELSE
                IF INDEX(S) > RESERVE(F) THEN RETURN
        ELSE
                IF THERE IS A T IN BASE(F) WITH
    \$ $\quad \operatorname{COND}(\mathrm{T})=\operatorname{READY}$ AND INDEX(T)>
\$ RESERVE(F) THEN
IF RESERVE(F) < INDEX(S) THEN CREATE
\$ AN ACTION AS FOLLOWS
$\operatorname{TARGET}(\operatorname{ACTION})=\operatorname{RESERVE}(\mathrm{F})$
ORIGIN (ACTION) $=F$
CMD (ACTION) $=\mathrm{UP}$













* this failure occurred when $F$ was not up
* this caused F to fail
* no standby left, F does not fail
* a stanby beyond the pointer failed
* at least two standbys before the failure
* failure was on-line, turn on first standby
this failure occurred when $F$ was not up
* this caused F to fail
no standby left, F does not fail
* a stanby beyond the pointer failed
* at least two standbys before the failure
* move pointer to next standby
RESERVE(F) is only standby
* failure was on-line, turn on reserve
* the FAIL part of the standard failure rule


## FAIL:

## $\operatorname{COND}(F)=$ FAIL

$F O M(F)=0$
CREATE AN ACTION AS FOLLOWS
TARGET (ACTION) $=$ CHIEF $(F)$
ORIGIN $(A C T I O N)=F$
CAUSE $($ ACTION $)=$ FAIL
FILE ACTION IN ACTS
FOR EACHS IN BASE(F) WITH COND = UP
\$ CREATE AN ACTION AS FOLLOWS
TARGET $(A C T I O N)=S$
ORIGIN (ACTION) $=F$
CMD (ACTION) = READY
FILE ACTION IN ACTS

* adjust attributes of $F$
* send action to chief
* send action to all subfunctions which are up

RETURN

* $F$ is sent an action to go to ready state
$\operatorname{COND}(F)=$ READY
$\operatorname{FOM}(F)=0$
FOR EACH S IN BASE(F) WITH COND = UP
\$ CREATE AN ACTION AS FOLLOWS
TARGET (ACTION) $=S$
ORIGIN $(A C T I O N)=F$
CMD (ACTION) $=$ READY
FILE ACTION IN ACTS
RETURN
* adjust attribute of $F$
* send action to all subordinates which are up
send action to all subordinates which are up

TABLE 6
DOWNWARD ACTIONS FOR STANDARD FAILURE RULE

### 6.1 Introduction

It is assumed that demand, $D$, for the GSP capacity will increase as a function of time and that the demand is known in advance. The minimum acceptable capacity, MC, is typically of the form $D+$ constant where the constant is a "cushion" designed to avoid penalty costs. The nominal capacity, NC, is a desirable level for ordinary purposes. As time increases NC - MC will typically decrease (refer to Figure 5).

At the initial time of the mission, capacity will be NC (0). As time increases failures will occur. Some of these failures will be in the support system and will not cause any decrease in capacity, others will cause such a decrease. For example, in Figure 5 the broken line gives the actual capacity of the system (AC). Capacity failures occur at points 1 and 3, subsystem failures at 2 and 4.

The intent of the simulation is to minimize the cost of maintenance. In order to do this, two sources of cost are considered: (1) the cost of the delivery system, and (2) the inventory cost of delivering capacity before it is needed. The major cost of the delivery system is independent of payload size, thus a trip with less than full payload is inefficient. There exists one or more size limits on the payload (e.g., weight, volume) and these are grouped under the heading size limit. When the first of these is reached, the payload is full. It is never possible to fly with a payload in which the size limit is exceeded.

The inventory cost is kept small by keeping the capacity between the minimal and nominal capacity. Because penalties occur when the capacity is below the minimum the requirement $A C \geqslant M C$ will be enforced. Thus, whenever the payload is full or the actual capacity drops to the minimum, a service mission is scheduled. Since it is desirable to keep the payload as large as possible and the capacity as close to nominal capacity as possible, two different policies arise. It is assumed that there exists a supply list which contains, in the order needed, the items necessary to bring the mission to nominal capacity.

### 6.2 Size Constrained Policy

When a failure occurs the proper replacement parts are loaded and a test is made to determine if the size limit has been reached. If so, the service mission is launched. If the size limit has not been reached, the current value of $A C$ is compared to MC. If AC = MC all available space is filled with modules from the supply list and the mission is launched. Under this policy, a service mission always flys a full payload but no relation between AC and NC can be drawn.

### 6.3 Capacity Constrained Policy

When a failure occurs the proper replacement parts are loaded and so are those items on the supply list which will bring the capacity up to NC at the time of failure. Size and capacity tests are then checked to see if a mission should be launched at the current time. If a size limit is not reached before the actual capacity drops to MC, the launch then takes place when AC = MC. This approach always brings the actual capacity of the system up to the nominal capacity but may produce less than full payloads.

### 6.4 An Additional Problem

Since failures come at distinct points in time with discrete jumps in capacity (and corresponding payload) it may be that at some point the payload is below the size limit and AC is above MC, however, at the next failure either AC is less than MC or the payload is greater than the size limit. Since minimal capacity is demand plus a cushion, the mission should be flown at once with no change of plans if AC falls below MC. In the case that payload has exceeded the size limit, a problem arises since the size limit is an absolute constraint, the mission cannot be flown if it is violated. In this case, the procedure will be to remove the last package(s) which caused the limit to be exceeded, mark these packages as first on the next service mission and fly the current mission.


TIME
FIGURE 5
CAPACITY CURVES

## 7. Conclusions and Recommendations

The problem of maintenance of space missions is a serious one and the method outlined herein is a useful tool for approaching this problem. The particular problem of GSP maintenance is a significant one and will be a good test of the method.

Development of tools to solve the GSP problem, starting with the sample system of Chapter 3, will provide significant progress toward the general method.

Among the problems which arise in the implementation of the method one in particular deserves special study. Just as many modules have the same features and can be built into the computer model by the generic module concept, so to many functions one duplicates. A concept of generic function would be useful in this respect (one estimate is that the final GSP model will have 30 identical XPDR GROUPs rather than the 3 of the simple system). In the general case it cannot be assumed that identical basic functions will belong to the same module types. Thus any concept of generic function must allow for the placement of specific IBF's in different modules.

It is recommended that work on this problem continue in two directions. First that a simulation package actually be built and run for GSP and second that work on the general method continue beyond the GSP version of this problem.

## REFERENCES

[1] C. Derman, Finite State Markovian Decision Processes, Academic Press, (New York) 1970.
[2] A. A. B. Pritsker and C. D. Pegden, Introduction to Simulation and SLAM, John Wiley and Sons (New York) 1979.
[3] J. Steincamp, 1.2 M XRO - OPSIM-, Unpublished Notes, Marshall Space Flight Center.
[4] S. Sugihara and H. Wong, System Availability Analysis Aerospace Report No. TOR-0066 (5129-01) -17, 1970.
NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM MARSHALL FLIGHT CENTER THE UNIVERSITY OF ALABAMA
AXIALLY SYMMETRIC MOTION OF A ROTATING STRATIFIED FLUID
PREPARED BY: Benjamin J. Martin, Ph. D.
ACADEMIC RANK: Professor
COLLEGE AND DEPARTMENT: Atlanta University Mathematics
NASA/MSFC:(OFFICE)(DIVISION)(BRANCH)
Space Sciences Laboratory Atmospheric Sciences Fluid Dynamics
MSFC COUNTERPART:
Dr. William Fowlis
DATE:
July 29, 1980
CONTRACT NO:
NGT-01-002-099
(University of Alabama)

## TABLE OF CONTENTS

Abstract ..... XII - 1
Acknowledgements . ..... XII-2
Nomenclature ..... XII - 3
Introduction ..... XII-5
The Governing Equations ..... XII-6
The Solution of the System ..... XII-8
Conclusions and Recommendations ..... XII- 13
References ..... XII-15

## By

Benjamin J. Martin, Ph. D. Professor of Mathematics Atlanta University Atlanta, Georgia

## ABSTRACT

A spherical model of the general circulation of the Earth's atmosphere is planned for Spacelab. In order to provide adequate support for the experiment, analytical and numerical studies and simulations must be done. Therefore, an analysis of an earlier effort by Joseph Pedlosky in the area of spherical modeling of geophysical fluid flow is examined. Possible extension of his work is considered and alternative approaches suggested. Recommendations are made for future investigations.

## ACKNOWLEDGEMENTS

The writer wishes to express his gratitude to NASA and to the University of Alabama for the opportunity to participate in this program. The writer also acknowledges the assistance of $\mathrm{D}_{\mathrm{r}}$. William Fowlis and $\mathrm{D}_{\mathrm{r}}$. Basil Antar in branching out into a new area, namely, geophysical fluid dynamics.

| $\Omega$ | magnitude of the angular velocity vector |
| :---: | :---: |
| R | radius of inner sphere |
| D | distance between inner and outer spheres |
| $\checkmark$ | kinematic viscosity |
| $\mathrm{V}_{0}$ | velocity scale factor |
| $E=\frac{\nu}{2 \Omega D^{2}}$ | Ekman number $\left(=\frac{1}{2} E_{V}\right)$ |
| $\varepsilon=\frac{V_{0}}{2 \Omega R}$ |  |
| $\Delta T$ | temperature difference between inner and outer spheres |
| $\alpha$ | coefficient of thermal expansion |
| $x$ | thermal diffusivity |
| $\sigma=v / x$ | Prandtl number |
| g | gravitational constant |
| $\beta=\frac{g \alpha \Delta \Delta T}{4 \Omega^{2} R^{2}}$ | thermal Rossby number $\left(=\frac{1}{4} \varsigma\right)$ |
| T | temperature (scaled) |
| p | pressure (scaled) |
| $p$ | density |
| $S=D / R$ | aspect rates |
| u | eastward velocity (scaled) |
| $r$ | northward velocity (scaled) |
| $\omega$ | vertical velocity (scaled) |

$$
\begin{array}{ll}
\varphi & \text { longitude } \\
\theta & \text { latitude } \\
\mu \equiv \sqrt{\frac{1}{2}|\sin \theta|} \\
\operatorname{sen} \theta \equiv \begin{cases}1, & \theta>0 \\
-1, & \theta<0\end{cases} \\
\phi^{\prime} \equiv \frac{d \phi}{d \theta}
\end{array}
$$

## 1. INTRODUCTION.

When the Space Shuttle becomes operational some of the flights will take into orbit a laboratory known as Spacelab. A principal use of Spacelab will be the exploitation of the low gravity environment of the shuttle for scientific and technological investigations. In particular, it will be possible to perform geophysical fluid flow model experiments in true spherical geometry. A radial dielectric body force, which is analogous to gravity, can be achieved over a volume of liquid held between two concentric spheres. However on the Earth's surface the dielectric force cannot be made large enough to dominate the effect of terrestrial gravity. Hence the low-gravity environment of Spacelab is the ideal location for these experiments.

In order to adequately design the experiment for Spacelab, preliminary theoretical and numerical work must be done. Specifically, the scientific design program for the Atmospheric General Circulation Experiment (AGCE) requires that some baroclinic stability studies in spherical geometry be tackled. This can be done only after a basic state has been established. Simplified models must be used since otherwise the equations are nonseparable.

The particular task here is the extension of some previous work done in spherical geometry by Joseph Pedlosky to the specific case of interest in AGCE. This involves changing the boundary conditions under which the basic equations are to be solved. For simplicity the linear, axially symmetric steady state solution is sought.

## 2. THE GOVERNING EQUATIONS.

The general governing equations for an incompressible fluid in dimensional form are

The Momentum Equation

$$
\begin{aligned}
& \text { Momentum Equation } \\
& \frac{\partial}{\partial t} \bar{v}+(\bar{v} \cdot \nabla) \bar{V}+2 \bar{\Omega} \times \bar{v}=-\frac{1}{\vec{p}} \nabla \bar{p}+\bar{g}+\nu \nabla^{2} \bar{V}
\end{aligned}
$$

The Thermodynamic Equation

$$
\frac{\partial}{\partial t} \tilde{\tau}+(\bar{V} \cdot \nabla) \tilde{T}=u \nabla^{2} \tilde{T}
$$

The Continuity Equation

$$
\nabla \cdot \bar{V}=0
$$

and The Equation of State for a liquid

$$
\tilde{\rho}=\rho_{t}\left(1-\alpha\left(\tilde{T}-T_{0}\right)\right)
$$

Thus we have six equations in six unknowns $\bar{V}=\left(\tilde{V}_{r}, \tilde{V}_{\theta}, \tilde{V}_{\varphi}\right), \tilde{T}, \tilde{\rho}$, and $\tilde{\rho}$. Note also that the equations are non-linear. For the model under considertron the steady-state axially-symmetric case is of interest so that the operators are two -dimensional involving only $r$ and $e$ in spherical geometry:

$$
\begin{aligned}
& (\bar{V} \cdot \nabla) \bar{V}+2 \bar{\Omega} \times \bar{V}=-\frac{1}{\bar{p}} \nabla \tilde{p}+\bar{q}+\nu \nabla^{2} \bar{V} \\
& (\bar{V} \cdot \nabla) \tilde{T}=x \nabla^{2} \tilde{T} \\
& \nabla \cdot \bar{V}=0 \\
& \tilde{\rho}=\rho_{0}\left(1-\alpha\left(\tilde{T}-\bar{T}_{c}\right)\right) .
\end{aligned}
$$

Introducing non-dimensional variables

$$
\begin{aligned}
& \tilde{r}=R+\delta R z, \delta=D / R \\
& \tilde{V}_{r}=\delta V_{0} u^{\prime}(\theta, z) \\
& \tilde{V}_{\theta}=V_{0} V(\theta, z)
\end{aligned}
$$

$$
\begin{aligned}
& \tilde{V}_{\varphi}=V_{0} u(\theta, z) \\
& \tilde{T}=T_{0}+\Delta T z+\left(\partial \Omega R V_{0} \mid g \alpha D\right) T(\theta, z) \\
& \tilde{p}=p_{0}-\rho_{0} g D z+\frac{1}{2} g \alpha D \Delta T \rho_{0} z^{2}+2 \Omega V_{0} R P_{0} P(\theta, z)
\end{aligned}
$$

$$
\begin{aligned}
& \text { the equations become } \\
& \varepsilon\left[\omega \frac{\partial u}{\partial z}+v \frac{\partial u}{\partial \theta}+\delta u \omega-u v \tan \theta\right]+(\omega \delta \cos \theta-v \sin \theta) \\
& =\delta^{2} E\left[\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial u}{\partial \theta} \cos \theta\right)-\frac{u}{\cos ^{2} \theta}\right]+E\left[\frac{\partial^{2} u}{\partial z^{2}}+2 \delta \frac{\partial u}{\partial z}\right] \\
& \varepsilon\left[\omega \frac{\partial v}{\partial z}+v \frac{\partial v}{\partial \theta}+\delta v \omega+u^{2} \tan \theta\right]+u \sin \theta=-\frac{\partial P}{\partial \theta}+2 \delta^{3} E \frac{\partial \omega}{\partial \theta} \\
& +\delta^{2} E\left[\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial u}{\partial \theta} \cos \theta\right)-\frac{v}{\cos ^{2} \theta}\right]+E\left[\frac{\partial^{2} v}{\partial z^{2}}+2 \delta \frac{\partial v}{\partial z}\right] \\
& \varepsilon \delta\left[\delta \omega \frac{\partial \omega}{\partial z}+\delta r \frac{\partial \omega}{\partial \theta}-\left(u^{2}+v^{2}\right)\right]-\delta u \cos \theta=-\frac{\partial P}{\partial z}+T-\partial \delta^{4} E w \\
& +\delta^{4} E\left[\frac{1}{\operatorname{coses}} \frac{\partial}{\partial \theta}\left(\frac{\partial \omega}{\partial \theta} \operatorname{cose} \theta\right)\right]+\delta^{2} E\left[\frac{\partial^{2} \omega}{\partial z^{2}}+2 \delta \frac{\partial \omega}{\partial z}\right] \\
& -2 \delta^{3} E\left[\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}(v \cos \theta)\right] \\
& \beta \omega+\varepsilon\left[\omega \frac{\partial T}{\partial z}+V \frac{\partial T}{\partial \theta}\right]=\frac{\partial N \beta}{V_{0} D}+\frac{\delta^{2} E}{\sigma}\left[\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial T}{\partial \theta} \cos \theta\right)\right] \\
& +\frac{E}{\sigma}\left[\frac{\partial^{2} T}{\partial z^{2}}+2 \delta \frac{\partial T}{\partial z}\right] \\
& \frac{\partial \omega}{\partial z}+2 \delta \omega+\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}(v \operatorname{cose})=0
\end{aligned}
$$

For $\varepsilon$ sufficiently small the above equations can be linearized

$$
\begin{aligned}
& \delta u \cos \theta-v \sin \theta=\delta^{2} E\left[\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial u}{\partial \theta} \cos \theta\right)-\frac{u}{\cos \theta \theta}\right] \\
& +E\left[\frac{\partial^{2} u}{\partial z^{2}}+\partial S \frac{\partial u}{\partial z}\right] \\
& u \sin e=-\frac{\partial p}{\partial \theta}+\partial \delta^{3} E \frac{\partial \omega}{\partial \theta}+\delta^{1} E\left[\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial v}{\partial \theta} \cos \theta\right)-\frac{V}{\cos ^{2} \theta}\right] \\
& +E\left[\frac{\partial^{2} v}{\partial z^{2}}+\partial \delta \frac{\partial v}{\partial z}\right] \\
& -\delta u \cos \theta=-\frac{\partial P}{\partial z}+T+\delta^{4} E\left[\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial \omega}{\partial \theta} \cos \theta\right)\right]+\delta^{2} E\left[\frac{\partial^{2} \omega}{\partial z \theta}+\partial \delta \frac{\partial \omega}{\partial z}\right] \\
& -\partial \delta^{4} E w-\partial \delta^{3} E\left[\frac{1}{\operatorname{cose}} \frac{\partial}{\partial \theta}(\operatorname{vas} \theta)\right] \\
& \beta \omega=\frac{\partial x \beta}{V_{0} D}+\frac{\delta^{2} E}{\sigma}\left[\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial T}{\partial \theta}(\cos \partial)\right]+\frac{E}{\sigma}\left[\frac{\partial^{2} T}{\partial z^{z}}+\hat{\partial} \delta \frac{\partial T}{\partial z}\right]\right. \\
& \frac{\partial \omega}{\partial z}+2 \delta \omega+\frac{1}{\cos \theta}(v \cos \theta)=0 .
\end{aligned}
$$

These are the equations of motion in terms of non-dimensional variables which are appropriate for linear, axially-symmetric, steadystate motion of an incompressible fluid in spherical geometry.
3. THE SOLUTION OF THE SYSTEM.

The system that is used here is the system as derived and simplified by Pedlosky [1] with some obvious corrections and a few notational changes

$$
\begin{aligned}
-V \sin \theta+\delta \omega \cos \theta= & \delta^{2} E
\end{aligned} \begin{aligned}
&\left.-\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial u}{\partial \theta} \cos e\right)-\frac{u}{\cos ^{2} \theta}\right] \\
&+E\left[\frac{\partial^{2} u}{\partial z^{2}}+\partial S \frac{\partial u}{\partial z}\right] \\
& u \sin \theta=-\frac{\partial p}{\partial \theta}+\delta^{2} E {\left[\frac{1}{\cos ^{2}} \frac{\partial}{\partial \theta}\left(\frac{\partial v}{\partial \theta} \cos \theta\right)-\frac{v}{\cos ^{2} \theta}\right] } \\
&+E\left[\frac{\partial^{2} v}{\partial z^{2}}+\partial S \frac{\partial v}{\partial z}\right]
\end{aligned}
$$

$$
\begin{aligned}
& -\delta u \cos \theta=-\frac{\partial P}{\partial z}+T+\delta^{4} E\left[\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial \omega}{\partial \theta} \cos \theta\right)\right] \\
& +\delta^{2} E\left[\frac{\partial^{2} \omega}{\partial z^{2}}+\partial \delta \frac{\partial \omega}{\partial z}\right] \\
& \beta \omega=\frac{\delta^{2} E}{\sigma} \frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left(\frac{\partial T}{\partial \theta} \cos \theta\right)+\frac{E}{\sigma}\left[\frac{\partial^{2} T}{\partial z^{2}}+2 \delta \frac{\partial T}{\partial z}\right] \\
& \frac{\partial \omega}{\partial z}+\partial \delta \omega+\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}(v \cos \theta)=0 \text {. }
\end{aligned}
$$

The boundary conditions are

$$
\begin{aligned}
& (u, v, w)=(0,0,0) \text { an } z=0 \text { and } z=1 \\
& T(\theta, 0)=H_{L}(\theta) \\
& T(\theta, 1)=H_{T}(\theta)
\end{aligned}
$$

In order to solve the system with reasonable effort, the solution is decomposed into three parts --the interior motion (the I-subscripted variables), the lower boundary layer corrections (the L-subscripted variables), and the upper boundary layer corrections. Therefore

$$
\begin{aligned}
& u(\theta, z)=u_{I}(\theta, z)+u_{T}(\theta, x)+u_{L}(\theta, y) \\
& v(\theta, z)=E v_{I}(\theta, z)+V_{T}(\theta, x)+v_{L}(\theta, y) \\
& w(\theta, z)=E^{1 / z}\left[\omega_{I}(\theta, z)+\omega_{T}(\theta, x)+\omega_{L}(\theta, y)\right] \\
& P(\theta, z)=P_{I}(\theta, z)+\delta E^{1 / 2}\left[P_{T}(\theta, x)+P_{L}(\theta, y)\right] \\
& T(\theta, z)=T_{I}(\theta, z)+E^{1 / 2} \sigma \beta\left[T_{T}(\theta, x)+T_{L}(\theta, y)\right]
\end{aligned}
$$

where $x=(1-z) E^{-1 / 2}$ and $y=z E^{-1 / 2}$. Since the $L$ and $T$ subscripted variables are signicant only near the boundary, it is required that they tend to zero for large $y$ and $x$.

The equations for the lowest order correction functions in the respective boundary layers are

Lower Boundary Layer

$$
\begin{align*}
-V_{L} \sin \theta & =\frac{\partial^{2} u_{L}}{\partial y^{2}}  \tag{3,1}\\
u_{L} \sin \theta & =\frac{\partial^{2} V_{L}}{\partial y^{2}} \\
-\delta u_{L} \cos \theta & =-\delta \frac{\partial P_{L}}{\partial y}+E^{1 / 2} \beta T_{L} \\
\omega_{L} & =\frac{1}{2} \frac{\partial^{2} T_{L}}{\partial y^{2}} \\
\frac{\partial \omega_{L}}{\partial y} & =-\frac{1}{\cos e} \frac{\partial}{\partial \theta}(3.3)
\end{align*}
$$

Upper Boundary Layer

$$
\begin{aligned}
&-V_{T} \sin \theta=\frac{\partial^{2} u_{r}}{\partial x^{2}} \\
& u_{L} \sin \theta=\frac{\partial^{2} V_{r}}{\partial x^{2}} \\
&-\delta u_{T} \cos \theta=S \frac{\partial P_{T}}{\partial x}+E^{1 / 2} \sigma \beta T_{r} \\
& \omega_{T}=\frac{1}{2} \frac{\partial^{2} T_{r}}{\partial x^{2}} \\
& \frac{\partial \omega_{T}}{\partial y}=\frac{1}{\cos e} \frac{\partial}{\partial \theta}\left(V_{T} \operatorname{cose}\right)
\end{aligned}
$$

Similarly equations for the interior solution are obtained to lowest order in $E$

$$
\begin{gather*}
-V_{I} \sin \theta=\frac{\partial^{2} u_{I}}{\partial z^{2}}  \tag{3,6}\\
U_{I} \operatorname{sen} \theta=-\frac{\partial P_{I}}{\partial \theta}  \tag{3.7}\\
T_{I}=\frac{\partial P_{I}}{\partial z} \\
\partial \sigma \beta E^{-1 / 2} \omega_{I}=\frac{\partial^{2} \Gamma_{I}}{\partial z^{2}} \\
\frac{\partial \omega_{I}}{\partial z}=0 \tag{3,10}
\end{gather*}
$$

where, following Pedlosky, we have assumed $\delta \ll E^{1 / 2}$ and have dropped all terms involving powers of $\delta$.

The Boundary Layer Corrections.
The system of equations for the lower boundary layer corrections can be easily solved. First the coupled equations (3.1) and (3.2) are solved. This allows the solution of equation (3.5), which then permits the solution of equation (3.6).

$$
\begin{aligned}
& u_{L}(\theta, y)=-u_{I}(\theta, \theta) e^{-\mu y} \cos \mu y \\
& V_{L}(\Theta, y)=U_{I}(e, 0) e^{-\mu y_{j}} s_{\text {gre }} e \sin \mu y \\
& \omega_{L}(\theta, y)=\frac{1}{\operatorname{cose}} \frac{\partial}{\partial \theta}\left[\frac{U_{1}(t, \theta) \sin \theta}{\sqrt{2} \mu} e^{-\mu y} \cos \theta \sin \left(\mu y+\frac{\pi}{4}\right)\right] \\
& T_{L}(\theta, y)=\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left[\frac{v_{x}(c, 0) \sin \theta}{i^{3} \mu^{3}} e^{-\mu y} \cos \theta \sin \left(\mu y-\frac{\pi}{4}\right)\right]
\end{aligned}
$$

where the boundary conditions have been used to get

$$
\begin{aligned}
& u_{I}(\theta, 0)+u_{L}(\theta, 0)=0 \\
& E V_{I}(\theta, 0)+V_{L}(\theta, c) \approx W_{L}(\theta, 0)=0 \\
& T_{I}(\theta, 0)+E \sim \Gamma T_{L}(\theta, 0) \sim T_{I}(\theta, 0)=1+H_{L}(\theta) \quad(3.11) \\
& \omega_{I}(\theta, 0)+\omega_{L}(\theta, c)=0
\end{aligned}
$$

03

$$
\omega_{I}(\theta, 0)=-\frac{1}{2 \cos a} \frac{\partial}{\partial \theta}\left(\frac{u_{I}(\theta, 0) \operatorname{sen} \cos \theta}{\mu}\right) \quad(3,12)
$$

Similarly, for the upper boundary layer corrections we have

$$
\begin{align*}
& u_{T}(\theta, x)=-e^{-\mu x} u_{I}(\theta, 1) \cos \mu x \\
& V_{T}(\theta, x)=e^{-\mu x} u_{I}(\theta, 1) \sin \theta \sin \mu x \\
& \omega_{T}(\theta, x)=-\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left[\frac{u_{I}(\theta, 1) \sin \theta}{\sqrt{7} \mu} e^{-\mu x} \cos \theta \sin \left(\mu x+\frac{\pi}{4}\right)\right] \\
& T_{T}(\theta, x)=-\frac{1}{\cos \theta} \frac{\partial}{\partial \theta}\left[\frac{u_{I}(0,1) s^{2} m \theta}{\mu^{3}} e^{-\mu x} \cos \theta \sin (\mu x-\pi / 4)\right] \\
& T_{I}(\theta, 1)=H_{T}(\theta) \\
& \omega_{I}(\theta, 1)=\frac{1}{2 \cos \theta} \frac{\partial}{\partial \theta}\left(\frac{u_{I}(\theta, 1) \operatorname{sqn} \theta \cos \theta}{\mu}\right) \tag{3.14}
\end{align*}
$$

The Interior Variables.
Equation (3.10) implies that $w_{i}$ is independent of $z$ therefore $\omega_{I}(\theta, 0)=\omega_{I}(\theta, 1)$. Hence $u_{i}(0,0)=-u_{I}(0,0)$. Assuming $\omega_{I}(\theta, c)=$ $\omega_{I}(\theta)=\omega_{I}(\theta, 1)$ to be known the system can be solved beginning with equation (3.9) working back to equation (3.6):

$$
T_{I}(\theta, z)=\sigma \beta E^{-1 / 2} \omega_{I}(\theta)\left(z^{2}-z\right)+H_{T}(\theta) z-H_{L}(\theta)(z-1)
$$

$$
\begin{aligned}
u_{I}(\theta, z)= & -\frac{1}{\sin \theta}\left[\sigma \beta E^{-1 / 2} \omega_{I}^{\prime}(\theta) \frac{\left(4 z^{3}-6 z^{2}+1\right)}{12}+H_{T}^{\prime}(\theta) \frac{\left(2 z^{2}-1\right)}{4}\right. \\
& -H_{L}^{\prime}(\theta)\left(\frac{\left.\left.\partial z^{2}-4 z+1\right)\right]}{4}\right. \\
V_{I}(\theta, z)= & \frac{1}{\sin ^{2} \theta}\left[\sigma \beta E^{-1 / 2} \omega_{I}^{\prime}(\theta)(\partial z-1)+H_{T}^{\prime}(\theta)-H_{L}^{\prime}(e)\right]
\end{aligned}
$$

(Note: $P(\theta, z)$ is not of interest except as needed to find the other dynamic variables and thus has been omitted.)

The system is completely solved when $U_{I}(\theta, c)$ is known (or, equivalently, $\omega_{I}(\theta)$ ). $\omega_{f}(\theta)$ has not yet been found but a second order differential equation for $\omega_{工}$ can be derived

$$
\omega_{I}(\theta)=\frac{\operatorname{sgn} \theta}{2 \cos \theta} \frac{d}{d \theta}\left\{\left[\frac{1}{1 /} \sigma \beta E^{-1 / 2} \frac{d}{d \theta} \omega_{I}(\theta)-\frac{1}{4}\left(H_{T}^{\prime}(\theta)+H_{L}^{\prime}(\theta)\right)\right] \frac{\cos \theta}{\mu \sin \theta}\right\} .
$$

This equation has not yet been solved.

## 4. CONCLUSIONS AND RECOMMENDATIONS.

(a) The technique of Pedlosky is a pertubation technique, however this is disguised in the analysis. The net result is that the order of approximation is different for each variable. It is recommended therefore that each variable be explicitly represented as a perturbation, egg., $\phi=\phi_{0}+E^{1 / 2} \phi_{1}+E \phi_{z}+\cdots$, and all solutions be found to the same order of approximation if possible or as nearly as is possible.
(b) Some of the assumptions of smallness made by Pedlosky are too stringent for AGCE. As many as possible should be relaxed, in particular, $\delta \ll E^{1 / 2} \quad$.
(c) The author derived the equations of motion for spherical geometry. The resulting equations differ non-trivially from those of Pedlosky. Unless these discrepancies can be adequately explained, all future work should be done using the author's equations.
(d) Some consideration should be given to the non-linear equations, possibly using the technique of Rahm and Walin [2].

1. Pedlosky, J., "Axially symmetric motion of a stratified rotating fluid in a spherical annulus of narrow gap', J. Fluid Mechanics, Vol. 36, part 2, 1969.
2. Rahm, Lars and Walin, Gosta, "On thermally forced stratified rotating fluids", J. Fluid Mechanics, Vol. 97, part 4, 1980.
3. Reid, William H., editor, Mathematical Problems in the Geophysical Sciences, American Mathematical Society (Providence, Rhode Island: 1971).
4. Fein, Jay S., Boundary Layers in Homogeneous and Stratified Rotating Fluids, The University Presses of Florida (Gainesville, Fla: 1978).

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM 

## A RANKING ALGORITHM FOR SPACELAB

## CREW AND EXPERIMENT SCHEDULING

Prepared by: Frank H. Mathis and Robert D. Grone

Refer to Section IX for report.

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA <br> PAYLOAD CREW TRAINING COMPLEX (PCTC) UTILIZATION AND TRAINING PLAN STATUS REPORT 

Prepared By:
Michael R. Self, Ed.D.
Academic Rank:
Assistant Research Professor
University and Department:University of Alabama in HuntsvilleDevelopmental Learning Program
NASA/MSFC:
(Laboratory)
(Division)
(Branch)
Systems Analysis and IntegrationOperations DevelopmentMan/Systems Integration
MSFC Counterpart:
H. H. Watters
Date:
August 1, 1980
Contract No.:NGT-01-002-099
by
Michael R. Self, Ed.D. Assistant Research Professor University of Alabama-Huntsville Huntsville, AL


#### Abstract

The first section describes the physical facilities that comprise the Payload Crew Training Complex (PCTC). The PCTC is located at MSFC, Huntsville, Alabama, and contains the Host Simulator; Experiment Simulators; Spacelab Aft Flight Deck, Experiment Pallet, and Experiment Rack mockups; the Simulation Director's Console; Payload Operations Control Center; classrooms; and supporting soft- and hardware.

The second section of the report establishes the parameters of a training philosophy for Payload Crew Training at the PCTC. Major parameters include: (1) The intent of training is not to be diagnostic, to select the "best" Payload Crew, but should be enhancing to the development of high levels of proficiency in every Payload Crew trainee; (2) The intelligence of each Payload Crew trainee is obviously to be respected. This implies that training at the PCTC, while highly procedural and sophisticated should tend to be less rigid than similar kinds of traditional high technology training; and (3) The function of trainee evaluation will be to determine how well training is being conducted. From the cooperative debriefings between the Payload Crew and PCTC trainers, should come ideas for improving the training sequence.

Section three discusses the development of the training plan. Included are discussions of preassessment, skill identification, procedures, material development, and evaluation options.


## INTRODUCTION

NASA finds itself on the threshold of a new era in space travel, namely repeated flights with a reuseable Space Transportation System (STS), i.e., the Space Shuttle. Additionally, they have diversified their payload configurations to include a larger portion of the scientific community not traditionally associated with NASA. For the first time too, the European Space Agency (ESA) has primary development responsibility for the Shuttle payload. The result is a new learning opportunity in the co-development of space missions and requires new thinking in training concepts.

One of the major distinctions between past NASA manned flights and this new era has been the changes in training philosophy and requirements. In the past everyone that has flown space missions has been a career astronaut with a scientific or technical background with flight training and experience. Additionally, all of the training tended to be centralized under NASA auspices. Yet in the presently scheduled flights, i.e., Spacelabs 1, 2, and 3, the Principle Investigators (PI) have primary responsibility for training the Payload Crew (PC) in how to conduct the PI's experiments. The PI's are research oriented, typically in a university setting, and literally distributed around the world.

These factors have precipitated the need for decentralized training. The effect of the move to decentralized training has been to: 1) reduce substantially the cost of training to NASA, since NASA does not have to reproduce all of the PI's equipment for training, and 2) add impetus to assuring overall flight competence prior to the actual flight via systems integration training. Concern with this second component is part of the issue addressed in this document. Another facet of the issue is how to best use the Payload Crew Training Complex (PCTC).

The PCTC was established to train the PC on those experiments that require interface with the onboard Spacelab computer, i.e., the Command and Data Management System (CDMS). In establishing the facility many questions arose as to how to achieve maximum utilization from the facility. Some of those questions are addressed herein.

## OBJECTIVES

The concern then has been to examine the parameters associated with decentralized training; which, as mentioned previously is a situation unique to the Space Shuttle/Spacelab missions, but a concept that appears to be one that NASA plans to pursue at least for the foreseeable future. The objectives then that derive from this goal are, at present, threefold:

1) To describe the PCTC and its training capabilities.
2) To define a training philosophy that might guide the PCTC trainers in developing a training plan.
3) To describe a procedure for the development and implementation of a training plan for training the PC at the PCTC.

What makes this report a status report is that training has not yet begun, so there are no evaluations of actual training procedures discussed. Furthermore, decisions on the best training approach are still in a deliberative state due to the fact that all of the experiments requiring CDMS interface, while known, have not been broken down into the step-by-step procedures level. This breakdown is basic to skill identification and the subsequent training procedures development.

## PCTC

The PCTC is located at MSFC, Huntsville, Alabama. The facility houses the Host Simulator (HS); Experiment Simulators (ES); the Spacelab Aft Flight Deck (AFD), Experiment Pallet, and Experiment Rack mockups; a Simulation Director's Console (SDC), a Payload Operations Control Center (POCC) console; classrooms; and supporting soft- and hardware.

The HS is the computer system that provides the capability to simulate the command, communications, control, and display functions of the Spacelab CDMS. This capability is necessary to execute experiment-unique simulations. It has the following features:

1) Two Data Display Systems (DDS) which can be simultaneously operated in support of one mission phase.
2) Reconfiguration and turnaround for the support of one training mission to another can be accomplished in eight hours.
3) Drives the SDC and POCC consoles for training and to approximate flight conditions.
4) Is modular in design and does not preclude expansion in size or computing speed.
5) Can simultaneously simulate multiple experiments, consistent with CDMS operating capabilities and provides standard analog-to-digital and digital-to-analog services for interfacing with simple experiment-unique control and display panels.
6) Allows for operation in a real time mode so that display time and command responses are commensurate with actual flight conditions.
7) Simulates symbolic pictures of the earth, sun, and sky scenes in a standard TV format for training in experiment pointing commands.
8) Protects against inadvertent damage to hardware and/or software by any operator or trainee.
9) Allows for environmental inputs to experiment models such as orbital state vector, vehicle attitude, and ephemeris of heavenly bodies.
10) Allows for detection and accumulation of characteristic data, such as PC errors, response times, and operating times.

The ES's include the hardware and software necessary to simulate experiments through the HS during training. A Data Display Unit (DDU) showing simulated data for experiment operations originates from either an instrument or Experiment Computer Applications Software (ECAS) simulation software package.

Experiment-unique command data is initiated through the DDS-Keyboard (KB). The CDMS is entered via $K B$ command, item, type, and/or function keys and acted upon, and thus generates the appropriate response within the simulation software. Each ES will send data and accept commands from certain key simulated control and display panels located in the experiment racks onboard the Spacelab mockup. Field of view imagery and alphanumeric data for display on payload-unique Cathode Ray Tubes (CRT's) are provided from the ES. Simulated POCC ground command inputs are available to the crew via the ES. The ES accepts and processes commands from the SDC during simulation. These commands control such simulation circumstances as simulation modes, control sequencing, fault initiation, and time cycling.

The SDC enables the Simulation Director and Assistant Simulation Director to monitor, control, and modify the simulation operation. It is composed of four racks containing two computer terminals, a hard copy machine, four large black and white Closed Circuit TV (CCTV) monitors, and a work surface which encompasses the CCTV camera controls. DDU monitors in rack 4 will display the same DDU displays as the trainee sees in the Spacelab module. Cameras for the TV monitors will be used thus: two inside the Spacelab module, one for the AFD and one for general surveillance of the high bay area. CCTV cameras are controllable in pitch and azimuth from hand controllers on the SDC work surface. Three video tape recorders may be used to record data from any one of several video sources: four CCTV cameras, four DDU's, and three Spacelab CCTV's used with scene generation. Hard copies of data seen on any of the CRT's are available by merely pressing a button on the hard copier after selecting the appropriate CRT.

The PCTC mockups include the Spacelab Module Interior, Shuttle AFD Interior, and associated equipment and pallets for part task activities with the capability for supporting payload crew training for any single Spacelab mission and the capability for reconfiguration to any payload.

## TRAINING PHILOSOPHY

A training philosophy is nothing more than an acknowledgement of the guiding ideals within which training will hopefully be implemented. The components of that philosophy may not actually be formally stated again, but they should be reflected in how the training is actually conducted. Philosophy statements describe thoughts or events that typically are extremely difficult to quantify for measurement, yet are important enough to be considered.

The training philosophy for the utilization of the PCTC hopefully incorporates the following parameters:

1) The intent of training is not to be diagnostic to select the "best" PC, but should facilitate the development of high levels of proficiency in every PC trainee.
2) The intelligence of each PC trainee is obviously to be respected. This implies that training at the PCTC, while highly procedural and sophisticated should tend to be less rigid than similar kinds of traditional high technology training.
3) The function of trainee evaluation will be to determine how well training is being conducted. From the cooperative debriefings between the PC and PCTC trainers should come ideas for improving the training sequence to help assure the highest quality of training and the maximum utilization of the PCTC's capabilities.

Specifically, those philosophical considerations include:

1) In order to proceed in a logical training sequence, critical skills and critical experiment sequences should be identified to better prepare the PC for the time when he actually encounters them.
2) There is a need for $P C$ assessment upon initiation of training at the PCTC. This assessment is necessary for the PCTC trainers to correctly determine how well they are conducting the training. Initial assessment should probably involve assessment of the PC's skills in utilizing the $D D U / K B$, understanding the Experiment Computer Operating System (ECOS), and communication skills.
3) While individual PC decision making ability is a major concern, it is not planned to "assess" those skills during PCTC training. Instead, the goal is to provide training experiences to facilitate those skills. For example, while PCTC training is highly procedural, in selecting the contingencies for use in training, some of the training scenarios are speculative in nature. They have been developed using the best information available, but in some cases the exact flow of the experiment will not be encountered until the actual flight. Therefore, a training goal should be for the PC to remain cognizant of the overall experiment goal(s). The value of this is to be able to periodically step back from the minutia of experiment procedures and conceptually think through what is supposed to be happening. This should enhance the likehood that the PC would make the best decision when, in flight, he actually encounters that option point in the experiment flow.
4) It follows from the above statement that a large segment of learning will not occur until the actual flight because of the sheer uniqueness of space flight; therefore, optimum operations proficiency may be approximated but never achieved. This becomes important in setting achievement criteria for training.
5) Training should proceed from the simple to the complex, both in skill acquisition and in experiment training. There is some thought about starting with the more complex because of their criticality and the need for ample training time. However, it is conversely believed that if the PC trainee is allowed to progress from simple to complex, and if an adequate amount of time is allowed for learning each experiment regardless of its complexity, that the PC trainee will be developing the necessary skills to more easily handle the complex experiment when he gets to it. For example, by taking the simple experiments first the trainee is getting familiar and comfortable with the $\mathrm{DDU} / \mathrm{KB}$, experiment layout, ECOS, various ECAS/DEP, and the language in which the procedures are written. Then he gathers experience in nominal operations, simple off-nominals, and simple contingencies. All the while he is developing his skill to recognize subtle differences on the DDU displays, learns to discriminate messages, and is enhancing his intuitive knowledge of all the systems. Then when he encounters the more complex experiment, his prior knowledge and experiment experience actually aid in reducing some of the complexity factors of the experiment.
6) There should be sufficient training opportunities in the training schedule to allow the PC to work in the various possible dyads that might actually be chosen to fly the various Spacelab missions. Again this is not done to gather information for selection purposes (since the Investigators Working Group (IWG) will actually make the choices), but rather to give the dyads themselves the opportunity to better learn each others style of operation and idiosyncrasies prior to the flight crew selection.
7) The focus of training at the PCTC is on CDMS controlled experiments. Non-CDMS experiments will only be addressed as they mesh into the timeline training, i.e., while the PC is conducting a CDMS supported experiment, his operations timeline might call for him to initiate a life sciences experiment, hence, he might walk over to the lifesciences rack and flip a dummy switch and then resume his CDMS procedures.

## TRAINING

The intent of this section is to discuss how the PCTC will actually be utilized during training. At present, however, all of the information necessary to make such a plan is incomplete. Therefore, this section represents an estimation of how that training might progress.

Traditionally, training is approached as the compilation of its sequential parts. That sequence is usually thought to be assessment, objectives, materials, procedures, and evaluation. That is generally how training has been approached in the present endeaver.

Assessment in this situation refers to pretraining assessment. Three questions have been explored here; 1) why have pretraining assessment, 2) what type assessment is mandated, and 3) who will do the assessment? The why was partially addressed in the previous philosophy section; specifically, the PCTC trainers need to determine how well they are conducting the training. In order to achieve that goal, it will be necessary to determine where the PC trainee is in his training when he arrives at the PCTC. This will allow the PCTC trainers to assess how the trainee has progressed during training and then devise ways to better conduct future training. The second aspect of the why question deals with determining entry level abilities at skills basic to CDMS operation. At present, five areas have been tentatively identified as being basic to all experiment operations, and as such will comprise the initial assessment/training period at the PCTC. The five areas are:

1. Keyboard operation; skill and location
2. CDMS entry; recognizing different DDU displays, functions, etc.
3. ECOS operations
4. Communications; transmitting acurately and quickly with POCC
5. Timeline operations; getting the total picture of when everything is supposed to run during the mission.

The type of assessment could be group or individual discussions or written exams to determine overall understanding of the various areas. This assessment could preceed, run concurrent with, or follow PC trainee demonstration of skill proficiency in those areas requiring manipulative skills. At present, our thinking is that a group of PCTC trainers could do the assessment, supported with statistical records supplied by the HS on key stroke rate, error, and ability to call up various pages and experiments.

The objectives of training would include the development of high operations proficiency and overall experiment competence. Operations proficiency would include the PC's ability to perform the five areas listed previously, their ability to follow experiment procedures in a timely fashion, and the identification of critical sequences in the experiment operations that require special attention, which is partially a function of the PC's decision making abilities.

Once proficiency in the five basic areas has been acquired, two objectives become primary. The first deals with the PC's successful mastery of experiment procedures. The second deals with developing proficiency in the skills necessary to carry out the first objective. These "critical" skills would be above and beyond those skills mastered during training in the five basic areas. At present,however, these skills have not been identified. This condition exists for two reasons. The
first is that the procedures for all the experiments have not been written, which precludes looking at them to identify critical sequences. The second reason is that skill identification has not been prioritized; that is, at present some feel there is questionable doubt as to whether such skills exist, and if they do, they would be developed when the procedures are learned and practiced during training.

At issue too is a philosophical question as to whether training, and hence mission operations, is simple or complex. Is it simple in that all the PC has to do is follow the written procedures which will caution for upcoming critical sequences. The notion here is that all the PC needs to do is develop very high proficiency at knowing and implementing the already spelled out procedures, and then relying on the POCC if questions arise. And while it may be true that the POCC may understand each experiment in more depth than the PC, the POCC has not necessarily developed the procedural knowledge to the degree that the PC has and may be of little substantive help to the PC in some unforseen situations.

The other view is that good scientific enquiry requires complex thinking skills. In this view the PC has the capability to conduct the experiment as he best sees fit, due to the fact that at that instant he possesses the most advanced conceptual understanding of the entire mission. Scientific enquiry demands sharp decision making skills, and as such these should be addressed during training at the PCTC. If this complex view is ascribed to, then training should afford direct or indirect opportunities to make decisions based on prior knowledge of and training in critical sequences or contingencies in each experiment.

Skill identification can start by asking and answering the following questions, pulled from Ely (1962):
"In the main, the definition of critical task characteristics is most effectively accomplished by considering the consequences of a particular activity if it were accomplished poorly or required excessive time for completion. Asking the following questions with respect to the various tasks is helpful for determining criticality.

Would below-minimum performance:
o lead to an accident?
o result directly in mission failure?
o be impossible to remedy within time constraints or not at all?
o be difficult to detect because of inadequate information feedback?
o recur over time in such a way as to produce a cumulative effect?
o contribute a large proportion of time to the total time required for some larger and critical function?

An affirmative reply to at least one of the above questions implies that the activity in question should be further considered in determining the critical behaviors for measurement."

In order to approach this issue, a skill matrix is being compiled from the available experiment procedures. Actually, two matrices are being developed. Both are modifications of existing skill charts (Rabideau, 1964; and Smode, Gruber, and Ely, 1962). Both are felt will yield useful information and will hopefully validate each other (the original charts and their modifications are in Appendix A). In the final evaluation, the skill matrices might very well show that training is in fact "simple" as described previously, but then at least there would be more support for that position.

Following skill identification, training procedures can be more fully articulated using the present training definitions. These definitions exist for each experiment and are described in three levels, from A to C. The descriptions progress from a general description (A), to typical off-nominals (B), and finally to step-by-step procedures (C). These skills should then be incorporated into the level A and B experiment descriptions. The level A, B, and C descriptions could then be utilized additionally by; 1) sending out the A and B descriptions as part of the PC's pre-arrival training materials, and 2) using all three levels during training to brief each experiment from and then to train from. " $A$ " descriptions should reflect time of training, the skills addressed, and where the training will take place. " B " descriptions should identify the training personnel, the possible off-nominals, and critical experiment stages; i.e., where contingencies exist or where the experiment could be aborted or drastically effected via operator error. " C " descriptions represent the training guide the PC might follow during training. " C " descriptions list nominal and off-nominal procedures.

After initiation to the experiment during the pretraining briefing and exposure to experiment procedures, but before going to the mockup for practice, the PC should be able to demonstrate the following:

1) What is the PC's knowledge of the experiment; communicated verbally?
2) What is the PC's knowledge of the possible off-nominals for this experiment and where will they occur?
3) What are the PC's perceptions of his anticipated responses to those off-nominals?
4) Can the $P C$ verbally walk through the experiment's nominal run?

Evaluation criteria follow from skill identification and should be fully planned for prior to the implementation of training. At this point, four recommendations are made for training:

1) Pertinent PCTC trainers should go through the PC training program prior to any PC trainee to debug and improve on the system.
2) If the modified Ely chart for skill identification is validated, then his three instrument and criteria measures in Appendix $B$ can be used to develop PCTC evaluation criteria and instruments.
3) Additionally, video tapes, PC record books, debriefing questionnaires, and data graphs of performance can be collected to trace training progress
4) Two additional types of measures have been located in our review that have application to the current evaluation needs. One is an effectiveness rating scale:

| Decision Speed |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $!$ | $\vdots$ | $\vdots$ | Slow |
| Extremely | Very | Acceptable |  | Very |
| Rapid | Fast |  |  | Slow |

Decision Quality


The second is an assessment tool of procedural operations:

## Task Step

Error \& Consequences

1. Turn power switch to ON position
2. Read external temperature
3. 
4. 

The final area addressed in this report is the matter of developing a training schedule. This can not be developed at present until the skill identification is completed. Additional parameters in formulating a schedule would be:

1) A need to determine the total time for experiment training, so that a training schedule can be developed.
2) How much time is needed for the dyads to work in each of the possible pairs.
3) Number of actual training locations within the PCTC for the $P C$ to be trained at.

## APPENDIX A

Modified skill identification charts used to identify PC skills for training definition.



TABLI: $5 \cdot 213$
Application of Task Data for Error, Frequcncy, and Achicvement Measures (Based on Smodc, Grubcr, and Ely, 1962)

Types of Tasks

| Pollow lRoutine, P're-established Checkout Sequence |
| :--- |
| Position a Set of Discrete/Scrial Controls | Position a Sct of Disercte/Serial Controls Oblain Informalion I'tum Dynamic Displays

Adjust Instrument or Other Jidulipment Item

- Track

名 Commumicale Information
Utilize Sitalus or Refrence Data
Directyonserve fixtermal livents
Compule
S.lerl Altion Vrom Among Alternative Choices

Fullow listablished procedures in Insjection,
Calibration, and/or Checkout.
Arepnire Information I'rom Among Multiple Sources - Makre Malliple Control Actions
 g Ilemedy, Hepair. lialicate, and/or Assemble 0 Complex linipment.

Carry Out Computational/Measurement
Sequences for Deterniuing Status
Make Decisions Based on Multiple Information Sources

TABLE 8-2A
Application of Task Data for Latency and Accuracy Mcasures
(Based on Smode, Gruber, and Ely, 1962)

Types of Tasks


| Locition a Sct of Disercte/Serial Conlrols |
| :--- |
| Ohtain luformation lirom Dymamic Display |

Aijust Instrument or Other Equipment Item
曷

TABLE 8-2C
Application of Task Data for Consumption, Physiological and Behavioral Judgment Measures (Based on Smode, Gruber, and Ely, 1962)


## REFERENCES

Cream, B. W., F. T. Eggemeier, and G. A. Klein. A strategy for the development of training devices. Human Factors, Vol. 20, April 1978, p. 145-158

Grodsky, M. A. The use of full scale mission simulation for the assessment of complex operator performance. Human Factors, Vol. 9, p. 341-348, 1968

Johnson, S. L. Retention and transfer of training on a procedural task interaction of training strategy and cognitive style. TLSP: Final Scientific Report, 1 October 1976 - 30 September 1977. 79N12704非

Meister, David. Human Factors in Operational System Testing: A Manual of Procedures. San Diego: Navy Personnel Research \& Development Center, 1978

Meister, D. and G. Rabideau. Human Factors Evaluation in System Development. New York: John Wiley \& Sons, Inc., 1965

Rabideau, G. Field Measurement of Human Performance in Man/Machine Systems. Human Factors, Vo1. 6, December 1964, p. 653-671

Smode, A. F., A. Gruber, and J. H. Ely. The Measurement of Advanced Flight Vehicle Crew Proficiency in Synthetic Ground Environments. Report MRL-TDR-62-2, Behavioral Science Laboratory, Aerospace Medical Division, Wright-Patterson AFB, February 1962

Tolcoh, M. A. Human Decision Making. Human Factors Engineering Seminar. Stanford, Conn.: Dunlap \& Associates, Inc., 1960

Wood, M. E. Improved crew member training through multi-media instruction. Proceedings from The Role of Man in Navigation, July 1-3, 1970, p. 477-485, 71A10524非

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA <br> SOME METHODS IN HIGH ENERGY COSMIC RAY MEASUREMENT 

Prepared By:
Academic Rank:
University and Department:
NASA/MSFC(Laboratory)(Division)(Branch)
MSFC Counterpart:
Date:
Contract No.:

Julian B. Shand, Jr., Ph.D.

Professor
Berry CollegeDepartment of Physics
Space Sciences
Space PhysicsAstrophysics
T. A. Parnell
August 1, 1980
NGT-01-002-099(University of Alabama)

# SOME METHODS IN HIGH ENERGY COSMIC RAY MEASUREMENT 

## By

Julian B. Shand, Jr., Ph.D. Professor of Physics Berry College Mt. Berry, Georgia

## ABSTRACT

The response of ion chamber detectors to cosmic rays of energy a few $\mathrm{GeV} / \mathrm{nucleon}$ and up shows a logarithmic rise with energy. The energy resolution has not been accurately defined. In Ref. 1 a calculation technique has been outlined which gives useful results for cosmic ray protons. The present study has applied the method to much heavier particles and higher energies.

In addition, a balloon flight will soon detect cosmic ray showers. A detector of acoustic waves produced by the showers will be carried, as will emulsion detectors. The present study has involved calibration of a flat plastic scintillation plate which will act as a trigger for the acoustic detector, and will determine the position of exit of the showers in order to expedite visual examination of the emulsions.
XV-1

## INTRODUCTION

Cosmic rays have been investigated since their discovery in 1911. Their flux in space is nearly independent of time and direction. They consist of nuclei, electrons, and electromagnetic radiation. The nuclei are mostly protons (around 90\%) and helium (around 9\%) but heavy nuclei are present also, being over represented as compared to solar abundances. The energy spectrum extends to the greatest energies known, reaching as much as $10^{20} \mathrm{eV}$ or 1 joule.

The origin of cosmic rays is a mystery, as is their acceleration to ultra high energy. Exploding galaxies, supernovae, pulsars, white dwarfs, and other exotic types of objects have been proposed as sources, but none have been clearly identified.

Measurement of the energy, knowledge of the energy resolution of the measuring devices, and identification of the various nuclear species are important for an understanding of the phenomena. A logarithmic rise in ionization of gas in ion chamber detectors, as cosmic ray energy increases, has long been observed, for energy a few $\mathrm{GeV} / \mathrm{nucleon}$ and up. The energy resolution has not been accurately defined. In Ref. I a procedure has been outlined for calculating resolution for the case of particles of charge 1 . This type information is needed also for the heavy nuclei in cosmic rays, and work on this problem is described in the following section.

Flat plates of plastic scintillator can serve several functions in a balloon-borne cosmic ray experiment. Rays passing through a plate produce light in the plate, which can be collected and measured by photomultiplier tubes. The signals from these tubes can indicate the passage of a "shower" of particles produced by collision of a high-energy cosmic ray with a nucleus. These signals can thus be used as a trigger for another detector designed specifically for measuring the energy of such showers. In addition, the relative strengths of the signals from the photomultiplier tubes can be used to determine the place of exit of the particles. This information facilitates examination of nuclear emulsions located next to the plastic scintillator, in order to find tracks in the emulsion created by the particles passing through.

## OBJECTIVES

Two main objectives were pursued. The first was to calculate the average energy actually deposited in an ion chamber by an ultra-high energy particle of large mass and charge. The method was to extend a calculational scheme already applied successfully to particles of charge 1.

The second objective was to calibrate a plate of plastic scintillator for measurement of the position of a cosmic ray shower passing through it. The method of calibration was to inject pulses of light at known positions on the plate and record the responses of photomultiplier tubes at the corner of the plate.

## On Relativistic Rise in Ionization

Relativistic charged particles passing through a gasfilled chamber lose energy mainly by excitation and ionization of the gas atoms (Ref. 4). Experiment shows a slow logarithmic increase in energy loss as particle energy increases. This is called the relativistic rise. The rise finally stops and energy loss becomes constant in what is called the Fermi plateau. Examination of the relativistic rise and knowledge of the energy resolution can allow one to determine the energy of the particle. This requires that charge be measured independently by other detectors through which the high energy particles pass. Typically a Cerenkov counter utilizing solid radiators such as lucite plastic is used for charge identification in the ultrarelativistic energy range. Thus for cosmic ray work the particular kind of nucleus striking the detector can be found (Ref. 5).

One way to look at the interaction between an incident particle and the electrons of an absorber is by representing the electromagnetic field of the particle by two terms (see Fano, Ref. 2). One, called the longitudinal term, is merely a static Coulomb interaction. The second term, called the transverse term, involves emission and reabsorption of virtual photons. Fano has treated the interaction in the lowest order (Born) relativistic approximation. The longitudinal interaction is effectively constant at high velocities. The transverse term, however, increases in strength as velocity increases. This is expected since the transverse field increases proportional, to $\gamma$, where

$$
\gamma=\left(1-\beta^{2}\right)^{-\frac{1}{2}} \text { and } \beta=v / c \text {. }
$$

This relativistic rise eventually stops because of polarization of the absorbing medium.

Fano found expressions for differential cross sections for ionization. By integrating over energy he found energy loss. This approach does not agree well with experiment for the case of thin gas absorbers (Ref. 1,3) because the energy loss occurs in a relatively few discrete collisions. Cobb et al. (Ref. 1) have used Fano's results up to a point but instead of integrating over energy have used Monte Carlo techniques and calculated energy losses in discrete collisions many times to get an average energy loss, a spectrum of energy loss, and the amount of relativistic rise. The results are in good agreement with experiment for incident particles of charge 1.

This author has adapted Cobb's method, which came from Fano, to the case of particles of large charge. Several difficulties were encountered:
A. Cobb's use of the letter $E$ in expressions was thought to be ambiguous: it could mean energy transfer to an electron, which would include the binding energy of the electron, as Fano seems to use it: or it could mean the kinetic energy of the electron after collision, as is suggested by Cobb's referring to the Rutherford cross-section as being

$$
\frac{2 \pi}{m v^{2}} \frac{1}{E^{2}}
$$

In this expression Rutherford did mean $E$ as kinetic energy of the electron.
B. Cobb's expression for the longitudinal differential cross-section for ionization included the factor (with units 1/energy)

$$
\delta\left(E-E_{i}\right)
$$

where $E_{i}$ is the electron binding energy. This has the form of the Dirac delta function, which has meaning when integrated. Since integration is not used in Cobb's method, it is not clear how to use it.
C. Cobb's expression for the dielectric constant $\varepsilon$ of the absorber gas contains the frequency $\omega$. This seems ambiguous: earlier in his paper the same variable had been used as the frequency of a virtual photon. However, Fano wrote that the $\omega$ in the dielectric constant expression is a root of the equation

$$
c^{2} q^{2}-\varepsilon(\omega)=0
$$

where $c$ is the speed of light and $q$ is momentum transfer to the electron. Since Cobb's treatment follows Fano, an uncertainty in interpretation is introduced.

Consultation with some experts on energy loss and a search of the literature have shed little light on these difficulties. Communication with the authors (Fano, Cobb et al.) will be attempted to resolve the ambiguities. This author has written computer code to calculate cross sections for ionization, using various combinations of interpretations, in order to be able to distinguish the correct ones. The work is still in progress.

In September 1980, a balloon flight will carry several cosmic ray experiments. Events of the highest energy will be of particular interest. Nuclear emulsions, X-ray film, and calorimetry will measure characteristics of these events. A plastic scintillator can be configured to determine the position of exit of cosmic rays. The rays cause the plastic to emit light. Most of the light is trapped in the thin plate of plastic by total internal reflection. When photomultiplier tubes are optically coupled to the corners of the plate, the tubes measure the amount of light reaching them. The amount varies inversely with the distance to the event, roughly. Several factors complicate the tube response, including edge effects, scratches, crazing of the plastic, inhomogeneity of the scintillating material, and probably others.

Preliminary calibration of the scintillator was done by this author. The method required a light-tight box. The scintillator plate, which measured $44^{\prime \prime}$ by $44^{\prime \prime}$ by $1 / 4^{\prime \prime}$, lay in the box on white paper which had a grid of horizontal and vertical lines drawn on it, with 10 cm spacing. A photomultiplier tube was attached with optical coupling grease to a specified position at one corner of the plastic plate. The photomultiplier tube was a $3^{\prime \prime}$, ten stage EMI, type 9708 KB . The output of the tube was amplified and referred to a multichannel analyzer which displayed pulse heights and allowed a determination of average pulse height.

Light pulses were obtained by sending electric pulses at frequency 10 Khz to a light-emitting diode (LED). The LED was mounted at the top of a shiny brass tube whose open end fit into a clear plastic disk. The disk, with LED, could be stuck to the plastic scintillator with optical grease and accurately located above an intersection of grid lines on the paper underneath. Thus positions of photomultiplier tube and LED were known to about $\frac{1}{2} \mathrm{~cm}$.

After pulses had been injected for 90 seconds and the average pulse height ascertained, the disk with the LED was removed and the plate wiped with dry cotton material and then with the same kind of material moistened with pure ethyl alcohol. All other procedures tried left some residue of film which scattered light out of the plate.

After the cleaning, the LED disk was regreased and repositioned. In this way a map was generated. Upon completion, the photomultiplier tube was moved to another corner and another map generated. Only one photomultiplier tube was ready at the time, so 3 other "dead" tubes were kept at the 3 other corners to absorb light as would happen in the final setup.

Checks were made of the reproducibility of the reading from a particular point. Deviations ranged from 0 to $12 \%$, averaging about 5\%.

A sample of experimental results is shown here. The positions refer to a Cartesian coordinate with the plastic scintillator plate defining the first quadrant, and the photomultiplier tube centered at $\mathrm{x}=5.9 \mathrm{~cm}, \mathrm{y}=5.9 \mathrm{~cm}$. Pulse heights along an edge and along an approximate diagonal are given. Positions are not integral multiples of 10 cm because visualization is easier than in the system actually used.

Position (cm)

| x | y | Pulse height | x | $y$ | Pulse height |
| ---: | :--- | ---: | ---: | ---: | ---: |
| 21.8 | 10 | 2740 | 21.8 | 20 | 1640 |
| 31.8 | 10 | 1260 | 31.8 | 30 | 720 |
| 41.8 | 10 | 740 | 41.8 | 40 | 410 |
| 51.8 | 10 | 450 | 61.8 | 50 | 280 |
| 61.8 | 10 | 300 | 71.8 | 60 | 190 |
| 71.8 | 10 | 240 | 81.8 | 70 | 180 |
| 81.8 | 10 | 180 | 91.8 | 90 | 160 |
| 91.8 | 10 | 130 | 101.8 | 100 | 130 |
| 101.8 | 10 | 110 |  |  |  |

At several positions in the scintillator pulse heights had local maxima or minima. Some of these could be identified with defects mentioned in the first paragraph under calibrating.

Concerning relativistic rise in ionization, an attempt should be made to locate the authors of Ref. 1 and correspond with them to resolve the difficulties of interpretation of their paper. In addition, the computer trial and error attempts to find the correct interpretation should be continued in case all else fails.

Calibration of the plastic scintillator as a determiner of position has been done. The experimental error indicates a positional error of nearly 5 cm can be expected. Since the calibration was done with light injected from outside, while the actual use will have light generated inside the plastic, a check of the numbers should be made with all 4 photomultiplier tubes in place, mounted in final form. Energetic radiation which will cause the plastic to scintillate should be injected from outside.

## REFERENCES

1. Cobb, J. H., W. W. M. Allison and J. N. Bunch, "The Ionization Loss of Relativistic Charged Particles in Thin Gas Samples and Its Use for Particle Identification," Nuclear Instruments and Methods 133, 1976.
2. U. Fano, "Penetration of Protons, $\alpha$ Particles, Mesons," Annual Review of Nuclear Science 13, 1963.
3. Crispin, A., and G. N. Fowler, "Density Effect in the Ionization Energy Loss of Fast Charged Particles in Matter," Reviews of Modern Physics, Vol. 12, No. 3, July 1970.
4. Jackson, J. D., Classical Electrodynamics, Wiley, New York, N. Y., 1962.
5. Gregory, J. C., and T. A. Parnell, "A Measurement of the Relativistic Rise in Xenon Filled Ionization Chambers for Cosmic Ray Iron," Papers of the Sixteenth International Cosmic Ray Conference, Vol. 12, 1979.

# nasa/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA 

## PHOTOCHEMISTRY OF MONODENTATE AND BIDENTATE

 CARBONATO COMPLEXES OF RHODIUM (III)Prepared By:Peter S. Sheridan
Academic Rank: Assistant Professor
University and Department: Colgate University Departmemt of Chemistry
NASA/MSFC:
(Laboratory)
Space Sciences Laboratory
(Division)
Materials Processing in Space
(Branch)
MSFC Counterpart:
Date:Robert J. Naumann
August 2, 1980
Contract No: ..... NGT-01-002-099

PHOTOCHEMISTRY OF RHODIUM (III)-CARBONATO COMPLEXES
by
Peter S. Sheridan Department of Chemistry Colgate University Hamilton, New York 13446

## ABSTRACT

A scheme for the photochemical fixation of water (eq 1) has been proposed.

$$
\begin{equation*}
2 \mathrm{H}_{2} \mathrm{O} \longrightarrow 2 \mathrm{H}_{2} \quad+\quad 0_{2} \tag{1}
\end{equation*}
$$

Fuel cells which generate electrical energy using the reverse of eq 1 are used in NASA spacecraft, so the solar-induced regeneration of $\mathrm{H}_{2}$ and $\mathrm{O}_{2}$ would allow the same materials to be recycled through the fuel cells, dramatically increasing energy output per pound of fuel carried aloft.

The proposed scheme involved a 5-step reaction sequence; the first step involves the 2 -electron reduction of a metal by a coordinated carbonate ligand, with corresponding oxidation of the carbonate to $\mathrm{CO}_{2}$ and $\mathrm{O}_{2}$.

Ligand field photolysis of trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} 0\right)\left(\mathrm{CO}_{3}\right)\right] \mathrm{ClO}_{4}$ (I), and $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{CO}_{3}\right)\right] \mathrm{ClO}_{4}$ (II) have been studied in the solid state and in aqueous solution at various pH values. Both salts are photoinert in the solid phase, but are quite photoreactive in aqueous solution. In solution, the monodentate ion (I) undergoes efficient isomerization to a mixture of cis and trans - $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} 0\right)\left(\mathrm{CO}_{3}\right)\right]{ }^{+}$, presumably with water exchange. A minor pH increase upon photolysis is evidence of inefficient carbonate $\left(\mathrm{CO}_{3}=\right)$ release, with formation of $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} 0\right)_{2}\right] \quad{ }^{3+}$.

In contrast, aqueous solutions of the bidentate carbonato complex, (II) undergo efficient pH decrease upon ligand field photolysis. Changes in the electronic spectrum ( $200-500 \mathrm{~nm}$ ) and pll changes indicate that the desired redox is occurring (eq 2)

$$
\begin{equation*}
\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{CO}_{3}\right)\right]^{+} \frac{\mathrm{h} v}{\mathrm{H}_{2} 0} \rightarrow\left[\mathrm{Rh}(\mathrm{en})_{2}\right]+\quad+\mathrm{CO}_{2}+\frac{1}{2} \mathrm{O}_{2} \tag{2}
\end{equation*}
$$

The pH increase is due to the aqueous behavior of $\mathrm{CO}_{2}$ (eq 3 ).

$$
\begin{equation*}
\mathrm{CO}_{2}+\mathrm{H}_{2} \mathrm{O} \longrightarrow \mathrm{H}^{+}+\mathrm{HCO}_{3}- \tag{3}
\end{equation*}
$$

A positive test for released carbonate ion supports this assignment. Work is proceeding to confirm this new photochemical reaction, with emphasis on gaining understanding of the exact nature of the rhodium photoproduct, the details of the photochemical mechanism, and to test the generality of photoredox reactions of bidentate carbonato complexes.

## ACKNOWLEDGEMENTS

I wish to thank Dr. Robert Naumann for allowing a stimulating summer of research, without memos or committees; it was great. Mr. Charles Carter and Mr. Marion Kent were instrumental in facilitating my application, and I thank them for their help and support. Ms. Margie Phillips typed this manuscript, for which I am most grateful. Finally, I owe a debt to Dr. Robert Snyder who found room in his Separations Processes Lab for an inorganic photochemist.

## NOMENCLATURE

$$
\begin{aligned}
& \text { LMCT - ligand-to-metal-charge-transfer } \\
& \text { - ethylenediamine }\left(\mathrm{H}_{2} \mathrm{NCH}_{2} \mathrm{CH}_{2} \mathrm{NH}_{2}\right) \\
& \text { K } \quad \text { - equilibrium constant ( } \mathrm{K}=4.3 \times 10^{-7} \text { for hydrolysis of } \mathrm{CO}_{2} \text { ) } \\
& \mathrm{CO}_{2}+\mathrm{H}_{2} \mathrm{O} \rightleftharpoons \mathrm{H}^{+}+\mathrm{HCO}_{3}{ }^{-} \\
& \text {hט - implies the reaction conditions included the incidence of } \\
& \text { light } \\
& \text { nm - nanometer }
\end{aligned}
$$

## LIST OF FIGURES

| Figure No | Title | Page |
| :---: | :---: | :---: |
| 1. | Electronic Absorption Spectra of trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$, | XVI-10 |
|  | $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+}$and the Corning 7-59 Colored G1ass Filter |  |
| 2. | Spectral changes induced by photolysis of trans-[Rh(en) $\left.2\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$ | XVI-12 |
| 3. | Spectral changes induced by photolysis of $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+}$ | XVI-16 |
| 4. | Acidity increase in photolyzed solutions of $\left\lceil\mathrm{R} h(\mathrm{en})_{2} \mathrm{CO}_{3}\right\rceil^{+}$as a function of photolysis time | XVI-19 |
|  | LIST OF TABLES |  |
| Table Number | Title | Page |
| 1. | pH changes induced by photolysis of trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$ | XVI-13 |
| 2. | pH changes induced by photolysis of $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+}$. | XVI-13 |

## INTROAHCTION

The photochemical behavior of Rh(III) complexes has attracted considerable interest ${ }^{1}$, as a variety of photoinduced aquations and stereochemical rearrangements have been observed. While exceptions exist, Rh(III) amine complexes characteristically undergo efficient photoinduced ligand loss/aquation, with rearrangement to a symmetric trans configuration. There is little evidence of any photoinduced reduction of the Rh(III) center, even upon photolysis of LMCT bands in such complexes as $\left[\mathrm{Rh}\left(\mathrm{NH}_{3}\right)_{5} \mathrm{I}\right]{ }^{2+2}$; the instability of the Rh(II) oxidation state and the photoreactivity of the ligand field excited states lead to a predominantly ligand field excitedstate chemistry for Rh(III) amines.

Our work on the photochemistry of rhodium complexes has led to a reaction sequence which, on paper, represents the photochemical cleavage of water into $\mathrm{H}_{2}$ and $\mathrm{O}_{2}$. It is a five step reaction sequence (eq 4-8); the net reaction is the desired water fixation.


The first step is the photoinduced oxidation of a $d^{6}$ metal center (Rh(III) here) by a carbonate ligand, generating $0_{2}, \mathrm{CO}_{2}$ and a coordinatively unsaturated $\mathrm{d}^{8}$ complex (eq 4). Oxidative-addition of $\mathrm{H}_{3} \mathrm{O}^{+}$then regenerates the Rh (III) center, creating a hydride ligand (eq 5). Photoinduced (or thermal) labilization of this pseudohalide ligand (eq6) would lead to instantaneous $\mathrm{H}_{2}$ generation (eq 7). Under the proper pH conditions, reaction of the photogenerated $\mathrm{CO}_{2}$ (eq 1) and the aquohydroxo complex (eq 8) would regenerate the carbonato starting material (eq 8 ).

This reaction sequence has many advantages over other currently studied water-photolysis schemes:
a) Each of the steps is either a well-known reaction (eq 7 and 8 ) or is a reaction for which closely analogous reactions have been studied (eq 4,5 , and 6).
b) No radical intermediates are involved, as the 2-electron transfer of the $d^{6}-d^{8}$ transition metal system allows complementarity.
c) The scheme is a self-contained, catalytic cycle, with only water consumed and $\mathrm{H}_{2}$ and $\mathrm{O}_{2}$ generated.
d) The scheme is quite flexible, as many variables can be adjusted.

A major difficulty with this scheme is the first step, as there is little precedence for such a photoinduced two-electron transfer. An additional problem is that reports on the photochemistry of carbonato complexes are surprisingly rare.

An apparent two-electron transfer was observed upon irradiation of alcohol solutions of some carbonato complexes of $\mathrm{Pt}(\mathrm{II})^{3}$. In the presence of nucleophiles ( $\mathrm{X}=\mathrm{C}_{2} \mathrm{H}_{4}, \mathrm{P} \phi_{3}, \mathrm{C} 0, \phi \mathrm{C} \equiv \mathrm{C} \phi$ ) photolysis of $\mathrm{PtL}_{2}\left(\mathrm{CO}_{3}\right)$ ( $\mathrm{L}=\mathrm{P} \phi_{3}$, As $\phi_{3}$ or $\mathrm{P}\left(\mathrm{CH}_{3}\right) \phi_{2}$ ) generates $\mathrm{PtL}_{2} \mathrm{X}_{\mathrm{n}}(\mathrm{N}=2$ for $\mathrm{X}=\mathrm{CO} ; \mathrm{n}=1$ for bulkier ligands), consistent with the in situ formation of $\mathrm{PtL}_{2}$. Carbon dioxide was released during the photolysis (one mole per mole of Pt ) and alcohols added to the photolysis solutions could be oxidized to aldehydes and ketones, confirming that the carbonato ligand was the two-electron reducing agent.

Such apparent two-electron transfers have also been observed upon photolysis of some oxalato complexes. Some $\mathrm{PtL}_{2}(\mathrm{ox})$ complexes, similar to the carbonato complexes just discussed, photolyze to give $\mathrm{CO}_{2}$ and similar $\mathrm{Pt}(0)$ photoproducts. 4 Irradiation of $\mathrm{Pd}\left(\mathrm{NH}_{3}\right)_{2}(\mathrm{ox})$ was reported ${ }^{5}$ to generate $\mathrm{CO}_{2}$ and a reactive species identified as $\mathrm{Pd}\left(\mathrm{NH}_{3}\right)_{2}$, which decomposed to give Pd black. Poulenc ${ }^{6}$ noted that the yellow $\mathrm{Rh}(\mathrm{py})_{3}$ (ox)Cl solid turned red in sunlight, and Gillard et. al ${ }^{7}$. have shown the reaction involves another photoinduced two-electron transfer

$$
\mathrm{Rh}(\mathrm{py})_{3}(\mathrm{ox}) \mathrm{Cl} \longrightarrow \quad " \mathrm{Rh}(\mathrm{py})_{3} \mathrm{Cl} "+2 \mathrm{CO}_{2}
$$

Work at the Space Sciences Laboratory this summer focused on this first step, the photolysis of Rh (III) carbonato complexes. Two complexes were chosed for study - trans - $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} 0\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$and $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{CO}_{3}\right)\right]^{+}{ }^{-}$ as they allowed a study of the affect of monodentate vs. bidentate coordination of the carbonate ion on the resulting photochemistry.

## OBJECTIVES

Research was directed toward an understanding of the following questions:

1. What is the photochemical behavior of Rh (III)-carbonato complexes?
2. Do the charge-transfer and ligand field states of these Rh (III) complexes display distinct photochemical reactivity, or does intersystem crossing lead to a common photoreactive state?
3. What affect does monodentate vs. bidentate coordination of the carbonato ligand have on the resulting photochemistry?
4. Are such complexes photosensitive in the solid state, or is an aqueous solution necessary for photoreactivity? In solution, does the pH affect the photochemical reaction.

The objective of the work was to determine if photoinduced two-electron reduction of a Rh(III) center could be induced by photolysis of the appropriate Rh (III)-carbonato complex.

## A. Preparation of Complexes

trans-Aquochlorobis(ethylenediamine)rhodium(III) Perchlorate
The trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]{ }^{+}$ion was prepared by the method of Rerek ${ }^{8}$. A sample of trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}(\mathrm{Cl})_{2}\right] \mathrm{NO}_{3} \cdot \mathrm{H}_{2} \mathrm{O}^{9}$ was anated with a five fold excess of $\mathrm{NaN}_{3}$ (under reflux) and the desired trans- [Rh(en) $\left.{ }_{2}\left(\mathrm{~N}_{3}\right)\right]^{\text {t }}$ was crystallized as the $\mathrm{BF}_{4}^{-}$salt upon addition of $\mathrm{NaBF}_{4}$. This diazido ion was dissolved in dilute $\mathrm{HClO}_{4}$, and $\mathrm{NaNO}_{2}$ was added; this generated $\mathrm{NO}^{+}$, destroying the coordinated azide.

$$
\mathrm{N}_{3}{ }^{-}+\mathrm{NO}^{+} \longrightarrow \mathrm{N}_{2} \mathrm{O}^{(\mathrm{g})}+\mathrm{N}_{2}(\mathrm{~g})
$$

At a neutral pH (adjusted with LiOH) the resulting trans- $\left[\mathrm{Rh}(\mathrm{en})_{2} \quad(\mathrm{OH})\left(\mathrm{H}_{2} \mathrm{O}\right)\right]^{2+}$ was precipitated as the perchlorate salt. A solution of this ion was acidified at pH 1 with $\mathrm{HClO}_{4}$, and $\mathrm{NaHCO}_{3}$ slowly added until a pH of 5.80 was reached. The perchlorate salt was precipitated by cooling in ice, followed by dropwise addition of ethanol. The absorption spectrum had $\lambda$ max at $347 \mathrm{~nm}(\varepsilon=105)$, consistent with published data.

## Carbonatobis(ethylenediamine)rhodium(III) Perchlorate

( $\left.\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right] \mathrm{C1O}_{4}$ was prepared by the method of Harris et. al. ${ }^{10}$ It displayed a single peak in the electronic spectrum at $327 \mathrm{~nm}(\varepsilon=250)$; this absorption coefficient is less than the literature value, but any impurities were shown not to interfere in the photochemistry. There were no detectable changes in the electronic absorption spectrum of this ion in solutions between pH 1 and pH 13 , clearly ruling out significant contamination of either of the monodentate ions, cis or trans- $\left(\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right) \mathrm{CO}_{3}\right)^{+}$.

Barium chloride solution
Aqueous solutions of barium chloride were prepared by dissolving $\mathrm{BaCO}_{3}(s)$ in $1 \mathrm{MHCl}(\mathrm{aq})$ until effervescence ceased, and a pH of $7(+1)$ was reached. The resulting solution was filtered and kept in a sealed flask to avoid $\mathrm{CO}_{2}$ contamination. It was boiled for ca. 20 minutes before use.

Other reagents were used as supplied by the manufacturers.

## B. Photolyses

Photolyses were performed on either solid samples of aqueous solutions of the complexes. Solid samples were finely ground with mortar and pestle, and then stuck on a piece of green tape ( $1 \mathrm{~cm} \times 3 \mathrm{~cm}$ ). This tape was suspended in a 1 cm fused silica cell (Beckman), connected to a microvolumeter (vide infra).

For the solution studies, a sample of the desired sample was dissolved in the appropriate aqueous solution at room temperature, and placed in a 1 cm fused
silica cell (Beckman). Electronic spectra were recorded with the sample in the same cell.

## C. Equipment

The irradiation source was a 1000 watt Mercury-Xenon Lamp (Hanovia Manufacturing Corporation) with compatible power supply and starter (Schoeffel LPS 255 power supply and Model 359 igniter). Photolyses were performed at 28.5 amps at 22 V . Use of the compatible Schoeffel high-intensity monochromator did not allow sufficient intensity to pass to the sample, so wavelength selection was affected with a Corning colored glass filter (C-7-59). This filter effectively absorbed all radiation below 280 nm , but allowed irradiation of the ligand field bands. The absorption spectra of the C-7-59 filter along with the spectra of the two complexes studied is given in Fig. 1.

Electronic spectra were recorded on a Beckman Acta CII spectrophotometer with an external 10 inch recorder (Beckman Model 1005). The pH measurements were done with a PHM 64 Research pH Meter (Radiometer/Copenhagen) with a GK 2321C Combined Electrode. Samples averaged about 3.0 ml for pH measurements.

A microvolumeter was prepared after a design of Stevenson ${ }^{10}$. A 1 ml graduated pipet (graduated in 0.01 ml units) was used, and the evolution of a gas was detected by monitoring the flow of water bubbles down the pipet. Photolysis of solid samples of $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{~N}_{3}\right)_{2}\right] \mathrm{BF}_{4}$, which releases $\mathrm{N}_{2}$, was used to check the sensitivity of the microvolumeter.

A. Trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$

## 1. Results

In the solid state, trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{CO}_{3}\right)\left(\mathrm{H}_{2} \mathrm{O}\right)\right] \mathrm{C1O} 4$ is photoinert. Prolonged photolyses ( 2 hrs .) with both filtered and unfiltered light did not cause any detectable release of gas. When the photolyzed solid was dissolved in water, it had an electronic spectrum indistinguishable from that of an unphotolyzed sample.

In aqueous solution, and in the absence of light, $\left(\mathrm{Rh}(\mathrm{en}){ }_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$ is relatively unreactive. At room temperature and at a solution pH between 6 and 10, there is a minor pH increase and a minor increase in the intensity of the ligand field band. This reaction is presumably the thermal aquation of the carbonate ion, but it was slow enough to have no measurable impact on the photochemical reactions.

Upon photolysis, the electronic spectrum undergoes pronounced changes (Fig. 2). Note that photolysis causes:
a) an increase in intensity of the ligand field band,
b) a blue shift of the peak maximum from 347 nm to about 332 nm ,
c) virtually no shift in the position of the low-energy edge of the charge-transfer band.

There appears to be an isosbestic point near 242 nm , although the spectra are nearly parallel at that point, so the presence of an isobestic is difficult to determine.

In addition to the spectral changes observed in Fig 2., photolysis causes the release of a base into solution. Data from a typical experiment are given in Table 1, where the pH changes observed in an unphotolyzed blank and photolyzed sample are listed. (The pH changes in Table 1 were recorded from the same experiment monitored in Fig 2.).

After extended photolysis, only minor spectral changes could be photoinduced. To determine the isomeric composition of this solution, 2 ml of conc HC1 was added to a 3 ml sample of photoproduct, and refluxed for 1 hour. The resulting solution had an electronic spectrum of a mixture of cis and trans $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{Cl}_{2}\right]^{+}$(ca. $63 \%$ cis and $37 \%$ trans).

The specific experiment described in Fig 2 and Table 1 was done with unfiltered light, but the same results were obtained using filtered light. The reaction was approximately 5 times more rapid using the more intense unfiltered irradiation, but the nature of the spectral and pH changes were independent of excitation wavelength.


TABLE 1

| total <br> photolysis time <br> (seconds) | sample <br> pH | blank <br> pH |
| :---: | :---: | :---: |
| Initial | 7.032 | 7.060 |
| 100 | 7.110 | 7.070 |
| 200 | 7.184 | 7.081 |
| 300 | 7.245 | 7.096 |
| 450 | 7.325 | 7.129 |
| 600 | 7.376 | 7.136 |
| 800 | 7.435 | 7.151 |
| 1000 | 7.470 | 7.184 |
| 1300 | 7.504 | 7.216 |
| 2000 | 7.540 | 7.220 |

Typical pH changes induced by Photolysis of Trans- $\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)^{+}$

TABLE 2

| total photolysis time (seconds) | $\begin{aligned} & \text { sample } \\ & \text { pH } \end{aligned}$ | blank pH |
| :---: | :---: | :---: |
| Initial | 9.905 | 9.916 |
| 100 | 9.411 | 9.862 |
| 200 | 8.557 | 9.789 |
| 300 | 7.729 | 9.721 |
| 400 | 7.375 | 9.616 |
| 550 | 7.120 | 9.508 |
| 705 | 7.002 | 9.380 |
| 900 | 6.911 | 9.277 |
| 1300 | 6.691 | 9.185 |
| 2200 | 6.495 | 9.086 |

## 2. Discussion

The available results force the conclusion that photolysis of trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{CO}_{3}\right)\left(\mathrm{H}_{2} \mathrm{O}\right)\right]+$ causes isomerization (eq ${ }^{9}$ ).
trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]+\quad$ cis- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$(9) The cis- $\left.\mathbb{R h}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$ion has an absorption maximum at 332 nm , at a significantly higher absorption coefficient than the trans isomer, consistent with the blue shift and increased absorption observed upon photolysis. The lack of change in the charge-transfer edge is also significant, as the intense charge-transfer peak is a LMCT band with carbonate ligand as the donor; the LMCT bands for the analogous $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}_{2}\right]^{+}\right.$ions begin at significantly higher energies. As photolysis causes no shift in that edge, it is clear that significant amounts of carbonate ion are not labilized during the photolysis.

The pH changes are not consistent with this proposed reaction, however, as the reaction of eq 9 should not lead to any pH changes. Calculation of the amount of $\mathrm{OH}^{-}$released after extensive photolysis shows that only a very minor amount of the Rh complex is releasing base to the solution. For example, during the experiment described in Fig 2 and Table 1, 1000 seconds of photolysis causes spectral changes indicating that about $80 \%$ of the starting material has undergone photoinduced isomerization. After those same 1000 sec. , only about $.0026 \%$ of the starting material has released an $\mathrm{OH}^{-}$into solution. We conclude that any pH changes are due to a minor side reaction which is not related to the dominant photoisomerization. The pH increase is apparent as soon as photolysis begins, suggesting that whatever reaction causes the pH increase, it is a primary photochemical process. The rate of pH increase remains relatively constant throughout the photolysis, even when little starting material remains to be photolyzed; this implies that there is also an inefficient base-releasing secondary photochemical reaction. Whether these pH changes are due to inefficient release of carbonate ion or an ethylenedimine chelate is not known at this time.

Further support for the proposed photoisomerization comes from the anation studies. Refluxing the photoproduct in HCl leads to the acid-catalyzed aquation of the carbonate ion, which would be followed by chloride anation of the diaquo ions. Since hydrolysis of the carbonate ion occurs without metaloxygen bond cleavage, and anations at Rh(III) centers are also known to be stereoretentive, the composition of the resulting mixture of cis and trans$\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{Cl}_{2}\right]+$ would reflect the stereochemical composition of the cis and trans- $\left[\mathrm{kh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$photoproduct mixture.

The best value for this composition - $63 \%$ cis and $37 \%$ trans-must be considered a preliminary result; further work will be done to determine whether photolysis of trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$generates a photostationary state of that composition. Such a result, which available information supports, would imply that the dominate photochemical reaction of cis- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$ is isomerization back to the trans_isomer.

The reactions described above occurred upon irradiation with filtered (ligand field) and unfiltered (ligand field and charge-transfer) light. This suggests a single photoreactive state; since no evidence of photoinduced redox behavior was observed, a ligand field photoreactive state is indicated.
(Presumably, photoinduced water exchange accompanies isomerization, although the necessary isotope exchange studies have not been done to verify this assumption). There is no evidence of $\mathrm{CO}_{2}$ release, which would suggest a photoredox process, even upon excitation of the LMCT bands. The rate of photoreaction was markedly increased when the LMCT bands were irradiated, implying that intersystem crossing, from the initially populated charge-transfer states to the photoreactive ligand field state, is efficient.

In summary, the photochemical behavior of the monodentate carbonate complexes is reminiscent of the behavior of analogous haloamine complexes of Rh(III). The LMCT states are photoinert, while the ligand field states are photoreactive, leading to ligand labilization and aquation, with concomitant stereochemical rearrangement about the Rh (III) center.


1. Results

In the solid state, $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{CO}_{3}\right)\right] \mathrm{ClO}_{4}$ is photoinert. Prolonged photolysis ( 2 hrs ) with both filtered and unfiltered light did not cause any detectable release of gas, and when a sample so photolyzed was dissolved in water, it had an absorption spectrum indistinguishable from that of an unphotolyzed sample.

In aqueous solution, and inthe absence of light, $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{CO}_{3}\right)\right]^{+}$ is very inert. At room temperature and at a solution pH between 6 and 10 , there was no detectable change in the absorption spectrum for over 48 hours. No pH changes, beyond those observed in the blank due to absorption of atmospheric $\mathrm{CO}_{2}$, were observed in aqueous solutions of $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{CO}_{3}\right)\right]^{+}$.

Upon photolysis, the electronic spectrum undergoes pronounced changes (Fig. 3). Note that photolysis causes:
a) a decrease in intensity of the ligand field band, with virtually no change in wavelength maximum
b) a minor shift in the low-energy edge of the charge-transfer band to higher energy.

Jsosbestic points are not observed in the spectrum during the initial stages of the photolysis, but an isosbestic point does develop at about 314 nm after the initial stages of the photolysis are complete.

In addition to the spectral changes observed in Fig 3, photolysis causes a marked decrease in the solution pH . Data from a typical experiment are given in Table 2. The pH changes recorded in Table 2 were observed from the same experiment monitored in Fig 3.

The presence of photoproduced carbonate ion in solution was checked by photolyzing a sample of $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{CO}_{3}\right)\right]^{+}$in a solution containing $\mathrm{Ba}^{2}+$

ions; formation of $\mathrm{BaCO}_{3}$ (s) would signal the presence of $\mathrm{CO}_{3}$ in solution. stream of $\mathrm{N}_{2}\left(\mathrm{~g}\right.$ ) was bubbled through a slightly acidic solution of $\left[\mathrm{Rh}(\mathrm{en})_{2}\right.$ $\left.\mathrm{CO}_{3}\right]+$ for 30 min . This solution was made basic ( pH 10 ) with $\mathrm{LiOH}(\mathrm{s})$, and bubbling was continued for another 15 min . This solution volume was doubled with the addition of a freshly boiled $\mathrm{BaCl}_{2}$ solution ( 0.5 M ), with $\mathrm{N}_{2}$ bubbling continued throughout these manipulations. Any $\mathrm{BaCO}_{3}(\mathrm{~s})$ which formed was allowed to precipitate and a clear solution was put into a closed, fused silica cell for photolysis. The formation of a precipitate was monitored visually, or by recording the absorption spectra at 550 nm , where only scattered light could lead to absorption.) Photolysis with unfiltered light rapidly led to precipitate formation (presumably $\mathrm{BaCO}_{3}$ (s) ), which did not occur in a blank sample held in the dark. Photolysis with the filtered light did not lead to immediate precipitate formation, but prolonged photolysis did lead to a measurable cloudiness. We attribute this to the less efficient formation of $\mathrm{BaCO}_{3}(\mathrm{~s})$.

In an effort to determine the stereochemical composition of the photoproduct, a sample of $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+}$was photolyzed with unfiltered radiation for 3200 sec. , and the photoproduct solution was then diluted with an equal volume of conc HC1. The solution immediately became yellow, and displayed a broad absorption band with local maxima at about 355 nm and 325 nm . Refluxing this solution for about 1 hour led to a yellow solution without a peak at 355 nm , but a more intense, clean peak centered at 331 nm ; this peak had an absorption coefficient of about $3 \times 10^{3} \mathrm{M}^{-1} \mathrm{~cm}^{-1}$.

## 2. Discussion

The photochemical reactivity of $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+}$has not yet been clearly elucidated; additional work is necessary to fully explain the phenomena observed. The following hypothesis, therefore must still be considered tentative.

Direct labilization of carbonate or ethylenediamine ligands can be ruled out as primary photochemical processes, as such reactions would lead to pH increases, not the observed pH decrease. Photoinduced isomerization, as observed for the monodentate complex (vide supra), may be occurring, but cannot account for the dramatic pH decrease; in addition, the photoinduced spectral changes do not support the idea of an isomerization to some other Rh(III) tetramine complex. The observed increase in acidity, along with the detection of free $\mathrm{C}_{3}=$ in solution, strongly implies that $\mathrm{CO}_{2}$ is a primary photoproduct. The $\mathrm{H}^{+}$(aq) and $\mathrm{CO}_{3}{ }^{=}$(aq) result from the rapid hydrolysis of the acidic oxide.

$$
\mathrm{CO}_{2(\mathrm{aq})}+\mathrm{H}_{2} \mathrm{O} \rightleftharpoons \mathrm{H}^{+}+\mathrm{HCO}_{3}^{-} \rightleftharpoons 2 \mathrm{H}^{+}+\mathrm{CO}_{3}{ }^{-}
$$

Independent detection of $\mathrm{H}^{+}$and $\mathrm{CO}_{3}{ }^{=}$in the photolyzed solutions supports the supposition of photogeneration of $\mathrm{CO}_{2}$.

Two possible routes for $\mathrm{CO}_{2}$ release seem likely:
a) Decarboxylation of coordinated $\mathrm{CO}_{3}=$ to generate $\mathrm{CO}_{2}$ and $\mathrm{OH}^{-}$
b) oxidation of coordinated $\mathrm{CO}_{3}=$ to give $\mathrm{CO}_{2}$ and $\frac{1}{2} \mathrm{O}_{2}$.

In case a), observed for the thermal reaction of the analogous Co(III) complex, the following reaction sequence would be anticipated:

$$
\begin{aligned}
{\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+}+\mathrm{H}^{+} } & \left.\longrightarrow \mid \mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3} \mathrm{H}\right]^{2+} \\
\mathrm{H}_{2} \mathrm{O}+\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3} \mathrm{H}\right]^{2+} \longrightarrow & \left.\left.\longrightarrow \mid \mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right) \mathrm{CO}_{3} \mathrm{H}\right)\right]^{2+} \\
{\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right) \mathrm{CO} \mathrm{O}_{3} \mathrm{H}\right]^{2+} } & \left.\longrightarrow \mid \mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)(\mathrm{OH})\right]^{2+}+\mathrm{CO}_{2}
\end{aligned}
$$

$\overline{\text { Net }}$| reaction |
| :--- |
| $\mathrm{H}_{3} \mathrm{O}^{+}+\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+} \longrightarrow\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)(\mathrm{OH})\right]^{2+}+\mathrm{CO}_{2}$ |

This process does release the weak acid, $\mathrm{CO}_{2}$, but consumes the strong acid $\mathrm{H}_{3} \mathrm{O}+$. Thus, the observed pH decrease would not be anticipated if such a process were occurring. In addition, reaction of the photoproduct of such a sequence with HCl would 1 ead to cis or trans- $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{Cl}_{2}\right]^{+}$; the observed product, with a single, intense peak at 331 nm is neither of the $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{Cl}_{2}\right]^{+}$ isomers.

Possibility b), the two electron oxidation of the carbonate ligand, is more consistent with the available information (and is the more fascinating photochemical reaction). Assuming concomitant two-electron reduction of the Rh(III) center, a reaction of the form (eq 10)

$$
\begin{equation*}
\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+} \longrightarrow\left[\mathrm{Rh}(\mathrm{en})_{2}\right]^{+}+\mathrm{CO}_{2}+\frac{1}{2} \mathrm{O}_{2} \tag{10}
\end{equation*}
$$

would be occurring.
Reaction 10 is consistent with a variety of observations. Plots of "moles of $\mathrm{H}^{+}$released" as a function of photolysis time showed two distinct patterns, depending upon the initial pH (Fig. 4). At pH 6 , the rapid initial generation of $\mathrm{H}^{+}$becomes less pronounced as the photolysis proceeds; at pH 10 , the initial photolysis period generates only minor amounts of $\mathrm{H}^{+}$, followed by a more rapid rate of $\mathrm{H}^{+}$generation. This is consistent with the release of a weak acid, as at pH 10 the $\left[\mathrm{HCO}_{3}{ }^{-}\right] /\left[\mathrm{CO}_{2}\right]$ ratio would be four orders of magnitude larger than at pH 6.

$$
\begin{align*}
& \mathrm{H}_{2} \mathrm{O}+\mathrm{CO}_{2} \rightleftharpoons \mathrm{HCO}_{3}^{-}+\mathrm{H}^{+}  \tag{11}\\
& \mathrm{K}=\frac{\left.\left[\mathrm{H}^{+}\right][\mathrm{HCO}]_{3}^{-}\right]}{\left[\mathrm{CO}_{2}\right]}  \tag{12}\\
& \frac{\mathrm{K}}{\left[\mathrm{H}^{+}\right]}=\frac{\left[\mathrm{HCO}_{3}^{-}\right]}{\left[\mathrm{CO}_{2}\right]} \tag{13}
\end{align*}
$$

Such a large concentration of bicarbonate ion would act as a buffer, resisting the generation of $\mathrm{H}^{+}$ions upon the addition of $\mathrm{C}_{2}$ to the solution. As shown in Fig. 4, $\mathrm{H}^{+}$increase was initially quite slow, until the buffering action of the bicarbonate ion was overcome. At pll 6, the bicarbonate ion concentration would be less formidable, and little buffcring action would be apparent. If we consider that for reaction 11 ,

$$
\left[\mathrm{H}^{+}\right]=\left[\mathrm{HCO}_{3}^{-}\right] \text {and }\left[\mathrm{CO}_{2}\right]=\left[\mathrm{CO}_{2}\right] \text { initial }-\left[\mathrm{H}^{+}\right]
$$

we can write

$$
K=\frac{\left|\mathrm{H}^{+}\right|^{2}}{\left[\mathrm{CO}_{2}\right]}=\frac{\mathrm{H}^{+2}}{\left[\left.\mathrm{CO}_{2}\right|_{\text {initia1 }}\left[\mathrm{H}^{+}\right]\right.}
$$



Solving for $\mathrm{H}^{+}$, one obtains:

$$
\begin{equation*}
\left.\left[H^{+}\right]=\left\{K\left(\mathrm{CO}_{2}\right] \text { initial }\right)^{\prime}+\frac{1}{4} K^{2}\right)^{\frac{1}{2}}-\frac{1}{2} K \tag{14}
\end{equation*}
$$

Evaluating for $\mathrm{H}^{+}$at increasing levels of $\left[\mathrm{CO}_{2}\right]$ initial (representing the $\mathrm{CO}_{2}$ gnerated by photolysis) gives a curve which closely models the pH 6 curve of Fig. 4.

After 2200 seconds of unfiltered irradiation of $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+}$, the spectrum shows that a significant amount of the starting material has photochemically reacted. The data from Table 2 can be used to show that the $\mathrm{H}^{+}$released at that time is equivalent to about $10 \%$ of the total Rh (III) concentration. This, again, is consistent with the photogeneration of a weak acid $\left(\mathrm{CO}_{2}\right)$. A more quantitative assessment is not yet possible, as the exact nature of the Rh -containing photoproduct is not known.

The nature of the Rh-containing photoproduct is a major mystery remaining after this summer's work. The observed spectral changes do not suggest the formation of any obvious Rh(III) tetraamine species, and refluxing the photoproduct in HC1 leads to an intensely yellow species which has, as yet, escaped identification. The spectral changes are consistent with the formation of the same product observed upon $\mathrm{BH}_{4}^{-}$reduction of cis- $\left[\mathrm{Rh}(\mathrm{en}){ }_{2} \mathrm{Cl}_{2}\right]^{+}$, suggestive of a two-electron reduction of the Rh(III) center, but further work would be necessary to corroborate this supposition.

In summary, work this summer has led us to the tantalizing hypothesis that ligand field and charge-transfer photolysis of a bidentate carbonato complex of Rh(III) leads to the two-electron oxidation of the carbonate ligand, with complementary reduction of the Rh(III) center. This result, if borne out by further studies, would represent a new type of photochemical reaction, and would encourage further work on the water-photolysis scheme (eq 4-8).

## 1. Conclusions

a) The monodentate carbonato complex, trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}(\mathrm{H} 20)\left(\mathrm{CO}_{3}\right)\right]^{+}$ efficiently undergoes photoisomerization:


There is no evidence of electron-transfer upon either ligand field or chargetransfer excitation. A very inefficient base-releasing photochemical reaction also occurs, but its nature has not been characterized. Preliminary evidence implies that the cis-aquocarbonato photoproduct undergoes photoinduced isomerization back to trans- $\left[\mathrm{Rh}(\mathrm{en})_{2}\left(\mathrm{H}_{2} \mathrm{O}\right)\left(\mathrm{CO}_{3}\right)\right]^{+}$.
b) Ligand field and charge-transfer photolysis of the bidentate carbonato complex, $\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}{ }^{+}$causes the release of $\mathrm{C}_{2}$ into solution by a reaction tentatively described as:

$$
\left.\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}\right]^{+} \longrightarrow \mathrm{h} \psi \longrightarrow \mathrm{Rh}(\mathrm{en})_{2}\right]^{+}+\mathrm{CO}_{2}+\frac{1}{2} \mathrm{O}_{2}
$$

If subsequent studies corroborate the occurrence of a two-electron transfer, it will represent a new class of photochemical reaction. Such a reaction is the crucial first step in a water-photolysis scheme (eq 4-8), so confirmation of an efficient two-electron transfer, induced by a single photon, will open a new area of transition metal photochemistry.

## 2. Recommendations

This summer's study has shown that a promising area of transition metal photochemistry can be found in the study of bidentate carbonate complexes. The nature of the Rh -containing photoproduct in the photolysis of $\left[\mathrm{Rh}(\mathrm{en})_{2} \mathrm{CO}_{3}{ }^{+}\right.$ must receive immediate attention, and the design of future experiments depends on the results of that study. Available evidence strongly suggests that a twoelectron redox process is involved, so the next steps must include:
a) a study to determine the factors which maximize the efficiency (quantum yield) of this redox reaction.
b) Studies must be initiated to increase the amount of the solar spectrum absorbed by the photoreactive complex. A logical approach would be to look at some substituted bipyridyl ligands or other aromatic ligands with extensive absorption inthe visible and near-IR portions of the spectrum. The use of intensely colored sensitizers should also be studied.
c) Parallel studies should explore the detailed chemistry of the watersoluble complexes of $\mathrm{Rh}(\mathrm{I})$. The nature of the four-coordinate $\mathrm{d}^{8}$ complexes, and the proposed $d^{6}$ hydrides must be determined. The thermal and photochemical properties of a variety of such hydride complexes must be studied to determine the conditions which would lead to a photoreactive hydride complex.

In summary, the discovery of a two-electron redox reaction, induced by a single photon, suggests that it may be possible to photochemically generate $\mathrm{H}_{2}$ from water without the complications inherent in the hydrogen atom radical. This should encourage further study of a carbonate-based system for the photochemical cleavage of water.

## REFERENCES

1. a) Adamson, A. W.; Fleischauer, P.S., Concepts in Inorganic Photochemistry (eds.), 1975, Wiley-Interscience, N.Y.: (Chapter 5, by Ford, P.C.; Hintze, R.E.; Petersen, J.D.)
b) Endicott, J.F.; Wong, C-L,; Inoue, T.; Natarajan, P., Inorg. Chem., 1979, 18,450
c) Ford, P.C.; Skibsted, L.H., Strauss, D., Inorg. Chem., 1979, 18, 3171.
d) Clark, S.; Petersen, J.D., Inorg. Chem., 1979, 18, 3394.
e) Martins, E.; Sheridan, P.S.; Inorg. Chem., 1978; 17, 2822, 3631.
2. Kelly, T. C.; Endicott, J.F., J. Phys. Chem., 1972, 76, 1937.
3. Blake, D. M.; Mersecchi, R., J. Chem. Soc.; D., (Chem. Comm.), 1971, 1045.
4. Blake, D. M.; Nyman, C. J., J. Amer. Chem. Soc., 1970, 92, 5359.
5. Mann, F. G.; Crowfoot, D. C.; Gattiker, D. C.; Wooster, N., J. Chem. Soc., 1935, 1642.
6. Poulenc, P., Ann. Chim. (France), 1935, 4, 634.
7. Addison, W. W.; Gillard, R.D.; Sheridan, P.S.; Tipping, L.R.H.; J. Chem. Soc., Da1ton, 1974, 709.
8. Johnson, S.A.; Basolo, Fi, Inorg. Chem., 1962, 1, 925.
9. Rerek, Mark, Master's Thesis, State University of New York at Binghamton, August, 1980.
10. Palmer, D.A.; vanEldik, R.; Kelm, H.; Harris, G.M., Inorg. Chem. 1980, 19, 1009.

1980

# NASA/ASEE SUMMER FACULTY RESEARCH FELLONSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA <br> CAICULATIONS AND MEASUREMENTS IN SUPPORT OF THE TRIGLYCINE SULFATE ZERO GRAVITY GROWTH EXPERIMENT 

| Prepared By: | John M. Springer, Ph.D. |
| :--- | :--- |
| Academic Rank: | Research Associate |
| University and Department: | Fisk University <br> Department of Physics |
| NASA/MSFC: <br> (Iaboratory) <br> (Division) <br> (Branch) | Space Sciences |
| MSFC Counterpart: | Materials Processing <br> Solid State |
| Date: | Roger Kroes, Ph.D. |
| Contract No.: | August 1, 1980 |

CAICULATIUNS AND MEASUREMENTS IN SUPPORT OF THE TRIGLYCINE SULFATE ZERO GROWTH GRAVITY EXPERIMENT

## By

John M. Springer, Ph.D. Research Associate in Physics Fisk University Nashville, Tennessee

## ABSTRACT

The solution growth of crystals of triglycine sulfate (TGS) in a zero gravity environment is a materials processing experiment planned for an early Spacelab mission. This experiment has the twin objectives of both studying the growth process by optical means and also returning the grown crystals to earth for characterization.

This paper describes two areas of investigation which are required for maximum utilization of the flight experiment results. The first area studied is the use of an ultramicroscope to characterize structural defects in the grown TGS crystals. Mie thoory has been applied to the light scattering expected from idealized crystal defects in order to predict the sensitivity of ultramicroscopy in the detection of small defects.

The second area of investigation has involved the measurement of optical properties of TGS solutions in water. This data is needed in order to interpret interferometric studies of mass transport through the TGS solution during the growth phase of the filight experiment. In par= ticular, refeactive index data has been obtained for a wide range of temperatures and concentrations of TGS solutions. Absorbance measurements in the ultraviolet have also been made for the same samples.
Figure No. Title Page
1 Schematic diagram of an ultramicroscope used to observe defect scattering ..... XVII-7
2 Automated ultramicroscope ..... XVII-8
3 Orientation of the intensity functions for right angle scattering ..... XVII-10
4 Refractive index vs Temperature for Solutions of TGS in Water ..... XVII-16
5 Refractive vs Concentration for Solutions of TGS in Water ..... XVII-17
6 Concentration vs Wavelength at 50\% Trans- mittance for Triglycine Sulfate in Water ..... XVII-18

## IIST OF TABLES

Table No. Title Page
1 Power scattered at $90^{\circ}$ from sphericaldefects for plane polarized incident beamof wavelength 400nm and intensity$1.6 \times 10^{7} \mathrm{~W} / \mathrm{m}^{2}$. XVII-132 Refractive index data for TriglycineSulfate Solution.XVII-15

One of the materials processing experiments that will be flown aboard an early Spacelab mission is the solution growth of triglycine sulfate (TGS). This experiment has two objectives, one being the observation of the crystal growth process itself by optical techniques such as Schlieren photography and interferometry, the other being the return of the zeromgravity grown crystals to earth for testing and comparison with their earth-grown counterparts.

This report describes two studies performed which are related to the TGS growth experiment. The first part discusses the use of a light scattering technique called ultramicroscopy to characterize triglycine sulfate crystals. The theory of Mie for the scattering of light by spheres is used to indicate the sensitivity of the ultramicroscope in the detection of volume defects. Although based on an idealization of the actual defects that would be present in a real crystal, this analysis is useful in analyzing the effectiveness of this technique.

The second part of the report presents data on the optical properties of solutions of TGS in water. This information on the refractive index and optical absorbance of TGS in solution is required for the analysis of the Schlieren and interferometric data which will be obtained during the crystal growth experiment on Spacelab 3.

## I. Crystal Characterization by Ultramicroscopy

A. Ceneral

Crystal characterization involves describing a crystal's structure and composition in such a way that those of its properties of relevance to the study at hand are well enough known to predict in principle the behavior of the crystal. One of the important features of a crystal are its departures from the perfect array of atoms by which its structure is commonly described. These departures from ideality, called defects, exist in every real crystal and range from the simple absence of a single atom from the lattice to dislocations involving the entire crystal.

The scattering of electromagnetic radiation incident on a crystal is one tool for detecting defects. The scattering and diffraction of X-rays yields information on structural features whose size is on the order of the lattice spacing. In the case of optically transparent crystals, the scattering of visible light can also reveal structural defects in a crystal, though not with the detail of X-ray topography. For a crystal whose optical properties are germane to its application, however, the measurement of visible light scattering is essential for its proper characterization.

One method being implemented to study TGS crystals by light scattering is that of the ultramicroscope. In this technique, illustrated in Figure 1, a beam of light focussed to a small diameter is used to illuminate the crystal. As the light passes through the crystal, it is scattered by inhomogeneities in the crystal's refractive index, such as are caused by certain types of defects. In order to observe the light scattered from the defects, a microscope is used to collect the light scattered at right angles to the beam. Small, isolated defects show up as points of light in the dark surrounding field of view. Regions of a crystal where the defect density is quite high will show nebulous or diffuse scattering regions. As described by Vand, 1 a small He-Ne laser makes an excellent source for the ultramicroscope. This method has already been used to study the relation between defects in TGS and its dielectric properties. ${ }^{2}$

A diagram of an ultramicroscope being constructed for the study of TGS is shown in Figure 2. This ultramicroscope will use photoelectric rather than photographic detection of scattering centers, and will be computer driven so a three dimensional map of the light scattering defect may be constructed.* In the remainder of this section of the report, an analysis is described in which the expected sensitivity of this ultramicroscope to defects is calculated. This sensitivity will be seen to depend on defect size and relative effective refractive index. * This is being constructed at Fisk University with the support of

NASA Grant \#NSG-8060.


FIGURE 1: Schematic diagram of ultramicroscope used to observe defect scattering.


FIGURE 2: Automated ultramicroscope

## B. Discussion

The model "defect" which has been used in this study is that of a homogeneous sphere of diameter $D$ and refractive index $n$. This spherical defect is assumed to be surrounded by the undisturbed TGS lattice. The absorption coefficient of the defect is assumed to be zero, so the refractive index is a real number. It is illuminated in the ultramicroscope by a plane electromagnetic wave of wavelength 632.8 nm in air.

The theory of scattering of electromagnetic waves by spheres is a classical problem, and was solved analytically by Mie. ${ }^{3}$ His solutions involve infinite sums of complex functions, but several sets of tables are available in the literature which make the evaluation $\mathrm{pf}^{\prime}$ the solutions for particular cases straightforward, if somewhat tedious. 4,5,6 Computer programs for this purpose have also been written, and would be useful if a large number of cases must be calculated.

Figure 3 defipes the geometry of the problem. According to the treatment by Boll, the scattered light intensity is represented by two intensity functions $i_{1}$ and $i_{2}$, given by

$$
\begin{equation*}
i_{1}=\frac{\lambda^{2}}{8 \pi^{2}} \left\lvert\, \sum_{n=1}^{\infty}\left\{A_{n} \frac{d P_{n}(x)}{d x}+B_{n}\left[x \frac{d P_{n}(x)}{d x}-\left(1-x^{2}\right) \frac{d^{2} P_{n}(x)}{d x^{2}}\right]\right\}^{2}\right. \tag{1}
\end{equation*}
$$

and

$$
\begin{equation*}
i_{2}=\frac{\lambda^{2}}{8 \pi^{2}}\left|\sum_{n=1}^{\infty}\left\{A_{n}\left[x \frac{d P_{n}(x)}{d x}-\left(1-x^{2}\right) \frac{d^{2} P_{n}(x)}{d x^{2}}\right]+B_{n} \frac{d P_{n}(x)}{d x}\right\}\right|^{2} \tag{2}
\end{equation*}
$$

where $P_{n}(x)$ is a Legendre polynomial of order $n$, with $x=-\cos (\theta)$. $\theta$ is the scattering angle from the incident direction. The coefficients $A_{n}$ and $B_{n}$ are complex Ricatti-Bessel and Hankel functions, and for the values to be considered here can be obtained from tables. 4 These coefficients depend on only two physically relevant parameters for their evaluation, a size parameter.

$$
\begin{equation*}
\mathcal{L}=\frac{\pi D}{\lambda} \tag{3}
\end{equation*}
$$

where $D=$ the scattering particle diameter
$\lambda=$ the wavelength of the incident light in the surrounding medium
and $m$, the refractive index of the scattering particle relative to the surrounding medium.


FIGURE 3: Orientation of the intensity functions for right
angle scattering.

For plane-polarized incident light, the intensity function $i_{1}$ is the component of scattered light polarized perpendicular to the electric vector of the incident beam, and $i_{2}$ represents the component of scattered light polarized parallel to the plane formed by the electric vector of the incident beam and its direction of propagation. The resulting scattered intensity per unit incident intensity and unit solid angle is for any direction $\theta, \phi$ given by

$$
\begin{equation*}
I=\frac{\lambda^{2}}{4 \pi^{2}}\left[\left(\cos ^{2} \phi\right) i_{1}+\left(\sin ^{2} \phi\right) i_{2}\right] \tag{4}
\end{equation*}
$$

In the ultramicroscope, the scattered light collected by the objective lens is nominally at a scattering angle $\theta=90^{\circ}$. If the incident light on the crystal is polarized with its electric vector perpendicular to the plane formed by the incident and scattered light directions of propagation, then $\varnothing=90^{\circ}$ also, so the total scattered intensity per unit solid angle is

$$
\begin{equation*}
I_{\Omega}=I_{0} \frac{\lambda^{2} i_{2}}{8 \pi^{2}} \tag{5}
\end{equation*}
$$

where $I_{0}$ is the incident intensity $\left(W / m^{2}\right)$.
A typical objective used on the microscope will have a numerical apperature of 0.25 . Taking into account the refraction that occurs at the surface of the TGS crystal as the scattered light passes into the air, an objective lens of N. A. 0.25 will collect a cone of light with a half-angle of $9^{\circ}$ in the crystal, corresponding to a solid angle of 0.00612 . So for this case, the scattered power entering the objective lens will be

$$
\begin{equation*}
P_{S}=I_{0} \frac{\lambda^{2}}{8 \pi^{2}}\left(6.12 \times 10^{-3}\right) i_{2} \tag{6}
\end{equation*}
$$

In the ultramicroscope the detector is a photodiode with an active area which is larger than the image of the smaller defects with which this analysis is concerned, so the scattered light power collected by the objective lens will all fall on the detector when the defect is centered in its field of view. $P_{S}$, therefore, is the relevant value which must be calculated and compared with the sensitivity of the detector and its amplifier.

## C. Calculations

The relevant physical parameters of a defect are its diameter D and refractive index $n$. TGS in solution at its saturation point at $25^{\circ}$ has a refractive index of approximately 1.38. The crystalline solid has three principal indices which range from 1.484 to $1.584 .{ }^{8}$ Values of the relative refractive index $m$ could be expected to vary typically from 0.6 for a void with absolute index 1 to 0.93 for a volume defect consisting of uncrystallized solution. The actual values of $m$ could of course be close to 1 for small distortions of the lattice. Values of the size parameter which are available in Boll's? tables range upward from 1 to 120. Since only small defects are of interest in finding the limit of detectable scattering, values of from l-10 were used in the calculations. In triglycine sulfate, this corresponds to a diameter of from approximately $1.2 \times 10-7$ meters to $1.2 \times 10^{-6}$ meters when the wavelength of the incident light is 632.8 nm .

The laser which will be used is a $5 \mathrm{~mW} \mathrm{He}-\mathrm{Ne}$ laser, which can be focussed to a $20 \times 10^{-6}$ mater diameter in the crystal. This gives an average beam intensity of $I_{o}=1.6 \times 10^{7} \mathrm{~W} / \mathrm{m}^{2}$. Using a refractive index of 1.58 for the crystalline TGS, the wavelength of the laser light in the crystal is $\lambda=400 \mathrm{~nm}$. These values give a scattered power from equation 6 of

$$
\begin{equation*}
P_{s}=I_{0} \frac{\lambda^{2}\left(6.12 \times 10^{-3}\right)}{8 \pi^{2}} i_{2}=\left(1.98 \times 10^{-10} 0_{\text {Watts }}\right) i_{2} \tag{7}
\end{equation*}
$$

Representative values of the scattered power collected by the • ultramicroscope for several values of and $m$ are shown in Table 1. The effective scattering dofect size $D$ is also given.

An evaluation of these values in terms of photoelectric detection can be made by comparison with the noise figure for a photodiode which is being evaluated in the ultramicroscope. The photodiode, a model PN-4OA from United Detector Technology, has a noise equivalent power (NEP) of $10^{-14}$ Watts, For spherical defects with a relative refractive index near unity (0.93), defects of this type would have to exceed 300 nm to be above the noise level of the detector. For voids, where the refractive index would be around 0.6 , a defect somewhat small ( 100 nm ) would scatter sufficiently to be observed.

TABLE 1

Power scattered at $90^{\circ}$ from spherical defects for plane polarized incident beam of wavelength 400 nm and intensity $1.6 \times 10^{7} \mathrm{~W} / \mathrm{m}^{2}$. Power is in Watts.

| $\alpha$ | Defect <br> Diameter <br> ( nm ) | Relative Refractive Index m |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 0.6 | 0.7 | 0.75 | 0.8 | 0.9 | 0.93 |
| 1 | 127 | $6.2 \times 10^{-14}$ | $2.9 \times 10^{-15}$ | $1.5 \times 10^{-15}$ | $6.6 \times 10^{-16}$ | $4.7 \times 10^{-17}$ | $1.2 \times 10^{-17}$ |
| 2 | 250 | $2.7 \times 10^{-12}$ | $1.2 \times 10^{-12}$ | $7.5 \times 10^{-13}$ | $3.2 \times 10^{-13}$ | $2.8 \times 10^{-14}$ | $7.5 \times 10^{-15}$ |
| 3 | 381 |  |  |  |  | $4.1 \times 10^{-14}$ | $1.1 \times 10^{-13}$ |
| 4 | 508 |  |  |  |  | $4.9 \times 10^{-14}$ | $1.3 \times 10^{-13}$ |
| 5 | 635 |  |  |  |  | $8.9 \times 10^{-14}$ | $3.2 \times 10^{-13}$ |
| 10 | 127 |  |  |  |  | $2.0 \times 10^{-12}$ | $1.6 \times 10^{-12}$ |

## II. Optical Measurements

Refractive index and ultraviolet absorbance measurements were made on solutions of triglycine sulfate in distilled water at concentrations ranging from 0.25 grams-TGS/100 grams $-\mathrm{H}_{2} \mathrm{O}$ to 32 grams-TGS $/ 100 \mathrm{grams}-\mathrm{H}_{2} \mathrm{O}$. Samples of various concentrations were obtained by making an initial solution of $32 \mathrm{~g} / 100 \mathrm{~g}$, dividing this in half and diluting one part to make a $16 \mathrm{~g} / 100 \mathrm{~g}$ solution, and then repeating this process to form a series of samples each one-half as concentrated as its predecessor. The resulting solutions were stored in sealed bottles. The samples were weighed during their preparation to a precision of $\pm 0.01$ grams.

Refractive index measurements were made using a Billingham and Stanley Model 60/ED precision Abbe refractometer. The sample and prism temperature was controlled by water circulated from a water bath in which the temperature was regulated to $\pm 0.01$ degree centigrade. The actual refractometer head temperature versus bath temperature was determined by calibration to compensate for heat loss in the connecting tubing. Temperature measurement was made using a Cole-Parmer Model 8502-25 thermistor thermometer, and a Yellow Springs Inc. Model 703 thermistor probe. A sodium lamp was used to illuminate the refractom meter, so the values obtained are for the sodium $D_{1}$ line, 589.6 nm . Values calculated from the critical angle measurements using the refractometer are given in Table 2. Graphs of $n_{d}$ vs temperature for various concentrations and $n_{d}$ vs concentration for different temperatures are shown in Figures 4 and 5 .

Transmission curves in the range of 200-300 nm were measured for the same samples used in the refractive index study. The curves were recorded on a Model 701 Heath/Schlumberger singlembeam spectrophotometer. All measurements were made for a 10 mm path length. This data will be used in experimentally determining mass diffusion coefficients for TGS in water. For this purpose, a graph of the wavelengths at which the present transmission of the solutions reaches $50 \%$ is useful. This graph is shown in Figure 6.

TABIE 2: Refractive Index Data for Triglycine Sulfate Solution - $\left(n_{d}\right)$

| Concentration$\mathrm{g} / 100 \mathrm{~g} \mathrm{H}_{2} \mathrm{O}$ | Solution Temperature - Degree Celsius |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 24.98 | 29.66 | 34.36 | 39.28 | 43.95 | 48.75 | 53.45 | 58.23 |
| 0.25 | 1.3329 | 1.33236 | 1.33170 | 1.33100 | 1.33018 | 1.32938 |  |  |
| 0.50 | 1.3331 | 1.33252 | 1.33191 | 1.33114 | 1.33052 | 1.32952 | 1.32858 | 1.32762 |
| 2.00 | 1.33381 | 1.33326 | 1.33259 | 1.33184 | 2.33108 | 1.33020 | 1.32929 |  |
| 2.00 | 1.3354 | 1.33476 | 2.33421 | 2.33337 | 1.33253 | 1.33176 | 2.33085 | 1.32989 |
| 4.00 | 1.3385 | 1.33791 | 1.33721 | 1.33643 | 1.33560 | 1.33471 | 1.33380 | 1.33292 |
| 8.00 | 1.34408 | 1.34343 | 1.34274 | 1.34190 | 1.34101 | 1.34009 | 1.33918 | 1.33819 |
| 26.00 | 1.35430 | 1.35352 | 1.35270 | 1.35188 | 1.35011 | 1.35011 | 1.34917 | 1.34815 |
| 32.00 | 1.3711 | 1.37038 | 1.36948 | 1.36879 | 1.36779 | 1.36681 | 1.36590 | 1.36494 |

XVII-15

FIGURE 4: Refractive index vs temperature for solutions of TGS in water. Each curve is marked with the concentration of the solution in grams of TGS per 100 grams water.


FIGURE 5: Refractive index vs concentration for solutions of TGS in water.



## REFERENCES

1. Vand, V. and K. Vedan, "The Laser as Light Source for Ultramicroscopy and Light Scattering by Imperfections in Crystals" J. Appl. Phys. 37, 2551-2557, 1960.
2. Mohler, R. and W. White, "Characterization of Volume Defects in Triglycine Sulfate by laser Light Scattering", J. Crystal Growth 28, 240-248, 1975.
3. Mie, G., Ann. Physik, 25, 377(1908).
4. Deirmendjian, D., Electromagnetic scattering on Spherical Polydispersions, American Llsevier Publishing Co., New York, N. Y., 1969.
5. Che, C., Clark, G. C., and S.W. Churchill, Tables of Angular Distribution Coefficients for Light-Scattering By Spheres, University of Michigan Press, Ann Arbor, Michigan, 1957.
6. Denman, H. H., et. al, Angular Scattering Functions for Spheres, Wayne State University Press, Detroit, Michigan, 1966.

## 7. Boll, R. H., Leacock, J. A., Clark, G. C., and S. W. Churchill, Tables of Light-Scattering Functions, The University of Michigan Press, Ann Arbor, Michigan, 1950.

8. Dion, I. M., et. al., Acta Crystallographica, 12, 259-264, 1959.

## NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

 MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMAFLUID DYNAMIC ANALYSIS OF THE SPACE SHUTTLE MAIN ENGINE HIGH PRESSURE OXIDIZER TURBOPUMP SLINGER SEAL
Prepared By:
Academic Rank:Assistant ProfessorUniversity and Department
NASA/MSFC:
(Laboratory) (Division) (Branch)Structures and PropulsionPropulsionTurbomachinery and Combustion Devices
MSFC Counterpart:
Loren Gross
Date: August 1, 1980
Contract No.:NGT-01-002-099(University of Alabama)

# FLUID DYNAMIC ANALYSIS OF THE SPACE SHUTTLE MAIN ENGINE HIGH PRESSURE OXIDIZER TURBOPUMP <br> SLINGER SEAL <br> By <br> Captain James S. Voss <br> Assistant Professor of Mechanics <br> United States Military Academy <br> West Point, New York 

## ABSTRACT

An analytical study has been conducted to clarify the details of the flow on the bladed side of a centrifugal type dynamic shaft seal utilized to contain liquid oxygen in the Space Shuttle Main Engine high pressure oxidizer turbopump.

The governing equations are solved to predict the pressure and temperature gradients and to aid in investigating the nature of the liquid-vapor interface.

Recommendations for design improvements are discussed.

## LIST OF FIGURES

Figure No. Title ..... Page
1 Cutaway view of slinger seal model ..... XVIII-19
2 Actual slinger configuration ..... XVIII-20
3 Slinger model with notation ..... XVIII-21

## NOMENCLATURE

a - Axial clearance between seal and housing (in)
b - Blade height (in)
C - Constant pressure specific heat (BTU/1bm- ${ }^{\circ} R$ )
$\mathrm{C}_{\mathrm{m}} \quad-\quad$ Moment coefficient (one side), $\tau / 1 / \rho \rho \omega^{\mathbf{2}}{ }_{\mathbf{r}}{ }^{5}$
$C_{m s} \quad$ - Smooth disk moment coefficient (one side)
$\mathrm{C}_{\mathrm{mv}} \quad$ - Vaned disk moment coefficient
$E_{C} \quad-\quad$ Eckert Number, $\omega^{\mathbf{2}} \mathbf{R}^{\mathbf{2}} / \mathbf{C} \boldsymbol{T}_{0}$
h - Enthalpy (BTU/1bm)
$k \quad$ - Thermal conductivity (BTU/ft.-hr. $-{ }^{0}$ R)
$K \quad$ - Fluid angular velocity ratio, $\beta / \omega$
$\dot{m} \quad-\quad$ Mass flow rate $(1 \mathrm{bm} / \mathrm{sec})$
P - Pressure ( $1 \mathrm{bf} / \mathrm{in}{ }^{2}$ )
$\overline{\mathrm{P}} \quad$ - Dimensionless pressure, $\mathrm{p} / \rho \omega^{\mathbf{2}} \mathrm{R}^{\mathbf{2}}$
$\operatorname{Pr} \quad-\quad$ Prandtl number, $\mathbf{c} \mu / \mathbf{k}$
$r \quad-\quad$ Radial coordinate, radial distance (in)
$\bar{r} \quad-\quad$ Dimensionless radial coordinate, $r / R$
R - Disk radius (in)
Rint - Liquid-vapor interface radius (in)
Re - Reynolds number, $\omega \mathbf{R}^{\mathbf{2} / \nu}$
T - Temperature ( ${ }^{\circ} \mathrm{R}$ )
$\overline{\mathrm{T}} \quad$ - Dimensionless temperature, $T / T_{0}$
$T_{0} \quad-\quad$ Seal fluid flow inlet temperature $\left({ }^{\circ} R\right)$
u - Radial velocity component (ft/sec)
$\bar{u} \quad$ - Dimensionless radial velocity component, u/R

| v | - | Tangential velocity component (ft/sec) |
| :---: | :---: | :---: |
| $v$ | - | Dimensionless tangential velocity component, $\mathrm{v} / \mathrm{R} \omega$ |
| w | - | Axial velocity component (ft/sec) |
| w | - | Dimenisonless axial velocity component, w/R $\omega$ |
| z | - | Axial coordinate, axial distance (in) |
| $\bar{z}$ | - | Dimensionless axial coordinate, z/R |
| $\beta$ | - | Angular velocity of fluid (rad/sec) |
| $\mu$ | - | Viscosity ( $1 \mathrm{bm} / \mathrm{ft}-\mathrm{sec}$ ) |
| $\nu$ | - | Kinematic viscosity ( $\mathrm{ft}^{2} / \mathrm{sec}$ ) |
| $\rho$ | - | Density ( $1 \mathrm{bm} / \mathrm{ft}^{3}$ ) |
| $\tau$ | - | Torque (ft-lbf) |
| $\omega$ | - | Angular velocity of disk ( $\mathrm{rad} / \mathrm{sec}$ ) |
|  |  | Subscripts denoting radial location on disk |

## INTRODUCTION

In high pressure oxidizer turbopumps it is necessary to have effective sealing systems to prevent contact between the turbine hot gas and the oxidizer being pumped. The centrifugal dynamic shaft seal, or slinger seal, can be a valuable part of such a sealing system.

In a slinger seal, a disk shaped impeller with vanes on one side and the other side smooth, rotates with the shaft to be sealed inside a close fitting, but non-contacting housing, as shown in Fig. 1. The liquid to be sealed is centrifuged to the outer periphery of the rotating disk, but because the vaned side has a steeper pressure gradient [1] there will be a resulting net pressure difference across the seal which limits the fluid flow. A gas introduced to the system or vaporized liquid will form an interface with the liquid on the vaned side of the seal.

The primary seal in the Space Shuttle Main Engine high pressure oxidizer turbopump uses this type of slinger with a downstream labyrinth as shown in Fig. 2. The actual seal is modeled by the simplified geometry indicated by Fig. 3.

Previous investigations have primarily been concerned with the more general topic of rotating disks. The theoretical work on free rotating disks done by von Kärmän [2], which was corrected by Cochran [3], and the analysis of the enclosed rotating disk by Schultz-Grunow [4] are fundamental to many of the later studies. Daily and Nece, with their experimental work on enclosed rotating disks [5], [6], [7], laid a base for much of the theoretical [8], [9] as well as experimental [10], [11], [12] work that followed.

A limited number of researchers have studied the utilization of enclosed rotating disks as centrifugal seals, with most of their work being experimental or the correlation of experimental data with simple models [13], [14], [15], [16], [17], [18]. A notable exception is the theoretical analysis of smooth slinger seals by Rosenthal and Reshotko [19], [20].

The purpose of this paper is to analyze the dynamics of the fluid on the vaned side of a slinger seal. The governing equations are established and reduced to a form solvable by numerical methods. Time constraints prevent the complete solution by matching of the boundary conditions and numerical analysis so this will be done as a future task.

Some of the important thermodynamic design considerations for this type seal are the radial pressure distribution, temperature distribution, and the liquid-vapor equilibrium interface. A solution is found for the radial pressure distribution and in order to look at the liquid-vapor interface a simplified analysis is used to obtain a restrictive temperature distribution.

## DEVELOPMENT OF EQUATIONS

The fluid will be treated as an incompressible, Newtonian fluid where body forces and radiation effects are negligible. Temperature variations will generally be small so the viscosity, thermal conductivity and specific heats can be treated at constants for a particular temperature range [21]. The flow is steady and axisymmetric. Using these assumptions and the usual notation for rotational flow, the governing equations are:

Continuity

$$
\begin{equation*}
\frac{\partial u}{\partial r}+\frac{u}{r}+\frac{\partial w}{\partial z}=0 \tag{1}
\end{equation*}
$$

Radial Momentum

$$
\begin{equation*}
u \frac{\partial u}{\partial r}-\frac{v^{2}}{r}+w \frac{\partial u}{\partial z}=-\frac{1}{e} \frac{\partial p}{\partial r}+v\left[\frac{\partial^{2} u}{\partial r^{2}}+\frac{\partial}{\partial r}\left(\frac{u}{r}\right)+\frac{\partial^{2} u}{\partial z^{2}}\right] \tag{2}
\end{equation*}
$$

Tangential Momentum

$$
\begin{equation*}
u \frac{\partial v}{\partial r}+\frac{u v}{r}+w \frac{\partial v}{\partial z}=v\left[\frac{\partial 2 v}{\partial r^{2}}+\frac{\partial}{\partial r}\left(\frac{v}{r}\right)+\frac{\partial 2 v}{\partial z^{2}}\right] \tag{3}
\end{equation*}
$$

Axial Momentum

$$
\begin{equation*}
u \frac{\partial w}{\partial r}+w \frac{\partial w}{\partial z}=-\frac{1}{e} \frac{\partial p}{\partial z}+v\left[\frac{\partial{ }^{2} w}{\partial r^{2}}+\frac{1}{r} \frac{\partial w}{\partial r}+\frac{\partial{ }^{2} w}{\partial z^{2}}\right] \tag{4}
\end{equation*}
$$

Energy

$$
\begin{align*}
& \rho C\left(u \frac{\partial T}{\partial r}+w \frac{\partial T}{\partial z}\right)=k\left(\frac{\partial^{2} T}{\partial r^{2}}+\frac{1}{r} \frac{\partial T}{\partial r}+\frac{\partial^{2} T}{\partial z^{2}}\right)+  \tag{5}\\
& \mu\left[2\left(\frac{\partial u}{\partial r}\right)^{2}+2\left(\frac{u}{r}\right)^{2}+2\left(\frac{\partial w}{\partial z}\right)^{2}+\left(\frac{\partial v}{\partial z}\right)^{2}+\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial r}\right)^{2}+\left(\frac{\partial v}{\partial r}-\frac{v}{r}\right)^{2}\right]
\end{align*}
$$

Equations (1)-(5) can be non-dimensionalized by using the following substitutions:

$$
\bar{r}=\frac{r}{R}, \bar{Z}=\frac{Z}{R}, \bar{u}=\frac{u}{\omega R}, \bar{v}=\frac{v}{\omega R}, \bar{w}=\frac{w}{\omega R}, \bar{P}=\frac{P}{\rho \omega^{2} R^{2}}, \bar{T}=\frac{T}{T_{0}}
$$

where $T_{0}$ is the fluid inlet temperature.
Using these expressions the differential equations become in dimensionless form:

Continuity

$$
\begin{equation*}
\frac{\partial \bar{u}}{\partial \bar{p}}+\frac{\bar{u}}{\bar{r}}+\frac{\partial \bar{w}}{\partial \bar{z}}=0 \tag{6}
\end{equation*}
$$

Radial Momentum

$$
\begin{equation*}
\bar{u} \frac{\partial \bar{u}}{\partial \bar{p}}-\frac{\bar{v}^{2}}{\bar{r}}+\bar{w} \frac{\partial \bar{u}}{\partial \bar{z}}=-\frac{\partial \bar{p}}{\partial \bar{r}}+\frac{1}{\operatorname{Re}}\left[\frac{\partial^{2} \bar{u}}{\partial \bar{r}^{2}}+\frac{1}{\bar{r}} \frac{\partial \bar{u}}{\partial \bar{r}}-\frac{\bar{u}}{\bar{r}}+\frac{\partial^{2} \bar{w}}{\partial \bar{z}^{2}}\right] \tag{7}
\end{equation*}
$$

Tangential Momentum

$$
\begin{equation*}
\bar{u} \frac{\partial \bar{v}}{\partial \bar{r}}+\frac{u \bar{v}}{\bar{r}}+\bar{w} \frac{\partial \bar{v}}{\partial \bar{z}}=\frac{1}{R e}\left[\frac{\partial^{2} \bar{v}}{\partial \bar{r}^{2}}+\frac{1}{\bar{r}} \frac{\partial \bar{v}}{\partial \bar{r}}-\frac{\bar{v}}{\bar{r}^{2}}+\frac{\partial^{2} \bar{v}}{\partial \bar{z}^{2}}\right] \tag{8}
\end{equation*}
$$

Axial Momentum

$$
\begin{equation*}
\bar{u} \frac{\partial \bar{w}}{\partial \bar{r}}+\bar{w} \frac{\partial \bar{w}}{\partial \bar{z}}=-\frac{\partial \bar{p}}{\partial \bar{z}}+\frac{1}{R_{e}}\left[\frac{\partial^{2} \bar{w}}{\partial \bar{r}^{2}}+\frac{1}{r} \frac{\partial \bar{w}}{\partial \bar{r}}+\frac{\partial^{2} \bar{w}}{\partial \bar{z}^{2}}\right] \tag{9}
\end{equation*}
$$

Energy

$$
\begin{aligned}
& \overline{\mathbf{u}} \frac{\partial \bar{T}}{\partial \bar{r}}+\bar{w} \frac{\partial \bar{T}}{\partial \bar{z}}=\frac{1}{\operatorname{PrRe}}\left(\frac{\partial^{2} \bar{T}}{\partial \bar{r}^{2}}+\frac{1}{\bar{r}} \frac{\partial \bar{T}}{\partial \bar{r}}+\frac{\partial^{2} \bar{T}}{\partial \overline{\mathrm{z}}^{2}}\right)+ \\
& \frac{E c}{R e}\left[2\left(\frac{\partial \bar{u}}{\partial \bar{r}}\right)^{2}+2 \frac{\bar{u}^{2}}{\bar{r}^{2}}+2\left(\frac{\partial \bar{w}}{\partial \bar{z}}\right)^{2}+\left(\frac{\partial \bar{v}}{\partial \bar{z}}\right)^{2}+\left(\frac{\partial \bar{u}}{\partial \bar{z}}\right)^{2}+\frac{\partial \bar{u} \partial \bar{w}}{\partial \bar{z} \partial \bar{r}}+\right. \\
& \left.\left(\frac{\partial \bar{w}}{\partial \bar{r}}\right)^{2}+\left(\frac{\partial \bar{v}}{\partial \bar{r}}\right)^{2}-\frac{\bar{v}}{\bar{r}} \frac{\partial \bar{v}}{\partial \bar{r}}+\frac{\bar{v}^{2}}{\bar{r}^{2}}\right]
\end{aligned}
$$

where

$$
\begin{equation*}
R_{\theta}=\frac{\omega R^{2}}{\nu}, \operatorname{Pr}=\frac{\mu C}{k} \quad, E c=\frac{\omega^{2} R^{2}}{C T_{0}} \tag{11}
\end{equation*}
$$

Equations (6)-(9) can be solved independent of equation (10) and are to be done by numerical analysis. The resulting flow field solution can then be used to solve equation (10) for the temperature distribution. The entire solution of these equations will be completed at a later time.

## RADIAL PRESSURE DISTRIBUTION

The sealing capability of a slinger will be determined by the radial pressure gradient. The exact nature of this gradient will be quite complex due to the complicated flow across the disk. As an example, see reference 22 for a detailed look at secondary flows in a centrifugal impeller. Other interrelated factors are the presence of a liquid-vapor interface [18], pressure variation between vanes [23], and fluid throughflow [24], although the experimental results of Due [25] indicated that there was no significant dependance on flow coefficient.

A simplified approach is necessary so it is assumed that once steady state operation is reached the fluid in any thin annulus moves as a solid mass rotating about its own axis with angular velocity $K \omega$ where $K$ is the ratio of the fluid angular velocity, $\beta$, to the disk angular velocity, $\omega$. The radial velocity will be small compared to the tangential velocity and will be neglected, so the radial momentum equation (2) becomes:

$$
\begin{equation*}
\frac{1}{r} v^{2}=\frac{1}{\rho} \frac{\partial p}{\partial r} \tag{12}
\end{equation*}
$$

Since the pressure is a function of the radius only, this equation reduces to an ordinary differential equation. Then substituting $K \omega r$ for the tangential velocity results in:

$$
\begin{equation*}
\frac{d p}{d r}=\frac{\rho(K \omega r)^{2}}{r} \tag{13}
\end{equation*}
$$

Integrating between any two radial points on the face of the slinger yields:

$$
\begin{equation*}
\Delta p=p_{2}-p_{1}=\frac{\rho}{2} K^{2} \omega^{2}\left(r_{2}^{2}-r_{1}^{2}\right) \tag{14}
\end{equation*}
$$

Other investigators [5], [12], [19], have obtained the same result by different approaches and their close correlation with their experimental results verifies its general applicability.

The value of $K$ must be determined to solve for the pressure gradient, but $K$ will be influenced by several variables including geometry of the seal, solidity, clearances, Reynolds number and radius. Since the fluid between the vanes on a slinger will rotate essentially with the slinger angular velocity, $K \approx 1$, but the fluid in the gap between the vanes and the housing will rotate at some lesser angular velocity. The regions could be considered separately but in reference 25 Due established equations for $K$ by correlating experimental data to equation (14) using a statistical regression analysis. The applicable equation from Due is:

$$
\begin{equation*}
K=0.951-0.282 \frac{r}{R}+0.2137 \frac{b}{a+b}+0.175(b) \tag{15}
\end{equation*}
$$

where $R$ is the seal radius, $a$ is the seal to housing gap width, $b$ is the blade height, and $r$ is the particular radial distance. Average values for the subject seal are: $a=0.08 \mathrm{in} ., b=0.25 \mathrm{in} .$, and $R=2.24 \mathrm{in}$. Using these values in equations (14) and (15) the radial pressure distribution can be determined.

## TEMPERATURE DISTRIBUTION

Due to the rotation of the slinger, energy is imparted to the fluid, manifested as a temperature rise and eventually as a phase change from liquid to vapor. This phase change causes a drastic change in the density of the fluid and a resulting smaller radial pressure change. To account for this effect, the radial temperature distirbution must be known.

A simple control volume analysis will be used to look at the enthalpy increase from one radial plane to another. Knowing the enthalpy and pressure defines the state at a point so the temperature, density or any other required properties can be found using a standard computer code [27] or from tables or charts [28]. The process will be assumed to be steady state, steady flow [26] and the radial distance between inlet and exit planes will be small so the change in kinetic and potential energies will be negligible. Heat transfer to the walls will be considered negligible. Continuity requires the mass flow at the inlet to equal that of the exit and with the other conditions as above, the one dimensional energy equation will be:

$$
\begin{equation*}
\dot{m} h_{1}=\dot{m} h_{2}+\dot{w} \tag{16}
\end{equation*}
$$

where $\dot{m}$ is the fluid mass flow through the seal, $h$ is the enthalpy at a particular radius and $\dot{W}$ is the shaft power absorbed by the fluid.

$$
\begin{equation*}
\dot{\mathrm{W}}=-\tau \omega \tag{17}
\end{equation*}
$$

with torque being defined for one side of a disk $[5],[12]$ as:

$$
\begin{equation*}
\tau=\frac{c_{m} \rho \omega^{2}\left(r_{2}^{5}-r_{1}{ }^{5}\right)}{4} \tag{18}
\end{equation*}
$$

so the change in enthalpy becomes:

$$
\begin{equation*}
h_{2}-h_{1}=\frac{c_{m} \rho \omega^{3}\left(r_{2}{ }^{5}-r_{1}^{5}\right)}{4 \dot{m}} \tag{19}
\end{equation*}
$$

The moment coefficient $C_{m}$ must be determined. For the smooth side of the seal the value of $C_{m s}$ is found to be 0.0008 for $R_{e}=10^{8}$ using the empirical
equations from reference 8. For this seal the flow is between regimes 3 and 4 as defined by Daily and Nece [5] so an average value of $C_{m s}$ from the two applicable equations is used. The disk is thin so no contributions due to the tip are considered.

For the vaned side of the disk the moment coefficient will be greater and will be approximated by the method derived by Thew [15] where

$$
\begin{equation*}
C_{m v}=C_{m s}+1.2 \frac{b+a}{R}(1-K)^{2}\left[\left(\frac{a+b}{0.8 a}\right)^{-1}\right]^{2}\left[1-\left(\frac{R_{i n t}}{R}\right)^{4}\right] \tag{20}
\end{equation*}
$$

The pressure and enthalpy distributions can now be calculated, yielding any other thermophysical properties needed at a radial point. The velocity ratio is found to be essentially constant for the smooth side, $K=0.44$, but varies with radius on the vaned side so $\mathrm{C}_{\mathrm{m}}$ also varies as shown. The mass flow of $0.13 \mathrm{lbm} / \mathrm{sec}$ and slinger angular velocity of $2932 \mathrm{rad} / \mathrm{sec}$ are obtained from engine balance data [30] . On the vaned side it is necessary to assume a value of Rint, determine the point at which the fluid begins vaporizing, and use this as the new assumed value of $\mathrm{R}_{\mathrm{in}}$, continuing until the assumed and determined values converge. For this seal a radius of 2.2 in. is where vaporization, therefore the interface begins.

Using the property values from reference 29 (their station 20) as being at the root of the smooth side of the slinger ( $r=1.7 \mathrm{in}$.) the radial property variations are calculated and summarized in Table 1.

## CONCLUSIONS AND RECOMMENDATIONS

A thorough literature survey has been completed with the most applicable work listed beginning on page XVIII-16. A set of non-dimensional partial differential equations have been established which describe the flow on the face of a slinger seal. The pressure and temperature distributions for the Space Shuttle Main Engine high pressure oxidizer turbopump primary slinger seal have been described and the location of the liquid-vapor interface determined.

The radial pressure distribution equation (14) should provide good results based on the experimental work confirming it, but the analysis of the temperature variation is approximate at best and too many simplifying assumptions are made to trust its validity. The smooth side values for $C_{m}$ are sufficiently grounded with experimental work but the equation for $\mathrm{C}_{\mathrm{mv}}$ has not been investigated enough to verify its generality. Due to these factors and the nature of the inital data used from reference 29 , it is believed that this simplified analysis can be used only to predict tendencies.

Further work needs to be done on vaned slinger seals. Solution of the governing equations and matching of all boundary conditions should be done to get a complete picture of the flow involved and to predict more accurate pressure and temperature distributions. An experimental program should be conducted to determine actual pressure and temperature values to correlate with the analytical work. No previous experimental work could be found through the literature survey which deals with liquid oxygen, high rotational speeds as found in modern turbopumps, or the energy transfer involved with the production of a liquid-vapor interface.

Recommendations based on the literature survey that could improve the sealing performance of the Space Shuttle Main Engine high pressure oxidizer turbopump slinger are:

- include circumferential tangs on the vaned side of the seal with matching grooves in the housing to stabilize the liquid-vapor interface, thereby reducing leakage [14].
- increase the number of vanes to 24 which has been established as optimum [14].
- utilize a non-wetting coating on the housing opposite the vaned side of the seal to reduce secondary flow that enhances leakage. [20].


## REFERENCES

1. Thew, M. T. and M. G. Saunders, "The Hydrodynamic Disk Seal", Proc. 3rd Int. Conf. on Fluid Sealing, British Hydromechanics Research Assoc., Paper H5, April, 1967.
2. Kármán, Th. von, "Uber laminare und turbulent Reibung", Zeitschrift angew. Math. Mech., Vol. 1, 233-252, 1921.
3. Cochran, W. G., "The flow due to a rotating disk", Proc. Cambridge Phil. Soc. 30, 365-375, 1934.
4. Schultz-Grunow, F., "Der Reibungswiderstand Rotierender Scheiben in Gehausen", Zeitschrift angew Math. Mech., Vol. 15, 191, 1935.
5. Daily, J. W. and R. D. Nece, "Roughness and Chamber Dimensional Effects on Induced Flow and Frictional Resistance of Enclosed Rotating Disks", MIT Hydrodynamics Laboratory Technical Report No. 27, May 1958.
6. Daily, J. W. and R. D. Nece, "Chamber Dimensional Effects on Induced Flow and Frictional Resistance of Enclosed Rotating Disks", Journal of Basic Engineering, Trans. ASME, Series D, Vol. 82, 217-232, 1960.
7. Daily, J. W. and R. D. Nece, "Roughness Effects on Frictional Resistance of Enclosed Rotating Disks", Journal of Basic Engineering, Trans. ASME, Series D, Vol. 82, 553-562, 1960.
8. Ketola, H. N. and J. M. McGrew, "Theory of the Partially Wetted Rotating Disk", Proceedings of the Third International Conference on Fluid Sealing, British Hydromechanics Research Association, Paper H4, April 3-4, 1967.
9. Ketola, H. N. and J. M. McGrew, "Pressure, Frictional Resistance, and Flow Characteristics of the Partially Wetted Rotating Disk", Journal of Lubricating Technology, Trans. ASME, Series F, Vol. 90, 395-404, 1968.
10. Due, H. F. and W. C. Keathly, "A Study of Pressure Prediction Methods for Radial Flow Impellers", Pratt and Whitney Aircraft Report FR-952, 13 April 1964.
11. Due, H. F. and W. E. Young, "Investigation of Pressure Prediction Methods for Radial Flow Impellers, Phase II", Pratt and Whitney Aircraft Report FR-1276, 8 March 1965.
12. Due, H. F. and W. E. Young, "Investigation of Pressure Prediction Methods for Radial Flow Impellers, Phase III", Pratt and Whitney Aircraft Report FR-1716, 4 Feb. 1966.
13. McGrew, J. M. and H. N Ketola, "Non-contacting Dynamic Seals", General Electric Report on "Study of Dynamic and Static Seals for Liquid Rocket Engines, Final Report, Jan. 1, 1965-Sept. 1, 1965".
14. Wood, G. M., D. V. Manfredi and J. E. Cygnor, "Centrifugal Types of Dynamic Shaft Seals", ASME Paper 63-WA-167, 1963.
15. Thew, M. T., "Further Experiments on the Hydrodynamic Disk Seal", Proc. 4th Int. Conf. on Fluid Sealing, British Hydro. Res. Assoc., Paper 39, April 1969.
16. McGrew, J. M. and H. N. Ketola, "Non-contacting Dynamic Seals", General Electric Report on "Study of Dynamic and Static Seals for Liquid Rocket Engines, Final Report, Apr. 1, 1964-0ct. 1, 1964."
17. Thew, M. T., "The Labyrinth/Hydrodynamic Disk Seal (LHDS)", Proc. 8th Int. Conf. on Fluid Sealing, British Hydromechanics Research Assoc., Paper B1, Sept. 1978.
18. Kraev, M. V., "Nature of Liquid-Gas Interface on Bladed Face of a Rotating Disk", Soviet Aeronautics, Vol. 19, No. 1, 116-118, 1976.
19. Rosenthal, R. L. and E. Reshotko, "Laminar Flow Analysis of a Plane Disk Slinger Seal", Case Institute of Technology FTAS/TR-66-14, Aug. 1966.
20. Reshotko, E. and R. L. Rosenthal, "Fluid Dynamic Considerations in the Design of Slinger Seals", Lubrication Engineering, Vol. 24, 303-314, July 1968.
21. Schlichting, H., "Boundary Layer Theory", 6th ed., McGraw-Hill, New York, 1968.
22. Eckardt, D., "Detailed Flow Investigations Within a High Speed Centrifugal Compressor Impeller", Journal of Fluids Engineering, Sept. 1976.
23. Wislicenus, G. F., "Fluid Mechanics of Turbomachinery", 2d ed., Dover, New York, 1965.
24. Daily, J. W., W. D. Ernst, and V. V. Asbedian, "Enclosed Rotating Disks with Superposed Throughflow: Mean Steady and Periodic Unsteady Characteristics of Induced Flow", MIT Hydrodynamics Laboratory Report No. 64, April 1964.
25. Due, H. F., "An Empirical Method for Calculating Radial Pressure Distribution on Rotating Disks", Journal of Engineering For Power , April 1966.
26. Van Wylen, G. J., and R. E. Sonntag, "Fundamentals of Classical Thermodynamics", 2d ed., SI Version, John Wiley and Sons, Inc., New York, 1976.
27. Hendricks, R. C., et al., "Gasp - A Computer Code for Calculating the Thermodynamic and Transport Properties for Ten Fluids: Parahydrogen, Helium, Neon, Methane, Nitrogen, Carbon Monoxide, Oxygen, Fluorine, Argon, and Carbon Dioxide", NASA TN D-7808, Feb. 1975.
28. "ASRDI Oxygen Technology Survey, Vol. I, Thermophysical Properties", NASA SP 3071, 1972.
29. Anderson, P. G., et al., "Fluid Flow Analysis of the SSME High Pressure Oxidizer Turbopump", Lockheed Missile and Space Company, Inc., HREC TR D568277, Aug. 1978.
30. "Rocketdyne SSME FMOF Engine Balance", Nov. 1978.


Figure 1. Cutaway view of slinger seal model


Figure 2. Actual slinger configuration


Figure 3. Slinger model with notation.

TABLE 1
SLINGER RADIAL PROPERTY DISTRIBUTIONS

Smooth Side

| Radius <br> (in.) | Velocity <br> Ratio | Moment <br> Coefficient | Enthalpy <br> $(\mathrm{BTU} / \mathrm{lbm})$ | Pressure <br> $\left(1 \mathrm{bf} / \mathrm{in}^{2}\right)$ | Temperature <br> $\left({ }^{\text {OR })}\right.$ | Density <br> $\left(1 \mathrm{bm} / \mathrm{ft}^{3}\right)$ |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| 1.7 | .44 | .0008 | -34.3 | 334 | 216 | 61.2 |
| 1.8 | .44 | .0008 | -32.5 | 361 | 219 | 60.5 |
| 1.9 | .44 | .0008 | -30.3 | 388 | 223 | 59 |
| 2.0 | .44 | .0008 | -27.5 | 416 | 230 | 58 |
| 2.1 | .44 | .0008 | -24.3 | 446 | 237 | 56 |
| 2.2 | .44 | .0008 | -20.6 | 476 | 245 | 54 |
| 2.24 | .44 | .0008 | -18.9 | 488 | 247 | 53 |

Vaned Side

| Radius <br> (in.) | Velocity <br> Ratio | Moment <br> Coefficient | Enthalpy <br> $(\mathrm{BTU} / 1 \mathrm{bm})$ | Pressure <br> $(1 \mathrm{bf} / \mathrm{in} .2)$ | Temperature <br> (OR) | Density <br> $\left(1 \mathrm{bm} / \mathrm{ft}^{3}\right)$ | Quality |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 2.24 |  |  | -18.9 | 488 | 247 | 53 | 0.0 |
| 2.22 | .88 | .0040 | -14.8 | 464 | 254 | 51 | 0.0 |
| 2.20 | .88 | .0039 | -11.2 | 442 | 258 | 48 | 0.0 |
| 2.18 | .88 | .0037 | -8.0 | 421 | 255 | 42 | 0.1 |
| 2.16 | .89 | .0036 | -5.4 | 402 | 253 | 35 | 0.2 |

NAS A/ASEE SUMMER FACULTY RESEARCH FELLOWS HIP PROGRAM

MARS HALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA

## VALIDATION OF THE S OLAR HEATING AND COOLING HIGH S PEED PERFCRMANCS (HISPER) COMPUTER CODE

## Prepared By:

Academic Rank:

University and Department

Donald B. Wallace, Ph. D., P.E.

Associate Professor

The University of Alabama in Huntsville
Department of Mechanical Engineering

Systems Analysis and Integration
Systems Engineering
Systems Requirement \& Definition
Robert F. Elkin

August 1, 1980
NGT-01-002-099
(University of Alabama)

# VALIDATION OF THE SOLAR HEATING AND COOLING HIGH SPEED PERFORMANCE (HI S PER) COMPUTER CODE 

## By

Donald B. Wallace, Ph.D., P. E. Associate Professor of Mechanical Engineering The University of Alabama in Huntsville Huntsville, Alabama 35809

## ABSTRACT

Computer simulations of solar energy systems are used to evaluate system performance, size system components and predict fossil fuel savings. The solar heating and cooling high speed performance (HISPER) computer code was previously developed to give a quick and accurate predictions. As a simplification of the TRNSYS program, it achieves its computational speed by not simulating detailed system operations or performing detailed load computations.

In order to validate the HISPER computer code for air systems the simulation was compared to the actual performance of an operational test site. The site selected was the Home Builders Association of Huntsville office building in Huntsville, Alabama. Solar insolation, ambient temperature, water usage rate, and water main temperatures from the data tapes for the site were used as input for the HISPER program.

The HISPER program was found to predict the heating loads and solar fraction of the loads with errors of less than ten percent. Good correlation was found on both a seasonal basis and a monthly basis. Several parameters (such as infiltration rate and the outside ambient temperature above which heating is not required) were found to require careful selection for accurate simulation.

## ACKNOWLEDGEMENTS

I would like to express my gratitude to the NASA/ASEE Summer Faculty Fellowship Program for providing the research opportunities this summer. Thanks is also extended to Doctor B. F. Barfield, Marion Kent, and Jim Dozier for their capable administrarion of the program. A special appreciation goes to my counterpart, Robert F. Elkin, who provided a lot of technical assistance and support in this project.

## TABLE OF CONTENTS

Abstract ..... i
Acknowledgements ..... ii
Introduction ..... 1
Objectives ..... 2
Operational Test Site ..... 2
Site Description ..... 2
Operational Test Site Data ..... 4
Site Operation ..... 6
HISPER Modifications ..... 6
Simulation Input Data ..... 10
Simulation Results ..... 10
Effect of TLH ..... 11
Effects of Control Logic ..... 11
Hot Water Percentages ..... 17
Conclusions and Recommendations ..... 18
References ..... 19

## INTRODUCTION

Solar energy collection systems are installed on dwellings in order to reduce the occupant's dependence on conventional forms of energy for heating and cooling. Due to the diffuse nature of solar energy and the variability in the weather, it is impractical to install a system large enough to provide all of the energy needs. A back-up system, using conventional energy, must be provided for supplying energy during long periods of cloudy weather. The solar energy system is ideally sized to maximimize the benefit to cost ratio (the ratio of annual solar energy provided divided by the life cycle cost of the solar energy system). Because of the large number of design parameters and the weather variability, the design optimization problem is a difficult one.

Computer programs have been developed to help predict solar heating and cooling (SHAC) performance. One widely used program is TRNSYS which was developed at the University of Wisconsin Solar Energy Laboratory [1]*. This is a general purpose program that allows the user to simulate a system by defining the performance of and the interaction between various types of components (collectors, storage, pumps, control logic, etc.). TRNSYS gives an accurate indication of the model performance but it does require a significant amount of computer time to run. In a design situation it is more desirable to have an analysis tool which can quickly and economically give reasonable answers. This allows more design iterations to be analyzed.

In response to the need for a rapid solar simulation computer program, Northrup Services, Inc. (under contract to NASA) developed HISPER - a high speed performance program [2]. HISPER is basically a simplification of the TRNSYS program in which some detailed system operation and load calculations are not performed (such as room temperature oscillations and ductwork heat losses). When HISPER was compared to TRNSYS for identical cases, the predicted system performance differed by less than 7 percent while computer time required by TRNSYS was approximately 30 times greater than the time required by HISPER.

The usefulness of any simulation is dependent on how well it predicts the performance of the actual system. This study provides this validation of the HISPER program by comparing the predicted results using actual weather data
*numbers in brackets refer to references at the end of this paper.
with the measured performance of the site during this same period. The Home Builders Association of Huntsville office building was selected as the site to be simulated.

This report first describes the Home Builders Association office building and its operation over a period from September, 1978 to April, 1979. Next, the basic organization of the HISPER program and the modifications necessary to run actual weather data is presented. Third, the simulation input data is discussed. The fourth part is a discussion of the results of the study including the effects of several critical design parameters. Finally the results of this study are discussed and recommendations made for the use of HISPER and for topics which need further study.

## OBJECTIVES

The primary objective of this study are to verify the general computational philosophy of the HISPER program by demonstrating that the program adequately predicts the performance of actual solar energy systems. This verification will result in more confidence in HISPER for use as a design tool.

Secondary objectives include: identification of the effects of parameter values on the simulation performance, identification of any desirable changes in the HISPER logic, and the establishment of a post-design review procedure in which the actual performance of a site can be compared with the predicted design value.

## OPERATIONAL TEST SITE

## Site Description:

The operational test site selected for use in this study is the Home Builders Association of Huntsville office building located in Huntsville, Alabama (see Figure 1). This building has two stories with a total floor area of approximately 210 square meters ( 2270 square feet). The building was designed by IBM (under contract to NASA) to be a single family dwelling but is being used as an office building at this location.

The south facing roof is sloped at $45^{\circ}$ and contains 30 solar collectors (manufactured by Solar Energy Products) for a gross area of 67 square meters ( 720 square feet). These collectors are single glazed and use air as the energy transport medium.


Figure 1. Office Building for the Home Builders
Association of Huntsville, Alabama

Figure 2 shows a schematic of the solar energy collection system. Air heated in the solar collectors first passes through an air/water heat exchanger so that the domestic water going to the hot water heater (not shown) may be preheated. Then the hot air is directed by the air handles either to the house or to the pebble bed storage, depending upon the energy needs of the house at that time. The pebble bed storage has a volume of 14 cubic meters ( 500 cubic feet) and contains 23 metric tons ( 25 tons) of river rock.

During periods when no solar energy is being collected, house air may be passed through the pebble bed thus providing energy to the house. If the temperature of the pebble bed storage is too low (below $32^{\circ} \mathrm{C}$ or $90^{\circ} \mathrm{F}$ ), a Westinghouse air-to-air heat pump provides the heat demanded by the house.

A complete description of the building and the solar energy system design and operation is given in various reports prepared by the designers - the IBM Corporation Federal Systems Division [3, 4, 5].

## Operational Test Site Data:

The Home Builders Association office building was equipped with 45 sensors to obtain data for performance evaluation. These include 27 temperature sensors, 7 electric power recorders, 5 flow rate sensors, a solar insolation indicator, and 5 others. Each of these sensors is scanned at 320 second intervals ( 5 minutes and 20 seconds) and their average value is recorded by a data logger. All of this information is then put on to magnetic tapes.

This data is used by IBM to prepare a monthly performance summary for the site. The performance summary includes the climatic conditions, house heating and hot water loads, and solar collection information. The monthly summaries also include any known problems (such as malfunctioning equipment) which occur during the month.

The period chosen for the comparison of HISPER and the actual system was October 1978 through April, 1979. This was chosen since both the 5 minute recorder data (on magnetic tape) and the monthly reports were available. Summer months were not included since the real system is put in "Summer Mode" in which the pebble bed storage is not used. HISPER does not have the capabilities of the 'Summer Mode'.

A review of the information on the data tapes indicated that the data tapes could not be used directly with the HISPER program. First, periodic sensor malfunctions were discovered (such as ambient temperature going from about $29^{\circ} \mathrm{C}\left(84^{\circ} \mathrm{F}\right)$ to $-20^{\circ} \mathrm{C}\left(-4^{\circ} \mathrm{F}\right)$ for a period of about 45 minutes and then back again). Secondly, gaps in the data of up to 30 hours was observed on several occasions. These problems were overcome by correcting the obviously erroneous data and inserting reasonable values for the missing data. This data manipulation will result in a small discrepancy in the IBM Monthly Reports and the HISPER predictions. Since the total amount of bad or missing data is


Figure 2. Actual System Configuration
very small, the discrepancy is insignificant.

## Site Operation:

A review of the test site data revealed several areas which could lead to significant errors. These relate to hot water usage and temperature control of the house.

Since the site is being used as an office building rather than a residence, the demand for hot water was quite low. When the hot water tank was set at $63^{\circ} \mathrm{C}$ ( $145^{\circ} \mathrm{F}$ ), before November 22,1978 , an average of only 30 kg ( 7.8 gallons) of hot water were used daily. After the hot water temperature was lowered to $51^{\circ} \mathrm{C}\left(123^{\circ} \mathrm{F}\right)$ on November 22, 1978 , the hot water demand rose to an average of 36 kg ( 9.5 gallons) a day. Very little water was used on the weekends.

The temperature of the water coming out of the water main fluctuates over the period of the year. Water main temperature varies from a high in August at about $25^{\circ} \mathrm{C}\left(77^{\circ} \mathrm{F}\right)$ to a low in February at about $9^{\circ} \mathrm{C}\left(48^{\circ} \mathrm{F}\right)$. This variation probably results from the water being transported in mains that are not buried deep enough to be unaffected by seasonal temperature variations.

Another problem is illustrated in Figure 3. March 24th and 25th were cold weekend days following a rather mild week (high ambient temperatures). Apparently someone had shut off the thermostat before leaving work on Friday so that when the workers returned on Monday, the building temperature was $8^{\circ} \mathrm{C}\left(46^{\circ} \mathrm{F}\right)$. There is other evidence that the thermostat settings had been adjusted over the period of investigation.

Figure 3 also shows that a significant variation in room temperature does occur on a daily basis. This may be due to a direct solar gain either through the windows or through the walls. The resulting heating in the occupied area will reduce the amount of solar or auxiliary energy required.

## HISPER MODIFICATIONS

The HISPER simulation program consists of a main program and three fundamental subroutines: MODEL, WETHR, and SOLAR. The purpose of each of these subprograms is:

1. MAIN - The main program reads the system parameters, performs the time integration, and controls the output of the program. The main program calls subroutine MODEL.


Figure 3. Ambient and Room Temperatures in March, 1979
2. MODEL - This subroutine calculates all of the energy flows in the system model in a given increment of time. It calculates collector performance, the dwelling heating load, the hot water load, and the flow of energy in the solar system (including in to and out of storage). It calls WETHR and SOLAR.
3. WETHR - This subroutine provides ambient temperature and horizontal solar insolation to MODEL. These values are read from weather tapes containing typical weather data.
4. SOLAR - This subroutine receives the total hourly horizontal insolation output by the WETHR subroutine and converts it to total insolation incident on the collector surface using one of four user determined methods.

The basic HISPER procedure, with the exception of WETHR, was not changed significantly. Some control logic was added to MODEL and different forms of output were provided for in MAIN. Subroutine SOLAR was not altered at all. Since the climatic data was from a new source, WETHR had to be changed completely.

HISPER was designed to read weather tapes which contain typical climatic data for a particular location. In order to validate HISPER predictions for an actual operational test site, it was necessary to modify HISPER to read the actual climatic conditions at that site. Sensors at the Home Builders Association office building include an ambient temperature sensor and a pyranometer to measure total solar insolation in the collector plane.

HISPER uses both insolation on a horizontal surface (for load calculations) and insolation on the collector plane (for solar collector performance) in its simulation procedure. Normally the horizontal insolation is provided by the weather tapes and HISPER uses one of four standard techniques to estimate the tilted insolation from the horizontal value (see references [1, 6, 7]). Since the site data includes only tilted insolation data, HISPER's WETHR subroutine had to be modified to reverse the standard procedures and predict the horizontal insolation value which will produce the observed tilted value when SOLAR is called.

HISPER is designed to simulate a typical domestic dwelling so the assumed hot water load represents what a normal family might use each day. Since the system being studied is an office building, the HISPER predicted daily usage ( $224 \mathrm{~kg}--59 \mathrm{gal}$ ) does not correspond well to the actual usage of about 33 kg ( 8.7 gallons) per day. Therefore HISPER was modified to read (in WETHR) actual water usage and true water main temperature from the site data.

The control logic used in HISPER is designed to provide the maximum amount of energy to the loads as possible. The main control features include:

1. Turning on the collector when ever energy can be gained.
2. Energy is extracted by the hot water preheat tank whenever the collector is on and the outlet temperature is greater than the preheat tank temperature.
3. Energy is provided to the house whenever the house has a load and the air inlet temperature, either from the collector or storage, is greater than the house temperature (returns air temperature is assumed to be room temperature).
4. Energy is not allowed to go to storage if the temperature from the collector is less than the temperature at the top of the storage bed.

In order to study the effects of actual control settings on the HISPER Simulation performance, the following features were added to the MODEL subroutine control logic.

1. A minimum temperature difference between the collector plate temperature and the bottom of the pebble bed storage bin before the collector would collect energy (note: the plate temperature was based on steady state flow conditions) Actual site required a $15^{\circ} \mathrm{C}\left(28^{\circ} \mathrm{F}\right)$ temperature difference [8].
2. A minimum temperature difference between the collector outlet temperature and the preheat tank for domestic hot water before energy could be extracted from the air and be put in the preheat tank. Actual site required $11^{\circ} \mathrm{C}\left(20^{\circ} \mathrm{F}\right)$ to turn on and no less than $2^{\circ} \mathrm{C}\left(3^{\circ} \mathrm{F}\right)$ before turning off [8].
3. A minimum temperature at the top of the pebble bed storage was necessary for the storage to be allowed to provide energy to the house. Actual site required a temperature from the storage of at least $32^{\circ} \mathrm{C}$ ( $90^{\circ} \mathrm{F}$ ). [8].
4. An air return temperature from the house that was lower than the nominal house temperature. Review of data tapes indicates that the return air is approximately $5^{\circ} \mathrm{C}\left(9^{\circ} \mathrm{F}\right)$ below room temperature. This is probably due to thermal stratification in the rooms.

The control logic as modified by these four features will be referred to as the "modified control logic."

## SIMULATION INPUT DATA

A variety of sources was used to develop the input data for the HISPER simulation program. The attempt was to make the model for the Home Builders of Huntsville office building as realistic as possible.

The manufacturer's data for the sunworks solar collectors was used since other tests of the collectors by Wyle Laboratories showed good correlation with the predicted performance. Manufacturers data was also used for the air to water heat exchanger used for preheating water.

Flow rates in the air ducts were taken from actual tests performed by IBM after the construction of the test site.

Internally generated heat loads were assumed to be $2890 \mathrm{~kJ} / \mathrm{hr}$. This corresponds to the average heat generation at 10 people, lights, 2 refrigerators, and three coffee pots. The storage losses (since the storage bed is in the building, is considered to be part of the internally generated heat.

The parameters used for heat load calculations are primarily based on a design analysis done at NASA/MSFC in 1977. After the Home Builders Association building was constructed, tests were conducted to determine actual heat loss coefficients. The results of the study indicate that the building load demand is commensurate with the $720 \mathrm{BTUH} /{ }^{\circ} \mathrm{F}$ design building loss coefficientarea product (UA).

The only significant difference between the NASA / MSFC design values and the values used in the HISPER simulation was for infiltration. The NASA/MSFC value was one half air change per hour which assumes very tight construction. The tests performed on the building were conducted at night with no employees present at the office building. ASHRAE [9] recommends design infiltration rates from one half to two air changes per hour, depending on the number of windows and doors in the building. Since the building is being used as an office, with the possibility of many people entering and leaving during the day, an air infiltration rate of 1.5 air changes per hour was used.

## SIMULATION RESULTS

HISPER was designed to give a reasonable prediction of solar performance of a home. This means that it should be able to predict the loads and the amount of solar energy delivered to the loads. The accuracy of the predictions was investigated with respect to several of the input parameters: infiltration rate, the outside ambient temperature above which the house is assumed to require no heat (TLH), and the effect of control logic. The domestic hot water predictions were also investigated.

Effect of TLH: (Temperature above which no energy is required by the house)
A house does not always require energy whenever the outside ambient temperature is less than the house temperature. This is because any house has some direct solar gain and a large amount of thermal capacitance. The direct solar gain statisfies some of the current house demand as well as being absorbed in the house for use at night. This effect can be seen in the room temperature fluctuations in Figure 3.

The data tapes for the Home Builders Association building were searched to determine what percentage of the time the heating/cooling system was on as a function of the outside temperature. Figure 4 shows this relationship. The increase in the percentages for the higher temperatures is due to cooling, not heating. From Figure 4, it is obvious that if the house is set at $22^{\circ} \mathrm{C}$ $\left(72^{\circ} \mathrm{F}\right)$, it actually does not require significant heating until well below that value, at say $10^{\circ} \mathrm{C}$.

Initial simulations were run assuming heating was not required over an ambient temperature of $10^{\circ} \mathrm{C}$, i. e. TLH $10^{\circ} \mathrm{C}\left(50^{\circ} \mathrm{F}\right)$. The HISPER results are shown on a monthly load basis in Figure 5 and on a daily load basis in Figure 6 (a). An infiltration rate of 1.1 was assumed so that the total house load for the period was correct. This value of TLH results in HISPER under predicting the actual loads on cold days (heavy loads) and over predicting on warm days (light loads).

A value of TLH of $7^{\circ} \mathrm{C}\left(45^{\circ} \mathrm{F}\right)$ and an infiltration rate of 1.5 air changes an hour was then used to produce the results in Figure 5 and Figure 6 (b). These values result in a very good load prediction on both a monthly and daily basis.

## Effects of Control Logic:

Once the actual loads were matched through the appropriate selection of TLH, the portion of the total load provided by solar energy was investigated. Figure 7 shows that a good correlation was achieved on a monthly basis.

A closer examination of component performance revealed that the temperature at the top of the storage was higher after periods of heavy load than HISPER predicted. (See Figures 8 and 9). This was due to the actual control logic which stopped energy withdrawal at temperatures of less than $32^{\circ} \mathrm{C}\left(90^{\circ} \mathrm{F} 0\right.$. The storage bed model of HISPER is programmed to supply, on demand, energy to the building anytime the temperature at the top of the storage is greater than the thermostat setting.

The temperature at the bottom of the pebble bed storage at the actual site often reached temperatures $5^{\circ} \mathrm{C}\left(9^{\circ} \mathrm{F}\right)$ below the nominal room temperature setting. This is probably due to the return air from the house being at a lower temperature than the room.

$$
\text { XIX - } 11
$$



Figure 4. Heating/Cooling System "on" Time


Figure 5. Predictions of Monthly Heating Load

(a) $\mathrm{TLH}=10^{\circ} \mathrm{C}$

(b) $\mathrm{TLH}=7^{\circ} \mathrm{C}$

Figure 6. Actual Versus Predicted Daily Heating Loads


Figure 7. Solar Percent of House Heating Load

(a) Actual Test Site Data

(b) FIISPER Prediction (Modified Control Logic)


Figure 8. December Storage 'I'emperatures (Top, Midde, \& Bottom)

(a) Actual Test Site Data

(b) HISPER Prediction (Modified Control Logic)

(c) HISPER Prediction (Stándard Control Logic)

Figure 9. January Storage Temperatures (Top, Middle, \& Bottom)

The four changes to the "control logic detailed in the HISPER Modification section were made with the following values:

1. Minimum temperature difference between collector plate and storage bottom: $7^{\circ} \mathrm{C}\left(13^{\circ} \mathrm{F}\right)$
2. Minimum temperature difference between collector output temperature and preheat tank temperature: $5^{\circ} \mathrm{C}\left(9^{\circ} \mathrm{F}\right)$
3. Minimum temperature from storage to house: $32^{\circ} \mathrm{C}\left(90^{\circ} \mathrm{F}\right)$
4. House return temperature: $5^{\circ} \mathrm{C}\left(9^{\circ} \mathrm{F}\right)$ below room temperature.

The effects of this modified control logic can be seen in Figures 7, 8 and 9. The temperature difference criteria for turning on the collector and preheat tank heat exchanger did not have a great effect on the results. The minimum temperature to the house requirement reduced the solar percentages since the storage could no longer be "drained." Lowering the return temperature from the house allowed the collectors to operate more efficiently, collect more energy and thus deliver more to the house. The last two effects tend to counteract one another thus resulting in still a very good prediction of solar percentages on a monthly basis (Figure 7). Figures 8 and 9 show that these changes do result in a significant improvement in the simulation of the pebble bed storage behavior.

## Hot Water Percentages:

Hot water loads were accurately predicted since almost all of the information required for the load calculation was read from the data tapes. The percentage of the domestic hot water was not predicted well by HISPER. HISPER predicted a larger solar fraction than actually existed.

The inaccurate prediction is a result of several factors. First, the water usage demand is so low in this application that energy losses are of a much greater magnitude than energy used in heating the water actually used. Therefore any error in heat loss parameters will have a large effect on the solar percentage.

The second effect is that HISPER only simulates heat lost from the preheat tank, not the standard hot water tank. Due to the low usage rate, most of the solar provided heat will be lost before it is used.

The solar percentages of the hot water load for the actual site and the HISPER predictions are shown in Table 1.

Table 1. Solar Percent of Hot Water Load

| Month | Actual Site <br> Percentage | HISPER Predicted <br> Percentage |
| :---: | :---: | :---: |
| October, 1978 | 52 | 99 |
| November,1978 | 49 | 90 |
| December,1978 | 64 | 91 |
| January, 1979 | 50 | 68 |
| February, 1979 | 43 | 69 |
| March, 1979 | 59 | 100 |
| April, 1979 | 64 | 100 |

## CONCLUSIONS AND RECOMENDATIONS

The major conclusion is that the HISPER solar simulation program seems to be a sufficiently accurate design tool. Errors of less than 10 percent were observed when comparing predicted results to the actual results. One caution, however, is that parameters such as infiltration and TLH (the temperature above which no heating is required) must be carefully selected to achieve good monthly predictions.

Although HISPER predictions have been validated against actual performance, several areas were discovered which need further attention:

1. The value of $7^{\circ} \mathrm{C}$ for TLH (temperature above which heating is not required) seems lower than the site data indicates it should be (Figure 4). This should be investigated for this site and others.
2. The validation process should be carried out for several other operational test sites.
3. The four control options previously discussed should be added to HISPER. This will give HISPER greater flexibility but will not significantly increase computer time required for simulation.
4. The water main temperature was observed to vary from $8^{\circ} \mathrm{C}$ to $25^{\circ} \mathrm{C}$ over the period of a year. This variation could significantly affect the predicted hot water load since HISPER currently assumes a constant water main temperature.

## REFERENCES

1. "TRNSYS - A Transient Simulation Program," Solar Energy Laboratory, University of Wisconsin, Engineering Experiment Station Report 38 , March 31, 1975.
2. Gray, J. C., "Solar Heating and Cooling High Speed Performance Program," Technical Note TN - 240-1661, Northrup Services, Inc., Huntsville, Alabama, July, 1976.
3. "System Design Package for IBM System One Solar Heating and Domestic Hot Water," DOE/NASA CR-150614, prepared by IBM Corporation, February, 1977.
4. "SIMS Prototype System 1 - Design Data Brochure," DOE/NASA CR - 150534, prepared by the IBM Corporation, January, 1978.
5. "Installation Package - SIMS Prototype System 1A," DOE/NASA CR - 150524, prepared by the IBM Corporation, December, 1976.
6. Duffie, J. A. and Beckman, W. A., Solar Energy Thermal Processes, John Wiley and Sons, Inc. , 1974.
7. Liu, B. Y. H. and Jordan, R. C., "The Interrelationship and Characteristic Distribution of Direct, Diffuse and Total Solar Radiation," Solar Energy, IV (3), July, 1960.
8. "SIMS Prototype System 1 Test Results - Engineering Analysis," DOE/NASA CR - 150522, prepared by the IBM Corporation, January, 1978.
9. ASHRAE Handbook - 1977 Fundamentals, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc., 1977.

## NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA <br> INCIPIENT FAILURE DETECTION (IFD) OF ANTI-FRICTION BALL BEARINGS

## Prepared By:

Academic Rank:
University and Department:

NASA/MSFC:
(Laboratory)
(Division)
(Branch)
MSFC Counterpart:

Date:
Contract No:

David Y. Wang, Ph.D.
Associate Professor
University of Wisconsin-Plattville Civil Engineering Department

Systems Dynamics Structural Dynamics Environmental Analysis

Luke A. Schutzenhofer Jess H. Jones

August 1, 1980
NGT-01-002-099
(University of Alabama)

# INCIPIENT FAILURE DETECTION (IFD) OF BALL BEARINGS 

by

David Wang, Ph.D.<br>Associate Professor of Civil Engineering University of Wisconsin - Platteville Platteville, Wisconsin

## ABSTRACT

Because of the immense noise dissipated during the operation of a large engine such as the SSME, the unique ball bearing signals are often buried by the overall signature of the machine. As a result, the most commonly used technique, pattern recognition (1)*, to detect and predict ball bearing failures is rendered useless.

In order to retrieve the buried bearing defect signals from the background noises, a survey of two improved techniques was conducted.

Signal averaging method (2) averages $n * *$ signals over a time duration equaling the known period of the ball bearing defect frequency. With this method, the desired ball bearing defect signal will be enhanced over the noise by a factor of $\sqrt{ } \mathrm{n}$.

A second method is to perform a further frequency analysis on the logarithmic spectrum. The purpose of this operation is to detect and separate different families of harmonics and sidebands, thereby giving information about modulating frequencies.

The above mentioned alternate IFD methods were carried out with a Norland 3001 waveform and data system. The obtained results are then compared with the power spectral density (PSD) of the overall machine signals and conclusions are made accordingly.

[^0]
## ACKNOWLEDGEMENTS

The author wishes to acknowledge contributions made to this project by his counterparts at NASA, Dr. L. A. Schutzenhofer and Mr. J. H. Jones. The author also wishes to express his appreciation to those of the Environmental Analysis Branch who provided help in preparing this report.

## NOMENCLATURE

| $f(t)$ | - complex signal |
| :---: | :---: |
| $S(t)$ | - periodic signal |
| $n(t)$ | - noise |
| S/N | - signal to noise ratio |
| $t_{k}$ | - beginning time of the kth repetitive signal |
| $\mathrm{t}_{s}$ | - sample time interval |
| T | - period |
| $\sigma$ | - rms of noise $n(t)$ |
| $f_{0}$ | - ball pass frequency over outer race |
| $f_{i}$ | - ball pass frequency over inner race |
| $f_{s}$ | - ball spin frequency |
| $\mathrm{D}_{\mathrm{w}}$ | - ball diameter |
| $\mathrm{d}_{\mathrm{m}}$ | - bearing pitch diameter |
| $\alpha$ | - contact angle |
| N | - number of rolling elements |
| R | - shaft rotative speed in Hz |

## LIST OF FIGURES

| Figure No. |  | Page |
| :---: | :---: | :---: |
| 1 | Narrow band frequency analysis initial, middle, and final stages | XX-14 |
| 2 | PSD (initial stage) | XX-15 |
| 3 | PSD (middle stage) | XX-16 |
| 4 | PSD (final stage) | XX-17 |
| 5 \& 6 | Zoom (initial stage) | XX-18 \& XX-19 |
| 788 | Zoom (middle stage) | XX-20 \& XX-21 |
| 9 \& 10 | Zoom (final stage) | XX-22 \& XX-23 |
| 11 | Three dimensional PSD | XX-24 |

Every operating machine will generate a combination of sound and vibration in the form of 'signatures' which are unique to the devices being evaluated. The ability to monitor, analyze, and use this signature to locate vibratory sources and predict future failures is the goal of an IFD system. Pattern recognition is the most commonly used technique in evaluating ball bearing conditions. It is performed by spectrum analyzing the output of an accelerometer signal over a specified freqeuncy range that includes the ball pass frequencies. An abnormal increase in amplitude of this 'signature' over some duration of time indicates deterioration of ball bearings.

During the operation of a large machine such as the SSME, the bearing 'signature' in the form of electronic signals is often buried by the large background noise. Therefore, it is necessary to extract the bearing defect signals from the combined signal before an evaluation of the bearings can be made. This report describes a signal averaging or summation technique to extract the repetitive bearing defect signal from a complex signal with a high level background noise.

A cepstrum (3) type of signal analysis was also carried out. The advantage of cepstrum over a spectrum analysis is that it can separate families of periodic sidebands around some fundamental frequencies. The purpose of this cepstrum analysis is to cross check its results with that obtained using the signature averaging method. The possibility of using cepstrum in evaluating ball bearing conditions was also explored.

## OBJECTIVE

The objective of this project is to seek a data reduction process in order to extract pertinent information from the time record of SSME during its tests. The extracted information is, then, used to predict condition of ball bearings installed in the high pressure oxygen turbopump (HPOTP) of the SSME. Of numerous available methods, signal averaging or summation analysis (3), (7), were chosen to be examined in this report.

## SIGNATURE EXTRACTION

A signal with known periodicity, $T$, is buried deep in noise. To extract that signal by averaging or summation, the total signal plus noise is divided into $m$ segments of duration $T$. Each segment is sampled at a number of points during every period. The samples are stored in a data storing device such as a magnetic tape which algebraically adds together all m samples.

A repetitive event of period $T$ will always be added in the same sample location. The resulting sum, thus obtained, will be $m$ times the average amplitude of that event. Events which do not repeat with the same period will result in a sum proportional to $\sqrt{ } \mathrm{m}$, therefore, the desired signal has been enhanced over the noise by a factor of $\sqrt{ } \mathrm{m}$.

The above described signal averaging process can be described as follows. Signal $f(t)$ is

$$
\begin{equation*}
f(t)=S(t)+n(t)---- \tag{1}
\end{equation*}
$$

the signal at kth repetition starts at time instant $t_{k}$, if taking $\mathrm{t}_{1}=0$, is

$$
\begin{align*}
f\left(t_{k}+i t_{s}\right) & =s\left(t_{k}+i t_{s}\right)+n\left(t_{k}+i t_{s}\right) \\
& =s\left(i t_{s}\right)+n\left(t_{k}+i t_{s}\right)--- \tag{2}
\end{align*}
$$

Since noise, $n(t)$, is random and the $m$ segments of samples are independent, summation of the $m$ noises will yield a mean square value of $m \sigma^{2}$ and a rms value of $\sqrt{ } \mathrm{mo}$. Hence, after adding the $m$ segments of the signal, equation (2), which indicates values at the ith location of the signal $f(t)$ within the period $T$ becomes

$$
\begin{aligned}
\sum_{k=1}^{m} f\left(t_{k}+i t_{s}\right) & =m S\left(i t_{s}\right)+\sum_{k=1}^{m} n\left(t_{k}+i t_{s}\right) \\
& =m S\left(i t_{s}\right)+\sqrt{ } m \sigma
\end{aligned}
$$

Therefore, the signal to noise ratio, after $m$ summations, is

$$
\left(\frac{S}{N}\right)_{\mathrm{m}}=\frac{\mathrm{mS}}{\sqrt{\mathrm{~m} \sigma}}=\sqrt{\mathrm{m}}\left(\frac{\mathrm{~S}}{\mathrm{~N}}\right)_{\mathrm{m}}
$$

A Norland 3001 waveform analyzer was used to carry out the signal extraction procedures described above. There are two ways to accomplish signal extraction. One is to stack segments of signal of given period $T$ by shifting the signal segment of duration $T$, without any gap, $m$ times successively. Another way is
XX-8
to convolute the signal, $f(t)$, with a train of $m$ units impulses at a time delay of the given period $T$. The latter procedure is described by the following operation.

$$
\begin{aligned}
a(t) & =\int_{-\infty}^{\infty} f(t-\varepsilon) \sum_{k=1}^{m}(\varepsilon-k T) d \varepsilon \\
& =\sum_{k=1}^{m} f(t-k T)
\end{aligned}
$$

Failures of bearings within the HPOTP were experienced during tests of the SSME. Spallings, damage cuts, pittings, and discoloration appear on both inner and outer races, rolling elements, and cages. The most serious failure seems to occur at bearings closest to the turbine of HPOTP. It was suspected from evidences of damages, that these failures were caused by excessive loading in the axial direction of the turbopump during the starting stage of the test. In order to be able to predict the oncoming bearing as well as other types of failures, data were collected from several strategically located accelerometers. These signals were usually low pass filtered at 10 KHz or under before storing on magnetic tapes for later analysis. Because of the low pass filtering of data, possibility of adopting various IFD techniques in the higher frequency range ( 30 KHz and higher) as diagnostic instruments is precluded.

The collected data in the form of time record is time averaged or summed and then transformed to the ferquency domain as power spectral density (PSD) diagrams. This allows one to relate specific periodic occurrence of the bearing to specific frequencies in the data.

In order to analyze the obtained results, ball pass frequencies over a defect were computed according to the following equations.

$$
\begin{aligned}
& f_{0}=\frac{N}{2} R\left(1+\frac{D_{W}}{d_{m}} \cos \alpha\right) \\
& f_{i}=\frac{N_{2}}{2} R\left(1-\frac{D_{W}}{d_{m}} \cos \alpha\right) \\
& f_{s}=R \frac{d_{m}}{D_{W}}\left(1-\left(\frac{D_{W}}{d_{m}}\right)^{2} \cos ^{2} \alpha\right)
\end{aligned}
$$

For the given bearing configuration, various ball pass frequencies and other relevant data (at $100 \%$ SSME power) are listed as follows:
$f_{0}=2570 \mathrm{~Hz}, f_{i}=3530 \mathrm{~Hz}, f_{\mathrm{s}}=2920 \mathrm{~Hz}$
$\mathrm{N}=13, \quad \mathrm{R}=469 \mathrm{~Hz}$
Static capacity of the bearing $\cong 3050$ lbs.

Results obtained from the test carried out in the Spring of 1980 are presented in figure 1 through figure 11. Figure 1 is a collection of three PSD's for the initial, middle, and final stage of the bearing life span. Narrow band spikes at shaft rotative speed and turbine blade frequencies appear in the diagram at 469 $\mathrm{Hz}, 1836 \mathrm{~Hz}$ and 3672 Hz , as expected. Note the growth of the narrow band spike at 2570 Hz region which indicates a possible fault on the outer race. Defects on the inner race and rolling elements are not obvious in this particular figure.

Figure 5 through figure 10 are the so-called zoom PSD's at the same time instants as those in figures 1 through figure 4. It is more desirable thatn the normal PSD's because it provides a considerably finer resolution over the portion of the spectrum that we are interested in. Again, possible defects on the outer race appear on these spectra in the neighborhood of 2600 Hz region. Finally, a useful technique in detecting the deterioration of bearings with respect to time is presented in the form of a three dimensional PSD with time as the third parameter.

## CONCLUSIONS AND RECOMMENDATIONS

A signal averaging or summation process has been described that is used ta extract the bearing defect signals form the time record of SSME. It appears that this procedure can also be applied to the zoom and three-dimensional PSD for finer resolution and better presnetation of the deterioration history of ball bearings. An electronic equipment of black box type based on this principle could be constructed and adopted as part of an "on-line" IFD system. One drawback in applying this technique is that the period, $T$, of a repetitive signal cannot be absolutely time-locked in an analog to digital converter because of the limitation of the sampling time interval. The result of this deficiency is the tendency of the summed time signal being smeared gradually. Signal to noise gain will be decreased accordingly.

Even though other techniques such as "acoustic emission" (10) and "shock pulse" have been proposed, they have not been proven successful. One alternative method in ball bearing defects detection is to utilize the bearing high frequencies (13), (14), especially for large machines. Natural frequencies of ball bearings are usually far beyond that of the noise level, therefore, less noise interference to the bearing signal. In order to adopt this method of approach, real time signal should be collected without any low pass filtering.

## REFERENCES

1. Taylor, J. I., "Determination of Anti-Friction Bearing Conditions by Spectral Analysis," Progress of Machinery Vibration Monitoring and Analysis Seminar, February 1977.
2. Randall, R. B., "Cepstrum Analysis and Gear Box Fault Diagnosis," BED Application Note No. 13-150.
3. Weichbrodt, B., "Mechanical Signature Analysis, A New Tool for Product Assurance and Early Fault Detection," Annals of Reliability and Maintainability," Vol. 5, AIAA, 1966.
4. Hower, D. R., "Signal Analysis Techniques for Vibration Diagnostics," 22nd Meeting of Mechanical Failure Prevention Group, October 1976.
5. Harris, T. A., "Rolling Bearing Analysis," Wiley, New York, 1966.
6. Bickel, H. J., "Real Time Spectrum Analysis," Sound and Vibration, Sol. 5, No. 3, March 1971.
7. Trimble, C.R., "What is•signal Averaging?" Hewlett_Packard Journal, April 1968.
8. Fieldhouse, K., "Techniques for identifying Sources of Noise," Sound and Vibration, Vol. 4, No. 12, 1970.
9. Bannister, R. L., and Donato, V., "Signature Analysis of Turbomachinery," Sound and Vibration, September 1971.
10. Balderston, H. L., "The Detection of Incipient Failure in Bearings," Materials Evaluation, June 1969.
11. Braun, S., "Signal Analysis for Rotating Machinery Vibrations," Pattern Recognition, Vol. 7, 1975.
12. Bendat, J. S., and Piersol, A. G., "Random Data Analysis and Measurement Procedures," Wiley-Interscience, New York, 1958.
13. Harting, D. R., "Incipient Failure Detection by Demodulated Resonance Analysis," Instrumentation Technology, September 1977.
14. Burchill, R. F., "Resonant Structure Techniques for Bearing Fault Analysis," 18th Meeting of Mechanical Failures Prevention Group, November 1971.
15. Randall, B., "Frequency Analysis," Bruel \& Kjar, September 1977, 2nd edition.


NARROW BAND FREQUENCY ANALYSIS - INITIAL, MIDDLE, AND FINAL STAGES
Figure 1
196 HPOT RAD 135 ACC S+ RJN

PSD (INITIAL STAGE)


PSD (MIDDLE STAGE;
193 HFOT RAD $: 35 \mathrm{ACC} \mathrm{S}+220$

PSD (FINAL STAGE)


ZOOM (INITIAL STAGE)




(ヨコVLS TVRIIJ) HOOZ

（GOVIS TVHIき）W！OOZ


# NASA/ASEE FACULTY PESEARCH FEIJOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER <br> THE UNIVERSITY OF ALABAMA <br> PUJSED EDDY CURRENT TESTING 

## Prepared By:

## Academic Rank:

University and Department:

NASA/MSFC:
(Laboratory)
(Division)
(Branch)
MSFC Counterpart:
Date:
Contract No.:

Gary L. Workman, Ph.D.
Pssociate Professor
Athens State College Department of Scientific Technologies

Matcrials and Processes
Engineering Physics
Nondestructive Testing
J. M. Knadler, III

August 1, 1980
NGT-01-002-099
(University of Alabama)

Gary L. Workman, PH.D. Associate Professor of Scientific Technologies Athens State College Athens, Alabama 35611

## ABSTRACT

The traditional nondestructive testing technique used for inspecting welds in large space structures, such as the "External Tank", has been radiographic testing, in conjunction with surface inspection methods, such as dye penetrant testing. In order to improve upon the economics and reliability of these critical inspections, new approaches to the test methods are under consideration. One approach is to replace the very expensive radiographic inspection with an automated eddy current testing technique. A description of pulsed eddy current and its relationship to multifrequency techniques if presented here, as well as a some preliminary results obtained from observing pulsed waveforms with apparatus and algorithms currently in use for ultrasonic testing of welds. It can be shown the pulsed eddy current techniques can provide similar results.

Eddy current testing has primarily been used for sorting of materials due to differences in conductivities, detecting differences in surface due to corrosion on cracks, and in some instances to detect flaws below the surface. Manual operation of the test proves to be very unreliable due to such test parameters as lift off, non-uniform conductivity, depth of penetration, surface treatment, and complex geometries of the test object. The use of a pulsed eddy current techniques can eliminate some of the noncritical paramcters affecting the eddy current signals and facilitate in the detection of critical paraneter such as flaws, subsurface voids, and corrosion.

## Acknowledgements

I wish to acknowledge Mr. John M. Knalder, III, for allowing me to conduct the research this summer and Mr. Ken Woodis for assisting me in setting up the computer interfaces. Also Mr. Ben Bankston and Mr. Claude Williamon, for implementing experimental fixtures. I also thank Miss Pearline Stone for typing the report.

## Introduction

The nondestructive inspection of large space structures requires that a large part of the inspection procedure be automated in order to reduce human errors and to reduce the amount of time and money spent in this critical phase of production. Since a large number of the procedures used for inspecting the External Tank are concerned with determining flaws in welds, there is a need to develop an inspection technique, which can be automated, to determine flaws in welds and structures with complex geometries. The techniques; whereby one generates eddy current in a metallic material and observes the changes in the circuit parameters due to material differences, has been chosen as one possible approach. The principles of eddy current testing has been treated in a number of publications (1-3). Of particular interest to the researcher is the book by Libby ${ }^{l}$ and the work of Dodd $^{3}$.

Eddy currents are generated in metallic specimens by alternating currents in coils in close proximity to the material. Some of the properties associated with eddy currents are:

1. They constitute a closed circuit; hence, in the material, they usually have a circular configuration
2. They penetrate into the material with decreasing
amplitude away from the surface.
3. Variations in part distances introduces electrical phase differences.
4. The magnitude of eddy current flow is proportional to the current in the test coil.
5. Both the electrical conductivity and the magnetic permeasbility affect the eddy current flow in the material.
6. Discontinuities which affect the eddy current flow are easily detected as the eddy current redistribute themselves around the discontinuity.

These general properties have enabled eddy current techniques to be useful for a number of applications. Some of the most notable are:

1. Measurements of electrical conductivity
2. Measurements of magnetic permeability
3. Sorting of metallic materials
4. Measurements of distances between components
5. Thickness determinations
6. Detections of voids on discontinuities in metallic materials

A number of commerical eddy current instruments are available for these applications. The reliability and sensitivity of each test depends upon many parameters, for instance:

1. Coil-to-conductor spacing and geometry
2. Choice of frequency for a particular conductor
3. Orientation of defects to coil
4. Sample geometry
5. Coil characteristics

Such a large number of parameters affecting an eddy current measurement accounts for the range and flexibility of the technique, as well as its pitfalls. For instance, in the determination of flaws in welds, one has to keep all other parameters constant so that only the interaction of the eddy currents with a flaw would be observed as a signal. Even with a judicious laboratory arrangement, this is very difficult to accomplish. When testing complex geometries, in a production environment, it is practically impossible to maintain that many parameters constant. Also, most materials exhibit minute changes in composition which can affect the eddy current measurements and this can be misinterpreted as a flaw. Since it can be so difficult to make a single eddy current measurement and interpret the results correctly, many researchers in the field are developing methods which take a number of measurements and solve simultaneously for different parameters which can change. The use of computers in both hardware control and data acquisition and analysis allows these approaches to become easily automated.
XXI-3

## Objectives

The primary objectives of this work is to determine the feasibility of using pulsed eddy current methods for nondestructive testing. As described in the report there are many ways of handling the data received from the eddy current test. Our approach has been to pulse an eddy current probe, use present FFT data processing and determine if the information contained in the frequency domain does allow one to solve for the properties of the material under inspection.

## Multiparameter Methods

The electrical quanities of interest in an eddy current measurement are usually a voltage or current and its phase relationship to the driving voltage or current. Eddy current circuits are most easily represented as simple RLC circuits and their signal responses represented as impedance of the material being inspected. Hence impedance plane plots are most often used to illustrate and differentiate between the different parameters affecting the eddy current measurement. For example, if we consider a simple coil and driver arrangement such as

then if $V_{0}=V_{0} \cos w t$ (i.e. a sinusoidal waveform) then

$$
\begin{aligned}
V & =V_{0}\left(\frac{z}{R+z}\right)=V_{0}\left|\frac{z}{R+z}\right| \\
& =V_{0} \sqrt{\frac{\left(R_{2}+2 \pi f L_{1}\right)^{2}}{\left(R_{1}+R_{2}+2 n f L_{1}\right)^{2}}}
\end{aligned}
$$

Notice that the measurement of a single voltage can determine only a single parameter or the combined effect of several parameters. In order to differentiate between the
parameters in an eddy current measurement, one has to synthesize an excitation signal which then can be decomposed into frequencies affected by the desired parameters. One approach is to use multifrequency excitation from which we can extract a voltage and a phase value for each frequency. For example, if we use two frequencies, then we have four measurements for each test and
$\mathrm{V}_{1}=\mathrm{P}_{11} \mathrm{~A}_{1}+\mathrm{P}_{12} \mathrm{~A}_{2}+\mathrm{P}_{13} \mathrm{~A}_{3}+\mathrm{P}_{14} \mathrm{~A}_{4}$
$\mathrm{~V}_{2}=\mathrm{P}_{21} \mathrm{~A}_{1}+\mathrm{P}_{22} \mathrm{~A}_{2}+\mathrm{P}_{23} \mathrm{~A}_{3}+\mathrm{P}_{24} \mathrm{~A}_{4}$
$\mathrm{~V}_{3}=\mathrm{P}_{31}{ }^{\mathrm{A}_{1}}+\mathrm{P}_{32} \mathrm{~A}_{2}+\mathrm{P}_{33} \mathrm{~A}_{3}+\mathrm{P}_{34} \mathrm{~A}_{4}$
$\mathrm{~V}_{4}=\mathrm{P}_{41} \mathrm{~A}_{1}+\mathrm{P}_{42} \mathrm{~A}_{2}+\mathrm{P}_{43} \mathrm{~A}_{3}+\mathrm{P}_{44} \mathrm{~A}_{4}$
where $A_{1}, A_{2}, A_{3}, A_{4}$, represent the material properties as coefficients in response to the eddy current source and each Pij represents the magnitude of that effect in each measurement. $\quad \mathrm{V}_{1}$ through $\mathrm{V}_{4}$ represents either a voltage or phase measurement. This procedure has been developed and successfully demonstrated by Dr. Dodd and his Union Carbide Corp. in Oak Ridge $(4,5)$. Although the parameters $A_{1}$ through $A_{4}$ appear to be linearly independent, experimentally it can often be difficult to determine unique solutions and thus to identify such effects as lift-off, voids, or corrosion. However, their approach was to synthesize known types of defects, liftoff's, etc into the laboratory calibration. Then, during an inspection procedure, their system uses a pattern recognition approach to go back and fit the experimental data
to the calibration data in order to determine what responses are contained in each measurement. To date, this approach has worked very well in the inspection of tubing used in nuclear reactor heat exchangers.

The frequencies chosen for this approach use a high frequency signal for surface effects and a lower frequency (say $1 / 10$ ) for multilayer identification, and even another lower frequency (say $1 / 100$ ) for detection of voids farther away from the surface.

## Pulsed Eddy Current Techniques

The multifrequency technique provides a workable solution to the eddy current testing problem. The major drawback to the method is that it is expensive. A multifrequency head using several coils will prove to be extremely costly. In addition to the data acquisition circuit, precise frequency filters are required to maintain a reliable signal-to-noise ratio for each phase and amplitude measurement. A promising technique, which should prove to be less expensive, is to use a pulsed excitation rather than the continuous alternating current as in the previous work. In reality pulsed eddy current and multifrequency eddy current provide similar information. For instance, a short duration pulse can be transformed into the frequency domain and the resulting wave form consists of a large number of frequencies. Hence, we can also produce multifrequency excitation using a pulse with a finite time period.

Pulsed eddy current data can be obtained in two different ways, although the two approaches are essentially equivalent. The pulsed eddy current experiment can be performed in a manner similar to the multifrequency experiment discussed above, the only deviation is using a finite pulse instead of several continuous wave frequencies. The interaction of the coil with the eddy currents is still dependent on the inherent frequency of interaction. Hence, one would still measure amplitudes and phases of the resulting signals and perform the data reduction in exactly the same way as before. The two methods are equivalent in that same sense.

The other approach one might take in using pulsed eddy currents is to actually transform the time domain data into the frequency domain and use the multifrequency analytical approach with these transformed values. Our approach this summer has been to obtain data in this mode in order to determine the feasibility of the approach. A number of algorithms currently in use for ultrasonic imaging applications one applicable, as well as a hardwired Fast Fourier Transform Analyzer (6). Hence we felt that this approach might proceed the quickest for a summer's work.

## Experimental Results

All of the measurements taken in this work were obtained with a Microtek model 101 ultrasonic pulsing system driving a Kervonics reflection eddy current coil. The pick-up coil signals were digitized by a 8-bit $A / D$ and stored in a Biomation Transient Analyzer.

Model \#8100. The data was then transferred to a PDP11/45 computer and analyzed via software generated by contractor personnel. An example of the data obtained here is shown in Figure 1-4. The pick-up signal itself is shown in Figure 1 and its FFT is shown in Figure 2. These observations were obtained statically with a Kervonics 40A probe (pick-up coil radius $=0.7 \mathrm{~cm}$ ) on aluminum.

Lift-off simulation is shown in Figure 3 and the resulting waveform when over a 20 mil surface defect in shown in Figure 4.

Note the changes which occur for the frequency domain data, particularly at 0.5, at 1.3 , and at 3 MHZ . In all three cases the pick-up signal changes somewhat but not as quantitative as given in the frequency data. Obviously then the analysis of a three frequency probe at $0.5,1.3$, and 3 MHZ should provide the same type of information, but would require three frequencies with separate pick-up coils. Since the scanning mechanism, which was expected to be used in this work, never did arrive; all the measurements were made using a specially inprovised mount for the probe. Since no motion occurred during the measurement, all tests were static.

## Conclusions and Recommendations

The pulsed eddy current technique can differentiate between various materials properties using FFT analysis. Since we were not able to obtain data while scanning over a specimin, that should be the next objective in this project. Also, in view of the information published by Dr. Dodd and his associates, the following recommendations can be made.

1. Improve the resolution of the $A / D$ converter presently being used by replacing it with at least a l2-bit $A / D$ converter.
2. Use waveform synthesis techniques to generate a pulse which contains those frequencies of interest for a particular inspection. A microprocessor generated pulse can do this.
3. Implement the pattern recognition algorithms at Marshall Space Flight Center. Upon request most of his software is available.

If pulsed eddy current is to be developed at MSFC, a full time effort should be provided. 'he nondestructive testing requirements for large spacecraft will be more casily implemented using either a pulsed or multi-frequency method.
-01 ESEC.


- 01 USEC.


1
$L$
$L$

山
ロ
$\square$
1
$I$
$z$
0
$๔$
5


Figure 3: FFT Signal for Lift-Off

$$
M A G N I T \text { U } \quad \text { I } E \quad F F T
$$



## REFERENCES

1. Libby, H. L., Introduction to Electromagnetic Nondestructive Test Methods, Krieger Pub., Huntington, N.Y., 1979.
2. "Eddy Current Inspection", Nondestructive Inspection and Quality Control, vol. 11 of Metals Handbook, ASM. Metals Park, ohio, 1976.
3. Dodd, C. V., Basic Principles of Eddy Current Testing, Internal document, Union Carbide Corp., Oak Ridee, TN, 1978.
4. Deeds, W. E., Dodd, C. V., Scott, G. W.; "Computer-Aided Design of Multifrequency Eddy Current Tests for Layered Conductors with Multiple Property Variations", ORNL/TM-6858, 1979.
5. Dodd, C. V. and Deeds, W. E., "Multiple Property Variations in Coaxial Cylindrical Conductors Determined with MultipleFrequency Eddy Currents".
6. Smith, G.R., "Interactive Signal Analysis and Ultrasonic Data Collection System"; Final Report, Contract MAS 8-32024.

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALBAMA 

## STABILITY OF THE STRATIFIED CYLINDRICAL ANNULUS FLOW

| Prepared By: | Basil N. Antar, Ph.D. |
| :--- | :--- |
| Academic Rank: | Associate Professor |
| University and Department: | University of Tennessee Space <br> Institute <br> Department of Engineering <br> Science and Mechanics |
| NASA/MSFC: <br> $\quad$ (Laboratory) <br> $\quad$ (Branch) | Space Sciences <br> Atmospheric Sciences |
| MSFC Counterpart: | F1uid Dynamics |
| Date: | W. W. Fowlis |
| Contract No.: | August 1,1980 |

## Prepared By:

Academic Rank:
University and Department:

NASA/MSFC:
(Laboratory)
(Division)
(Branch)
MSFC Counterpart:
Date:
Contract No.:

Basil N. Antar, Ph.D.
Associate Professor
University of Tennessee Space Institute
Department of Engineering Science and Mechanics

Space Sciences
Atmospheric Sciences
F1uid Dynamics
W. W. Fowlis

August 1, 1980
NAT-01-002-099
(University of Alabama)

Basil N. Antar Assistant Professor of Engineering

Science and Mechanics The University of Tennessee Space Institute Tullahoma, Tennessee 37388


#### Abstract

The linear stability analysis for the stratified flow between two rotating circular cylinders is formulated. Two approaches for the stability analysis are presented. The first approach results in an algebraic eigenvalue problem, while the second results in an initial value problem for the perturbation function. The advantages and disadvantages of both approaches are discussed and a preferable numerical solution technique is outlined.


# Basil N. Antar Assistant Professor of Engineering Science and Mechanics The University of Tennessee Space Institute Tullahoma, Tennessee 37388 

## ABSTRACT

The linear stability analysis for the stratified flow between two rotating circular cylinders is formulated. Two approaches for the stability analysis are presented. The first approach results in an algebraic eigenvalue problem, while the second results in an initial value problem for the perturbation function. The advantages and disadvantages of both approaches are discussed and a preferable numerical solution technique is outlined.

## INTRODUCTION

It is generally believed that the existence of flucuations in the circulation of the atmosphere can be attributed to the instability of the stationary dynamical state without fluctuations. Partial confirmation of this belief has recently come from detailed analysis of the numerical general circulation models. However, no exact laboratory experiments have been conducted of the atmospheric circulation in order to test this hypothesis. The absence of such a laboratory model is basically due to the difficulty arising from the large earth gravity field which is ever present in any terrestrial experiment. Recently hopes for the possibility of setting up an experiment model of the earth's global circulation has been rekindled again through the availability of the gravity free environment of the forthcoming missions of Spacelab. Indeed an experiment to model and study atmospheric fluctuations arising from stability considerations in a true spherical geometry is being considered by NASA for a future Spacelab Mission [1].

The existence of fluctuations in the flow inside a rotating cylindrical annulus in which the inner cylinder is cooled while the outer cylinder is heatcd have been observed experimentally by Fultz et al. [2] and Fowlis and Hide [3]. These fluctuations were found to resemble in a remarkable way those that are observed in the atmosphere although the geometries are different. Due to this striking resemblence between the waves in the annulus experiments and the waves in the atmosphere, the flow ficld in the rotating cylindrical annulus has been studied extensively (see Hide and Mason [4] for an up to date discussion on the annulus problem).

One of the major findings common to both the analytical and experimental studies of the annulus problem is that waves of different azimuthal wave lengths appear only for specific value of the governing parameters of the problem. This fact is usually shown by plotting a curve in the Taylor-Froude number plane which separates the wave regime from the no-wave (symmetric) regime. Such a curve has been obtaincd experimentally by Fowlis and Hide [3]. However, in the theoretical treatment of the exact annulus problem only few points on that curve have been obtained, originally by Williams [5] and subsequently by Quon [6]. These theoretical confirmations
have been obtained through the complete, three-dimensional, numerical solution of the governing equations for the annulus geometry. These numerical solutions of the exact problem consume excessive amounts of computer time to perform. For this reason the regime diagram which was obtained experimentally have not been reproduced in full through numerical solutions, since the amounts of computer time needed to perform these calculations are beyond the scope of present day computers.

In this report we will outline a different numerical method for obtaining the regime diagrams for the annulus problem. Since the existence of the observed fluctuations in the annulus is due to instability of the basic dynamical state, it then seems reasonable that a stability analysis of the basic state, symmetric annulus flow should predict under what conditions the fluctuations will appear. Now, it is well known that stability analysis requires far less calculations than the total solution of the governing system of equations. Hence, such an analysis may be efficiently used to generate the desired regime diagrams. In this report two methods of stability analysis will be presented and the advantages and disadvantages of each method will be discussed in detail. Furthermore, a recommendation will be given on the best numerical implementation of the techniques in question.

## STABILITY ANALYSIS

As indicated in the introduction there are basically two analytical approaches that may be used to verify the experimental regime diagrams. The first approach is to integrate the initial-boundary value problem starting from a state of no motion and following the solution in time. This approach requires the soltuion of the full governing equation in three-dimensions and for long enough time to discern whether fluctuations will appear or not. However, the computation time that may be necessary to implement this approach could be prohibitive. Quon [6] has indicated that as much as 150 hours of computer time on the Univac 1108 is necessary to establish only one point in the regime diagram using this approach.

The alternative approach which is subsequently outlined is to use linear stability analysis to establish the regime diagrams. This approach may be implemented in one of two ways both of which will be discussed at length.

The linear stability analysis starts by first formulating the governing equation for the annulus flow in cylindrical geometry. These equations are the Navier-Stokes equations, the energy equation and the mass conservation equation. The Navier-Stokes equations in cylindrical geometry in a rotating reference frame, are the following:

$$
\begin{align*}
& u_{t}+u u_{r}+\frac{v}{r} u_{\varphi}+w u_{z}-\frac{v^{2}}{r}-2 \Omega v=-\frac{1}{\rho} P_{r} \\
&+v\left(\nabla^{2} u-\frac{2}{r^{2}} v_{\varphi}-\frac{u}{r^{2}}\right)  \tag{1}\\
& v_{t}+u v_{r}+\frac{v}{r} v_{\varphi}+w v_{z}+\frac{u v}{r}+2 \Omega u=-\frac{1}{\rho} P_{\varphi}  \tag{2}\\
&+v\left(D^{2} v+\frac{2}{r^{2}} u_{\varphi}-\frac{v}{r^{2}}\right) \\
& w_{t}+u w_{r}+\frac{v}{r} w_{\varphi}+w w_{z}=-\frac{1}{\rho} p_{z}+v D_{w}^{2}+\alpha g T \tag{3}
\end{align*}
$$

where

$$
\nabla^{2} f=f_{r_{r}}+\frac{1}{r_{e}} f_{q \phi}+f_{z z}+\frac{1}{r} f_{r}
$$

and

$$
f_{a}=\partial f / \partial a
$$

In Eq. (3) the Boussinesq approximation has been used. In the above equations $u, v$ and $w$ are the radial ( $r$ ), aximuthal ( $\phi$ ) and axial ( $z$ ) velocity components, respectively. $p$ and $\rho$ are the pressure and density; $\nu$ is the kinematic viscosity; $\alpha$ is the volumetric expansion coefficient and $T$ is the temperature. $\Omega$ is the angular velocity of both cylinders. In conjunction with above system the energy and mass conservation equations are needed which are given by

$$
\begin{align*}
& T_{t}+u T_{r}+\frac{v}{r} T_{\phi}+w T_{z}=k \nabla^{2} T  \tag{4}\\
& u_{r}+\frac{1}{r} v_{\phi}+w w_{z}+\frac{u}{r}=0 \tag{5}
\end{align*}
$$

where $\kappa$ is the thermal diffusivity of the fluid.

Next, all the field variables are separated into a stationary mean part and a perturbation part in the following way:

$$
\begin{aligned}
& u=u_{0}+u^{\prime} \\
& v=v_{0}+v^{\prime} \\
& w=w_{0}+w^{\prime} \\
& p=p_{0}+p^{\prime} \\
& T=T_{0}+T^{\prime}
\end{aligned}
$$

When the above variables are substituted into the governming system of equations (1)-(5) and using the fact that the mean variables ( $f_{0}$ ) themselves satisfy equations (1)-(5) a system of equations governing the perturbation functions results. However, this latter system of equations is as complicated as the original system with the added complexity of having the basic state functions appearing as coefficients in the equations. This means that in order to solve for the perturbation functions one needs to solve for the basic state first.

If it is further assumed that the perturbation functions are very small, ie. infinitesimal then it is possible to linearize the system of equation for the perturbations. These equations will be of the following form after dropping the primes:

$$
\begin{align*}
& u_{t}+u_{0} u_{r}+\frac{v_{0} u_{p}+w_{0} u_{z}}{}+u u_{0_{r}}+w u_{0}-\frac{2 v_{0} v}{r}-2 \Omega v  \tag{6}\\
&=-\frac{1}{\rho} p_{r}+v\left(\nabla^{2} u-\frac{2}{r_{2}} v_{\varphi}-\frac{u}{r^{2}}\right) \\
& \begin{aligned}
v_{t}+u_{0} v_{r}+\frac{v_{0}}{r} v_{\varphi}+w_{0} w_{z} & +u v_{0_{r}}+w v_{0}+\frac{1}{r}\left(u_{0} v+u v_{0}\right)+2 \Omega u \\
& =-\frac{1}{p_{r}} p_{\varphi}+v\left(\nabla^{2} v+\frac{2}{r^{2}} u_{\varphi}-\frac{v}{r^{2}}\right)
\end{aligned} \\
& w_{t}+u_{0} w_{r}+\frac{v_{0}}{r} w_{\varphi}+w_{0} w_{t}+u v_{0_{r}}+w w_{o_{z}}=-\frac{1}{\rho} p_{z}+v \nabla^{2} w+\alpha g T \tag{7}
\end{align*}
$$

$$
\begin{align*}
& T_{t}+u_{0} T_{r}+\frac{v_{0}}{r} T_{\varphi}+w_{0} T_{t}+u T_{o_{r}}+w T_{o_{z}}=k D^{2} T  \tag{9}\\
& u_{r}+\frac{u}{r}+\frac{1}{r} v_{4}+w_{z}=0 \tag{10}
\end{align*}
$$

In the above equations the stationary basic state variables are functions of $r$ and $z$ alone are assumed to be known everywhere.

EIGENVALUE PROBLEM
We continue with the stability analysis by looking for solutions to Eqs. (6)-(10) for the perturbation functions with the known stationary basic state for some known initial small values for the perturbations. The solution will proceed by first formulating the eigenvalue problem. Eqs. (6)-(10) can be simplified by noting that they are linear and the coefficients are independent of the azimuthal direction $\phi$. Under these conditions the perturbation functions admit the following form of solution:

$$
\begin{gather*}
{[u, v, w, \pi, p]=[\hat{u}(r, z), \hat{v}(r, z), \hat{w}(r, z), \theta(r, z), \pi(r, z)] \times}  \tag{11}\\
\exp [i k(\varphi-c t)]
\end{gather*}
$$

The solution form (11) represents a wave of wavelength $2 \pi / k$ which is propagating in the azimuthal direction $\phi$ with a speed c. Upon substituting the functional form (11) into the perturbation equations (6)-(10), the following equations result

$$
\begin{align*}
-i k c \hat{u}+u_{0} \hat{u}_{r}+\frac{v_{0}}{r} i k \hat{u}+w_{0} \hat{u}_{z} & +\hat{u} u_{o_{r}}+\hat{w} u_{0} z-2 \frac{v_{0} \hat{u}}{r}-2 \Omega \hat{v} \\
& =-\frac{1}{\rho} \hat{p}_{r}+v\left(\nabla^{2} \hat{u}-\frac{2 k}{r^{2}} \hat{v}-\frac{\hat{u}}{r^{2}}\right) \\
-i k c \hat{v}+u_{0} \hat{v}_{r}+\frac{h_{1}}{r} v_{0} \hat{v}+w_{0} \hat{v}_{z} & +\hat{u} v_{0_{r}}+\hat{w} v_{0}+\frac{1}{r}\left(u_{0} \hat{v}+\hat{u} v_{0}\right)+2 \Omega \hat{u}  \tag{13}\\
& =-\frac{i k}{\rho_{r}} \hat{p}+v\left(\nabla^{2} \hat{v}+i \frac{k}{} \hat{r^{2}}-\frac{\hat{v}^{2}}{r^{2}}\right)
\end{align*}
$$

$$
\begin{align*}
&-i k c \hat{w}+u_{0} \hat{w}_{r}+\frac{i k_{r} v_{0} \hat{w}+w_{0} \hat{w}_{z}+\hat{u} w_{o_{r}}+\hat{w} w_{0 z}=}{}-\frac{1}{\rho} \hat{p}_{z}  \tag{14}\\
&+v \nabla^{2} \hat{w}+\alpha g \theta \\
&-i k c \theta+u_{0} \theta_{r}+i \frac{k v_{0}}{r} \theta+w_{0} \theta_{z}+\hat{u} T_{o_{r}}+\hat{w} T_{o_{z}}=k \nabla^{2} \theta  \tag{15}\\
& \hat{u}_{r}+\frac{\hat{u}}{r}+\frac{i k \hat{v}^{r}}{r}+\hat{w}_{z}=0 \tag{16}
\end{align*}
$$

In the present work the perturbation functions will be zero at the boundaries.

Now Eqs. (12)-(16) may be written in the following operator form:

$$
\mathscr{L}\left\{\begin{array}{c}
\hat{u}  \tag{17}\\
\hat{v} \\
\hat{w} \\
\hat{p} \\
\hat{\theta}
\end{array}\right\}=c \mathcal{L}_{2}\left\{\begin{array}{c}
\hat{u} \\
\hat{v} \\
\hat{w} \\
\hat{p} \\
\theta
\end{array}\right\}
$$

where $\mathscr{L}_{1}$ and $\mathscr{L}_{2}$ are linear partial differential operators which involve the basic state $u_{0}, v_{0}, w_{0}, T$ and $k$ Since Eq. (17) is homogeneous with Romofene8us boundary condiction then it forms an eigenvalue problem for the eigenvalue c.

Normally the system (17) is solved through the method of separation of variables which will reduce the eigenvalue problem to an ordinary differential eigenvalue problem. Such a problem may be solved either by the shooting method or the matrix method (see Antiar [7] for a discussion on the application of these techniques). However, in the present case we can not use the method of separation of variables since the coefficients in the operators are not separable. Thus, the partial differential eigenvalue problem (17) has to be solved as it stands.

A logical way to handle the system (17) is to use the finite difference method especially since the coefficients are given as a set of numbers at specific points in the domain. This is due to the fact that the basic state solution is obtained through a finite difference method. Without going too much into the details of the finite differencing procedure, if the derivatives in the operators $d_{1}$ and $\mathscr{L}_{2}$ are
approximated by finite differences, the equation (17) will be transformed to the following alegebraic problem

$$
\begin{equation*}
[A]\{\bar{x}\}+c[B]\{\bar{x}\}=0 \tag{18}
\end{equation*}
$$

where [A] and [B] are banded matrices and is a vector representing the value of the field variables at the discrete points in the solution domain. Taking the inverse of [B] in (18), the equation may be rewritten as

$$
\begin{equation*}
\{[D]-c[I]\}\{x\}=0 \tag{19}
\end{equation*}
$$

where $[\Delta]=[B]^{-1}[A]$. Clearly (19) then is an algebraic eigenvalue problem which may be solved through matrix manipulations.

Although in principal the technique is straight forward, in reality the transformation of Eq. (17) into Eq. (18) is quite involved and will not be discussed here. The major difficulty in formulating Eq. (19) is that the vector $\bar{x}$ is very large. In fact if one chooses to discretize Eq. (17) through 10 points in each of the $z$ and $r$ directions the vector $x$ will have 400 elements. However, we know from experience that 10 points for the discretization is not enough and a reasonable choice will be 50 points. For this latter value the vector $x$ will possess $10^{4}$ elements. This size of the vector will make the solution of the eigenvalue problem very lengthy. Another approach which might be more suitable for the present problem is discussed next.

## INITIAL VALUE PROBLEM

The second alternative for obtaining a solution to the perturbation equations (6)-(10) is to assume a solution which possesses a sinusoidal form in the azimuthal direction, $\phi$, ie.

$$
\begin{equation*}
f(r, z, \varphi, t)=\hat{f}(r, z, t ; k) \sin k \varphi \tag{20}
\end{equation*}
$$

This form of the solution is admissible due to the linearity of the equations and also the independence of the coefficient function of the azimuthal direction $\phi$. Upon substituting the functional form (20) into eqs. (6)-(10), the following linear system of equations results:

$$
\begin{align*}
& \hat{u}_{t}+u_{0} \hat{u}_{r}+i k \frac{v_{0}}{r} \hat{u}+w_{0} \hat{u}_{z}+\hat{u} u_{0 r}+\hat{w} u_{0}-\frac{2 v_{0}}{r} \hat{v}-2 \Omega \hat{v} \\
& =-\frac{1}{\rho} \hat{p}_{r}+v\left(D^{2} \hat{u}-2 \cdot \frac{k_{k}}{r_{2}}-\frac{\hat{u}_{2}}{r^{2}}\right)  \tag{21}\\
& \hat{v}_{t}+u_{0} \hat{v}_{r}+i k_{0} v_{0} \hat{v}+w_{0} \hat{u}_{z}+\hat{u} v_{0 r_{r}}+\hat{w} v_{0 z}+\frac{1}{r}\left(u_{0} \hat{v}+\hat{u} v_{0}\right)+2 \Omega \hat{u} \\
& =-\frac{i k}{\rho r} \hat{b}+v\left(D^{i} \hat{v}+\frac{i, k}{r^{2}} \hat{u}-\frac{\hat{v}}{r^{2}}\right)  \tag{22}\\
& \hat{w}_{t}+u_{0} \hat{w}_{r}+\frac{i k v}{r} \hat{w}^{r}+w_{0} \hat{w}_{z}+\hat{u} w_{o r}+\hat{w} w_{0}=-\frac{1}{8} \hat{p}_{z}+\nu D \hat{w}^{n}+\alpha g \theta  \tag{23}\\
& \theta_{t}+u_{0} \theta_{r}+\frac{i k}{r} v_{0} \theta+w_{0} \theta_{z}+\hat{u} T_{0_{r}}+\hat{w} T_{0 z}=k \nabla^{2} \theta \tag{24}
\end{align*}
$$

The set (21)-(24) constitute a linear partial differential set depicting an initial value problem. Thus a solution to the system may be obtained once the proper initial conditions are given. The initial conditions in the is case are given again as a wave function with a specified wavenumber $k$ and a specified amplitude function.

The stability or instability of the perturbations is resolved by examining the behavior of the amplitude functions
with time. If for a given wavenumber $k$, the amplitudes appear to grow with time, then the perturbations are said to be unstable. If, on the other hand, the perturbation amplitudes seem to decay with time then the pertubations are said to be stable. As for the eigenvalue problem, however, the sign of the imaginary part of $c$ determines stability or instability. In this case if the sign of $c_{i}$ is negative, then the perturbations are said to be stable, otherwise it is unstable.

Now, due to the complexity of the coefficients of the system (21)-(24), no closed form solution is possible. Thus one has to resort to numerical solutions. In this case it seems logical to use finite difference techniques for both the time and the two spatial directions. Conceptually, a numerical solution for this system should be obtained without difficulty due to the linearity of the system.

## REFERENCES

1. Fowlis, W. W. and G. H. Fichtl, "Geophysical Fluid Flow Model Experiments in Spherical Geometry." Proc. Third NASA Weather and C1imate Program Science Review, NASA Conf. Pub1. 2029, Pap. No. 32, 1977.
2. Fultz, D., R. R. Long, G. V. Owens, W. Bohan, R. Kaylor and J. Weil, "Studies of Thermal Convection in a Rotating Cylinder with Some Implications for LargeScale Atmospheric Modeling," Meteor. Monogr., No. 21, 1959.
3. Fowlis, W. W. and R. Hide, "Thermal Convection in a Rotating Annulus of Liquid: Effects of Viscosity on the Transition Between Axisymmetric and Non-Axisymmetric Flow Regimes," J. Atmos. Sci., Vol. 22. 1965.
4. Hide, R. and P. J. Mason, "Sloping Convection in a Rotating Fluid," Advances in Physics, Vo1. 24, 1975.
5. Williams, G., "Baroclinic Annulus Waves," J. Fluid Mech. Vol. 49, 1971.
6. Quon, C., "A Mixed Spectral and Finite Difference Model to Study Baroclinic Annulus Waves," J. Comp. Phys. Vol 20, 1976.
7. Antar, B. N., "On the Solution of Two-Point Linear Differential Eigenvalue Problems," J. Comp. Phys. Vol. 20, 1976.

## NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACEFLIGHT CENTER THE UNIVERSITY OF AIABAMA <br> ON ROSEN'S THEORY OF GRAVITY AND COSMOLOGY

Prepared By:
Academic Rank:
University and Department:
NASA/MSFC:
(Laboratory)
(Division)
(Branch)
MSFC Counterpart:
Date:
Contract No.:

Dr. Ronnie C. Barmes
Associate Professor
Lambuth College Physicai Science Department

Space Sciences
Cyrogenics
Gravitation

Dr。Peter Eby
August 1, 1980

NGT-01-002-099
(University of Alabama)

## ON ROSEN'S THEORY OF GRAVITY AND COSMOLOGY

By

Dr. Ronnie C. Barnes Associate Professor of Astronomy Lambuth College Jackson, TN 38301<br>ABSTRACT

Formal similarities between general relativity and Rosen's bimetric theory of gravity have been used to analyze various bimetric cosmologies. The following results have been found:
(1) Physically plausible model universes which have a flat static background metric, a Robertson-Walker fundamental metric and which allow co-moving coordinates do not exist in bimetric cosmology.
(2) It is difficult to use the Robertson-Walker metric for both the background metric and $g_{\mu \nu}$ and require that $g_{\mu \nu}$ and $\gamma_{\mu \nu}$ have different time dependences because there are not enough bimetric field equations to simultaneously fix the time dependences.
(3) A consistency relation for using co-moving coordinates in bimetric cosmology was derived.
(4) Certain spatially flat bimetric cosmologies of Babala (1975) were tested for the presence of particle horizons.
(5) An analytic solution for Rosen's (1978) $\mathrm{k}=+1$ model was found.
(6) Rosen's singularity free $k=+1$ model arises from what appears to be an arbitary choice for the time dependent part of $\gamma_{\mu \nu}$ 。

## ACKNOWLEDGEMENTS

It isa.pieasure to acknowledge the support of Dr. Peter Eby who suggested the investigation of particle horizons in bimetric cosmology and for several interesting discussions about Rosen's theory.
Thanks should also be given to Dr. Ted Fay and to his wife, Ann, for providing a home away from home and thus making this research easier to accomplish.

## I. INTRODUCTION

In 1973 and 1974 Nathan Rosen proposed a bimetric theory of gravitation. The origin of its name arises from Rosen's assumption that two tensor fields ( $g_{\mu \nu}$ and $\gamma_{\mu \nu}$ ) exist at each point $\left(x^{1}, x^{2}, x^{3}, x^{4}\right)$ of spacetime. Rosen $(1973,1974,1978)$ has given the following interpretations for $g_{\mu \nu}$ and $\gamma_{\mu \nu}$ :
As in general relativity, $g_{\mu \nu}$ is the fundamental metric tensor of Riemannian geometry. The motion of material bodies and photons is detecmined by $g_{\mu \nu}$. Moreover, those aspects of spacetime which can be observed by using clocks and measuring rods are contained in $g_{\mu \nu}$.
On the other hand, $\gamma_{\mu \nu}$ may be thought of as representing the spacetime metric that one would obtain if all the mass-energy in the universe was removed. $\gamma_{\mu \nu}$ is called the "background" metric. According to Rosen (1978), " $\gamma_{\mu \nu}$ describes a geometry of spacetime not directly observable at present."
Bimetric gravitation has a number of interesting features:

1. It is a metric theory which, if suitable boundary conditions are chosen, agrees exactly with general relativity in the post-Newtonian limit (Lee, et. al. 1976).
2. It does not appear to predict the existence of black holes (Rosen 1973, 1974).
3. Considerably larger (5-6X) masses for neutron stars are allowed by bimetric gravitation than by general relativity (Rosen and Rosen 1975).
4. If the background metric is assumed to be Minkowskian $\left(\gamma_{\mu \nu}=\eta_{\mu \nu}\right)$, then it is possible to express gravitational energy and momentum density via a conserved tensor and the field equations and resulting calculations are simpler than in general relativity (Rosen 1973,1974).
5. Homogeneous bimetric cosmological models which have acceptable values for the Hubble parameter $H_{o}$ and the deceleration parameter $q_{0}$ have been found and studied by Babala ( $k=0,1975$ ), Goidman and Rosen ( $k=-1$, 1977), and Rosen ( $k=+1$, 1978).
6. At the present time the only observation known to pose difficulties for Rosen's theory is the period change $|\dot{\mathrm{P}} / \mathrm{P}|=1.2 \times 10^{-7} \mathrm{yr}^{-1}$ determined by Taylor et.al. (1976) for the binary pulsar PSR $1913+16$. Will and Eardley (1977) analyzed binary star motion using
bimetric theory and, assuming that gravitational radiation from a source is described by a retarded potential solution of the wave equation, have shown that bimetric gravitation allows dipole radiation. (However, also see Rosen 1978b for a different view.) Will and Eardley concluded that Rosen's theory is consistent with observations only if PSR $1913+16$ contains either (a) two very heavy neutron stars of nearly equal mass or (b) a very light neutron star whose companion is a rapidly rotating white dwaif or helium main sequence star.

## II. MOTIVATION--COSMOLOGICAL HORIZONS

One of the most intriguing predictions of general relativity is its prediction of "event horizons" in spacetime. These are boundaries which separate observeable events from those not observeable. The best known example of an event horizon is that which surrounds a spherical black hole and which happens to arise at the same value of the radial coordinate, $r=2 \mathrm{GM} / \mathrm{c}^{2}$, for which the Schwarzchild metric

$$
\begin{equation*}
d s^{2}=\left(1-\frac{2 G M}{\pi c^{2}}\right) d t^{2}-\frac{d r^{2}}{1-2 G M / \pi c^{2}}-r^{2} d \theta^{2}-r^{2} \sin ^{2} \theta d \phi^{2} \tag{1}
\end{equation*}
$$

has a singularity.
However, the existence of an horizon in spacetime is independent of the choice of a coordinate system and, in fact, it is possible to have event horizons even when spacetime is perfectly homogeneous. Rindler (1956), in his classic paper on cosmological horizons, showed that homogeneous cosmology permits universes with particle horizons, event horizons or both.
The presence of particle horizons (a surface in 3-space which divides all fundamental particles into two classes: those which have been observed by a given "fundamental observer" and those which have not been observed) has troubled cosmologists for the following reason. Particle horizons represent regions which have been and still are causally disconnected from other parts of the universe. However, according to the precepts of canonical cosmology, all fundamental observers must find identical physical conditions (of course disregard small scale local inhomogenities) at the same instant of cosmic time. The obvious and difficult question is: How do regions which are causally disconnected achieve such a remarkable degree of physical identity?
Of course, one can always postulate special initial conditions, but this approach seems more in the spirit of theology than of
science. A more challenging approach is to assume that the universe began with anisotropies present and that through the actions of certain mechanisms homogeneity naturally arose.(Misner, Thorne and Wheeler 1973, ch. 30). For this approach to be successful it is essential that the following two requirements be met:
(1). A satisfactory mechanism for converting initial anisotropies into isotropy be found.
(2). Whatever the mechanism, there must be sufficient time to create the degree of homogeneity and isotropy now observed
It is possible to use the $3^{\circ} \mathrm{K}$ microwave background radiation to place some constraints on the time available for producing homogeneity (Misner, Thorne and Wheeler 1973, ch. 30). These constraints are so severe--the universe must be already homogeneous as early as the epoch $\mathrm{z}=1000$ (only 600,000 yr after creation in an Einstein-de Sitter model ) that no entirely satisfactory chaotic cosmology has yet been found.
Therefore, because particle horizons pose an important problem for cosmology, it is important to analyze bimetric cosmology in order to discover if these horizons exist in bimetric models as they do in general relativity.

## III. BIMEIRIG COSMOLOGY WITH THE ROBERTSONWALKER METRIC.

Nowadays it is popular to use Robertson-Walker spacetime $d s^{2}=d t^{2}-\frac{R^{2}(t)}{\left(1+k r^{2} / 4\right)^{2}}\left(d r^{2}+r^{2} d \theta^{2}+r^{2} \sin ^{2} \theta d \phi^{2}\right)$
for cosmological investigations. Here ( $r, \theta, \phi$ ) form a set of "comoving" spherical coordinates, $t$ is cosmic time, $R(t)$ is the scale factor function, and $k$ is the constant of space curvature.
Rindler (1956) derived equations which allow one to test for horizons in Robertson-Walker spacetime. In models which have $R=0$ a particle horizon will exist at time $t_{0}$ if

$$
\begin{equation*}
\int_{0}^{t_{0}} \frac{d t}{R(t)} \quad \text { converges. } \tag{3}
\end{equation*}
$$

Some cosmologies never pass through $R=0$ and begin their motion as $t \rightarrow-\infty$. In these models a particle horizon will exist at $t_{0}$ if $\int_{-\infty}^{t_{0}} d t / R(t)$ converges.
For an event horizon there is only one test, viz.,
$\int_{t_{0}}^{\infty} d t / R(t)$ converges.
Although the above tests are, at least in principle, straightforward there do arise problems when one attempts to use them in bimetric cosnology. The principle difficulty arises from the

> XXIII-5
from the fact that none of the published bimetric cosmological models are consistent with the Robertson-Walker line element for $g_{\mu \nu}$ and, at the same time, consistent with the form used for $\gamma_{\mu \nu}$ in solving the bimetric field equations. For example, Babala (1975) initially assumes

$$
\begin{align*}
& d s_{I}^{2}=g_{\mu \nu} d x^{\mu} d x^{\nu}=e^{2 \Phi} d t^{2}-e^{2 \psi}\left(d r^{2}+r^{2} d \theta^{2}+r^{2} \sin ^{2} \theta d \phi^{2}\right)  \tag{6}\\
& d s_{I I}^{2}=\gamma_{\mu \nu} d x^{\mu} d x^{\nu}=d t^{2}-d r^{2}-r^{2} d \theta^{2}-r^{2} \sin ^{2} \theta d \phi^{2} \tag{7}
\end{align*}
$$

where $\Phi$ and $\Psi$ are functions of $t$. By redefining the time coordinate it is possible to transform eq. (6) into RobertsonWalker form

$$
\tau \equiv \int e^{\Phi} d t \quad ; \quad d s_{I}^{2}=d \tau^{2}-e^{2 \Psi(\tau)}\left(d r^{2}+r^{2} d \theta^{2}+r^{2} \sin ^{2} \theta d \phi^{2}\right)
$$

From equations given by Babala it is easy to show that his models require
$\tau=e^{\Phi_{0}} \tanh t+k$
where $\Phi_{0}$ and $k$ are obtained from initial conditions. The simplest case to analyze is $k=\Phi_{0}=0$ and then one finds

$$
\begin{align*}
& \tau=\tanh t \\
& t=\frac{1}{2} \ln \frac{1+\tau}{1-\tau}  \tag{12}\\
& e^{\Psi(\tau)}=e^{\Psi_{0}} \frac{1}{1-\tau^{2}}\left(\frac{(1+\tau}{1-\tau}\right)^{\lambda^{\prime}} \tag{13}
\end{align*}
$$

where $\lambda^{\prime}$ is a constant of integration related to the value of $\dot{\Psi}_{0}$. According to Babala, expanding models are given by $\lambda^{\prime}>1, \lambda^{\prime}=1$ $-1<\lambda^{\prime}<1$. The simplest case to study is the one with $\lambda^{\prime}=1$ because eq. (13) simplifies to form easy to integrate.
If the Rindler tests are applied, then one finds that this model contains both a particle horizon and an event horizon at $\tau_{0}$ because both
$\int_{-\infty}^{t_{0}} \frac{d t}{R(t)} \rightarrow \int_{-1}^{\tau_{0}} e^{-\Psi_{0}}(1-\tau)^{2} d \tau=\frac{e^{-\Psi_{0}}}{3}\left(8-\left(1-T_{0}\right)^{3}\right)$
$\int_{t_{0}}^{\infty} \frac{d t}{\bar{R}(t)} \rightarrow \int_{T_{0}}^{1} e^{-\Psi_{0}}(1-\tau)^{2} d \tau=\frac{e^{-\Psi_{0}}}{3}\left(1-T_{0}\right)^{3}$
converge.

However, it is possible to find bimetric models without particle horizons. An example is afforded by choosing $\lambda^{\prime}=2$. In this model the particle horizon integral logarithmically diverges as $\tau \rightarrow-1$.

$$
\begin{equation*}
\left.\int_{-1}^{\tau_{0}} e^{-\psi_{0}\left(\frac{(1-\tau)^{3}}{1+\tau}\right.} d \tau=\left(8 \ln (1+\tau)-8(1+\tau)+2(1+\tau)^{2}-\frac{1}{3}(1+\tau)^{3}\right) e^{-\psi_{0}}\right]_{-1}^{\tau_{0}} \tag{16}
\end{equation*}
$$

The difficulty which results when the time coordinate is redefined lies in the fact that while eq. (8) succeeds in converting eq. (6) to Robertson-Walker form it also forces the background metric to be time dependent. However, in deriving eqs. (10)-(13), gotten by substituting eqs. (6) and (7) into the bimetric field euations, Babala assumed that $\gamma_{\mu \nu}$ was independent of time. Bimetric cosmologies published by Goldman and Rosen (1977) and Rosen (1978) are subject to the same criticism. However, in the latter paper Rosen assumes that $\gamma_{\mu \nu}$ is time dependent at the outset and that $g_{\mu \nu}$ is given by eq. (6). In this case the effect of making a new time coordinate via eq. (8) is to change the time dependence of $\gamma_{\mu \nu}$ away from the form used in setting up the field equations.
The preceding considerations suggest that the Robertson-Walker line element be introduced at the start rather than making an attempt to transform to it at a later stage. Unfortunately, this procedure does not work if one uses a flat, static background metric. On substituting eqs. (2) and (7) into the bimetric field equations (see Rosen 1973,1974 for more detail):

$$
\begin{equation*}
N_{\mu \nu}=\frac{1}{2} \gamma^{\alpha \beta} g_{\mu \nu \mid \alpha \beta}-\frac{1}{2} \gamma^{\alpha \beta} g^{\lambda_{\rho}} g_{\mu \lambda \mid \alpha} g_{\nu \rho \mid \beta}=-8 \pi \kappa\left(T_{\mu \nu}-\frac{1}{2} g_{\mu \nu} T\right) \tag{17}
\end{equation*}
$$

and evaluating the $(4,4)$ component one finds

$$
\begin{equation*}
0=-4 \pi n \rho \tag{18}
\end{equation*}
$$

for a matter dominated universe with neglible pressure. Apparently, only empty Robertson-Walker bimetric cosmologies are allowed by using eqs. (2) and (7).
The simplest way to produce non-empty bimetric Robertson-Walker models is to let $\gamma_{\mu \nu}$ be time dependent. A.straightforward method would be to let $\gamma_{\mu \nu}$ have a Robertson-Walker form but with a different time dependence than $g_{\mu \nu}$, viz.,

$$
\begin{align*}
& d s_{I}^{2}=d t^{2}-\frac{R^{2}(t)}{\left(1+k r^{2} / 4\right)^{2}}\left[d r^{2}+r^{2} d \theta^{2}+r^{2} \sin ^{2} \theta d \phi^{2}\right]  \tag{19}\\
& d s_{\text {II }}^{2}=d t^{2}-\frac{S^{2}(t)}{\left(1+k r^{2} / 4\right)^{2}}\left[d r^{2}+r^{2} d \theta^{2}+r^{2} \sin ^{2} \theta d \phi^{2}\right] . \tag{20}
\end{align*}
$$

However, this approach is not completely satisfactory because
there are not enough field equations to simultaneously determine both $R(t)$ and $S(t)$.
Rosen (1978) made an interesting attempt to fix the time dependence of $\gamma_{\mu \nu}$ by using the cosmological principle. It is well known, e.g., see Robertson and Noonan 1968, ch. 14, 347, that theonly non-static metric which satisfies the perfect cosmological principle is the spatially flat de Sitter line element

$$
\begin{equation*}
d s^{2}=d t^{2}-e^{\alpha t}\left(d r^{2}+r^{2} d \theta^{2}+r^{2} \sin ^{2} \theta d \phi^{2}\right) \tag{21}
\end{equation*}
$$

When $\gamma_{\mu \nu}$ is identified with de Sitter's metric it is possible to solve Rosen's field equations for $R(t)$. Rosen, however, rejected the resulting model because it has a big-banglike singularity at $t \rightarrow \infty$. Assuming that the time dependence of $\gamma_{\mu \nu}$ is $\cosh t / a$ rather than $\exp 2 t / a$ and that both $g_{\mu \nu}$ and $\gamma_{\mu \nu}$ are spacetimes of constant positive curvature, Rosen was able to derive an expanding, closed and singularity free universe.

However, it should be pointed out that these goals were reached by making what seems to be an arbitary choice for $S(t)$. Also Rosen fails to point out that if $S(t)=\cosh t / a$, then $\gamma_{\mu \nu} n o$ longer satisfies the perfect cosmological principle because the Hubble parameter of this metric

$$
\dot{s} / s=\frac{1}{a} \tanh t / a
$$

is not independent of time.
During the course of this fellowship the author attempted to develop a satisfactory mehod for fixing the time dependent part of $\gamma_{\mu \nu}$. One approach, which looked promising, was suggested by the fact that in general relativity one can start with a rather general spacetime

$$
\begin{equation*}
d s^{2}=e^{\nu(r, t)} d t^{2}-e^{\mu(r, t)}\left(d r^{2}+r^{2} d \theta^{2}+r^{2} \sin ^{2} \theta d \phi^{2}\right) \tag{23}
\end{equation*}
$$

and derive the Robertson-Walker line element by making use of a "consistency relation" which stems from the use of co-moving coordinates (McVittie, 1965, ch.8).
Since no published work about bimetric consistency relations now exists, the author hoped to (1) derive a consistency relation for the use of co-moving coordinates and (2) use it, if possible, to pin down the time dependence of $\gamma_{\mu \nu}$.

The first step was accomplished for the following case. If one assumes that $\gamma_{\mu \nu}$ and $g_{\mu \nu}$ generate spherically symmetric spacetimes

$$
\begin{align*}
& d S_{I}^{2}=g_{\mu \nu} d x^{\mu} d x^{\nu}=e^{\nu(r, t)} d t^{2}-e^{\mu(r, t)}\left(d r^{2}+r^{2} d \theta^{2}+r^{2} \sin ^{2} \theta d \phi^{2}\right)  \tag{24}\\
& d s_{I I}^{2}=\gamma_{\mu \nu} d x^{\mu} d x^{\nu}=d t^{2}-e^{\lambda(r, t)}\left(d r^{2}+r^{2} d \theta^{2}+r^{2} \sin ^{2} \theta d \phi^{2}\right) \tag{25}
\end{align*}
$$

and then substitutes these into the bimetric field equations, then one finds the following non-vanishing components

$$
\begin{align*}
N_{11}=\frac{1}{2}\left[\mu_{r r}-\lambda_{r r}\right. & \left.+\left(\mu_{r}-\lambda_{r}\right)\left(\frac{2}{r}+\frac{1}{2} \lambda_{r}\right)\right] e^{\mu-\lambda}+\frac{1}{2}\left[\lambda_{t t}-\mu_{t t}+\frac{3}{2}\left(\lambda_{t}-\mu_{t}\right)\right] e^{\mu}  \tag{26}\\
& +\frac{1}{8} \lambda_{t}^{2} e^{2 \mu-\lambda_{-N}}-\frac{1}{8} \lambda_{t}^{2} e^{\lambda_{+\omega}} \tag{27}
\end{align*}
$$

$N_{22}=r^{2} N_{11}$
$N_{33}=r^{2} N_{11} \sin ^{2} \theta$
$N_{41}=-\frac{1}{2}\left[\nu_{r r}+\left(\frac{2}{r}+\frac{1}{2} \lambda_{r}\right) \nu_{r}\right] e^{\nu-\lambda}+\frac{1}{2}\left[\nu_{t t}+\frac{3}{2} \lambda_{t} \nu_{t}\right] e^{\nu}+\frac{3}{8} \lambda_{t}^{2}\left[e^{\mu-\lambda}-e^{\lambda+2 \nu-\mu}\right]$
$N_{H}=\frac{1}{4}\left[\lambda_{t r}+\lambda_{t}\left(\nu_{r}+\lambda_{r}-\mu_{r}\right)\right] e^{\nu}-\frac{1}{4}\left[\lambda_{t r}-\lambda_{t}\left(\nu_{r}+\lambda_{r}-\mu_{r}\right)\right] e^{\mu-\lambda}$.
The use of co-moving coordinates requires that, in spherically symmetric spacetime,
$T^{14}=(\rho+p) \frac{d x^{1}}{d s} \frac{d x^{4}}{d s}-g^{14} p$
vanish since $d x^{\prime} / d s=0$ and $g^{H}=0$. However, these conditions also require $\mathrm{T}^{14}-\mathrm{g}^{14} \mathrm{~T} / 2=0$ and that means $\mathrm{N}_{1} 4=0$. Thus, the bimetric consistency relation is

$$
\begin{equation*}
\left[\lambda_{t r}+\lambda_{t}\left(\nu_{r}+\lambda_{r}-\mu_{r}\right)\right] e^{\nu}=\left[\lambda_{t r}-\lambda_{t}\left(\nu_{r}+\lambda_{r}-\mu_{r}\right)\right] e^{r-\lambda} \tag{32}
\end{equation*}
$$

For comparison, the analogous general relativistic equation is

$$
\begin{equation*}
\mu_{r t}=\frac{1}{2} \mu_{t} \nu_{r} \tag{33}
\end{equation*}
$$

As outlined in McVittie's textbook, it is possible to integrate eq. (33) and, by requiring that the ( 1,1 ) and ( 2,2 ) components of the Einstein tensor yield the same cosmological fluid pressure, one arrives at the Robertson-Walker line element.
Unfortunately, when the same procedure is followed in bimetric theory, one discovers that $N_{11}$ and $N_{22}$ already give the samedifferential equation for pressure. Thus, unlike general
relativity, the consistency relation for co-moving coordinates does not lead to appreciable progress and, in particular, no new information is gained about the time dependence of $\gamma_{\mu \nu}$.
IV. AN ANALYTIC SOLUTION FOR ROSEN'S $1978 \mathrm{k}=+1$ MODEL

As mentioned in the previous section, Rosen (1978) found a closed space expanding model which does not have a singularity. During the course of this summer's work the author found an analytic solution to the bimetric field equations of this model.

$$
\begin{align*}
& \ddot{\Phi}+\frac{3}{a}\left(\tanh \frac{t}{a}\right) \dot{\Phi}+\frac{3}{a^{2}} \tanh ^{2} \frac{t}{a} \sinh (2 \psi-2 \Phi)=-4 \pi e^{\Phi+3 \psi}(\rho+3 p) \\
& \ddot{\psi}+\frac{3}{a}\left(\tanh \frac{t}{a}\right) \dot{\psi}-\frac{1}{a^{2}} \tanh ^{2} \frac{t}{a} \sinh (2 \psi-2 \Phi)=4 \pi e^{\Phi+3 \psi}(\rho-p) \\
& \dot{\rho}+3(\rho+p)\left(\dot{\psi}+\frac{1}{a} \tanh \frac{t}{a}\right)=0 \tag{36}
\end{align*}
$$

If one investigates a zero pressure model and sets $g_{44}=1$, then these equations simplify and it is possible to use 44 the integral of eq.(36)

$$
\rho^{e^{3 \psi} \cosh ^{3} \frac{t}{a}=A \quad ; A \text { is a constant of integration }, ~}
$$

to set up a single second order differential equation which is linear in $\Psi$

$$
\begin{equation*}
\ddot{\psi}+\frac{3}{a} \dot{\psi} \tanh t / a=\frac{8 \pi A}{3} \operatorname{sech}^{3} \frac{t}{\alpha} . \tag{38}
\end{equation*}
$$

By using the classical variation of parameters method it is possible to arrive at the following general solution to eq.(38):

$$
\begin{array}{r}
\Psi(t)=\left(\frac{8 \pi}{3} A \frac{t}{a}+\frac{1}{2} a B\right)\left[\operatorname{sech} \frac{t}{a} \tanh \frac{t}{a}+\tan ^{-1}(\sinh t / a)\right]+ \\
\frac{8 \pi A}{3} \operatorname{sech} \frac{t}{a}-\frac{8 \pi A}{3} \int \tan ^{-1}\left(\sinh \frac{t}{a}\right) \frac{d t}{a}+C . \tag{39}
\end{array}
$$

In this solution $B$ and $C$ are constants of integration and the integral can be carried out by expanding $\tan ^{-1}(\sinh t / a)$ in a power series.
Although eq. (39) is not as simple as one would like, it is nevertheless, sufficiently simple to analyze for possible singularities in cosmic time. By letting $t \rightarrow-\infty, t \rightarrow 0$ and $t \rightarrow \infty$ and looking at each term in (39) one can verify that Rosen's conclusion about the absence of a singularity in this model is correct.

## V. SUMMARY

Because the formalism of Rosen's bimetric theory is so similiar to general relativity one would think that it would be a simple matter to set up bimetric cosmology by using the Robertson-Walker metric for $g \mu \nu$. However, if such is attempted, then one finds that, there are too few bimetric field equations to simultaneously fix the time dependence of $g_{\mu \nu}$ and $\gamma_{\mu \nu}$.
Other results are:
(1) Certain flat bimetric cosmologies of Babala have particle horizons; others do not. Apparently, the bimetric field equations (by themselves) are not strong enough to prevent the occurence of cosmological particle horizons, although strangely enough, they do exclude black holes.
(2) An empty universe results if one tries to use a flat background metric along with a Robertson-Walker fundamental metric.
(3) Rosen's singularity free universe of positive spatial curvature (Rosen 1978) arises from what seems to be a an arbitary choice for the time dependent part of $\gamma_{\mu \nu}$ 。
(4) A consistency relation for using co-moving coordinates in bimetric cosmology was derived.
(5) An analytic solution was found for the field equations which describe Rosen's (1978) universe of positive space curvature.

## REFERENCES

Babala, D. 1975, J. Phys. A., 8, 1409.
Goldman, I. and Rosen, N. 1977, Ap.J., 212, 602.
Lee, D.I., Caves, C.M., Ni, W-T, and Will, C.M. 1976, Ap.J., 206, 555.
Mc Vittie, G.C. 1962, General Relativity and Cosmology (U。 of Ill. Press: Urbana).
Misner, C.W., Thorne, K.S. and Wheeler, J.A. 1973, Gravitation (Freeman Co.: San Francisco).
Rindler, W. 1956, MNRAS, 116, 662.
Robertson, H.P. and Noonan, T.W. 1968, Relativity and Cosmology, (W.B. Saunders Co.: Philadelphia).

Rosen, N. 1973, J. Gen. Rel. Grav., 4, 435. 1974, Ann. Phy. (N.Y.), $84,455$.
1978, J. Gen. Rel. Grav., 2, 339。
1978b, Ap.J., 221, 284.
Rosen, N. and Rosen, J. 1975, ApoJ. 202, 782.
Taylor, J.H., Hulse, R.A., Fowler, L.A., Gullahorn, G.E. and Rankin, J.M. 1976 Ap.J. (Lett.), 206, L53. Will, C.W. and Eardley, D.M. 1977 Ap. J. (Iett.), L91.

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA <br> MEASUREMENTS OF THE SAMPLING VOLUME OF AN INFRARED LASER DOPPLER VELOCIMETER FOR THE STUDY OF FEASIBILITY OF SIMULTANEOUS THREE-DIMENSIONAL VELOCITY PROFILE MAPPING 

| Prepared By: | Robert L. Bond, Ph.D. |
| :--- | :--- |
| Academic Rank: | Associate Professor |
| University and Department: | University of Arkansas <br> Department of Electronics <br> and Instrumentation |
| NASA/MSFC: <br> (Laboratory) <br> (Division) <br> (Branch) | Electronics and Control |
| MSFC Counterpart: | Optical and R.F. Systems <br> Optics |
| Date: | James W. Bilbro |
| Contractor No.: | August l, l980 |

# MEASUREMENTS OF THE SAMPLING VOLUME OF AN INFRARED LASER DOPPLER VELOCIMETER FOR THE STUDY OF FEASIBILITY OF SIMULTANEOUS THREE-DIMENSIONAL VELOCITY PROFILE MAPPING 

## By

Robert L. Bond
Associate Professor of Electronics \& Instrumentation University of Arkansas Little Rock, Arkansas

ABSTRACT

This project is a continuation and configuration of the experimental work performed during the summer of 1979. Data obtained after the submission of last year's report gave strong indication that it would be possible by spatial masking to code information onto an outgoing laser beam which would yield three-dimensional velocity information when analyzed at an interferometer.

The work continued into this year has involved detailed, high resolution incoherent mapping of the sampling volume to determine the three-dimensional profile of the volume of the beam from which information will be obtained by scattering.

## ACKNOWLEDGEMENTS

I wish to express my gratitude to Mr. James W. Bilbro for his continuous support and many suggestions offered throughout the performance of this work. I also appreciate Mr. Bilbro's assignment of Miranda James to assist me during the data-taking phase of this work. Miranda has substantially facilitated this normally time-consuming and tedious process. I also wish to thank Dr. Bob Barfield for his very competent management of the Summer Faculty Program.

## LIST OF FIGURES

| Figure | Title | Page |
| :---: | :---: | :---: |
| 1 | Optical Block Diagram | XXIV. 8 |
| 2 | Trace along x-axis through unfocused laser beam -- $500 \mu \mathrm{~m}$ aperture | XXIV. 9 |
| 3 | Trace along z-axis of beam focused at 47 feet -- $25 \mu \mathrm{~m}$ pinhole | XXIV. 10 |
| 4 | Trace along z-axis of beam focused at 47 feet -- comparison with and without wire filter -- $50 \mu \mathrm{~m}$ pinhole | XXIV. 11 |
| 5 | Trace along x-axis -- comparison with and without wire filter -- $50 \mu \mathrm{~m}$ pinhole | XXIV. 12 |
| 6 | Trace along z-axis -- comparison with and without wire filter -- 50 $\mu \mathrm{m}$ pinhole | XXIV. 13 |
| 7 | Comparison of before focus traces in x -axis | XXIV. 14 |
| 8 | Comparison of after focus traces in x-axis | XXIV. 15 |
| 9 | Comparison of before focus traces in $y$-axis | XXIV. 16 |
| 10 | Comparison of after focus traces in y -axis | XXIV. 17 |
| 11 | Traces (x) in focal plane for the total beam and for left and right halves | XXIV. 18 |
| 12 | Traces (y) in focal plane for the total beam and for left and right halves | XXIV. 19 |
| 13 | Traces (x) 2.5 inches after focus for the total beam and for left and right halves | XXIV. 20 |
| 14 | Traces (y) 2.5 inches after focus for the total beam and for left and right halves | XXIV. 21 |

Marshall Space Flight Center has been involved in the application of Laser Doppler Velocimeters (LDVs) since their inception by Yeh and Cummins ${ }^{1}$ in 1964. Subsequent research has applied LDVs to the study of aircraft wake vortices, 2,3 atmospheric turbulence measurements, 4 remote intensity fluctuations, 5 dust devil measurements, 6 and presently to severe storms. 7

Except for the earliest work in wind tunnels, most applications have used $\mathrm{CO}_{2}$ lasers to increase the signal-to-noise ratio - an advantage resulting from the atmospheric window at 10.6 micrometers. Present work involves pulsed $\mathrm{CO}_{2}$ laser to further increase the signal-to-noise ratio. Future work will utilize the advantages of off-axis telescopes and, possibly, lasers operating at wavelengths which more nearly match atmospheric windows.

Each LDV has certain limitations on the velocity vectors which it can measure. The early LDV of Yeh and Cummins, for example, could easily measure flow velocities perpendicular to the incident optical plane. However, it could not measure velocity components along the optical axis. Subsequent systems, such as those used to measure aircraft wake vortices, measure the components along the optical direction at the expense of the orthogonal components. A need exists for an LDV which is capable of reading all orthogonal velocity components. Such an LDV could be used to map the three-dimensional characteristics of a remote flow field.

Several developments in LDVs point to the possibility of coding sufficient information into a laser beam so that the returned radiation will contain three-dimensional velocity component data. These developments include the classical heterodyne, fringe, and the more recent time-of-flight LDVs. It is possible that a combination of these approaches would yield the three-dimensional information necessary to map a three-dimensional flow field. The most likely impediments appear to be instrumentation artifacts which may obscure the signal of one or more velocity components or produce an unacceptable signal-to-noise ratio.
XXIV. 1

The purpose of this work has been to obtain a detailed understanding of the sampling volume of the infrared laser Doppler velocimeter assembled during the 1979 Summer Faculty Fellowship Program. Detailed, high resolution maps were to be made of both the coherent and incoherent sampling volumes. The response of the sampling volume to spatial masking within the telescope was also to be determined. The response of the system to various objects moving through the sampling volume was also to be determined.

## EXPERIMENTS

The infrared laser Doppler velocimeter (Figure 1) assembled during the Summer Faculty Fellowship Program last summer was returned to its operational status. The system was improved by modifying the external mirrors to permit mechanical rather than electrical positioning. This step was necessary to maintain more precise alignment of the entire laser Doppler velocimeter (LDV).

The detailed study of the coherent sampling volume yielded essentially the same results as those obtained last summer. The double maximum present last summer in the Doppler Signal Power vs. Secondary Mirror Position curves was not as pronounced. The cause of this phenomenon was isolated to the z-scan mechanism of the secondary mirror. As the secondary mirror is moved along the optical (z) axis, the focal point projected onto the detector assembly (27-50 feet away) moves erratically about in the $x-y$ plane of the detector. Of course, similar wandering of the beam occurs relative to a fixed position on the rotating sandpaper wheel.

Ironically, a similar phenomenon occurs when the rotating sandpaper wheel located on a carriage on the optical bench is moved along the $z$-axis of the beam. In this case the "wandering of the beam" is only apparent. A fixed point on the rotating wheel is actually moving in the $x-y$ plane as the $z-$ traverse is made. Obviously, data entailing movement of both the wheel and secondary mirror contains a compound error resulting from both phenomena.

Data taken by scanning the secondary mirror, first in one direction, then in the reverse direction, is reasonably repeatable but is not symmetric relative to the maximum power point in the scan. This asymmetry is in part due to the $x-y$ motion introduced into the outgoing beam.

The same data taken by scanning the rotating wheel along the optical axis, first in one direction, then in the reverse direction, is approximately as repeatable but with much less asymmetry.

When the signal power (as measured by a specturm analyzer) was plotted against the voltage (as measured on an oscilloscope) there were discontinuities on the semi-log plot. The causes of these points which did not fall on the straight line plot were never determined. There is a remote possibility
that they resulted from an ailing Honeywell specturm analyzer which failed to function after this data was taken.

The incoherent data verified and accentuated the conclusions arrived at from the coherent data - namely that the beam wanders about in the $x-y$ plane as it rotates in an irregular arc about the optical axis. An example of the phenomenon is as follows. With the secondary mirror adjusted to focus at the near end of the optical bench let us assume the focal point is on the optical axis. As the secondary mirror is adjusted to move the focal point toward the far end of the optical bench, the beam first swings toward the left as it swings upward and finally swings back toward the right as it reaches its maximum height at the far end of the optical bench. This general pattern was repeated for many experiments. Occasionally there was some fine detail in the arc but the trend remained the same. It was determined that the relationship between the secondary mirror position and the focal point position on the optical axis was highly repeatable and agreed with theoretical predictions.

The incoherent sampling volume was measured by moving an aperture throughout the double cone defined by the focused beam from the laser. Figure 2 is a traverse in the $x$-axis of the laser beam without secondary mirror or wire filter. This x-axis profile is the standard for all additional work which was taken with the beam focused between 27 and 50 feet by the secondary mirror. A $500 \mu \mathrm{~m}$ aperture was selected for the traverse of the unfocused beam to keep the power density within the limits of the pyroelectric detector and lock-in amplifier used for power detection. Apertures of suitable size were always used for taking data since neutral density filters operable at $10.6 \mu \mathrm{~m}$ were not available. Attempts to use existing flats of $\mathrm{Ge}, \mathrm{CaF}_{2}$ or $\mathrm{BaF}_{2}$ were futile because of significant artifacts introduced into the data via diffraction and multiple reflections. Notice the asymmetry in the beam which is present to some extent in all data.

Initial data was taken with a 1.0 mm aperture as was done last year. It was determined that this aperture did not produce sufficient resolution for a laser beam that has a theoretical focal diameter of approximately 40 mil . The resolution would be particularily dismal if one were interested in detecting auxillary maxima about the center of the beam. As a result, the 1.0 mm aperture attached to a Coherent Radiation power head was replaced with pinholes of either 25 or $50 \mu \mathrm{~m}$ attached to a pyroelectric detector. Depending upon the laser
output power, the $50 \mu \mathrm{~m}$ aperture was frequently so large that it produced slight saturation of the pyroelectric detector or lock-in amplifier. Therefore, much of the data was taken with the $25 \mu \mathrm{~m}$ aperture when the laser power was too large for the other size aperture. The Coherent Radiation power head could not be used because it did not have sufficient sensitivity with small apertures and because its frequency response was below the 4 Hz minimum response of the lock-in amplifier and accompanying rotating aperture. Though data were taken at many z-axis positions, those which follow were taken at a nominal 47 feet.

Figure 3 is a trace along the optical axis obtained by moving the aperture along the optical bench with the secondary mirror set for a nominal distance of 47 feet. Notice the asymmetry relative to the nominal focal plane of the nominal focal plane is caused by variation in time of the $x-y$ position of the focal point. The beam was originally positioned in $x, y$ and $z$ to provide maximum power at the nominal focal plane.

Figure 4 compares the trace along the z-axis taken with and without the wire filter in place. There is a natural reduction in total power when the wire is in place but otherwise the two traces are similar.

Additional information was obtained by $x$ and $y$ traverses of the sampling volume at the nominal focus and at distances before and after the nominal focal plane. Figure 5 compares the $x$-axis traverse taken at the nominal focal plane with and without the wire filter in position. Figure 6 presents the same information for the $y$-axis. Other than the normal reduction in power produced when the wire is introduced, the two traces are similar for both the $x$ and $y$ traverses.

Figure 7, in combination with Figure 8, represents several $x$ traverses in the nominal focal plane and in several planes before and after the nominal focal plane. Figure 7 compares the $x$ trace in the nominal focal plane to the $x$ traces taken at nominal half ( 5 in.$)$ and quarter ( 7 in .) power points before focus. Figure 8 compares the $x$ trace in the focal plane to the $x$ trace taken at nominal half power (3 in.) and quarter power (4 in.) points after focus.

Figure 9 and 10 present corresponding traces for the $Y$ axis.

The introduction of the wire filter just prior to the secondary mirror of the telescope splits the outgoing laser beam into two distinct nearly equal components - one left and one right. Each component is made up of a large lobe of radiation separated into several smaller diffraction lobes whose intensity decreases rapidly as one progresses from the $z$ axis outward - to right for right hand lobe - to left for left hand lobe. Data indicates that the $x$ and $y$ tracings taken at the nominal focal plane for the whole beam is not the simple intensity sum of the two individual half beams. It is speculated that the system is sufficiently abberated that the individual lobes each form a non-detailed blur in the region of the focal plane. When the two lobes are simultaneously present at the focal plane, they coherently interfer to form an interferance pattern with its maximum on axis and usually with side lobes.

Figure 11 presents a trace along the x axis at the nominal focal plane for the total beam and for the left and right hand halves of the beam. Either the right or left half of the beam is removed by placing a mask near the telescope. The edge of the mask coincides with the center of the shadow of the wire filter. The introduction of only one half the beam at a time is guaranteed by viewing the pattern with liquid crystal paper.

Figure 12 presents a trace along the $y$ axis. The $x-y$ position of the focal point has moved in this plot.

Figure 13 and 14 present the same data for a nominal half power point 2.5 inches after the nominal focal plane.

The normal sampling interval for the $x, y$ traverses is 2.5 inches, which provides that there is no overlap of sample areas whether a 25 or $50 \mu \mathrm{~m}$ aperture is used. For the data with the unfocused beam a $500 \mu \mathrm{~m}$ (approximately 20 mil ) aperture sampled every 25 mils.

The incoherent sampling volume of the infrared laser Doppler velocimeter assembled in the 1979 Summer Faculty Fellowship Program was found to be heavily dependent upon the configuration of the outgoing laser beam. In general when only half the beam is allowed to focus at the focal plane, the $x, y$ traverses indicate a broader beam with nominal side lobes. However, if the total beam is permitted to impinge on the focal plane there is substantial narrowing of the beam with increased maximum intensity and an increase in side lobes.

The double maximum indicated by the coherent sampling volume appears to be an artifact of the scanning process not a fundamental entity of the sampling volume.

It is not known at this point how significant any of these effects are on the actual Doppler signal and its analysis. It is recommended that this information be obtained by first "calibrating" the laser output followed by automated data taking procedures to minimize the very apparent time factor which permeates all the data.


FIGURE 1. OPTICAL BLOCK DIAGRAM


FIGURE 2. TRACE ALONG X-AXIS THROUGH UNFOCUSED LASER BEAM -- $500 \mu \mathrm{~m}$ APERTURE


FIGURE 3. TRACE ALONG z-AXIS OF BEAM FOCUSED AT 47 FEET - $25 \mu \mathrm{~m}$ PINHOLE

- WIRE FILTER IN
- WIRE FILTER OUT


FIGURE 4. TRACE ALONG z-AXIS OF BEAM FOCUSED AT 47 FEET -COMPARISON WITH AND WITHOUT WIRE FILTER -- $50 \mu \mathrm{~m}$ PINHOLE


FIGURE 5. TRACE ALONG X-AXIS -- COMPARISON WITH AND WITHOUT WIRE FILTER -- $50 \mu \mathrm{~m}$ PINHOLE



FIGURE 7. COMPARISON OF BEFORE FOCUE TRACES IN $x$-AXIS

FIGURE 8. COMPARISON OF AFTER FOCUS TRACES IN $x-A X I S$


FIGURE 9. COMPARISON OF BEFORE FOCUS TRACES IN x-AXIS

FIGURE 10. COMPARISON OF AFTER FOCUS TRACES IN Y-AXIS


FIGURE 11. TRACES (x) IN FOCAL PLANE FOR THE TOTAL BEAM AND FOR LEFT AND RIGHT HALVES

- TOTAL BEAM
$\triangle$ RIGHT SIDE OF BEAM
- LEFT SIDE OF BEAM


FIGURE 12. TRACES ( y ) IN FOCAL PLANE FOR THE TOTAL BEAM AND FOR LEFT AND RIGHT HALVES



1. Yeh, Y. and H. Z. Cummins, Appl. Phys. Lett 4, 176 (1964).
2. Huffaker, Robert M., et al. Development of a Laser Doppler System for the Detection, Tracking, and Measurement of Aircraft Wake Vortices, Final Report. FAA-RD-74-213, March 1975.
3. Bilbro, J. W., et al. Laser Doppler Velocimeter Wake Vortex Tests, Final Report. NASA TM-X-64988, March 1976.
4. Polge, R. J. and J. W. Bilbro, Proceedings of Southeastcon '78, IEEE, April 1978.
5. Kennedy, L. Z., and J. W. Bilbro, App1. Opt. 15, 2008 (1976).
6. Bilbro, J. W., et al. Laser Dust Devil Measurements, NASA TN D-8429, February 1977.
7. Bilbro, James W. and William W. Vaughn, Bull. Am. Meteorological Soc. 59, 1095 (1978).

## NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA <br> A MONTE CARLO APPROACH TO TOLERANCE ANALYSIS OF A NEAR-DIFFRACTION LIMITED OPTICAL IMAGING SYSTEM

## Prepared By:

Academic Rank:
University and Department:

NASA/MSFC:
(Laboratory)
(Division)
(Branch)
MSFC Counterpart:
Date:
Contract No. :
W. Harold Davenport

Assistant Professor
Columbus College Department of Mathematics

Electronics and Control Optical and R.F. Systems Optics

Donald B. Griner
August 8, 1980
NGT-01-002-099
(University of Alabama)

# A MONTE CARLO APPROACH TO TOLERANCE ANALYSIS OF A NEAR-DIFFRACTION LIMITED OPTICAL IMAGING SYSTEM 

by
W. Harold Davenport, Ph.D. Assistant Professor of Mathematics Columbus College Columbus, Georgia


#### Abstract

Using the theory of Don Knuth, a computer program has been developed to generate random numbers on the Sigma $V$ computer. A subroutine gives a random real number from a uniform distribution on the interval $(0,1)$. A normal distribution with mean zero and variance 1 is also obtained. In my NASA/ASEE report for 1979, a statistical description of 5 possible errors is given. These 5 errors are $S$ which is $X$-axis tilt; $S_{2}$ which is $Y$-axis tilt; $S_{3}$ which is $X$-decenter; $S_{4}=Y$-decenter; ${ }^{2}$ and $S_{5}$ which is despace. ${ }^{3}$

Combining the above results a quality function, $Q\left(S_{1}, S_{2}\right.$, $\ldots, S_{5}$, is obtained via a computer program. Values of $Q$ are given in fractional wavelengths. The (variables) errors $S_{1}, \ldots .$. $\mathrm{S}_{5}$ are assumed to be independent of each other with possibly different statistical descriptions. Consequently, $Q$ is a linear function.


## ACKNOWLEDGEMENTS

The author would like to express his gratitude to Dr. Bob Barfield for directing a very successful summer program. I would also like to express my appreciation to my MSFC counterpart, Don Griner, for his consultations and guidance during the summer; to Charlie Jones for the many technical discussions concerning this project; and to Julie Taylo for her programming assistance.

## LIST OF FIGURES

Figure No. Title Page

1
Ritchey-Chretian Optical System ..... XXV-5
Spot Diagram No Aberrations ..... XXV-6
Spot Diagram With 0.l Arc Min of ..... XXV-7Tilt About X-AxisHistogram of Average RMSOPDXXV-8

## LIST OF TABLES

Table No. Title Page
1 Zernike Coefficients With 0.l Arc Min ..... XXV-9Tilt About the X-Axis
2 Zernike Coefficients With no ..... XXV-10Aberrations

## I. INTRODUCTION

A Monte Carlo approach is used to determine wavefront error on a two mirror telescope. A computer program that traces N rays through an optical system and calculates the root mean square of the optical path difference (RMSOPD) is used for this analysis. An expanded version of this program will also give the Zernike coefficients. In this study only five errors are used, X-tilt, Y-tilt, X-decenter, Y-decenter, and despace of the secondary mirror. A statistical description is obtained for these five errors. Assuming a uniform probability of any one of these errors, we obtain a wavefront error.

The critical ingredient in a Monte Carlo method is a good random number generator. Knuth has developed a theory for an algorithm that generates a uniformly distributed set of pseudorandom numbers. This algorithm was programmed in FORTRAN IV and used to obtain a uniform distribution of random numbers on ( 0,1 ) and a normal distribution of mean zero and variance of 1.

## II. OBJECTIVES

To develop a statistical simulation that uses the Monte Carlo techniques in the form of a computer program. The RSS method of tolerance analysis then is to be compared with the Monte Carlo method.
III. TOLERANCE ANALYSIS OF A NEAR-DIFFRACTION LIMITED OPTICAL SYSTEM

A near-diffraction limited optical imaging system is an optical system designed such that root mean square of the optical path difference (RMSOPD) is less than or equal . $25 \lambda$. In this analysis the wavelength ( $\lambda$ ) used was 632.8 nm . A ray trace computer program written in FORTRAN IV for the Sigma V computer was used. This program can trace $N(N=101$ to 1000 ) rays through an optical system and calculates RMSOPD. The optical path difference (wavefront errors) is obtained by subtracting the chief ray length from the remaining ray lengths. The chief ray is always traced by the program; ignoring the central obscuration it is used as our reference. A spherical wavefront on the image plane (focal plane) is used as the reference. The optical system is a modified Cassegrain or Ritchey-Chretian design. The exit pupil for this design is back of the secondary mirror (see Figure l). The entrance pupil is the primary mirror.

The ray trace program has a graphics capability of plotting the spot. Figure 2 is a spot diagram of the program when no aberrations are present. Page XXV-10 gives the output that corresponds to this case. Recall that aberration describes the lack of homocentricity in a pencil of rays (input) or the deviation of a wave surface from
a spherical form. In Figure 3 we have a spot diagram with 0.1 arc min of tilt about the X -axis. The output that corresponds to this case is on page XXV-9.

The primary lens (mirror) will remain fixed. The secondary has six degrees of freedom but is symmetric with respect to rotation; hence, we only consider:
(i) X -axis tilt ( S )
-.95 to .95 arc min
-1.7 to 1.7 arc min (centroid correction) -2.9 to 2.9 arc min (best chief ray)
(ii) Y-axis tilt ( $\mathrm{S}_{2}$ )

Same range of numbers as X -axis tilt
(iii) X-axis decenter $\left(\mathrm{S}_{3}\right)$
-165 to $165 \mu \mathrm{~m}$
-300 to $300 \mu \mathrm{~m}$ (centroid correction)
-530 to $530 \mu \mathrm{~m}$ (best chief ray)
(iv) Y-axis decenter ( $\mathrm{S}_{4}$ )

Same range of values as X -axis decenter
(v) Despace ( $\mathrm{S}_{5}$ )
-10.9 to 10.9
When the optical system is perturbed by one of the above five perturbations, it is the exit pupil wavefront that is operated upon.

## IV. MONTE CARLO METHOD

A random number generator using the theory of Don Knuth (see reference 7) is used to give a uniform distribution on ( 0,1 ). Also see reference 2, page $240 f$, for a FORTRAN listing of this random number computer program. Let $Q=Q\left(S_{1}, S_{2}, \ldots, S_{5}\right)$ be the quality function where $S_{i}$, $i=1, \ldots, 5$ are the crrors given above. A Monte Carlo apprōach is used to obtain Figure 2. A random number is selected from the random number generator corresponding to a uniform distribution. This number is then scaled to the maximum range of the possible error for $S_{i}$. The five errors $S_{i}$ are assumed to be independent of each other. ${ }^{\text {Hence, }}$ one to five different uniformly distributed random numbers are obtained to give the histogram of Figure 2. The data for each of the histograms was formatted as follows:

Run No.

$S_{2}$
$S_{3} \quad S_{4}$
$S_{5}$ RMSOPD AVG. RMSOPD

One thousand runs were made and the average RMSOPD for these 1000 runs was plotted on the $Y$-axis.

One thousand runs were made where $X-Y$ tilt and $X-Y$ decenter were given an equal probability of occurring. The tolerance range used for each of these four errors was the range that would give a . $25 \lambda$ RMSOPD wavefront error. If we take the root sum square (RSS) of these four values we get $0.5 \lambda$. From Figure 4 we see that the average RMSOPD is $0.302 \lambda$. Hence, this would imply for this scenario that the RSS technique is a more rigid tolerance than a Monte Carlo technique. So an optical system designed with respect to the RSS method would be better than one designed relative to the tolerances given by a Monte Carlo method.

Taking the RSS of two errors with a value of $.25 \lambda$ we get $0.354 \lambda$. From Figure 4 we see that the Monte Carlo method used on any combination of errors X-tilt, Y-tilt, X-decenter, and Y -decenter taken two at a time will give a range of $.158 \lambda$ to $.226 \lambda$.
V. AN OPTICAL SYSTEM VIEWED AS AN AUTOMATON

Let $A=(S, \Sigma, M)$ be an automaton

$\Sigma=\{\sigma / \sigma=$ a unit of measure of x -axis tilt of the secondary mirror in arc min\}.
Let N be some positive integer such that $\sigma^{N}=0$. To obtain $\sigma^{2}=\sigma \sigma$ we define $\sigma^{2}$ to be two units of tilt, e.g., let $\sigma=1$ arc min then $\sigma^{2}$ is two arc min, etc. With this binary operation $\Sigma$ makes a semigroup with zero which acts like an identity. Note $\Sigma$ is semigroup isomorphic to $\left(Z^{\prime}{ }^{+}\right) .\left(Z_{N}{ }^{+}\right)$is the semigroup of integers modulo $N$ with respect to addition.
$S=$ states of the system which will be the Zernike coefficients
 are the Zernike polynomials.

$$
\begin{aligned}
& \text { Now for } M_{\sigma} \varepsilon M, \sigma \varepsilon \Sigma, \text { define } \\
& M_{\sigma}: S \longrightarrow S \\
& A_{i} \longrightarrow A_{i} M_{\sigma}=A_{i}^{\prime}
\end{aligned}
$$

$$
\text { so } M_{\sigma} \sim\left(\begin{array}{lll}
A_{1} & \cdots & A_{k} \\
A_{1}^{\prime} & \ldots & A_{k}^{\prime} \\
A_{1} & \cdots & A_{k}
\end{array}\right)
$$

Choose $P_{1}$,..., $P_{k}$ such that they are orthogonal with respect to the inner product

$$
\begin{gathered}
\left(P_{i}, P_{j}\right)=\int P_{i} P_{j} \\
X^{2}+Y^{2} \leq 1
\end{gathered}
$$

Thus, if $M_{\tau}=\left(\begin{array}{lll}A_{1} & \cdots A_{k} \\ A_{1}^{\prime} & \ldots & A_{k}^{\prime} \\ A_{1} & \cdots & A_{k}\end{array}\right)$ then

$$
\begin{aligned}
& M_{\sigma} M_{\tau}=\left(\begin{array}{lll}
A_{1} \ldots A_{k} \\
A_{1}^{\prime \prime \prime} & \ldots A_{k}^{\prime \prime}
\end{array}\right) \\
& \text { where } A_{i}^{\prime \prime \prime} \\
& A_{i}^{\prime} \\
& =A_{1}^{\prime}+A_{i}^{\prime \prime}
\end{aligned}
$$

## VI. RESULTS AND DISCUSSION

The usual method of combining wavefront errors is the root sum square (RSS). For example, using the five errors above and combining them as follows:

$$
\left(s_{1}^{2}+s_{2}^{2}+s_{3}^{2}+s_{4}^{2}+s_{5}^{2}\right)^{\frac{1}{2}}
$$

Also, another method is to combine those by a Monte Carlo technique. If $S_{5}$, despace, is not considered then we obtain Figure 4. Hence, an optical system designed to operate within the tolerance criterion of the RSS method would also satisfy the tolerance conditions of the Monte Carlo method.

The maximum bounds of the error on focus were approximately -10 to $10 \mu \mathrm{~m}$. The focus proved to be a critical error in this analysis and further study is required.

Problems to consider:

1. Develop algorithms for an optical system considered as an automaton.
2. Extend this Monte Carlo method of tolerance analysis to more complicated optical systems.
3. Develop an algorithm and program it in FORTRAN for a Zernike coefficient program (see Davenport, NASA/ASEE summer report, 1979).

PRIMARY MIRROR


FIGURE 1 RITCHEY-CRETIAN OPTICAL SYSTEM


Figure 2. No Aberrations

average rmsopd in wavelengths


FIGURE 4 HISTOGRAM OF AVERAGE RMSOPD UNIFORM DISTRIBUTION OF 1 DDE POINTS.

```
ST 2.4M CONFIGURATION, FIELD PT. .100, .000
    101 RAYS
```

| RMS OPD | .2432 |
| ---: | ---: |
| STREHL | -1.334 |
| CENTROID: |  |
| $X$ | -.3229 |
| $Y$ | .0000 |
| $Z$ | .0000 |
| CHIEF RAY: |  |
| $X$ | -.3351 |
| $Y$ | .0000 |

RANK 10
$P 1=-.1079021 E-01$

| P2 | $\begin{aligned} & \text { ZERNIKE } \\ & \text { P3 } \end{aligned}$ | COEFFICIENTS | P5 | P6 |
| :---: | :---: | :---: | :---: | :---: |
| $\times$ TILT | Y TILT | FOCUS | ASTIG(45) | ASTIG(0) |
| -1.4602? | -.264879E-07 | -.496520E-02 | .163329E-18 | -.642378E-04 |
| P7 | P8 | P11 | P12 |  |
| Y COMA | $x$ COMA | 3D-ORDER | 5TH-ORDER |  |
| . 182050E-0? | -. 495005 | .141003E-04 | .170063E-05 |  |
|  |  | DESIGN | AMETERS |  |


| \# | D | E DESIGN PARAMETERS | R | QN |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 10.0000000000000 | -1.00228977203300 | -1102.50000000000 | -1.00000 |
| 2 | DUMMY |  |  |  |
| 3 | -490.000000000000 | -1.49599512243700 | -135.464362850970 | -1.00000 |
| 4 | DUMMY |  |  |  |
| 5 | 640.000000000000 | -1.00000000000000 | 1000000000.00000 | 1.00000 |

```
                                    R1 = 120.000
                                    R4 = 699.507000000000
                                    EE = .340000 FL = 5760.00
```

*STOP* 0
Table 1. Zernike Coefficients with 0.1 Arc Min
Tilt About the X-Axis

ST 2. 4 M COMFIGURATION, FIELD PT. .000, .000 101 RAYS

| RMS OPD | $.7330 E-05$ |
| ---: | :--- |
| STREHL | 1.000 |
| CENTROID: |  |
| $X$ | .0000 |
| $Y$ | .0000 |
| $Z$ | .0000 |
| CHIEF RAY |  |
| $\chi$ | .0000 |
| $Y$ | .0000 |

RANK 10
$P 1=.3218636 E-04$


## REFERENCES

1. Barakat, Richard, Optimum Balanced Wavefront Aberrations for Radially Symmetric Amplitude Distributions: Generalizations of Zernike Polynomials, Journal of the Optical Society of America, Vol. 70, No. 6, June 1980, pp. 739-742.
2. Forsythe, George E., Michael A. Malcolm, and Cleve B. Moler, Computer Methods for Mathematical Computations, Prentice-Hall, 1977.
3. Freund, John E., Modern Elementary Statistics, 3rd ed., Prentice-Hall, 1967.
4. Hammersley, J. M., and D. C. Handscomb, Monte Carlo Methods, Wiley-Methuen and Co. LTD, 1964.
5. Ginzburg, A., Algebraic Theory of Automata, Academic Press, 1969.
6. Klein, Miles V., Optics, Wiley, 1970.
7. Knuth, Don, The Art of Computer Programming: Seminumerical Algorithms, Vol. 2, Addison-Wesley, 1969.
8. Koch, Donald G., A Statistical Approach to Lens Tolerancing, Society of Photo-Optical Instrumentation Engineers, Vol. l47, Computer-Aided Optical Design (1978), pp. 71-82.
9. Kovach, Ladis D., Computer-Oriented Mathematics, Holden-Day, 1964.
10. Scalzo, Frank, and Rowland Hughes, Elementary Computer-Assisted Statistics, rev. ed., Van Nostrand, 1978.
11. Kuo, Shan S., Computer Applications of Numberical Methods, Addison-Wesley, 1972.

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER <br> THE UNIVERSITY OF ALABAMA 

## POWER DISTRIBUTION FOR ELECTRON BEAM WELDING

| Prepared by: | Eugene Edwards, M.S.E.E. |
| :--- | :--- |
| Academic Rank: | Assistant Professor |
| University and Department: | Alabama A.\& M. University <br> Department of Engineering and <br> Engineering Technology, |
| NASA/MSFC: |  |
| $\quad$(Laboratory) <br> $\quad$ (Division) <br> $\quad$ Branch) | Materials and Processes <br> Process Engineering <br> Metals Processes |
| NASA Counterpart: | Arthur C. Nunes, Ph.D. |
| Date: | August 1, 1980 |
| Contract No.: | NGT-01-002-099 <br> (University of Alabama) |

# POWER DISTRIBUTION FOR ELECTRON BEAM WELDING 

By

Eugene Edwards, M.S.E.E.<br>Assistant Professor of Electrical Engineering Alabama A\&M University<br>Normal, Alabama

## ABSTRACT

Electron beams are being used to apply heat to metals for welding purposes.

In this report, the power distribution in the beam is analyzed. Experimental and digital computer techniques are used to evaluate the radial distribution of power detected by a wire probe circulating through the beam. Analytical models developed at MSFC of the Electron Beam-Workpiece interation contain two weld loss parameters in a function of the radial power distribution. The results from this research are intended to yield information on the mechanism by which these losses take place.

Additional information on the second loss mechanism is provided by measuring metal weight losses during welding. The power wasted on these weld losses can thus be estimated and distributed to obtain a more accurate weld penetration computation.

## ACKNOWLEDGEMENTS

I am most grateful to ASEE, NASA-MSFC (Mr. Marion Kent), and the University of Alabama (Drs. Barfield and Karr) for making it possible for me to obtain this research fellowship.

I especially would like to recognize Dr. Arthur Nunes, my NASA Counterpart, for outstanding contributions toward the success of this work. Also at MSFC's Materials and Processes Laboratory, I acknowledge Paul H. Schuerer (Division Chief) and Melvin McIlwain (Branch Chief, EH42). The technical assistance of Pete Smith (EH15), Clifton Green (EH42) and Howard Novak (EH42) was appreciated, as well as the inspiration of Willibald Prasthofer (EH44) and Carl Wood (EH44). Many thanks to Doris Flowers and Beverly Robinson for typing and reviewing this paper, respectively.

## LIST OF FIGURES

Figure No. Title Page
1 Illustration of Travel Direction ..... XXVI- 9
2 EB Gun, Simplified Illustration ..... XXVI-10
3 Weld Illustrations ..... XXVI-11
4 Typical Minicomputer Output on Beam Power and Weld Speed ..... XXVI-14Cross Section of a Spherical Beam Generation System.XXVI-17
Reduction of the Current Density Difference Between Center and Margin of an Irradiated Area of Diameter (D) by an Increased Electron Beam Diameter of $\mathrm{d}^{\prime} \mathrm{F}$ ..... XXVI-18
Creating a constant power density $\mathrm{P}^{\prime \prime}(x)$ on a Length 1-d'F by Time Linear Deflection of an Electron Beam d"F Over the Length 1 ..... XXVI-18
8 Experimental Setup ..... XXVI-20
9 Signals Triggering the Scope Due to Noise Present During Pumping of the Vacuum Chamber. ..... XXVI-2210
Wire Probe Characteristics ..... XXVI-25
Typical Beam Current Configuration (Actual Photo) ..... XXVI-26
Peak Beam Current for Two Pulse Signals ..... XXVI-27
Single Pulse Signal Distributed ..... XXVI-27
Radial Current Distribution (2 Dimensions) ..... XXVI-29
Radial Current Distribution (Profile) ..... XXVI-30

## LIST OF TABLES

Table No. Title Page
1 Power Relations for Evaporation Loss. ..... XXVI-15
2 Digital Computer Results on Current Density ..... XXVI-28

## NOMENCLATURE

AV - accelerating voltage
BC - beam current
BF - beam focus
D - diameter of the irradiated target
F - in welder's computer program, focus
FIL - filament
I - beam current
Ie - emission current
OK - degrees Kelvin
L - length of the wire probe
P - perveance
$P_{C} \quad$ - power lost to the crucible by conduction
$P_{d} \quad$ - power dissipated
$\mathrm{Pe}_{\mathrm{e}}$ - power loss for electrons not reaching the target
Peff - effective power
$\mathrm{P}_{\mathrm{g}}$ - power of the gun
$P_{i} \quad$ - power developed in the electron beam
$P_{1}, P_{1}$ - convergence points
$P_{L} \quad$ - power loss due to latent heat
$P_{n} \quad$ - loss of power to the evaporant
$\mathrm{Pr}_{\mathrm{r}}$ - power lost by heat radiation
$P_{v} \quad$ - power lost by electrons not striking the target
$\mathrm{P}_{\mathrm{x}} \quad$ - loss of power due to X -ray production
POT . - potentiometer
R - brightness of the beam
T - cathode temperature
$V_{a} \quad$ - acceleration voltage
$d_{a}$ - cathode diameter at point (a)
$d_{f}$ - beam's spot diameter
$\mathrm{d}_{\mathrm{k}}$ - cathode diameter at $k$
$d_{m} \quad$ - smallest diameter of the beam
$\mathrm{e} \quad$ - electron charge of $1.6 \times 10^{-9}$ coul.
$\mathrm{eV}_{0}$ - probable thermal energy
$j_{F} \quad$ - current density of the beam at the target
$j_{k} \quad$ - current density for beam current (J) at the cathode
$k \quad$ - constant, such as the $k$ th term, also $K$, as a constant
$k_{b} \quad$ - Boltzman constant
p - power density at the target
p'F - high power density
$\alpha$ - alpha, space-charge-limited current flow function
$\alpha_{L}$ - aperture at the target
$\gamma$ - gamma, aperture angle of marginal rays between the anode
$\bar{\gamma}$ - aperture of marginal rays at the position of the smallest beam cross section
$\Gamma$ - gamma, thickness of the wire probe
$\omega$ - omega, width of the wire probe
$\varnothing$ - phi, angle of current signal for computer input
$\pi-\mathrm{pi}, \simeq 3.14$
$\boldsymbol{P}$ - rho, beam target distance from motor-plate to some point on the probe
$\theta$ - theta, aperture angle of the cathode beam generator

## OBJECTIVES

The objective of this study is to investigate the power distribution of the electron beam in order to understand the loss mechanisms associated with the beam-workpiece interation. Another objective of this research is to apply techniques that will yield distribution and profiles of the beam such that power losses may be carefully examined.

## INTRODUCTION

## Electron Beam Welder (Sciaky Bros. Inc., Chicago, Illinois)

In an electron beam process, the part is welded by the Kinetic energy of a dense beam of accelerated electrons. As the electrons strike the workpiece, their Kinetic energy is converted to the thermal energy causing the metal to melt and fuse.

The device which emits, accelerates, and focuses the electrons on the workpiece is referred to as the electron beam gun. The gun can be mounted on a carriage to provide several axes of motion, or it can be held stationary while the work is driven, or both the gun and the work can be driven. See Figure 1 below for a simplified illustration.


FIGURE 1 - Illustration of Travel Direction

The gun assembly is housed in a chamber which is pumped down to a fine vacuum level, usually to about $1 \times 10^{-4}$ torr. or less. By operating in a vacuum the life of the filament (Figure 2) is protected and the electrons can be focused into a fine dense beam. The filament which is heated to a high temperature would be destroyed if exposed to air at the operating temperature. If there were a sufficient concentration of loose molecules in the chamber, the electrons would collide with these molecules. This would cause the beam to be diffused and the electrons to lose some of their Kinetic energy.

In Figure 2, the voltage of the bias allows the beam current to be changed independent of the accelerating potential (within the operating limits of the gun). Irrespective of the bias level, the maximum beam current cannot exceed the maximum current range of the gun. This range is determined by the mating elements installed in the gun. Maximum current is obtained at 0 bias. As the bias is increased, the current is reduced.


The spacing between the filament, cathode, and anode determines the internal resistance of the gun. This resistance is non-linear and varies directly with the distance between these elements. As the distance is increased, the beam current is decreased.

The model E-S9460 gun has 2 major ranges; 250 mA and 500 mA . A mating filament and cathode are provided for each range. In each of these 2 ranges, additional ranges can be obtained by inserting a spacer between the 2 main sections. The same anode is used in both major ranges.

Ranges of Gun E-S9460

| Range | 500 mA | Range $:$ | 250 mA |
| ---: | :--- | ---: | :--- |
| Filament: | C-K2501 | Filament: | C-G8615 |
| Cathode | B-K2501 | Cathode | B-H844 |
| $* * A n o d e$ | $:$ | B-M5740 | **Anode $:$ |

For additional ranges use: For additional ranges use: 350 mA W/0.050" spacer

150 mA W/0.100" spacer 100 mA W/0.200" spacer 50 mA W/0.500" spacer

For application requiring shallow penetration with narrow welds at low currents, a small hole cathode is used.

Electron beam welds posses several unique characteristics:
(1) Welds are produced in many cases 15 to 20 times narrower than comparable TIG/MIG welds. Notice the dotted isothermal lines of typical Tungsten Arc welding compared to Electron Beam welding (Figure 3).

(A) TUNGSTEN ARC

(B) ELECTRON BEAM

FIGURE 3 - Weld Illustrations
**The same cathode is used in both major ranges.
(2) Less thermal energy is transferred to the unmelted base metal and thus reducing residual stresses, warping, and increasing the strength of the wetd.
(3) Absorption of atmospheric gasses by the weld is eliminated by the vacuum system, thus highly oxidizable materials can be welded (Zirconium, Beryllium, etc.).

## POWER RELATIONS

The output power from the electron beam gun follows the general power formula of the form

$$
\begin{equation*}
P_{g}=V_{a} I_{e}, \tag{1}
\end{equation*}
$$

where, $P_{g}$ is the gun power
$V_{a}$ is the acceleration voltage
and $I_{e}$ is the emission current.
Power, $P_{d}$, is lost as the electrons are taken off of the anode. This dissipated power must be subtracted from the gun power.

Since all of the electrons will not reach the welding target the gun power becomes even smaller. The beam focus plays an important role in allowing more electrons to reach the target. Now, the effective power becomes

$$
\begin{equation*}
P_{e f f}=P_{g}-P_{d}-P_{e}, \tag{2}
\end{equation*}
$$

where $\mathrm{P}_{\mathrm{e}}$ is the power lost due to electrons not reaching the target.
For certain applications, a low precision gun with a diffused output beam may allow a little more than $50 \%$ of the emitted electrons to reach the target [1],

Energy is lost when the electrons collide with system components and target samples. Smith, reference [1], attribute some of this energy loss to:
o excitation and ionization of electrons, by electronelectron interaction, and displacement of atoms with the lattice by energy trnasferred to individual atoms. This occurs at energies of a few-hundred KEV.
o X-ray production by resonance absorption occurs at the high energy range. These losses are minor.
o inadequate focusing method creating minor loss of electrons from the beam.

0 vapor between the welding sample and gun can cause minor losses due to collisions.

0 back scatter of electrons reflected produces the main source of loss (25\%).

It is generally safe to assume that the effective power that is converted to heat the sample is

$$
\begin{equation*}
P_{\text {eff }}=0.60 P_{g} . \tag{3}
\end{equation*}
$$

Note that equation Eq. (3) is approximately the equation for average power. Information on evaporation losses will be presented later in this paper.

The important relationship between beam power and quality of the weld penetration can be obatained for various metals, metal sizes, and welding speeds. A computer routine is available that gives results similar to those shown in Figure 4 (NASA-MSFC).

A summary of the power-energy distribution is given by Hunt and Hughes [2] as the process heat balance:

$$
\begin{equation*}
P_{i}=P_{v}+P_{r}+P_{L}+P_{n}+P_{x}+P_{c} \tag{4}
\end{equation*}
$$

where, $P_{\mathbf{i}}$ is the power developed in the electron beam by acceleration
to the anode voltage.
$P_{V}$ is the power lost by electrons striking other than the target; e.g., filament-to-anode current (negligible),
$P_{r}$ is the power lost by heat radiation from the liquid evaporant surface (hard to determine).
$P_{L}$ is the power loss due to the latent heat of evaporation for the evaporant (available for most materials).
$P_{n}$ is the loss of power to the evaporant by ionization and secondary electron generation (due to vapor density and electron back-scattering).
$P_{x}$ is the power loss due to production of $X$-rays (negligible).
$P_{c}$ is the power loss to the crucible by conduction (very high).
Table I (reference [2]) on page XXVI-15 shows data on power requirements of typical sources evaporating various materials from a 4-inch-diameter water cooled copper crucible, fed from the bottom with rods of evaporant. This data implies that there appears to be an optimum beam density above and below which evaporation rate decreases for a given level of electron beam power at the evaporant surface.

XXVI-13

2219 Aluminum- -Plate Thịckness . 75 Inches
50\% Nonconductive Losses- $-10 \%$ to Nailhead
. 08 Inch Root Width


FIGURE 4 - Typical Minicomputer Output on Beam Power and Weld Speed

## TABLE 1

| Material | Power In | Evap. Rate | Specific Power <br> Requirement |  |
| :--- | :---: | :---: | :---: | :---: |
| Zirconium | 72 kw | $2.57 \mathrm{lb} / \mathrm{hr}$ |  | $28.0 \mathrm{kwh} / \mathrm{lb}$ |
| 20\% Cr S.S. | 70 kw | $9.38 \mathrm{lb} / \mathrm{hr}$ | $7.5 \mathrm{kwh} / \mathrm{lb}$ |  |
| Ti-6A1-4V | 70 kw | $5.17 \mathrm{lb} / \mathrm{hr}$ | $13.5 \mathrm{kwh} / \mathrm{lb}$ |  |
| Aluminum-Bronze | 80 kw | $8.96 \mathrm{lb} / \mathrm{hr}$ | $13.1 \mathrm{kwh} / \mathrm{lb}$ |  |
| Inconel 600 | 70 kw | $5.06 \mathrm{lb} / \mathrm{hr}$ | $13.8 \mathrm{kwh} / \mathrm{lb}$ |  |

## TABLE 1 - Power Relations for Evaporation Loss

## EQUATIONS OF THE BEAM CURRENT AND POWER DENSITY

When the E.B. gun is operating, its beam current follows the equation [3]

$$
\begin{equation*}
I=P V_{a}^{3 / 2} \tag{5}
\end{equation*}
$$

where, $V$ is accelerating voltage and $P$ is the perveance. Schiller, Heisig, and Lênk [3] used the equation below for perveance ( $P$ ) as well as the following equations including a deviation of the power density $\left(p_{f}\right)$ :

$$
\begin{equation*}
P=\text { perveance }=15 \times 10^{-6} \frac{(1-\cos \theta)}{\alpha^{2}}, \tag{6}
\end{equation*}
$$

where, $\alpha$ is a function of the space-charge-limited current flow depending on the relation,
$\alpha=f \frac{d_{k}}{d_{a}}$, between the cathode diameter
$d_{k}$ and the anode bore diameter $d_{a}$. The $f$ implies function of.
$\theta$ is the aperture angle of the cathode beam generator or the convergence angle of marginal rays in the beam generator (See Figure 5).

$$
\mathrm{eV}_{0}=\text { probable thermal energy }=\mathrm{k}_{\mathrm{b}} T=\frac{\mathrm{T}}{11,608} \mathrm{ev}, \text { (7) }
$$

where,
$V_{0}$ is thermic electron potential
e is the charge of $1.6 \times 10^{-19}$ coul.
$k_{b}$ is the Boltgmann constant in $\frac{\mathrm{ev}}{\mathrm{deg}}$.
$T$ is the cathode temperature ${ }^{\circ} \mathrm{K}$.
A Maxwellian energy distribution is assumed with usual emission temperatures, perhaps $3000^{\circ} \mathrm{K}$.

$$
\begin{equation*}
R=\text { brightness of the beam }=\frac{j_{k} V_{a}}{\pi V_{0}}=\frac{j_{k} e V_{a}}{\pi k_{b} T_{k}} \text {, } \tag{8}
\end{equation*}
$$

where,
$j_{k}$ is the current density for beam current (J) at the cathode.

$$
\begin{equation*}
p=\text { power density at target }=j_{F}^{\prime} \cdot V_{a} \tag{9}
\end{equation*}
$$

where,
$j^{\prime}{ }_{F}$ is the current density of the beam at the target and equals $R \pi \alpha_{L}^{2}$. Here, $\alpha_{L}$ is the aperture at the target.
An arrangement of Equations (7), (8), and (9) yields (a high power density $\mathrm{p}_{\mathrm{F}}$ )

$$
\begin{equation*}
p_{F}=\frac{j_{k}}{V_{0}} \alpha_{L}^{2} \cdot v_{a}^{2} \tag{10}
\end{equation*}
$$



FIGURE 5 - Cross Section of a Spherical Beam Generation System of Pierce-type (Schematically)
$\mathrm{d}_{\mathrm{k}} \quad=$ diameter of cathode
$d_{a} \quad=$ diameter of anode bore
$\theta \quad=$ aperture angle of the beam generator convergence angle of marginal rays in the beam generator

| $\gamma \quad=$ | aperture angle of marginal rays behind <br> the anode, considering scattering effect |
| ---: | :--- |
| $P_{i}^{\prime} \quad P_{1} \quad=$ | convergence points |
| $d_{m} \quad=$ | smallest diameter of the beam consider |
|  | the intrinsic space charge |$\quad$| $\bar{\gamma} \quad$ | aperture of marginal ray at the position <br>  <br> of the smallest beam cross section |
| ---: | :--- |



$$
\begin{aligned}
& d_{F}^{\prime}=D \\
& j_{F}^{\prime}\left(r=\frac{D}{2}\right)=0.37 j_{F}^{\prime}(r=0) \\
& d_{F}^{\prime \prime}=3 D \\
& j_{F}\left(r=\frac{D}{2}\right) \approx 0.9 j_{F}(r=0)
\end{aligned}
$$

FIGURE 6 - Reduction of the Current Density Difference Between Center and Margin of an Irradiated Area of Diameter D by an Increased Electron Beam Diameter $d_{F}^{\prime \prime}$


FIGURE 7 - Creating a Constant Power Density $p^{\prime \prime}(x)$ on a Length $1-d_{F}^{\prime}$ by Time Linear Deflection of an Electron Beam d ${ }_{F}^{\prime \prime}$ over the Length 1. (The range over which power density decreases to $0.63 \mathrm{p}^{\prime \prime}(\mathrm{x}=0)$ is $\left.=\mathrm{d}_{\mathrm{F}}^{\prime}\right)$

Reference [3] also used Figures 6 and 7 to show that time-linear deflection is one coordinate results in a power density distribution (deflection $X= \pm .5$ ) of constant values in certain ranges. Power outside this range decreases and is a power loss. The term $d_{F}$ is the beam's spot diameter and $D$ is the target diameter (irradiated).

## TECHNIQUES FOR MEASURING AND EVALUATING THE ELECTRON BEAM

This section explains the experimental and computer aided approaches used to evaluate the beam configuration.

The initial problem task for this section is that of determining what available method and apparatus to use to take measurements of the beam. The rotating probe (or wire) approach was chosen. This method utilizes a small heat protected motor to turn the wire probe through the beam. The signal obtained on the probe is fed into a storage oscilloscope. Pictures of the signal for various operating conditions of the system are taken from the scope and analyzed. The technique has the following advantages:
(a) simple
(b) apparatus available and not refined
(c) beam current is measured
(d) beam diameter is obtainable

The wire probe did not obtain appreciable damage from the beam and adequately allowed suitable measurements to take place.

Adjusting the accelerating voltage (thus power output) appears not to change the shape nor the general dimensions of the beam.

The experimental setup used to measure the beam is shown in Figure 8. The oscilloscope was used to monitor and record not only the beam signals but general signals (noise) in the vacuum. During pump down the oscilloscope was triggered by noise signals equivalent to those shown in Figure 9.

Since this electron beam welding machine receives its instructions from a computer controlled system, the sample computer program for a typical set of data inputs is shown on page XXVI-23.

When the beam comes into contact with the rotating wire probe, the electrons that are not backscattered or secondary-emitted are received as the current of the beam (perhaps $50 \%$ ). The current through the wire has the common variation in magnitude at various angles of emergence. The surface contour of the current is obtained from the electron distribution. Care must be taken in determining the needed current signal configuration and associate information instead of bad data. The collecting wire probe has a symmetrically fixed amount of wire extending from two sides of the plate affixed to the


FIGURE 8 - EXPERIMENTAL SETUP

(a) Noise in the E.B. Welder Recorded at 2 Milliseconds Fullscale (above)

(b) Noise in the E.B. Welder Recorded at 5 Milliseconds Fullscale (above)

FIGURE 9 - Signals Triggering the Scope Due to Noise Present During Pumping of Vacuum Chamber

## TYPICAL DATA RESULTS

```
\rho=2.833 INCHES
\omega=.048 INCHES
\Gamma=.003 INCHES
BF (60 KV) 5.00
FIL 3.50
AV 30 KV
BC 50 MA
THE SAMPLE COMPUTER PROGRAM FOR THE WELDER IS SHOWN BELOW.
```

$\left.\begin{array}{ll}\hline & \text { FILE RECORD } \\ \text { NASA EB WELDER }\end{array}\right]$ PAGE 1

A TABLE OF COMPUTER INPUT DATA ( $\delta, I, E T C . .$.$) IS DEVELOPED FROM FIGURE 11$ FOR USE ON THE DIGITAL MINICOMPUTER.
rotor of the motor (as was shown in Figure 8). The motor was shielded from the beam and from heat by a thin sheet of metal. The current collected was sent through a load resistor and the output from this load to the oscilloscope.

In order to adequately measure the beam current, the wire probe dimensions were carefully measured (Figure 10). The beam was also precisely applied to a target area of the probe at a chosen radius from the center of the plate of the motor.

Figure 11 shows the typical current signal and can be plotted versus angle position. With the given beam characteristics or welding operation status, the signal profile is always of a similar shape. To avoid possible aggravation of the signal, it should be saved on the scope as soon as a clear configuration is observed. The scope will automatically trigger to the signal prior to being available for saving the signal. Interferences during signal generation, such as surfaces disturbances not allowing adequate distribution of beam electrons, were considered minor.

Two cases representing the output current signal are in Figures 12 and 13. The first case (Fig. 12) is the peak beam current for two pulse signals. The second case (Fig. 13) is an expanded version of one of the pulse signals distributed with a wider profile. Several sweep time divisions of the scope were investigated for many conditions until satisfactory signals were realized.

The current vs. probe angle was plotted and used as inputs for the computer routine that calculates the current density per unit radius. Additional inputs to the computer included (a) the distance from the center of the plate of the motor to a target point on the wire, (b) the width of the wire probe (receptor), (c) the total beam current, and (d) the number of current-angle data points.

As can be seen from Table 2 (page XXVI-28) and Figures 14 and 15, a current density profile can be obtained. The radial distribution is modulated by a strong periodic error signal of a sort known to arise from small errors (about $0.1 \%$ ) in data. Work is continuing to obtain current density profiles.

The measurements taken from this experimental setup show that much more beam power is lost than can be explained from measurements taken from the wire probe, unless very large amounts of energy are carried away per secondary emission or backscattered electrons. This justifies the need to continue investigating the evaporation power loss and other power losses accounting for this reduction of total beam power. This phenomenon occurs quite often when samples are being welded. Some information is available on secondary emission losses from a 1974 study conducted at NASA-MSFC.


```
\(\mathrm{L} \simeq 8^{\prime \prime}\)
\(\Gamma=\) Thickness .003"
\(\omega=\) Width .048"
\(\rho=\) Beam Target Distance 3-3/4"
```

WELDER OPERATION STATUS:
Beam Focus - BF ( 60 Kv ) 5.00
Filament Setting - Fil 3.50
Accel. Voltage, Beam Current - 30 Kv , BC 50 ma

Actual size of the Wire Probe ( $\Gamma$ not shown)

FIGURE 10 - Wire Probe Characteristics


FIGURE 11 - Typical Beam Current Configuration (actual photo)

## EXPERIMENTAL RESULTS



9Z-Inxx
FIGURE 12 - Peak Beam Current for Two Pulse Signals


FIGURE 13 - Single Pulse Signal Distributed

THE CHARACTERISTICS OF THE TUNGSTEN WIRE ARE
$\rho$, THE DISTANCE FROM THE CENTER OF THE MOTOR PLATE TO A TARGET POINT ON THE WIRE
$\omega$, THE WIDTH OF THE WIRE
$\Gamma$, THE THICKNESS OF THE WIRE
THE E.B. WELDER WAS COMPUTER CONTROL FOR AN OPERATION STATUS FOR VALUES OF

BF, BEAM FOCUS
FIL, FILAMENT SETTING
AV, ACCELERATING VOLTAGE
BC, BEAM CURRENT
TYPICAL DATA RESULTS ARE SHOWN ON THE NEXT PAGE.

## TABLE 2

| ANGLE (DEG) | CURRENT <br> (MA) | RADIUS <br> (IN) | CURRENT DENSITY (AMPS/SQ IN) |
| :---: | :---: | :---: | :---: |
| . 017 | 26.32 | 8.40541E-04 | -832.032 |
| . 031 | 28.15 | 1.53275E-03 | 1153.54 |
| . 053 | 28.97 | 2.62051E-03 | -571.086 |
| . 074 | 31.79 | 3.65883E-03 | 301.177 |
| . 093 | 33.61 | 4.59827E-03 | . 0935703 |
| . 11 | 35.43 | 5.43882E-03 | 214.231 |
| . 134 | 37.25 | 6.62548E-03 | -3.99142 |
| . 154 | 39.08 | 7.61436E-03 | 104.89 |
| . 178 | 40.89 | 8.80103E-03 | 75.7133 |
| . 211 | 42.72 | . 0104327 | 66.7768 |
| . 248 | 44.54 | . 0122622 | -57.2598 |
| . 266 | 46.36 | . 0131522 | 246.266 |
| . 327 | 48.18 | . 0161683 | 6.38338 |
| . 386 | 50 | . 0198856 | 45.3961 |

Ratio of Total Current Calculated to Total Beam Current. 1

TABLE 2 - Digital Computer Results on Current Density


FIGURE 14 - Radial Current Distribution (2 dimensions)


FIGURE 15: Radial Current Distribution (profile)

## CONCLUSIONS

At the present time we are able to capture by experimental techniques the current configuration for an electron beam. The computer routines allow us to theoretically determine the beam characteristics on the workpiece for a given power profile. Techniques developed and described in this report can be valuable for gathering information needed to assess the nature of weld power losses.

Future work should be done in order to relate accurately the total beam characteristics and power distribution to weld geometry. However, we have successfully completed the experimental task of obtaining the signal configuration which determines the cross sectional power distribution in the electron beam.

The results of Figures 11 and 13 compare with those obtained by Sayegh and Dumonte [4] [5].

An analysis for obtaining the radial current density and the profiles in Figures 14 and 15 is available but small errors in data (or deviations from a radial distribution) apparently cause strong periodic error modulations in the radial distribution.

## RECOMMENDATIONS

1. Techniques for obtaining more information on electron reflection should be investigated.
2. Controlled use of the welding system such that the time wasted in equipment venting and pump down is minimized.
3. Advanced studies should be done on the trajectories of the beam (both gaussian and nongaussian type).
4. Curvilinear distribution analyses should be considered for further study.
5. Relate the diameter of the beam and the power distribution on a welding sample in items 3 and 4 above.
6. Determine if the target surfaces cause any disturbance in the distribution of the beam's electrons on the wire probe.
7. Determine the optimum setting for the E.B. welder in order to obtain the optimum power density profile.

## REFERENCES

1. Smith, H. R., "Principles of Electron Beam Technology," 2nd Electron Beam Processing Seminar (R. M. Silva, ed.), Universal Technology, Dayton, Ohio, June 1972.
2. Hunt, Charles d'A. and Hughes, J. L., "Principles of Electron Beam Evaporation," 2nd Electron Beam Processing Seminar (R. M. Silva, ed.), Universal Technology, Dayton, Ohio, June 1972.
3. Schiller, S., Heisig, U., and Lenk, P., "Theory and Development of Electron Beam Sources," 2nd Electron Beam Processing Seminar (R. M. Silva, ed.), Universal Technology, Dayton, Ohio, June 1972.
4. Sayegh, G. and Dumonte, P., " 100 KW EB Gun for Welding-Design, Realization and Applications," 4th International Electron Beam Processing Seminar (R. M. Silva, ed.), Universal Technology Corporation, Dayton, Ohio, April 1976.
5. Sayegh, G. and Dumonte, P., "Theoretical and Experimental Techniques for Design of Electron Beam Welding Guns, Effects of Welding Conditions on Electron Beam Characteristics, "3rd Electron Beam Processing Seminar Proceedings, Universal Technology Corporation, Dayton, Ohio, March 1974.

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM 

MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA

MEASUREMENT OF VELOCITY FIELDS IN FLUID DYNAMICS EXPERIMENTS

| Prepared by: | Danie1 E. Fitzjarrald |
| :--- | :--- |
| Academic Rank: | Research Associate <br> University and Department |
| Geophysical Fluid Dynamics <br> Institute <br> Florida State University <br> $\quad$ (Laboratory) <br> $\quad$ (Division) | Space Sciences Laboratory |
| MSFC Counterpart: | Atmospheric Sciences <br> Fluid Dynamics |
| Date: | George H. Ficht1 |
| Contract No.: | August 6, 1980 |

## by

Daniel E. Fitzjarrald Geophysical Fluid Dynamics Inst. Florida State University Tallahassee, FL

ABSTRACT

In many fluid dynamics experiments in which the motion has both spatial and temporal variations, it is necessary to measure the entire two-dimensional field of velocity at an instant. Two methods to accomplish this measurement have been investigated. In both cases the emphasis has been on using a microcomputer-video system in the analysis of data. The first mothod is the traditional one in which tracers are introduced into the fluid and their position compared at two closely spaced times. The second method involves scattering coherent light in the fluid and obtaining motion by analyzing the multipleexposed speckle pattern recorded on photographic film. Both of these velocity field methods can be used in the geophysical fluid flow and space processing experiments in progress at NASA.

1. Introduction

In many fluid dynamics experiments both spatial and temporal variations of the flow patterns are important. This is particularly true in experiments modelling geophysical flows and in space-processing experiments, both in progress at NASA. With conventional instrumentation the experimentalist must rely on the use of fine wire probes, laser-Doppler measurements, or other similar techniques that are all limited to a single point in space. It is possible to move the measuring point around in the fluid, but the technique is never satisfactory when there are significant time variations in the flow field.

Streak photography has often been used in such cases, i.e., taking photographs of the displacement of tracers in the velocity field. Data analysis is quite cumbersome, however, whenever there is much variation in the flow field, or when a dense coverage of data points is required.

It is the goal of this summer's research to investigate a promising alternative to conventional velocity field measurement. The technique to be investigated will be termed light-speckle techniques, and all are closely related to streak photography. The key in each case is that the tracers and time intervals are chosen so that optical correlation techniques can be used to recover velocities from the data.
2. Light-Speck1e Velocity Field Measurement

Although there are several variations in details, these techniques are all basically the same. Light from scatterers that move with the fluid is recorded at closely-spaced time intervals on photographic film. The optical pattern from the first instant is slightly changed by the fluid motion at succeeding time intervals. Thus, when the multiple-exposed negative is interrogated by a narrow beam of coherent light, the correlation between patterns recorded at succeeding times is indicated by a pattern of regular fringes. The spacing of the fringes tells the amount of displacement, i.e., the time interval and the orientation of the fringes tells the direction of displacement. Thus, by illuminating the fluid with
a thin sheet of light and viewing it at a right angle to the plane of illumination, an entire two-dimensional field of velocity can be obtained at an instant.

### 2.1 Generating the Optical Pattern

The optical pattern that is to be displaced by the fluid can be generated either by using the coherent nature of laser light with very small scatterers, or with white light and larger scatterers. The former method has been used by Simpkins and Dudderer (1978) and is based on the fact that laser light appears grainy when diffusely reflected off a surface. The grainy appearance is caused by constructive and destructive interference of coherent light scattered off roughness elements the same size as the wavelength. These "objective speckles" are a function of the size of the scatterers, while the "subjective speckles" seen by the observer or recorded on film are dependent on the aperture size of the lens through which they are viewed. When the displacement of the recorded subjective speckles due to fluid motion is between one and twenty speckle diameters the correlation will yield a good fringe pattern. The fringes will also be enhanced if more than two optical patterns are recorded on the same negative.

A useable optical pattern can also be obtained by using larger scatterers illuminated by white light (Chiang and Asundi, 1979), or by a laser (Meynart, 1980). In the latter case the spatial coherence of the laser light is used only to obtain a thin light sheet of high power density. Again, multiple exposures increase the visibility of the fringe.

### 2.2 Interrogating the Negative

Once the multip1c-exposed negative has been obtained, there are two methods to recover the velocity field. Illuminating the negative point-by-point with a small coherent beam yields fringes that come from a small portion of the flow field where the velocity was constant. The negative is thus interrogated at each point in the field that the velocity is required.

A whole-field method of analysis can be obtained by illuminating the entire negative and using a spatial filtering setup (see Meynart, 1979) that consists of two lenses (Fourier transformers) and an off-axis pinhole
(spatial filter). The result is the whole velocity field described in isolines of the component of velocity depending on the orientation of the spatial filtering hole.

There are advantages to both these interrogation methods. The first yields results that are more precise, with the limits of precision depending on the quality of the fringes that are obtained. The wholefield method is simpler and much quicker.
3. Use of Microcomputer in Analysis

In order to speed up analysis to make possible the evaluation of fast-changing geophysical and spaceprocessing fluid flows, it is desirable to automate the data analysis as much as possible. Using a micro-computer-TV system it should be possible to accomplish the point-by-point interrogation of the multipleexposed negatives automatically. The microcomputer will be programmed to count the fringes, determine orientation, and then move the film transport on to the next position. In addition, the TV system can determine velocity by simply subtracting successive positions of large scatterers to serve as a check on the velocities obtained using the speckle techniques.
4. Planned Work and Work Accomplished to Date

At this writing the experimental work is just getting into progress. Multiple exposures have been taken in a small water flume and thermal convection cells illuminated by a 1 watt $C W$ argon laser. The pulse repetition rate and flow rate have been varied to determine optimum conditions for obtaining good fringes.

In addition considerable program development work has been accomplished on the TV-microcomputer system. Picture enhancement and pattern recognition programs have been written and debugged, so that they can be put together to measure fringes and track individual scatterers.
5. Conclusions

The light-speckle velocity field techniques described above offer tremendous possibilities toward the understanding of geophysical and space-processing fluid flow experiments. The ability to obtain an entire two-dimensional velocity field at an instant makes these techniques extremely powerful.

The experiments and computer program development in progress should demonstrate the utility of these lightspeckle techniques.

## References

Chiang, F. P., and A. Asundi, 1979. White light speckle method of experimental strain analysis. Applied Optics, 18, 409-411.

Meynart, R., 1980. Equal velocity fringes in a Rayleigh-Benard flow by a speckle method. Applied Optics 19, 1385-1386.

Simpkins, P. and T. Dudderer, 1978. Laser speck1e measurements of transient Benard convection. J. Fluid Mech., 89, 665-671.

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM MARSHALL FLIGHT CENTER THE UNIVERSITY OF ALABAMA

ON THE CONTROL AND STABILITY OF THE PINHOLE-OCCULTER FLEXIBLE BOOM FACILITY

PREPARED BY:
ACADEMIC RANK:
COLLEGE AND DEPARTMENT:

NASA/MSFC: (OFFICE) (DIVISION) (BRANCH)

MSFC COUNTERPART:
DATE:
CONTRACT NO:

Michael Greene, PhD
Assistant Professor
University of Alabama
Department of Electrical Engineering

Preliminary Design
Subsystems Design
Navigation and Control
J. R. Parker

July 15, 1980
NGT-01-002-099
(University of Alabama)

## ACKNOWLEDGMENTS

The author wishes to thank the members of the Navigation and Control Branch of the Subsystems Design Division of the Program Development Preliminary Design Office. Special thanks to Dave Shultz, branch chief; Joe Parker, my counterpart; Chris Rupp, Joe Howell, and, P. .D. Nicaise, perpetrator of practical jokes. A special thanks goes to Jenny Reisz for all the excellent and timely programming work and to Joe Dabbs, project leader for pinhole-occulter.

## TABLE OF CONTENTS

Page
Introduction ..... 2
Background. ..... 2
Results ..... 4
Conclusions and Recommendations ..... 7
References ..... 9
List of Figures Title Page
Figure 1: Pinhole occulter facility ..... 11
Figure 2: Pinhole occulter facility configuration ..... 12
Figure 3: Boom frequency as a function of tip mass and boom length ..... 13
Figure 4: Simplified block diagram of pinhole control system ..... 14
Figure 5: Reduced block diagram of pinhole control system ..... 14
Figure 6: Open loop root location versus tach. gain ..... 15
Figure 7: Root locus for uncompensated system as a function of forward gain ..... 17
Figure 8: Root locus for compensated system as a function of forward gain ..... 17
Figure 9: Error response of boom tip to thruster firing and gravity gradients. ..... 19
Figure 10: Step response ..... 19
Figure 11: Frequency response of fixed boom and rate compensated boom ..... 20
Figure 12: Frequency responsc of the uncompensated gimbal driven boom system ..... 21
Figure 13: Frequency response of the compensated gimbal driven boom system ..... 22
Figure 14: Frequency response of the closed loop control ratio ..... 23

ON THE CONTROL AND STABILITY OF THE PINHOLE-OCCULTER FLEXIBLE BOOM FACILITY

By<br>Michael Greene, PhD<br>Assistant Professor of Electrical Engineering University of Alabama<br>Tuscaloosa, AL


#### Abstract

The dynamics and control of the Shuttle based, gimbal mounted, boom pointing system for use in the x-ray pinhole experiments were studied. The overall system stability was studied, and root locus compensation was designed to achieve stability and pointing stability requirements. This single axis control study found the compensated system to be stable and capable of pointing to within 6.5 arc sec while the Shuttle rocks in its $0.1^{\circ}$ deadband. The system can follow step (position) inputs with zero error and ramp (velocity) inputs with a steady state error coefficient of $4.5 \mathrm{sec}^{-1}$.


## Introduction

The pinhole-occulter system is a Space Shuttle based experiment for the production of hard X-ray images taken primarily from the Sun. The system is basically a pinhole camera utilizing a deployable $50-\mathrm{m}$ flexible boom for separating the pinhole from the recording devices located in the Shuttle as seen in Figure 1. At the distal end of the boom from the Shuttle is a 50 kg mask containing two sets of pinholes and two chronograph shields as seen in Figure 2. At the proximal end the detectors are located and mounted, along with the boom, to a gimbal pointing system (either IPS or AGS) mounted in the Shuttle Payload Bay.

The mask of Figure 2 must be pointed at the $X$-ray source with a pointing stability of less than 10 arc sec to align the axes of the detectors with the pinholes and shields. Failure to do so will result in a blurring of the images on the detectors and a loss of resolution. Being a Shuttle based experiment, the system will be subjected to the disturbances of the Shuttle. The worst of these is thruster firing for orbit correction; Shuttle uses a bang-bang thruster control system to maintain orbit to within $\pm 0.1^{\circ}$. Other disturbances include man motion, motion induced by other systems, and gravity gradient torques.

The control system of the pointing mount senses both position and velocity of the mask tip and uses these to accurately point the boom. The AGS employing perfect sensors was used in this study of the single axis ( $y$-axis) control system. The extension of this study to a three-axis system will be the topic of a later study.

## Background

The structural analysis of the boom [ $1,2,3$ ] revealed the seven mode shapes and natural frequencies shown in Table 1. These natural frequencies varied with both boom length and tip mass as seen in Figure 3. For a $50-\mathrm{m}$ boom with a 50 kg tip mass, the first two natural frequencies are $1 \mathrm{rad} / \mathrm{sec}$ for deflections about the x and y axes, respectively. The third natural frequency is $17.1 \mathrm{rad} / \mathrm{sec}$ for z axis motion. For x and y motions the next two natural frequencies are $17.4 \mathrm{rad} / \mathrm{sec}$ and $59.9 \mathrm{rad} / \mathrm{scc}$, respectively.

For $y$-axis control (torques about $y$-axis), the equations of motion for the boom, assuming a damping factor of 10 percent, are [1]:

$$
\begin{align*}
& \ddot{\eta}_{1}+.2 \dot{\eta}_{1}+\eta_{1}=-3687.93 \ddot{\Pi}_{2}  \tag{1}\\
& \dddot{\eta}_{2}+.2 \dot{\eta}_{2}+\eta_{2}=0 \tag{2}
\end{align*}
$$

$$
\begin{align*}
& \dddot{n}_{3}+3.42 \dot{\eta}_{3}+292.4 \eta_{3}=-.018396 \ddot{\theta}_{2}  \tag{3}\\
& \ddot{\eta}_{4}+3.46 \dot{\eta}_{4}+300 \eta_{4}=1250.734 \ddot{\theta}_{2}  \tag{4}\\
& \ddot{\eta}_{5}+3.46 \dot{n}_{5}+300 \eta_{5}=0  \tag{5}\\
& \ddot{n}_{6}+11 \dot{\eta}_{6}+3014 \eta_{6}=52.3905 \ddot{\theta}_{2}  \tag{6}\\
& \ddot{\eta}_{7}+11 \dot{\eta}_{7}+3014 \eta_{7}=0  \tag{7}\\
& \theta_{\mathrm{TIP}}=2.895 \times 10^{-4}{ }_{\eta_{1}}-8 \times 10^{-10} \eta_{3}+5.31 \times 10^{-5} \eta_{4} \\
& \tag{8}
\end{align*}
$$

Taking the Laplace transform of equations (1) through (7) and inserting into equation (7), we obtain:

$$
\begin{align*}
0_{\mathrm{TIP}}(\mathrm{~S})= & {\left[\frac{-1.07 \mathrm{~S}^{2}}{\mathrm{~S}^{2}+.2 \mathrm{~S}+1}+\frac{1.5 \times 10^{-11} \mathrm{~S}^{2}}{\mathrm{~S}^{2}+3.42 \mathrm{~S}+292.4}+\frac{.07 \mathrm{~S}^{2}}{\mathrm{~S}^{2}+3.465+300}\right.} \\
& \left.+\frac{1.57 \times 10^{-3} \mathrm{~S}^{2}}{5^{2}+11 \mathrm{~S}+3014}+1\right] 0_{2}(\mathrm{~S}) \tag{9}
\end{align*}
$$

Ignoring terms 2 and 4 and reducing, the ratio of input motion to output motion for the boom is obtained:

$$
\begin{align*}
\frac{{ }^{\theta} \mathrm{TIP}^{(S)}}{\theta_{2}(\mathrm{~S})}= & \frac{-.07\left(\mathrm{~S}^{2}-2.865-14.28\right)}{\mathrm{S}^{2}+.2 \mathrm{~S}+1}+\frac{.07 \mathrm{~S}^{2}}{\mathrm{~S}^{2}+3.48 \mathrm{~S}+300}  \tag{10}\\
& =\frac{-.0289\left(\mathrm{~S}^{3}+688.17 \mathrm{~S}^{2}-2194.12 \mathrm{~S}^{1}-10380.62\right)}{\mathrm{S}^{4}+3.67 \mathrm{~S}^{3}+301.69 \mathrm{~S}^{2}+63.47 \mathrm{~S}^{1}+300} \tag{11}
\end{align*}
$$

In its factored form, this transfer function for the boom is:

$$
\begin{equation*}
G_{B}(S)=\frac{-.0289(S+2.63)(S-5.88)(S+671)}{\left(S^{2}+.2 S+1\right)\left(S^{2}+3.46 S+300\right)}=\frac{N_{B}(S)}{D_{B}(S)} \tag{12}
\end{equation*}
$$

The block diagram for the system is shown in Figure 4. The pointing mount uses a position plus integral control with a velocity feedforward [4]. For the sake of clarity, the velocity feedforward will be considered separately from the forward transfer function of the Gimbal Pointing System which is:

$$
\begin{equation*}
G_{G}(S)=\frac{K_{p}(S+1.2)}{S} \tag{13}
\end{equation*}
$$

The velocity feedforward and the boom comprise a loop which can be reduced to a new forward transfer function, $G_{B S}(S)$, to treat the system as a unity feedback system as seen in Figure 5. The new forward transfer function is given by:

$$
\begin{align*}
& G_{B S}(S)=\frac{G_{B}(S)}{1+K_{R} S_{B}(S)}=\frac{N_{B}(S) / D_{B}(S)}{1+K_{R} \mathrm{KN}_{B}(S) / D_{B}(S)}  \tag{14}\\
& G_{B S}(S)=\frac{N_{B}(S)}{D_{B}(S)+K_{R} S N_{13}(S)} \tag{15}
\end{align*}
$$

From equation (15) we can see that the open loop eigenvalues can be arbitrarily set by adjustment of the rate gain, $\mathrm{K}_{\mathrm{R}}$. The closed loop eigenvalues will then be set by means of the compensator, $G_{C}(S)$, and the forward gain, $\mathrm{K}_{\mathrm{P}}$.

## Results

The open loop eigenvalues were set by adjustment of the rate feedforward gain, $\mathrm{K}_{\mathrm{R}}$. The plot of these roots (eigenvalues) as $\mathrm{K}_{\mathrm{R}}$ is increased from zero is shown in Figure 6. At a gain between 0.1 and 0.2 , the system goess unstable. At a gain of 0.07 the open loop roots are:

$$
S_{1,2}=-.14 \pm j .98
$$

ànd,

$$
S_{3,4}=-1.23 \pm \mathrm{j} 17.38
$$

Yielding the new transfer function:

$$
\begin{equation*}
G_{B S}(S)=\frac{-.02896(S+2.63)(S-5.88)(S+671)}{\left(S^{2}+28 S+.99\right)\left(S^{2}+2.46 S+303\right)} \tag{16}
\end{equation*}
$$

The major result of this rate compensation (which is built into the pointing mount) is to increase the damping ratio of the first quadratic root and decrease the damping of the second. This gives the basic system transfer function:

$$
\begin{equation*}
G_{S Y S}(S)=\frac{K^{1}(S+1.2)(S+2.63)(S-5.88)(S+671)}{S\left(S^{2}+.28 S+.99\right)\left(S^{2}+2.46 S+303\right)} \tag{17}
\end{equation*}
$$

Setting $G_{c}(S)=1$, we can again plot the root locus for the system as $K^{1}$ is varied ( $\mathrm{K}^{1}=-.02896 \mathrm{~K}_{\mathrm{p}}$ ). This plot is shown in Figure 7. At a gain between -0.2 and -0.03 the system goes unstable and oscillates; the system is unstable for all values of positive gain. At a gain of -0.2 the dominate roots have a time constant of about 8 seconds and would be extremely underdamped with $\xi_{=}=0.01\left(\mathrm{~S}_{1,2}=-0.13 \pm \mathrm{j} 12.3\right)$. This is not acceptable;
compensation is required.

The compensation chosen was:

$$
\begin{equation*}
G_{c}(S)=\frac{\left(S^{2}+2.46 S+303\right)(S+2)}{(S+671)(S+10)^{2}} \tag{18}
\end{equation*}
$$

This compensator cancels the open loop roots at $s=-1.23 \pm j 17.38$ and the zero at $S=-671$. If the compensator does not cancel the complex roots exactly it is better to undersize the real part of the compensator roots than oversize them. The response will contain terms of $\mathrm{Ae}^{-1.23} \mathrm{COS}$ $(17.38 t+\phi)$ but $A$ will be extremely small due to the proximity of the zeroes to these roots.

The basic idea of the compensation is to cause the roots nearest the imaginary axis (first mode) to move to the left as gain is increased while higher order modes (roots) move to the right. This allows for a high value of the gain which produces low steady state ramp errors and a wide passband. Ideally a type 2 system is desired but due to the zero in the right half $s$-plane, this is impossible.

Using the compensator of equation (18), a new $G_{\text {total }}$ is obtained:

$$
\begin{equation*}
G_{\text {total }}(S)=\frac{K^{1}(S+1.2)(S+2)(S+2.6)(S-5.88)}{S\left(S^{2}+.28 S+.99\right)(S+10)^{2}}=\frac{K^{1} N(S)}{D(S)} \tag{19}
\end{equation*}
$$

With unity feedback, the closed loop control ratio is:

$$
\begin{equation*}
\frac{C}{\mathrm{R}}=\frac{\mathrm{G}}{1+G}=\frac{K^{1} N(S) / D(S)}{1+K^{1} N(S) / D(S)} \tag{20}
\end{equation*}
$$

which reduces to

$$
\begin{equation*}
\frac{C(S)}{R(S)}=\frac{K^{1} N(S)}{D(S)+K^{1} N(S)} \tag{21}
\end{equation*}
$$

The plot of the closed loop roots is seen in Figure 8. At a gain of approximately -16 , system becomes unstable. At a gian of -12 , the dominant root is $S=-0.75$ and the ramp error coefficients is $K_{1}=4.5 \mathrm{sec}^{-1}$. For a ramp input of $0.01^{\circ}$, the steady state error for the system would be 8 arc sec.

The error response of the system to the Shuttle's thruster firing and drift due to gravity gradient is seen in Figure 9. The peak error response is 3.6 arc sec as the Shuttle drifts through its limit cycle and 6.5 arc sec when one thruster is firing. The maximum gimbal torque required to oppose the thruster acceleration is $17.9 \mathrm{NT}-\mathrm{m}$ while the maximum torque during drift is $8.3 \mathrm{NT}-\mathrm{m}$.

The step response of the system to a $1^{\circ}$ step is shown in Figure 10. Within 5 sec of the step, the tip angle is within 104 percent of steady state. The peak overshoot is 42 percent occurring at 7 sec . The response is within $\pm 1$ percent of steady state within 16 sec . Steady state error to step commands is zero. The torque required at the gimbal motors for this response exceeds the rated $33 \mathrm{NT}-\mathrm{m}$ during the first 12 sec . The torque required to initially accelerate the boom is up to 182 NT-m/degree. This limits step commands to less than 0.18 degree.

The frequency responses for the boom and rate compensated boom are seen in Figure 11. The responses of the fixed boom is a series of single order zeroes and quadratic poles starting at 1 radian/sec ( .16 Hz ). The boom is itself stable having an 18 dB gain margin and a $35^{\circ}$ phase margin. The effect of the gimbal system rate compensation is to increase the damping ratio of the first mode and decrease the damping of the second. At a $K_{r}=0.07$, stability is unaffected.

The frequency response of the gimbal mounted boom system is shown in Figure 12 for a gain of 16. At this gain, the system is unstable having a gain margin of -50 dB . And a phase margin of nearly $-90^{\circ}$. For a stable system, the gain would have to be reduced drastically and would yield a slow, low bandpass system with high ramp following errors.

The frequency response of the compensated system is shown in Figure 13 for a gain of 16 . At this value of gain, the system is just unstable with both a phase margin and a gain margin of zero. Reducing the gain of 2.4 dB to a gain of 12 , the system becomes stable and has the closed loop frequency response of Figure 14. The system bandpass is 2.5 Hz and it will follow up to 0.2 Hz signal without error or phase shift. This is the basic reason the system can tolerate the drift of the Shuttle in its deadband and the firing of one thruster. The drift of the Shuttle is very nearly a 0.01 Hz signal while the turnaround at the deadband is like a 1 Hz signal. The system response to thruster firing is not frequency (dynamics) limited but torque limited by the gimbal motors.

## Conclusions and Recommendations

As compensated, the gimbal mounted boom system is stable with fast dynamics and low following errors. The system is relatively unaffected by the drift of the Shuttle in its deadband and by one thruster firing (maximum error 6.5 arc sec ). The system is limited by the torque produced by the gimbal motors not by dynamics.

This study only considered motions about the $y$-axis, not $x$-axis or z . Motion about the x -axis should be very similar to that of the y axis except modes shapes 2 and 5 would predominate. The control system for x -axis should be nearly identical to that of the y -axis. Motion about the z-axis is twisting motion of primary the third mode (see Table 1). This control system should be much simpler than either x or y control.

This study assumed perfect motion detectors for the feedback loops. Good detectors [6] are available for $x, y$, and $z$ motions but their sensitivities and accuracies and sensitivities need to be added to the control simulation as disturbance.

The dynamic as well as probabilistic characteristics of each of these disturbances as well as man motion need to be characterized and incorporated into the control model to determine the pointing accuracies and pointing stability of the gimbal mounted boom. Future work should include a stochastic analysis of the control system concerning probabilistic disturbances and the minimization of their effects on pointing stability. The results of such a study would be a three-axis model of the system compensated for stability as well as compensated for disturbance induced pointing variations.

## References

1. Sperry Support Services, Final Report Contract \#NAS8-33588, "Sperry Interactive Graphics System (SIMS) and its Application to Control Simulation of the Pinhole X-Ray/Coronograph Experiment," 1980.
2. Snowdon, J. C. "Vibration and Shock in Damped Mechanical Systems," John Wiley and Sons, New York, N. Y., 1968.
3. Thomson, W. T. "Vibration Theory and Applications," Prentice-Hall, Englewood Cliffs, N. J., 1965.
4. Sperry Flight Systems, Report on Contract \#NAS8-33413, "Annular Suspension Pointing System - Gimbal System Design Report - Volume I Data Package for Preliminary Design Review," Data Requirement 8, 1979.
5. D 'Azzo, J. J. and Houpis, C. H. "Linear Control System Analysis and Design, Conventional and Modern." McGraw-Hill, New York, N. Y., 1975.
6. Visidyne, Final Report Contract \#NAS 8-33697, "Pinhole X-Ray/ Coronograph Optical Systems Concept Definitions Study," 1980.

TABLE 1. PINHOLE BOOM FREQUENCIES AND MODE SHAPES

| $\mathrm{L}=50 \mathrm{METERS}$, TIP MASS $=505.7 \mathrm{KG}$ |  |
| :---: | :---: |
| FREO (HZ) | MODE SHAPE |
| 1 |  |
| 1 |  |
| 17.1 |  |
| 17.4 |  |
| 17.4 |  |
| 54.9 |  |
| 54.9 |  |



Figure 1. Pinhole occulter facility.

## CONFIGURATION



Figure 2. Pinhole occulter facility configuration.


Figure 3. Boom frequency as a function of tip mass and boom length.


Figure 4. Simplified block diagram of pinhole control system.


Figure 5. Reduced block diagram of pinhole control system.





Figure 9. Error response of boom tip to thruster firing and gravity gradients.




XXVIII-20


Figure 12. Frequency response of the uncompensated gimbal driven boom system.
(s)


XXVIII-23
NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAMMARSHALL SPACE FLIGHT CENTERTHE UNIVERSITY OF ALABAMAFFT METHODS IN SIGNAL PROCESSING OFTHE COAL INTERFACE DETECTOR RADAR

| Prepared By: | Darko Kajfez, Ph. D. |
| :---: | :---: |
| Academic Rank: | Professor |
| University \& Department | University of Mississippi Department of Electrical Engineering |
| NASA/MSFC : |  |
| (Laboratory) | Electronics \& Control |
| (Division) | Optical \& RF Systems |
| (Branch) | Communications \& Tracking |
| MSFC Counterpart: | Billy R. Reed |
| Date: | Auqust 1, 1980 |
| Contract No.: | $\begin{aligned} & \text { NGT-01-002-099 } \\ & \text { (University of Alabama) } \end{aligned}$ |

# FFT METHODS IN SIGNAL PROCESSING OF THE COAL INTERFACE DETECTOR RADAR 

## BY

Darko Kajfex, PH. D. Professor of Electrical Engineering University of Mississippi University , MS 38677

ABSTRACT
The FM radar for the coal interface detector, operating in the frequency band 2 to 4 GHz , is intended for the display of thicknesses between 2 cm and 20 cm . Because of such a short range, the thickness information is contained in the very few lowest spectral components of the output signal. To overcome this inconvenience, the Fourier series of the output signal has been augmented to approximate a Fourier integral. This modification in the signal processing resulted in a higher spectral density, which in turn enabled an easier identification of the interface position in the laboratory. The orientation and spacing of the receiving and transmitting antennas is found to have an important influence on the system performance. A field test in the coal mine is planned in the future.

## ACKNOWLEDGEMENTS

I appreciate the opportunity of continuing the work on the challenging problem of the CID radar during my second year of NASA/ASEE Summer Faculty Research Fellowship Program. My NASA counterpart, Mr. Billy R. Reed provided an excellent support in all aspects of the project. Mr. Ivan A. Burroughs was very helpful in programming of the signal processing procedure, and Mr. Edmund H. Gleason provided the expertise in the selection and utilization of the various electronic instruments and subsystems. The typing of the report was performed by Joyce Gray.

## INTRODUCTION

Coal interface detector (CID) radar is designed to measure the thickness of the coal layer in the longwall mining method. The frequency of the radar transmitter is modulated in a saw tooth fashion between 2 and 4 GHz . The thickness of the coal layer is measured by observing the frequency difference between the signals reflected from the front and rear surface of the coal. The radar output signal, which contains this difference frequency, is analyzed on a digital Fourier Analyzer, which then provides a display of the reflection amplitude vs. the coal thickness.

The basic principle of the CID radar has been under investigation at the MSFC for several years. The difficulty in implementing the radar lies in the presence of a multitude of erroneous signals, which can be misinterpreted as the rear surface return signals. Also, the frequency of the reflected signal is only several times $(<10)$ higher than the modulation frequency. Therefore, it is difficult to determine which of the individual spectral components belongs to the position of the rear surface.

The presence of the multiple reflections may be somewhat reduced by using long cable sections between the individual components of the system, and by placing absorbing materials around the antennas. The second difficulty of distinguishing between the inidvidual components in the spectrum can be greatly reduced by processing the time-domain signal before taking the Fast Fourier Transform. Examples of measured wave-forms and their spectra are included in the report.

## OBJECTIVES

This report describes the work performed in the second period of the NASA/ASEE Summer Faculty Research Fellowship Program. The radar configuration is somewhat different from the one described in the last year's report (Reference (1)). The difference lies in the fact that a direct detector receiver has been utilized instead of the mixer-type receiver. The objective of the investigation described in the present report is to achieve the radar configuration which is suitable for application not only in the laboratory, but in the actual coal mine situation. Much of the effort is directed toward determining the most suitable signal processing procedure which produces a clear display of the coal thickness.

## CIRCUIT DESCRIPTION



Fig. 1. Block diagram of the CID radar using the detector receiver
The transmitting part of the CID radar in Fig. l consists of the Hewlett Packard Function Generator Model 3312, the Wiltron Sweep Generator Model 610, the Avantek Transistor Amplifier Model 406-3M, and the Scientific Atlanta Ridged Horn Antenna Model 40M-2.0. The lenoth of the coaxial cable between the amplifier and the antenna is $\ell_{l}=1 \mathrm{~m}$. The receiving part of the radar consists of an identical antenna, as in the transmitting part, and a Narda Detector Mount Model 4503. The detected signal is amplified and filtered in the Ithaco Dual Filter Model 4302. The filter is set for pulse operation (Bessel type response) in the lowpass mode.

The operating data are as follows:

As indicated in Fig. 1, the received signal consists of two parts: One part comes from the wave which was reflected from the front surface of the coal, and the other part is a weaker signal from the wave which penetrated the coal and was then reflected from the rear interface. If the coal dielectric constant is denoted by $\varepsilon_{r}$ the sweep period by $T$, the coal layer thickness by d, and the RF deviation by $B$, the frequency difference between the front and rear reflection is

$$
\begin{equation*}
f_{d}=\frac{4 B d \sqrt{\varepsilon_{r}}}{c T} \tag{1}
\end{equation*}
$$

c. is the velocity of light in free space. For the data listed above, a coal layer of thickness $d=6 "(15.24 \mathrm{~cm})$ and the dielectric constant of $\varepsilon_{r=5}$ would produce a difference frequency approximately 355 Hz .

When such a mixture of the two microwave signals is applied to the square law detector, their difference frequency (here 355 Hz ) is obtained. This signal is interrupted by the repetition rate of the sweep generator. It is afterwards amplfiied and filtered before being delivered to the Fourier Analyzer Hewlett Packard Model 5451B, which performs the signal processing and displays the result.

The radar system described above is similar in concept to the radar described in Reference (2). The main difference of the present system is the use of a direct detector system instead of the heterodyning receiver, and the use of a digital signal processing system instead of an analog one.

## SIGNAL PROCESSING

The tests in the laboratory are performed on slabs of "artificial coal". These are slabs of different thicknesses, made from a mixture of dielectric and conducting materials, so that the dielectric constant and the loss tangent approximate the values of the actual coal from the mine. In the particular measurement to be described, the slab thickness 15.5 cm ( $6^{\prime \prime}$ ) and of the cross section $61 \times 61 \mathrm{~cm}$ (2'x2') was backed by an aluminum foil for better rear surface reflection. The received signal at the input to the Fourier Analyzer is shown in Fig. 2.

The observed signal is periodic with the period $\mathrm{T}=25.6 \mathrm{~ms}$. The sawtooth waveform which modulates the transmitter is shown in Fig. 3. The characteristics of the transmitter are such that the modulation voltage 2.4 V produces an output of 2 GHz and 4.8 V produces an output frequency of 4 GHz . It is seen from Figures 2 and 3 that the detected signal displays an even symmetry within each half period. This is to be expected, because the transmitted frequency attains the same value twice within each period : once on the upswing, and once on the downswing of the modulation signal.

The signal shown in Fig. 2 has been sampled with the time intervals 50 us, the total number of sample points being 1024. The waveform shown includes two full periods of the signal. If now only one period is retained (first 512 points) and the Fast Fourier Transform of this waveform is computed, the resulting spectrum is such as shown in Fig. 4. Each spectral component makes one point in the plot. Due to an automatic plotting procedure, the points are interconnected by straight lines, creating the appearance of a continuous spectrum. The spectrum actually consists of discrete frequencies, separated from each other. by the repetition frequency 39 Hz . According to (l), the beat frequency for the $6^{\prime \prime}$ artificial coal is 355 Hz . From Fig. 4 it can be seen that the two longest spectral components are components number one and number eight. Thus, by neglecting the first component which belongs to the repetition frequency, the frequency of the return signal is estimated to be eight times the repetition frequency, or 312 Hz (instead of 355 Hz ).

There are several inconveniences associated with the display such as in Fig. 4. In the first place, the repetition frequency usually comes out to be the strongest component, which fact makes it difficult to read the thickness of the thin coal layers, where spectral components are very close to the repetition frequency. Furthermore, the components immediately next to the observed return frequency often have a small amplitude, so that the signal is not always readily identifyable as in Fig. 4, where the seventh, eighth, and ninth spectral components are all large. It was decided that a difficulty in locating the return signal is caused by the fact that the spectral components are too far apart





FIG. 4 FOURIER SERIES OF THE PERIODIC SIGNAL


FIG. 5 HANN TIME WINDOW
from each other. In order to create a more densely populated spectrum, advantage has been taken of the digital signal processing in the following manner.

The original sampled signal from Fig. 2 consists of 1024 points. The sample extends over two periods, each period has a half-period symmetry. Therefore, the basic information is contained within 256 points. Thus, a time window is created, of the length 256 sample intervals, as shown in Fig. 5. This is the so-called Hann window of the shape $1-\operatorname{COS} x$. When the sampled signal is multiplied with the window, the result looks very much like a single burst of oscillations shown in Fig. 6. The total sample length is still 1024 points, but $75 \%$ of the sample points are forced to be zero. The Fast Fourier Transform performed on these 1024 sample points will have much larger spectral dnesity from the spectrum in Fig. 4, and thus will allow more accurate reading of the coal thickness.

Before the spectrum is displayed, another operation is performed. The radar return signal is redundant in the sense that the half period which corresponds to the modulation upswing is a mirror image of the other half period which corresponds to the modulation downswing. Therefore, another Hann window is created, shifted for 256 points. Again the sampled signal is multiplied with this second window, and the Fast Fourier Transform is completed. The resulting spectrum is then multiplied with the previously evaluated spectrum. The procedure is equivalent to taking the convolution of the two half periods in the time domain. This insures that the information contained in each half period contributes equally to the final result. The amplitude of the resulting spectrum is then displayed in a logarithmic scale, as shown in Fig. 7.

The spectrum in Fig. 7 is dense enough for each strong spectral component to be easily identified. When comparing this spectrum with the one in Fig. 4, it is noticed that now the return signal has the largest amplitude, and is thus easily recognized. It appears that the amplitude of this return is about 20 dB larger than other spectral components on either side. However, it should be remembered that the spectrum in Fig. 7 has been obtained by multiplying the two spectra emanating from each of the two half periods, so that the resulting product is made of the squared amplitudes.

Because of this, the number of decibels in Fig. 7 must be divided by two, so that the actual difference between the largest lobe and the next largest lobe is only about 10 dB , and not 20 dB .

The frequency corresponding to the rear surface reflection in Fig. 7 is 310 Hz , instead of the predicted value 355 Hz . The discrepancy may have been caused by an inaccurate knowledge of the frequency deviation $B$ and the dielectric constant $\varepsilon_{r}$ which both enter into equation(l).


FIG. 6 RECEIVED WAVEFORM MULTIPLIED WITH THE HANN WINDOW


FIG. 7 FOURIER TRANSFORM DISPLAY

When a sinusoidal signal is interrupted by the square pulses, the specturm of such a modulated signal contains not only the carrier frequency, but also a number of sidelobes which are decreasing in amplitude according to sin $x / x$ function. The first sidelobe of this function is 13.5 dB smaller than the maximum. The radar return signal is switched each time the modulation waveform changes its slope. Therefore, the observed spectrum will contain considerable sidelobes on each side of the carrier frequency ( 355 Hz in the example discussed above. These sidelobes may be misinterpreted as apparent radar returns. One way of reducing these sidelobes is to use the Hand window, such as shown in Fig. 5. The sidelobes of the sinusoidal signal modulated with the Han window are 37 dB smaller than the maximum. This is the reason that the sampled signal in Fig. 6 has been multiplied by the Han window.

One unavoidable consequence of reducing the sidelobes by using the Han window is the doubling of the spectral width of the main lobe, when compared with the rectangular window (4) should two close surfaces be distinguished in the radar display, their difference in frequency should be at least half width of the main lobe. For the sinusoidal signal multiplied with the Han window, the first zero appears at double the value of the $\sin x / x$ function:

$$
\pi \Delta f \frac{T}{2}=2 \pi
$$

The frequency difference between the main lobe maximum and the first zero is

$$
\begin{equation*}
\Delta f=\frac{4}{T} \tag{2}
\end{equation*}
$$

If two signals are reflected from two very closely located surfaces their frequencies may differ for less than $\Delta f$, and cannot be distinguished on the display. Substituting (2) into (1), we find the minimum resolution distance between the two reflections in the coal

$$
\begin{equation*}
d_{\text {res }}=\frac{2 c}{B \sqrt{\varepsilon_{r}}} \tag{3}
\end{equation*}
$$

The resolution distance is independent of the repetition frequency. The only way to obtain a finer resolution is to increase the frequency deviation For the system with $B=2 \mathrm{GHz}$, the resolulion distance is $\mathrm{d}_{\mathrm{res}}=6.7 \mathrm{~cm}$.

When two reflection surfaces are dloser to each other than dres, their main spectral lobes will merge into a single lobe, and the two returns will not be distinguishable on the radar display. One way to reduce the resolution distance is to use the rectangular window instead of the Hann window. This would split the resolution distance in half at the expense of increasing the sidelobes. It follows from the previous discussion that large sidelobes are undesirable. Therefore, the fact that the frequency deviation of the radar is 2 GHz sets on a priori limit of 6.7 cm as the smallest recognizable coal thickness ( 3.3 cm with the rectangular window).

## ANTENNA CONSIDERATIONS

Little is known about the design of microwave antennas suitable for sensing the material properties in close range. The majority of microwave antennas are designed to produce a given electromagnetic radiation in the far field region. In the early experiments with the CID radar, the transmitting and receiving antennas have sometimes been placed less than one wavelength apart from each other or from the coal interface. It has been empirically found afterwards that these distances must be increased to several wavelengths.

One important property of the antenna to be used for the CID radar is low dispersion. The time delay for the signal, to pass from the antenna terminal to an obstacle in front of the antenna and back to the terminal, should be independent of frequency. One of the popular wideband antennas, the log periodic array, is ill suited for the task, as pointed out in [2].

Another important property of the pair of antennas to be used in the radar is a high decoupling when they are placed side by side. By looking at Fiourel, it becomes obvious that a direct propagation from the transmitting to the receiving antenna creates a third wave impinging upon the detector. The difference frequency of that wave and the front surface reflection may produce an erroneous "return" signal on the radar display.

Without making any theoretical investigation, and without even comparing the performance of various commercially available antennas, the pair of ridged horn antennas has been assigned to do the job. The orientation and the distance between them has been experimentally adjusted to give the best performance as follows.

Figure 8 depicts two possible relative positions of the receiving and transmitting antennas, denoted parallel and collinear. It has been found experimentally that the collinear orientation offers higher isolation between the antennas. The distance D has also been varied, and the value $\mathrm{D}=25 \mathrm{~cm}$ appears to be sufficient. The measurement has been conducted when the antenna-to-coal distance is $l_{\mathrm{a} 1}=l_{\mathrm{a} 2}=60 \mathrm{~cm}$.

Another unexpected critical adjustment proved to be the angle 0 in figure 9. This angle measures the departure of the coal surface from making a right angle with the antenna system. It has been found experimentally that the return signal from the rear surface drops considerably when the difference $1_{a 1}-1_{a 2}$ becomes larger than 2 cm . For $\mathrm{D}=24 \mathrm{~cm}$ and $\mathrm{b}=10 \mathrm{~cm}$, this implies that the angle between the coal surface and the antenna boresight should not depart from the right angle for more than $\theta=3.3$ degrees.


Fig.8. polarization of antennas:
(a) PARALLEL
(b) collinear


FIG. 9. INCLINATION OF THE COAL SURFACE WITH RESPECT TO THE ANTENNA SYSTEM

## CONCLUSIONS

The CID radar system which has been developed has certain advantages over the previously utilized systems. The digital processing of the signal on a Fourier Analyzer System enables more convenient display which seems to improve the separation between the actual return and other unintentionally created signals. When also proper care is exercised in antenna positioning, the system seems to be capable of giving meaningful real-time results in a coal mine environment.

The improvement in the resolution distance may be achieved only by increasing the frequency deviation of the system. Further improvements on the antenna system may also bring an enhancement of the desired return signal and a reduction of the undesired signals.

## REFERENCES

1. Kajfez, D., "Digital Processing of the Spectrum Emanating from the Coal Interface Detector (CID) Radar," University of Alabama in Huntsville, Report on NASA/ASEE Summer Faculty Research Fellowship Program, Huntsville, 1979.
2. Owen , T. E., Tranbarger, O., "Vol. II - Investigation and Development of a High Resolution FM-CW Radar System for Residual Coal Thickness Measurements," Southwest Research Institute, San Antonio, Texas, June l, 1977.
3. $* * *$, "Fourier Analyzer Training Manual," Hewlett Packard Application Note 140-0.
4. Beauchamp, K. G., "Signal Processing, New York : John Wiley, 1973, p. 368.

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA 

EXCHANGE OF STABILITY FOR A COLUMN OF FLUID WITH VARIABLE RAYLEIGH NUMBER UNDER FREE BOUNDARY CONDITIONS

| Prepared By: | Letitia Korbly, Ph.D. |
| :---: | :---: |
| Academic Rank: |  |
| University and Department: | University of Alabama in Birmingham |
| NASA/MSFC: |  |
| (Laboratory) | Space Sciences |
| (Division) | Atmospheric Sciences |
| (Branch) | Fluid Dynamics |
| MSFC Counterpart: | George H. Fichtl |
| Date: | August 15, 1980 |
| Contract No.: | NGT'-01-002-099 |
|  | (University of Alabama) |

EXCHANGE OF STABILITY FOR A COLUMN OF FLUID WITH VARIABLE RAYLEIGH NUMBER UNDER FREE BOUNDARY CONDITIONS
by
Dr. Letitita Korbly
Department of Mathematics
University of Alabama in Birmingham
Birmingham, Alabama

## ABSTRACT

Motion starts in a column of fluid subject to an adverse temperature gradient when the gravitational forces are large enough to overcome those of viscosity. This point is governed by the critical (smallest) Rayleigh number.

When the column is heated rapidly from below and not well mixed the Rayleigh number varies with height.

For such a system with harmonic boundary conditions the principle of exchange of stabilities is shown. Attempts to show this principle under rigid and semirigid boundary conditions failed.

The Rayleigh number was assumed to follow the nondimensional distribution $P(2)=R\binom{2+z_{0}}{z_{0}}^{-\alpha}, \alpha=4 / 3$ or $\alpha=3 / 2, z_{0}=10^{-5}, 0 \leq z \leq 1$. The lowest values for at which there would be convection was estimated at $8.512 \times 10^{8}$ for $\alpha=4 / 3$ and $4.7023 \times 10^{9}$ for $\alpha=3 / 2$. The eigenfunction was approximated in both cases.

This work has application to both columns of air in the Earth's atmosphere and the GFFC experiment to be flown in Spacelab 3.

## INTRODUCTION

The Navier Stokes equations have classically been used to describe the motion of fluid in the presence of temperatore gradients.

In the Benard problem, the fluid is assumed to be incompressible and at rest; it is assumed to be uniformly heated from below. The vertical temperature distribution is assumed to be linear. The equations are then linearized by looking at temperature and velocity perturbations.

The equations are then reduced to a single sixth order linear partial differential equation with constant coefficients in the vertical velocity perturbation $\omega$. Fourier transform in the horizontal space variables has reduced the number of independent variables. Only time, $t$, and height, $z$, remain. The equation is then transformed in $t$ and non-dimensionalized. It becomes:

$$
\left[\left(D^{2}-a^{2}\right)\left(D^{2}-a^{2}-\sigma\right)\left(D^{2}-a^{2}-p \sigma\right)+R a^{2}\right] W(z)=0 \quad 0 \leq 2 \leq 1
$$

Here the vertical velocity perturbation is

$$
w=e^{\sigma t+i\left(a_{1} x+a_{2} y\right)} \backslash N(z)
$$

Here $D W=W$ and $a^{2}=a_{1}{ }^{2}+a_{2}{ }^{2}$.
The principle of exchange of stabilities holds if convection modes which neither grow nor decay with time are time independent. If $\mathbb{R}$, the Rayleigh number, and $p$, the Prandial number, are constant, exchange of stabilities can be shown to hold under both free and rigid boundary conditions.

If is allowed to vary with height, what will happen? We were able to show the principle of exchange of stabilities holds under free, but not rigid or semirigid, boundary conditions.

Having this result, the onset of convection was investigated for certain distribution of $R$ which model what occurs in the summer atmosphere below the inversion. The values for used were

$$
R(z)=R\left(\frac{z_{2}}{z_{0}}\right)^{-\alpha}
$$

$$
z_{0} \leq 2 \leq 8, \alpha=4 / 3 \operatorname{sen} \alpha=3 / 2 \text {. }
$$

Here $d$ is the height of the inversion, $2_{0}$ is surface roughness length, and $\alpha$ is empirically chosen. This is nondimensionalized, and the critical (smallest) value of $R$ for which convection takes place is estimated.

The results of the calculations were checked by determining the temperature gradient at the surface, $\left.\frac{\partial \theta}{\partial z}\right|_{y=z_{0}}=\beta_{0}$ This determined $\frac{\partial \theta}{\partial z}=\beta_{0} z^{-4 / 3}$ which was then integrated to give a temperature drop of roughly $15^{\circ}$ in the first meter above the surface. This figure checks reasonably well with the data.

## EXCHANGE OF STABILITIES

These calculations show that the principle of exchange of stability holds for a column of fluid heated from below if the Rayleigh number is variable when the boundary conditions are harmonic. The Boussinesq approximations are used. The transverse velocities, $u$ and $v$, the pressure, $p$, and the temperature perturbation, $\theta$, are eliminated. The nondimensional equation is

$$
\nabla^{2}\left(D_{t}-\nabla^{2}\right)\left(p D_{t}-\nabla^{2}\right) \omega=-\nabla_{h}^{2} R \omega
$$

Here $\nabla^{2} \omega=\omega_{x x}+\omega_{y y}+\omega_{z z} ; \quad \nabla_{n}^{z} \omega=\omega_{x x}+\omega_{y y}$ $\rho=\nu / K$ is the Brandt number, and the Rayleigh number, $R=g \alpha \beta d^{Y} /(\gamma k)$, depends on the height, $z$.

Proceeding with the analysis into normal modes, set

$$
w=e^{\sigma t+i\left(a_{1} x+a_{2} y\right)} W(z)
$$

$$
\begin{aligned}
& \text { The operator then becomes: } \\
& * \quad\left(D^{2}-a^{2}\right)\left(D^{2}-a^{2}-o\right)\left(D^{2}-a^{2}\right) w=-R a^{2} \backslash N
\end{aligned}
$$

Here $a^{2}=a_{1}{ }^{2}+a_{2}^{2}$ and $D W=W^{\prime}$.
The harmonic boundary conditions are:

$$
\begin{aligned}
& W(0)=0=W(1) \\
& W^{\prime \prime}(0)=0=W^{\prime \prime}(1) \\
& W^{\prime \prime \prime}(0)=0=W V^{\prime \prime}(1)
\end{aligned}
$$

Multiplying (*) by $\mathbb{V}$ and integrating forms the following equation.
$x_{*} \int_{0}^{1}\left(D^{2}-a^{2}\right)\left(D^{2}-a^{2}-\sigma^{\prime}\left(n^{2}-a^{2}-y, \sigma\right) W V / \bar{V}=-\left.\int_{0}^{1} R a^{2} \backslash N\right|^{2} d z\right.$
The following integral occurs in the analysis of (**):

$$
I_{1}=-S_{0}^{1}\left(D^{2}-a^{2}\right) W W W_{z}=S_{0}^{1}|D W|^{2} d z+\left.a^{2} S_{0}^{1} W\right|^{2} d z
$$

The boundary conditions, $W(0)=0=W(1)$, make the integrated terms vanish. Set $S_{0}^{1}\|W\|^{2} d_{z}=\|W\|^{2}$ and $S_{0}^{1}|D W V|^{2} d_{z}+a^{2} S_{0}^{1}|w \nu|^{2} d_{z}=\|\nabla W\|^{2}$

Both $\|W\|$ and $\|\nabla W\|$ arc strictly positive unless $W \equiv 0$ under the boundary conditions. So $I_{1}=\|\nabla \mathrm{W}\|^{2}$ is strictly positive unless $\backslash N=0$

The next integral,

$$
\begin{aligned}
I_{2} & =S_{0}^{1}\left(D^{2}-a^{2}\right)^{2} W W N d z=-\int_{0}^{1} D\left(D^{2}-a^{2}\right) W D W N+a^{2}\left(D^{2}-a^{2}\right) W \mathcal{W} d_{2} \\
& =S_{0}^{1}\left|\left(D^{2}-a^{2}\right) W\right|^{2} d z
\end{aligned}
$$

Again the integrated terms vanish because $W(0)=0=W(1)$ and $W^{\prime \prime}(0)=0=W^{\prime \prime}(1)$. (The same result can be obtained for this integral under rigid boundary conditions.)

Finally

$$
\begin{aligned}
I_{3} & =-S_{0}^{1}\left(D^{2}-a^{2}\right)^{3} W \bar{w} d z \\
& =S_{0}^{\prime}\left(D^{2}-a\right)^{2} D w D \bar{w}+a^{2}\left(D^{2}-a^{2}\right)^{2} W \bar{W} d z \\
& =-S_{0}^{\prime}\left(D^{2}-a^{2}\right)^{2} W\left(D^{2}-a^{2}\right) \bar{W} d z
\end{aligned}
$$

Again the integrated terms vanish under both free and rigid boundary conditions.

$$
\begin{aligned}
I_{3}= & S_{0}^{1}\left|\left(D^{2}-a^{2}\right) D W\right|^{2} d z+a^{2} S_{0}^{1}\left|\left(D^{2}-a^{2}\right) W\right|^{2} d z \\
& +\left.\left(D^{2}-a^{2}\right) D W\left(D^{2}-a^{2}\right) \bar{W}\right|_{0} ^{1}
\end{aligned}
$$

Under free boundary conditions the integrated terms vanish so

$$
I_{3}=\int_{0}^{1}\left|\left(D^{2}-a^{2}\right) D W\right|^{2} d z+\left.a^{2} \int_{0}^{1}\left(D^{2}-a^{2}\right) W\right|^{2} d z=\left\|\nabla\left(D^{2}-a^{2}\right) W\right\|^{2}
$$

When (**) is multiplied by -1 , it becomes $-\int_{0}^{1}\left(D^{2}-a^{2}\right)^{3} W \bar{W} l_{z}+(p+1) \sigma \int_{0}^{1}\left(D^{2}-a^{2}\right)^{2} W \bar{W} d z-\rho \sigma^{2} \int_{0}^{1}\left(D^{2}-a^{2}\right) W \bar{W} d z$ $\int_{0}^{1}=a^{2} S_{0}^{1} R|\omega|^{2} d z$ Set $J=S_{0}^{\prime} R \mid \omega l^{2} d_{z}$. If $I_{1}, I_{2}, I_{3}$ and $J$ are substituted in the equation above it becomes

$$
p I_{1} \sigma^{2}+(p+1) I_{2} \sigma+I_{3}-J=0 .
$$

Since $P$, the Prandtl number, is always positive, and $I$, is positive unless $\omega=0$, the quadratic formula can be applied to solve for $\sigma$. This gives

$$
\sigma=\left[-(p+1) I_{2} \pm \sqrt{(p+1)^{2} I_{2}^{2}-4 p I_{1}\left(I_{3}-J\right)}\right] / p I_{1}
$$

To show that exchange of stabilities holds, we must show that if $\operatorname{Re} \sigma=0$, then $\operatorname{Im} \sigma=0$.

There are two cases where Re $\sigma=0$
(i) $\quad 4 甲 I_{1}\left(I_{3}-J\right)=0$ and
(ii) $(p+1) I_{2}=0$ and $-4 p I,\left(I_{3}-J\right) \prec 0$.

If (i) holds $\sigma=0$, if (ii) holds, $\bar{I}_{2}=0$ so

$$
\int_{0}^{1}\left|\left(D^{2}-a^{2}\right) W\right|^{2} d z=0
$$

This implies

$$
\left(D^{2}-a^{2}\right) W \equiv 0
$$

The last implies that

$$
I_{1}=-\int_{0}^{1}\left(D^{2}-a^{2}\right) W \bar{W} d z=0
$$

In both cases Re $\sigma=0$ implies $\sigma=0$. This completes the proof of the principal of exchanges of stabilities for operators whose Rayleigh number is variable under free boundary conditions.

## References

Chandrasekhar, S. "Hydrodynamic and Hydromagnetic Stability" Oxford University Press, 1961.

## ANALYSIS INTO NORMAL MODES

Since the principle of exchange of stabilities holds for the variable Rayleigh number problme, we attempt to solve the steady state problem for certain functions, $\mathbb{R}$

$$
R(z)=R\left(\frac{z}{z_{0}}\right)^{-\alpha}, \quad \quad z_{0}=10^{-5}, \alpha=1 / 3 \Omega \alpha=3 / 2,
$$

The equation is

$$
\left[\left(\nabla^{2}\right)^{3}-R\left(\frac{z}{z_{0}}\right)^{-\alpha} \nabla_{h}^{2}\right]_{u}=0, \quad z_{0} \leq z \leq 0
$$

with the harmonic boundary conditions

$$
\begin{aligned}
& u(0)=0=u(1) \\
& u^{\prime \prime}(0)=0=u^{\prime \prime}(1) \\
& u^{(u)}(0)=0=u^{(v)}(1)
\end{aligned}
$$

Here $\nabla_{u}^{2}=u_{x x}+u_{y y}+u_{z z}$ and $\nabla_{h}^{z} u=u_{x *}+u_{y y}$
If we set $\}=z-z$, the equation becomes

$$
\left.\left(\nabla^{2}\right)^{3} u=R\left(\frac{2+z_{4}}{z_{0}}\right)^{-\alpha} \nabla_{h}{ }^{2} u, \quad 0 \leq\right\} \leq Q-z_{0}
$$

with harmonic boundary conditions at 0 and $d-z_{0}$. The equation is normalized by setting

$$
\begin{aligned}
& x=x /\left(d-z_{0}\right) \\
& y=y /\left(d-z_{0}\right) \\
& z=\xi /\left(d-z_{0}\right)
\end{aligned}
$$

and becomes

$$
\left(\nabla^{2}\right)^{3} u=R\left[z_{0}^{\alpha} /\left(d-z_{0}\right)^{4}\right]\left(z+z_{0}\right)^{-\alpha} \nabla_{h}^{2} u, \quad 0 \leq z \leq 1
$$

with harmonic boundary conditions at 0 and 1 . Let

$$
R_{0}=R z_{0}^{\alpha} /\left(d-z_{0}\right)^{4}
$$

The equation reduces to

$$
\left(\nabla^{2}\right)^{3} u=R_{0}\left(z+z_{0}\right)^{-\alpha} \nabla_{4}^{2} u
$$

If it is assumed that

$$
u=\sum_{n=0}^{\infty} e^{i\left(a_{1} x+a_{2} y\right)} A_{n} \sin \pi n z
$$

with $a^{2}=a_{1}^{2}+a_{2}^{2}$, the equation reduces to the ordinary differential equation:

$$
\left(D^{2}-a^{2}\right)^{3} u=-R_{0}\left(z+z_{0}\right)^{-a} a^{2} u
$$

If $A_{n}=0, n \Sigma 2$, we can look for a first approximation to the critical (smallest) $R_{0}$ for which the equation can be satisfied. This will be approximated by finding the
smallest ' $R_{0}$ such that

$$
0=\int_{0}^{1}\left[\left(D^{2}-a^{\prime}\right)^{3}+1 R_{0}\left(z+z_{0}\right)^{-\alpha} a^{z}\right] \sin n=\sin \pi z d z
$$

That is

$$
\left(\pi^{2}+a^{2}\right)^{3} S_{0}^{1} \sin ^{2} \pi z d z=' R_{0} a^{2} S_{0}\left(z+z_{0}\right)^{-2} \sin ^{2} \pi z d z
$$

Set $\quad I_{1,1}^{\alpha}=S_{0}^{\prime}\left(z+z_{0}\right)^{-\alpha} \sin ^{2} n z d z$
and note that $\quad \int_{0}^{1} \sin ^{2} n 2=1 / 2$.
This yields

$$
\begin{aligned}
& \text { yields } \\
& \left.R_{0}=\left(\pi^{2}+a^{2}\right)^{3} /\left(2 a^{2} I_{1,1}^{\alpha}\right)=\left(1 / 2 I_{11}^{\alpha}\right)\right)\left(\pi^{2}+a^{2}\right)^{3} / a^{2}
\end{aligned}
$$

It is easy to see that ' $R_{0}$ has a minimum when ${ }^{\prime} R_{0}{ }^{\prime}=0$, that is, when $a^{2}=\pi^{2} / 2$. Then

$$
' R_{0}=2717 \% /\left(8 I_{1,1}^{\alpha}\right)
$$

Since numerical integration yields the values

$$
士_{1 / 1}^{4 / 3}=1.760401
$$

and

$$
I_{1,1}^{3 / 2}=2.157553 .
$$

The first estimate, ${ }^{R_{0}}$, of $R_{0}$ is

$$
\begin{array}{lll}
' R_{0}=186.25045^{\prime}, & \text { if } & \alpha=4 / 3, \\
' R_{0}=152.3743, & \text { if } & \alpha=3 / 2 .
\end{array}
$$

The first estimate of $R$, then is

$$
8.667841 \times 10^{8}, \quad \text { if } \quad \alpha=4 / 3
$$

and

The second approximation is made by assuming $A_{n}=0, n \geqslant 3$.
The problem is to find the smallest number, ${ }^{2} R_{0}$, and a pair of numbers $A_{1}$ and $A_{2}$, with
$O=\int_{0}^{1}\left[\left(D^{2}-a^{0}\right)^{3}+R_{0}\left(z+z_{0}\right)^{-x} a^{2}\right]\left(A_{1} \sin \pi z+A_{2} \sin 2 \pi z\right) \sin n z d z$
and
$0=\int_{0}^{1}\left[\left(D^{2}-a^{2}\right)^{3}+{ }^{2} R_{0}(2+2)^{-\alpha} a^{2}\right]\left(A_{1} \sin \pi z+A_{2} \sin 2 \pi z\right) \sin 2 \pi^{2} d z$.
That is

$$
\begin{aligned}
A_{1}\left(\pi^{2}+a^{2}\right)^{3}\left[\int_{0}^{1} \sin ^{2} n 2 d z\right]=a^{2} R_{0}\left[A_{1}\right. & \int_{0}^{1}\left(2+z_{0}\right)^{-\alpha} \sin ^{2} \pi z d z \\
& \left.+A_{2} S_{0}^{1}(2+2)^{-\alpha} \sin 2 \pi z \sin \pi 2 d a\right]
\end{aligned}
$$

and

$$
\begin{array}{rl}
A_{2}\left(\pi^{2}+a^{2}\right)^{3}\left[\int_{0}^{1} \sin ^{2} 2 \pi z d z\right]=a^{2} & 2 R_{0}\left[A_{1} S_{0}^{1}\left(z+z_{0}\right)^{-\alpha} \sin n z \sin 2 \pi z d z\right. \\
& +A_{2} S_{0}^{1}\left(2+z_{0}\right)^{-\alpha} \sin ^{2} 2 \pi z d z .
\end{array}
$$

Set $\quad I_{j, k}^{\alpha}=S_{0}^{\prime}\left(z+z_{0}\right)^{-\alpha} \sin j \pi z \operatorname{sink} \pi z d z$.
Note that $I_{j, k}^{\alpha}=I_{k, j}^{\alpha}$ and that $S_{0}^{\prime} \sin ^{2} j \pi z d z=V_{2}$.
This means that

$$
A_{1}\left(\pi^{2}+a^{2}\right)^{3} / 2=a^{2} \quad{ }^{2} R_{0}\left[A_{1} I_{1,}^{\alpha}+A_{2} I_{1,2}^{\alpha}\right]
$$

and

$$
A_{2}\left(4 \pi^{2}+a^{2}\right)^{3} / 2=a^{2}{ }^{2} R_{0}\left[A_{1} I_{1,2}^{\alpha}+A_{2} I_{2,2}^{\alpha}\right] .
$$

The numbers $I_{j, k}^{\alpha}$ are found to any desired degree of accuracy by Simpson's rule.

The problem can be written now as an eigenvalue problem in $V^{2} R_{0}$ with eigenvector ( $A_{1}, A_{2}$ ).

$$
\begin{aligned}
& \left(1 / 2 R_{0}\right) A_{1}=2 a^{2} /\left(\pi^{2}+a^{2}\right)^{3}\left[I_{1,1}^{\alpha} A_{1}+I_{1,2}^{\alpha} A_{2}\right] \\
& \left(1 / 2 R_{0}\right) A_{2}=2 a^{2} /\left(4 \pi^{2}+a^{2}\right)^{3}\left[I_{1,2}^{\alpha} A_{1}+I_{2,2}^{\alpha} A_{2}\right]
\end{aligned}
$$

At the value of $a^{2}$ corresponding to the smallest value of $R_{o}$, the function $1 /{ }^{2} R_{o}$ takes on its largest value. This will be the maximum of the largest eigenvalue of all the matrices,

$$
M=\left(\begin{array}{ll}
m_{11} & m_{12} \\
m_{21} & m_{22}
\end{array}\right)
$$

where $\quad m_{j, k}=\left(2 a^{2} /\left(j^{2} \pi^{2}+a^{2}\right)\right.$ in $I_{j, k}^{\alpha}$.
$M$ is not symmetric.
It is convenient to set $b=a^{2} / \pi^{2}$. Then

$$
M_{j, k}=\left(2 b /\left(j^{2}+b\right)^{3}\right) I_{j, k}^{\alpha} / \pi^{4} .
$$

Let $r=b /(1+b)^{2}$ and $s=b /(4+b)^{3}$. $r$ reaches its maximum when $b=1 / 2$; then $r=4 / 27$ and $s=4 /(27)^{2}$. $s$ reaches its maximum when $b=2$; then $s=1 / 108$ and $r=2 / 27$.

It is clear that the entries of $M$ are positive since $I_{j}^{\alpha} k$ is the integral of a positive function and the entries of the second row are much smaller than those of the first row.

$$
\alpha=4 / 3
$$

$\alpha=8 / 2$

$$
\begin{array}{ll}
I_{1,1}^{4 / 3}=1.760401 & I_{1,1}^{3 / 2}=2.157553 \\
I_{1,2}^{4 / 3}=1.177827 & I_{1,2}^{3 / 2}=1.667493 \\
I_{2,2}^{4 / 3}=2.593355 & I_{2,2}^{3 / 2}=3.446855
\end{array}
$$

These numbers are accurate to the last place, though some of them are extrapolated to determine the last place.

The problem then is to find the largest eigenvalue, $\lambda$, of

$$
2\left(\begin{array}{lll}
r & I_{1,1}^{\alpha} & r I_{1,2}^{\alpha} \\
s & I_{1,2}^{\alpha} & s I_{2,2}^{\alpha}
\end{array}\right)
$$

Then $\lambda / \pi^{4}$ will be the largest eigenvalue of $M$ and ${ }^{2} R_{0}=\pi^{v} / \lambda$

$$
0=\lambda^{2}-2\left(r I_{1,1}^{\alpha}+s I_{2,2}^{\alpha}\right) \lambda+4 \operatorname{sis}\left(I_{1,1}^{\alpha} I_{2,2}^{\alpha}-\left(I_{1,2}^{\alpha}\right)^{2}\right)
$$

It is difficult to maximize $\lambda$ as a function of $a^{2}$ analytically. Instead, using the values of $I j_{j}$ obtained by numerical intergration, shown in the tables, values of $\lambda$ were calculated in a range between $b=.5$, and $b=2$.

When $a^{2}=\pi / 2, b=y / 2$. This is the value of both the critical Rayleigh number in the constant coefficient case $(\alpha=0)$ and the value of $b$ corresponding to the first estimate ${ }^{\prime} R$, of $R$. we did earlier.

The calculation was done as follows. For $\alpha=4 / 3, \lambda$ and $2 R_{0}$ were evaluated at numbers

$$
b=0.00001,0.10001,0.20001, \cdots, 1.90001
$$

The largest value of $\lambda$ occurred when 0.50001 . $\lambda$ and ${ }^{2} R$. were then evaluated at

$$
b=0.50001,0.5100,0.5200, \ldots, 0.59001
$$

The maximum value of $\lambda$ this time occurred when $b=0.5 / 001$. So $\lambda$ and $\geqslant R_{\text {. were evaluated at }}$

$$
b=0.50101,0.50201, \ldots, 0.51801,0.51501
$$

This process was continued until the machine produced no differences in the sixth decimal place for ${ }^{2} R_{0}$. This represents eight significant digits. The numbers If, are known only to seven significant digits. The calculadion is not very sensitive to changes in $b$ (and therefore $\mathrm{b}^{2}$ ) when is near its critical value, but it should be quite sensitive to changes in $I_{j, k}^{\alpha}$.

When $\alpha=4 / 3$, the critical value was established at

$$
b=0.50904, \quad a^{2}=5.024023
$$

Then

$$
2 R_{0}=183.5693
$$

and

$$
2 R=8.520189 \times 10^{8}
$$

The eigenvector is

$$
\binom{A_{1}}{A_{2}}=\binom{0.9996605}{0.0260555}
$$

At

$$
\begin{array}{ll}
b=0.50904, & \lambda=0.5306395, \\
r=0.1481322 & \text { and } \quad s=0.0055526
\end{array}
$$

The corresponding critical values for $\alpha=3 / 2$ were

$$
\begin{array}{ll}
b=0.51208 \\
2 R_{0}=148.9276 \\
{ }^{2} R=4.709315 \times 10^{7}
\end{array} \quad a^{2}=5.054027
$$

The eigenvector is

$$
\binom{A_{1}}{A_{2}}=\binom{0.9995444}{0.0301839}
$$

At $b=0.51208, \quad \lambda=0.6540702$

$$
r=0.1481199 \quad \text { and } \quad s=0.0055^{-} 745^{-}
$$

If we compare ' $R_{0}$, the first estimate of $R_{0}$ with ${ }^{2} R_{0}$, the second estimate of $R_{0}$, we have $\alpha=3 / 2$

| ${ }^{1} R_{0}$ | 186.7505 | 152.3743 |
| :--- | :--- | :--- |
| ${ }^{2} R_{0}$ | 183.5693 | 188.9226 |

The numbers $\mathcal{R}_{0}$ converge quickly to $R_{0}$. If a third


$$
\begin{array}{lll}
I_{1,3}^{\alpha} & 0.832954 & 1.289302 \\
I_{2,3}^{\alpha} & 1.871715 & 2.791557 \\
I_{3,3}^{\alpha} & 3.182768 & 4.442455
\end{array}
$$

This problem is analogous to the last. We are looking for a function
$W(2)=A_{1} \sin \pi z+A_{2} \sin 2 \pi z+A_{3} \sin 3 \pi z$
with

$$
S_{0}^{1}\left[\left(D^{2}-a^{2}\right)^{3} W+a^{2} \quad 3 R_{0}\left(z+z_{0}\right)^{-\alpha} W\right] \sin j \pi z d z=0
$$

Let $t=b /(q+b)^{3}$. Then we must find the largest eigenvalue of the matrix, $m^{3}$

$$
m^{3}=2\left(\begin{array}{ccc}
r I_{1,1}^{\alpha} & r I_{1,2}^{\alpha} & r I_{1,3}^{\alpha} \\
s I_{1,2}^{\alpha} & s I_{2,2}^{\alpha} & s I_{2,3}^{\alpha} \\
t I_{1,3}^{\alpha} & t I_{1,3}^{\alpha} & t I_{3,3}^{\alpha}
\end{array}\right)
$$

The factor $t$ is much smaller than $s$, just as $s$ is much smaller than $r$. The determinate of $m^{j}$ is then $O(r s t)$. The sum of the determinate of the principal submatrices is $O(r s)$. The trace is $O(r)$.

Since the determinate is the product of the eigenvalues, the trace is the sum of the eigenmodes, and the sum of the partial products of the eigenvalues is the sum of the determinate of the principal submatrices, it is easy to see that $m^{3}$ has one relatively large eigenvalue, $\lambda$, and that its other eigenvalues are near 0 . The large eigenvalue should be $O(r)$.

Because the large eigenvector is so isolated and because we have a good guess ( $2 r I_{1,1}$ ) for it, the power method was used to calculate $\lambda$ and the eigenvector. This converged. rapidly, giving a value of $\lambda$ that depended on $b$.
$\lambda$ was maximized exactly as in estimating ${ }^{2} R_{0}$.
For $\alpha=4 / 3$, the maximum occurred when $b=0.50970$ and $a^{2}=5.030537$. and ${ }^{3} R=8.512375 \times 10$ . Then, the third estimate, $3 \boldsymbol{R}_{0}=183.3936$

$$
\left(\begin{array}{l}
A_{1} \\
A_{2} \\
A_{3}
\end{array}\right)=\left(\begin{array}{l}
0.999656 \\
0.026135 \\
0.001981
\end{array}\right)
$$

For $\alpha=3 / 2$, the occurred when $b=0.5132$, and $\quad a^{2}=5.06508$.

Then ${ }^{3} R_{0}=148.2001$
and $\quad{ }^{3} R=4.702310 \times 10^{9}$

The eigenvector is

$$
\left(\begin{array}{l}
A_{1} \\
A_{2} \\
A_{3}
\end{array}\right)=\left(\begin{array}{l}
0.999537 \\
0.030308 \\
0.002520
\end{array}\right)
$$

The three estimates of the eigenvalue converged quickly.

| ${ }^{1} R_{0}$ | 186.75 | 152.39 |
| :--- | :--- | :--- |
| $2 R_{0}$ | 183.57 | 148.93 |
| ${ }^{3} R_{0}$ | 183.39 | 148.70 |

It should be noted that none of these values, $\lambda, R_{0}$, or the eigenvector are very sensitive to changes of $b$ or $a^{2}$ near the critical value.

## REFERENCES

Chandraskhar, S. 'Hydrodynamic and Hydromagnetic Stability", 0xford University Press, 1961.

Gantmacher, F. R. "The Theory of Matrices", Vol. II, Chelsea, New York, 1959.

Velarde, M. G. and Christian Normond "Convection", Vol. 243 \#1, Scientific American, July 1980.

$$
\text { dor }=
$$

Conclusions and Recommendations

Work like this in the case of variable viscosity, thermal conductivity and gravity has atmospheric app1ications as well as applications to the space program. This would be applicable to the GFFC experiment to be flown in Spacelab 3, especially if rigid boundary conditions were handled. At least some of this is presently feasible.

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM <br> MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA 

IMPROVEMENTS TO THE NASAP CODE

| Prepared By: | David Perel, Ph.D. |
| :--- | :--- |
| Academic Rank: | Assistant Professor |
| University and Department | University of South Carolina <br> College of Engineering |
| NASA/MSFC: |  |
| Division: | Program Development Office |
| MSFC Counterpart: | Carl E. Colley |
| Date: | August 5, 1980 |
| Contract No.: | NGT 01-002-099 of Alabama |

## IMPROVEMENTS TO THE NASAP CODE

BY

David Perel<br>Assistant Professor of Engineering University of South Carolina Columbia, South Carolina

## ABSTRACT

The object of this work is the modification and improvement of the FORTRAN code, NASAP, written by the author. The purpose of NASAP is the transformation of CAD-generated NASTRAN input data to input data for DESAP II and/or DESAP I. The latter programs are used for structural optimization, and in the initial design phase are often more useful than NASTRAN, which is used for structural analysis.

NASAP is designed to operate interactively, and to this end, modifications and refinements have been made which simplify the interactive participation of the engineer, thus reducing chances for error as well as time spent at the computer terminal. Redundancies have been eliminated, and a tabular form of input has been incorporated into NASAP.

Currently, NASAP is being expanded to include iterative design cycles involving DESAP II and/or DESAP I.

## ACKNOWLEDGMENT

The author thanks Carl Colley for his help in this project.

## INTRODUCTION

The Preliminary Design Office currently has the capability of creating NASTRAN bulk data by use of an interactive graphics system implemented on a Computervision system. This method of creating NASTRAN input data is far more efficient in terms of time and accuracy than manual preparation of the data. The CADDS software associated with the Computervision creates NASTRAN input data. This data is then transmitted to a main computer for NASTRAN execution in the batch mode. The output of the program is essentially the stress and displacement fields of the structure being analyzed.

DESAP I and DESAP II are also structural programs, but differ from NASTRAN in that they optimize the structure with respect to some variable, usually weight (i.e., minimizes the weight) subject to user-supplied limits on stress and displacement (DESAP I), and stress and buckling (DESAP II).

The input required for DESAP I and II is rather lengthy, but much of it, including the geometric description of the structural model, may be derived from the CADDS-generated NASTRAN input data. This is the purpose of NASAP.

The current version of DESAP I allows constraints on stress and displacement, while DESAP II allows constraints on stress and buckling. In order to generate a design satisfying constraints on stress, displacement, and buckling, it is necessary to use both DESAP I and II in an iterative cycle. Development of this capability is currently underway.

## OBJECTIVE

The objective of this project is the improvement and expansion of NASAP in the following ways: (1) elimination of redundancies; (2) simplification of user-supplied input; and, (3) incorporation of the capability of iterative design cycles involving DESAP I and DESAP II.

## REPORT

The body of this report consists mainly of the NASAP program. Its major subroutines are self-explanatory. An overall flow chart, Figures la-1d, outlines the main features of the program. The program is based on the following structural equivalences between NASTRAN and DESAP I and II:

NASTRAN
ROD
BAR
QDMEM
QUAD2
TRIA2
SHEAR

DESAP I and II
BAR
BEAM
Plane Stress Quadrilateral
Plate Quadrilateral
or Triangle
Quadrilateral Shear Panel

This is reflected in the five subroutines in NASAP called BAR, BEAM, QDMEM, QT, and SHEAR. These subroutines deal only with the structural elements their names imply, and perform ordering and manipulation of NASTRAN data via the subroutines CANDP, STOR, PAIR, VEC, and ID.

Figure 2 illustrates the design cycles that are possible using NASAP with DESAP I and DESAP II. The main feature of these cycles is the modification of file 13 (DESAP II) via. calculated design variables and/or tolerances from DESAP I, and the corresponding route for file 14 (DESAP I).

When an acceptable design has been achieved using both DESAP I and DESAP II, there is no guarantee that this design is an optimum, i.e., that for the given structure it minimizes weight subject to the given constraints on stress, displacement, and buckling. There may exist other designs of lower weight also satisfying the constraints. This situation cannot be remedied without incorporating the three constraints in a single program.


LEGRND:
$\rightarrow$ TO FILE IS (DESAP -II)
NASAP


SIBROUTTANE


FIG.1F


FIG. IC


FIG. Id


DESIGN CYCLES

FIG. 2
*** FILE MOBIFICATEONT INVOLVE CHANGES IN DESIGN VARIABLES AND VARIOUS TOLERANCES













$\qquad$
$\qquad$
$\qquad$

SUBROUTINE CANDP ENTRYPOINT COO663


| 40101 | 10 |  | SUBFUUTINE CANUPIKUE, NCCARD, NPCARD, COUT1, COUT2,POUT1, POUT21 | 000030 |
| :---: | :---: | :---: | :---: | :---: |
| codus | 2* |  | COMMUN/BSCH/HUMLL8LENSIART(16) | 000030 |
| LOIU4 | $3 *$ |  | COMMUN/HANOEL/KVCMIO (20), KVCPIU(100), MAXMID, MAXPID, NUMHIO, NUMPID | Unco 30 |
| C0104 | 4. | 6 (tix ${ }^{\text {a }}$ | CHANGE GROUF, NO. OF ELEMENIS. | 400030 |
| C0105 | 5 * |  | INTEGEF LOUT1IIOC,6) | 0 COO 30 |
| 10100 | 6* |  | ULMENSLON COUT2 (100,3) | 000030 |
| $\begin{aligned} & 60106 \\ & 40107 \end{aligned}$ | 7* | C ** | CHANGE GROUP. NO. OF PID'S INTEGER POUT | 000030 |
| U0114 | 9* |  | UIMENSION POUT2(25,5) | 400030 |
| 00110 | 10* | C ${ }^{\text {a }}$ + |  | -00u30 |
| 00111 | 11* |  | UIMEASION DUMI201 | 000030 |
| 00111 | 12* | C $\# 4$ | KDE. IMPUT. $S=R O U .6=8 A R, E T C$. | 000030 |



$\square$
$\square$

$\qquad$
$\qquad$
$\qquad$
$\qquad$
-. SUBROUTINE COUNT ENTFY POINT LOGO5. 3



ELURD
$-$

## -FOR,S ELURU,ELORO <br> OFOR,S ELURU,ELORD HSA E3 -07130/8U-09:45:23 120,1

SUBROUTINE ELOHD

STORAGE USEU: CODE(1) OUC165; DATAIUI OCOO73: BLANK COMMON(2) DOOUOU

- EXTERNAL KEFERENCES (BLOCK, NAMEI

UUS NERK3S


| 1.0101 | 10 |  |  | C00017 |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{array}{r} 00101 \\ 20103 \end{array}$ | $\begin{aligned} & 2 * \\ & 3 * \\ & \hline \end{aligned}$ | 6 ** | CHAPME GROUP. NO. OF ELEMENTS INTELER ELDAT11100.231 | cocol 7 <br> 400017 |
| 60104 | $4 *$ |  | UIMENSION ELUAT21100,6) | COCO17 |
| $\mathrm{CCLO}_{4}$ | 5 | c $4 *$ |  | cood 17 |
| LOLOS | 64 |  | INTEGER EIOMAX, HOLDI(23) | c00017 |
| L0106 | 7* |  | UIMENSION HOLD 2101 | 600017 |
| 40106 | 8* | C | ORLERS ELDAT AFAAYS IN ASCENDING OROER OF IEL (ELOATI (L, 1) | 000017 |
| 00100 | 9* | c | HUM. INPUL. HO . UF POWS IN ELOAT, AND ELOAT2 | 000017 |
| L0100 | 10* | $c$ | eluati, eloat 2e on infut, element data arrays. on output, ele nus. have been | cocol 7 |
| L0106 | 12* | 6 | PUI IN ASCENGING ORDER. | 000017 |
| L0100 | 12* | $c$ | NCOL 1,NCOL2.INPUT.NO.OF COLS. IN ELDATI AND ELDATZ | LCuO17 |
| U01u7 | 13" |  |  | 100017 |
| Lol13 | $14 \%$ | 151 | NENC=NENU-1 | coug2 |
| 60110 | 154 | $c$ | PRINT H1, NENU | c00023 |
| LO110 | 16* | 6111 | FORMATI NENUこ',I3) | 400023 |
| -01111 | 17\% |  | IF (AEND.EO. 1 ) 6010160 | 600025 |
| UC113 | 18* |  | EIDFAXEELDAT111.1) | U00027 |
| $\mathrm{CCl}_{4}$ | 19* |  | HRQmMx $=1$ | coun3 |
| -c115 | - ل* |  | LO $152 \mathrm{I}=2$, NLNO | 000041 |
| LOL20 | c1* |  | IFIELDATI(1, HI, GI.EIOMAX) 60 10 153 | 000941 |
| 40122 | 22* |  | 60 T0 152 | COLO45 |
| Cct2) | 23* | 153 | EIUMAX EELDAIU1Lell | 100047 |
| 61124 | 24* |  | HROWMX $=1$ | cou050 |
| 00125 | 25\% | 152 | CONTINUE | cocos |
| -4125 | 264 | c | PRINT 222 , NILNU, NROUMX, EIDMAX | c00055 |
| -10125 | 27* | C222 | FORHAT, NENU, NRCWMAX, EIUMAX $=$, 3 (2X, I 31$)$ | ca0055 |
| L0125 | <8* | c | INTEFCHAHGE HOW IIH LARGEST EID WITH ROw NEND | c00055 |
| L6127 | 29* |  | IFINFOMMX.EQ.NENO, GO 170151 | 100055 |
| L6131 | 30* |  | UO $154 \quad J=1$, NCOL 1 | L00060 |
| 00134 | 31* |  | HOLV! ${ }_{\text {d }}$ | 009105 |
| L0135 | 32* |  | ELCAIIINENO,JJ=ELDATI(NROWMX,J) | 000106 |
| col30 | 33* |  | ELOAT1(NROWMX, J) $=$ HOLDI(J) | 0.00110 |




## SUBROUIINE FORHZ ENIRY POINI GOO141.

STORAUE USEU: CODE 111 000160: DATA(G) TUO131: BLANK COMAONI2) 000000

EXIERMAL REFERENGES (BLOCK. NAME)

| W0c3 | Nrous |
| :---: | :---: |
| $\begin{array}{r} 0, J 4 \\ 0005 \end{array}$ | $\begin{aligned} & \mathrm{NIO} 15 \\ & \mathrm{NIO} 2 \mathrm{~S} \\ & \hline \end{aligned}$ |
| udio 4047 | NROUS NERR3s |



| 40101 | 1* |  | SUBROUTINE FORMZIJJ,KQAD, TYPE, ISTOP, VEID, MAX) | 400006 |
| :---: | :---: | :---: | :---: | :---: |
| colot | 2* | C ${ }^{\text {* }}$ | JI.INPUT.INDEX OF KOAD | 000006 |
| 40101 | 3* | C ${ }^{\text {a }}$ | KOAD.INPUT.CONTAINS CONST. CODES | 000006 |
| U0161 | 4* | C ** | TYPE.INPUT.TYPE OF CURRENT ELEMENT | 000006 |
| 10102 | 5* | c ${ }_{\text {* }}$ | ISTOP, OUTPUTENU. OF ELS P FOR CURRENT CONSTE COOE | 000006 |
| 00101 | 6* | C * | VEID.OUTPUT. Io NUMBERS OF ELEMENTS | 000006 |
| 00101 | 7* | C * | MAX.JNPULECONIALAS MAXIHUK NO. Of ELEMEMIS | 000006 |
| colcl | 8 * | C ** |  | 000006 |
| 40101. | 9* | c $\# *$ | CHANGE GROUP. NO. OF ELEMENIS | 000006 |
| volus | 10* |  | INTEGER VEID(100) | 000006 |
| 00103 | 11* | C ${ }^{\text {¢ }}$ |  | L00006 |
| UC104 | 12* |  | INTEGER KOAD(2), TYPE(2) | 000006 |
| U0105 | 13* |  | $1012=1,100$ | 000006 |
| cujio | 14* | 1 | VEIOIIこ | 000006 |
| 1012 | 15* |  | WRIIE lG, 21 IYPE KOADU11 | 406007 |
| L0121 | 16\% | 9 | FORHAT $2 \times$, 'EMTER NO. OF ', A6, A2, ELS. FOR CONST. COUE', 14) | c00024 |
| 00122 | 17\% |  | REAU (5,711SIOP . | counz4 |
| L0125 | 18* |  | IFIISIOP.GE,MAXI WRIIE $6,2 \mathrm{U}$ ) TYPE, TYPE, MAX | 000032 |
| U0141 | 19* | 20 |  | 400061 |
| 60141 | 21* |  |  | Onan' 1 |
| 00142 | c1年 |  | 2. REEMIER NO. OF ELSE.1. | 400061 |
| 40144 | 22* |  | LFIISTOP.GE.MAX) REAO (5,7IISTOP | 00c061 |
| 00146 | 23* | 7 | FORMAI 11 | C00073 |
| 0147 | 24* |  | WRITE 16,10)ISIOP, TYPE | 000073 |
| C0150 | 65* | 10 | FORMAT12X, 'EMTER THE', 14, 1X, A6, A2.' ELEMENT LD NOS. FOR CONST. COO | 000112 |
| 40156 | 26* |  | IE NT, 1',/5x, 'MAKE THEM THE SAME AS NASTRAN NOS. $\%$ | 400112 |
| 60150 | 27\% |  | 25 X - THEY MAY BE PRINIEU ON A LINEISI SEPARATEO_BY A SPACEISI AND/O | 000112 |
| L0.250 | 28* |  | 3R CUMMA') | 400112 |
| 00157 | 29* |  | HEAU (5,7)(VEIU(K), $K=1,15 T U P)$ | 400112 |


| FURH2 |  |  | OR1E 673080 | PAGE | 2 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 00165 | 3U* | Retukn | 000125 |  |  |
| 60160 | 31* | END | 606165 |  |  |
| ENU FOR |  |  |  |  |  |




$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$








SUBROUIINE NOUPT ENIRY POINT LOU4ST

STORAGE USED: CODE(1) OOG472: DATAIU) 015130: BLAMK COMHON(2) OOCOOO
COMAON BLOCKS:
COUS $8 \mathrm{ACH} \mathrm{UCOO}_{4}$

EXIEGNAL REFERENCES (BLOCK, NAME)

| 0004 | PdREWS |
| :---: | :---: |
| $\begin{array}{r} i v i s \\ w u v \end{array}$ | $\begin{aligned} & \text { NKDUS } \\ & \text { NIOI } \end{aligned}$ |
| - نJ 7 | NIOCs |
| U-17 | N103s |
| $\begin{array}{r} 0011 \\ 0012 \end{array}$ | NmDUS <br> NERR3s |



| $\begin{aligned} & 60101 \\ & 60102 \\ & \hline \end{aligned}$ | $\begin{aligned} & 1 * \\ & 2 * \end{aligned}$ |  | SUBFOUTINE NODPT COMMCN/BACH/PUM(2B),NSTART(18) | 000000 100000 |
| :---: | :---: | :---: | :---: | :---: |
| L0103 | 3* | C ** | CHANGE GKOUP. NO. OF DESAP BAR ELS. INASTRAN RODI | vodoud |
| LC104 | 4* |  | INTEGER CROD (1) ${ }^{\text {de2 }}$ | cougoo |
| Lu104 | 5* | - ** | CHATIGE GKOUP. NO. OF NODAL POINTS | coleos |
| 00105 | 6* |  | - NILCER SHARI 3 30U, Y) | cooguo |
| $\begin{aligned} & 00100 \\ & 20100 \end{aligned}$ | 7* | C** |  | coonou conoun |
| $\begin{aligned} & 41147 \\ & 40107 \end{aligned}$ | 9* $10 *$ | c $\# *$ |  | Lonoogo <br> 400040 |




| NOUPI |  |  |  | DAIE 073080000353 | PAGE 4 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 60342 | 125* |  | 0061 L2E1. NUMNP |  |  |  |
| $60345$ <br> LO347 | $\begin{aligned} & 126 \% \\ & 127 \% \\ & \hline \end{aligned}$ |  | IF, NGDPTIIL2,1).EO.CROD (L1,(4)) GO TO 62 609061 | $400353$ |  |  |
| $\begin{aligned} & 60350 \\ & 60353 \end{aligned}$ | 128* | $\begin{aligned} & 62 \\ & 63 \end{aligned}$ | $\begin{aligned} & \text { UO } 63 \mathrm{~L} 3=5,7 \\ & \text { NOUR } 11(L 2,13)=1 \end{aligned}$ | $\begin{aligned} & 100360 \\ & 000365 \end{aligned}$ |  |  |
| 40355 40357 | 130* | 61 | CONTINUE | 000375 |  |  |
| L0357 | 1314* | 64 | CONIINUE | 000375 |  |  |
| 60361 | 133* | $c$ | PRINT 200 | 000375 |  |  |
| co361 <br> 40361 | $\begin{aligned} & 134 * \\ & 135 * \end{aligned}$ | $\begin{aligned} & \mathrm{C} 200 \\ & \mathrm{C} \quad \geqslant \% \end{aligned}$ | FORMAT: FINAL OUTPUT OF NODPT."I <br> COMBINES NOOPIL ANO SHARZIC FORH DESAP NODAL PT .DAIA | $\begin{array}{r} 400375 \\ 000375 \end{array}$ |  |  |
| C0363 | 136* |  | WRITE112.9031 | U00375 |  |  |
| 141365 | 1374 | 903 | FORMAT1' COHM NODAL POIKT DATA FROM NOOPT', | 000403 |  |  |
| 40366 | 138\% | 959 | DO 960 K=1, NUMNP | C00403 |  |  |
| L0371 | 139* | 960 |  | 000410 |  |  |
| 40371 | 14.7 | C960 |  | 000410 |  |  |
| 60400 | 141\% | 901 | FORMAI $1715.3 F 10.3,15.510 .31$ | 000436 |  |  |
| $\begin{array}{r} 00407 \\ -0416 \\ \hline \end{array}$ | $\begin{array}{r} 142 * \\ 143 * \\ \hline \end{array}$ |  | RETURN END | 000436 |  |  |
| ENL FOR |  |  |  |  |  |  |
| OHOG.P URDER |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

$\qquad$
$\qquad$
$\qquad$
$\qquad$

-'

HSA E3 - URULF,ORDER

## SUBROUTINE ORUER ENIFY POINT GOCSJZ

STORAGE USEU: CODE(1) DOOSG1: DATAIOI 000570: BLANK COMMONI2I DODOOO
COMMON BLOCKS:

| U0゙, 3 | $\triangle A C H$ | 600044 |
| :---: | :---: | :---: |
| 1004 | MOZART | cou334 |
| - Uus | HANLEL | -Cu1 |

- EXIERNAL REFERENCES (BLOCK, NAMEI


| 60101 | 14 | SUBEMULIAE OKOLR | 400002 |
| :---: | :---: | :---: | :---: |
| u0103 | 2* |  | LOCOU2 |
| 00103 | 34 | 1/AKVCMDI2¢), KVCPILIDGOI, MAXMID, MAXPIU, NUMMIU, NUMPIC. | -00ncz |












$\qquad$
$\qquad$
$\square$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$



$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$




| 00322 | 114* | 302 |  | 000377 |
| :---: | :---: | :---: | :---: | :---: |
| C0322 | 115* | 1) READ 5,7$)$ (GEOPR21J12, ${ }^{\text {a }}$ (12=2,61 |  | 000377 |
| 60323 | 116* |  |  | 000377 |
| 00326 | 117* | 283 | CONTINUE <br> WRITE(12.96) | 000411 |
| 40334 | 118* |  |  | $0_{00411}$ |
| co332 | 119* | 96 | FORHATI' COMM GEOMETRIG PROPERTIES FROM QDHEMI' | 000422 |
| 60333 | 120t |  | 00299 J10 1 , NUHGEO | 000422 |
| 00330 | 121* | 299 | WRITE(12,303IINDEX(J10), (GEOPR2(J10,J11), J12=1,6) <br> FORMAT115,6F10.01 | cou422 |
| L0346 | 1<2* | 103 |  | 000440 |
| 00346 | 1<3* | C ** | GENERATE ELEMENT LOAD NULTIPLIERSWRITEIS. | 000440 |
| L0347 | 124* |  |  | 000440 |
| LO351 | 125* | 111 | FORMATIIUX, ELEMENT LOAD MULTIPLIERS FOR CONST. CODE 1')DO $115 \mathrm{SG}=1,4$ | 000451 |
| L11354 | 1<6* |  |  | 000451 |
| LO355 | 127* |  | $00115 \times 7=1,4$ | C00451 |
| 40364 | 128* | 115 |  | 000451 |
| 40363 | 129* |  | ט0 $120 \mathrm{k}=1,4$ | 000461 |
| C0366 | $130 \%$ |  | MRIIE (6.112) YAR(k8) | 000464 |
| L0371 | 131* | 112 | FORMATI2X, 'FUR LOAU ', A1,' ENTER X, Y, 2 ELEMENT LOAD FRACTIONS, IN | 000472 |
| 10371 | 132\% |  | ITHAT OROER DEFAILL = ALL ZEROS ${ }^{1} 1$ | 000472 |
| 60372 | 133* |  | READ 5,7 ( DUMM(K111, K11 $=1,3$ ) | 000472 |
| 00375 | 134* |  | IFIOUMM(X).EGellelgo 10 - 20 | 000502 |
| 60377 | 135* |  | $00113 \mathrm{k} 11=2,4$ | c00511 |
| 6040: | 136* | 113 |  | 006511 |
| UC404 | 137* | 120 | CONTINUE | L00515 |
| 10406 | 1387 |  | HRIIE112.1117) | 000515 |
| 40410 | $137 \%$ | 1117 | FORMAT' COHM ELEMENT LOAD MULTIPLIERS FROM GDMEMI' UO $121 \quad k 9=1,4$ | 000525 |
| $6{ }^{6} 411$ | 140* |  |  | 000525 |
| CO414 | 141* | 121 | WRITE(12,116)(EMUL (K9,K10),K10=1,4) | 000525 |
| 10423 | 142* | 116 | FORMAT (4F10.0) | 100541 |
| 0.0423 | 143* | C \# ${ }^{\text {a }}$ | GENERATE ELEMENT DATAHRITE 6 , 2 II) | 000541 |
| 00424 | 144: |  |  | 000541 |
| 00426 | 145* | 127 |  | 000551 |
| 60427 | 146* |  |  | 000551 |
| C0432 | 147* |  |  | 000551 |
| 60433 | $148 \%$ |  | WRITE16, 130) IX | 000552 |
| 40436 | 149* | $13 n$ | FORMAI(2X,'ENTER OESIGN VARIABLE NO., IUV,FOR EL.NO.', I5) READ $15,71 C O D M D R E K R O K E L$ | 000560 |
| C0437 | 1513. |  |  | 000560 |
| 00442 | 151* |  | WRITE16,1311IX <br> FORMAT $2 x^{\prime}$. ENIER DESIGN VARIABLE FRACIION.FFG.FOR EL.NO.'.I4.'. DE | 000566 |
| 60445 | 1564 | 131 |  | 000574 |
| 40445 | 153* |  | 1FAULT=1.' | L00574 |
| 60446 | 154* |  | LLDAT 2 CRROHE $11=1$. | 000574 |
| L0447 | 155* |  | READ (5,7)DUM | 000576 |
| 10452 | 156* |  | LECDUMENE, | 100604 |
| L0454 | 157\% |  | ELDATZ(KROW, こ) $=0$. | 400610 |
| 10455 | 158* |  | $\text { WR1TE }(6+132) 1 x$ | 400611 |
| $\mathrm{LO}_{4} \mathrm{G}$ <br> 10461 | $\begin{aligned} & 159 * \\ & 160 \% \\ & \hline \end{aligned}$ | 132 | KEAOL5,7)ELOAT2(MRONE3) | 000617 000617 |
| 10464 | 161* |  | WRITE(6,133)IX | 400625 |
| 60467 | 162\% | 133 | FORMAL $2 x$, ENTER EFC FOR EL NO. ${ }^{\circ}$. | 1.00633 |
| L0470 | $103 *$ |  | ELUAT2(KHOW, $+1=1$. | 400633 |
| 10471 | 164* |  | REAU15, ${ }^{\text {d }}$ IOUM | 100635 |
| 10474 10476 | 165* |  | IFIOUM.NE.O.) ELDAT2(KROW,4) $=U_{U M}$ HRITE(6,134)IX | $600643$ |
| U05J1 | 167* | 134 | FORHATI2X, 'EIHER SIRESS PRINTOUT CODE,KS,FOR EL.NO.',I4, ', DEFAULT | L00655 |
| 0.501 | 168* |  | $1=3 \% 1$ | 1406655 |
| 40502 | 167* |  | COOMIOIKROW, yl $=3$ | 1:00655 |
| 60503 | 170* |  | KEAL $5,7110 \cup M$ | 000657 |




|  |  |  |
| :---: | :---: | :---: |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

-1 ________
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$



| 40175 | 59* |  | CALI PAIRIKKODE, CODMID, PGOHIEISIOP, NUH | 100131 |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{aligned} & 40170 \\ & 60177 \end{aligned}$ | $\begin{aligned} & 63 * \\ & 61 * \end{aligned}$ |  | $\begin{aligned} & \text { NPAR(3)=NUMMAT } \\ & \text { WRIE }(6, S D I) \\ & \hline \end{aligned}$ | $\begin{array}{r} 000144 \\ 000146 \\ \hline \end{array}$ |
| C0201 | 62* | 501 | FORMATI2X, ENTER CODE FOF INCOMPATIBLE DISPL. MODES: 0 MEANS USE $T$ | 000153 |
| c0201 | 63* |  | IHEM: 1 MEANS DO NOT USE THEM: | 000153 |
| C0202 | 04* |  | REAO(5,7INPAR(6) | 600153 |
| 00205 | 65\% |  | HRITE 12.7951 | C00161 |
| 00207 | 06* | 795 | FORMAII COMM ELEMENT CONTROL FROM GDMEM2'1 | 000166 |
| 00210 | 67* |  |  | C00166 |
| 60213 | 68 * | 504 | FORMAT(6I5) | 000210 |
| c0213 | 69* | C * | SET UP MAIERIAL PROPERIIES DATA | 000210 |
| L0214 | 70* |  | LO 83 JIU=1, NUMMAT | 000210 |
| 60217 | 71* |  |  | 000211 |
| 40220 | 72* |  | MTPRO1 ${ }^{\text {dia }}$, $21=1$ | 000212 |
| CO221 | 73* |  | DO 84 J11 $=1$, NUMMAT | 000217 |
| 40224 | 74* |  | IFIMATCHIJII, I).EO.INDEX(J10)IGO TO 85 | 400217 |
| 40220 | 75* |  | 601084 | 000222 |
| L0227 | 16* | 85 | NROW = MATCH(J11,2) | 000224 |
| 00236 | 17* |  | 60.10-86 | 000225 |
| UC231 | 78* | 84 | continue | 000231 |
| 10233 | 79* | 86 |  | 000231 |
| LU234 | 80\% |  | MTPRO3(Jlo.1) $=0$. | 000234 |
| C0235 | 61\% |  | MTPRO3, $110,21=$ MA 121 (NROW, 11 | 000235 |
| 40236 | 82\% |  | MTPRO3(JIO, 3 ) $=$ MAT12(NROW,3) | 000237 |
| 10237 | 83* |  | HTPRO31-10,4) $=0$. | 000241 |
| L024 | $84 *$ |  | MTPRO3(J10,5) = MAT12(NROW, 8 ) | 000242 |
| L. 241 | 85* |  |  | 000244 |
| 40242 | -6* |  | MTPRO3 (J10, 7 ) = MATİ(NROW, IU) | CDO246 |
| 00243 | 87\% |  | IX =INDEX(JIE) | 000250 |
| 00244 | 88* | 83 | CONTINUE | 000254 |
| 00244 | 69* | c ${ }^{\text {¢ }}$ | MAIERIAL PRQPEKIIES DATA IS DONE | 600254 |
| 40246 | 90* |  | WRITE112,349) | 000254 |
| 40256 | 41\% | 349 | FORMAT1' COMM MAIERIAL PROPEEILES FROM OOMEM ${ }^{\text {C* }}$ | 400266 |
| 00251 | 92 * |  | UO 350 J12=1, fiUMmat | U0C266 |
| 20254 | 93* | 350 |  | 000266 |
| 40254 | 94* |  | 1), $113=1,71$ | 000266 |
| 10270 | 95\% | 310 |  | c00314 |
| 10270 | 96* | C | GEnEKATE ELEmENt load multipliers | 400314 |
| (27) | 97* |  | WRITE10.111) | cou314 |
| 40273 | 98* | 111 | FORMAIIIUX, 'LLEMENT LOAD MULTIPLIERS FGR CONSt. Cook 2\% | u06325 |
| 10274 | 99* |  | 10 O $115 \quad k 6=1,4$ | 000325 |
| 00277 | 100 |  | ט0 115 $k 7=1,5$ | uco 325 |
| 1030 | 101\% | 115 | EMUL $1 \mathrm{~K} 6, \mathrm{~K} 71-2$. | 600325 |
| 40305 | 162* |  | U0 $120 \mathrm{~K} 8=1,4$ | L00335 |
| $\mathrm{CO}_{0}$ | 153* |  | URITE(G, 112)VAR(kg) | 100340 |
| L0313 | 104* | 112 | FORMAIIZX, 'FUR LOAD ', Al, ' ENTER PRESSURE ARO X, Y,Z ELEMENT LOAD F | cou346 |
| 40313 | 105: |  | IRACTIOIS IN THAT ORDER. LEFAULI ALL ZEROS.1 | CJU346 |
| L0314 | 1.6* |  | REAL (5,7) (OUMM(K11), K11 $=1,4$ ) | L00346 |
| 00317 | 167* |  | 1FIOUMM11.Er.V.IGU 10 120 | U0L356 |
| LO 321 | 108* |  | D0 113 K 11=2,5 | 400365 |
| 4 L 324 | 157* | 113 | EMUL (K8, K11)=OUMM(K11) | C04365 |
| L0326 | 11〕* | 120 | CONTINUE | 600371 |
| L0330 | 111* |  | WRITE112,1117) | U04371 |
| U0332 | 112* | 1117 | FORMAT ${ }^{\text {- }}$ COMM ELEMENI LOAD MULIIPLIERS FROM GOMEMZ') | (100401 |
| 40333 | 113* |  | U0 12 $6 ¢=1,4$ | 000401 |
| 60330 | 114* | 121 | WRITE(i2,116)(EMUL(K),K1C),K10=1,5) | LCU401 |
| 10345 | 115* | 116 | FORMAT!5F10.6) | 000415 |







##  

$\qquad$

$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$

SUBROUTINE SHEAR ENIRY POINT DOU527


EXtefnal heferences iblock, names









HSA ES－U7130／8こ－09：47：12（14，
$\qquad$
SUBFOUTINE VEC ENIRY POINT 000053

EXIERMAL HEFERENCES GLOCK，NAMEI
－LUJ4 NERK3S
$\qquad$


| 41101 00101 | 1＊ |  | SUDFOUTINE VECIMID，PIDI | cob010 |
| :---: | :---: | :---: | :---: | :---: |
| 60101 | $3 \%$ | C＊＊ | OF THE MID ANO PID NO．S | 600010 |
| ＋0101 | 4＊ | ع ${ }_{\text {＊}}$＊ | PIUIMIO）．INPUT，IF ALREAOY IH KYCPIDIKYCMIUI，RETURN． | 600010 |
| C0101 | 5＊ | C＊＊ | If Not，Put in place of first zero． | L0U010 |
| Uc123 | 6＊ |  | COMMCN／HANDEL／KVCMID（20），KVCPID（100），MAXMID，MAXPID，NUMMID．NUMPID | 100010 |
| L0104 | 7＊ |  | INTEGER PID | 000010 |
| LClu4 | －${ }^{\text {a }}$ | C | PRIAT 599 | 000010 |
| C0144 | 9＊ | C599 | FORMATi IM IN VEC DOING MY THING＇I | jocodo |
| 60104 | 10＊ | $c$ | PRINT OQU，MIC，PID | 000010 |
| 40104 | 11＊ | ceut | FORMAT（＇MID AND PID IN VEC＇，2IS） | c00010 |
| $\mathrm{COHO}_{5}$ | 12＊ |  | $1056 \mathrm{I}=1, \mathrm{MaXNI}^{10}$ | 400010 |
| －011u | 13 \％ |  | IFIKVCMIU（I）．EL．C）GO TO 5 | c00010 |
| 10112 | 14＊ |  | IF | 100012 |
| 00114 | $15 *$ |  | 601050 | c00025 |
| 40115 | 16＊ | 5 | KVCMIU1II $=$ M1L | 600017 |
| C0116 | 174 |  | 6010100 | 000020 |
| 00117 | 18＊ | 50 | COMIINUE | 000026 |
| LO121 | 19＊ | 140 | CONTINUE | co0026 |
| C0121 | てい＊ | $c$ | PRILI 101 | UC0026 |
| L012c | c1＊ | 101 | FORMATI＇at start of znu loup in vec． | 000026 |
| 00123 | 224 |  | VO $150.1=1, \mathrm{MAXPIU}$ | 000026 |
| 60126 | 234 |  | IFIKVCPIOIII．EG．3） 60 to 105 | 400026 |
| 10134 | －4＊ |  | IF（PIW．EO．KVCPIU！I）JGO TO 5un | 000030 |
| C0132 | ＜5＊ |  | 60 1015 S | 000033 |
| 00133 | ＜6＊ | 105 | KVCPIU（I）＝PIL | L00035 |
| $\begin{array}{r} 60134 \\ 60135 \end{array}$ | 27\％ | 15u | GO TO 50 continue | $000036$ $000042$ |

rec

$\qquad$
$\qquad$
$\square$
$\qquad$
$\qquad$
$\qquad$
$\qquad$
$\qquad$

## CONCLUSIONS AND RECOMMENDATIONS

Currently NASAP transforms CAD-generated NASTRAN data to data for DESAP I or DESAP II with minimum input from the engineer. The program has been thoroughly checked.

An effort is underway to implement the design cycles shown in Figure 2. Such a capability will greatly enhance the effectiveness of DESAP I and DESAP II by allowing the simultaneous satisfaction of stress, displacement, and buckling constraints.

## REFERENCES

1. Kiusalaas, J. and Reddy, G. B.; DESAP II: "A Structural Design Program with Stress and Buckling Constraints" Department of Engineering Science and Mechanics, The Pennsylvania State University, University Park, PA, April 1976.
2. Kiusalass, J. and Reddy, G. B.; DESAP I: "A Structural Design Program with Stress and Displacement Constraints" Department of Engineering Science and Mechanics, The Pennsylvania State University, University Park, PA, April 1976.

# NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM 

## MARSHALL SPACE FLIGHT CENTER THE UNIVERSITY OF ALABAMA

# DEPLOYABLE STRUCTURE DESIGN FOR THE SCIENCE 

 AND APPLICATIONS SPACE PLATFORMPrepared by: Henry W. Stoll, Ph.D.

Academic Rank:
University and Department:
NASA/MSFC:
Division:Branch:
Special Projects
MSFC Counterpart: James K. Harrison
Date: August 1, 1980
Contract No.: NGT-01-002-099(University of Alabama)

DEPLOYABLE STRUCTURE DESIGN FOR THE SCIENCE AND APPLICATIONS SPACE PLATFORM

By<br>Henry W. Stoll<br>Associate Professor of Mechanical Engineering Univeristy of Wisconsin - Platteville Platteville, Wisconsin

## ABSTRACT

NASA has long been interested in deployable structure technology as a means for achieving efficient spacecraft packaging for launch. An important application currently under study is the Science and Applications Space Platform (SASP). In this study, basic concepts regarding deployable structures design, including systematic design/classification schemes and a "deployability" criterion, are proposed for use in synthesis, analysis and evaluation of alternative deployable structure designs. Using design guidelines based on SASP requirements and the basic concepts developed, a variety of new designs are synthesized, and these along with previously proposed designs are analyzed and evaluated. Recommendations and conclusions regarding optimal deployable structure design are made based on insights gained in the study.

NASA has long been interested in deployable structure technology as a means for achieving efficient spacecraft packaging for launch. An important application currently under study by NASA involves the Science and Applications Space Platform (SASP). A representative configuration (Fig. 1) has been evolved which satisfies the SASP objectives of providing accommodation for the simultaneous operation of several payloads of various science and application disciplines. The configuration incorporates three independently oriented platform arms, dedicated to celestrial, solar, and Earth viewing, respectively. Each platform arm provides interface provisions for payload carriers at two locations, and each location has provisions for two interfaces on opposite sides of the platform arm. The platform arms utilize deployable structures which allow compaction for launch in the Orbiter cargo bay.

A variety of deployable structure concepts have been proposed over the years (see Ref. [1]* for example) to meet different spacecraft packaging needs. In addition, several specific design concepts have been proposed for use in SASP, some of which are presently under development by NASA/MSFC. There does not exist, however, any clear-cut agreement regarding a preferred design approach for SASP, nor does it appear that all viable design alternatives have been identified. This report documents a 10 -week, one-man effort, conducted between late May 1980 and August 1980, to take a fresh look at the deployable structure design problem in light of SASP design goals and requirements. Emphasis in this study has been placed on the mechanism aspects of the deployable structures problem. Structural considerations have been addressed only in so far as they impact the mechanism problem.

## II. OBJECTIVES

The principal objectives of this study are (1) to completely define the deployable structure design problem in light of SASP design goals and requirements; (2) develop basic concepts relating to design morphology, systematic synthesis, deployment method, and performance criteria such as deployability, compactability, etc. which can be used to assist in the synthesis, analysis, and evaluation of deployable structure designs; (3) identify alternative deployable structure designs which appear to hold promise for the SASP application;

[^1](4) analyze and evaluate new and previously proposed designs; (5) investigate various detail design considerations relating to structural dimensions, control of structural stiffness and damping, etc.; and (6) summarize findings in the form of recommendations for optimal design.
III. DESIGN GUIDELINES

For the purposes of this study, it is assumed that, to be acceptable, a deployable structure design should conform to the SASP design goals and requirements as currently envisioned in reference [2]. These requirements are summarized as follows:
(1) State-of-the-Art Constraints. The design should utilize current design techniques and materials in order to minimize the need for development of new technologies and the use of unproven space operations.
(2) Astronaut Participation. Deployment utilizing an astronaut in EVA mode should be avoided. The astronaut may be utilized in a back-up role.
(3) Retract Capability. In order to provide maximum flexibility and contengencies to conduct mission operations, it is desirable, but not necessary, that a retract capability of the deployed structure on orbit be provided.
(4) Payload Utilities Accommodation. The platform structure should facilitate means for distribution of payload utilities in the form of a separate electrical harness to each payload carrier interface (Fig. 1). Each harness will consist of $4-1 / 0$ power cables plus other electrical and communication leads.
(5) Launch Packaging. The envelope of the platform arm in its compacted form must be such that the three platform arms can be accommodated within the Orbiter cargo bay with each of the arms connected to the SASP platform support module.
(6) Misalignment and Distortion. The maximum misalignment between the payload carrier interface and the platform arm mechanical interface at the rotary joint should not exceed +1.0 degree. This alignment requirement is to include effects of fabrication tolerances, joint mechanical dead band, and interface misalignment. Dynamic stability of the platform arm alignment will be held to $\pm 0.1$ degree, including effects such as cyclic thermal distortion, structural distortion due to environmental and induced loads, and mechanical dead band.
(5) The deployment drive should be small, mechanically simple, and straight forward in its operation.
(6) A retract capability is desirable.
C. Deployed Structure Considerations
(1) The deployed and rigidized structure should have a minimum of dead-band or other unwanted relative motion.
(2) Structural properties should not degrade or otherwise vary due to fluctuating loads or other adverse conditions inherent in the space environment.
(3) Mounting points of the deployed structure should lie in a plane normal to the longitudinal axis of the structure in order to permit interfacing of the structure with the platform support module and the payload carrier interface.
D. General Considerations
(1) The overall deployable structure concept should be mechanically simple and free from unnecessary design complexities.
(2) Structural links and connections should be easily manufacturable from materials suitable for space applications.
(3) The design should be insensitive to manufacturing error and tolerance build-up; close tolerances should be avoided.

## IV. Basic Concepts

The purpose of this section is to develop some basic concepts which can be used to assist in the synthesis, analysis, and evaluation of deployable structure design. As a starting point, it is useful to note that deployable structures operate both as mechanisms during the deployment stage and as structures after they are fully deployed. Conversion of the deployable structure from a mechanism into a structure is facilitated by hardening or freezing one or more of the kinematic pairs which connect the structural links together. It should also be noted that, depending on the deployable structure design, deployment

TABLE 1. Possible Link-Joint Combinations to Give a Specified Mobility Based on Equation (1)

| $m=0$ |  | $m=1$ |  | $m=2$ |  | $m=3$ |  |
| :---: | :---: | :---: | ---: | ---: | ---: | ---: | ---: |
| $n$ | $j_{1}$ | $n$ | $j_{1}$ | $n$ | $j_{1}$ | $n$ | $j_{1}$ |
| 3 | 3 | 4 | 4 | 3 | 2 | 4 | 3 |
| 5 | 6 | 6 | 7 | 5 | 5 | 6 | 6 |
| 7 | 9 | 8 | 10 | 7 | 8 | 8 | 9 |
| 9 | 12 | 10 | 13 | 9 | 11 | 10 | 12 |
| 11 | 15 | 12 | 16 | 11 | 14 | 12 | 15 |

Examination of the deployable, three-sided planar truss structure of Figure $2 a$ shows it to be unsatisfactory with regard to interfacing and mounting needs. Altering link proportions and mounting point locations gives the configuration of Figure 2b. This design is unsatisfactory structurally, but its square ends meet interfacing and mounting requirements. from Table 1, the next possible mechanism with a mobility of 1 has $n=6, j_{1}=7$. Using this combination results in the foursided planar truss structure depicted in Figure 2c. This configuration is acceptable structurally and meets interfacing and mounting needs. Also, like the three-sided truss (Figure 2a), it requires only one imput motion for constrained deployment and can be converted to a structure by freezing only one joint.

Efforts to synthesize deployable structure designs utilizing more complex arrangements and numbers of links and joints resulted generally in designs which, if they were acceptable structurally, turned out to be various combinations of the three and four-sided planar structures characterized by Figure 2a and 2c. This insight suggests that the simplest deployable structure designs consist of a series of repeating cells, each of which are based on very simple planar structures. This conclusion is substantiated by the fact that most practical designs proposed to date are applications of this approach.
B. Morphological Analysis. Having decided that the simplest and therefore most practical deployable structures are made up of a series of repeating cells, the next step is to carefully examine the morphology of the simple planar structure or cell. Figure 2 shows that each simple planar structure is composed of three different structural links, namely the longeron, the lateral, and the diagonal. Consider now that each of these links can be constructed to have one of five different properties. A link can be rigid in which case its length is fixed; it can be telescoping in which case its length is
changeable; it can fold at its midspan such as the longeron of Figure 2 a or the diagonal member of Figure 2c; it can be flexible in which case it is rigid in tension but buckles elastically in compression; or it can be separable in which case parts of the link are not contiguous in the compacted mode.

Different joint types can also be used to connect the various structural links. Observing the single-degree-of-freedom joint restriction limits joint type selection to pinned joints (revolute pairs) and sliding type joints (prismatic pairs). A screw pair can be substituted for a slider to facilitate actuation. In the case of flexible diagonal links, ball and socket joints and wrapping pairs (e.g., cable and drum, belt and pulley, etc.) can be utilized as long as the structure remains properly constrained. Also, properly designed flexural pivots or plastic hinges can be used in place of pinned or ball and socket joints. Finally, similar to the separable link, joints can be made detachable to facilitate more efficient compaction of the structure.

It should be noted that, kinematically speaking, the telescoping link and sliding pair are identical. However, from the standpoint of visualizing deployable structure design, the ability of the telescoping link to change length is an important property and hence it is treated separately from other sliding pairs. In mobility determinations, the telescoping link should be considered as two rigid links joined by a sliding pair. Similarly, the folding link is acutally two rigid links connected by a revolute pair. But, for the purposes of visualizing deployable structure design, it is treated as a particular link type. As will be shown subsequently, this deviation from conventional kinematic thinking leads to a powerful systematic synthesis procedure.
C. Deployment Methods. Deployment of the structure can be facilitated (1) manually using astronauts in EVA, (2) mechanically using remote manipulators, teleoperators, etc.; or (3) automatically using a self-contained drive system. All deployment methods other than the automatic methods may be classified as "erectable" methods. Automated deployment methods may be further subdivided as follows:

Case 1. The entire, multi-cell deployable structure is designed as a linkage having a mobility of 1. Hence, only one link need be driven to produce constrained motion during deployment.

Case 2. Actuating links in each cell are interconnected such that actuation of one link produces quasi-simultaneous deployment of all cells.
angles, and other parameters of mechanisms which give insight as to whether a mechanism is a good one or a poor one are commonly used. Examples of these include mechanical advantage, transmission angle, and pressure angle. Many of these have a number of features in common, including the fact that most can be related to the velocity ratios of the mechanism and therefore can be determined solely from the geometry of the mechanism. However, most also depend upon some knowledge of the application of the mechanism, especially of which are the input and output links.

In the case of deployable structures, a more general "index of merit" is required since there are no clearly defined output links. A survey of the literature indicates that the index proposed by Denavit [4] may be useful as a measure of deployability. This index of merit is the determinant ( $\Delta$ ) of the coefficients of the simultaneous equations relating the dependent velocities of a mechanism. When this determinant becomes small, the mechanical advantage also becomes small and the deployability of the structure is reduced. This same determinant also appears in the denominator of the dependent accelerations and all other quantities which require taking derivatives of the loop-closure equation. Consequently, it is true in general that, if this determinant is small, the deployment mechanism will function poorly in all respects - force transmission, motion transformation, sensitivity to manufacturing errors, and so on.

The deployability index ( $\Delta$ ) proposed above depends on the deployable structure geometry, link dimensions, and on which link is driven to effect deployment (i.e., rotation of a link, contraction of a telescoping member, etc.). Consequently, the deployability index can be used to determine the most suitable actuation method for a particular design and to compare deployability potential of alternative designs.
F. Design Complexity. Design complexity is a nebulous concept which is difficult to quantify and yet it has an important bearing on the acceptability of alternative deployable structure designs. In general, it is safe to say the the probability of achieving high deployment reliability as well as meeting deployed structure alignment and stiffness design goals will decrease with increasing design complexity. When examined carefully, design complexity is found to involve a variety of factors such as manufacturing methods, materials, space, and economics. In the case of deployable structures, these considerations correlate well with the number of kinematic links, $n$, the number of joints, $j 1$, and the mobility, $m$, of the structure. Mobility is involved because deployment system complexity increases with increasing degrees-of-freedom of the

Examination of this approach shows that, although the four considerations proposed are valid, they are not independent of each other in every case. Hence, the real driver for systematic synthesis in this approach is found to be the number of different link type combinations which are available. That is, the five link types (rigid, fold, telescoping, flexible, and separable) taken in groups of three (lonqeron, lateral/diagonal, and diagonal) results in a total of $5^{3}=125$ different possible combinations. Further variatons are possible by using different combinations of pinned and sliding joints as well as other joint types.

Possible link type combinations were organized by grouping the three-sided planar truss links (Fig. 2a) in the order of longeron-diagonal-diagonal and grouping the four-sided planar truss in the order of longeron-lateral-diagonal. Each link type was given the letter designation shown in Table 2. This notational convention gives the following 25 column by 5 row array of possible link type combinations for both the three and foursided planar truss cell.

In addition to systematic synthesis, the organizational scheme of Table 2 suggests a convenient method for classifying or coding various deployable structure concepts. Employing the notation of Table 2 gives the following alpha-numeric code,

|  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

a) MSFC-HYAxID [6]


ClASSIFICATION: $1-4-\operatorname{AaC}-2$
Mosicirí: $\quad m=1 /$ cecl
comppctacuiry : $\frac{L}{V} \approx\left(\frac{h}{2+w}\right)\left(\frac{l / h}{l / h+1}\right)$
Cumpsermy <cere*:

$$
m=5, R=6,5=1
$$

b) MSFC - SC1550RS [11]


CLASSIPICATIAN: $1-4-B C A-1,2$, OR 3
moblury: $m=1$ for t-cell stricture
Cumpoctarenery: $\quad \frac{L}{V}=\left(\frac{k}{3+w}\right)\left(\frac{\ell / h}{\sqrt{(2 / h)^{2}+1}}\right)$
Cospecxity/CELL ${ }^{4}: m=9, R=11,5=1$
C) MDAC - FOLDAEE TNVSS [7]


CLASSIFICATIUN: $\quad 1-4-B A C-3$
MOBMTVX ${ }^{*}=$ MIV $3 / C E L L$

Coneractafury: $\frac{L}{v} \simeq\left(\frac{d z}{4+w}\right)\left(\frac{l}{h}\right)$

d) MSFC-PD EXPתNIAADEE TRUSS [5]


CLASSIFICATION: 1-3-BAA-1
MOBlLNY*: $m=1$ for $(A+1)$ ccel sreveruee
compnctachery: $\frac{c}{V} \approx\left(\frac{\hbar}{3+w}\right)\left(\frac{\lambda / h}{\sqrt{\frac{1}{4}(1 / h)^{2}+1}}\right)$
Complex)ry/ct<t : A1 $=6, R=6,5=0$
NOTE: -..- "Seissoes" Dertovenient Liwhabe

* moviliv AND COMPIEYITY ARE COLCUEATED FOR menaene sthucture.

Figure 3. Previously Proposed Deployable Structure Concepts for SASP

$$
\mathrm{xXxIf}-15
$$



$$
m=1 \text { fie eviry } 2 \text { crles }
$$

a) DIAGONAC/LATINNE SLIDEK (1-4-AAA).

b) LONGETON /DIAGUNAS SEIDEA (1-リ-ARA).

c) Lonction (Lateral slidec (1-4-anA).

d) CQuIVICENT 1-4-ACA DESIGN

Figure 4. Alternative 1-4-AAA Designs

c) $1-4-B A D-3 D E S 1 C N$

* Mobility ! Compactability are calculated sore planar structure

Figure 5. New Deployable Strucutre Concepts for SASP Application


Loop Closure Equations:

Figure 6. Deployability Index Calculation for MSFC-Hybrid Deployable Structure

Use of equation (3) is illustrated in Figure 6. Here it is seen that one cell of the MSFC-Hybrid design (Fig. Ba) contains two loop-closures and hence, its deployability index is the product of the "index of merit" or $\Delta$ calculated for each loop-closure. Since this design utilizes Case 2 deployment, the deployability index for the cell is also the deployability index for the structure as a whole.

Examination of the large variety of deployable structore design concepts synthesized in this study shows that each design is made up of different combinations of basic mechanisms or linkages, where each linkage defines a particular loopclosure. The basic linkages comprising the selected designs of Figures 3 and 5 are presented in Figure 7. The deployability index ( $\Delta$ ) for each possible input motion to these linkages has been derived as a function of the link dimensions 1 , $h$, and $d$ and the degree of deployment, $x^{*}=x / 1$, where $0 \leq x^{*} \leq 1$. In addlion, each deployability index has been plotted as an function of percent deployment.

The results depicted in Figure 7 can be used in conjunction with equation (3) to gain considerable insight regarding detail design of deployable structures. Remembering

It should be noted that a large $1 / \mathrm{h}$ ratio (or for a given h, a large $l$ dimension) not only improves deployability, but also decreases the number of cells required to produce $a$ required deployed length, L. Hence, design complexity of the overall structure will be reduced by using large $1 / h$ ratios.
B. Compactability Analysis. Based on the design guidelines, a high degree of compactability is not needed in the SASP application. Nor is high compactability particularly desirable in this application because of complications arising due to handling of the utility cables. It is informative, however, to examine the sensitivity of compactability to changes in structural dimensions. Using the ratio of deployed length to compacted volume (L/V) as a measure of compactability, it is found that, for all designs, the magnitude of $L / V$ and therefore, compactability, is increased as the $1 / \mathrm{h}$ ratio is made larger. This is at least partly due to the reduction in the number of cells produced by a large l dimension. It should be noted that, for designs where $L / V$ ratio involve complicated functions of l/h, compactability is improved by making $h$ small as well as by making l large.
C. Rigidization Considerations. It has been stated previously that the deployable structure can be converted into a structure by hardening or freezing joints until the mobility is less than or equal to zero. The question is where to place the locking device, i.e., which joints should be hardened or frozen.

In a system where energy is not dissipated, the products of deflection and forces are constant (equal to the energy)

$$
\begin{align*}
& \text { energy in }=\text { energy out } \\
& \mathrm{F}_{\mathrm{a}} \cdot \delta_{\mathrm{a}}=\mathrm{F}_{\mathrm{b}} \cdot \delta_{\mathrm{b}}  \tag{4}\\
&=\mathrm{T}_{\mathrm{c}} \cdot \theta_{\mathrm{c}}=\mathrm{T}_{\mathrm{d}} \cdot \theta_{\mathrm{d}}
\end{align*}
$$

where $a, b, c$, and $d$ are arbitrary points on the deployable structure, $F$ is the force in the direction of the deflection $\delta$, and $T$ is the torque in the direction of the angular deflection $\theta$. The product of velocity and force is also constant (equal to the power)

$$
\begin{align*}
& \text { power in }=\text { power out } \\
& \mathrm{F}_{\mathrm{a}} \bullet \mathrm{~V}_{\mathrm{a}}=\mathrm{F}_{\mathrm{b}} \bullet \mathrm{~V}_{\mathrm{b}}=\mathrm{T}_{\mathrm{c}} \bullet \omega_{\mathrm{c}}=\mathrm{T}_{\mathrm{d}} \bullet \omega_{\mathrm{d}} \tag{5}
\end{align*}
$$

where $V$ is the linear velocity and $\omega$ is the angular velocity.

These new concepts, as well as other previously proposed designs, were analyzed and evaluated.

Althouqh no clear-cut, "best design" emerged from this effort, several insights into deployable structure design for the SASP application did become apparent. These insights are summarized by the following conclusions and recommendations:
(1) It is important to keep the deployable structure design as simple as possible in order to achieve required reliability goals. Poor deployability associated with Case 1 deployment, randomness and other "uncertainties" associated with Case 2 deployment, and extra weight and complexity associated with Case 4 deployment make Case 3 deployment appear most appropriate for the SASP application. Designs which minimize the number of deployment drives and are, at the same time simple, should be preferred.
(2) Both deployability and compactability are enhanced by large $1 / \mathrm{h}$ ratios. In addition, the number of cells which must be deployed to achieve a desired deployed length is reduced by large $1 / \mathrm{h}$. It is therefore recommended that the largest $1 / \mathrm{h}$ ratio consistent with structural and spacecraft requirements be used.
(3) Because of the large number and wide distribution of joints in a deployable structure, it is likely that the dynamic behavior of the structure can be tailored through selective hardening of various joints and use of joints designed to have specific damping properties. It is recommended that the feasibility of such an approach be further investigated and verified.
(4) Although much of the work in this study was directed toward the SASP application, many of the results are applicable to a wide variety of deployable structure needs. The systematic design procedure developed led to a number of promising deployable structure concepts which were rejected, in this study, because they did not fit the SASP application. Since man first ventured into space, deployable structure technology has been an integral part of the spacecraft "packaging" problem. It appears that there will be many future needs for deployable structures in space. In view of this, and to avoid future duplication of effort, it is recommended that the material developed in this study, as well as that developed in other studies concerned with deployable structures, be consolidated into a Deployable Structures Handbook. This handbook should be designed in such a way that it can be used to help identify the best and most appropriate deployable structure approach for a particular application.

```
NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM
    MARSHALL SPACE FLIGHT CENTER
        THE UNIVERSITY OF ALABAMA
    PAYLOAD MAINTENANCE COST MODEL
    FOR THE SPACE TELESCOPE
```


## Prepared By:

Academic Rank:
University and Department:

William L. White, Ph. D.
Associate Professor
Jackson State University Department of Mathematics

NASA/MSFC:
Laboratory:
Division:
Branch:
MSFC Counterpart:
Date:
Contract No.:

Systems Analysis and Integration
Systems Engineering
Flight Systems Analysis
Frank Pizzano
August 1, 1980
NGT-01-002-099
(University of Alabama)

# PAYLOAD MAINTENANCE COST MODEL 

FOR THE SPACE TELESCOPE

## BY

WILLIAM L. WHITE

# ASSOCIATE PROFESSOR OF MATHEMATICS 

JACKSON STATE UNIVERSITY
JACKSON, MISSISSIPPI 39217

## ABSTRACT

The research project focuses on the development of an optimum maintenance cost model for the Space Telescope for a fifteen (15) year mission cycle. A review of various documents and subsequent updates of failure rates and configurations were made. The reliability of the Space Telescope for one year, two and one-half years, and five years were determined using the failure rates and configurations. The failure rates and configurations were also used in the maintenance simulation computer model which simulates the failure patterns for the fifteen (15) year mission life of the Space Telescope. Cost algorithms associated with the maintenance options as indicated by the failure patterns were developed and integrated into the model.

## INTRODUCTION

This report presents the summer's activities directed toward determining the most feasible maintenance policy for the Space Telescope for a fifteen (15) year mission cycle. These include cost algorithms associated with part replacements and/or complete refurbishment. Participating in the summer's activities with me were Frank Pizzano, my counterpart, Rodney Stewart, Allen Forney, Molly Anderson, and Carol Cleveland.

Two basic tasks were adopted which were as follows:
(1) To update a maintenance simulation computer model which simulates the failure patterns for the fifteen (15) year mission cycle of the Space Telescope.
(2) To develop cost algorithms for the maintenance options according to the failure patterns to ascertain the optimum maintenance policy.

In order to accomplish the tasks the activities performed were as follows:
(1.1) A detailed study of the maintenance simulation computer model.
(1.2) Updates of failure rates were made for various parts whenever necessary.
(1.3) Model configuration updates were made where revisions in the specifications called for a reconfiguration.
(1.4) The reliability of the Space Telescope was determined for one year, two and one-half years, and five years.
(1.5) Updates were made in the computer program where model configuration changed.
(2.1) Cost Data were collected and organized for parts whose prices were known or available information would allow a price estimate.
(2.2) Cost Algorithms were written involving a base-line algorithm, i. e. calculates cost if time schedule is standard and extended algorithm, i. e. takes base-line figure and adds a certain percentage according to the relationship between the time left in the new schedule and the time left in the standard schedule.
(b) Update the maintenance simulation computer model with respect to component/systems configurations, failure rates, and orbit replaceable unit cost estimates.
(c) Run the program and obtain the following cost comparisons against the specification plan of having three (3) in-orbit maintenance actions during the fifteen (15) year mission operating life of the Space Telescope:
(1) Assume in-orbit corrective maintenance actions only for the fifteen (15) year mission operating life with frequencies as dictated by the maintenance simulation computer model.
(2) Assume in-orbit corrective (Plus Preventive) maintenance actions only for the fifteen (15) year mission. (The preventive actions would involve components subject to degradation with age such as batteries.)
(3) Assume in-orbit corrective actions plus ground return actions with frequencies as dictated by the maintenance simulation computer model.

Reliability versus cost will be compared for each case using the Shuttle cost constants plus orbit replaceable unit cost estimates to obtain an indication of a preferable maintenance policy.

## RELIABILITY OF THE SPACE TELESCOPE

The Space Telescope consist of three major subsystems: An Optical Telescope Assembly, A Support System Module, and The Scientific Instruments. The support system module encloses the optical telescope assembly and scientific instruments and also provide all interfaces with the shuttle orbiter.

The maintenance of the space telescope will involve either in-orbit maintenance or ground return maintenance or both. The items which will be maintained in orbit will be called orbit replaceable units. Each item is given a part number, a string number, and an orbit replaceable unit number for identification. The orbit replaceable units are configured on trays (i.e., A component in which all parts have the same orbit replaceable unit number), in such a way that they are easily accessible to the maintenance crew. The items which will be replaced on ground return will be called ground replaceable units.

The components of the optical telescope assembly were identified and failure rates associated with each component were obtained from references $[1,4]$ to determine the reliability of each item for one year, two and one-half years, and five years.

However, in the case of the DF-224 computer which consist of three memory units required and six provided with three active and three inactive, the reliability according to reference [3] is given By:

$$
\begin{align*}
R(t)= & e^{-3 \lambda t}+\frac{3 \lambda}{\lambda^{\prime}} e^{-3 \lambda t}\left(1-e^{-\lambda^{\prime} t}\right)+\frac{3 \lambda\left(3 \lambda+\lambda^{\prime}\right)}{2 \lambda^{\prime^{2}}} e^{-3 \lambda t}\left(1-e^{-\lambda^{\prime} t}\right)^{2} \\
& +\frac{3 \lambda}{6 \lambda^{\prime 3}}\left(3 \lambda+\lambda^{\prime}\right)\left(3 \lambda+2 \lambda^{\prime}\right) e^{-3 \lambda t}\left(1-e^{-\lambda^{\prime} t}\right)^{3} \tag{5}
\end{align*}
$$

Where $\boldsymbol{\lambda}$ is the failure rate and $\lambda^{\prime}$ is the standby failure rate, i.e., $\lambda^{\prime}=$ $0.1 \lambda$. The central processor unit and the in-put-output unit are the other components of the DF-224. In these cases there are three provided with one required so the reliability formula is given by:

$$
\begin{equation*}
R(t)=e^{-\lambda t}+\frac{\lambda}{\lambda^{\prime}} e^{-\lambda t}\left(1-e^{\lambda^{\prime} t}\right)+\frac{\lambda\left(\lambda+\lambda^{\prime}\right)}{2 \lambda^{\prime 2}} e^{-\lambda t}\left(1-e^{-\lambda^{\prime} t}\right)^{2} \tag{b}
\end{equation*}
$$

The program for the computation of reliability of each component is given below.

```
Rea1 L1, L2
Double Precision A, B, Al, Bl, C, D, PM, PCP,T
Read (1,5) L1, L2, T
A = DEXP (-3.*L1*T)
B =1. - DEXP (-L2*T)
C = 3. *Ll + L2
Al = DEXP (-Ll*T)
B1 = 1. - DEXP (-L2*T)
D = 3.*L1 + 2.*L2
PM = A+3.*L1/L2*A*B+3.*L1*(3.*L1+L2)*A*B**2/2./L2**2
PM = PM+3.*L1/6.L2**3*C*D*A*B**3
PCP = A1+L1/L2*A1*B1
PCP = PCP+L1*(L1+L,2)/2.L2**2*A1*B1**2
Write (1,6) L1, L2, T, PM, PCP
5 Format (3 E15.6)
6 Format (1X, 5E15.6)
Stop
End.
```

The reliability of the support systems module for one year, two and one -half years, and five years is given in Table [2].

A breakdown of the base-line algorithm includes cost associated with maintenance analysis and planning, management of the or bit replaceable units, ground maintenance facility, and orbiter launch. The total expenditures being the sum of the cost for the maintenance analysis and planning, orbit replaceable unit management, ground maintenance facility and orbiter launch. The extended algorithm takes the base-line figure and adds a certain percentage according to the relationship between the time left in the new schedule and the time left in the standard schedule. The cost per part is used to determine the total cost for parts being replaced at a downstate.

## MAINTENANCE SIMULATION MODEL

The maintenance simulation computer model is programmed to test each item once every 72 hours for failure. The failure rate of each item is used to determine the probability that the item has failed. With a random number generator the item is either failed or not failed by comparing a random number with the probability (computed) of failure. Each item is given a part number, a string number, and an orbit replaceable unit number for identification, as items are failed a $\log$ is kept of the failures by specifically identifying the failed item. This is done so the computer will know when the number of failed items result in a downstate. The Shuttle is then called for maintenance action. The time of each failure is also logged. The tests can be continued every 72 hours for a 15 year mission cycle. Each 15 year mission cycle test is called a pass. The model has the capabilities of making a number of passes sufficient to predict the mean time to failure with reasonable confidence. Plots of the failures versus time can also be made. This information together with the cost algorithm which is integrated in the model will give the total cost relative to maintenance of the Space Telescope according to the standard schedule as well as cost relative to the simulated schedule.

## CONCLUSION

A maintenance policy for the Space Telescope is a must if the Space Telescope is to serve the needs of the scientific community effectively. The policy must optimize both cost for maintenance and availability for use by the scientific community. Therefore, the policy must be based on all available resources and information. Recognizing that many parts of the Telescope are of new design, very little is known other than simulated testing concerning the reliability thus the task of defining a policy becomes even more difficult. To this end, generic failure rates are used to determine the reliability of each item. Various configurations of the Space Telescope are continuously being updated in an effort to simulate a model which will more closely parallel the actual performance of the Space Telescope. The reliability of the Space Telescope with updated configurations was determined for one year, two and one-half years, and five years. The hope is that reliability as indicated by the model will provide a basis for a maintenance policy according to a particular time schedule. Due to the lack of acquired data relative to the actual item performance of the Space Telescope the model simulation approach seems appropriate. This approach has been used with much success in the past.

## REFERENCES

1. Contractor's Reliability Allocations
2. IBM Nas 5-25630, "SI C \& DH Reliability Prediction Report" November 20, 1980
3. Memorandum EF01, "Reliability Analysis and Evaluation of DF-224/NSSC-II Computers for Space Telescope," October 1978.
4. MIL-HDBK-217 B/C, "Reliability Prediction of Electronic Equipment."

## TABLE 2

## SUMMARY OF SUPPORT SYSTEMS MODULE ORBIT

REPLACEABLE UNIT (ORU) RELIABILITY STUDY

## Reliability (Probability of Success)

|  | R (lyear) | $\mathrm{R}(21 / 2$ years) | R (5 years) |
| :--- | :--- | :--- | :--- |
| Support Systems <br> Module with present <br> ORU capability | 0.828343 | 0.214951 | $0.0_{5} 3446$ |
| Support Systems <br> Module with degraded <br> mission capability | $0.949195 *$ | $0.540912 *$ | $0.03^{3} 3066 *$ |
| Reliability enhance- <br> ment with limited <br> LMSC suggested <br> replacement ORU's <br> (20 ORU's Total) | 0.909337 | 0.909793 | 0.378308 |
| Reliability enhance- <br> ment with MSFC <br> suggested replacement <br> ORU's <br> (12 ORU's Total) | 0.890368 | $0.966236 *$ | $0.883826 *$ |

NOTE: Present ST Design includes approximately 37 ORU types with 69 total ORU's.

TABLE 4

## SUMMARY OF TOTAL SPACE TELESCOPE SYSTEMS RELIABILITY STUDY

## Reliability (Probability of Success)

R (1 year)
R(2 1/2 years)
R (5 years)
(1) Probability of the Space Telescope functioning successfully in accordance with SPEC and including at least 3 of 6 science functions.
(2) Item (1) above except including mission 0.794267
0.290403
0.0454 degradation allowances.

End of Document


[^0]:    * Number within the parentheses refers to the references at the back of the report.
    ** Number of samples taken.

[^1]:    *Numbers in brackets refer to references at end of paper.

