# FINAL REPORT 

THE ASEE-UNIVERSITY OF MARYLANDCATHOLIC UNIVERSITYNASA SUMMER FACULTY FELLOWSHIP PROGRAM JUNE 12 TO AUGUST 18, 1967

GODDARD SPACE FLIGHT CENTER GREENBELT, MARYLAND


## FINAL REPORT

THE ASEE-UNIVERSITY OF MARYLANDCATHOLIC UNIVERSITY-NASA SUMMER FACULTY FELLOWSHIP PROGRAM

June 12 to August 18, 1967

April 1968

Goddard Space Flight Center
Greenbelt, Maryland

## PRECEDING PAGE BLANK NOT FILMED.

## INTRODUCTION

The GSFC Faculty Fellowship Program is a cooperative effort between Goddard, the American Society of Engineering Education (ASEE), the University of Maryland, and the Catholic University of America.

Under this program university faculty members come to Goddard for a period of ten weeks during which they engage actively in research and at the same time participate in seminars related to their research.

The objectives of this program are many; some of them are listed below:

1. Stimulation of schools to become interested in the research problems confronting Goddard.
2. Creation of interest on the part of the university faculty to continue their research after completing the formal program.
3. Stimulation of our people professionally through associations with the university faculty and through participation in the program's seminars.
4. Establishment of closer ties with the Universities.

This report marks the completion of the third year of this program at Goddard, and represents the progress in one of the objectives of the Summer Faculty Program.


## SUMMER FELLOWSHIP PARTICIPANTS

## ADMINISTRATION

| Dr. B. Fang | Co-director | Professor, S. Sc. Dpt., CU |
| :--- | :--- | :--- |
| Dr. A. Marcovitz | Co-director | Professor, EE Dept., <br> Univ. of Md. |
| Dr. J. Morakis | Co-director | Sr. Eng., Advanced <br> Development Div., GSFC |
| Mr. J. Reese | Co-director | Head, Employee Development <br> Branch, GSFC |
| Mr. T. Morris | Conference Coordinator | Conference and Institutes <br> Division, Univ. of Md. |

# SECOND-YEAR FELLOWS PARTICIPATING IN THE 1967 ASEE-NASA 

 SUMMER INSTITUTEFELLOW

Dr. John R. Bolte
Mr. David A. Dalgleish

Mr. Billie Joe Graham

Dr. Frank P. Kuhl, Jr.

Mr. Richard L. Martin

Dr. George E. McCluskey, Jr.

Dr. Harlan J. Perlis

Mr. Frederick A. Rojak

Dr. Theodore G. Toridis

Dr. John P. Uldrick

ACADEMIC RANK AND AFFILIATION

Associate Professor<br>Department of Physics<br>San Diego State College

Chairman, Electronics-Electrical
Engineering Technology Department Spring Garden Institute
Associate Professor
Science Department
U. S. Naval Academy
Assistant Professor
Electrical Engineering
Union College
Instructor
Electrical Engineering DepartmentUniversity of Maryland
Assistant Professor, Department
of Mathematics-Astronomy
Lehigh University
Associate Professor
Department of Electrical EngineeringRutgers - The State University
Associate Professor
Department of Electrical Engineeringand Physics
Lowell Technological Institute
Assistant Professor
Engineering Department
The George Washington University
Associate Professor
Engineering Mechanics DepartmentClemson University

FELLOW
Mr. William A. Barrett

Dr. David Cheng

Dr. Bernard Chern

Dr. Richard A. Elco

Dr. Carl L. Gruber

Mr. David M. Koenig

Mr. Nicholas Kyriakopoulos

Mr. Paul A. Lux

Mr. Jack F. Morris

Mr. David F. Nyman

ACADEMIC RANK AND AFFILIATION

Instructor<br>Department of Electrical Engineering<br>Union College

Assistant Professor
Electrical Engineering Department
University of Missouri
Assistant Professor
Physics Department
University of Florida
Assistant Professor
Electrical Engineering
Carnegie Institute of Technology
Assistant Professor
Electrical Engineering
South Dakota School of Mines and Technology

Instructor
Chemical Engineering Department Catholic University

Assistant Professor
Engineering Department
George Washington University
Assistant Professor
Electrical Engineering Department
Sacramento State College
Instructor
Electrical Engineering Department University of Missouri

Assistant Professor
Systems Analysis
Miami University

| Dr. John B. Peatman | Assistant Professor <br> Electrical Engineering <br> Georgia Institute of Technology |
| :--- | :--- |
| Dr. Harvey Rexroad | Professor <br> Physics Department <br> University of West Virginia |
| Dr. Linus E. Schrage | Assistant Professor <br> Industrial Engineering Department <br> Stanford University |
| Dr. Philip M. Seal | Professor and Head <br> Engineering Department <br> Norwich University |

## ACADEMIC RANK AND AFFILIATION

Assistant Professor Electrical Engineering Georgia Institute of Technology

Physics Department<br>University of West Virginia

Assistant Professor Industrial Engineering Department Stanford University

Professor and Head
Engineering Department
Norwich University
Page
A SURVEY ON SOME MATHEMATICAL MODELS OF THE VERY LOW FREQUENCY WAVE PROPAGATION ..... 1.
David H. S. Cheng
A THEORETICAL STUDY OF ELECTRIC FIELDS IN THE LUNAR ATMOSPHERE ..... $45 /$
Richard A. Elco
A STUDY OF THE FREQUENCY AND STABILITY CHARACTERISTICS OF A CO 2 LASER ..... $63 \sqrt{ }$
B. H. Graham
RELATIVE PHASE DIFFERENCE ANALYSIS PROGRAM. ..... 77 N
D. M. Koenig
AN INVESTIGATION OF METHODS FOR THE EXPERIMENTAL DETERMINATION OF PCM SIGNAL TO NOISE RATIO ..... 89
Jack F. Morris
COMPUTER SYSTEMS FOR REAL-TIME APPLICATIONS. ..... 105
David F. Nyman
ON THE USE OF ADAPTIVE TECHNIQUES TO IMPROVE ANTENNA SERVO ACCURACY ..... 123
H. J. Perlis
RETRODIRECTIVE ANTENNAS AND THE MULTIPATH PROBLEMS ASSOCIATED WITH THEIR USE ON THE PROPOSED ORBITING DATA RELAY NETWORK ..... $145 \sqrt{ }$
P. M. Seal
TRANSFER FUNCTIONS OF THE ROSMAN I ANTENNA STRUCTURE ..... 161T. G. Toridis and Yu Chen
SPATLAL LIGHT MODULATION EMPLOYING A FERRO-MAGNETIC FLUID. ..... 185
J. P. Uldrick and P. A. Lux
ABSTRACTS ..... 215

# A SURVEY ON SOME MATHEMATICAL MODELS OF THE VERY LOW FREQUENCY WAVE PROPAGATION 

by
David H. S. Cheng ${ }^{\dagger}$

## INTRODUCTION

## Historical Background

The development of the theory of the very low frequency wave propagation may be divided roughly into two periods. The first period was ushered in with the introduction of long distance communications by means of the wireless at the turn of the century. The underlying principle in almost all the earlier analyses was the diffraction concept advanced by Rayleigh in 1904 [1].

In such a diffraction model, the existence of the ionosphere was ignored, and the earth was considered as a homogeneous conducting sphere. Among the earlier investigators were Poincare [2], MacDonald [3], Nicholson [4], Sommerfeld [5], Love [6], March [7], Rybczynski [8], and others. Most of the theoretical investigations made use of the theory of the zonal harmonics, either by a direct summation or by forming an integral from the series, to explain the facts concerning the propagation of radio waves around the earth's curvature. However, the series then formulated was difficult to sum because of its slow convergence for the radio problem.

Poincare [2] and Nicholson [4] attempted to replace the series by an integral and then obtained an approximate value for the integral by means of the calculus of residues. Their scheme, though substantially sound, was exceedingly elaborate, and their approximations became invalid in the vicinity of the antipodes of the transmitter.

The method used by MacDonald [3] was to approximate the original series by a modified series which in turn was replaced by an integral. This procedure was questioned from a purely mathematical point of view, though worked in most practical circumstances.

Love [6] undertook the laborious task of evaluating the series for the radio problem by numerically adding the terms of the series taken in groups. His memoir contained a complete bibliography of the problem up to 1915.

[^0]The results of the aforementioned authors did not explain the propagation phenomenon round the earth's curvature; nevertheless, a correct conclusion was reached that the diffracted field decayed exponentially from the transmitter.

Sommerfeld [5] investigated the problem using a planar model and concluded that the field decayed approximately as $1 / \sqrt{d}$ or at most $1 / \mathrm{d}$ where d is the distance of propagation from the transmitter, which checked fairly well with the observed facts. Upon his suggestion, March [7] reformulated the problem using a spherical model. He cast his solution in an integral in order to obtain an answer more consistent with observed facts. Indeed, he found that the field decayed approximately as $1 / \sqrt{\theta} \sin \theta \sim 1 / \mathrm{d}$ where $\theta$ is the angular distance between the transmitter and the receiver. March concluded that the waves were propagated without exponential attenuation contrary to the results obtained by other investigators. His method was severely questioned. However, von Rybczynski [8] extended March's method and pointed out that the exponential factor was neglected by March.

In reference to these investigations, Smith [9] remarked that "many years of work of the most famous mathematicians were inadequate to derive this task, which at the first glance seems very simple, from that state with respect to which Nicholson said that out of all mathematical problems it is the only one that has caused such a great divergence of opinions." It was against this background that Watson [10], in 1918, succeeded in his first of two classical papers on VLF wave propagation in making the final conclusion that we could not explain the fact of the long distance propagation of long waves around the terrestrial sphere using only the diffraction model.

The second period was initiated in 1919 by the work of Watson [10]. In the second of his two papers on the radio problem, Watson postulated a concentric plasma of high conductivity to account for the ionosphere which is mainly responsible for the VLF propagation over great distances. Watson succeeded in reformulating the integral used by others by means of now the well known Watson transformation. Instead of working with the slowly convergent series he was able to replace it by a rapidly converging series. He introduced the earth ionosphere waveguide concept. The radio waves excited by a vertical Hertzian dipole propagated in the atmosphere between the earth and the ionosphere in a way similar to the microwave in modern wave guides.

Among the earlier investigators in the second period in the development of the theory of the propagation of LF and VLF waves, aside from Watson, are Rydbeck [11] , Van der Pol and Bremmer [12], Eckersley [13], Bremmer [14], Kendrick [15], Weyrick [16], Brekhovskikh [17], and Al'pert [18 ], Fock [19],
and others. The results of their analyses indicated needs for further refinement of the model such as the introduction of terrestrial sphere stratification, anistropy, continuous ionospheric stratification, polarization coupling and anisotropy at the ionosphere and the effect of ions, both heavy and light.

## 2. Recent Investigations

Among the most recent and active investigators are Budden [20], Wait [ 21], Johler [22], and Kranushkin [23].

Wait is a proponent of the mode theory (or the residue theory). To account for the behaviors of the fields in the presence of an ionosphere, he considered successively the ionosphere as sharply bounded, gradual, and then stratified; isotropic and the anisotropic; uniform and non-uniform. The stumbling block of the mode theory was the transcendental model equation in Hankel functions of complex order for determining the modes of the propagating waves. In its solutions, various approximations were employed. In view of this difficulty, Johler attempted the solution of the VLF propagation problem using only the original zonal harmonic series. With the improved summation method, the series could readily be summed on a large computer in spite of a large number of terms needed. Budden, on the other hand, advocated a direct and numerical solution of the full-wave equation taking into account of the coupling given arisen by the anisotropic and non-uniform nature of the ionosphere.

Kranushkin resolved the problem using the theory of non-self-adjoint differential operator eigenfunction expansion. He called his method the normal wave solution. He considered the ionosphere as stratified media with the boundaries being weakly angle dependent.

We will consider in the next section the four mathematical models currently being used; namely,
i. the full-wave solution,
ii. the mode theory,
iii. the zonal harmonic series,
iv. the normal wave method.

## 3. Theoretical Background

Most of the theoretical considerations of problems involving the propagation of radio waves are based on the assumption that the electromagnetic field satisfies Maxwell's equation in the form

$$
\begin{aligned}
& \nabla \times \overrightarrow{\mathrm{E}}=-\mathrm{i} \omega \mu_{0} \overrightarrow{\mathrm{H}} \\
& \nabla \times \overrightarrow{\mathrm{H}}=\mathrm{i} \omega \epsilon \overrightarrow{\mathrm{E}}
\end{aligned}
$$

In the above equations, a time harmonic factor $\exp (i \omega t)$ is understood; $\overrightarrow{\mathrm{E}}$ and $\overrightarrow{\mathrm{H}}$ are the complex representations of the electric and magnetic field vectors; $\mu_{0}$ and $\epsilon$ are the permeability and permittivity of the material medium. For most non-ferromagnetic media, the permeability equals to that of the free space. For the ionosphere, the permittivity is a tensor. If $\epsilon$ is put equal to $\epsilon_{0} \kappa$, where $\epsilon_{0}$ is the permittivity of the free space, then $\kappa$, the dielectric tensor can be written as

$$
\begin{gathered}
\kappa=\left(\begin{array}{lll}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{array}\right)-\frac{x}{U\left(U^{2}-Y^{2}\right)} \\
\times\left(\begin{array}{cc}
U^{2}-\ell^{2} Y^{2} & i n Y U-\ln Y^{2} \\
-i n Y U-\ln Y^{2} & -i m Y U-\ln Y^{2} \\
i m Y U-\ln Y^{2}-m^{2} Y^{2} & i \ell Y U-m n Y^{2} \\
-i \ell Y U-m n Y^{2} & U^{2}-n^{2} Y^{2}
\end{array}\right)
\end{gathered}
$$

The standard URSI notation is used here. X is the square of the ratio of plasma frequency to wave frequency; $U=1-i Z$, where $Z$ is the ratio of the electron collision frequency to the wave frequency; $Y(\ell, m, n)$ are the components of the vector $\vec{Y}$ referred to some rectangular cartesian coordinates. $Y$ is the magnitude of the ratio of the gyromagnetic frequency of electrons to the wave frequency in the direction of the earth's magnetic field.

The dielectric tensor is asymmetric. It implies that the ionosphere is an anisotropic medium. The changes in the electron density and collision frequency and their variations with height contributing to the inhomogeneity to the property of the ionosphere of the VLF wave propagation.

In all the problems considered in the sequel, the primary source is a Hertzian vertical dipole.

## II. THE DIFFRACTION MODEL

We will consider first the diffraction problem together with Van der Pol and Bremmer.

An electric dipole is situated at $Q$, a distance $b$ away from the center of the sphere of radius a. The observation point is at $P$, a distance $r$ away from the center of the sphere. The points $Q$ and $P$ are separated by a distance R.

The electromagnetic fields, $\overrightarrow{\mathrm{E}}$ and $\vec{H}$ satisfy Maxwell's equations. A time factor $\exp (-i \omega t)$ is used. The analytic problem consists of finding an electric Hertzian vector potential function $\pi \vec{r}$, where $\vec{r}$ is a position vector, such that it satisfies the following differential equations and boundary conditions:
(1) $\begin{aligned} & \left(\nabla^{2}+\mathrm{k}_{1}^{2}\right) \pi=0 \quad(\mathrm{r}>\mathrm{a}) \text {, }\end{aligned}$

$$
\left(\nabla^{2}+k_{2}^{2}\right) \pi=0 \quad(r<a)
$$

(2) $\frac{\partial}{\partial r}(\mathrm{r} \pi)$ and $\mathrm{k}^{2} \pi$ are continuous at $\mathrm{r}=\mathrm{a}$,
(3) A singularity at $Q$ of the form $\frac{e^{i k_{1} R}}{i k_{1} R}$ which is further called the primary
field.

The electric and magnetic fields are given by

$$
\begin{aligned}
& \mathrm{E}_{\mathrm{r}}=\mathrm{k}^{2}+\frac{\partial^{2}}{\partial \mathrm{r}^{2}}(\mathrm{r} \pi), \\
& \mathrm{E}_{\theta}=\frac{1}{\mathrm{r}} \frac{\partial^{2}}{\partial \mathrm{r} \partial \theta}(\mathrm{r} \pi), \\
& \mathrm{H}_{\phi}=-\frac{\mathrm{ic}}{\omega} \mathrm{k}^{2} \frac{\partial \pi}{\partial \theta}
\end{aligned}
$$

and

$$
\mathrm{k}_{1,2}^{2}=\frac{\epsilon_{1,2} \omega^{2}+\mathrm{i} \tau_{1,2} \omega}{\mathrm{c}^{2}} .
$$

## 1. The Zonal Harmonic Series

The solution of the total Hertzian potential function in the form of the zonal harmonic series for $r>a$ is,

$$
\begin{align*}
\pi_{t o t} & =\frac{e^{i k_{1} R}}{i k_{1} R} \\
& +\sum_{n=0}^{\infty}(2 n+1) R_{n} \frac{\psi_{n}\left(k_{1} a\right)}{\zeta_{n}^{(1)}\left(k_{1} a\right)} \zeta_{n}^{(1)}\left(k_{1} b\right) \zeta_{n}^{(1)}\left(k_{1} r\right) P_{n}(\cos \theta) \tag{1.1a}
\end{align*}
$$

and for $r<a$ is

$$
\begin{equation*}
\pi_{t o t}=\frac{k_{1}^{2}}{k_{2}^{2}} \sum_{n=0}^{\infty}(2 n+1)\left(R_{n}+1\right) \frac{\psi_{n}\left(k_{1} a\right)}{\psi_{n}\left(k_{2} a\right)} \zeta_{n}^{(1)}\left(k_{1} b\right) \psi_{n}\left(k_{2} r\right) P_{n}(\cos \theta) \tag{1.1b}
\end{equation*}
$$

where

$$
\begin{aligned}
\zeta_{n}^{(1)}(x) & =\sqrt{\frac{\pi}{2 x}} H_{n+1 / 2}^{(1)}(x)=x^{n}\left(-\frac{1}{x} \frac{d}{d x}\right)^{n}\left(\frac{e^{i x}}{i x}\right), \\
\zeta_{n}^{(2)}(x) & =\sqrt{\frac{\pi}{2 x}} H_{n+1 / 2}^{(2)}(x)=x^{n}\left(-\frac{1}{x} \frac{d}{d x}\right)^{n}\left(\frac{i e^{-i x}}{x}\right), \\
\psi_{n}(x) & =\frac{1}{2}\left\{\zeta_{n}^{(1)}(x)+\zeta_{n}^{(2)}(x)\right\} \\
& =\sqrt{\frac{\pi}{2 x}} J_{n+1 / 2}(x)=x^{n}\left(-\frac{1}{x} \frac{d}{d x}\right)^{n}\left(\frac{\sin x}{x}\right) .
\end{aligned}
$$

( $\mathrm{H}=$ Hankel function, $\mathrm{J}=$ Bessel function) and where

$$
\begin{equation*}
R_{n}=\frac{-\left[\frac{1}{x} \frac{d}{d x} \log \left\{x 4_{n}(x)\right\}\right]_{x=k_{1} a}+\left[\frac{1}{x} \frac{d}{d x} \log \left\{x 4_{n}(x)\right\}\right]_{x=k_{2} a}}{\left[\frac{1}{x} \frac{d}{d x} \log \left\{x \zeta_{n}^{(1)}(x)\right\}\right]_{x=k_{1} a}-\left[\frac{1}{x} \frac{d}{d x} \log \left\{x 4_{n}(x)\right\}\right]_{x=k_{2} a}} \tag{1.2}
\end{equation*}
$$

It is well known that

$$
\begin{align*}
& \frac{e^{i k_{1} R}}{i^{k_{1} R}}=\sum_{n=0}^{\infty}(2 n+1) \zeta_{n}^{(1)}\left(k_{1} r\right) \psi_{n}\left(k_{1} b\right) P_{n}(\cos \theta) \quad(r>b)  \tag{1.3a}\\
& \frac{e^{i k_{1} R}}{i^{k_{1} R}}=\sum_{n=0}^{\infty}(2 n+1) \zeta_{n}^{(1)}\left(k_{1} b\right) \psi_{n}\left(k_{1} r\right) P_{n}(\cos \theta) \quad(r<b) \tag{1.3b}
\end{align*}
$$

By substituting (1.3b) into (1.1a), (1.1a) can be expressed explicitly as a zonal harmonic series. Furthermore, on the surface of the earth where $r=s$, it can be reduced to the following

$$
\begin{equation*}
\pi_{t o t}=\frac{i}{(k a)^{3}} \sum_{n=0}^{\infty}(2 n+1) \frac{\zeta_{n}^{(1)}\left(k_{1} b\right)}{\zeta_{n}^{(1)}\left(k_{1} a\right)} \frac{1}{N_{n}} P_{b}(\cos \theta) \tag{1.4}
\end{equation*}
$$

where

$$
\begin{align*}
\mathrm{N}_{\mathrm{n}}(\mathrm{x}, \mathrm{y}) & =\left[\frac{1}{\mathrm{x}} \frac{\mathrm{~d}}{\mathrm{dx}} \log \left\{\mathrm{x} \zeta_{\mathrm{n}}^{(1)}(\mathrm{x})\right\}\right]_{\mathrm{x}=\mathrm{k}_{1} \mathrm{a}} \\
& -\left[\frac{1}{\mathrm{y}} \frac{\mathrm{~d}}{\mathrm{dy}} \log \left\{\mathrm{y} \psi_{\mathrm{n}}(\mathrm{y})\right\}_{\mathrm{y}}\right]=\mathrm{k}_{2} \mathrm{a} \tag{1.5}
\end{align*}
$$

In the reduction to Eq. (1.4), use is made of the Wronskian:

$$
\begin{equation*}
\psi_{n}(x) \zeta_{n}^{(1)}(x)-\zeta_{n}^{(1)}(x) \psi_{n}^{\prime}(x)=\frac{i}{x^{2}} \tag{1.6}
\end{equation*}
$$

## 2. The Watson Transformation and the Residue Series

Following Watson, we transform (1.4) into an integral over $n$ leading to

$$
\begin{equation*}
\pi_{\text {tot }}=\frac{1}{\left(k_{1} a\right)^{3}} \int_{c_{1}} \frac{n d n}{\cos (n \pi)} \frac{\zeta_{n-1 / 2}^{(1)}\left(k_{1} b\right)}{\zeta_{n-1 / 2}^{(1)}\left(k_{1} a\right)} \frac{1}{N_{n-1 / 2}} P_{n-1 / 2}[\cos (\pi-\theta)] \tag{1.7a}
\end{equation*}
$$

where $C_{1}$ is the contour which enclosed the positive real axis in a counterclockwise direction. By a judicial manipulation of $\mathrm{C}_{1}$, Eq. (1.7a) is equivalent to the following:

$$
\begin{align*}
\pi_{t o t} & =\frac{1}{\left.k_{1} a\right)^{3}} \int_{L} \frac{n d n}{\cos (n \pi)} \frac{\zeta_{n-1 / 2}^{(1)}\left(k_{1} b\right)}{\zeta_{n-1 / 2}^{(1)}\left(k_{1} a\right)} \frac{1}{N_{n-1 / 2}} P_{n-1 / 2}[\cos (\pi-\theta)] \\
& +\int_{-i \infty}^{i \infty} f(n) d n \tag{1.7b}
\end{align*}
$$

where $f(n)$ is the integrand of (1.7a). The line integral in (1.7b) is usually considered negligible. L is the contour encloses all the poles in the integrand of (1.7a) other than the ones due to $\cos (n \pi)$.

The first term in (1.7b) enables us to obtain a physical interpretation. Let us expand $1 / \cos (\mathrm{n} \pi)$ as

$$
\begin{equation*}
\frac{1}{\cos n \pi}=2 e^{i \pi n} \quad \sum_{m=0}^{\infty} e^{\mathrm{i} \pi m(2 n+1)} \tag{1.8}
\end{equation*}
$$

We can write (1.7b) over the integration path $L$ as

$$
\begin{equation*}
\pi_{\mathrm{tot}} \cong \sum_{\mathrm{m}=0}^{\infty} \pi_{\mathrm{m}} \tag{1.9}
\end{equation*}
$$

where

$$
\begin{align*}
& \pi_{\mathrm{m}}=\frac{2 \mathrm{e}^{\mathrm{i} \pi_{\mathrm{m}}}}{\left(\mathrm{k}_{1} \mathrm{a}\right)^{3}} \int_{\mathrm{L}} \operatorname{ndn} \mathrm{e}^{\mathrm{i} \pi_{\mathrm{n}}(2 \mathrm{~m}+1)} \frac{\zeta_{\mathrm{n}-1 / 2}^{(1)}\left(\mathrm{k}_{1} \mathrm{~b}\right)}{\zeta_{\mathrm{n}-1 / 2}^{(1)}\left(\mathrm{k}_{1} \mathrm{a}\right)} \frac{1}{\mathrm{~N}_{\mathrm{n}-1 / 2}} \\
& \times \mathrm{P}_{\mathrm{n}-1 / 2}[\cos (\pi-\theta)] \tag{1.10}
\end{align*}
$$

The integral in (1.10) can now be represented as the sum of the residues of all the enclosed poles in $L$, and it is written as

$$
\begin{align*}
\pi_{m}=\frac{4 \pi i e^{i \pi_{m}}}{\left(k_{1} a\right)^{3}} & \sum_{s=0}^{\infty} n_{s} e^{i \pi_{n_{s}}(2 m+1)} \frac{\zeta_{n_{s}-1 / 2}^{(1)}\left(k_{1} b\right)}{\zeta_{n_{s}-1 / 2}^{(1)}\left(k_{1} a\right)} \\
& \frac{1}{\left(\frac{\partial N_{n-1 / 2}}{\partial n}\right)_{n=n_{s}}} \quad P_{n_{s}-1 / 2}[\cos (\theta-\pi)] \tag{1.11}
\end{align*}
$$

Upon using a well-known asymptotic formula for the spherical harmonics of complex order with positive imaginary part, (if $\theta$ is not too near 0 or $\pi$ ),

$$
\begin{align*}
& \mathrm{P}_{\mathrm{n}-1 / 2}[\cos (\pi-\theta)]=\frac{1}{\sqrt{2 \pi \mathrm{n} \sin \theta}} \mathrm{e}^{-\mathrm{in}(\pi-\theta)+\mathrm{i} \pi / 4}  \tag{1.12}\\
\pi_{\mathrm{m}} \cong & \frac{2 \sqrt{2 \pi}}{\left(\mathrm{k}_{1} \mathrm{a}\right)^{3}} \frac{\mathrm{e}^{\mathrm{i} \pi(\mathrm{~m}+3 / 4)}}{\sqrt{\sin \theta}} \\
& \times \sum_{\mathrm{s}=0}^{\infty} \frac{\zeta_{\mathrm{n}_{\mathrm{s}}-1 / 2}^{(1)}\left(\mathrm{k}_{1} \mathrm{~b}\right)}{\zeta_{\mathrm{n}_{\mathrm{s}}-1 / 2}^{(1)}\left(\mathrm{k}_{1}\right)} \frac{\mathrm{n}_{\mathrm{s}}^{1 / 2}}{\left(\frac{\partial \mathrm{~N}_{\mathrm{n}-1 / 2}}{\partial \mathrm{n}}\right)_{\mathrm{n}=\mathrm{n}_{\mathrm{s}}}} \mathrm{e}^{\mathrm{in} \mathrm{~s}}(\theta+2 \pi \mathrm{~m}) \tag{1.13}
\end{align*}
$$

The last factor in (1.13) becomes

$$
\begin{equation*}
e^{i n_{s}\left(\theta+2 \pi_{m}\right)}=e^{\left(i \alpha_{s}-\beta_{s}\right)\left(\theta+2 \pi_{m}\right)} \tag{1.14}
\end{equation*}
$$

where

$$
\mathrm{n}_{\mathrm{s}}=a_{\mathrm{s}}+\mathrm{i} \beta_{\mathrm{s}} \quad\left(\beta_{\mathrm{s}}>0\right)
$$

Hence (1.13) represents a wave travelling round the sphere $m$ times before reaching the receiver with a phase velocity and attenuation determined by $\alpha_{s}$ and $\beta_{s}$. Because of the fact that a wave travelling $m$ times round the earth becomes greatly attenuated, only the $m=0$ term contribute significantly to the fields. Thus,

$$
\begin{align*}
\pi_{\text {tot }} & \cong \pi_{0} \\
& =\frac{2 \sqrt{2 \pi} \mathrm{e}^{i 3 \pi / 4}}{\left(k_{1} a\right)^{3} \sqrt{\sin \theta}} \\
& \times \sum_{s=0}^{\infty} \frac{\zeta_{n_{s}-1 / 2}^{(1)}\left(k_{1} b\right)}{\zeta_{n_{s}-1 / 2}^{(1)}\left(k_{1} a\right)} \frac{n_{s}^{1 / 2}}{\left(\frac{\partial N_{n_{s}}-1 / 2}{\partial n}\right)_{n=n_{s}}} e^{i n_{s} \theta} \tag{1.15}
\end{align*}
$$

## 3. The Eigenfunction Method

Van der Pol and Bremmer also consider the possibility of casting the radio wave propagation as an eigenvalue problem. In this case Eq.'s (1.1a) and (1.1b) are determined by the following considerations:
(1) $\left(\nabla^{2}+\mathrm{k}_{1}^{2} \lambda_{\mathrm{j}}^{2}\right) \phi_{\mathrm{j}}=0 \quad(\mathrm{r}>\mathrm{a})$

$$
\left(\nabla^{2}+\mathrm{k}_{2}^{2} \lambda_{\mathrm{j}}^{2}\right) \phi_{\mathrm{j}}=0 \quad(\mathrm{r}<\mathrm{a})
$$

(2) $\mathrm{k}^{2} \phi_{\mathrm{j}}$ and $\frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \phi_{\mathrm{j}}\right)$ are continuous at $\mathrm{r}=\mathrm{a}$.
(3) $\mathrm{k}_{1}^{2} \int_{\mathrm{r}=\mathrm{n}}^{\infty} \phi_{\mathrm{j}}^{2} \mathrm{~d} \tau+\mathrm{k}_{2}^{2} \int_{\mathrm{r}=0}^{\mathrm{a}} \phi_{\mathrm{j}}^{2} \mathrm{~d} \tau=1$
where $\phi_{j}$ 's are the eigenfunctions corresponding to the eigenvalues $\lambda_{j}$ 's. The primary field pertaining to a point source leads to a three dimensional Dirac function $\frac{4 \pi i \partial(R)}{k_{1}}$ at $Q$.

The eigenfunctions are of the form

$$
\phi_{n, j}= \begin{cases}A_{n, j} \zeta^{(1)}\left(\lambda_{n, j} k_{1} r\right) P_{n}(\cos \theta) & (r>a)  \tag{1.16}\\ B_{n, j} \phi_{n}\left(\lambda_{n, j} k_{2} r\right) P_{n}(\cos \theta) & (r<a)\end{cases}
$$

where n is a positive integer. The total field satisfying conditions (1) through (3) becomes

$$
\begin{align*}
\pi_{t o t}= & \frac{4 \pi i}{k_{1}} \sum_{n=0}^{\infty} \sum_{j=0}^{n} \frac{A_{n, j}^{2}}{\left(1-\lambda_{n, j}^{2}\right)} \zeta_{n}^{(1)}\left(\lambda_{n, j} k_{1} b\right) \\
& \zeta_{n}^{(1)}\left(\lambda_{n, j} k_{1} r\right) P_{n}(\cos \theta) \quad(r>a)  \tag{1.17a}\\
\pi_{\text {tot }}= & \frac{4 \pi i}{k_{1}} \sum_{n=0}^{\infty} \sum_{j=0}^{n} \frac{A_{n, j} B_{n, j}}{\left(1-\lambda_{n, j}^{2}\right)} \zeta_{n}^{(1)}\left(\lambda_{n, j} k_{1} b\right) \\
& \phi_{n}\left(\lambda_{n, j} k_{2} r\right) P_{n}(\cos \theta) \quad(r<a) \tag{1.17b}
\end{align*}
$$

where

$$
\begin{align*}
& A_{n, j}^{2}= \frac{(2 n+1)}{2 \pi k_{1}^{2} a^{3}\left\{\zeta_{n}^{(1)}(x)\right\}^{2}\left[\left(\frac{d^{2}}{d x^{2}}+\frac{1}{x} \frac{d}{d x}\right) \log \left\{x \zeta_{n}^{(1)}(x)\right\}\right.} \\
&\left.-\frac{k_{1}^{2}}{k_{2}^{2}}\left(\frac{d^{2}}{d y^{2}}+\frac{1}{y} \frac{d}{d y}\right) \log \left\{y \psi_{n}(y)\right\}\right] \tag{1.18a}
\end{align*}
$$

and

$$
\begin{align*}
& \mathrm{B}_{\mathrm{n}, \mathrm{j}}^{2}=\left.\frac{(2 \mathrm{n}+1)}{2 \pi \mathbf{k}_{2}^{2} a^{3}\left\{\psi_{\mathrm{n}}(\mathrm{y})\right\}^{2}\left[\frac { \mathrm { k } _ { 2 } ^ { 2 } } { \mathrm { k } _ { 1 } ^ { 2 } } \left(\frac{\mathrm{~d}^{2}}{\mathrm{dx}} \mathrm{x}^{2}\right.\right.}+\frac{1}{\mathrm{x}} \frac{\mathrm{~d}}{\mathrm{dx}}\right) \log \left\{\mathrm{x} \zeta_{\mathrm{n}}^{(1)}(\mathrm{x})\right\} \\
&\left.-\left(\frac{\mathrm{d}^{2}}{\mathrm{dy}}+\frac{1}{\mathrm{y}} \frac{\mathrm{~d}}{\mathrm{dy}}\right) \log \left\{\mathrm{y} \psi_{\mathrm{n}}(\mathrm{y})\right\}\right] \tag{1.18b}
\end{align*}
$$

The difference between the earlier method and the method of the eigenfunctions is clear. In the latter the summation is done over the integers whereas in the former it is done over the complex roots of the modal equation.

To sum up, we have seen in this section that there are now clearly three different methods used in the solution of the VLF propagation problems. They are (a) the zonal harmonic series method; (b) the residue series method; and (c) the eigenfunction method. In the next two decades, most of the analyses are done following these three methods.

## THE EARTH-IONOSPHERE MODEL

To account for the effects on the VLF wave propagation due to the presence of the ionosphere, additional boundary conditions were introduced. The ionosphere was first considered as a sharply bounded homogeneous medium, then as stratified media. Magneto-ionic theory was also introduced to take care of both the isotropic case where the earth magnetic field was ignored and the anisotropic case where the earth magnetic field is included. The additional conditions introduced led to a great deal of complexity; nevertheless, the mathematical models employed had their genesis in the diffraction model.

## The Mode Theory

Among the most active proponent of the mode theory is J. R. Wait. We shall look at the problem with him.
A. Isotropic, uniform, sharply bounded ionosphere. The ionosphere is treated as a concentric homogeneous sharply bounded ionosphere. The source is a verticle dipole. The fields are expressed in terms of a Hertzian vector which has only a radial component $U$. In the region between the surface of the earth ( $r=a$ ) and the lower edge of the ionosphere ( $\mathrm{r}=\mathrm{c}$, or $8+\mathrm{h}$ ),

$$
\begin{align*}
& \mathrm{E}_{\mathrm{v}}=\left(\mathrm{k}^{2}+\frac{\partial^{2}}{\partial \mathbf{r}^{2}}\right)(\mathrm{ru}), \\
& \mathrm{E}_{\theta}=\frac{1}{\mathbf{r}} \frac{\partial^{2}}{\partial \mathbf{r} \partial \theta}(\mathrm{ru})  \tag{1.1}\\
& \mathrm{H}_{\phi}=-\mathrm{i} \epsilon \omega \frac{\partial u}{\partial \theta} \\
& \mathbf{E}_{\phi}=\mathrm{H}_{\mathrm{r}}=\mathrm{H}_{\theta}=0
\end{align*}
$$

The time factor $\exp (i \omega t)$ is understood. $\epsilon$ and $\mu$ are the electric constants and $\mathrm{k}=(\epsilon \mu)^{1 / 2} \omega$. Subscripts g and i are added to the field quantities where reference is made to the region $r<a$ and $r>c$, respectively. To account for the singularity introduced by the source at $r=b$, the Hertzian potential satisfies the inhomogeneous wave equation,

$$
\begin{equation*}
\left(\nabla^{2}+\mathrm{k}^{2}\right) \mathrm{V}=\mathrm{C} \frac{\delta(\mathrm{r}-\mathrm{b}) \delta(\theta)}{2 \pi \mathrm{r}^{2} \sin \theta} \tag{1.2}
\end{equation*}
$$

for $a<r<a+h$, where the $\delta$ 's are the dirac delta functions. The field in the region $a<r<a+h$ is written as the sum of two parts $U_{e}$ and $U_{s}$, where the latter is a solution of the homogeneous wave equation and the former satisfies the singularity due to the dipole source.

$$
\begin{equation*}
U_{s}=\frac{i k c}{4 \pi} \sum_{q=0}^{\infty}(2 q+1)\left[A_{q} h_{q}^{(2)}(k r)+B_{q} j_{q}(k r)\right] \times P_{q}(\cos \theta) \tag{1.3}
\end{equation*}
$$

where $\mathrm{j}_{\mathrm{q}}(\mathrm{kr})$ and $\mathrm{h}_{\mathrm{q}}^{(2)}(\mathrm{kr})$ are spherical Bessel function of the first kind and Hankel function of the second kind respectively; and $\mathrm{P}_{\mathrm{q}}(\cos \theta)$ is the Legendre function. And

$$
U_{c}=\frac{i k c}{4 \pi} \sum_{q=0}(2 q+1) P_{q}(\cos \theta) \begin{cases}j_{q}(k r) h_{q}^{(2)}(k b) & \text { for } r<b  \tag{1.4}\\ h_{q}^{(2)}(k r) j_{q}(k b) & \text { for } r>b\end{cases}
$$

The Hertzian potentials $U_{g}$ and $U_{i}$ are solutions of the homogeneous wave equations

$$
\begin{equation*}
\left(\nabla^{2}+k_{g}^{2}\right) U_{g}=0 \quad \text { for } 0 \leq r \leq a \tag{1.5}
\end{equation*}
$$

and

$$
\left(\nabla^{2}+k_{i}^{2}\right) U_{i}=0 \quad \text { for } r \geq c
$$

To satisfy the boundedness of $U$ at $r=0$ and $r=\infty$, the solutions are

$$
\begin{equation*}
\mathrm{U}_{\mathrm{g}}=\frac{i \mathrm{kc}}{4 \pi} \sum_{\mathrm{q}}(2 \mathrm{q}+1) \mathrm{P}_{\mathrm{q}}(\cos \theta) \mathrm{a}_{\mathrm{q}} \mathrm{j}_{\mathrm{q}}\left(\mathrm{k}_{\mathrm{q}} \mathrm{r}\right) \tag{1.6}
\end{equation*}
$$

and

$$
\begin{equation*}
U_{i}=\frac{i k c}{4 \pi} \sum_{\mathrm{q}}(2 \mathrm{q}+1) \mathrm{P}_{\mathrm{q}}(\cos \theta) \mathrm{b}_{\mathrm{q}} \mathrm{~h}_{\mathrm{q}}^{(2)}\left(\mathrm{k}_{\mathrm{i}} \mathrm{r}\right) \tag{1.7}
\end{equation*}
$$

where $a_{q}$ and $b_{q}$ are constants of integration. For the boundary conditions, Wait introduced the impedance concepts; thus, for $q^{\prime}$ th term in the series they are

$$
\begin{array}{ll}
\mathrm{E}_{\theta}^{(q)}=-\mathrm{Z}_{\mathrm{g}}^{(\mathrm{q})} \mathrm{H}_{\phi}^{(q)} & \text { at } \mathrm{r}=\mathrm{a} \\
\mathrm{E}_{\overparen{(q)}}^{(\mathrm{q})} \mathrm{Z}_{\mathrm{i}}^{(\mathrm{q})} \mathrm{H}_{\phi}^{(\mathrm{q})} & \text { at } \mathrm{r}=\mathrm{c} \tag{1.9}
\end{array}
$$

where

$$
\begin{equation*}
Z_{g}^{(q)}=\frac{1}{i \in \omega} \frac{\frac{\partial}{\partial r}\left[r j q\left(k_{q} r\right)\right]}{r j_{q}\left(k_{q} r\right)} \tag{1.10}
\end{equation*}
$$

and

$$
\begin{equation*}
Z_{i}^{(q)}=-\frac{1}{i \in \omega} \frac{\frac{\partial}{\partial r}\left[r b_{q}^{(2)}\left(k_{i} r\right)\right]}{r b_{q}^{(2)}\left(k_{i} r\right)} \tag{1.11}
\end{equation*}
$$

These can be rewritten upon replacing kr by x as

$$
\begin{array}{ll}
\frac{1}{x} \frac{\partial}{\partial x}(X U)=i\left(Z_{g}^{(q)} / \eta\right) U & \text { for } x=k a \\
\frac{1}{x} \frac{\partial}{\partial x}(X U)=-i\left(Z_{i}^{(q)} / \eta\right) U & \text { for } x=k c \tag{1.13}
\end{array}
$$

The solution for $\mathrm{a} \leq \mathrm{r} \leq \mathrm{c}$ is

$$
\begin{equation*}
\mathrm{U}=\frac{\mathrm{ikc}}{4 \pi} \sum_{\mathrm{q}=0}^{\infty}(2 \mathrm{q}+1) \mathrm{h}_{\mathrm{q}}^{(2)}(\mathrm{kb}) \mathrm{h}_{\mathrm{q}}^{(1)}(\mathrm{kr}) \times \mathrm{P}_{\mathrm{q}}(\cos \theta) \frac{\mathrm{F}_{\mathrm{q}}}{\mathrm{D}_{\mathrm{q}}} \tag{1.14}
\end{equation*}
$$

where

$$
\begin{gather*}
F_{q}=\left[1+R_{g}^{(q)} \frac{h_{q}^{(1)}(k a) h_{q}^{(2)}(k r)}{h_{q}^{(2)}(k a) h_{q}^{(1)}(k r)}\right]\left[1+R_{i}^{(q)} \frac{h_{q}^{(2)}(k c) h_{q}^{(1)}(k b)}{h_{q}^{(1)}(k c) h_{q}^{(2)}(k b)}\right]  \tag{1.15}\\
D_{q}=1-R_{g}^{(q)} R_{i}^{(q)} \frac{h_{q}^{(1)}(k a) h_{q}^{(2)}(k c)}{h_{q}^{(2)}(k a) h_{q}^{(1)}(k c)}  \tag{1.16}\\
R_{g}^{(q)}=\frac{\frac{d}{d x}\left[l u \times h_{q}^{(1)}(x)\right]_{x=k a}-i Z_{g}^{(q)} / \eta}{\frac{d}{d x}\left[l u \times h_{q}^{(2)}(x)\right]_{x=k a}-i Z_{g}^{(q)} / \eta} \tag{1.17}
\end{gather*}
$$

$$
\begin{equation*}
R_{i}^{(q)}=-\frac{\frac{d}{d x}\left[\ell u \times h_{q}^{(2)}(x)\right]_{x=k c}+i Z_{i}^{(q)} / \eta}{\frac{d}{d x}\left[\ell u \times h_{q}^{(1)}(x)\right]_{x=k c}+i Z_{i}^{(q)} / \eta} \tag{1.18}
\end{equation*}
$$

By means of the Watson transformation, the solution becomes

$$
\begin{equation*}
\mathrm{U}=-\mathrm{ikc} \sum_{\nu} \frac{\nu+\frac{1}{2}}{\sin \nu \pi} h_{\nu}^{(1)}(\mathrm{kb}) \mathrm{h}_{\nu}^{(1)}(\mathrm{kr}) \frac{\mathrm{F}_{\nu}}{\mathrm{D}_{\nu}^{\prime}} \mathrm{P}_{\nu}[\cos (\pi-\theta)] \tag{1.19}
\end{equation*}
$$

where $D_{\nu}^{\prime}=\partial / \partial \nu D_{\nu}$. All the values in the summand are to be evaluated at the poles of $f(\nu)$ which are the roots of the model equation

$$
\begin{equation*}
\mathrm{D}_{\nu}=0 \tag{1.20}
\end{equation*}
$$

If use is made of the Airy integral as expounded by V. A. Fock in place of the Hankel function, the radial component of the electric field can be written in the form

$$
\begin{equation*}
\mathbf{E}_{\mathrm{r}} \cong \frac{\mathrm{e}^{-\mathrm{ika} \theta}}{\mathrm{a}(\theta \sin \theta)^{1 / 2}} \mathrm{~V} \tag{1.21}
\end{equation*}
$$

apart from a constant factor.
And

$$
\begin{equation*}
V=e^{i \pi / 4}\left(\frac{x}{\pi}\right)^{1 / 2} \oint \frac{e^{-i x t}\left[w_{1}(t-y)+A(t) w_{2}(t-y)\right]}{\left[w_{1}^{\prime}(t)-q w_{1}(t)\right][1-A(t) B(t)]} d t \tag{1.22}
\end{equation*}
$$

where

$$
\begin{equation*}
\mathrm{x}=(\mathrm{ka} / 2)^{1 / 2} \theta, \quad \mathrm{y}=(2 / \mathrm{ka})^{1 / 3} \mathrm{k}(\mathrm{r}-\mathrm{a}) \tag{1.23}
\end{equation*}
$$

Now the contour is to enclose the complex poles tn's which are solutions of

$$
\begin{equation*}
1-A(t) B(t)=0 \tag{1.24}
\end{equation*}
$$

where

$$
\begin{gather*}
A(t)=-\left[\frac{w_{1}^{\prime}\left(t-y_{0}\right)+q_{i} w_{1}\left(t-y_{0}\right)}{w_{2}^{\prime}\left(t-y_{0}\right)+q_{i} w_{2}\left(t-y_{0}\right)}\right] .  \tag{1.25}\\
B(t)=-\left[\frac{w_{2}^{\prime}(t)-q w_{2}(t)}{w_{1}^{\prime}(t)-q w_{1}(t)}\right] .  \tag{1.26}\\
y_{0}=(2 / k a)^{1 / 3} k h, \quad q_{i}=-i(k a / 2)^{1 / 3} Z / \eta_{0}, \quad \eta_{0}=120 \pi
\end{gather*}
$$

and

$$
q=-i(k a / 2)^{1 / 3}\left(\frac{i \epsilon_{0} \omega}{\sigma_{g}+i \epsilon_{g} \omega}\right)\left(1-\frac{i \epsilon_{0} \omega}{\sigma_{g}+i \epsilon_{g} \omega}\right)^{1 / 2}
$$

The residue or mode series is

$$
\begin{equation*}
V=-2(\pi x)^{1 / 2} e^{-i \pi / 4} \sum_{n=0}^{\infty} \frac{e^{-i x t_{n}}\left[w_{1}\left(t_{n}-y\right)+A\left(t_{n}\right) w_{2}\left(t_{n}-y\right)\right]}{\left[w_{1}^{\prime}\left(t_{n}\right)-q w_{1}\left(t_{n}\right)\right]\left[\frac{\partial}{\partial t} A(t) B(t)\right]_{t=t_{n}}}(1 \tag{1.27}
\end{equation*}
$$

It may also be written as

$$
\begin{equation*}
V=\frac{4(\pi x)^{1 / 2}}{y_{0}} e^{-i \pi / 4} \sum_{n=0}^{\infty} e^{-i x t_{n}} G_{n}(\hat{y}) G_{n}(y) \Lambda n \tag{1.28}
\end{equation*}
$$

where $\hat{y}=(2 / \mathrm{ka})^{1 / 3} \mathrm{k} \mathrm{Z}_{0}$. The functions $G_{n}$ are height functions and they are normalized to unity for $y$ or $y_{0}$ equal to zero. Explicitly

$$
\begin{gather*}
G_{n}(y)=\frac{f\left(t_{n}, y\right)}{f\left(t_{n}, 0\right)}  \tag{1.29}\\
f\left(t_{n}, y\right)=w_{1}\left(t_{n}-y\right)+A\left(t_{n}\right) w_{2}\left(t_{n}-y\right) \tag{1.30}
\end{gather*}
$$

The function $\Lambda_{n}$ is a modal excitation factor. It is normalized so that, in the limit of zero curvature and perfectly reflecting boundaries, it becomes unity for all modes. Explicitly,

$$
\begin{equation*}
\Lambda_{n}=\frac{y_{0}}{2}\left[\left(t_{n}-q^{2}\right)-\frac{\left(t_{n}-y_{0}-q_{i}^{2}\right)\left[w_{2}^{\prime}\left(t_{n}\right)-q w_{2}\left(t_{n}\right)\right]^{2}}{\left[w_{2}^{\prime}\left(t_{n}-y_{0}\right)+q_{i} w_{2}\left(t_{n}-y_{0}\right)\right]^{2}}\right]^{-1} \tag{1.31}
\end{equation*}
$$

b. Anisotropic Ionosphere. For the case of an anisotropic ionosphere where the earth magnetic field is not ignored, the electromagnetic field components in the homogeneous space ( $a \leq r \leq c$ ) are expressed in terms of purely TM and TE waves. They are derived from two scalar functions $U$ and $V$ as follows:

$$
\begin{align*}
& E_{r}=\left(\frac{\partial^{2}}{\partial r^{2}}+k^{2}\right)(r U) \\
& E_{\theta}=\frac{1}{r} \frac{\partial^{2}}{\partial \theta \partial r}(r U)-\frac{i \mu \omega}{r \sin \theta} \frac{\partial}{\partial \phi}(r V) \\
& E_{\phi}=\frac{1}{r \sin \theta} \frac{\partial^{2}}{\partial \phi \partial r}(r U)+\frac{i \mu \omega}{r} \frac{\partial}{\partial \theta}(r V)  \tag{1.32}\\
& H_{r}=\left(\frac{\partial^{2}}{\partial r^{2}}+k^{2}\right)(r V) \\
& H_{\theta}=\frac{1}{r} \frac{\partial^{2}}{\partial \theta \partial r}(r V)+\frac{i \in \omega}{r \sin \theta} \frac{\partial}{\partial \phi}(r U) \\
& H_{\phi}=\frac{1}{r \sin \theta} \frac{\partial^{2}}{\partial r \partial \phi}(r V)-\frac{i \in \omega}{r} \frac{\partial}{\partial \theta}(r U)
\end{align*}
$$

The functions $U$ and $V$ satisfy

$$
\left(\nabla^{2}+k^{2}\right)\left\{\begin{array}{l}
U  \tag{1.33}\\
V
\end{array}\right\}=0
$$

in source free region. The source again is a vertical dipole. The primary field of such a source is of a purely TM character. It is derived from a single scalar function $U_{e}$ satisfying the following inhomogeneous wave equation,

$$
\begin{equation*}
\left(\nabla^{2}+\mathrm{k}^{2}\right) \mathrm{U}_{\mathrm{e}}=\mathrm{c} \frac{\partial(\mathrm{r}-\mathrm{b}) \partial(\theta)}{2 \pi \mathrm{r}^{2} \sin \theta} \tag{1.34}
\end{equation*}
$$

The boundary conditions, in terms of surface impedance, are

$$
\left.\begin{gather*}
\frac{\partial}{\partial x}(x U)=i \Delta_{g} x U \\
\frac{\partial}{\partial x}(x \bar{V})=\left(i / \Delta_{g}\right) x \bar{v} \tag{1.35}
\end{gather*}\right|_{x=k a}
$$

where

$$
\Delta_{\mathrm{g}}=\mathbf{Z}_{\mathrm{g}} / \eta, \quad \mathrm{U}=\mathrm{U}_{\mathrm{e}}+\mathrm{U}_{\mathrm{s}}, \quad \nabla=\eta \mathrm{V}
$$

and

$$
\left.\begin{align*}
& \frac{\partial}{\partial x}(x U)=\Delta_{11} \frac{\partial}{\partial x}(x \bar{V})-i \Delta_{12}(x U) \\
& -i(x \bar{V})=\Delta_{21} \frac{\partial}{\partial x}(x \bar{V})-i \Delta_{22}(x U) \tag{1.36}
\end{align*}\right|_{x=k c}
$$

where

$$
\Delta_{11}=Z_{11} / \eta, \quad \Delta_{12}=Z_{12} / \eta, \quad \Delta_{21}=Z_{21} / \eta
$$

and

$$
\Delta_{22}=Z_{22} / \eta
$$

The radial field in the concentric waveguide region written in the matrix form is

$$
\left[\begin{array}{c}
\mathbf{E}_{\mathbf{r}}  \tag{1.37}\\
\eta \mathbf{H}_{\mathrm{r}}
\end{array}\right] \cong \frac{\mathrm{e}^{-\mathrm{ika} \theta}}{\mathrm{a}(\theta \sin \theta)^{1 / 2}}\left[\begin{array}{c}
\mathbf{F}_{\mathrm{e}} \\
\mathbf{F}_{\mathrm{h}}
\end{array}\right] \frac{\mathrm{i} \mu_{0} \omega \mathbf{I d s}}{4 \pi}
$$

where

$$
\left[\begin{array}{l}
F_{e}  \tag{1.38}\\
F_{h}
\end{array}\right]=-2(\pi x)^{1 / 2} e^{-i \pi / 4} \sum_{n=0}^{\infty} \frac{e^{-i x t_{n}\left[w_{1}\left(t_{n}-y\right)+A\left(t_{n}\right) w_{2}\left(t_{n}-y\right)\right]}}{\left[w_{1}^{\prime}\left(t_{n}\right)-q w_{1}\left(t_{n}\right)\right]\left[\frac{\partial}{\partial t} A(t) B(t)\right]_{t_{n}}}\left[\begin{array}{l}
1 \\
0
\end{array}\right]
$$

The roots $t_{n}$ 's are determined from the matrix equation

$$
\begin{equation*}
[\mathrm{A}(\mathrm{t})][\mathrm{B}(\mathrm{t})]=\mathrm{I} \tag{1.39}
\end{equation*}
$$

where

$$
A(t)=\left[\begin{array}{ll}
\| R_{\| I} & { }_{\perp} R_{\|}  \tag{1.40}\\
{ }_{\| 1} R_{\perp} & { }_{\perp} R_{\perp}
\end{array}\right] \exp \left[-i(4 / 3)\left(y_{0}-t\right)^{3 / 2}-i \pi / 2\right]
$$

and

$$
\begin{align*}
& B(t)=\left[\begin{array}{cc}
-\left[\frac{w_{2}^{\prime}(t)-q w_{2}(t)}{w_{1}^{\prime}(t)-q w_{1}(t)}\right] & 0 \\
0 & -\left[\frac{w_{2}^{\prime}(t)-\hat{q} w_{2}(t)}{w_{1}^{\prime}(t)-\hat{q} w_{1}(t)}\right]
\end{array}\right]  \tag{1.41}\\
& q=-i(k a / 2)^{1 / 3} \Delta_{g}, \text { with } \Delta_{g}=\left(\frac{i \epsilon_{0} \omega}{\sigma_{g}+i \epsilon_{g} \omega}\right)^{1 / 2}\left(1-\frac{i \epsilon_{0} \omega}{\sigma_{g}+i \epsilon_{g} \omega}\right)^{1 / 2}
\end{align*}
$$

and

$$
\hat{\mathrm{q}}=-\mathrm{i}(\mathrm{ka} / 2)^{1 / 2} \Delta_{\mathrm{g}}^{\mathrm{h}}, \text { with } \Delta_{\mathrm{g}}^{\mathrm{h}}=\left(\frac{\sigma_{\mathrm{g}}+\mathrm{i} \epsilon_{\mathrm{g}} \omega}{\mathrm{i} \epsilon_{0} \omega}-1\right)^{1 / 2} .
$$

${ }_{\| 1} R_{\| 1},{ }_{1} R_{\| I},{ }_{\|} R_{\perp}$, and ${ }_{1} R_{\perp}$ are the reflection coefficients, where $\|$ and refer to the direction of the electric vector, relative to the plane of incidence, of the incident (first suffix) and reflected (second suffix) wave.

## The Zonal Harmonic Series

The stumbling block of the mode theory is the rather formidable transcendental modal equation to solve. Recently Wait and Spies made detailed computation of the characteristics of the least attenuated modes. However, a single mode is not adequate to describe the complete field satisfying Maxwell's equations at either intermediate and great distances for daytime models of the lower ionosphere. In the case of a nighttime model even more modes are needed at all distances.

With the speed of the modern digital computers, the argument once held against the zonal harmonic solution of the VLF propagation problem is no longer convincing. Granted that the zonal harmonic series converges rather slow. Johler showed that with improved summation technique, the results of the computations indicate that full rigor can be achieved with comparative ease at frequencies less than about $50 \mathrm{kc} / \mathrm{s}$, leaving only the assumed model for the transmitter and the propagation medium and avoiding the complications of the Watson transformation.

For a simple model of the terrestrial sphere of radius a surrounded by concentric ionosphere from $r=c$ to $r=\infty$, the field components are expressed in terms of the Hertzian potential functions, $\pi^{\theta}$ and $\pi^{\mathrm{m}}$ as follows,

$$
\begin{align*}
& \mathrm{E}_{\mathrm{r}}=\frac{-1}{\mathrm{rb} \sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial \pi^{\mathrm{e}}}{\partial \theta}\right) \\
& \mathrm{E}_{\theta}=\frac{1}{\mathrm{rb}} \frac{\partial^{2}}{\partial \mathrm{r} \partial \theta}\left(\mathrm{r} \pi^{\mathrm{e}}\right) \\
& \mathrm{H}_{\phi}=\frac{1}{\mathrm{~b}} \frac{\mathrm{k}^{2}}{\mu_{0} \mathbf{i} \omega} \frac{\partial \pi^{\mathrm{e}}}{\partial \theta}  \tag{2.1}\\
& \mathrm{E}_{\phi}=\frac{\mu_{0} \mathrm{i} \omega}{\mathrm{~b}} \frac{\partial \pi^{\mathrm{m}}}{\partial \theta} \\
& \mathrm{H}_{\mathrm{r}}=\frac{-1}{\mathrm{rb} \sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial \pi^{\mathrm{m}}}{\partial \theta}\right) \\
& \mathrm{H}_{\theta}=\frac{1}{\mathrm{rb}} \frac{\partial^{2}}{\partial \mathrm{r} \partial \theta}\left(\mathrm{r} \pi^{\mathrm{m}}\right)
\end{align*}
$$

where $\pi^{\theta}$ and $\pi^{\mathrm{m}}$ satisfy the wave equations

$$
\left(\nabla^{2}+k^{2}\right)\left\{\begin{array}{l}
\pi^{\mathrm{e}}  \tag{2.2}\\
\pi^{\mathrm{m}}
\end{array}\right\}=0
$$

The media of propagation are characterized by their electric constants in the wave number, $k$. Thus, for air, with index of refraction, $\eta$, or dielectric constant $\epsilon_{1}=\eta_{1}^{2}$

$$
\mathrm{k}^{2}=\mathrm{k}_{1}^{2}=\frac{\omega^{2}}{\mathrm{c}^{2}} \epsilon_{1}
$$

or

$$
\begin{equation*}
\mathrm{k}_{1}=\frac{\omega}{c} \eta_{1} \tag{2.3}
\end{equation*}
$$

For the ground

$$
\begin{equation*}
\mathrm{k}=\mathrm{k}_{2}=\frac{\omega}{\mathrm{c}} \sqrt{\epsilon_{2}-\mathrm{i} \frac{\sigma \mu_{0} \mathrm{c}^{2}}{\omega}} \tag{2.4}
\end{equation*}
$$

where $\sigma$ is the ground conductivity and $\epsilon_{2}$ is the relative dielectric constant.
The wave number of the ionosphere is rather intricate.

$$
\begin{equation*}
\mathrm{k}=\mathrm{k}_{2}=\frac{\omega}{\mathrm{c}} \eta_{0, \mathrm{e}}^{\mathrm{i}, \mathrm{r}} \tag{2.5}
\end{equation*}
$$

There are four distinct propagation components with the complex index of refraction. $\eta_{0, \mathrm{e}}^{\mathrm{i}, \mathrm{r}}$ corresponding to ordinary wave ( 0 ), and extraordinary wave (e), upgoing wave (i), and downgoing wave (r). They are related to the roots of a Booker type quartic in the parameters, $\zeta$, which is the result of the simultaneous solution of the Langevin equations of charged particles and Maxwell's equations. The index of refraction, $\eta$,

$$
\begin{equation*}
\eta^{2}=\zeta^{2}+\sin ^{2} \phi \tag{2.6}
\end{equation*}
$$

where $\phi_{i}$ is the angle of incidence on the lower ionosphere plasma.

The boundary conditions are

$$
\left.\begin{gather*}
\mathrm{k}_{1}^{2}\left(\pi_{0}^{\mathrm{e}}+\pi_{1}^{\mathrm{e}}\right)=\mathrm{k}_{2}^{2}\left(\pi_{2}^{\mathrm{e}}\right) \\
\frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \pi_{0}^{\mathrm{e}}+\mathrm{r} \pi_{1}^{\mathrm{e}}\right)=\frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \pi_{2}^{\mathrm{e}}\right)  \tag{2.7}\\
\pi_{1}^{\mathrm{m}}=\pi_{2}^{\mathrm{m}} \\
\frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \pi_{1}^{\mathrm{m}}\right)=\frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \pi_{2}^{\mathrm{m}}\right) \\
\mathrm{k}_{1}^{2}\left(\pi_{0}^{\mathrm{e}}+\pi_{1}^{\mathrm{e}}\right)=\mu_{0} \mathrm{i} \omega \mathrm{Q}_{\mathrm{em}}^{-1} \frac{1}{\mathrm{r}} \frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \pi_{3}^{\mathrm{m}}\right) \\
\frac{1}{\mathrm{r}} \frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \pi_{0}^{\mathrm{e}}+\mathrm{r} \pi_{1}^{\mathrm{e}}\right)=\mu_{0} \mathrm{i} \omega \mathrm{Q}_{\mathrm{me}}^{-1} \pi_{3}^{\mathrm{m}}  \tag{2.8}\\
\pi_{1}^{\mathrm{m}}=\mathrm{Q}_{\mathrm{me}} \frac{1}{\mathrm{r}} \frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \pi_{3}^{\mathrm{e}}\right) \\
\mu_{0} \mathrm{i} \omega \frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \pi_{1}^{\mathrm{m}}\right)=\mathrm{Q}_{\mathrm{em}} \mathrm{k}_{3}^{2} \pi_{3}^{\mathrm{e}}
\end{gather*}\right|_{\mathrm{r}=\mathrm{c}}
$$

where in the quasi-longitudinal approximation,

$$
\begin{gather*}
\mathrm{Q}_{\mathrm{me}}=\frac{\mathrm{E}_{\phi}}{\mathrm{E}_{\theta}}= \pm \frac{\mathrm{k}_{3} / \mathrm{k}_{1}}{\sqrt{k_{3}^{2} / \mathrm{k}_{1}^{2}-\rho}} \\
\mathrm{P}_{\mathrm{me}}=\frac{\mathbf{E}_{\mathrm{r}}}{\mathrm{E}_{\theta}}=\frac{-\sin \phi_{\mathrm{i}}}{\zeta} . \tag{2.9}
\end{gather*}
$$

$\rho=\sin ^{2} \phi_{2}$ is the complex direction sine squared of propagation, and the minus sign ( - ) is taken with the ordinary wave, $\mathrm{k}_{3,0}$ and the plus sign ( + ) is taken with the extraordinary wave $\mathrm{k}_{3, \mathrm{e}}$. The solutions are

$$
\begin{align*}
& \pi_{2}^{e}=\frac{1}{k_{2}^{2} r b} \sum_{n=0}^{\infty}(2 n+1) a_{n}^{e} \psi_{n}\left(k_{2} r\right) P_{n}(\cos \theta) \\
& \pi_{2}^{m}=\frac{1}{k_{2}^{2} r b} \sum_{n=0}^{\infty}(2 n+1) a_{n}^{m} \psi_{n}\left(k_{2} r\right) P_{n}(\cos \theta) \\
& \pi_{3}^{e}=\frac{1}{k_{3}^{2} r b} \sum_{n=0}^{\infty}(2 n+1) d_{n}^{e} \zeta_{n}\left(k_{3} r\right) P_{n}(\cos \theta) \\
& \pi_{3}^{m}=\frac{1}{k_{3}^{2} r b} \sum_{n=0}^{\infty}(2 n+1) d_{n}^{m} \zeta_{n}\left(k_{3} r\right) P_{n}(\cos \theta) \tag{2.10}
\end{align*}
$$

The constants $a_{n}^{e}, b_{n}^{e}$ and the like are determined by the boundary conditions.

For a stratified ionosphere model, the matching for boundary conditions could be a rather formidal task. (For details see Reference 22.)

## The Full-Wave Theory

It is seen that when coupling is introduced into the wave propagation problem, no matter what mathematical model one uses whether it is based on the residue series or the zonal harmonic series, one often resorts to the digital computer for numerical computations. Budden is one among many of the English schools who rely more and more on the direct solution of the full wave equation using computers. This approach is formally of complete generality and does not depend on the assumption of a special ionospheric model.

In deriving the differential equations, Budden uses the cartesian coordinates with the $z$-axis directed vertically upwards. A plane wave is incident at angle $\theta_{i}$ to the verticle on the ionosphere from below. Let $s=\sin \theta_{i}, c=\cos \theta_{i}$. Let the ionosphere be stratified such that in each of which the medium is considered homogeneous. For the waves in each stratum $\partial / \partial x()=-i k s() ;$ $\partial / \partial \mathrm{y}()=0$. Then Maxwell's equations give,

$$
\begin{gather*}
-\frac{d E_{y}}{d z}=-i k H_{x} \\
\frac{d E_{x}}{d z}+i k s E_{z}=-i k H_{y} \\
-i k s E_{y}=-i k H_{z}  \tag{3.1a}\\
-\frac{d H_{y}}{d z}=\frac{i k}{\epsilon_{0}} D_{x} \\
\frac{d H_{x}}{d z}+i k s H_{z}=\frac{i k}{\epsilon_{0}} D_{y} \\
-i k s H_{y}=\frac{i k}{\epsilon_{0}} D_{z} \tag{3.1b}
\end{gather*}
$$

From the magneto-ionic theory, after the Z-components of the fields are eliminated, we can write the following matrix equation,

$$
\begin{equation*}
\frac{\mathrm{de}}{\mathrm{dZ}}=-\mathrm{ikTe} \tag{3.2}
\end{equation*}
$$

where

$$
e=\left(\begin{array}{c}
E_{x}  \tag{3.3}\\
-E_{y} \\
H_{x} \\
H_{y}
\end{array}\right)
$$

and T is a $4 \times 4$ matrix whose elements are related to the parameters describing the ionosphere, such as the electron density, the electron collision frequency, the earth's magnetic field and others.

The matrix $T$ has four characteristic roots or eigenvalues $q_{i}(i=1,2,3,4)$ which satisfy the characteristic equation

$$
\begin{equation*}
\operatorname{det}(T-q I)=0 \tag{3.4}
\end{equation*}
$$

where I is the $4 \times 4$ unit matrix.
In order to solve the matrix equation, it is convenient to introduce a new column matrix f, thus

$$
f=\left(\begin{array}{c}
f_{1}  \tag{3.5}\\
f_{2} \\
f_{3} \\
f_{4}
\end{array}\right)
$$

such that

$$
\begin{equation*}
\mathrm{e}=\mathrm{S} \mathrm{f} \tag{3.6}
\end{equation*}
$$

where $S$ is some $4 \times 4$ matrix whose elements are functions of $z$ to be determined.

Assuming that $S$ is non-singular so that its inverse $\mathrm{S}^{-1}$ exists. Substitution of (3.6) in (3.2) and premultiplication of $S^{-1}$ gives

$$
\begin{equation*}
f^{\prime}+i k S^{-1} T S f=-S^{-1} S^{\prime} f \tag{3.7}
\end{equation*}
$$

where the dash indicates the derivative with respect to z .
The $S$ is so chosen as to make $f_{i}$ the only element of $f$ appearing on the left hand side of the $i^{\prime}$ th equation in (3.7). This means that $S^{-1}$ TS must be a diagonal matrix, this can be done if the characteristic roots $q_{i}$ of $T$ are all distinct. It can then be shown that

$$
S^{-1} T S=\left[\begin{array}{cccc}
q_{1} & 0 & 0 & 0  \tag{3.8}\\
0 & \mathrm{q}_{2} & 0 & 0 \\
0 & 0 & \mathrm{q}_{3} & 0 \\
0 & 0 & 0 & \mathrm{q}_{4}
\end{array}\right]=\Delta
$$

and Eq. (3.7) becomes

$$
\begin{equation*}
f^{\prime}+i k \Delta f=-S^{-1} S^{\prime} f \tag{3.9}
\end{equation*}
$$

Equations (3.9) are the four, first order, coupled differential equations. For a homogeneous medium $S^{\prime}=0$ and (3.9) becomes

$$
\begin{equation*}
f^{\prime}+i k \Delta f=0 \tag{3.10}
\end{equation*}
$$

or

$$
\begin{equation*}
f_{i}^{\prime}+i k q_{i} f_{i}=0 \quad(i=1,2,3,4) \tag{3.11}
\end{equation*}
$$

The independent characteristic waves in a homogeneous medium are determined by

$$
\begin{equation*}
f_{i}=e^{-i k q_{i} z} \quad(i=1,2,3,4) \tag{3.12}
\end{equation*}
$$

Now, premultiply (3.9) by any $4 \times 4$ matrix, say $F$, (3.9) then can be written as

$$
\begin{equation*}
(\mathbb{F f})^{\ell}+\left(\mathrm{ikF} \triangle-\mathrm{F}^{\prime}\right) \mathrm{f}=-\mathrm{FS} \mathrm{~S}^{-1} \mathrm{~S}^{\prime} \mathrm{f} \tag{3.13}
\end{equation*}
$$

Choose F so that

$$
\begin{equation*}
F^{\prime}=\mathrm{ikF} \Delta, \tag{3.14}
\end{equation*}
$$

and

$$
\begin{equation*}
(\mathrm{Ff})^{\prime}=-\mathrm{FS}^{-1} \mathrm{~S}^{\prime} \mathrm{f} \tag{3.15}
\end{equation*}
$$

and integrate (3.15) in the form

$$
\begin{equation*}
\mathrm{f}=\mathrm{F}^{-1}-\mathrm{F}^{-1} \int^{z} \mathrm{FS}^{-1} \mathrm{~S}^{\prime} \mathrm{fd} z \tag{3.16}
\end{equation*}
$$

From (3.14) $F$ may be taken as the diagonal matrix whose diagonal elements are $\exp \left(i k \int^{z} q_{i} d z\right)$; hence, if $f_{0}$ is the fundamental solution of (3.10) namely

$$
\begin{equation*}
f_{0 i}=\exp \left[-i k \int^{z} q_{i} d z\right] \tag{3.17}
\end{equation*}
$$

(3.16) becomes

$$
\begin{equation*}
\mathrm{f}=\mathrm{f}_{0}-\mathrm{f}_{0} \int^{\mathrm{z}} \mathrm{f}_{0}^{-1} \mathrm{~S}^{-1} \mathrm{~S}^{\prime} \mathrm{f}_{0} \mathrm{~d} \mathrm{Z} \tag{3.18}
\end{equation*}
$$

The evaluation of $S^{-1} S^{\prime}$ and its properties can be found in Budden's book. The thing of interest to us here is that we have a set of first order differential equations (3.15), which can be integrated numerically choosing various ionospheric models. A work of this kind has been undertaken by Budden and his coworkers in the Cavandish Laboratory. Their method of attack is to assume some plausible law for the variation of electron density and collision frequency with height in the ionosphere, and to calculate the reflecting properties of this model ionosphere for long and very long waves. This is repeated for a series of models, and those models are selected which have properties most closely resembling the experimental results.

## The Normal Wave Solution

In all the previous methods of solution of the VLF wave propagation problem, numerical checks were made by assuming a prior knowledge of the parameters used. The selections of the parameters though based on sound and scientific judgment, were, to say the least, arbitrary. Furthermore, these methods could account for only homogeneous tracks (night or day) up to now. The diurnal changes in both the phase and the amplitude of the VLF waves remained mostly unexplained. In view of the inadequacy in these approaches, Krasnushkin proposed a mixed method. He suggested to use the data known in the short range propagation to evaluate some of the unknown parameters. This is basically an inverse method. With the unknown parameters determined, a direct method is used to evaluate the fields at great distances. In doing so, some of the difficulties encountered in evaluating the roots of a transcendental equation as in the mode theory could be circumvented.

The mathematical models he used are based on the theory of non-self adjoint differential operator eigenfunction expansions. He called these functions normal waves.
a. Stratified ionosphere. In a spherically stratified media, let the dielectric tensors be as follows,

$$
\begin{gather*}
\left\|\epsilon_{\mathrm{k} \|}=\right\| \begin{array}{ccc}
\epsilon_{\mathrm{rr}}^{\mathrm{k}} & 0 & 0 \\
0 & \epsilon_{\theta \theta}^{\mathrm{k}} & \epsilon_{\theta \phi}^{\mathrm{k}} \\
0 & \epsilon_{\theta \phi}^{\mathrm{k}} & \epsilon_{\theta \theta}^{\mathrm{k}}
\end{array} \|  \tag{4.1}\\
\mathrm{r}_{\mathrm{k}-1} \leq \mathrm{r} \leq \mathrm{r}_{\mathrm{R}} \quad \mathrm{k}=0,1,2 \cdots, \mathrm{~N}
\end{gather*}
$$

where, for the zero layer ( $k=0$ ), or ( $r_{-1}=0, r_{0}$ )

$$
\begin{equation*}
\epsilon_{\mathrm{rr}}^{0}=\epsilon_{\theta \theta}^{0}=\epsilon^{\prime}+\mathbf{i} \frac{4 \pi \sigma_{0}}{\omega} \mathbf{i} \epsilon_{\theta \phi}^{0}=0 \tag{4.2}
\end{equation*}
$$

for the first layer $(k=1) ;\left(r_{0}=a, r_{1}\right)$

$$
\begin{equation*}
\epsilon_{\mathrm{r} r}^{\prime}=\epsilon_{\theta \theta}^{\prime}=1 \quad \epsilon_{\theta \phi}^{\prime}=0 \tag{4.3}
\end{equation*}
$$

and for the other $n-1$ layers ( $k=2,3, \ldots N)$

$$
\left(r_{1}, r_{2}\right) ;\left(r_{2}, r_{3}\right) \cdots \cdots\left(r_{N-1}, \quad r_{N}=\infty\right)
$$

the tensor $\left\|\epsilon_{\mathrm{k}}\right\|$ is as follows,

$$
\begin{align*}
& \epsilon_{\mathrm{rr}}=1+\frac{\mathbf{i} \omega_{0}^{2}}{\omega\left(\mathrm{~V}_{\mathrm{eff}}-\mathrm{i} \omega\right)}, \\
& \epsilon_{\theta \theta}=1+\frac{\mathrm{i} \omega_{0}^{2}\left(\mathrm{~V}_{\mathrm{eff}}-\mathrm{i} \omega\right)}{\omega\left[\left(\mathrm{V}_{\mathrm{eff}}-\mathrm{i} \omega\right)^{2}+\omega_{\mathrm{H}}^{2}\right]}  \tag{4.4}\\
& \epsilon_{\theta \phi}=-\frac{\mathrm{i} \omega_{0}^{2} \omega_{\mathrm{H}}}{\omega\left[\left(\mathrm{~V}_{\mathrm{eff}}-\mathrm{i} \omega\right)^{2}+\omega_{\mathrm{H}}^{2}\right]}
\end{align*}
$$

where

$$
\begin{align*}
& \omega_{0}^{2}=\frac{4 \pi \mathrm{Ne}^{2}}{\mathrm{~m}} \\
& \omega_{\mathbf{H}}=-\frac{\mathrm{e} \mathrm{H}_{0}}{\mathrm{mc}} \tag{4.5}
\end{align*}
$$

The components of the field in each layer are

$$
\begin{gathered}
\frac{1}{\mathrm{r} \sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \mathbf{H}_{\phi}^{\mathrm{k}}\right)=-\frac{\mathrm{i} \omega}{\mathrm{c}} \epsilon_{\mathrm{rr}}^{\mathrm{k}} \mathrm{E}_{\mathrm{r}}^{\mathrm{k}}+\frac{4 \pi}{\mathrm{c}} \mathrm{I}_{\mathrm{r}}^{\mathrm{k}}, \\
\frac{1}{\mathrm{r}} \frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \mathrm{H}_{\phi}^{\mathrm{k}}\right)=\frac{\mathrm{i} \omega}{\mathrm{c}}\left[\epsilon_{\theta \theta}^{\mathrm{k}} \mathrm{E}_{\theta}^{\mathrm{k}}+\epsilon_{\theta \phi}^{\mathrm{k}} \mathrm{E}_{\theta}^{\mathrm{k}}\right], \\
\frac{1}{\mathbf{r}}\left[\frac{\partial}{\partial \mathbf{r}}\left(\mathrm{r} \mathrm{H}_{\theta}^{\mathrm{k}}\right)-\frac{\partial \mathrm{H}_{\mathrm{r}}^{\mathrm{k}}}{\partial \theta}\right]=-\frac{i \omega}{\mathrm{c}}\left[\epsilon_{\phi \theta}^{\mathrm{k}} \mathrm{E}_{\theta}^{\mathrm{k}}+\epsilon_{\phi \phi}^{\mathrm{k}} \mathrm{E}_{\phi}^{\mathrm{k}}\right],
\end{gathered}
$$

$$
\begin{gather*}
\frac{1}{\mathrm{r} \sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \mathrm{E}_{\phi}^{\mathrm{k}}\right)=-\frac{i \omega}{\mathrm{c}} \mathrm{H}_{\mathrm{r}}^{\mathrm{k}} \\
-\frac{1}{\mathrm{r}} \frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \mathrm{E}_{\phi}^{\mathrm{k}}\right)=\frac{i \omega}{\mathrm{c}} \mathrm{H}_{\theta}^{\mathrm{k}} \\
\frac{1}{\mathrm{r}}\left[\frac{\partial}{\partial \mathrm{r}}\left(\mathrm{r} \mathrm{E}_{\theta}^{\mathrm{k}}\right)-\frac{\partial \mathbf{E}_{\mathrm{r}}^{\mathrm{k}}}{\partial \theta}\right]=\frac{i \omega}{\mathrm{c}} H_{\phi}^{k} \tag{4.6}
\end{gather*}
$$

They are subject to the following conditions:

1. Tangential components $\mathrm{E}_{\theta}, \mathrm{E}_{\phi}, \mathrm{H}_{\theta}, \mathrm{H}_{\phi}$ are continuous on the spherical surfaces $r_{R}=$ const., $R=1,2, \ldots, N$.
2. In the center of the earth $\left(r=r_{-1}=0\right)$ the fields are bounded and
3. At infinity they tend to zero.

Now, introduce the Hertzian potential function

$$
\left|\begin{array}{l}
\mathrm{B}\left(\mathrm{r}_{1} \theta\right)  \tag{4.7}\\
\mathrm{A}\left(\mathrm{r}_{1} \theta\right)
\end{array}\right|
$$

where

$$
\begin{gathered}
\mathrm{H}_{\phi}=\frac{1}{\mathrm{r}} \frac{\partial \mathrm{~B}}{\partial \theta} ; \quad \mathrm{E}_{\phi}=\frac{1}{\mathrm{r}} \frac{\partial \mathrm{~A}}{\partial \theta} \\
\mathrm{E}_{\mathrm{r}}=\frac{\mathrm{i}}{\mathrm{k}_{0} \epsilon_{\mathbf{r} \mathbf{r}} \mathrm{r}^{2} \sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial \mathrm{~B}}{\partial \theta}\right), \\
\mathrm{E}_{\theta}=-\frac{\mathbf{i}}{\mathrm{k}_{0} \epsilon_{\theta \theta} \mathrm{r}} \frac{\partial^{2} \mathrm{~B}}{\partial \mathbf{r} \partial \theta}-\frac{\epsilon_{\theta \phi}}{\mathrm{r} \epsilon_{\theta \theta}} \frac{\partial \mathrm{A}}{\partial \theta},
\end{gathered}
$$

$$
\begin{gather*}
\mathrm{H}_{\mathrm{r}}=-\frac{\mathrm{i}}{\mathrm{k}_{0} \mathrm{r}^{2} \sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial \mathrm{~A}}{\partial \theta}\right), \\
\mathrm{H}_{\theta}=\frac{\mathrm{i}}{\mathrm{k}_{0} \mathrm{r}} \frac{\partial^{2} \mathrm{~A}}{\partial \mathrm{r} \partial \theta} \tag{4.8}
\end{gather*}
$$

where $k_{0}=\omega / c$ is the wave number of the free space.
Eq. (4.6) can be expressed in terms of A and B

$$
\ell_{r}^{(k)}\left|\begin{array}{l}
B_{k}  \tag{4.9}\\
A_{k}
\end{array}\right|+l_{\theta}^{(k)}\left|\begin{array}{l}
B_{k} \\
A_{k}
\end{array}\right|=\frac{4 \pi}{c} r^{2}\left|\begin{array}{c}
I_{r} \\
0
\end{array}\right|
$$

where

$$
\ell_{\mathbf{r}}^{(\mathrm{k})}=\left|\begin{array}{cc}
\epsilon_{\mathbf{r} \mathbf{r}}^{\mathrm{k}} \mathbf{r}^{2} \frac{\partial}{\partial \mathrm{r}}\left[\frac{1}{\epsilon_{\theta \theta}^{\mathrm{k}}}\right]+\mathrm{k}_{0}^{2} \epsilon_{\mathbf{r} \mathbf{r}}^{\mathrm{k}} \mathrm{r}^{2} & -\mathrm{i} \mathbf{k}_{0} \epsilon_{\mathbf{r} \mathbf{r}}^{\mathrm{k}} \mathbf{r}^{2} \frac{\partial}{\partial \mathbf{r}}\left[\frac{\epsilon_{\theta \phi}^{\mathrm{k}}}{\epsilon_{\theta \theta}^{\mathrm{k}}}\right]  \tag{4.10}\\
\frac{\mathrm{i} \mathrm{k}_{0} \epsilon_{\theta \phi}^{\mathrm{k}}}{\epsilon_{\theta \theta}^{\mathrm{k}}} \mathrm{r}^{2} \frac{\partial}{\partial \mathbf{r}} & \mathbf{r}^{2} \frac{\partial^{2}}{\partial \mathbf{r}^{2}}+\mathrm{k}_{0}^{2} \mathbf{r}^{2} \epsilon_{\phi \phi}^{\mathrm{k}}+\frac{\left(\epsilon_{\theta \phi}^{\mathrm{k}}\right)}{\epsilon_{\theta \theta}^{\mathrm{k}}} .
\end{array}\right|
$$

and

$$
\ell_{\theta}^{(k)}=\left|\begin{array}{cc}
\frac{1}{\sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial}{\partial \theta}\right) & 0  \tag{4.11}\\
0 & \frac{1}{\sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial}{\partial \theta}\right)
\end{array}\right|
$$

The continuity conditions for $\mathrm{E}_{\theta}, \mathrm{H}_{\theta}, \mathrm{E}_{\phi}, \mathrm{H}_{\phi}$ at $\mathrm{r}=\mathrm{r}_{\mathrm{k}}$ become

$$
\begin{gather*}
\Delta^{(k)}\left|\begin{array}{c}
B_{R} \\
A_{R}
\end{array}\right|_{r=r_{R}}=\Delta^{(k+1)}\left|\begin{array}{c}
B_{k+1} \\
A_{k+1}
\end{array}\right|_{r=r_{K}} \\
\left|\begin{array}{c}
B_{R} \\
A_{R}
\end{array}\right|_{r=r_{R}}=\left|\begin{array}{c}
B_{k+1} \\
A_{k+1}
\end{array}\right|_{r=r_{K}} \tag{4.12}
\end{gather*}
$$

where

$$
\Delta(\mathrm{k}) \left\lvert\, \begin{array}{cc}
\frac{1}{\epsilon_{\theta \theta}^{\mathrm{k}}} \frac{\partial}{\partial \mathbf{r}} & -\frac{\mathrm{i} \mathrm{k}_{0} \epsilon_{\theta \phi}^{\mathrm{k}}}{\epsilon_{\theta \theta}^{\mathrm{k}}}  \tag{4.13}\\
0 & \frac{\partial}{\partial \mathbf{r}}
\end{array}\right.
$$

Now decompose (4.7) into the product

$$
\left|\begin{array}{l}
\mathrm{B}  \tag{4.14}\\
\mathrm{~A}
\end{array}\right|=\left|\begin{array}{l}
\mathrm{Y}(\mathrm{r}) \\
\mathrm{Z}(\mathrm{r})
\end{array}\right| \psi(\theta)
$$

Then $\ell_{\mathrm{r}}^{(\mathrm{k})}$ will act only on $\left|\begin{array}{l}\mathrm{Y} \\ \mathrm{Z}\end{array}\right|$ and $\ell_{\theta}$ only on $\theta$. After replacing $\partial / \partial \mathrm{r}$ and $\mathrm{d} / \mathrm{dr}$ one obtains the differential operator $\mathrm{L}_{\mathrm{r}}$ defined by $\ell_{\mathrm{r}}(\mathrm{k})$. Similarly, one obtains $L_{\theta}$ defined by $\ell_{\theta}(k)$. Then one has a single operational equation:

$$
\mathrm{L}_{\mathrm{r}}\left|\begin{array}{l}
\mathrm{B}  \tag{4.15}\\
\mathrm{~A}
\end{array}\right|+\mathrm{L}_{\theta}\left|\begin{array}{l}
\mathrm{B} \\
\mathrm{~A}
\end{array}\right|=\frac{4 \pi}{\mathrm{c}} \mathrm{r}^{2}\left|\begin{array}{l}
\mathrm{I}_{\mathrm{r}} \\
0
\end{array}\right|
$$

To find the solution of (4.15), we will find first the solution of the homogeneous equation

$$
L_{r}\left|\begin{array}{c}
\mathrm{B}  \tag{4.16}\\
\mathrm{~A}
\end{array}\right|+\mathrm{L}_{0}\left|\begin{array}{l}
\mathrm{B} \\
\mathrm{~A}
\end{array}\right|=0
$$

From (4.14), we have

$$
\frac{\mathrm{L}_{\mathrm{r}}\left|\begin{array}{l}
\mathrm{Y}(\mathrm{r}) \\
\mathrm{Z}(\mathrm{r})
\end{array}\right|}{\left|\begin{array}{l}
\mathrm{Y}(\mathrm{r}) \\
\mathrm{Z}(\mathrm{r})
\end{array}\right|}=-\frac{\mathrm{L}_{\theta} \psi(\theta)}{\psi(\theta)}
$$

or

$$
\begin{align*}
& \mathrm{L}_{\mathrm{r}}\left|\begin{array}{c}
\mathrm{Y}(\mathrm{r}) \\
\mathrm{Z}(\mathrm{r})
\end{array}\right|=\mathrm{x}\left|\begin{array}{l}
\mathrm{Y}(\mathrm{r}) \\
\mathrm{Z}(\mathrm{r})
\end{array}\right|  \tag{4.17}\\
& \mathrm{L}_{\theta} \psi(\theta)+\mathrm{x} \psi(\theta)=0 \tag{4.18}
\end{align*}
$$

The solution of (4.17) exists for eigenvalues $\mathrm{x}_{\mathrm{j}}$ 's $(\mathrm{j}=1,2, \ldots)$ corresponding to the eigenfunctions $\left.\begin{aligned} & Y_{j} \\ & Z_{j}\end{aligned} \right\rvert\,$

Equation (4.18) is the usual Legendre equation

$$
\begin{equation*}
\frac{1}{\sin \theta} \frac{\mathrm{~d}}{\mathrm{~d} \theta}\left(\sin \theta \frac{\mathrm{~d} \psi_{\mathrm{j}}}{\mathrm{~d} \theta}\right)+\nu_{\mathrm{j}}\left(\nu_{\mathrm{j}}+1\right) \psi_{\mathrm{j}}=0 \tag{4.19}
\end{equation*}
$$

where

$$
\begin{equation*}
\nu_{\mathrm{j}}\left(\nu_{\mathbf{j}}+1\right)=\mathbf{x}_{\mathrm{j}} \tag{4.20}
\end{equation*}
$$

For the case of a verticle dipole source placed ar $r=b$ we may write the solution of (4.15) as follows,

$$
\left|\begin{array}{l}
B  \tag{4.21}\\
A
\end{array}\right|=-\frac{\pi P}{c b^{2}} \quad \sum_{j=1}^{\infty}\left|\begin{array}{l}
Y_{j}(r) \\
Z_{j}(r)
\end{array}\right| \frac{Y_{j}(b)}{N_{j} \sin \nu_{j} \pi} P \nu_{j}[\cos (\pi-\theta)]
$$

where $P$ is the dipole moment of the source and

$$
\begin{equation*}
N_{j}=\int_{0}^{\infty} \frac{Y_{j}^{2}}{\epsilon_{r r} r^{2}} d r+\int_{0}^{\infty} \frac{Z_{j}^{2}}{r^{2}} d r \tag{4.22}
\end{equation*}
$$

Using the asymptotical approximation of the Legendre function, and expanding $1 / \sin \nu_{j} \pi$, we write (4.21) as follows,

$$
\begin{align*}
\left|\begin{array}{l}
B \\
A
\end{array}\right| & \cong \frac{2 P}{{c b^{2}}^{2}} \sqrt{\frac{\pi}{2 \sin \theta}} \sum_{j=1}^{\infty}\left|\begin{array}{l}
Y_{j}(r) \\
Z_{j}(r)
\end{array}\right| \frac{Y_{j}(b) W_{j}^{1 / 2}}{N_{j}} \\
& \cdot\left\{\sum_{n=0}^{\infty} \exp \left[i(n+1) 2 \pi W_{j}-i\left(W_{j}^{*} \theta+\frac{\pi}{4}\right)\right]\right. \\
& \left.+\exp \left[i 2 \pi n W_{j}+i\left(W_{j}^{*} \theta+\frac{\pi}{4}\right)\right]\right\} \tag{4.23}
\end{align*}
$$

where

$$
W_{j}^{*}=W_{j}+\frac{1}{2}
$$

The separate terms in the braces of (4.23) denote the signals reaching the receiver after passing round the earth $n$ times. Because of the attenuation only the waves with $n=0$ arriving at the receiver along the shortest path should be considered. Thus,

$$
\left|\begin{array}{l}
B  \tag{4.24}\\
A
\end{array}\right| \cong \frac{2 P}{\mathrm{cb}^{2}} \sqrt{\frac{\pi}{2 \sin \theta}} \sum \mathrm{Y}_{\mathrm{j}}(\mathrm{~b})\left|\begin{array}{l}
\mathrm{Y}_{\mathrm{j}}(\mathrm{r}) \\
\mathrm{Z}_{\mathrm{j}}(\mathrm{r})
\end{array}\right| \frac{\mathrm{e}^{\mathrm{i}\left(\mathrm{~W}_{\mathrm{j}}^{*} \theta+\frac{\pi}{4}\right)}}{\mathrm{N}_{\mathrm{j}} \sqrt{W_{j}}}
$$

The field components are

$$
\begin{aligned}
& H_{\phi}=A(\theta) i \sum_{j=1}^{\infty} Y_{j}(r) B_{j}(\theta), \\
& E_{r}=-\frac{A(\theta) i}{k_{0} r} \sum_{j=1}^{\infty} W_{j} Y_{j}(r) B_{j}(\theta),
\end{aligned}
$$

$$
\begin{align*}
& \mathbf{E}_{\theta}=\frac{A(\theta)}{k_{0}} \sum_{j=1}^{\infty} \frac{d Y_{j}(r)}{d r} B_{j}(\theta), \\
& E_{\phi}=A(\theta) i \sum_{j=1}^{\infty} Z_{j}(r) B_{j}(\theta), \\
& H_{r}=\frac{A(\theta) i}{k_{0} r} \sum_{j=1}^{\infty} W_{j} Z_{j}(r) B_{j}(\theta),  \tag{4.25}\\
& H_{\theta}=-\frac{A(\theta)}{k_{0}} \sum_{j=1}^{\infty} \frac{d Z_{j}(r)}{d r} B_{j}(\theta)
\end{align*}
$$

where

$$
\begin{gathered}
A(\theta)=\frac{2 P}{c^{2} r} \sqrt{\frac{\pi}{2 \sin \theta}} \\
B_{j}(\theta)=\frac{Y_{j}(b)}{N_{j}} W_{j}^{1 / 2} \exp \left[i\left(W_{j}^{*} \theta+\frac{\pi}{4}\right)\right]
\end{gathered}
$$

For calculating the field in the atmospheric layer, a further simplification of formula (4.2) is needed. For the details involved, see reference [23].
b. Stratified medium with relief. To account for the diurnal variations of the field, Krasnushkin proposed another model to take into consideration of the slow variations of the dielectric tensor (4.1) along the path of radio waves. He assumed that

$$
\begin{equation*}
\|\epsilon\|=\mathrm{f}(\mathrm{r}, \mu \theta, \mu \phi) \tag{4.26}
\end{equation*}
$$

where $90^{\circ}-\theta$ and $\phi$ are latitude and longitude and $\mu$ is some small parameter. The operational equation now takes the following form

$$
\mathrm{L}_{\mathrm{r}}\left|\begin{array}{l}
\mathrm{B}  \tag{4.27}\\
\mathrm{~A}
\end{array}\right|+\mathrm{L}_{\theta \phi}\left|\begin{array}{l}
\mathrm{B} \\
\mathrm{~A}
\end{array}\right|=\frac{4 \pi}{\mathrm{c}} \mathrm{r}^{2}\left|\begin{array}{c}
\mathrm{I}_{\mathrm{r}} \\
0
\end{array}\right|_{\theta}
$$

This equation is not separable due to dependence of $L_{r}$ on $\theta$ and $\phi$. However, for small $\mu$ one may expand $L_{r}$ in powers of $\mu$. The first approximation yields the following eigenvalue equation:

$$
L_{\mathrm{r}}\left|\begin{array}{l}
\mathrm{Y}  \tag{4.28}\\
\mathrm{Z}
\end{array}\right|=\mathrm{x}(\mu \theta, \mu \phi)\left|\begin{array}{l}
\mathrm{Y} \\
\mathrm{Z}
\end{array}\right|
$$

From whence one obtains the set of eigenvalues $\mathrm{x}_{\mathrm{j}}(\mu \theta, \mu \phi), j=1,2, \ldots \infty$ corresponding to the eigenfunctions

$$
\left|\begin{array}{l}
\mathbf{Y}_{\mathrm{j}}(\mathrm{r} ; \mu \theta \mu \phi)  \tag{4.29}\\
\mathbf{Z}_{\mathrm{j}}(\mathrm{r} ; \mu \theta \mu \phi)
\end{array}\right|
$$

Thus we assume that

$$
\left|\begin{array}{l}
\mathrm{B}  \tag{4.30}\\
\mathrm{~A}
\end{array}\right|=\sum_{\mathrm{j}} \Phi_{\mathrm{j}}(\theta \phi)\left|\begin{array}{l}
\mathrm{Y}_{\mathrm{j}}(\mathrm{r} ; \mu \theta, \mu \phi) \\
\mathrm{Z}_{\mathrm{j}}(\mathrm{r} ; \mu \theta, \mu \phi)
\end{array}\right|
$$

For $\Phi_{\mathrm{j}}(\theta \phi)$ one obtains the equation

$$
\begin{gather*}
\frac{1}{\sin \theta} \frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial \Phi_{\mathrm{j}}}{\partial \theta}\right)+\frac{1}{\sin \theta} \frac{\partial^{2} \Phi_{\mathrm{j}}}{\partial \phi^{2}}+\mathrm{x}_{\mathrm{j}} \Phi \\
=\frac{4 \pi}{\mathrm{cN}_{\mathrm{j}}} \int_{0}^{\infty} \frac{\mathrm{I}_{\mathrm{r}} \mathrm{Y}_{\mathrm{j}}}{\epsilon_{\mathrm{rr}}} \mathrm{dr} \tag{4.31}
\end{gather*}
$$

If one seeks the solution in the form

$$
\begin{equation*}
\Phi_{\mathrm{j}}=\mathrm{U}_{\mathrm{j}}(\theta, \phi) \exp \left[\mathrm{i} \mathrm{~S}_{\mathrm{i}}(\theta, \phi)\right] \tag{4.32}
\end{equation*}
$$

where $U_{j}$ is a slowly varying function of $\theta$ and $\phi$. After substituting (4.32) into (4.31), one finally obtains for $\left|\begin{array}{l}B \\ A\end{array}\right|$ the following expressions:

$$
\begin{align*}
\left|\begin{array}{l}
\mathrm{B} \\
\mathrm{~A}
\end{array}\right|=\frac{2 P}{\mathrm{cb}^{2}} \sqrt{\frac{\pi}{2 \sin \theta}} & \sum Y_{j}(b)\left|\begin{array}{l}
Y_{j}(r) \\
Z_{j}(r)
\end{array}\right| \frac{W_{j}^{-1 / 2}}{\sqrt{\mathrm{~N}_{\mathrm{j}}(0)} \sqrt{\mathrm{N}_{\mathrm{j}}(P)}} \\
& \cdot \exp [\mathrm{i} S(\theta, \phi)] \tag{4.33}
\end{align*}
$$

where normalization factors $N_{j}(0)$ and $N_{j}(P)$ are evaluated at the sender and the receiver points. The path of the wave is defined by the eikonal function $s$, evaluated from the equation

$$
\begin{equation*}
\left(\frac{\partial \mathbf{S}_{\mathrm{j}}}{\partial \theta}\right)^{2}+\frac{1}{\sin ^{2} \theta}\left(\frac{\partial \mathbf{S}_{\mathrm{j}}}{\partial \phi}\right)^{2}=\mathbf{x}_{\mathrm{j}}(\mu \theta, \mu \phi) \tag{4.34}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathrm{S}=\int_{0}^{\mathrm{P}} \sqrt{\mathbf{x}_{\mathrm{j}}} \mathrm{~d} \mathrm{~S} \tag{4.35}
\end{equation*}
$$

If the geodetic line is followed, then

$$
\begin{equation*}
\mathbf{S}=\int_{\theta_{0}}^{\theta_{\mathbf{P}}} \sqrt{\mathbf{x}_{\mathbf{j}}} \mathrm{d} \theta \tag{4.36}
\end{equation*}
$$

After substituting (4.36) into (4.33), one finally obtains the radial component of the electric field

$$
\begin{align*}
E_{r} & =\sqrt{\frac{W}{\sin \theta}} \exp \left[i \frac{\pi}{4}\right]\left\{\sum_{j=0} n_{j}(0) n_{j}(P) \exp \left[i \int_{\theta_{0}}^{\theta_{P}} \sqrt{x_{j}} d \theta\right]\right. \\
& \left.+\sum_{K=0} n_{K}(0) n_{K}(P) \exp \left[i \int_{\theta_{0}}^{\theta_{P}} \sqrt{x_{R}} d \theta\right]\right\} \tag{4.37}
\end{align*}
$$

## SUMMARY

So far we have seen four mathematical models in the study of the VLF wave propagation; namely (1) waveguide mode theory, (2) zonal harmonic series, (3) eigenfunction method, and (4) full wave theory. Of the four, the last stresses more on the computational aspects of the problem whereas the first three give analytic expressions as solutions of the problem. Models one and three are more elegant in form, each is capable of explaining the propagation problem in simple and tractable terms using mostly one or two modes. Model (1) is based on the concept of the classical residue series, whereas model (2) is formulated on the modern theory of the non-self adjoint operators. The stumbling block in the mode theory is the rather formidable transcendental equation whose solution is necessary to account for the various modes needed in evaluating the electromagnetic fields. There is a corresponding transcendental equation in the eigenvalue method, however, Frasnushkin showed that the difficulty can be circumvented by means of what he called the mixed method; whereby he used the data of the short range fields and some of the known parameters to determine the unknown parameters. Then he used the known and the determined parameters to evaluate the long range propagation phenomenon.

In contrast to the mode theory model and the eigenfunction model, the zonal series model relies only on the original harmonic series for its solution. The speed of the electronic computers offsets the objection of the direct summation of the slowly convergent harmonic series. A complete solution is thus made possible leaving only the assumed model for the transmitter and the propagation medium. The advantage of a descriptive discussion that a simple formula affords is lost. Perhaps when all the details of the attributes of the ionosphere are included, graphical solutions based on numerical computation are the only answers. In which case, the full wave theory model as is handled by Budden may be the best approach. Unfortunately, Budden's is a planar model. To satisfy the long range propagation, the earth's curvature must be taken into consideration.

The abundant experimental materials cumulated over the past fifty years in the field of VLF propagation have established many facts that the current theories are still incapable of explaining, such as

1. the diurnal and the seasonal variations of the long distance VLF field;
2. the dependence of both the field intensity and the phase on a heterogenous path of various illuminations;
3. the effect of the solar flare on both the field strength and the phase.

It appears that some work could be done on the formulation of a new model which does not depend only harmonically on time in order to resolve some of the aforementioned observed facts.

## BIBLIOGRAPHY

1. Rayleigh, Lord, "On the Acoustic Shadow of a Sphere," Phil. Trans. Roy. Soc., London A203. 86-110 (1904).
2. Poincare, H., "Sur la diffraction des ondes electriques: a propos d' un article de M. Macdonald," Proc. Roy. Soc., London A72, 42-52 (1904).

Poincare, H., "Sur la diffraction des ondes hertgiennes," Rend. Cir. Mat., Palermo, t. 29, p. 169 (1910).
3. MacDonald, H. M., "The Bending of Electric Waves Round a Conducting Obstacle," Pts. I, II, III, IV, Proc. Roy. Soc., 71, 251 (1903); Proc. Roy. Soc., 72, 59 (1904); Roy. Soc., Phil. Trans. A210, 113 (1911); Proc. Roy. Soc. 90, 50-61 (1914).
4. Nicholson, J. W., "On the Bending of Electric Waves Round a Large Sphere," Pts. I, II, III, IV, Phil. Mag., Ser. 6, 19, 516 (1910); 20, 157 (1910); 21, 62 281 (1911); 19, 757 (1910).
5. Sommerfeld, A., "Über die Ausbreitung der Wellen in der drahtlosen Telegraphie," Ann. Phys. 28, 665 (1909).
6. Love, A. E. H., "The Transmission of Electric Waves Over the Surface of the Earth," Phil. Trans. A215, 105-141 (1915).
7. March, H. W., "Über die Ausbreitung der Wellen der drahtlosen Telegraphie auf der Erdkugel," Ann. Physik. Vierte Folge XXXVII (37), 29-50 (1912).
8. von Rybczynski, W., "Über Ausbreitung der Wellen der drahtlosen Telegraphie auf der Erdkugel," Ann. Physik, Vierte Folge, XLI (41), 191-208 (1913).
9. Smith, F. E., "The Travel of Wireless Waves," Jour. of I.E.E., London, 73, 574 (1933).
10. Watson, G. N., "The Transmission of Electric Waves Round the Earth," Proc. Roy. Soc., London ACXV (95), 546 (1919).

Watson, G. N., "The Diffraction of Electric Waves by the Earth," Proc. Roy. Soc., London AXCV (95), 83 (1918).
11. Rydbeck, O. E. H., "On the Propagation of Waves in an Inhomogeneous Medium," Chalmers Tek Hőgsk, Handl. No. 74 (1948).
12. Van der Pol, Balth, and H. Bremmer, "The Diffraction of Electromagnetic Waves From an Electrical Point Source Round a Finitely Conducting Sphere, with Applications to Radiotelegraphy and the Theory of the Rainbow," Phil. Mag. (7) 24, Pt. I, 141 (July 1937); Pt. II, 825 (Nov. 1937); Supp. J. Sci., 25, $\overline{817}$ (June 1938).
13. Eckersley, T. L., "Long Wave Transmission, Treated by Phase Integral Methods," Proc. Roy. Soc., A, 137, 158 (1932).
14. Bremmer, H., Terrestrial Radio Waves. N. Y. Elsevier Publ. Co., Inc. (1949).
15. Kendrick, G. W., "Radio Transmission Formula," Phys. Rev. 31, 1040 (1928).
16. Weyrich, R., "Zur Theorie der Ausbreitung elektromagnetischer Wellen lüngs der Erdoherfläche," Annal. der Phys. 85, 552 (1928).
17. Brekhovskikh, L. M., "Waves in Layered Media." N. Y. Academic Press (1960).
18. Al'pert, Ya. L., "On the Trajectories of Rays in a Magneto-active Ionized Medium - The Ionosphere." Izvestiya Akad. Nauk, SSSR, Seriya Fizisheskaya, 12, 241 (1948).
19. Fock, V. A., "Electromagnetic Diffraction and Propagation Problems." N.Y. Pergamon Press (1965).
20. Budden, K. G., "The Reflection of Very Low Frequency Radio Waves at the Surface of a Sharply Bounded Ionosphere with Superimposed Magnetic Field." Phil. Mag. 42, 833 (1951).

Budden, K. G., "The Propagation of Very Low Frequency Radio Waves to Great Distances." Phil. Mag. 44, 504 (1953).

Budden, K. G., "The Numerical Solution of Differential Equations Governing Reflection of Long Radio Waves From the Ionosphere." Proc. Roy. Soc., A, 227, 516 (1955).

Budden, K. G., "Radio Waves in the Ionosphere." London, Cambridge University Press (1961).
21. Wait, J. R., "Introduction to the Theory of VLF Propagation," Proc. IRE, 1624-1647 (1962).

Wait, J. R., "Electromagnetic Waves in Stratified Media." N. Y. The Macmillan Co. (1962).

Wait, J. R., "The Mode Theory of VLF Radio Propagation for a Spherical Earth and a Concentric Anisotropic Ionosphere." Cand. Jour. Phys. 41, 299-315 (1963).

Wait, J. R., and K. P. Spies, "Characteristics of the Earth-Ionosphere Waveguide for VLF Radio Waves, ${ }^{\text {" NBS Tech. Note No. } 300 \text { (1964). }}$
22. Johler, J. R. and L. A. Berry, "Propagation of Terrestrial Radio Waves of Long Wavelength - Theory of Zonal Harmonics with Improved Summation Techniques." Jour. Res. 66D, 737-768 (1962).

Johler, J. R., "Radio Wave Reflections at a Continuously Stratified Plasma with Collisions Proportional to Energy and Arbitrary Magnetic Induction." Proc. Int. Conf. Ionosphere, 436-445 (1962).
23. Krasnushkin, P. E., "On the Propagation of Long and Very Long Radio Waves Around the Earth, " Nuovo Cimento, suppl. XXVI, Ser. X, 50-111 (1962).

Krasnushkin, P. E., and N. A. Yablochkin, "Theory of Propagation of Very Long Waves." English version issued by the Computer Center, The Acad. Sci. USSR, Moscow (1963).

# A THEORETICAL STUDY OF ELECTRIC FIELDS IN THE LUNAR ATMOSPHERE 

by
Richard A. Elco*


#### Abstract

Lunar atmospheric electric fields are a result of charge separation and $V \times B$ interactions of the interplanetary magnetic field and the lunar ionosphere which is created by photoionization and charge-exchange ionization of the accreted neutral atmosphere. The interplanetary magnetic field provides a mechanism for transfer of momentum from solar wind protons to slow ions which increases the loss rate of accreted ions from the lunar ionosphere. A collisionless hydromagnetic model of the partially ionized gas flow on the moon yields the ordinary collisionless hydromagnetic fluid equations. Since the solar wind velocity upstream of the moon is supersonic, a shock wave forms on the moon. However, for a steady state shock, the shock must be aligned with the interplanetary magnetic field. The possible geometry of the shock wave is discussed. The maximum increase in the total density and magnetic field across the shock is three times the solar wind values. Therefore, the lunar atmosphere is extremely tenuous, having a total particle density of no greater than $50 / \mathrm{cm}^{3}$. Space charge electric fields in the lunar ionosphere are two orders of magnitude smaller than the $\vec{V} \times \vec{B}$ electric fields which are the order of ten millivolts per meter.


## INTRODUCTION

The sources of the lunar atmospheric and surface electric fields can be grouped into the following three categories. The first, is charge separation in the lunar ionosphere. This mechanism is of primary importance if the ionosphere is in diffusive equilibrium and the charge separation is due to the different scale heights for the ionic and electronic components of the ionosphere. The second source is a dynamic or $\vec{V} \times \vec{B}$ interaction which results from the motion of the lunar ionosphere in the interplanetary magnetic field. The third source is a surface electron sheath which exists to balance the solar wind current incident in the lunar surface with the photoelectron current emitted from the lunar surface (Singer and Walker (1962)). The nature of the lunar electric fields is then directly related to the dynamics of the lunar ionosphere. The primary

[^1]physical processes which maintain the lunar atmosphere and ionosphere are (1) surface recombination of solar wind protons incident on the lunar surface and subsequent thermal emission of neutral hydrogen atoms, and (2) photoionization and charge exchange ionization of the accreted neutral hydrogen atoms. The lunar atmosphere is essentially a collisionless mixture of energetic solar wind protons, energetic electrons, thermal hydrogen ions, thermal or low energy electrons, and thermal hydrogen atoms.

Model lunar atmospheres and ionospheres, in which the effects of the interplanetary magnetic field in the solar wind are neglected, have been developed by Weil and Barasch (1963) and Hinton and Toeusch (1964). These models predict neutral and ion particle densities several orders of magnitude greater than the solar wind density within one to two thousand kilometers of the lunar surface. The loss mechanisms considered in these models were thermal escape and elastic scattering by solar wind protons. However, as was pointed out by Harwit and Hoyle (1962) and Bernstein et. al., (1963), the interplanetary magnetic field can provide a mechanism for transferring momentum from the solar wind protons to the slow ions generated by photoionization and charge exchange and hence increase the loss rate of ions from the lunar ionosphere. Other studies by Beard (1965) (1966) indicate that a collisionless hydromagnetic shock wave is formed by the interaction of the solar wind and the interplanetary magnetic field with cometary atmospheres which are, in some respects, not vastly different from the lunar atmosphere (Gold (1966)). Therefore, the presence of the interplanetary magnetic field in the solar wind can increase the rate at which ions are swept out of the lunar atmosphere and also give rise to the formation of a hydromagnetic shock wave which can drastically affect the lunar atmosphere.

If the solar wind together with the lunar atmosphere is considered as a mixture of interacting particles, where the interaction is due to the ionization of the flow neutral hydrogen atoms and subsequent transfer of momentum from the solar wind protons to the ions by the magnetic field, then a collisionless hydromagnetic model is appropriate for the description of the flow in the vicinity of the moon. In this analysis it is assumed that the moon does not have a significant dipole magnetic field. A multi-component hydromagnetic fluid model is developed and the hydromagnetic jump conditions are presented. Since the solar wind velocity far upstream of the moon is supersonic, it is likely that a stationary shock wave does form on the moon. The consequences of the shock wave are discussed, especially as it affects the particle densities and electric fields in the lunar atmosphere.

## The Collisionless Hydromagnetic Fluid Equations

The collisionless hydromagnetic fluid equations which are given below are derived in the same manner as for the ordinary two fluid (electrons and ions)
collisionless model except that, in this model, momentum exchange by processes other than collisions is taken into account. The exact process by which momentum is exchanged need not be specified, however, some possible processes are wave interactions between slow and energetic ions due to a two stream instability or magneto-ionic wave interactions due to local disturbances of the magnetic field when photo or charge exchange ionization occurs. Since the gas mixture is made up of energetic and slow electrons, ions and neutrals, a three component mixture can be used. The gas is assumed to be electrically neutral.

The continuity equations of each component are - for the ions

$$
\begin{equation*}
\nabla \cdot\left(m_{i} n_{i} \vec{w}_{i}\right)=-\frac{\partial m_{i} n_{i}}{\partial t}+n_{a} v_{i} m_{i} \tag{1}
\end{equation*}
$$

electrons

$$
\begin{equation*}
\nabla \cdot\left(m_{e} n_{e} \vec{w}_{e}\right)=-\frac{\partial m_{e} n_{e}}{\partial t}+n_{a} v_{i} m_{e} \tag{2}
\end{equation*}
$$

and neutrals

$$
\begin{equation*}
\nabla \cdot\left(m_{a} n_{a} \vec{w}_{a}\right)=-\frac{\partial m_{a} n_{a}}{\partial t}-n_{a} v_{i} m_{a} \tag{3}
\end{equation*}
$$

The $v_{i}$ is the photoionization rate. No terms involving charge exchange ionization are involved because the net creation rate of neutrals or ions by charge exchange is zero. The velocity $\overrightarrow{\mathrm{w}}_{\mathrm{K}}$ is the mass center velocity of the $\mathrm{K}^{\text {th }}$ component, i.e., $n_{e} \vec{w}_{e}=n_{e}^{s} \vec{V}_{e}^{s}+n_{e}^{*} \vec{V}_{e}^{*}$ where the $s$ and asterisk refer respectively to the slow and energetic electrons and $\overrightarrow{\mathrm{V}}$ is the particle velocity. A combination of Equations (1) through (3) yields the continuity equation for the fluid as a whole,

$$
\begin{equation*}
\nabla \cdot(\rho \overrightarrow{\mathbf{u}})=-\frac{\partial \rho}{\partial \mathbf{t}} \tag{4}
\end{equation*}
$$

in which

$$
\begin{equation*}
\rho=m_{i} n_{i}+m_{e} n_{e}+m_{a} n_{a} \tag{5}
\end{equation*}
$$

and

$$
\begin{equation*}
\rho \overrightarrow{\mathrm{u}}=\mathrm{m}_{\mathrm{i}} \mathrm{n}_{\mathrm{i}} \overrightarrow{\mathrm{w}}_{\mathrm{i}}+\mathrm{m}_{\mathrm{e}} \mathrm{n}_{\mathrm{e}} \overrightarrow{\mathrm{w}}_{\mathrm{e}}+\mathrm{m}_{\mathrm{a}} \mathrm{n}_{\mathrm{a}} \overrightarrow{\mathrm{w}}_{\mathrm{a}} \tag{6}
\end{equation*}
$$

The equation of motion for the fluid is then

$$
\begin{equation*}
\rho\left(\frac{\partial \overrightarrow{\mathrm{u}}}{\partial \mathrm{t}}+(\overrightarrow{\mathrm{u}} \cdot \nabla) \overrightarrow{\mathrm{u}}\right)=-\nabla \mathrm{P}-\frac{\nabla \mathrm{B}^{2}}{2 \mu_{0}}+\frac{(\overrightarrow{\mathrm{B}} \cdot \nabla) \overrightarrow{\mathrm{B}}}{\mu_{0}} \tag{7}
\end{equation*}
$$

where $P$ is the total pressure, i.e.,

$$
\begin{equation*}
P=P_{i}+P_{e}+P_{a} \tag{8}
\end{equation*}
$$

This form of the equation of motion which includes the neutrals is obtained for the following reason. During the charge exchange process a slow neutral atom exchanges an electron with an energetic proton which then becomes an energetic neutral, therefore the result is a net increase in momentum of the neutral component and a net decrease in momentum of the ion component (Biermann (1966)).

The generalized Ohms Law for the fluid (Holt and Haskell (1965)) is

$$
\begin{equation*}
\frac{\partial \vec{J}}{\partial t}-\frac{e}{m_{e}} \nabla P_{e}-\frac{n_{e} e^{2}}{m_{e}}\left(\vec{E}+\overrightarrow{\mathrm{u}} \times \vec{B}-\frac{1}{n_{e} e} \vec{J} \times \vec{B}\right)=\frac{e}{m_{i}} \vec{A}_{i}-\frac{e}{m_{e}} \vec{A}_{e} \tag{9}
\end{equation*}
$$

where $\vec{J}$ is the current density, e the magnitude of the electron charge and $\vec{A}_{i}$ and $\vec{A}_{e}$ are the rates of change of momentum of the ions and electrons due to interaction with neutrals. The rate of change of the ion momentum $\overrightarrow{\mathrm{A}}_{i}$ is negative. It is also most likely that the electrons also experience a net loss of momentum during each charge exchange which is the order of $m_{e} / m_{i} \vec{A}_{i}$. Therefore it is assumed that the R.H.S. of Equation (9) is zero which implies that the momentum exchanged in this way does not contribute to a current in the fluid.

The energy equation for the fluid is

$$
\begin{equation*}
\rho \frac{\mathrm{D}}{\mathrm{Dt}}\left(\frac{\mathrm{P}}{\rho}+\epsilon+\frac{\mathrm{u}^{2}}{2}\right)=\overrightarrow{\mathrm{E}} \cdot \overrightarrow{\mathrm{~J}} \tag{10}
\end{equation*}
$$

in which $D / D t=\partial / \partial t+\vec{u} \cdot \nabla$ and $\epsilon$ is the specified internal energy. If it is assumed, following Chandrasekhar (1960), that only the energy of charged particle motion transverse to the magnetic field contributes to the internal energy of the ionized portion of the fluid, then the specific internal energy $\epsilon$ is

$$
\begin{equation*}
\epsilon=\frac{n_{e} W_{\perp}^{e}+n_{i} W_{\perp}^{i}+\frac{3}{2} n_{a} k T_{a}}{\rho} \tag{11}
\end{equation*}
$$

where $W_{\perp}$ is the energy due to motion transverse to the magnetic field in a frame of reference stationary with respect to the mass center fluid velocity $\vec{u}, T_{a}$ is the neutral gas temperature and $k$ the Boltzmann constant. It is convenient to rewrite (11) as

$$
\begin{equation*}
\epsilon=\bar{W}+\frac{1}{2} \frac{\mathrm{n}_{\mathrm{a}} \mathrm{k} \mathrm{~T}_{\mathrm{a}}}{\rho} \tag{12}
\end{equation*}
$$

where

$$
\begin{equation*}
\bar{W}=\frac{n_{e} W_{\perp}^{e}+n_{i} W_{1}^{i}+n_{a} k T_{a}}{\rho} \tag{13}
\end{equation*}
$$

The partial pressures of the ion, electron and neutral gases are

$$
\begin{equation*}
P_{i}=n_{i} W_{1}^{i}, \quad P_{e}=n_{e} W_{1}^{i} \text { and } P_{a}=n_{a} k T_{a} \tag{14}
\end{equation*}
$$

Then, from Equations (8) and (13) the total pressure is

$$
\begin{equation*}
\mathbf{P}=\rho \overline{\mathbf{W}} \tag{15}
\end{equation*}
$$

The energy equation can then be written as

$$
\begin{equation*}
\rho \frac{\mathrm{D}}{\mathrm{Dt}}\left(2 \overline{\mathrm{~W}}+\frac{1}{2} \frac{\mathrm{n}_{\mathrm{a}} \mathrm{kT}_{\mathrm{a}}}{\rho}+\frac{\mathrm{u}^{2}}{2}\right)=\overrightarrow{\mathrm{E}} \cdot \overrightarrow{\mathrm{~J}} \tag{16}
\end{equation*}
$$

which indicates that the effective ratio of specific heats is somewhere between 2 and $5 / 3$ depending on the value of $n_{a}^{k} T_{a}$ with respect to the values of $n W_{\perp}$ for the electrons and ions.

## The Lunar Hydromagnetic Shock Wave

Studies of collisionless hydromagnetic shock waves have been made using nonlinear wave mechanisms and various forms of the hydromagnetic approximations of the plasma equations (Beard (1965), Sen and Spero (1967)), to describe the shock structure and the initial and final plasma states. The hydromagnetic approach is used in the following analysis. While the hydromagnetic model does not provide information about the details of the shock wave structure, it does yield, in a straight forward way, the necessary conditions for shock wave formation and the jump equations across the shock wave.

Using the continuity equation, (4), the equation of motion (7), the generalized Ohms Law (9) and the field equations $\nabla \times \vec{E}=-\partial \vec{B} / \partial t$ and $\nabla \times \vec{H}=\vec{J}$ to determine allowed discontinuities in the fluid flow yields the jump conditions derived by Sen and Spero (1967) which are

$$
\begin{gather*}
{\left[\rho \mathrm{u}_{\mathrm{n}}\right]=0}  \tag{17}\\
{\left[\rho \mathrm{u}_{\mathrm{n}} \overrightarrow{\mathrm{u}}_{\mathrm{t}}-\frac{\mathrm{B}_{\mathrm{n}} \overrightarrow{\mathrm{~B}}_{\mathrm{t}}}{\mu_{0}}\right]=0}  \tag{18}\\
{\left[\rho \mathrm{u}_{\mathrm{n}}^{2}+\mathrm{P}+\frac{\mathrm{B}_{\mathrm{t}}^{2}}{2 \mu_{0}}\right]=0}  \tag{19}\\
{\left[\mathrm{P}_{\mathrm{e}}+\frac{\mathrm{B}_{\mathrm{t}}^{2}}{2 \mu_{0}}\right]=0}  \tag{20}\\
{\left[\frac{m_{e}}{\mu_{0} \mathrm{e}} \frac{\partial \vec{B}_{\mathrm{t}}}{\partial \mathrm{t}}+\frac{\mathrm{B}_{\mathrm{n}} \overrightarrow{\mathrm{~B}}_{\mathrm{t}}}{\mu_{0}}\right]=0} \tag{21}
\end{gather*}
$$

and

$$
\begin{equation*}
\left[\vec{E}_{t}\right]=0, \quad\left[B_{n}\right]=0 \tag{22}
\end{equation*}
$$

where the subscript $n$ refers to the component of a vector variable normal to the shock wave while $t$ refers to those components which are in the plane of the shock wave. The symbol [ ] is defined as

$$
[\mathrm{f}(\mathrm{n})]=\mathrm{f}\left(0_{+}\right)-\mathrm{f}\left(0_{\ldots}\right)
$$

where the plane of the shock wave is located at $n=0$.
As a consequence of Equations (21) and (22), Sen and Spero (1967) show that no steady-state hydromagnetic shock wave can exist if $\mathrm{B}_{\mathrm{n}} \neq 0$. This implies that the surface of a steady-state hydromagnetic shock wave must be aligned with the magnetic field. Therefore, a shock wave formed by the flow of plasma around a sphere in a magnetic field perpendicular to the free stream velocity will be semicylindrical instead of hemispheric in shape where the axis of the semi-cylinder is aligned with the magnetic field vector. This conclusion agrees qualitatively
with the results obtained by Ludford and Singh (1963) in their study of the crossed field flow of a perfectly conducting, incompressible fluid over a non-conducting sphere. However, it is obvious that the semi-cylindrical shock surface cannot extend indefinitely on each side of the sphere, therefore, non-steady shock waves must exist on each side of the steady shock wave as is illustrated in Figure 1.


Figure 1-Illustration of combination steady state and non-steady state hydromagnetic shock wave on the moon due to interaction of lunar ionosphere with solar wind and interplanetary magnetic field.

In view of the complexity and uncertainty implied by the above discussion of the shock wave geometry, the following assumptions are made in order to obtain an estimate of the post shock conditions which are relevant to the lunar electric fields. First, the interplanetary magnetic field is assumed to be perpendicular to the free stream velocity of the solar wind and secondly, the steadystate shock wave surface is assumed to be a semi-cylinder aligned with the magnetic field as is shown in Figure 2. The local shock coordinates are the cartesian system $x, y$ and $z$ and the subscripts 1 and 2 refer to the pre-shock and post-shock states respectively. The magnetic field normal to the shock is then zero, i.e., $B_{x}=0$, and also $B_{y}=0$.


Figure 2-Simplified geometry used in analysis of lunar hydromagnetic shock wave and post shock atmosphere.

The jump equations (Eqs. (17) to (21)) can now be simplified to

$$
\begin{gather*}
{\left[\rho \mathbf{u}_{\mathrm{x}}\right]=0}  \tag{23}\\
{\left[\mathrm{u}_{\mathrm{y}}\right]=0}  \tag{24}\\
{\left[\mathbf{u}_{\mathrm{z}}\right]=0}  \tag{25}\\
{\left[\rho \mathbf{u}_{\mathrm{x}}^{2}+\mathbf{P}+\frac{\mathbf{B}_{z}^{2}}{2 \mu_{0}}\right]=0} \tag{26}
\end{gather*}
$$

and

$$
\begin{equation*}
\left[\mathbf{P}_{\mathrm{e}}+\frac{\mathbf{B}_{z}^{2}}{2 \mu_{0}}\right]=0 \tag{27}
\end{equation*}
$$

Taking the curl of the Ohms Law (Eq. 9) and solving for the jump equation yields

$$
\begin{equation*}
\left[B_{z} u_{x}\right]=0 \tag{28}
\end{equation*}
$$

Also the energy equation yields the following jump equation

$$
\begin{equation*}
\left[2 \overline{\mathrm{~W}}+\frac{1}{2} \frac{\mathrm{n}_{\mathrm{a}} \mathrm{kT} \mathrm{~T}_{\mathrm{a}}}{\rho}+\frac{\mathrm{u}_{\mathrm{x}}^{2}}{2}+\frac{\mathrm{B}_{\mathrm{z}}^{2}}{\rho \mu_{0}}\right]=0 \tag{29}
\end{equation*}
$$

where it is assumed that $\nabla \mathrm{P}_{\mathrm{e}} \times \nabla \mathrm{B}_{\mathrm{z}}=0$. Assuming that $2 \overline{\mathrm{~W}} \gg 1 / 2 \mathrm{n}_{\mathrm{a}} \mathrm{kT} \mathrm{T}_{\mathrm{a} / \rho}$, i.e., the ratio of specific heats is two, and that the solar wind Mach number is very large compared to unity, the following limiting solutions are obtained:

$$
\begin{aligned}
\mathrm{B}_{\mathrm{z} 2} & =3 \mathrm{~B}_{z 1} \\
\mathrm{u}_{\mathrm{x} 2} & =\frac{\mathrm{u}_{\mathrm{x} 1}}{3} \\
\rho_{2} & =3 \rho_{1}
\end{aligned}
$$

and

$$
\begin{equation*}
u_{y 2}=u_{y 1} \tag{30a,b,c,d}
\end{equation*}
$$

The hydromagnetic shock wave standoff distance $\Delta$, is approximately given by the ordinary fluid shock standoff distance which is

$$
\begin{equation*}
\Delta \approx \frac{\mathrm{R}_{0}\left(\rho_{1} / \rho_{2}\right)}{1+\sqrt{2 \rho_{1} / \rho_{2}}} \tag{31}
\end{equation*}
$$

where $R_{0}$ is the radius of the moon. The maximum value of the stagnation point density calculated assuming $\mathrm{P} \propto \rho^{2}$ is approximately given by

$$
\begin{equation*}
\rho_{\mathrm{s}} \approx \frac{3}{2} \rho_{2} \approx 4.5 \rho_{1} \tag{32}
\end{equation*}
$$

This result indicates that radial variations near the stagnation point are quite small. However, it should be pointed out that, both the standoff distance and stagnation point density, as calculated, use the assumption that the net gas flux at the surface of the sphere is zero. If the sphere is losing mass, the net surface flux is not zero and both the surface density and standoff distance are increased. The ratio of the mass flux through the shock to the net surface flux is orders of magnitude less than unity in the case of the moon (Hinton and Taeusch (1964)) however, for example, in the case of comets, this ratio can be much greater than one.

The individual particle densities and flux densities behind the shock wave can now be determined. Because of the formation of the shock wave, no slow neutral atoms or slow hydrogen ions can be present upstream of the shock, otherwise the free stream flow would be perturbed ahead of the shock by ionization of the neutral atoms and a stationary shock would not exist. Therefore the pre-shock flux is composed of only solar wind protons and electrons. Then the particle density behind the shock is

$$
\begin{equation*}
3 \mathrm{n}_{\mathrm{i} 1}^{*}=\mathrm{n}_{\mathrm{i} 2}^{*}+\mathrm{n}_{\mathrm{i} 2}^{\mathrm{s}}+\mathrm{n}_{\mathrm{a} 2}^{*}+\mathrm{n}_{\mathrm{a} 2}^{\mathrm{s}} \tag{33}
\end{equation*}
$$

assuming that $n_{e} \approx n_{i}, m_{i}=m_{a}$ and neglecting terms the order of $m_{e} / m_{i}$. Since no charge exchange occurs ahead of the shock or in the shock, the energetic neutral density behind the shock is zero, i.e., $\mathrm{n}_{\mathrm{a} 2}^{*}=0$. Then (33) in terms of the total ion density is

$$
\begin{equation*}
3 n_{\mathrm{i} 1}^{*}=\mathrm{n}_{\mathrm{i} 2}+\mathrm{n}_{\mathrm{a} 2}^{\mathrm{s}} \tag{34}
\end{equation*}
$$

The mass flux density through the shock is conserved, therefore

$$
\begin{equation*}
\Phi_{1}=n_{\mathrm{i} 1}^{*} \mathrm{~V}_{\mathrm{i} 1}^{*}=\mathrm{n}_{\mathrm{i} 2}^{*} \mathrm{~V}_{\mathrm{i} 2}^{*}+\mathrm{n}_{\mathrm{i} 2}^{\mathrm{s}} \mathrm{~V}_{\mathrm{i} 2}^{\mathrm{s}}+\mathrm{n}_{\mathrm{a} 2}^{\mathrm{s}} \mathrm{~V}_{\mathrm{a} 2}^{\mathrm{s}} \tag{35}
\end{equation*}
$$

where, again, terms the order of $m_{e} / m_{i}$ are neglected $V$ is the particle velocity normal to the shock and $\Phi_{i}$ is the solar wind flux density.

At the surface of the moon the total mass flux is zero or

$$
\begin{equation*}
\Phi_{\mathrm{i} 0}^{\mathrm{S}}=\mathrm{n}_{\mathrm{a} 0}^{\mathrm{s}} \mathrm{U}_{\mathrm{a} 0}^{\mathrm{s}}=-\mathrm{n}_{\mathrm{i} 0} \mathrm{U}_{\mathrm{i} 0}-\mathrm{n}_{\mathrm{a} 0}^{*} \mathrm{U}_{\mathrm{a} 0}^{*} \tag{36}
\end{equation*}
$$

where $\Phi_{i 0}^{s}$ is the slow ion flux density and $U$ is the radial velocity in the plane $\phi=\pi / 2$ (See Figure 2). The velocity $U_{a 0}^{s}$ is the thermal velocity of the hydrogen atoms emitted from the lunar surface. Since (1) $U_{a 0}^{s}$, for a temperature of $400^{\circ} \mathrm{K}$, is much smaller than the solar wind velocity and, (2) the slow ion density at the surface is zero, the following relation is obtained:

$$
\begin{equation*}
\frac{\mathrm{n}_{\mathrm{a} 0}^{\mathrm{s}}}{\mathrm{n}_{\mathrm{i} 0}^{*}}=\left|-\frac{\mathrm{U}_{\mathrm{i} 0}^{*}}{\mathrm{U}_{\mathrm{a} 0}^{\mathrm{s}}}-\frac{\mathrm{n}_{\mathrm{a} 0}^{*} \mathrm{U}_{\mathrm{a} 0}^{*}}{\mathrm{n}_{\mathrm{i} 0}^{*} 0 \mathrm{U}_{\mathrm{a} 0}^{\mathrm{s}}}\right| \gg 1 \tag{37}
\end{equation*}
$$

The total particle density at the lunar surface $n_{0}$ is then given by

$$
\begin{equation*}
\mathrm{n}_{0} \approx \mathrm{n}_{\mathrm{a} 0}^{\mathrm{s}} \tag{38}
\end{equation*}
$$

which is the boundary condition on $n_{a 0}^{s}$ at the stagnation point since $n_{0}$ is known from Equation (32). Therefore, the stagnation point slow neutral particle flux density is orders of magnitude less than the solar wind flux and, from Equation (35), the total ion flux density behind the shock is approximately equal to the solar wind flux density or

$$
\begin{equation*}
\mathrm{n}_{\mathrm{i} 2} \mathrm{~V}_{\mathrm{i} 22} \approx \Phi_{1} \tag{39}
\end{equation*}
$$

The $\mathrm{n}_{\mathrm{i} 2}$ and $\mathrm{V}_{\mathrm{i} 2}$ are

$$
n_{\mathrm{i} 2}=3 n_{\mathrm{i} 1}^{*}-n_{\mathrm{a} 2}^{\mathrm{s}}
$$

and

$$
\begin{equation*}
V_{i 2}=\frac{\Phi_{1}}{3 n_{i 11}-n_{a 2}^{s}} \tag{40a,b}
\end{equation*}
$$

where the only unknown is $\mathrm{n}_{\mathrm{a} 2}^{\mathrm{s}}$. However, since $\mathrm{n}_{\mathrm{a} 0}^{\mathrm{s}}$ is known, it is possible to obtain $n_{a 2}^{s}$ from a solution of the equation of motion and the continuity equation for the slow ions.

A detailed solution for the motion of slow ions behind the shock is beyond the scope of this paper, however, an approximate solution can easily be obtained if it is assumed that the slow neutrals are emitted from the surface at the thermal velocity in only the radial direction. Further, since collisions do not occur, the radial velocity is assumed to be relatively constant, i.e., $\mathrm{U}_{\mathrm{a}}^{\mathrm{s}} \approx \mathrm{U}_{\mathrm{a} 0}^{\mathrm{s}}$. The continuity equation for the slow neutrals then becomes

$$
\begin{equation*}
\frac{1}{r^{2}} \frac{\partial}{\partial r}\left(r^{2} n_{a}^{s}\right)=-\frac{n_{a}^{s}}{U_{a 0}^{s}}\left(v_{i}+\Phi_{i}^{*} \sigma_{c x}\right) \tag{41}
\end{equation*}
$$

where $\Phi_{i}^{*}$ is the energetic proton flux density which is assumed to be constant, $\sigma_{c x}$ is the charge exchange cross section and $v_{i}$ is the photoionization rate. The solution to (41) is

$$
\begin{equation*}
n_{a}^{s}=\frac{n_{0}}{\left(r^{2} / R_{0}^{2}\right)} \exp \cdot\left\{\frac{\left(v_{i}+\Phi_{i}^{*} \sigma_{c x}\right) R_{0}}{U_{a}^{s}}\left(1-r / R_{0}\right)\right\} \tag{42}
\end{equation*}
$$

and the ion particle density for $\phi=\pi / 2$ and $\theta=0$ is given by

$$
\begin{equation*}
n_{i} \approx n_{0}-n_{a}^{s} \tag{43}
\end{equation*}
$$

if it is assumed that $\rho \approx \rho_{2}$ behind the shock. The neutral hydrogen and proton densities as a function of radius are shown in Figure 3 for $v_{i}=5 \times 10^{-7} / \mathrm{sec}$, $\sigma_{c x}=2 \times 10^{-15} \mathrm{~cm}^{2}, \Phi_{\mathrm{i}}^{*}=10^{9} / \mathrm{cm}^{2} \mathrm{sec}, \mathrm{n}_{0}=3 \times 10 / \mathrm{cm}^{3}$ and $\mathrm{U}_{\mathrm{a} 0}^{\mathrm{s}}=4 \times 10^{5}$ $\mathrm{cm} / \mathrm{sec}$ (Hinton and Taeusch (1964)). The $\theta$ variation of $n_{\mathrm{a}}^{\mathrm{s}}$, to a first approximation goes as the $\cos \theta$. The normal ion velocity and $n_{a 2}^{s}$ along the shock (Equation (40b)) as functions of $\theta$ are shown in Figure 4.

## Electric Fields in the Lunar Atmosphere

An estimate of the electric field in the vicinity of the stagnation point can be obtained by evaluating the generalized Ohms Law at $\theta=0$ assuming $\rho \approx \rho_{2}$ behind the shock. This yields

$$
\begin{equation*}
\overrightarrow{\mathrm{E}}=-\overrightarrow{\mathrm{u}} \times \overrightarrow{\mathrm{B}}-\frac{1}{\mathrm{n}_{\mathrm{e}} \mathrm{e}} \nabla \mathrm{P}_{\mathrm{e}} \tag{44}
\end{equation*}
$$


-. Figure 3-The thermal hydrogen atom and ion number densities behind the shock wave as functions of radius for $\phi=\pi / 2$ and $\theta=0$ and $n_{i 1}=10 / \mathrm{cm}^{3}$.


Figure 4-lon velocity and thermal hydrogen atom density immediately behind the shock wave as a function of $\theta$, for $\phi=\pi / 2 n_{\mathrm{i} 1}=10 / \mathrm{cm}^{3} \phi_{1}=10^{9} / \mathrm{cm}^{2}$ sec. Shock wave position is $R_{s} \approx\left(R_{0}+\Delta\right) / \cos \theta$.

The electron partial pressure $\mathrm{p}_{\mathrm{e}}$, from Equation (14), can be expressed as

$$
\begin{equation*}
P_{e}=n_{e}^{2}\left(\frac{B_{z}}{n_{e 2}}\right)\left(\frac{W_{12}}{B_{z 2}}\right) \tag{45}
\end{equation*}
$$

assuming that $B / n_{e}$ and $W_{\perp}^{e} / B$ are constants (Chandrasekhar (1960)). Then, since $n_{\mathrm{e}} \approx \mathrm{n}_{\mathrm{i}}$ and

$$
W_{\perp 2}^{\mathrm{e}}=\frac{1}{\mathrm{n}_{\mathrm{e} 2}}\left(\mathrm{P}_{\mathrm{e} 1}-\frac{4 \mathrm{~B}_{1}^{2}}{\mu_{0}}\right)
$$

form Equation (27), the electric field is

$$
\begin{equation*}
\overrightarrow{\mathrm{E}}=-\overrightarrow{\mathrm{u}} \times \overrightarrow{\mathrm{B}}-\frac{\mathrm{B}_{1}^{2}\left(\frac{\mu_{0} \mathrm{P}_{\mathrm{e} 1}}{\mathrm{~B}_{1}^{2}}-4\right)}{\mu_{0} \mathrm{e}\left(\mathrm{n}_{\mathrm{i} 2}\right)^{2}} \nabla \mathrm{n}_{\mathrm{i}} \tag{46}
\end{equation*}
$$

The mass center fluid velocity $\overrightarrow{\mathrm{u}}$, for $\phi=\pi / 2$ and $\theta$ equal to zero, has only a radial component which is approximately given by

$$
\begin{equation*}
u_{r} \approx 3 u_{x 2}\left(1-R_{0}^{2} / r^{2}\right) \tag{47}
\end{equation*}
$$

The radial and tangential values of the electric field are

$$
\mathrm{E}_{\phi} \approx 3 \mathrm{u}_{\mathrm{x} 1} \mathrm{~B}_{1}\left(1-\mathrm{R}_{0}^{2} / \mathrm{r}^{2}\right)
$$

and

$$
\begin{equation*}
\mathrm{E}_{\mathrm{r}} \approx-\frac{2 \mathrm{n}_{0} \mathrm{~B}_{1}^{2}\left(\frac{\mu_{0} \mathrm{P}_{\mathrm{e} 1}}{\mathrm{~B}_{1}^{2}}-4\right)}{\mu_{0} \mathrm{e}\left(\mathrm{n}_{\mathrm{i} 2}\right)^{2} \mathrm{R}_{0}}\left(\frac{\mathrm{R}_{0}}{\mathrm{r}}\right)^{3} \tag{48a,b}
\end{equation*}
$$

at $\theta=0$ in the $\phi=\pi / 2$ plane. The orders of magnitude of $E_{\phi}$ and $E_{r}$ are

$$
\mathrm{E}_{\phi}=0\left(\mathrm{u}_{\mathrm{x} 1} \mathrm{~B}_{1}\right) \approx 10^{-2} \text { volts/meter }
$$

and

$$
E_{r}=0\left(\frac{n_{0} B_{1}^{2}}{\mu_{0} \mathrm{e}\left(\mathrm{n}_{\mathrm{i} 2}\right)^{2} \mathrm{R}_{0}}\right) \approx 10^{-5} \text { to } 10^{-4} \text { volt s/meter }
$$

Plots of $\mathrm{E}_{\mathrm{r}}$ and $\mathrm{E}_{\phi}$ are presented in Figure 5.
The influence of the lunar gravitational field on the motion of the ionized lunar atmosphere and on the atmospheric electric fields is extremely slight. Inclusion of gravitational forces adds the following term to Equation (44)

$$
\vec{E}_{\text {grav. }}=-\frac{m_{e}}{e}\left(\frac{n_{i}-n_{e}}{n_{e}}\right) \vec{g}
$$

which has a maximum possible value of $10^{-11}$ volts/meter at the lunar surface. Elsewhere in the atmosphere the space charge is quite small, therefore, $\overrightarrow{\mathrm{E}}_{\mathrm{grav}}$. is essentially zero.


Figure 5-The radial and $\phi$ components of the lunar atmospheric electric fields as a function of radius for $\phi=\pi / 2$ and $\theta=0$.

At the lunar surface, photo emission and secondary emission gives rise to an outward flux of electrons, most of which must be reflected back to the surface by an electric field. This radial electric field is provided by the net positive surface charge due to the decrease of surface electrons which are involved in photo emission. Next to the surface is an electron sheath which, as calculated by Singer and Walker (1962), is the order of several centimeters thick. To maintain the proper boundary conditions for the atmospheric electric field, which is radially inward, the sheath contains a surface density of electrons in excess of those needed to balance the positive surface charge. The electric fields within ten centimeters of the surface are the order of 10 volts/meter and decrease very rapidly to the value of $10^{-4}$ volts/meter a few meters from the surface.

## Conclusions

It is very probable that a shock wave of some kind is present on the moon and measurements by Ness (1966) have tentatively verified the existence of a lunar wake which might have been generated by a shock wave. The semicylindrical geometry of the shock, which has been proposed, results from the
condition that the magnetic field normal to shock must be zero for a steady-state shock wave. The steady-state semi-cylindrical shock wave cannot, of course, extend indefinitely along the magnetic field lines therefore, a non-steady shock structure must exist on each side of the steady shock.

The application of the hydromagnetic jump conditions across the shock yields an increase in the total density and magnetic field of approximately three times the solar wind values. The stand-off distance of the shock is the order of 0.2 times the lunar radius. The collisionless hydromagnetic model does not provide a description of the internal structure of the shock. However, it can be concluded that some form of enhanced ionization process, such as the impact ionization suggested by Beard (1966), operates in the shock because no neutrals can be present upstream of the shock. The only way in which neutrals arriving at the shock from the lunar surface can be reflected by the shock is by ionization and acceleration back into the flow behind the shock. The thickness of the shock is probably related to the ionization processes in the shock.

The primary conclusions which are reached from this analysis are: (1) that most of the ions on the lunar atmosphere are swept along with the solar wind and interplanetary magnetic field, therefore, the lunar atmosphere is extremely tenuous, having a total particle density of no greater than $50 / \mathrm{cm}^{3}$, and (2) the electric field in the lunar atmosphere has a tangential component proportional to $\overrightarrow{\mathrm{u}} \times \overrightarrow{\mathrm{B}}$ which is the order of $10^{-2}$ volts/meter and a radial component, proportional to the gradient of the electric partial pressure, which is the order of $10^{-4}$ volts/ meter.

## ACKNOWLEDGMENT

The research described in this paper was conducted during the author's tenure as a ASEE-NASA summer Faculty Fellow at Goddard Space Flight Center. The topic of electric fields in the lunar atmosphere was suggested by Drs. J. P. Heppner and T. Aggson of the Goddard Space Flight Center. The author also wishes to acknowledge the discussions with Dr. W. F. Hughes of Carnegie-Mellon University on the shock wave geometry.

## REFERENCES

1. Beard, D. B., "Magnetohydrodynamic Shocks of Geophysical Interest," J. Geophys. Res., 70, 4181-4189, 1965.
2. Beard, D. B., "The Theory of Type I Comet Tails," Planet. Space Sci. 14, 303-311, 1966.
3. Bernstein, W., et. al., "The Lunar Atmosphere and the Solar Wind," Icarus, 2, 233-248, 1963.
4. Biermann, L., "The Magnetohydrodynamic Aspects of the Interaction of the Solar Wind with Comets," in The Proceedings of the Jet Propulsion Laboratory Conference on the "Solar Wind," edited by R. J. Mackin, Jr., and M. Neugebauer, Symposium Publications Division of Pergamon Press, New York, 365-371, 1966.
5. Chandrasekhar, S., "Plasma Physics," The University of Chicago Press, 78, 1960.
6. Gold, T., "The Magnetosphere of the Moon," in The Proceedings of the Jet Propulsion Laboratory Conference on the "Solar Wind," edited by R. J. Mackin, Jr., and M. Neugebauer, Symposium Publication Division of Pergamon Press, New York, 381-391, 1966.
7. Harwit, M., and F. Hoyle, "Plasma Dynamics in Comets." II. Influence of Magnetic Fields, Astrophys. J., 135, 875-882, 1962.
8. Hinton, F. L., and D. R. Taeusch, "Variation of the Lunar Atmosphere with the Strength of the Solar Wind," J. Geophys. Res., 69, 1341-1347, 1964.
9. Holt, E. H., and R. E. Haskell, "Foundations of Plasma Dynamics," The Macmillan Co., New York, 177, 1965.
10. Ludford, G. S. S., and M. P. Singh, "The Motion of a Non-Conducting Sphere Through a Conducting Fluid in a Magnetic Cross-Field," Proc. Camb. Phil. Soc., 59, 615, 1963.
11. Ness, N. F., "A Probable Observation of the Wake of the Moon," in The Proceedings of the Jet Propulsion Laboratory Conference on the "Solar Wind," edited by R. J. Mackin, Jr., and M. Neugebauer, Symposium Publication Division of Pergamon Press, New York, 393-400, 1966.
12. Sen, A. K., and D. M. Spero, "Jump Conditions for a Two-Fluid Plasma," Phys. Fluids, 10, 1597-1599, 1967.
13. Singer, S. F., and E. H. Walker, "Photoelectric Screening of Bodies in Interplanetary Space," Icarus, 1, 7-12, 1962.
14. Weil, H., and M. L. Barasch, "A Theoretical Lunar Atmosphere," Icarus, 1, 346-356, 1963.

# A STUDY OF THE FREQUENCY AND STABILITY CHARACTERISTICS OF A $\mathrm{CO}_{2}$ LASER 

by
B. J. Graham*

## INTRODUCTION

For deep space communication applications the $\mathrm{CO}_{2}$ laser must be stable with respect to power and frequency output. A study of the output of existing lasers has been made to determine the design parameters of $\mathrm{CO}_{2}$ lasers which will give a stable output.

## LASER DESIGNS

Figure 1 is a schematic of the first laser studied. This laser had internal gold mirrors sealed by a bellows arrangement for external adjustment within a 1.2 meter tube of $1-3 / 16$ inch inside diameter. The mirrors at each end of the tube were nearly $100 \%$ reflecting with the exception of a small hole (various mirrors had holes with diameters from 1 mm to 5 mm ) in the output mirror. The system was basically a flow system of $\mathrm{CO}_{2}, \mathrm{~N}_{2}$, He and many combinations of partial pressures of these gases were tried. A glass jacket around the discharge tube was used for air cooling. The laser, rf excited by a Johnson Viking 80 watt transmitter, had an output of approximately 3 watts. A second laser very similar to the first was dc excited by electrodes in each end. The external jacket in this case was used for either air or water cooling.

From the spectral data taken from these lasers a second basic design evolved (Figure 2).

This laser had a discharge tube sealed at both ends by salt windows set at Brewster's angles, and with the mirrors externally mounted. The tube was only 30 cm long and contained an external glass jacket for cooling by circulation of air or a liquid dielectric. The mirrors were readily interchangeable and various configurations of radius of curvature of the mirrors and size of the output holes

[^2]
Figure 1-CO $\mathrm{CO}_{2}$ Laser-RF Excited
Loose bushing to
allow free expansion -allow

Figure 2
were tried. All of these lasers were rf excited but the power output was low ( 50 milliwatts) compared to the longer lasers. Rf excitation was accomplished by placing parallel electrodes along the sides of the laser tube. The electrodes and plasma tube formed a capacitive system which was tuned to the resonant frequency of the rf exciter by addition of an inductance coil and adjustable tuning capacitor.

## SPECTRAL CHARACTERISTICS

The laser action for all the lasers used was from either the P- or R-branch rotational transitions of the $00^{\circ} 1-10^{\circ} 0$ vibrational bands of $\mathrm{CO}_{2}$ (Reference 1). The spectral output of the lasers either rf or dc excited, when the mirrors were sealed within the tube, was essentially the same and was quite unstable. All observations where made with the Cary White 90 infrared spectrophotometer.

The laser action varied from a single transition to as many as six transitions simultaneously but these were seldom repeatable. Figures 3, 4, and 5 are typical of many spectra taken of both the rf and dc excited lasers with sealed windows. On some runs spectral lines were separated by as much as $36 \mathrm{~cm}^{\mathbf{- 1}}$. It was determined that the fluctuations of input power, tube pressure, and thermal effects all contributed to the instabilities.

The shorter lasers with external, isolated mirrors did not support as many different spectral lines. In almost every run frequency stability had greatly improved over the previous lasers. As many as four transitions were observed at once, but all transitions were tightly grouped with the spread between lines never greater than $2 \mathrm{~cm}^{-1}$. Figure 6 shows a very sharp single spectral line from one of the short lasers. Figure 7 is the same spectral line but under a higher resolution and dispersion. The peaks shown here are much too close to be spectral lines. They are probably some of the more dominant modes within the $P(20)$ line brought out by the higher resolution of the spectrophotometer.

## MODES

The laser cavity supports not only the spectral radiation, but the electromagnetic modes which are characteristic of the resonant cavity. For a confocal resonator the resonant condition is given by (Reference 2):

$$
4 b / \lambda=2 q+(m+n+1)
$$



Figure 3-Spectrum of the Radiation from a D.C. Excited $\mathrm{CO}_{2}$ Laser Containing $\mathrm{N}_{2}$ and $\mathrm{CO}_{2}$. Note the Number of Spectral Lines and the Separation of one Line from the Group.


Figure 4-Spectrum of the Radiation From a D.C. Excited $\mathrm{CO}_{2}$ Laser Containing $\mathrm{CO}_{2} \mathrm{Only}$


Figure 5-RF Excited $\mathrm{CO}_{2}$ Laser With Laser Action On the Signal (P22) Line Only. The Laser Contained $\mathrm{CO}_{2}$ Only.

Figure 6-Very Sharp Laser Line ( $\mathrm{P}(20)$ ) Under Moderate Resolution. This Line Can, With Some Adjustment, Be Regularly Obtained From The Short $\mathrm{CO}_{2}$ Lasers With the Isolated Mirrors.

P(20) - $10.59 \mu$
17
0
0
0
0

1
1
2
0
0
0
0
0

Figure 7-Under High Resolution and Dispersion Some Modes Can Be Distinguished In a Single Spectral Live ( $\mathrm{P}(20)$ ) From the Short $\mathrm{CO}_{2}$ Laser
where $b$ is the mirror separation, $\lambda$ the wavelength, $q$ the number of half-waves along the length of the laser, and $m$ and $n$ designate the mode. For a laser resonator of 50 cm length the separation of the electromagnetic modes is of the order of $5 \times 10^{-3} \mathrm{~cm}^{-1}$. This means that as many as 20 electromagnetic modes can appear under the width of a single spectral line (natural line width $\sim 0.10$ $\mathrm{cm}^{-1}$ ). An experiment was designed (Figure 8) to look at these individual modes within a spectral line. In the experiment, a Lansing scanning interferometer was modified by replacing the flat Fabry-Perot type mirrors with a confocal arrangement of nera $100 \%$ reflecting gold mirrors, each with a 0.25 mm diameter hole drilled in its center. Radiation is coupled into the cavity through the hole in one mirror. When the cavity mirrors are critically aligned, for a given wavelength, i.e. the wavelength of an electromagnetic mode, the cavity will be in resonance and radiation from the output hole will be increased. When the length of the cavity is changed, this wavelength will not resonate and the output energy is reduced. Part of the incident radiation is picked off by a beam splitter and sent into an infrared spectrophotometer. This will make it possible to determine the wavelength of the spectral line from the laser, although it cannot clearly resolve the modes found within a spectral line.

One of the mirrors of the Lansing interferometer is backed by a piezoelectric translator which is driven by a linear ramp generator. The total translation in this particular arrangement is 1.2 microns which is a little more than one-tenth of the incident wavelength ( 10.6 microns) from the $\mathrm{CO}_{2}$ laser.

Figure 9a shows an oscilloscope sweep of the laser signal coupled out of the interferometer when the mirrors are stationary. Only the instrument noise appears in this trace. However, Figure 9b shows a sweep of the signal as the length of the cavity is shortened by 1.2 microns. The spikes on this trace are the electromagnetic modes that are brought into resonance as the cavity length is changed.

At this writing there has not been a detailed analysis of these traces. This could be a tedious task because the spikes along the trace not only represent the modes reinforced within the cavity but also the degenerate modes.

Figure 10a shows a sweep of the modes, and Figure 10b shows a sweep after the laser mirrors have been adjusted very slightly. This figure shows additional modes being reinforced and strengthened and could thus give an indication of what goes on inside the laser itself.

This interferometer arrangement should make it possible to completely analyze the closely spaced electromagnetic modes of the $\mathrm{CO}_{2}$ laser. It is hoped that this knowledge will aid in the design and development of a highly stable laser.


Figure 8


Figure 9(a) -Signal Plus Noise From $\mathrm{CO}_{2}$ Laser Radiation Through the Apparatus in Figure 8. (b )-Spikes in the Sweep Represent Laser Modes as They Brought Into Resonance as the Resonator Cavity is Shortened.


Figure 10(a)-Spikes in the Sweep are Laser Modes as They are Brought Into Resonance By Shortening the Resonator Cavity (Figure 8). (b)-The End Mirrors of the Laser are Slightly Adjusted and Some Modes are then Reinforced Over Others.

The design efforts on the short lasers with external, isolated mirrors were successful in producing a more monochromatic source. Study of this design shows that more needs to be understood about the following areas:
a. excitation mechanism of the laser action,
b. the modes within the laser cavity and how to control them,
c. the dependence of the lifetime of a closed laser tube upon the gases used and the preparation of the laser tube.

The experimental set up in Figure 8 is the first time a cavity resonator has been used in the infrared wavelength region. This has much promise as a means of analysis of the modes of lasers in the infrared region. This technique has good possibilities of application in the visible radiation region also.

## REFERENCES

1. C. K. N. Patel, "Continuous - Wave Laser Action on Vibrational - Rotational Transitions of $\mathrm{CO}_{2}$," Physical Review, Vol. 136, No. 5A, 30 Nov. 1964, pp. A1187-A1193.
2. G. D Boyd and J. P. Gordon, "Confocal Multimode Resonator for Millimeter Through Optical Wavelength Masers," The Bell System Technical Journal, March 1961, pp. 489-505.
3. T. K. McCubbin, Jr., Ronald Darone, and James Sorrell, "Determination of Vibration - Rotation Line Strengths and Widths Using a $\mathrm{CO}_{2}-\mathrm{N}_{2}$ Laser," Applied Physics Letters, to be published.

## RELATIVE PHASE DIFFERENCE ANALYSIS PROGRAM

by
D. M. Koenig*

## INTRODUCTION

Oscillator-pair data, which is relative phase difference ( $8 \theta$ ), is recorded continuously on strip charts (see Figure 1).


Figure 1

This curve is sampled at times $t_{i}, i=1,2, \ldots N_{3}$ where $t_{i}=(i-1) \Delta t$ with $\Delta t$ usually one hour. The temperature at these times is also sampled; thus a set of N data triples is produced: $\left(\mathrm{T}_{\mathrm{i}}, \delta \theta_{\mathrm{i}}, \mathrm{t}_{\mathrm{i}}\right), \mathrm{i}=1,2, \ldots \mathrm{~N}$.

The data are discontinuous because of scale shifts and crossovers. These points of discontinuity are coded (values are assigned to the variable $L_{i}$ ) appropriately so each sample point now has four numbers associated it: $t_{i}, T_{i}$, $\delta \theta_{i}, L_{i}$. For every data point, $\mathrm{T}_{\mathrm{i}}, \delta \theta_{\mathrm{i}}, \mathrm{L}_{\mathrm{i}}$ are punched onto one IBM card.

These cards are read by the program and the data is corrected for the discontinuities. The corrected data is printed out for checking purposes. The second phase consists of fitting the data to a double polynomial in time and temperature, i.e., the coefficients $\mathrm{A}_{\mathrm{k}}$ are found in

$$
\delta \theta_{i}=A_{1}+A_{2} t_{i}+A_{3} t_{i}^{2}+A_{4} T_{i}+A_{5} T_{i}^{2}+e_{i}
$$

[^3]by minimizing
$$
\sum_{i=1}^{N} e_{i}^{2}
$$

This is done by a standard least squares procedure which requires the solution of a $5 \times 5$ linear system.

The estimated standard deviation s, where

$$
s^{2}=\frac{1}{N} \sum_{i=1}^{N}\left(e_{i}\right)^{2}
$$

is calculated and each $e_{i}$ is compared to 3 s . If $\left|e_{i}\right|>3 \mathrm{~s}$, the ith point is rejected. After all of the $e_{i}$ are compared to $3 s$ the coefficients $A_{k}$ are recalculated using only those points passing the rejection test.

The last phase consists of an autocorrelation of the residuals $e_{i}$ to show the presence of any periodic behavior.

## DETAILED DESCRIPTION OF SUBPROGRAMS

## A. Overall Configuration

The program consists of seven subprograms entitled as PRG1, PRG2, etc. The overall flow chart is shown in Figure 2. The statement "transfer to PRG1" means transfer to the appropriate routine, carry out the necessary operations and then return to the point of transfer.

## B. PRG1: Data Corrector Routine

## 1. Critical Point Coding

Considering the case of increasing data, discontinuities of three types can occur and these are illustrated below in Figure 4.


Figure 2


PRG 4: REDUCE THE MATRIX BY THE GANSS-JORDAN METHOD AND DETERMINE THE COEFFICIENTS $A_{k}$

```
RETURN
```

PRG 5: CALCULATE THE $\mathrm{e}_{\mathrm{j}}, \overline{\mathrm{e}}, \overline{\mathrm{e}}^{2}$ AND SUBTRACT ē FROM EACH $\mathrm{e}_{\mathrm{i}}$ RETURN


Figure 3


Figure 4

Often one or more points are missing during a shift. To account for this, a certain value is given to the code variable which is called LABJ, where J refers to the Jth point. The coding system is given in Table 1.

Table 1

| Type of Point | Code |
| :--- | :---: |
| last good point before shift | 1 |
| bad or missing point | 3 |
| first good point after crossover | 5 |
| the very last good point | 6 |

An example of the point coding is shown in Figure 5. If a point does not fall into any of the classifications of Table 1 it is left uncoded. Thus for each data point one IBM card is prepared containing a temperature, relative phase difference, and a code (if it is a critical point).

## 2. Overall Procedure for PRG1

PRG1 reads and processes the data cards one at a time. On reading a card, and thus a data triple, values are temporarily assigned to the variables TMP, RPD, LABJ which stand for temperature, relative phase difference, and label.


Figure 5

Depending on LABJ, a new value is assigned to RPDB which is the corrected relative phase difference. The values of TMP, RPDB, and LABJ are assigned to $\operatorname{DTI}(J)$, OMEGA( $J$ ), and LAB(J) where $J$ is the point index. Having stored the corrected data, another card is read in , new values are assigned to TMP, RPD, LABJ, J is stepped up by one, and the cycle is repeated.

## 3. Correction for Crossovers

Each point is corrected by adding to it the value of a variable called AOF or "add-on-factor." Initially AOF is given the value zero. After each crossover AOF is increased by $1 \not \varnothing \emptyset \emptyset$. Since the data triples (TMP, RPD, LABJ) are read in sequentially, AOF is changed whenever LABJ has the value 5.
4. Correction for Shifts

If a point's code is 1 , it is the last good point before a shift. To correct the next good point a straight line is passed through the previous data points back to a shift or a crossover. This straight line is used to estimate the value of relative phase difference at the next good point had there been no shift. The difference between the actual and the estimated relative phase difference is subtracted from AOF and is used to correct the current point. The value of AOF is left unmodified until another crossover or shift occurs. To calculate the intercept and slope of the straight line a least squares method is used.

## 5. Modification for Decreasing Data

For decreasing data the add-on-factor becomes a subtract-off-factor. To account for this, AOF is multiplied by a factor FAC which is either plus or minus one. The value of FAC is read in at the beginning of the program and is set to plus one for increasing data.
6. Flow Diagram for PRG1



## C. PRG2: Subtraction of Mean Temperature

The mean temperature is calculated and then subtracted from each individual temperature.
D. PRG3: Polynomial Fitting Routine

The primary task of this program is to determine the coefficients $A_{k}$ in the double polynomial. The method of least squares is used and is derived as follows:

$$
\begin{aligned}
& \delta \theta_{i}=A_{1}+A_{2} T_{i}+A_{3} T_{i}^{2}+A_{4} t_{i}+A_{5} t_{i}^{2}+e_{i} \\
& \frac{\partial}{\partial A_{k}} \sum_{i}^{n} e_{i}^{2}=0 \\
& \sum e_{i} \frac{\partial e_{i}}{\partial \mathrm{~A}_{\mathrm{k}}}=0 \quad k=1,2,3,4,5 \\
& e_{i}=\delta \theta_{i}-A_{1}-A_{2} T_{i}-A_{3} T_{i}^{2}-A_{4} t_{i}-A_{5} t_{i}^{2} \\
& \frac{\partial e_{i}}{\partial A_{1}}=-1, \quad \frac{\partial e_{i}}{\partial A_{2}}=-T_{i}, \quad \frac{\partial e_{i}}{\partial A_{3}}=-T_{i}^{2}, \\
& \frac{\partial e_{i}}{\partial A_{4}}=-t_{i} \quad \frac{\partial e_{i}}{\partial A_{5}}=-t_{i}^{2} \\
& \sum \delta \theta_{i}=N_{1}+A_{2} \sum \mathrm{~T}_{\mathrm{i}}+\mathrm{A}_{3} \sum \mathrm{~T}_{1}^{2}+\mathrm{A}_{4} \sum \mathrm{t}_{\mathrm{i}}+\mathrm{A}_{5} \sum \mathrm{t}_{\mathrm{i}}^{2} \\
& \sum \delta \theta_{i} T_{i}=A_{1} \sum T_{i}+A_{2} \sum T_{i}^{2}+A_{3} \sum T_{i}^{3}+A_{4} \sum t_{i} T_{i}+A_{5} \sum t_{i}^{2} T_{i}
\end{aligned}
$$

$$
\begin{aligned}
& \sum \delta \theta_{i} \mathrm{~T}_{\mathrm{i}}^{2}=\mathrm{A}_{\mathrm{i}} \sum \mathrm{~T}_{\mathrm{i}}^{2}+\mathrm{A}_{2} \sum \mathrm{~T}_{\mathrm{i}}^{3}+\mathrm{A}_{3} \sum \mathrm{~T}_{\mathrm{i}}^{4}+\mathrm{A}_{4} \sum \mathrm{t}_{\mathrm{i}} \mathrm{~T}_{\mathrm{i}}^{2}+\mathrm{A}_{5} \sum \mathrm{t}_{\mathrm{i}}^{2} \mathrm{~T}_{\mathrm{i}}^{3} \\
& \sum \delta \theta_{\mathrm{i}} \mathrm{t}_{\mathrm{i}}=\mathrm{A}_{1} \sum \mathrm{t}_{\mathrm{i}}+\mathrm{A}_{2} \sum \mathrm{t}_{\mathrm{i}} \mathrm{~T}_{\mathrm{i}}+\mathrm{A}_{3} \sum \mathrm{t}_{\mathrm{i}} \mathrm{~T}_{\mathrm{i}}^{2}+\mathrm{A}_{4} \sum \mathrm{t}_{\mathrm{i}}^{2}+\mathrm{A}_{5} \sum \mathrm{t}_{\mathrm{i}}^{3} \\
& \sum \delta \theta_{\mathrm{i}} \mathrm{t}_{\mathrm{i}}^{2}=\mathrm{A}_{1} \sum \mathrm{t}_{\mathrm{i}}^{2}+\mathrm{A}_{2} \sum \mathrm{t}_{\mathrm{i}}^{2} \mathrm{~T}_{\mathrm{i}}+\mathrm{A}_{3} \sum \mathrm{t}_{\mathrm{i}}^{2} \mathrm{~T}_{\mathrm{i}}^{2}+\mathrm{A}_{4} \sum \mathrm{t}_{\mathrm{i}}^{3}+\mathrm{A}_{5} \sum \mathrm{t}_{\mathrm{i}}^{4}
\end{aligned}
$$

These five equations are solved for the coefficients by means of a Gauss Jordan reduction routine (PRG4).

## E. PRG4: Gauss-Jordan Reduction Routine

The procedure by which the linear system, constructed in PRG3, is solved is standard (see p. 224), Southworth and DeLeeuw, "Digital Computation and Numerical Methods," McGraw-Hill, 1965), and will not be discussed here. The title of the subroutine is GSRDN (A, B, X, N, DET). A is a doubly subscripted variable representing the elements of the coefficient matrix, $B$ is a singly subscripted variable representing the input matrix and $X$ is another singly subscripted variable representing the unknown matrix,

$$
\text { (A) }\{X\}=\{B\} .
$$

$N$ is the dimension of (A) which for the case discussed in $D$ is 5. DET is the value of the determinant of (A).
F. PRG5: Calculation of Error at Each Point

At each point $e_{i}, i=1, \ldots n$ is calculated, where

$$
e_{i}=\delta \theta_{i}-A_{1}-A_{2} T_{i}-A_{3} T_{i}^{2}-A_{4} t_{i}-A_{5} t_{i}^{2}
$$

at those points for which $\operatorname{LAB}(\mathrm{I})$ is equal to $3, e_{i}$ is set equal to zero. The average of the errors

$$
\bar{e}=\frac{1}{n} \sum_{i=1}^{n} e_{i}
$$

and the average of the errors squared

$$
\frac{1}{n} \sum_{i=1}^{n} e_{1}^{2}
$$

is calculated.

## G. PRG6: Rejection of Bad Points

The standard deviation of the $e_{i}$ is calculated

$$
s=\frac{1}{n} \sqrt{\sum e_{i}^{2}-\left(\sum e_{i}\right)^{2} / n}
$$

and then all the $\left|e_{i}\right|$ are compared to 3 s . If $\left|e_{i}\right|$ is greater than 3 s the value of $e_{i}$ is set to zero and $\operatorname{LAB}(\mathrm{I})$ to 3 . The procedure is repeated using the remaining non-zero $e_{i}$ where $n$ is the number of non-zero $e_{i}$. The flow diagram for this routine is shown below.


Figure 6

# - N68-25 8 8 

# AN INVESTIGATION OF METHODS FOR THE EXPERIMENTAL DETERMINATION OF PCM SIGNAL TO NOISE RATIO 

by
Jack F. Morris

## INTRODUCTION

The percentage of recovery of digital PCM data from a given magnetic tape can in general be related to the signal-to-noise ratio (SNR) at the telemetry system receiver. Occasionally, however, data that might normally be recovered is lost through processing errors or equipment malfunction. A study was proposed by the Data Processing Branch, the object of which was twofold: the telemetry link itself was to be investigated for the sources and nature of the noise at each stage of the data path through the receiver system to the analog data tape, and secondly, since the probability of bit error is directly related to the analog data SNR, it was proposed that one or more methods might be developed by which an estimate of data recovery might be made, independent of the processing line.

Spectral characteristics of the receiver system noise and SNR's throughout the system were studied experimentally using the ground stations set up in the Information Processing Division (IPD), and in the Network Test Training Facility (NTTF). The effect on data recovery of varying the modulation index and of varying the IF and post-detection bandwidths were also investigated.

In the second phase of the work several of the classical methods of signal analysis were investigated for applicability to PCM analog data. Computer programs and digital techniques were developed for obtaining signal amplitude probability distributions over any desired segment of analog data, for obtaining bit-by-bit statistics within the frame sync pattern, and for obtaining a spectral analysis of the data. Each of these programs will yield a SNR for the analog data, based upon the statistics developed in the program. An estimate of the probability of bit error, and therefore an estimate of the quality of the data, or the recovery expected, can thus be made for any given PCM analog data rape.

[^4]In an attempt to identify the sources of noise and the nature of the noise in the PCM telemetry receiver system, the following approaches were taken:

1. Magnetic tape recorded data, in the OGO-D satellite format were transmitted through the collimating antenna to the Satan antenna at NTTF. The received data were recorded at a series of known signal levels, calculated from a reference receiver system noise figure.
2. A clock signal, at a 64 kilobit persecond rate, was transmitted, received, and recorded in the same manner. The purpose of this was to provide a known signal, with a relatively simple frequency spectrum, to facilitate noise spectrum analysis.
3. Using the receiver system in IPD, simulated data were inserted into the RF path of the system. The modulation index and the signal level were both fully controllable, and the effect on the output data due to various inputs was investigated. For low signal conditions an apparent decrease in SNR improvement through the PM demodulator was noted, as was expected, and for high signal levels the SNR improvement approached the theoretical value.
4. In an investigation of the effect of tape recording on data quality, simulated data of known quality were recorded from the output of the IPD receiver system. The tapes were then played back into the IPD data analysis equipment. In each case few, if any, more errors were observed on playback than when recording, and the maximum increases in error rate represented possibly a 0.1 db degradation in SNR through recording.

## RESULTS OF RECEIVER SYSTEM INVESTIGATION

The primary, and surprising, result obtained in working with the transmission of simulated data through the system at NTTF is that data quality is not always reliably represented by the signal level indication on the system receiver. Clean simulated data were recorded at 60 inches per second (ips) by IPD people. These tapes were at data rates of 1,8 , and 64 kilobits per second (kbps). The simulated data were transmitted through the collimating antenna, received by the system's Satan antenna, and recorded on magnetic tape. The OGO-D satellite Operations Plan specifications were used for receiver and recorder settings. Four minutes of data were recorded at each of the signal
levels calculated for SNR as indicated in Table I below. The recorded data were then played back through the signal analysis line in the Processor Development Branch, IPD. A non-uniform discrepancy was noted to exist between calculated SNR (and, therefore, bit error rate) and the observed bit error rate, as indicated:

Table I
Calculated and Measured SNR

| SNR Calculated <br> by NTTF, in db | SNR measured <br> by IPD, in db | Error <br> in db |
| :---: | :---: | :---: |
| 4.85 | 3.00 | 1.85 |
| 7.30 | 4.90 | 2.40 |
| 9.65 | 6.80 | 2.85 |
| 11.90 | 8.60 | 3.30 |
| 13.90 | 10.20 | 3.70 |

Several possible explanations exist for the discrepancies noted, none of which has been fully investigated at this time. One of these possibilities is in the nature of the simulated data: because of the bandwidth limitations of magnetic tape recording, it is difficult to obtain a satisfactory simulation of satellite data by recording and playing back the output of a signal generator. Another problem area is in the calculation of SNR using the IF bandwidth and system noise figure as the basis. Experimentally measured noise power in the IF seems to be slightly higher than the value calculated using the noise figure. It has been proposed that perhaps the IF filter cutoff is not sufficiently sharp that simply using the half-power bandwidth will determine total IF noise power to the desired accuracy. Further investigation of each of these problems is planned.

A spectral analysis of the recorded clock signal plus noise revealed that the sky noise plus receiver noise in the system does have the "white" or uniform frequency distribution so often assumed in the literature pertaining to signal analysis. Observation on a spectrum analyzer, and measurement of noise power using various ranges of post-detection filter bandwidth, established the uniform nature of the noise in the 136 MHz range utilized by the telemetry system.

Using the receiver system in IPD, experimental values of SNR improvement through the PM demodulator were obtained, along with results indicating essentially negligible degradation of the data through tape recording. The absolute noise power in the receiver IF measured in the absence of any signal and with manual control of the AGC. Signal levels were then adjusted for various SNR's,
referred to the information bandwidth, and a frame sync pattern bit error count was made. Results were consistent over many runs and several days, and the values obtained are shown in Table II below:

Table II
SNR Improvement through PM Demodulator

| SNR in <br> IF Filter | Post-detection <br> SNR | SNR <br> Improvement |
| :---: | :---: | :---: |
|  |  |  |
| 1.8 db | 2.5 db | 0.7 db |
| 4.8 | 6.1 | 1.3 |
| 7.8 | 9.6 | 1.8 |
| 10.8 | 12.7 | 1.9 |

A modulation index of 1.25 was used for this work, which would yield a SNR improvement of 1.9 db for large signals, according to the theory. Although the last two measurements may be suspect, because of the very large number of frames of data which must be processed for valid statistical results, it should be noted that throughout all of this investigation the results obtained for data with SNR greater than about 9 db agreed in every case with theoretical predictions.

To complete the receiver system phase of this initial investigation, the signals described above were recorded on magnetic tape and then played back to determine the amount of degradation of SNR in the recording process. For each recording the IF was six times the information rate and post detection bandwidth was ten times the information rate. Tape speed was 30 ips . In the worst case the degradation amounts to about 0.1 db , which is essentially negligible. Results appear in Table III.

Table III
Comparison of Error Rates

| SNB <br> (db) | Pre-recording <br> error rate <br> (bit errors/frames) | Post-recording <br> error rate |
| :---: | :---: | :---: |
|  |  |  |
| (bit errors/frames) |  |  |

Several more data points were taken, at different post detection bandwidths, and the results were similar to the above. A slight improvement was observed in some cases and a degradation in other cases when operating the system with narrowed IF and post detection bandwidths. It appears at this writing that a slight improvement in data recovery can be obtained, using narrow band filter settings. The improvement is very small, however, and is very much a function of the tuning and adjustment of the receiver. When the receiver was tuned off of the carrier frequency by as little as 10 kHz the degradation in performance began to be noticeable: 200 errors in 27,000 frames, as opposed to 165 errors for the wider IF bandwidth data in the table above. The trade off between the system accuracy required to obtain a very slight improvement and the minor degradation experienced using the wider band system settings definitely favors the latter.

## PHASE II - PCM ANALOG DATA ANALYSIS <br> STATEMENT OF THE PROBLEM

Satellite telemetry data recorded on instrumentation type analog magnetic tape at the worldwide network of STADAN stations are mailed to Goddard Space Flight Center (GSFC) for processing. The telemetry data acquired from the satellites by the ground stations are received in less than an ideal form that includes blankouts, weak signals and noise. Phase II of this investigation was concerned with the study of PCM signal and noise characteristics of actual flight data recorded on magnetic tape, and the development of computer techniques and procedures for automatic signal evaluation.

Most of the theoretical investigations of PCM SNR which have appeared in the literature have been based upon an ideal square wave type signal corrupted by white Gaussian amplitude noise. Under these assumptions, it is easily shown that the square of the mean value of the signal-plus-noise represents the signal power and the variance about the mean represents the noise power. That is:

$$
\frac{\text { mean squared }}{\text { variance }}=\frac{\mathrm{m}^{2}}{\sigma^{2}}=\frac{\text { signal power }}{\text { noise power }}=\text { SNR }
$$

Thus, when experimentally measured values of signal mean and variance are used as an estimator of SNR, in the statistical sense, there are standard statistical procedures by which the true SNR may be obtained with any level of confidence. All that is required by these procedures is that a larger and larger number of samples be used to obtain the $\mathrm{m}^{2} / \sigma^{2}$ estimator of NSR as the confidence level is raised on the true value.

When the actual flight data signal-plus-noise is considered, however, it is seen that the signal is not ideal and the noise is not strictly an amplitude noise. Rather, the signal portion of the waveform represents an ideal square wave which may have been pre-filtered before transmission, has been subjected to. atmospheric fading and phase shift, has been further filtered by the IF and postdetection filters of the receiver system, and again filtered by the transfer characteristics of the record and playback electronics of the tape system. A typical flight data waveform is represented by Figure 1 below:


Figure 1-Typical PCM flight data waveform.

The filtered signal waveform is in addition corrupted by amplitude noise, by phase noise or phase jitter, by wow and flutter of the tape recording system and by other influences which may be considered to be noise, such as the spin modulation introduced by the rotation of the satellite.

The processing of PCM analog data is further complicated by the fact that it must generate its own timing signals, or clock. The data tape is played into a bit synchronizer and signal conditioner which must lock onto the data bit stream with an internal voltage controlled oscillator (VCO). The VCO frequency is corrected automatically for small changes in signal frequency, but cannot readily synchronize with data of low SNR. Thus, for the low SNR condition a reliable data-rate clock, derived from the VCO, may not be available.

## SUMMARY OF THE WORK

In an attempt to obtain a measurement of PCM data SNR, auto-correlation, cross-correlation, and ensemble averaging of the analog bit stream were considered. Except for an adaptation of the cross-correlation technique, from which a synchronizing signal was obtained, the corelation process did not lend itself to application to this problem. The reason for this lies in the nature of the data format, as explained in the following section.

Synchronous and random sampling procedures were used to acquire statistics on the data from which signal amplitude probability distributions, signal mean and variance, and spectral characteristics of the signal were determined through digital techniques. The amplitude distribution program was designed to yield a SNR for the data from the mean squared signal, representing signal-plus-noise power, and the mode, representing the signal power alone. The signal mean and variance program was designed to yield average and/or bit-by-bit statistics, from which SNR was calculated directly. Agreement between the two independent programs was found to be consistent, although a calibration routine is required before agreement is complete.

A spectral analysis program was written utilizing the same data samples as above and an adaptation of the Cooley-Tukey fast Fourier transform (FFT) algorithm. This program was designed to yield a SNR based upon the knowledge that the second harmonic of the signal bit rate is essentially non-existent. Any frequency components in the neighborhood of the second harmonic may therefore be interpreted as noise, and a noise power figure obtained.

## RESULTS OF PCM ANALOG SIGNAL INVESTIGATION

The data format of the OGO-D satellite was the format used for all of the signal investigation. This format is typical of many PCM systems, and consists of a "frame" of 128 nine-bit "words". Each frame has three known or programmed words used to locate the frame and synchronize the words for computer manipulation. These three words are termed the "frame sync pattern", and in the OGO format are the first three words of the frame, one shown schematically in Figure 2.

Now consider that the analog signal may be represented by

$$
f(t)=S(t)+N(t),
$$



Figure 2-PCM data format for OGO satellite.
where $S(t)$ is a periodical signal with zero mean, and $N(t)$ is a random noise with zero mean and without any periodic component. Then the auto-correlation function of $f(t)$ is given by

$$
\begin{aligned}
\phi_{\mathrm{ff}}(\tau) & =[\mathrm{S}(\mathrm{t})+\mathrm{N}(\mathrm{t})][\mathrm{S}(\mathrm{t}+\tau)+\mathrm{N}(\mathrm{t}+\tau)] \\
& =\phi_{\mathrm{SS}}(\tau)+\phi_{\mathrm{NN}}(\tau)+\phi_{\mathrm{NS}}(\tau)+\phi_{\mathrm{SN}}(\tau) \\
& =\phi_{\mathrm{SS}}(\tau)+\phi_{\mathrm{NN}}(\tau)
\end{aligned}
$$

under the conditions of zero mean signal and noise.
In a practical situation $\phi_{\mathrm{NN}}(\tau)$ tends to zero at some sufficiently large value of $\tau$, such that for uncorrelated signal and random noise the function

$$
\phi_{\mathrm{ff}}(\tau)=\phi_{\mathrm{SS}}(\tau), \text { for large } \tau,
$$

as indicated schematically in Figure 3. From the defining equation it is seen that $\phi_{f \mathrm{f}}(\tau=0)$ represents the mean squared signal-plus-noise, or the total signal-plus-noise power. $\phi_{\mathrm{SS}}(\tau=0)$ represents the signal power alone, and its value may be determined at some $\tau-\tau_{0}$ such that $\phi_{N N}\left(\tau_{0}\right)=0$ as indicated in Figure 3.

One way in which the auto-correlation process might be applied to PCM data, in an attempt to determine the signal component and the noise component of the data waveform, would be to treat the known frame sync pattern as a periodic function in the presence of noise. To do this requires that the data be considered purely random, as well as the true noise component. The limited usefulness of the method becomes apparent when it is noted that the frame sync pattern represents only 3 of 128 words in the frame, or less than $3 \%$ of the total signal. Thus, even though measurement of $\phi_{\mathrm{SS}}\left(\tau_{0}\right)$ may be practical, the errors
introduced by any periodic data would in all likelihood be greater than the measurement. That is, it would only require that 3 data words be constant for several consecutive frames to introduce a $100 \%$ error in the signal power measurement.


Figure 3-Autocorrelation function of a periodic signal plus random noise.

Another approach is to consider the frame sync pattern as the periodic component and the noise in the frame sync words as the random component. This requires gating out the data, and considering only the first 27 bits of each frame, as shown below:


Figure 4-Schematic of conditions for auto-correlating the frame sync words.

This latter technique was not pursued in the course of the investigation. The timing problem associated with gating out the frame sync patters and essentially connecting them end to end coherently was not solved during the time allotted. Further work might be done in this area of a manageable solution to the timing difficulty can be found.

The most useful technique discovered during this work was an adaptation of an amplitude probability distribution program. A procedure was developed under which either random samples or synchronous samples of the bit stream were collected, the probability of each amplitude increment calculated, and the probability as a function of amplitude plotted by digital computer. Figure 5 shows a representative probability curve for a data waveform such as that in Figure 1:


Figure 5-Typical amplitude probability function for a PCM data waveform.

The mean squared value of the distribution is proportional to the total signal-plus-noise power when the mean value is zero. That is,

$$
\begin{aligned}
& \overline{v^{2}(t)}=E\left(v^{2}\right)=\int_{-\infty}^{\infty} v^{2} p(v) d v \\
& E\left(v^{2}\right)=[E(v)]^{2}+\sigma^{2}
\end{aligned}
$$

where $\sigma^{2}$ is the variance, or second moment about the mean. When

$$
[E(v)]^{2}=0, \text { Then } E\left(v^{2}\right)=\sigma^{2}
$$

is the power delivered to a one ohm resistor across which $v$ is the voltage.
Under the assumption that the mode E of the distribution is related to the signal power, the SNR was expressed as

$$
\text { SNR }=\frac{\text { signal power }}{(\text { signal }+ \text { noise power })-(\text { signal power })}
$$

or

$$
\mathrm{SNR}=\frac{\mathrm{kE}}{\mathrm{E}^{2}} \frac{\sigma^{2}-\mathrm{kE}^{2}}{}
$$

where k is a proportionality constant, determined by calibration.

This program was used with samples of actual flight data for which error statistics had previously been obtained. Accurate agreement over a range of SNR's from approximately 4 db to 10 db was obtained using a value of $\mathrm{k}=0.78$, where the constant was chosen to yield a SNR of 20 db for clean simulated data played back from magnetic tape. The program generated consistent results when used with random samples of data, synchronous samples of data, or signal generator waveforms, and is considered a successful technique for obtaining SNR's.

An oscilloscope display of the PCM flight data reveals a slight pattern sensitivity in the data waveform. That is, certain bit patterns in the data seem to introduce anomalies in the waveform that are not observed with different bit combinations. A Polaroid photograph of several consecutive sweeps through the same data word is reproduced in Figure 6. Several repeating high and low values of the waveform are clearly seen in this picture.


Figure 6-Polaroid photo of PCM data waveform.

To investigate the apparent pattern sensitivity of the waveform, a digital computer program was developed which uses a cross-correlation technique to locate the frame sync pattern in a stream of data samples obtained synchronously, one or more samples per information bit. When the frame sync pattern is found, the samples from each bit within the pattern are stored separately.

The mean and variance about the mean for each bit in the frame sync can be computed, as well as the mean of the 27 bits in the pattern taken as a whole. Other statistics using these samples are easily generated with minor program changes.

It was found from this program that the pattern sensitivity was clearly revealed by the samples. Figure 7 represents the mean value, by bit position, obtained for the 27 bits of the frame sync pattern on a typical flight data tape of approximately 12 db SNR. The $3 \sigma$ value, where $\sigma$ is the standard deviation, is shown on each bit and represents the noise level that would results in an error rate of approximately $5 \times 10^{-3}$

In this program the SNR is derived from the ratio of mean squared to variance, where in this case the mean is determined at a constant phase point on each bit by synchronous sampling. The variance is taken about the signal bit mean, and is proportional to the noise power. It was found that the variance of the noise power from bit to bit is essentially negligible. That is, the noise contribution at each bit is the same, regardless of the mean value of the signal bit. This can be seen in the relatively uniform heights of the $3 \sigma$ points in Figure 7. Thus for high noise conditions certain bits would tend to be in error more often than others. That this is indeed the case is borne out by bit error statistics developed independently of this program by IPD.

A valuable feature of the amplitude distribution and cross-correlation programs discussed above is that they are applicable without modification to the study of the PCM bit synchronizer output SNR. All that is necessary is that the output of the "integrate and dump" circuitry be made available for use as a signal source. Amplitude distributions, mean and variance, and SNR figures for the output data may be compared over whole data tapes, or over any desired segment of data. This feature should contribute significantly to an evaluation of bit synchronizer performance.

The last program developed during this work was an adaptation of the Cooley-Tukey fast Fouruer transform (FFT) algorithm. The program will yield a frequency spectrum with a maximum width corresponding to one half the sampling rate. The resolution is a function of the number of samples processed for any given spectral "display". There is some danger of losing or distorting spectral information near the highest frequencies, and the sampling rate should therefore be chosen somewhat higher than twice the maximum frequency of interest.
3)


When the PCM analog data is applied to a conventional swept frequency spectrum analyzer, or to a real time analyzer, it is seen that the second harmonic component of the signal is essentially zero. The spectral null near the second harmonic in quite broad, and since the signal waveform is ideally a square wave type signal, which has undergone linear filtering, it should be expected that this would be true. Taking advantage of the null in the signal spectrum, and assuming that the frequency components near the second harmonic are due to noise alone, a digital filter is placed around any desired band of frequencies near the signal null. The digital filter is used to find the noise power density, and therefore the total noise power in any desired bandwidth. The FFT program will also find the total signal plus noise power, and the SNR may be obtained for the block of data from which the samples were taken.

Calibration of the digital noise filter is required for each different sample rate, data rate, or frequency resolution used in an analysis. The calibration is essentially a program input which determines the frequency width and location of the filter, and is easily determined from the data being analysed. Excellent agreement between results obtained with the programs previously discussed and the results of the FFT program has been observed. In addition, it should be noted that the FFT program has the capability of examining as small a segment of data as desired, which is of benefit in an evaluation of burst noise or other localized problems. For this reason the FFT is a useful addition to the library of data analysis programs.

## DISCUSSION OF FUTURE WORK

Calibration and refinement of the digital techniques developed during this initial investigation will be continued, using simulated data and flight data with known bit error statistics.

Additional planned future work includes further analytical and experimental study of the telemetry receiving system, the analog tape recorder, and the analog to digital processing line. A digital model of the bit synchronizer and signal conditioner will be developed and studied. An all digital model of the data gathering and reduction process might be considered the goal.

## ACKNOWLEDGEMENTS

This report is a result of the ASEE-NASA Summer Faculty Fellowship Program. The author extends his thanks to his Research Colleague, Mr. Bernard Narrow, of the Data Processing Branch; to Mr. Richard C. Lee, who shared
his office and contributed much valuable assistance; and to the many others in the Information Processing Division, without whose cooperation and ready assistance the Program could not have been as fruitful.

## BIBLIOGRAPHY

1. Y. W. Lee, Statistical Theory of Communication. New York: John Wiley \& Sons, Inc., 1960
2. W. B. Davenport and W. L. Root, Random Signals and Noise. New York: McGraw-Hill, 1958
3. M. Javid and E. Brenner, Analysis, Transmission and Filtering of Signals. New York: McGraw-Hill, 1963
4. J. C. Hancock, An Introduction to the Principles of Communication Theory. New York: McGraw-Hill, 1961
5. C. E. Gilchrist, "Signal to Noise Monitoring". Jet Propulsion Laboratory Space Program Summary No. 37-27, Vol. IV, pp. 169-184.
6. R. B. Kerr, "On Signal and Noise Level Estimation in a Coherent PCM Channel". IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-4, July, 1966.
7. C. M. Thomas, "Maximum Likelihood Estimation of Signal-to-Noise Ratio". USCEE Report, February, 1967. (AD 460 835)
8. C. M. Thomas and R. M. Gagliardi, "Data Rejection in PCM Telemetry Using SNR Estimation". 1967 National Telemetry Conference Jounal, pp. 235-239.
9. W. C. Lindsey, "Phase-Shift-Keyed Signal Detection with Noisy Reference Signals". IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-2, July, 1966.

# PRECEDING PAGE BLANK NOT FLLMED. 

COMPUTER SYSTEMS FOR
REAL-TIME APPLICATIONS
by
David F. Nyman*


#### Abstract

The STADAN tracking sites are currently operating with very little automatic control of the equipment involved with tracking a satellite. The Advanced Development Division at Goddard Flight Center is currently studyng the possibility of using a computer to control the tracking station equipment. This report will describe the functions of a computer and its associated monitor program with regard to real-time operation. Specific references will be made to the SIGMA 5 computer, a product of Scientific Data Systems, Inc., which has been proposed for this automation task.


## INTRODUCTION

This report will attempt to demonstrate how the hardware and software capabilities of the SIGMA 5 computer can be utilized to develop a real-time control system for tracking station automation. When considering software capabilities, it should be emphasized that this is a preliminary report. The MONITOR program for the SIGMA 5 is currently undergoing product evaluation by SDS personnel who are independent of its developers. Consequently, SDS systems applications personnel do not yet know all the nuances of the BATCH MONITOR. However, from experience with other SDS monitor programs and an investigation of the preliminary BATCH MONITOR manual, the probable way that it will operate was determined.

## REAL-TIME SYSTEMS

In order to explore the real-time capabilities of the SIGMA 5 computer, it is neccessary to define the characteristics of a real-time system. Basically, a real-time system contains an evaluation and control center (the computer) which

[^5]responds to conditions existing within the system's environment. The control center performs processing based on these conditions such that commands to control the system are generated. To perform this task, several extensions of normal computer batch processing become necessary.

## Multiprogramming

Since an environment is being controlled, it is necessary to have fast response to changes in the environment. This requirement places a great emphasis on the need for rapid input-output (I/O) to the computer. Further, there are situations when this requirement results in having more than one program share the computer.

During a peak load, many messages may be seeking the use of the computer. The time between the arrival of messages may become so small that the processing time for the various messages may be greater than the inter-arrival time. It then becomes necessary to process the messages in parallel, i.e., have more than one program share the computer.

Multiprogramming can be accomplished because the computer's functional units are able to operate independently and simultaneously. For example an input operation can be carried on simultaneously with a central processing unit (CPU) operation. If a program calls for data from an input file, the program may not be using the CPU while the information is being transferred to memory. Therefore, the CPU can be used by a second program.

## Interrupts

While multiprogramming allows programs to share the computer through utilization of overlapping functions it becomes necessary in real-time systems for some messages to seize control of the computer regardless of which function the computer is performing. Thus, an interrupt control feature which seizes the computer and loads the correct program for message processing must be included in a real-time system.

The following factors must be provided for effective interrupt control:
A. Some priority measure must be attached to each message so that the computer processes the most urgent request.
B. The status of the program currently in control of the CPU must be saved when an interrupt occurs.
C. The subprogram needed to process the interrupt must be identified and loaded from a file into high speed memory if it is not already present in memory.
D. On occurrence of an interrupt of lower priority than the program currently in control of the CPU, the interrupt message should join a queue for later processing.
E. At the end of interrupt processing, control must be returned to the program that was current at the time of the interrupt.
F. Interrupts should be able to be triggered (initiated) from several devices, e.g., analog to digital converters, input-output channels, the central processing unit, and timers.

## Memory Protection

Having several programs which are being processed in paralled on the computer results in a problem of interaction which must be avoided. One program must be prevented from referring to memory locations used by another program. Such interaction would obviously impair the validity of either program's results.

Further, if programs are to use a common subroutine, then the subroutine's logic should be such that intermediate results are saved when a program using the common routine is interrupted. When control is returned to the interrupted program, the intermediate results are restored and processing can continue in the common routine from the point in the logic at which the interrupt occurred. This type of programming makes the common routine re-entrant and is accomplished by keeping the routine's instruction areas separate from the associated data area. When an interrupt occurs, the contents of the data area can be placed in temporary storage and the data area can be initialized for processing of the interrupting program. When the interrupt processing is complete, the temporary storage is loaded into the data area prior to resumption of processing on the interrupted program.

## Relocatability

The scope of the system defines another characteristic of real-time computer applications. If all programs necessary for control can be resident in core at the same time, then, no provision for relocation of programs has to be made. In larger systems, programs may be called from file memory only
when they are needed for processing certain messages. Since many combinations of programs may occupy core memory at different times, it is necessary for programs to be relocatable, i.e., different memory locations are used for storing the programs with each loading. Further, if memory space is at a minimum, it may become necessary to section a program such that the only logical portions of the program which are currently being executed are in memory. The rest of the program is in library file storage and will be entered into core storage when needed for processing.

$$
;
$$

## Timers

The most catastrophic occurrence in a real-time system is for a program to gain control of the computer and either go into a loop or experience a hardware failure such that the program never finishes its task. This results in locking out of all other processing tasks. To avoid this, the real-time system is provided with an internal timer which can be set so that no program can maintain control over the central processing unit in excess of the specified time.

Another type of timer can be used to initiate interrupts at specified intervals. This interrupt could cause a sequence of programs to be initiated at these time intervals and thus, give the affect of cycling.

## III. THE SIGMA 5 COMPUTER

The SIGMA 5 computer is one manufacturer's attempt at answering the requirements for a real-time system. Realizing the rapid response time necessary in real-time systems, computer hardware has been designed to meet these requirements.

## Input-Output

Since a variety of messages may desire access to the computer concurrently, a multiplexor input-output processor (s) (MIOP) is available which will handle up to 32 input-output devices simultaneously. A maximum I-O rate of 500,000 bytes per second can be handled by each MIOP. The NTTF contract calls for two MIOPs.

A selector input-output processor (SIOP) is available but not specified in the NTTF contract. This processor is designed to quickly handle large amounts of data in a single burst. One input-output processor may be selected from as many as thirty-two devices. The transfer rate is approximately $4,700,000$ bytes per second. This type of input-output is excellent for file processing but seems to have little application at the tracking station.

## Working Memory

The core memory for the SIGMA 5 computer is organized in modules. Each module is capable of handling $4,8,12$ or 16,000 words of memory. A maximum of four modules can be attached to the system. The NTTF contract calls for two modules, each with 16,000 words of memory.

Each memory module has six-way access. This allows the module to be directly attached to both the central processing unit (CPU) and the input-output processors (IOP). A particular module can only be accessed by one device at a time; however, two modules could be active concurrently by having one attached to the CPU and the other attached to an IOP. Thus, the computer could be performing an input or output operation using one memory module while simultaneously performing calculations using information from the other module.

By interleaving the addresses of memory words on the two modules, further simultaneity of processing can be achieved. Interleaving refers to the technique of assigned odd addresses in one memory module and even addresses in the other module.

When a program is being executed, segments of the program logic may be executed in a sequential fashion, i.e., one instruction will be retrieved from an even location and executed and the next instruction will be retrieved from an odd location and executed, the next instruction from an even location, etc.

The first consequence of interleaving is that, when the program is executing sequentially, the IOP would never be locked out of a memory module for a time greater than the cycle time.* This is true since all the even core addresses are on one memory module and all the odd core addresses are on the alternate memory module. Thus the CPU is accessing alternate memory modules when processing the program and the idle memory module is available for IOP selection.

[^6]The execution of some programming instructions requires an alteration of core memory. Storing data from a register into core memory is an example of such an instruction. If the address of the memory location of where the data is to be stored is found in one memory module and the address of the next program instruction is in another module, interleaving increases processing speed since data can be stored while the next instruction is simultaneously retrieved and brought to the CPU.

## Real-Time Capabilities-Interrupts

The SIGMA 5 computer uses interrupts to initiate its real-time capabilities. These are arranged such that when an interrupt is triggered, the computer automatically executes the instruction stored at a unique core memory location associated with the interrupt level. Thus, by storing a program control instruction (such as an "Exchange Program Status Double-word" instruction - XPSD) at each of these interrupt locations, the routine to handle each particular interrupt can seize control of the CPU.

It was observed when describing real-time systems that priorities must be assigned to messages. This is accomplished through hardware on the SIGMA 5. Each interrupt level has a unique core location associated with it and follows the rule; the lower the core location assigned to the interrupt level, the higher the priority that is associated with the interrupt.

Interrupts can be classified as internal and external. Internal interrupts are triggered by such things as expiration of an internal interval timer, by a program request for input or output, etc. External interrupts are triggered by messages from devices on the MIOPs. Thus, interrupt levels over which the programmer has control can be grouped into clock, I-O, and external. On the SIGMA 5 computer specified in the NTTF contract, there are two levels of interrupts associated with the clock group, two levels associated with the I-O group, and thirty-two levels associated with the external group.

The interrupt level can exist in any of several conditions during the execution of a real-time job. This allows assignment of priorities to vary somewhat even though they are tied to a fixed core location. Interrupt states can be defined using the following terms:

1. A group of interrupts may be inhibited or active.
2. An interrupt level may be enabled or disabled.
3. In addition to being either enabled or disabled, an interrupt may be in one of four mutually exclusive states; disarmed, armed, waiting, or active.

When an interrupt level is enabled, the interrupt is waiting its turn for time on the CPU to execute the program associated with the interrupt. When it will be processed will depend upon its priority level. If an interrupt is disabled, it is removed from the priority sequence and must be enabled by a program instruction before it can gain control of the computer. When a particular interrupt level is disabled a lower priority interrupt is permitted to gain control of the computer.

When an interrupt level is disarmed and a signal is received on that level, no access to the computer is permitted and no record is kept of the signal that attempted to trigger the interrupt. This mode of operation could be used during testing of single programs so that there would be no interference with the test from outside sources.

An interrupt level can be armed through program commands such that signals to that interrupt level will be accepted and the level will move to the waiting state.

Once in the waiting state, the interrupt level must satisfy four conditions before the program associated with the interrupt level gains control of the CPU . These conditions are:

1. The level must be enabled.
2. The CPU must be at an interruptible point in its cycle.
3. The group inhibit must be off.
4. No higher priority interrupt level is in control of the CPU or is in the waiting state and enabled.

Once the interrupt gains control of the CPU, the interrupt level is said to be active. The interrupt level remains in the active state until it is interrupted by a higher priority interrupt or until the program associated with it is completed and control is turned back to a supervisory routine. This can be done by having the program execute a "Load Program Status Word" (LPSD) or a "Write Direct" (WD) instruction at the logical end of the program. At this point the interrupt level reverts back to either armed or disarmed state.

The disarmed, armed, waiting, and active states of an interrupt level are mutually exclusive. If an interrupt level is in waiting or active state and a signal to the interrupt level is effected it is lost. There is no queuing of requests for a given interrupt level. However, through the use of the waiting state, there is a queue of interrupts which are on different priority levels.

## Real-Time Capabilities - Multi-Programming

According to Scientific Data System's personnel, there are no multiprogramming capabilities as described in Section II of this report available on the SIGMA 5 computer. It is not possible for more than one program to share the same functional units of the computer.

## Real-Time Capabilities - Computer Control

Because of external interrupts there is an overhead time for changing programs. SIGMA 5 hardware makes provision for rapid change of computer control and data areas.

The control of the computer is centralized in a series of registers which comprise 64 bits of information. These registers are referenced collectively as the program status double word. The entire contents of this double word can be saved and reinitialized on five micro-seconds by a single program instruction. Theoretically then, control of the computer can be switched from one program to another if five micro-seconds.

The registers in the status double word that are of interest for the purposes of this report are:

1) IA Introduction address

RP Register pointer
CI
II $\}$ Interrupt group inhibits
EI
MS Master/slave mode
WK Write key
The IA is a seventeen bit register which contains the address of the next program instruction to be executed. When a program is being executed, this register is increased by one each time a program instruction is retrieved from storage. Thus, the instructions are executed in sequence except when a branch in the program logic is encountered. When this takes place, the core location to which the program is to branch is transferred to the IA.

Since the IA register controls which instructions are executed, this is the register which must be changed when a new program is to be executed due to an interrupt. This change is accomplished by having an "Exchange Program Status Doubleword" (XPSD) instruction located at the core location associated
with each interrupt level. The operand of the XPSD instruction contains the location of a new program status doubleword which is associated with the interrupt. This program status doubleword is exchanged with the current program status doubleword. Thus, the IA changed and the first instruction of the program associated with the interrupt is executed.

In addition to control being switched, some provision must be made for protecting the data associated with the old program and for initializing the data areas for the new program. To a degree, this is also accomplished through the program status doubleword.

There are groups of sixteen general purpose registers contained in the CPU. The NTTF contract calls for two such groups but up to sixteen such groups are available on the SIGMA 5. These general purpose registers can be used as accumulators, index registers (seven of the sixteen registers), temporary storage, control information, counts, pointers, etc. This information is relevant to a particular program and must be saved when the program is interrupted.

Within the program status doubleword, the register pointer (RP) is used to point to the block of sixteen registers currently in use. When the program status doubleword is loaded by the XPSD instruction, a new value for the RP can be loaded which will point to a new register block. Thus, volatile data associated with the interrupted program doesn't have to be stored but can be preserved by merely having the register pointer changed. Since the execution time for the XPSD instruction is five microseconds, the entire enrivonment of the computer can theoretically be changed in this amount of time.

In reality, with the NTTF system, this five microsecond switching will not be possible if the MONITOR system is used. The MONITOR requires a register block for its own use and therefore, only one register block is available for application programs. Thus, if the entire register block is used for each application program, the register block will have to be stored each time one program is interrupted and another program is loaded. The register block will then have to be initialized for the new program.

The storage and initialization of the entire register block could be accomplished by the first routine in the program associated with the interrupt. A "push down" instruction can store the contents of the registers in a group of storage locations reserved for this purpose. Thus a queue of the contents of the registers for interrupted programs is formed. A "load register" instruction can be executed so that the contents of an initialization area are loaded into the general purpose registers. The complete switching of all sixteen registers would take approximately 36.5 microseconds.

Protection of data areas that are not in the general purpose registers but are in core memory can be accomplished by the use of memory wire protect. This allows memory to be protected in 512 word pages. A two bit hardware "lock" in the CPU is associated with each of these pages. The program status double word contains a two bit write key (WK) which is checked against the lock before core memory can be altered by a program instruction. If a violation of the write protection oceurs, the memory is not altered and control of the computer is sent to location $\mathrm{X}^{\prime} 40^{\prime}$ where a branch to a programmed procedure to handle this error should start.

## IV. MONITOR PROGRAMS

As an introduction to the real-time capabilities of the SIGMA 5 computer, the characteristics of real-time systems were described. It will help with the exposition of the BATCH MONITOR if monitor programs are also explained in general.

The monitor program for a computer is provided as a controlling program for smooth hardware operation. The program normally consists of five major sections which are the supervisor, translators, the load routine, the foreman, and the input-output control system.

## Supervisor

At least a portion of the supervisor program is in core memory at all times. This section of the monitor program provides for communication between the programmer and the computer system and for recovery from errors during program execution.

Communication between the programmer and the computer is accomplished by a series of control cards. Each program can be identified and the monitor program processing to be performed on the program can be defined by these control cards. Key words which the supervisor program recognizes initiate a chain of events. Examples of these events are:

1. Dynamic assignment of input-output devices to the files used by the application program;
2. Calling of the assembler or compiler programs into memory;
3. Calling the loader routine into memory;
4. Keeping accounting data on program translation and execution times.

Uninterrupted processing can be achieved since the supervisor handles error conditions.

The application programs are stacked together as jobs to be done by the computer. When one program's processing is completed, control is returned to the supervisor program so that the next program job can be called and started. Should an abnormal or error ending occur within the application program, a branch to the monitor supervisor routine is attempted. If the supervisor program has not been destroyed as a result of the error, the supervisor program is able to recover and move on to the next program job in the sequence.

This arrangement of the job stream of programs to be executed has given rise to the term "batch shop" processing, i.e., a batch of independent programs are grouped together for processing and the monitor's supervisor routine handles the housekeeping necessary for processing of these programs. It is an obvious extension of this processing for a scheduler routine to be programmed as part of supervisor routine. This enables the program jobs to be run to be assigned a priority level and the scheduler routine will cause the highest level priority programs to be executed first.

The scheduler should not be confused with the priority interrupt levels as described for the SIGMA 5 computer. The interrupt priority levels are associated with hardware and real-time applications. The scheduler is associated with software and batch operation.

## Translators

These programs are used to translate source language programs into object or machine code. Assembly programs are characterized by the fact that they generally translate one executable source program statement into one machine language instruction. Any computer of a significant size would have an assembly program associated with it so that programmers would not have to program the computer in machine language. The assembler program for the SIGMA 5 computer is called SYMBOL.

Compiler programs are characterized by the fact that they translate one executable source program statement into several machine language instructions. A compiler program attempts to standardize a source program language for several computers. A program written in a compiler language could be translated (with minor changes) into the machine language for several different computers. Examples of compiler type translators are FORTRAN, COBOL, ALGOL, and PL-1.

The translator programs are part of the monitor system. They would be stored on some auxiliary mass storage device such as disk storage and called into high-speed core memory only when they are needed to convert a source program into machine language.

## Loader

The loader routine is called into high-speed core memory from a mass storage device by the supervisor program when it is necessary to load machine language programs into high-speed memory. This could be done immediately after translation of the source program into machine language. This is the "translate and go" mode of operation and is excellent for testing program logic. If errors in logic are found in the program, corrections can be made to the program in source language. The program is then translated, loaded, and run with test data again. The "translate and go" mode of operation does require a rapidly executing translator so that all the computer time is not wasted on translations.

The loader routine will also load machine language programs which have been generated previously and punched on cards or stored in auxiliary mass storage. This would be the normal mode of operation when the programs are operational, i.e., the debugging phase has been completed. Having the programs in object form eliminates the time-consuming translation from source language to machine language each time the program is executed.

A major function of the loader is to bring all subroutines and library routines which were requested by the main application program into memory and link them to the application program. The loader will also perform the relocation of the routines brought into memory.

## Foreman

This monitor routine is responsible for the correct functioning of the computer during application programs for either real-time or batch processing applications. Obviously then, it must be resident in core memory at the time of execution of the programs.

In batch processing and real-time applications, this routine is responsible for the operation of the input-output devices. It gives the actual command that initiate movements of the input-output mechanisms. Further, the routine monitors the performance of device. If error occurs during the input-output
operation (e.g., a parity error when reading data), the FOREMAN routine would cause the I-O device to perform the operation again.

Occasionally programs will be "trapped" during execution. This can result from several situations. An example would be a trap resulting from a division by zero. A trap could also be caused by a floating point data word which overflows its exponent range. When traps occur, it is the responsibility of the FOREMAN routine to perform the appropriate logic associated with the trap condition.

In real-time applications, an addition function is demanded of the FOREMAN routines. The routines must also handle external and internal interrupts. The interrupt service routines and the linkage to programs connected to the interrupt level are considered part of the FOREMAN.

## IOCS

These routines are common to all input-output requests. To alleviate the tedious task of writing these routines each time a reference is made to a file of input or output data, these routines are supplied in the MONITOR package and can be referenced by a single source statement in the application program. These routines interact extensively with the FOREMAN routines for I-O control. They become an integral part of the program and thus, they are resident in core during the execution of the application program.

## V. SDS BATCH MONITOR

The BATCH MONITOR for the SIGMA 5 computer will be available in October, 1967. This monitor will have a 4,000 word resident core load. The supervisor, translator, IOCS, and loader sections will be provided. The programs necessary for the real-time FOREMAN control will have to be written by the user. These real-time capabilities are termed to be FOREGROUND programs. Batch processing can be accomplished when the SIGMA 5 is not in the real-time mode. This batch processing is termed to be BACKGROUND programming.

It will have to be determined whether the foreground programs to handle various equipment interrupts will need to be resident in core or whether they can be stored in disk storage and called into core only when needed due to a particular interrupt. This decision will depend upon the quantity and size of programs necessary to perform all requirements for the NTTF station, the size of computer memory, and the speed of response to various interrupts.

Further, a basic decision must be made concerning the language in which the foreground tasks are to be programmed. This will have an effect on storage and timing requirements. The remainder of the report will discuss the advantages and disadvantages of the METASYMBOL and FORTRAN languages and how the interrupt routines are made part of the FOREMAN routine in the monitor package.

## Meta-Symbol

META-SYMBOL is the assembler program that will be available with the BATCH MONITOR in October, 1967. This language gains much of its power from the fact that user defined macro instructions are easily generated during an assembly. By the use of a "procedure directive" instruction (PROC), user defined macro instructions can be declared for use later in the executable program. The user can define a general routine to perform some processing function which is recurrent within his program. When the particular processing function is needed in the executable program, the name of the procedure directive is given in the command field (in the same way as a source language operation code would be given) and the field elements to be used in the procedure are supplied as arguments. This source coding causes machine language instructions for the routine to be generated in line with the machine language for the rest of the program. The generated machine code will have the correct addresses for the specified field elements.

META-SYMBOL also has the advantage of being able to reference monitor subroutines which are supplied with the BATCH MONITOR. These routines are routines which are helpful to the real-time systems programmer. They include routines that: 1) can set interval timers so that a fixed cycle of operations may be set in progress due to a timer interrupt: 2) provide for dynamic allocation of interrupt levels: 3) control reentrance to subroutines by moving context data from the subroutine to a dynamic storage area and returning the data to the subroutine after the higher priority interrupt has been serviced. These features and others are supplied with the BATCH MONITOR and are resident in core. A reference to them by a CAL instruction within the source program causes a branch to the monitor routines to be generated.

By choosing to program the foreground routines in META-SYMBOL, the user would gain the advantage of being able to begin writing and debugging routines immediately. The advantages of procedure directives would enable good systems programmers to provide a powerful assembler. MONITOR subroutines offer a convenience in dealing with real-time system demands but should be investigated in detail to see if they are efficient enough for the time requirements of the NTTF tracking system.

The disadvantage of META-SYMBOL as a programming language for the tracking programs is the fact that it ties the programs to the SIGMA 5 computer. Should another computer be specified for other tracking sites, then the programs would have to be re-written. For this reason, the advantages of compiler level programming for the foreground programs must be investigated.

## FORTRAN

The delivery shedule for the FORTRAN compiler leaves something to be desired when a real-time application is being considered. FORTRAN is scheduled to be delivered in four versions beginning in November, 1967 and ending in January, 1969. The special features needed to make FORTRAN easily connect with a real-time operating system will not be available until the January, 1969, version. The most glaring omission from the earlier versions will be the lack of a reentrant feature for subroutines. The monitor's capabilities cannot be utilized for this feature since no communication between the generated user program and the BATCH MONITOR routines is supplied as it is with the META-SYMBOL assembler.

Of course, the user can write his own routines so that any subroutine can become re-entrant. Upon being interrupted, it would be the responsibility of the FOREMAN program to save the program status double word, the current general purpose registers, the contents of location $\mathrm{X}(4 \mathrm{~F})$ and any volatile data that might be affected by the program servicing the interrupt. The last requirement would refer to any common user written subroutines that service interrupts, and FORTRAN library programs such as SINE, COSINE, etc. Since earlier versions of FORTRAN will not have separate areas for data and instructions in the library subroutines, it may be necessary to duplicate library routines and include some type of software pointer in the user written FOREMAN routine so that an available library routine for a particular function can be assigned to each unique interrupt message.

The obvious advantage of a compiler language over an assembler language is the ease of writing and debugging the program. In this case, some of this advantage is lost because of the delivery dates of the real-time FORTRAN compiler. On the other hand, more compact code is generated when good systems programmers use an assembler language as opposed to a compiler language. A study must be made to determine whether the real-time control program's memory requirements are loose enough to generate code by FORTRAN programming. Further, the scope and consequences of the project must be determined. If there is a high probability that the final real-time control package will find its way into all STADAN stations, then it would be convenient to have
the programs in FORTRAN so that the control computer can be rebid. Finally, FORTRAN coding offers the advantage of being easily changed when the need arises due to various tracking experiments that will be developed in a testing and training facility.

## System Generation

Regardles of which language is used to generate the foreground tasks associated with real-time control, they must be added to the delivered BATCH MONTTOR system becoming a permanent part of the resident core load and/or an addition to the portion of the monitor that resides on the disk during operation. These new routines would be added to the monitor during a systems generation computer run.

It will be recalled from the section on monitors in general that the programmer can communicate with the monitor through a series of control cards. Thus the monitor is able to expand itself by recognizing the control cards and adding the programs that follow to itself.

The procedure to expand the monitor by attaching routines to interrupt locations would be as follows. The routines themselves would be in a source language (META-SYMBOL or FORTRAN). These would be translated to machine language by the appropriate assembler or compiler program. The machine language programs would be stored on an element output (EO) file such that each had a unique name and each was a unique object module.

The load routine which was described in the above section would now be used to create a new monitor by loading each object module into memory and attaching it to an interrupt location.

Specifically, the ! LOAD control would be used. The element output file which was created during the translation phase above would now be referred to as the element input (EI) file by the appropriate parameter on the ! LOAD card. Each object module that was to be loaded from the EI file would be specified in the parameter argument list. Other files such as the System Library (SYSLIB) could be searched to have certain programs contained thereon added to the new monitor.

Another control card(s) following the ! LOAD card would attach the object modules to a given interrupt location. The : INTR card specifies the name of the object module and the hexadecimal address of the interrupt with which it is to be associated. The function of this card is to load the XPSD instruction
in the correct memory location for the interrupt level. Thus, during program execution, an interrupt at this level will send control to the object program to handle the interrupt.

## VI. CONCLUSION

It is difficult to determine exactly how each instruction or option within the BATCH MONITOR control language will operate or what its effect will be when used in conjunction with other control instructions without having the benefit of a computer system to test combinations. Hopefully, this report does clarify some of the underlying principles behind real-time control systems and that these principles will serve as a guide when actual writing and testing of the necessary programs becomes a reality.

## BIBLIOGRAPHY

Flores, Ivan. Computer Software: Programming Systems for Digital Computers. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1965.

Martin, James. Programming Real-Time Computer Systems. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1965.

SDS SIGMA SYMBOL and META-SYMBOL Reference Manual.
SDS SIGMA 5 Computer Reference Manual.

# PRECEDING PAGE BLANK NOT FILMED. 

# ON THE USE OF ADAPTIVE TECHNIQUES TO IMPROVE ANTENNA SERVO ACCURACY 

by

H. J. Perlis*

## SUMMARY

The basic antenna tracking system is analyzed under various conditions. The study suggests some desirable relationships to maintain within the framework of the existing configuration. In order to improve performance under different tracking environments an executive form of adaption is recommended. This employs available external information in a scheme to generate an adjustment vector in the tracking servo system.

## INTRODUCTION

Under the sponsorship of the 1967 ASEE-NASA Faculty Fellowship program, the author initiated a study to improve the tracking accuracy of existing servo systems. Since the systems considered are currently operational and various large antenna systems have been operating for over 10 years, there is a limited amount of design and operating data available. ${ }^{1,2,3}$ Hence, the work started with a review of the readily available design and operating manuals and reports. This was combined with visits and talks with tracking station personnel of the STADAN-Rosman I and the MSFN-Goddard NTTF facilities.

Although it is claimed that sophisticated simulation and analysis schemes are available, quite reasonable analytical results are obtained with the use of a mixture of simplified linear lumped models, Bode diagrams, and some "rule-of-thumb" folklore. This is made possible by a combination of state-of-the-art specifications, careful system alignment and calibration, the large masses in the system, and the use of highly accurate sensors and powerful hydraulic servo motors. The delivery schedule, the low-pass nature of the system, and the lack of reliable numerical values for stochastic parameters makes the employment of elementary design and error analysis methods understandable.

[^7]The servo system design in auto-track mode was reviewed. The nature of the sources of auto-track error was analyzed. The objective of this work was the development of ideas for an executive-type adaptive adjustment schedule. Within the basic framework of a fixed configuration a set of adjustable parameters was selected for each axis and a set of information parameters was selected. The information parameter vector is employed in an adjustment computer to generate the adjustment parameter vector.

## THE SYSTEM ANALYSIS OF THE TRACKING SERVO

The antenna tracking servo system studied was of the X-Y type, specifically the MSFN 30 -foot system. ${ }^{4}$ To simplify the analysis, the following discussion is addressed to a single channel. Since the channels operate independently and are of similar composition, the $X$ channel is discussed because it has the greatest load.

At low frequencies, one can treat the load as a pure inertia J (reflected back to the motor shaft), and the hydraulic servo actuator subsystem can be represented by a torque constant $K_{m}$. A linear, low-frequency model of the tracking servo is given by Figure 1. The transfer functions $G_{c_{p}}(s)$ and $G_{c_{r}}$ ( $s$ ) are the position and rate operational amplifiers with their associated networks. The motor-toantenna gear reduction is given by $N_{g}$, and the tachometer gain is shown as $K_{t}$. The input closed-loop position command is $\theta_{i}$. It includes both a desired angle signal $\theta_{\mathrm{x}}$ and a contaminating noise signal n . The system is subject to a wind disturbance. This is introduced as a torque signal $T_{w}$ which reflects back to the motor shaft.

The output antenna position $\theta_{\mathrm{A}}$ is expressed in terms of the responses of all of the inputs. Thus, in Laplace transform notation

$$
\begin{equation*}
\theta_{\mathrm{A}}(\mathrm{~s})=\left(\frac{\theta_{\mathrm{A}}}{\theta_{\mathrm{i}}}(\mathrm{~s})\right) \theta_{\mathrm{x}}(\mathrm{~s})+\left(\frac{\theta_{\mathrm{A}}}{\theta_{\mathbf{i}}}(\mathrm{s})\right) \mathrm{N}(\mathrm{~s})+\left(\frac{\theta_{\mathrm{A}}}{\mathrm{~T}_{\mathrm{W}}}(\mathrm{~s})\right) \mathrm{T}_{\mathrm{W}}(\mathrm{~s}) \tag{1}
\end{equation*}
$$

let

$$
\begin{gather*}
\mathrm{F}_{1}(\mathrm{~s}) \equiv \frac{\theta_{\mathrm{A}}}{\theta_{\mathrm{i}}}(\mathrm{~s})  \tag{2}\\
\mathrm{F}_{2}(\mathrm{~s}) \equiv \frac{\theta_{\mathrm{A}}}{\mathrm{~T}_{\mathrm{W}}}(\mathrm{~s})  \tag{3}\\
\mathrm{E}(\mathrm{~s}) \equiv \theta_{\mathrm{x}}(\mathrm{~s})-\theta_{\mathrm{A}}(\mathrm{~s}) \tag{4}
\end{gather*}
$$



Then,

$$
\begin{equation*}
E(s)=\left(1-F_{1}(s)\right) \theta_{x}(s)-F_{1}(s) N(s)-F_{2}(s) T_{W}(s) \tag{5}
\end{equation*}
$$

At this point, one might consider some linear optimization schemes. If a semi-free configuration is assumed along with stochastic signals for $\theta_{\mathrm{x}}, \mathrm{n}$ and $T_{w}$, a Wiener filter can be found to minimize the RMS value of the error. ${ }^{5}$ A more realistic approach would be to recognize the fact that $\theta_{x}$ is deterministic and can be approximated by a time function. ${ }^{6}$ Thus, the extension to the Wiener filter for a restricted class of mixed inputs could be used. ${ }^{7}$

Since the system being analyzed actually exists, it would seem more appropriate to consider a fixed configuration studied under some simplifying conditions. Thus, the following two situations will be separately analyzed:
(a) $\dot{\theta}_{\mathrm{x}}$ large
(b) $\dot{\theta}_{x}$ small

Case (a) could represent a low altitude orbiting space vehicle with a large $\theta_{y}$. In such a case the error in Equation (5) would be dominated by the first term. That is, the system error e ( $t$ ) could be approximated by the actuating error $e_{a}(t)$, and a classical deterministic analysis would follow. ${ }^{8}$

Case (b) could represent a high altitude, slowly moving space vehicle. In this case the noise and wind disturbance terms would dominate Equation (5). The problem could be treated as a fixed configuration Phillips'-type optimization involving some adjustable parameters. ${ }^{9}$

In any event one must work with Figure 1 and develop the expressions for $F_{1}(\mathrm{~s})$ and $\mathrm{F}_{2}(\mathrm{~s})$.

Thus,

$$
\begin{equation*}
F_{1}(s)=\frac{G(s)}{1+G(s)} \tag{6}
\end{equation*}
$$

where

$$
\begin{equation*}
G(s)=\frac{\frac{1}{N g} G_{c_{p}}(s) G_{r}(s)}{s} \tag{7}
\end{equation*}
$$

and

$$
\begin{equation*}
G_{r}(s)=\frac{\frac{K_{m}}{J} G_{c_{r}}(s)}{s+K_{t} \frac{K_{m}}{J} G_{c_{r}}(s)} \tag{8}
\end{equation*}
$$

Since a low frequency study is being employed which ignores the high frequency dynamics of the antenna and the hydraulic actuator, the $G_{c_{r}}$ (s) also can be treated as a constant. That is,

Let

$$
\begin{equation*}
K_{r} \equiv \frac{K_{m}}{J} G_{c_{r}}(s) \tag{9}
\end{equation*}
$$

Thus,

$$
\begin{equation*}
G_{r}(s)=\frac{K_{r}}{s+K_{t} K_{r}} \tag{10}
\end{equation*}
$$

In the auto-track mode the $G_{c_{p}}(s)$ is of the form

$$
\begin{equation*}
G_{c_{p}}(s)=\frac{K_{p}\left(s+\omega_{1}\right)}{s} \tag{11}
\end{equation*}
$$

Thus, the over-all system reduces to a simple unity negative feedback Type 2 system. The open-loop transfer function is given by

$$
\begin{equation*}
G(s)=\frac{\frac{1}{N g} K_{r} K_{p}\left(s+\omega_{1}\right)}{s^{2}\left(s+K_{t} K_{r}\right)} \tag{12}
\end{equation*}
$$

or in time-constant form

$$
\begin{equation*}
G(s)=\frac{K_{a}\left(1+s / \omega_{1}\right)}{s^{2}\left(1+s / \omega_{2}\right)} \tag{13}
\end{equation*}
$$

where

$$
\begin{gather*}
\omega_{2}=K_{t} K_{r}  \tag{14}\\
K_{a} \equiv \lim _{s \rightarrow 0}\left[s^{2} G(s)\right]=\frac{1}{\operatorname{Ng}} \frac{K_{p}}{K_{t}} \omega_{1} \tag{15}
\end{gather*}
$$

For $s=j \omega$ one can sketch the Bode asymptotic amplitude response of $G(j \omega)$.

$$
\begin{equation*}
G(j \omega)=\frac{K_{a}\left(1+\mathbf{j} \omega / \omega_{1}\right)}{(\mathbf{j} \omega)^{2}\left(1+\mathbf{j} \omega / \omega_{2}\right)} \tag{16}
\end{equation*}
$$

Figure 2 is a representation of the Bode asymptotic amplitude response for a typical stable system of this form. In general, the lead network corner $\omega_{1}$ is greater than 1 radian/sec., and for stability the $\omega_{2}$ given by Equation (14) must be larger than $\omega_{1}$. For a reasonable relative stability $\omega_{c}$ is normally located between $\omega_{1}$ and $\omega_{2}$.

Simple geometry will show that

$$
\begin{equation*}
K_{a}=\omega_{c} \omega_{1} \tag{17}
\end{equation*}
$$

Also, for maximum Phase Margin it is easy to show that the crossover frequency $\omega_{c}$ should be at the geometric mean of the two corner frequencies. That is,

$$
\begin{equation*}
\omega_{\mathrm{c}}=\sqrt{\omega_{1} \omega_{2}} \tag{18}
\end{equation*}
$$

Let

$$
\begin{equation*}
\mathrm{k} \equiv \frac{\omega_{2}}{\omega_{1}},>1 \tag{19}
\end{equation*}
$$

Then, by combining Equations (13), (17), and (18)

$$
\begin{equation*}
\mathbf{G}(\mathbf{s})=\frac{\sqrt{\mathbf{k}} \omega_{\mathrm{c}}^{2}\left(\mathrm{~s}+\omega_{\mathrm{c}} / \sqrt{\mathbf{k}}\right)}{\mathbf{s}^{2}\left(\mathrm{~s}+\sqrt{\mathrm{k}} \omega_{\mathrm{c}}\right)} \tag{20}
\end{equation*}
$$



If a normalized s-plane is defined as the $\lambda$-plane and

$$
\begin{gather*}
\mathbf{a} \equiv \sqrt{\mathbf{k}}  \tag{21}\\
\mathbf{s} / \omega_{\mathbf{c}} \equiv \lambda \tag{22}
\end{gather*}
$$

then

$$
\begin{equation*}
G(\lambda)=\frac{a\left(\lambda+\frac{1}{a}\right)}{\lambda^{2}(\lambda+a)} \tag{23}
\end{equation*}
$$

For such a system by standard geometric and trigonometric relationships one can show the constellation of poles and zeros in the $\lambda$-plane by Figure 3. From this diagram it follows that

$$
\begin{align*}
& \zeta=\frac{a-1}{2}  \tag{24}\\
& \eta=\sqrt{1-\zeta^{2}} \tag{25}
\end{align*}
$$

Since the closed-loop has a real pole at $\lambda=-1$, the system cannot be treated as a simple second order system or even a dominant pole system. That is, the dipole separation must be considered. A fractional closed-loop dipole separation taken with respect to the pole at $\lambda=-1$ is given by

$$
\begin{equation*}
\mathrm{DS}=\frac{\mathrm{a}-1}{\mathrm{a}} \tag{26}
\end{equation*}
$$

In the real frequency domain the Phase Margin can be written as

$$
\begin{equation*}
P M=2 \tan ^{-1} a-90^{\circ} \tag{27}
\end{equation*}
$$

As k increases the Phase Margin and the damping ratio $\zeta$ increase, and the dipole separation and the normalized damped natural frequency $\eta$ decrease. These variations are given in Table 1 for the interval $1 \leq a \leq 3$.

Figure 3-The Open and Closed-loop Poles and Zeros in the $\lambda$-Plane

Table 1

| k | a | $\mathrm{PM}\left({ }^{\circ}\right)$ | $\zeta$ | $\eta$ | $\mathrm{DS}(\%)$ |
| :--- | :--- | :---: | :--- | :--- | :---: |
| 1 | 1. | 0. | 0. | 1. | 0. |
| 2 | 1.414 | 19.4 | 0.207 | 0.976 | 29.2 |
| 3 | 1.732 | 30. | 0.366 | 0.930 | 42.1 |
| 4 | 2. | 36.8 | 0.5 | 0.865 | 50. |
| 5 | 2.236 | 41.8 | 0.618 | 0.785 | 56.4 |
| 6 | 2.449 | 45.4 | 0.725 | 0.686 | 58.0 |
| 7 | 2.646 | 48.6 | 0.823 | 0.565 | 62.2 |
| 8 | 2.828 | 51.0 | 0.914 | 0.405 | 64.5 |
| 9 | 3. | 53.0 | 1. | 0. | 66.7 |

The unit step response of this third-order system can be found via standard inversion tables. ${ }^{10}$ Using normalized time $\tau$, the response over the open interval $1<a<3$ is given by

$$
\begin{gather*}
-\theta_{\mathrm{A}}(\tau)=1+\frac{(\mathrm{a}-1)}{2(1-\zeta)} \mathrm{e}^{-\tau}+\frac{\mathrm{a}}{\eta}\left[\frac{\frac{1}{\mathrm{a}^{2}}-\frac{2 \zeta}{\mathrm{a}}+1}{2(1-\zeta)}\right]^{1 / 2} \mathrm{e}^{-\zeta \tau} \sin (\eta \tau+\alpha) \\
\tau \geq 0 \\
1<\mathrm{a}<3  \tag{28}\\
a=\tan ^{-1} \frac{\eta}{\frac{1}{\mathrm{a}}-\zeta}-\tan ^{-1} \frac{\eta}{1-\zeta}-\tan ^{-1} \frac{\eta}{-\zeta}
\end{gather*}
$$

where

$$
\begin{equation*}
\tau \equiv \omega_{c} t \tag{29}
\end{equation*}
$$

The response given by Equation (28) was computed for $k=1.5,2.0,2.5, \ldots$, 8.5 using a digital computer at Rutgers University. Figure 4 is a plot of the response family for $k=2,4,6,8$. The effect of the dipole separation is quite apparent. For low damping ratio the dipole separation is small, and the system resembles the simple second order case. Increasing the damping ratio lowers the peak overshoot. The effect of the zero is a tendency to increase rise time.

Figure 4-A Family of Unit Step Responses vs Normalized Time

The effect of the dipole is to "smear out" the backswing. Since the gain, dipole position, etc. are all dependent upon $k$, the system has the same "response time" (time to the first peak) $t_{p}$ which is approximately

$$
\begin{equation*}
t_{p} \approx \frac{3.0}{\omega_{c}} \tag{30}
\end{equation*}
$$

This is less than the equivalent second-order case because of the zero's effect on the leading edge response. The dipole effect on the backswing can be seen in Figure 4. It is interesting to note the approximate common crossing points at $\omega_{\mathrm{c}} \mathrm{t} \approx 1.5$ and 4.5 which are the results of the presence of the dipole. Table 2 gives a more complete listing of peak overshoot PO and normalized ( $\pm 5 \%$ ) settling time $\omega_{c} \mathrm{t}_{\mathrm{s}}(5 \%)$. The backswing smearing causes the settling time to increase with k for $\mathrm{k}>4.5$.

Table 2

| $k$ | $\operatorname{PO}(\%)$ | $\omega_{\mathrm{c}} \mathrm{t}_{\mathrm{s}}( \pm 5 \%)$ |
| :---: | :---: | :---: |
| 1.5 | 80.0 | 25.8 |
| 2.0 | 67.6 | 13.8 |
| 2.5 | 58.8 | 10.7 |
| 3.0 | 52.3 | 7.7 |
| 3.5 | 47.2 | 7.8 |
| 4.0 | 43.2 | 7.4 |
| 4.5 | 39.9 | 5.3 |
| 5.0 | 37.1 | 5.4 |
| 5.5 | 34.8 | 5.6 |
| 6.0 | 32.8 | 5.7 |
| 6.5 | 31.0 | 5.8 |
| 7.0 | 29.5 | 6.0 |
| 7.5 | 28.2 | 6.1 |
| 8.0 | 27.0 | 6.3 |
| 8.5 | 25.9 | 6.5 |

Throughout the above work, it is seen that, under the maximum Phase Margin constraint, the system is completely described by $k$ and $\omega_{c}$. From the point-of-view of maximum steady-state position error, one can write

$$
\begin{equation*}
\mathrm{e}_{\mathrm{ss}_{1_{\max }}}=\frac{\ddot{\theta}_{\mathrm{x}_{\max }}}{\mathrm{K}_{\mathrm{a}}} \tag{31}
\end{equation*}
$$

or

$$
\begin{equation*}
\mathrm{K}_{\mathrm{a}_{\text {min }}} \frac{\ddot{\theta}_{\mathrm{x}_{\text {max }}}}{\mathrm{e}_{\mathrm{ss}_{1_{\text {max }}}}} \leq \frac{\omega_{\mathrm{c}}^{2}}{\sqrt{\mathrm{k}}} \tag{32}
\end{equation*}
$$

Thus, to minimize static error one requires a high value of low frequency gain. This is reflected in expression of Equation (31). This will be the dominant consideration in the Case (a) situation.

In the Case (b) situation $E(s) \approx \mathrm{E}_{1}$ (s) where

$$
\begin{equation*}
E_{1}(s)=-\left(F_{1}(s) N(s)+F_{2}(s) T_{W}(s)\right) \tag{33}
\end{equation*}
$$

$F_{1}(s)$ is the same as before and is given by Equation (6). $F_{2}(s)$ is related to $F_{1}$ (s) by the following

$$
\begin{equation*}
F_{2}(s)=\frac{1}{G_{c}(s)} F_{1}(s) \tag{34}
\end{equation*}
$$

where

$$
\begin{equation*}
G_{c}(s)=N_{g} K_{m} G_{c_{r}}(s) G_{c_{p}}(s) \tag{35}
\end{equation*}
$$

Both $n(t)$ and $T_{W}(t)$ are ergodic stochastic signals, and, if the following assumptions are made,

$$
\begin{equation*}
\langle n(t)\rangle=\left\langle T_{w}(t)\right\rangle=\left\langle n(t) T_{w}(t+\tau)\right\rangle=0 \tag{36}
\end{equation*}
$$

the spectral density expression is given by

$$
\begin{equation*}
\Phi_{e_{1} e_{1}}(s)=F_{1}(-s) F_{2}(s)\left[\Phi_{n n}(s)+\frac{\Phi_{W W}(s)}{G_{c}(-s) G_{c}(s)}\right] \tag{37}
\end{equation*}
$$

Using the low frequency approximation of Equation (9),

$$
\begin{equation*}
\mathrm{G}_{\mathrm{c}}(\mathrm{~s})=\mathrm{K}_{\mathrm{W}} \omega_{1} \frac{\left(1+\mathrm{s} / \omega_{1}\right)}{\mathrm{s}} \tag{38}
\end{equation*}
$$

where

$$
\begin{equation*}
\mathrm{K}_{\mathrm{w}}=\mathrm{N}_{\mathrm{g}} \mathrm{~J} \mathrm{~K}_{\mathrm{r}} \mathrm{~K}_{\mathrm{p}} \tag{39}
\end{equation*}
$$

ForG (s) written in the following form

$$
\begin{equation*}
G(s)=\frac{K_{a}\left(1+s / \omega_{1}\right)}{s^{2}\left(1+s / \omega_{2}\right)} \tag{40}
\end{equation*}
$$

it follows that

$$
\begin{align*}
& F_{1}(s)=\frac{K_{a}\left(1+s / \omega_{1}\right)}{s^{3} / \omega_{2}+s^{2}+K_{a} / \omega_{1} s+K_{a}}  \tag{41}\\
& \frac{F_{1}(s)}{G_{c}(s)}=\frac{\frac{K_{a}}{K_{W} \omega_{1}} s}{s^{3} / \omega_{2}+s^{2}+K_{a} / \omega_{1} s+K_{a}} \tag{42}
\end{align*}
$$

where

$$
\begin{align*}
& \frac{\mathrm{K}_{\mathrm{a}}}{\mathrm{~K}_{\mathrm{W}} \omega_{1}}=\frac{1}{\mathrm{~N}_{\mathrm{g}}^{2} \mathrm{~K}_{\mathrm{t}} \mathrm{~K}_{\mathrm{r}} \mathrm{~J}}  \tag{43}\\
& \mathrm{~K}_{\mathrm{a}}=\frac{1}{\mathrm{~N}_{\mathrm{g}}} \frac{\mathrm{~K}_{\mathrm{p}}}{\mathrm{~K}_{\mathrm{t}}} \omega_{1} \tag{44}
\end{align*}
$$

Therefore,

$$
\begin{equation*}
\mathrm{E}_{1}(\mathrm{~s})=\frac{-1}{\left(\frac{\mathrm{~s}^{3}}{\mathrm{~K}_{\mathrm{a}} \omega_{2}}+\frac{\mathrm{s}^{2}}{\mathrm{~K}_{\mathrm{a}}}+\frac{\mathrm{s}}{\omega_{1}}+1\right)}\left[\left(1+\mathrm{s} / \omega_{1}\right) \mathrm{N}(\mathrm{~s})+\frac{\mathrm{s}}{\mathrm{~K}_{\mathrm{W}} \omega_{1}} \mathrm{~T}_{\mathrm{W}}(\mathrm{~s})\right] \tag{45}
\end{equation*}
$$

The noise can be considered to be approximately "white". The literature of wind torque disturbance on large antenna makes use of the following approximation (at least as a first order) for the spectral density ${ }^{11,12}$

$$
\begin{equation*}
\Phi_{\mathrm{WW}}(\mathrm{~s})=\frac{\mathrm{B}_{\mathrm{W}}^{2}}{\left(1-\mathrm{s}^{2} / \nu^{2}\right)}, \frac{(1 \mathrm{~b}-\mathrm{ft})^{2}}{\mathrm{rad} . / \mathrm{sec} .} \tag{46}
\end{equation*}
$$

Where $\mathrm{B}_{\mathrm{W}}^{2}$ is proportional to the fourth power of the wind velocity. Using this and the following form for the noise

$$
\begin{gather*}
\Phi_{\mathrm{nn}}(\mathrm{~s})=\mathrm{B}_{\mathrm{n}}^{2}>0, \frac{(\mathrm{rad} .)^{2}}{\mathrm{rad} . / \mathrm{sec} .}  \tag{47}\\
\Phi_{\mathrm{e}_{1} \mathrm{e}_{1}}(\mathrm{~s})=\frac{\mathrm{B}_{\mathrm{n}}^{2}\left[\frac{\mathrm{~s}^{4}}{\omega_{1}^{2} \nu^{2}}-\left(\frac{1}{\nu^{2}}+\frac{1}{\omega_{1}^{2}}\left[1+\frac{\mathrm{B}_{\mathrm{W}}^{2}}{\mathrm{~B}_{\mathrm{n}}^{2} \mathrm{~K}_{\mathrm{W}}^{2}}\right]\right) \mathrm{s}^{2}+1\right]}{\left[\frac{\mathrm{s}^{4}}{\mathrm{~K}_{\mathrm{a}} \omega_{2} \nu}+\frac{1}{\mathrm{~K}_{\mathrm{a}}}\left(\frac{1}{\nu}+\frac{1}{\omega_{2}}\right) \mathrm{s}^{3}+\left(\frac{1}{\mathrm{~K}_{\mathrm{a}}}+\frac{1}{\omega_{1} \nu}\right) \mathrm{s}^{2}+\left(\frac{1}{\nu}+\frac{1}{\omega_{1}}\right) \mathrm{s}+1\right][*]} \tag{48}
\end{gather*}
$$

Where [ $*$ ] denotes the property that this factor is identical to the other denominator factor with $s$ replaced by ( $-s$ ).

By Parseval's Theorem

$$
\begin{equation*}
\left\langle e_{1}^{2}(t)\right\rangle=\frac{1}{2 \pi j} \int_{-j^{\infty} \infty}^{j \infty} \Phi_{e_{1} e_{1}}(s) d s \tag{49}
\end{equation*}
$$

Substituting Equation (48) into (49) results in a standard Phillips' type integral with $n=4$. These integrals are tabulated. ${ }^{5}$ The resultant expression can be written in terms of a set of normalized coefficients.

$$
\begin{gather*}
\frac{2\left\langle\mathrm{e}_{1}^{2}(\mathrm{t})\right\rangle}{\nu \mathrm{B}_{\mathrm{n}}^{2}}=\frac{\frac{1-\mu \mathrm{k}}{1+\mathrm{k} \beta}-\frac{\frac{\gamma}{\beta}+\beta+1+\frac{1}{\mu}}{1+\beta}}{\frac{1+\beta}{1+\mathrm{k} \beta}+\frac{1+\mathrm{k} \beta}{\mu \mathrm{k}(1+\beta)}-\left(1+\frac{1}{\mu}\right)}  \tag{50}\\
\mathrm{B}_{\mathrm{n}}^{2}>0
\end{gather*}
$$

Where

$$
\left.\begin{array}{l}
\mathrm{k}=\omega_{2} / \omega_{1}  \tag{51}\\
\mu=\omega_{\mathrm{c}} / \nu \\
\beta=\omega_{1} / \nu \\
\gamma=\left(1+\frac{\mathrm{B}_{\mathrm{W}}^{2}}{\mathrm{~K}_{\mathrm{W}}^{2} \mathrm{~B}_{\mathrm{n}}^{2}}\right)
\end{array}\right\}
$$

To obtain a feel for the variations of Equation (50), digital computer runs were made at Rutgers for various sets of normalized coefficients. The maximum Phase Margin approximation was not made in the above work. However, to make sense, $\beta$ should be in the range $\mu / \mathrm{k}<\beta<\mu$. A convenient value for $\beta$ is the geometric mean of the two limits. That is,

$$
\begin{equation*}
\beta=\frac{\mu}{\sqrt{\mathrm{k}}} \tag{52}
\end{equation*}
$$

This is identical to the maximum Phase Margin condition.
The results of digital computer runs are tabulated in Table 3 for sets of $\gamma$ and $\mu$ values against k , using the maximum Phase Margin $\beta$. The following trends are noted:
(a) for all cases a minimum exists at $\mathrm{k}=4$
(b) with no wind the RMS error is dependent of $\mu$
(c) with wind the RMS error decreases with a decrease in $\gamma$
(d) with wind the RMS error goes down with an increase in $\mu$
(e) for a given $\mu$ and k the error decreases with a decrease in $\gamma$.

These normalized quantities called $\Gamma$ in the table must be multiplied by a factor. That is,

$$
\begin{equation*}
\left\langle e_{1}^{2}(t)\right\rangle=\frac{\nu}{2} B_{n}^{2} \Gamma(k, \mu, \beta, \gamma) \tag{53}
\end{equation*}
$$

Table 3
$\Gamma(\gamma=1$.

| k | 0.25 | 0.5 | 1 | 2 | 4 |
| ---: | :---: | :--- | ---: | ---: | ---: |
| 2 | 4.83 |  |  |  |  |
| 4 | 4.00 |  |  |  |  |
| 6 | 4.14 | $\longrightarrow$ |  |  | $\longrightarrow$ |
| 8 | 4.38 | - |  |  |  |
| 10 | 4.62 |  |  |  |  |

$\Gamma(\gamma=1.5)$

| $\mu$ | 0.25 | 0.5 | 1 | 2 | 4 |
| ---: | :---: | :---: | :---: | :---: | :---: |
| 2 | 6.14 | 5.93 | 5.54 | 5.14 | 4.93 |
| 4 | 5.22 | 5.02 | 4.67 | 4.32 | 4.11 |
| 6 | 5.47 | 5.24 | 4.87 | 4.51 | 4.28 |
| 8 | 5.82 | 5.58 | 5.18 | 4.79 | 4.54 |
| 10 | 6.19 | 5.92 | 5.50 | 5.08 | 4.81 |

$\Gamma(\gamma=2$.

| $k$ | 0.25 | 0.5 | 1 | 2 | 4 |
| ---: | :---: | :---: | :---: | :---: | :---: |
| 2 | 7.46 | 7.04 | 6.24 | 5.45 | 5.03 |
| 4 | 6.44 | 6.03 | 5.33 | 4.63 | 4.23 |
| 6 | 6.80 | 6.35 | 5.61 | 4.87 | 4.42 |
| 8 | 7.28 | 6.78 | 5.99 | 5.20 | 4.70 |
| 10 | 7.76 | 7.22 | 6.38 | 5.54 | 5.00 |

$\Gamma(\gamma=2.5)$

| $\mu$ | 0.25 | 0.5 | 1 | 2 | 4 |
| ---: | :---: | :---: | :---: | :---: | :---: |
| 2 | 8.77 | 8.14 | 6.95 | 5.76 | 5.13 |
| 4 | 7.66 | 7.05 | 6.00 | 4.95 | 4.34 |
| 6 | 8.13 | 7.45 | 6.34 | 5.24 | 4.56 |
| 8 | 8.73 | 7.98 | 6.80 | 5.62 | 4.87 |
| 10 | 9.33 | 8.52 | 7.26 | 6.00 | 5.19 |

$\Gamma(\gamma=3$.

| $k$ | 0.25 | 0.5 | 1 | 2 | 4 |
| ---: | ---: | :---: | :---: | :---: | :---: |
| 2 | 10.08 | 9.25 | 7.66 | 6.07 | 5.23 |
| 4 | 8.88 | 8.06 | 6.67 | 5.27 | 4.46 |
| 6 | 9.46 | 8.55 | 7.08 | 5.60 | 4.70 |
| 8 | 10.19 | 9.18 | 7.61 | 6.03 | 5.03 |
| 10 | 10.90 | 9.81 | 8.14 | 6.46 | 5.37 |

$\Gamma(\gamma=3.5)$

| k | 0.25 | 0.5 | 1 | 2 | 4 |
| ---: | :---: | ---: | :---: | :---: | :---: |
| 2 | 11.40 | 10.35 | 8.36 | 6.38 | 5.33 |
| 4 | 10.10 | 9.08 | 7.33 | 5.59 | 4.57 |
| 6 | 10.79 | 9.66 | 7.81 | 5.97 | 4.84 |
| 8 | 11.64 | 10.39 | 8.42 | 6.45 | 5.19 |
| 10 | 12.47 | 11.11 | 9.02 | 6.92 | 5.56 |

$$
\Gamma(\gamma=4 .)
$$

| k | $\mu .25$ | 0.5 | 1 | 2 | 4 |
| ---: | :---: | :---: | :---: | :---: | :---: |
| 2 | 12.71 | 11.46 | 9.07 | 6.69 | 5.43 |
| 4 | 11.31 | 10.10 | 8.00 | 5.90 | 4.69 |
| 6 | 12.12 | 10.76 | 8.55 | 6.33 | 4.97 |
| 8 | 13.09 | 11.59 | 9.22 | 6.86 | 5.36 |
| 10 | 14.04 | 12.41 | 9.90 | 7.38 | 5.75 |

It has been shown experimentally that $\nu$ is a constant and is independent of wind velocity. ${ }^{11}$ Typical values used are $\nu=0.11$ radians/second ${ }^{11}$ and $\nu=1.0$ radians/second. ${ }^{12}$

If $\nu=1.0 \mathrm{rad} . / \mathrm{sec}$. is assumed, a possible desirable set of nominal conditions might be:

$$
\left.\begin{array}{c}
\mathrm{k} \approx 4  \tag{54}\\
\omega_{\mathrm{c}} \geq 2.0 \mathrm{rad} . / \mathrm{sec} \\
\mathrm{PM} \approx 35^{\circ} \\
\gamma \leq 2.0
\end{array}\right\}
$$

For this set

$$
\begin{equation*}
\mathrm{K}_{\mathrm{W}} \geq\left(\frac{\mathrm{B}_{\mathrm{W}}^{2}}{\mathrm{~B}_{\mathrm{n}}^{2}}\right)^{1 / 2} \tag{55}
\end{equation*}
$$

The effective wind stiffness constant ( $\mathrm{lb}-\mathrm{ft} / \mathrm{rad}$ ) $\mathrm{K}_{\mathrm{w}}$ is given by Equation (39). If one assumes that $\omega_{1}, \mathrm{~K}_{\mathrm{p}}, \mathrm{K}_{\mathrm{t}}$, and $\mathrm{K}_{\mathrm{r}}$ are adjustable, it should be possible to obtain a compromise for each set of $B_{w}^{2}$ and $B_{n}^{2}$ in the Case (b) situation. Similarly, if the situation is of the Case (a) type an appropriate $K_{a}$ given by Equation (15) can be found to satisfy Equations (17) and (32).

## THE USE OF EXECUTIVE CONTROL

Based upon the work of the previous sections, the notion of self adaption is brought to mind. Since the tracking servo does not have to be self-contained, available outside information can be used to help to adjust the system in accordance with the objectives set up in the previous section. Such a system could use the concepts of executive control. ${ }^{13}$ That is, both measured values and prediction data can be used to develop the needed Case (a)/Case (b) decision and the associated set of required adjustments.

Typical predictions available by computation or from past experience are:

1. Minimum expected range
2. Maximum expected $Y$
3. Maximum expected X
4. Expected AGC level

Typical measurements available for up-dating:

1. RMS value of wind velocity
2. Average value of AGC

The predictions, measurements, together with nominal relationships and values for the system parameters can be used to develop the required adjustment vector. A conceptual block diagram of the executive-control adaptive tracking servo system is given in Figure 5. In the limited time available to devote to this project it was not possible to specify the complete details for the adjustment computer. However, it is felt that the proper directions and much of the basic system analysis has been performed. Thus, this work can be used as a start toward the realization of antenna tracking servo self-adaption as part of an overall station automation program.


Figure 5-The Over-all Executive Control Conceptual Block Diagram

## CONCLUSION

The overall system of an antenna tracking servo was analyzed in detail. The study was divided into two parts: (a) the case in which requirements for antenna acceleration dominate, and (b) the case in which wind gust loading and tracking receiver noise dominate. The study indicated the desirability for a frequency corner ratio $k$ or 4. The present system uses a $k=6$. This corresponds to a $\zeta \approx 0.7$, a desirable number for a simple second order system. The presence of a dipole suggests a $\zeta \approx 0.5$, which ironically happens to be the optimum value for least integral squared error in the unconstrained simple second-order system. The increasing of effective wind stiffness constant as a function of increased RMS wind velocity was found to be desirable. Throughout the work, it was felt desirable to maintain the maximum Phase Margin constraint.

## ACKNOWLEDGMENT

The writer wishes to express his appreciation to the members of the Control System Section of the Antenna Systems Branch of NASA Goddard Space Flight Center for their assistance. In particular, he wishes to acknowledge the time and effort expended by Mr. George C. Winston. In addition, the author wishes to thank Mr. Victor Tarassov, one of his Rutgers students, for his efforts in performing the digital computer calculations.

## REFERENCES

1. "Design Analysis - Unified S-Band System for Apollo Network, Volume 1, Parts I and II," Engineering Report on Contract NAS 5-9035, Collins Radio Co., Dallas, Texas, 1964.
2. J. M. Stephenson, et. al., "Design Criteria for a Large Multi-Purpose Tracking Antenna," WDL Technical Report 1368, Philco Western Development Laboratories, Palo Alto, California, 1961.
3. "Automatic Checkout and Adjustment Methods for the Servo and Control Systems of Large Steerable Antennas," Final Report on Contract NAS 510175, Sylvania Electronic Systems, Waltham, Mass., 1967.
4. "Unified S-Band Servo Subsystem (30-foot Antenna)," NASA Technical Manual ME-1529, 1966.
5. S. S. L. Chang, Synthesis of Optimum Control Systems, McGraw-Hill Book Co., New York, 1961.
6. C. H. Looney, Jr., and D. J. Carlson, "Coverage Diagrams for X-Y and Elevation-Over-Azimuth Antenna Mounts," NASA TND-2963, Goddard Space Flight Center, Greenbelt, Md., 1965.
7. L. A. Zadeh and J. R. Ragazzini, "An Extension of Wiener's Theory of Prediction," Journal of Applied Physics, 21, No. 7, pp. 645-655 (1950).
8. G. J. Thaler and R. G. Brown, Analysis and Design of Feedback Control Systems, 2nd Edition, McGraw-Hill Book Co., New York, 1960.
9. H. M. James, N. B. Nichols, and R. S. Phillips, Theory of Servomechanisms, Rad. Lab. Series, 25, McGraw-Hill Book Co., New York, 1947.
10. M. F. Gardner and J. L. Barnes, Transients in Linear Systems, Vol. 1, John Wiley and Sons, New York, 1942.
11. J. W. Titus, "Wind-Induced Torques Measured on a Large Antenna," NRL Report 5549, U. S. Naval Research Laboratories, Washington, D. C., 1960.
12. G. C. Newton,' Jr., L. A. Gould, and J. F. Kaiser, Analytical Design of Linear Feedback Controls, John Wiley and Sons, New York, 1961.
13. R. Staffin, "Executive-Controlled Adaptive Systems," D.E.E. Thesis Polytechnic Institute of Brooklyn, New York, 1959.

# RETRODIRECTIVE ANTENNAS AND THE MULTIPATH PROBLEMS ASSOCIATED WITH THEIR USE ON THE PROPOSED ORBITING DATA RELAY NETWORK 

by

P. M. Seal*

## INTRODUCTION

This report is based on a study of the characteristics of an Orbiting Data Relay Network, ODRN, designed to provide two-way communication between ground stations and target vehicles orbiting the earth. Essential background material is contained in the Final Report submitted by RCA to NASA on March 22, 1967, entitled "Orbiting Data Relay Network."

The two-way communication is to be provided by including a synchronous satellite as a link between the ground station and the target vehicle. A large part of the RCA Report is devoted to a study of the types of antennas which might be used in such a synchronous satellite, the final recommendation being that a self-focusing or retrodirective antenna array be used.

Considerable time was spent in studying the literature available on retrodirective arrays and the first part of the report deals with a bibliography of this literature along with a brief description of the highlights of each reference.

One problem in connection with the use of retrodirective arrays is the fact that the array will receive two signals from the target vehicle, one a direct signal and the other a signal reflected from the earth's surface. This problem, known as the multipath problem, is particularly serious when the target vehicle antenna is omni-directional, and the RCA report discusses this particular problem in considerable detail. A large part of the summer was spent in studying various aspects of this problem, the aim being to determine to what extent the beam steering of the retrodirective array would be affected by the reflected wave. The second part of this report deals with the results of this study and the conclusions reached.

[^8]
## BIBLIOGRAPHY ON RETRODIRECTIVE ANTENNA ARRAYS

## Books:

Microwave Scanning Antennas, edited by R. C. Hansen, Academic Press, Volume II, Array Theory and Practice, 1966.

Chapter 1, The Theory of Antenna Arrays, by R. S. Elliott.

This chapter was studied mainly as a review of antenna arrays in general, and of scanning arrays in particular. Of special interest was the section on the scanned array, which shows that by varying the relative phase angle between currents fed to the various elements in the array, a beam may be produced which scans over a wide area in space, the antenna array itself being stationary.

Volume III, Array Systems, 1966
Chapter 5, Self-Phased Arrays, by Donald L. Margerum.
This chapter covers the subject of self-phased arrays very extensively, including a detailed discussion of the principle of operation of such arrays, and a description of the various types of these arrays which have been investigated. There is also an extensive set of references at the end of the chapter.

## Periodicals:

The references listed below are all articles which appeared in the March, 1964 IEEE Transactions on Antennas and Propagation. This was a special issue on active and adaptive antennas.

Self-Phasing Array Antennas, by Skolnik and King.
This article describes the principle of operation of the self-phasing antenna, including the Van Atta Array, and discusses how noise and signals coming from more than one source affect the operation of the array. It also gives a number of applications.

Self-Focusing Array Research Model, by Sichelstiel, Waters, and Wild.
Describes the results obtained on a research model containing 25 subapertures at first, then extended to 64 subapertures. Results showed that good tracking was obtained in the presence of noise, even when the signal-to-noise ratio
was only 2 db . Good results were also obtained when signals from two sources separated by a few degrees were picked up by the array, even when the power in one signal was 12 db below the power in the other.

Electronically Adaptive Antenna Systems, by Ghose.
A feature of this article is that it analyzes in some detail the effect of noise on a self-focusing antenna on a statistical basis. It also emphasizes the importance of having three mutually orthogonal antenna arrays on space vehicles whose attitudes can change in an unpredictable manner, thus making the antenna system insensitive to the direction of arrival of the signal.

Retrodirective Array Using the Heterodyne Technique, by Pon.
Describes a 4-element array using a single mixer and local oscillator and shows its advantages over the Van Atta Array.

An Active Retrodirective Array for Satellite Communications, by Andre and Leonard.

One feature of this article is that it discusses the effects of phase errors introduced by the amplifiers, mixers and transmission lines interconnecting the amplifiers. A phase error of up to $\pm 30$ degrees produces a "worst case" loss of 1.2 db in the beam output. No mention was made, however, of the possible effect on the direction of the beam that might be caused by a phase error, especially if the error were in the same direction for all elements.

Also of interest was the statement that it was possible to "fabricate a circulator coupled tunnel-diode amplifier weighing approximately one ounce at S-band." It was also claimed that less than one watt of primary power was required for a 200-element array.

Spherical Retrodirective Array, by Rutz-Phillip.
This article emphasizes the fact that a spherical array will accept a wave from any direction in space, amplify it and return it in the direction from which it came, the amplitude of the reradiated signal being independent of the direction. It is therefore ideal for use on satellites.

This array makes use of a tunnel-diode frequency converter, and one statement of interest is that "one diode functions simultaneously as amplifier, frequency translator, modulator and phase inverter." No circulator is required to separate the incident wave from the reradiated wave.

The elements of the array do not need to be equally spaced, but the spacing between elements should be below one wavelength. Each element of the array should be circularly polarized, but the orientation of the axes can be arbitrary.

> A Satellite Data Transmission Antenna, by Belfi, Rothenberg, Schwartzman, Tilley and Wills.

This article discusses the advantages of using a retrodirective antenna array in a data-collecting satellite such as TIROS IV. Upon command from a ground station, such an antenna would transmit its information back to the ground station through a beam automatically pointed in that direction. The conclusion is that "the hybrid-phasing-matrix antenna is the optimum system for this satellite configuration in terms of performance, weight, power drain and reliability." Several advantages are listed, one being that "It will not respond to jamming or an incorrectly coded signal."

## A Phase-Locked Receiving Array for High-Frequency Communications Use,

 by Svoboda.One conclusion from this article is that "a phase-locked array can, under most circumstances, give interference rejection comparable to that of a conventionally phased array." However, a disadvantage is that its operation is affected by interference and noise in the phase detector. Of possibIe interest in connection with the multipath problem is Appendix II, entitled "Effect of Interfering Coherent Monochromatic Signal on Performance of Phase-Locked Array."

## Satellite Communications Relay System Using a Retrodirective Space Antenna, by Gruenberg and Johnson.

The array system described is considered to be quasi-passive in the sense that very little power is required for its operation. The total power requirement for a 10,000 -element array is claimed to be only one-half of a milliwatt, and "could be derived entirely from the ground." Several modulation techniques are described and a fairly extensive set of references is included.

## THE MULTIPATH PROBLEM

As mentioned in the introduction, a large part of the RCA Final Report on ODRN is devoted to multipath effects. These effects are especially serious in the emergency mode, where the antenna on the target vehicle is omnidirectional. Under such conditions the indirect signal reaching the spacecraft after being reflected from the earth's surface can be quite large compared with the direct signal. In the RCA report, several pages of Section 4.0, particularly pages 4-38 to $4-46$, and all of Appendix D, are devoted to a discussion of this problem.

One of the initial goals in this summer's study was to determine the effects of multipath propagation on the beam steering of the retrodirective array. While this goal was not completely realized, by any means, a few things were accomplished, and the remainder of this report is devoted to these accomplishments.

## Maximum Angle Between Direct and Reflected Waves.

One of the properties of the retrodirective antenna array is that when it receives signals from two different directions, it will retransmit each signal in the direction from which it was received. In the case of multipath propagation, one of these signals is the direct signal and the other, the reflected signal.

All derivations contained in the RCA report assume that the angle between the paths taken by the direct and the reflected waves is negligibly small. The RCA report also shows that, in general, the reflected wave is scattered and arrives at the spacecraft from a wide area of the earth's surface. However, it also shows that the energy in the reflected wave tends to be concentrated in the components arriving from the specular point, that is, the point at which the angle of incidence is equal to the angle of reflection.

On this basis, a study was made to determine the maximum angle between the path taken by the direct wave and the path taken by a reflected wave arriving from the specular point. Figure 1 shows the multipath geometry involved, the angle being referred to as $\Delta \theta$.

Equations were derived from this figure, as shown on next page, and $\Delta \theta$ was calculated as a function of the angle of incidence $\theta$, for the case of a target vehicle assumed to be 400 miles above the earth.

Results are shown in the lower graph of Figure 2. For this height, the maximum value of $\Delta \theta$ is about 1.6 degrees and occurs when $\theta \doteqdot 60$ degrees, or when the target aspect angle $\Theta \doteqdot 70$ degrees.


Figure 1-Multipath Geometry.
!


Figure 2-Curves for Obtaining Maximum $\Delta \theta$ and Maximum Fading Rate.

Although calculations were not carried out for other target heights, it is probable that the maximum value of $\Delta \theta$ |will be roughly proportional to target height.

Equations used for calculating the angle $\Delta \theta$ in terms of $\theta$ were obtained from the trigonometrical relations of Figure 1 and are listed below:

From triangle OTP:

$$
\begin{align*}
\operatorname{Sin} \alpha_{1} & =\frac{a}{a+r} \operatorname{Sin} \theta  \tag{1}\\
\phi_{1} & =\theta-\alpha_{1}  \tag{2}\\
r_{1} & =(a+h) \frac{\operatorname{Sin} \phi_{1}}{\operatorname{Sin} \theta} \tag{3}
\end{align*}
$$

From triangle OSP:

$$
\begin{align*}
\operatorname{Sin} \alpha_{2} & =\frac{a}{a+d} \operatorname{Sin} \theta \\
\phi_{2} & =\theta-\alpha_{2} \\
\mathbf{r}_{2} & =(a+d) \frac{\operatorname{Sin} \phi_{2}}{\operatorname{Sin} \theta} \tag{6}
\end{align*}
$$

From triangle OTS:

$$
\begin{equation*}
r^{2}=(a+h)^{2}+(a+d)^{2}-2(a+h)(a+d) \operatorname{Cos}\left(\phi_{1}+\phi_{2}\right) \tag{7}
\end{equation*}
$$

From triangle TPS:

$$
\begin{equation*}
\operatorname{Sin} \Delta \theta=\frac{r_{1}}{r} \operatorname{Sin} 2 \theta \tag{8}
\end{equation*}
$$

$$
\begin{equation*}
\Theta=2 \theta+\Delta \theta-\alpha_{1}=\theta+\phi_{1}+\Delta \theta \tag{9}
\end{equation*}
$$

## Differential Doppler Frequency or Fading Rate, $\Delta \mathrm{f}$.

The fading rate is defined on page 4-43 of the RCA report as the difference between the doppler shift in the direct signal and the doppler shift in the reflected signal. The report states that "the differential doppler is typically about 1 kHz for orbit heights up to 400 miles." No equations were shown, however, and no explanation was given to show how this figure was determined.

The equations derived for obtaining $\Delta \theta$ can be easily extended to obtain equations for calculating $\triangle f$, and these equations are included below.

From these equations, calculations were made to determine $\Delta f$ as a function of $\theta$ for the case of a target 400 miles above the earth traveling at a velocity of 5 miles per second in the plane made by the target, the spacecraft and the center of the earth. The frequency of the signal transmitted from the target was assumed to be 2000 MHz .

Results of these calculations are shown in the upper graph of Figure 2. For the example chosen, the maximum value of $\triangle \mathrm{f}$ is about 8.2 kHz and occurs when $\theta$ is about 55 degrees, or $\Theta$ is about 63 degrees. Note that when the target is directly below the spacecraft, or $\theta=0, \Delta f=0$ since the target's component of velocity in the direction of the spacecraft is zero. When $\theta=90$ degrees, the line between the target and the spacecraft is tangent to the earth, making the doppler shift in the direct signal just equal to the doppler shift in the reflected signal. Hence $\Delta f=0$ at this point also.

Equations used for calculating $\triangle \mathrm{f}$ in terms of $\theta$ were obtained by making use of Figure 1 and the equations for calculating $\Delta \theta$. These are summarized below:

$$
\begin{align*}
& \frac{d r}{d t}-\frac{d\left(r_{1}+r_{2}\right)}{d t}  \tag{10}\\
& \Delta f=\frac{}{c} x f,
\end{align*}
$$

where

$$
\begin{aligned}
f & =\text { frequency of the transmitted signal } \\
\mathbf{c} & =\text { velocity of propagation of signal in free space } \\
& =\text { velocity of light. }
\end{aligned}
$$

Assuming that the target $T$ in Figure 1 is traveling at a velocity $\nu$ parallel to the earth and in the plane of the figure,

$$
\begin{gather*}
\frac{d r}{d t}=\nu \operatorname{Sin} \Theta,  \tag{11}\\
\frac{d\left(r_{1}+r_{2}\right)}{d t}=\left(\frac{d r_{1}}{d \theta}+\frac{d r_{2}}{d \theta}\right) \times \frac{d \theta}{d \phi} \times \frac{d \phi}{d t}, \tag{12}
\end{gather*}
$$

where

$$
\begin{equation*}
\frac{\mathrm{d} \phi}{\mathrm{dt}}=\frac{\nu}{\mathrm{a}+\mathrm{h}} \tag{13}
\end{equation*}
$$

and

$$
\begin{equation*}
\frac{\mathrm{d} \theta}{\mathrm{~d} \phi}=\frac{1}{\frac{\mathrm{~d} \phi_{1}}{\mathrm{~d} \theta}+\frac{\mathrm{d} \phi_{2}}{\mathrm{~d} \theta}} \tag{14}
\end{equation*}
$$

Using equations 1 and 2, and differentiating with respect to $\theta$,

$$
\begin{equation*}
\frac{\mathrm{d} \phi_{1}}{\mathrm{~d} \theta}=1-\frac{\mathrm{a}}{\mathrm{a}+\mathrm{h}} \times \frac{\operatorname{Cos} \theta}{\operatorname{Cos} \alpha_{1}} \tag{15}
\end{equation*}
$$

Similarly, using equations 4 and 5,

$$
\begin{equation*}
\frac{\mathrm{d} \phi_{2}}{\mathrm{~d} \theta}=1-\frac{\mathrm{a}}{\mathrm{a}+\mathrm{d}} \times \frac{\operatorname{Cos} \theta}{\operatorname{Cos} \alpha_{2}} \tag{16}
\end{equation*}
$$

By differentiating equations 3 and 6 with respect to $\theta$,

$$
\begin{equation*}
\frac{\mathrm{dr}}{\mathrm{~d} \theta}=\frac{(\mathrm{a}+\mathrm{h})}{\operatorname{Sin} \theta}\left[\operatorname{Cos} \phi_{1} \times \frac{\mathrm{d} \phi_{1}}{\mathrm{~d} \theta}-\operatorname{Sin} \phi_{1} \operatorname{Cot} . \theta\right] \tag{17}
\end{equation*}
$$

$$
\begin{equation*}
\frac{\mathrm{dr}}{\mathrm{~d} \theta}=\frac{(\mathrm{a}+\mathrm{d})}{\operatorname{Sin} \theta}\left[\operatorname{Cos} \phi_{2} \times \frac{\mathrm{d} \phi_{2}}{\mathrm{~d} \theta}-\operatorname{Sin} \phi_{2} \operatorname{Cot} . \theta\right] \tag{18}
\end{equation*}
$$

## Fading Bandwidth and Coherent Bandwidth

The meaning of the term "fading bandwidth" and a discussion of how it differs from "fading rate" is discussed on page 4-44 of the RCA report, and also on page D-13 of Appendix D of the same report. Similarly "coherent bandwidth" is discussed on page 4-44 and at greater length on page D-20 of Appendix D. Further discussion of the relation between these two bandwidths is contained on page 4-45 of the report and again on page 4-61.

The only reason for commenting further on these two bandwidths is to call attention to the fact that the magnitudes assigned to them in the RCA report are all based on "moderate sea roughness." Yet on page D-19 of Appendix D a statement is made that the term used as a measure of sea roughness "could vary from $1 / 3$ to 3 times" the "typical" value over different terrains.

Furthermore, Equations $D-19 b$ and $D-21$ for fading bandwidth, $B_{f}$, as given on pages $D-18$ and $D-19$ of Appendix $D$, shows that $B_{f}$ varies inversely as the term $\mathrm{T} / \sigma$, whereas the equation for the coherent bandwidth, $\mathrm{B}_{\mathrm{coh}}$, given on page $\mathrm{D}-20$ shows that $\mathrm{B}_{\text {coh }}$ varies directly as the square of $\mathrm{T} / \sigma$, where $\mathrm{T} / \sigma$ may be considered as a "roughness factor." (Perhaps "smoothness factor" would be a better term, since it should be noted that the smoother the sea is, the greater is the term T/G.)

To show how $\mathrm{B}_{\mathrm{f}}$ and $\mathrm{B}_{\text {coh }}$ vary with the factor $\mathrm{T} / \sigma$, the curves of Figure 3 were drawn. Consider as an example the case of a target 100 miles above the earth, and the target aspect angle $\Theta=0$. For "normal" roughness, $B_{f}$ is approximately 25 kHz and $\mathrm{B}_{\mathrm{coh}}$ is about 15 kHz . However, under a "smoother" sea, where $T / \sigma$ is increased 3 times, $B_{f}$ drops to 8.3 kHz whereas $B_{c o h}$ increases to about 140 kHz . On the other hand, for a "rougher" sea, where T $/ \sigma$ is $1 / 3$ of the normal value, $\mathrm{B}_{\mathrm{f}}$ increases to 75 kHz while $\mathrm{B}_{\text {coh }}$ drops to 1.7 kHz . No attempt will be made here to discuss the significance of these figures, but presumably they should be taken into account in attempting to predict the quality of transmission under varying conditions.

NOTES ON A MEMORANDUM FROM P. SEALMAN OF AIRBORNE INSTRUMENT LABORATORIES TO P. BARRITT, NASA HEADQUARTERS JULY 1, 1967

This Memorandum proposes a frequency-hopping technique to take advantage of the time delay between the direct signal and the reflected signal. This technique is well described in the memorandum and a typical example of how it might be used is given. However, Equation 4 on page 5 of the memo, which is the equation used for determining the time delay, appears to be in error, mainly due to the fact that the term $h_{1}$ in the equation is assumed to be a constant independent of the grazing angle, $\psi$. This error leads to an incorrect value for the time delay when $\psi=10$ degrees, which in turn leads to an incorrect value for the number of frequencies required, N , in the example used.

A correct expression for the time delay is derived below, making use of Figures 4 and 1, as well as some of the equations already derived from Figure 1.

When the correct expression is used in the example given in the memorandum, it is found that for a target height of 200 nautical miles, or 230 statute miles, and a grazing angle $\psi$ of 10 degrees,

$$
\mathrm{h}_{0}=147 \text { statutemiles }
$$

making $t_{d}=274$ microseconds, instead of the 440 microseconds given in the example.

This in turn will make the required value of $\mathrm{N}=10$ instead of 6 and make the hopping rate $4000 \mathrm{hops} / \mathrm{second}$ instead of 2400 . It will also change the value of the bandwidth B shown at the bottom of page 8 from 2.4 MHz to 4.0 MHz , and the value described at the top of page 9 as "under one megacycle" to an actual value of 1.53 MHz .

It should be emphasized that this correction does not necessarily invalidate the use of the frequency hopping technique, but does appear to make its use a little more difficult than the example would indicate.

Derivation of Equation for Time Delay, $t_{d}$, Between
Direct and Reflected Signal.
With reference to Figure 4,

$$
\begin{equation*}
\mathrm{t}_{\mathrm{d}}=\frac{\mathrm{r}_{1}+\mathrm{r}_{2}-\mathrm{r}}{\mathrm{c}}, \tag{19}
\end{equation*}
$$



Figure 3-Curves Showing how Fading Bandwidth and Coherent Bandwidth Vary with Roughness of Sea.


Figure 4-Geometry Relating $r_{1}+r_{2}$ to $r$ and $h_{0}$ to $h$.
where $\quad t_{d}=$ time delay in seconds,

$$
\mathrm{c}=\text { velocity of light in miles } / \mathrm{sec}, \text { if } \mathrm{r}_{1}, \mathrm{r}_{2}, \text { and } \mathrm{r} \text { are given in miles. }
$$

If the distance $r$ to the spacecraft is very large compared to the target height $h, r$ and $r_{2}$ may be assumed to be parallel lines. Making this assumption, Figure 4 shows that

$$
\begin{align*}
r_{1}+r_{2}-\mathbf{r}=\mathrm{cd} & =\mathrm{Tc} \cos \theta \\
& =2 \mathrm{~h}_{0} \cos \theta  \tag{20}\\
& =2 \mathrm{~h}_{0} \sin \psi \\
\text { Hence } \mathrm{t}_{\mathrm{d}} & =\frac{2 h_{0}}{c} \operatorname{Sin} \psi . \tag{21}
\end{align*}
$$

As shown in Figure 4, $h_{0}$ is the height of the target above the tangent to the earth at the specular point $P$. Hence $h_{0}$ is a function of $\psi$.

The relation between $h_{0}, h$, and the grazing angle $\psi$ may be found from equations already derived from Figure 1, with the help of Figure 4. From Figure 1 we had

$$
\begin{equation*}
r_{1}=(a+h) \frac{\operatorname{Sin} \phi_{1}}{\operatorname{Sin} \theta}=(a+h) \frac{\operatorname{Sin} \phi_{1}}{\operatorname{Cos} \psi} . \tag{22}
\end{equation*}
$$

But Figure 4 shows that

$$
\begin{equation*}
h_{0}=r_{1} \operatorname{Sin} \psi \tag{23}
\end{equation*}
$$

Hence

$$
\begin{equation*}
h_{0}=(a+h) \operatorname{Sin} \phi_{1} \operatorname{Tan} \psi \tag{24}
\end{equation*}
$$

where, as already shown,

$$
\begin{equation*}
\phi_{1}=\theta-\alpha_{1}=90^{\circ}-\psi-\alpha_{1} \tag{25}
\end{equation*}
$$

and

$$
\begin{align*}
\operatorname{Sin} \alpha_{1} & =\frac{a}{a+h} \operatorname{Sin} \theta \\
& =\frac{2}{a+h} \operatorname{Cos} \psi \tag{26}
\end{align*}
$$

## CONCLUSIONS

The information contained in this report should not be considered as a solution to the multipath problem, by any means. In fact, it does not come close to answering the question of the effect of the reflected wave on the steering of the beam put out by a retrogressive antenna array. It does, however, contain a fairly extensive bibliography of articles on this type of antenna array. It also calls attention to some factors over and above those which are contained in the RCA report on the Orbiting Data Relay Network, factors which should receive careful attention in a continued study of the multipath problem. In this report, it is hoped that it will be of some value.

TRANSFER FUNCTIONS OF THE ROSMAN I ANTENNA STRUCTURE
by
T. G. Toridis* and Yu Chen**

The purpose of this project is to determine analytically the transfer functions associated with the structural part of the Rosman I Antenna. The antenna control system consists of various blocks which are interconnected to form a servo loop. A very simplified representation of the system is illustrated in Figure 1.


Figure 1-Simplified Representation of the Servo Loop

Each of the above blocks may be subdivided into a number of subsystems to obtain more sophisticated models representing the antenna system. If the transfer functions for each block or subsystem have been computed they may then be used to determine the characteristics of the control system. A detailed examination of the antenna structure indicates that it can conveniently be subdivided into several substructures. This is illustrated in Figure 2, on the following page:

[^9]

Figure 2-Simplified Representation of the Structural Loop

Each substructure may be replaced by one or more points or lumped masses in a simplified model representing the entire structure. The location of these points may be chosen to coincide with certain key locations of the antenna such as the drive system, the $X$ and $Y$ bearings, the reflector dish, the feedbox, etc. In this manner, results obtained by using the theoretical model can be compared with experimental values representing direct measurements obtained by means of recording apparatus mounted on the antenna structure. The points selected for this purpose are listed in Table I, below:

Table I
Lumped Mass Points

| Points 1 and 2: | X-Wheel Bearings |
| :--- | :--- |
| Points 3 and 4: | Hydraulic Drive |
| Point 5: | X-Wheel Counterweight |
| Points 6, 7, 8, 9: | Reflector Dish |
| Point 10: | Y-Wheel Counterweight |
| Points 11 and 12: | Y-Wheel Bearings |
| Point 13: | Feedbox |

The mass of the structure is lumped on each of the above 13 points. Considering a translational degree of freedom along each of the $x, y$, and $z$ directions of a Cartesian coordinate system a simplified model of 39 degrees of freedom is obtained (the $x$ and $y$ directions coincide with the directions of the $X$-wheel and $Y$-wheel shaft respectively, and $z$ represents the vertical direction). The masses associated with each degree of freedom are given in Table II ( $M=$ diagonal mass matrix in lb-sec ${ }^{2} / \mathrm{in}$.).

## Table II

Numerical Values of Lumped Masses (lb-sec ${ }^{2} / \mathrm{in}$ )

$$
\begin{aligned}
& m_{1,1}=m_{2,2}=m_{3,3}=666 \\
& m_{4,4}=m_{5,5}=m_{6,6}=666 \\
& m_{7,7}=m_{8,8}=m_{9,9}=252 \\
& m_{10,10}=m_{11,11}=m_{12,12}=252 \\
& m_{13,13}=m_{14,14}=m_{15,15}=1200 \\
& m_{16,16}=m_{17,17}=m_{18,18}=88 \\
& m_{19,19}=m_{20,20}=m_{21,21}=88 \\
& m_{22,22}=m_{23,23}=m_{24,24}=88 \\
& m_{25,25}=m_{26,26}=m_{27,27}=88 \\
& m_{28,28}=m_{29,29}=m_{30,30}=300 \\
& m_{31,31}=m_{32,32}=m_{33,33}=156 \\
& m_{34,34}=m_{35,35}=m_{36,36}=156 \\
& m_{37,37}=m_{38,38}=m_{39,39}=96
\end{aligned}
$$

In order to simulate the actual behavior of the structure by the simplified model the various masses would have to be interconnected by springs in a complicated manner. The determination of the proper connections is a rather difficult matter if a close analogy at the higher frequency range is desirable between the actual structure and its simplified model. Since the static properties of the actual structure has been previously determined by the Martin SB038 program it is desirable to use this available and realistic information somehow. This is found possible by appropriately choosing such key locations on the structure that coincide with existing grid points used in the SB038 program. A small program is written to generate from the tape which has the complete influence coefficient matrix of the structure ( $645 \times 645$ ) a $39 \times 39$ influence coefficient matrix for the simplified model used in this study. This is accomplished by applying successively
a unit load at each degree of freedom and letting the program compute the resulting deflections at all 39 points.

The influence coefficient matrix obtained in this manner is not found to be perfectly symmetrical. However, in order to be able to use a simple eigenvalue routine available in the Goddard computer library, the influence coefficient matrix has been forced into a symmetrical form. In addition, because of the fact that the diagonal terms of the matrix do not always predominate over the offdiagonal terms, it has been found necessary to re-estimate some of them to improve the conditioning of the matrix. The resulting matrix is given below in the form of successive rows, row 1 through row 39 , where $D$ is used to denote the influence coefficient matrix.

Having determined the mass and influence coefficient matrices of the simplified model, the differential equations governing its dynamical motion may be written in matrix form as

$$
\begin{equation*}
\mathrm{DM} \ddot{\mathrm{X}}+\mathrm{X}=\mathrm{D} Q \tag{1}
\end{equation*}
$$

where $X$ and $\ddot{X}$ represent the displacement and acceleration vectors, respectively, and $Q$ denotes the force vector. Premultiplying Eq. (1) by the transpose of the mass matrix, $M^{T}=M$, and letting $A=M D M$ ( $A$ is a symmetric matrix) the resulting equation is

$$
\begin{equation*}
A \ddot{X}+M X=M D Q \tag{2}
\end{equation*}
$$

## FREE VIBRATIONS

The natural frequencies and modes of vibration of the system may be obtained by setting the right hand side of Eq. (2) equal to zero, i.e.,

$$
\begin{equation*}
A \ddot{X}+M X=0 \tag{3}
\end{equation*}
$$

Let

$$
\begin{equation*}
X=W^{j} \sin \omega_{j} t \tag{4}
\end{equation*}
$$

where $\omega_{j}$ and $W^{j}$ represent the $j$-th natural frequency and mode of vibration, respectively. Substituting Eq. (4) into Eq. (3) and cancelling out sin $\omega$ t

$$
\begin{equation*}
-\omega_{j}^{2} A W^{j}+M W^{j}=0 \tag{5}
\end{equation*}
$$

ROWS OF D MATRIX


 $0 \infty$
0
0
1
0
0 $0_{0} 0$ $0.28568593 \mathrm{E}-06$
$-0.33807699 \mathrm{E}-08$
$0.13034829 \mathrm{E}-05$
$0.21709830 \mathrm{E}-09$
$0.38598579 \mathrm{E}-07$ $0.85830259 E-09$
$0.13034819 E-05$
$0.10749099 E-07$
$0.46246190 E-08$
$0.13034819 F-05$ $0.22805686 E-05$
$0.33806 .550 \mathrm{E}-09$
$-0.75605665 \mathrm{E}-07$
$0.22823656 \mathrm{E}-05$
$0.14101020 \mathrm{E}-06$
$0.13021597 E-05$
$-0.200046995-06$
$0.15 .678310 E-07$
$-0.88778333 E-06$
$-0.59545439 F-08$ $0.52964439 \mathrm{E}-09$
$0.62832950 \mathrm{E}-06$
$0.22823656 \mathrm{~F}-05$
$-0.10124600 \mathrm{E}-09$
$-0.75507332 \mathrm{E}-06$
 0.107478 A E
2
$-3.4138329 \mathrm{EE}-10$
$-0.5757249 \mathrm{E}-05$

 $-0.10353299 \mathrm{E}-08$
$-0.14093400 \mathrm{E}-09$
 00
0
10
10
0
0
0
0
0
0
0
0
0
0
0
0
0
10

 $\begin{array}{rr}0.22823656 E-05 & 0.17685989 E-09 \\ 0.47081940 E-06 & 0.22423656 E-05 \\ -0.23407556 E-06 & -0.11405499 E-05\end{array}$
$-0.21297999 E-09$
$0.28515480 \mathrm{E}-06$
$0.28515480 \mathrm{E}-06$
$-0.53468799 \mathrm{E}-09$



## 0．13669290E－09

## 0． $28515480 E-08$




## $0.85546440 E-0$ $-0.21152499 \mathrm{E}-0$ $0.28515480 \mathrm{E}-0$

$0.17726760 E-09$
$-0,20256 ? 00 E-06$

 －3．23707565E－06 0．28515480E－06

## 0 1 0 0 0 0 0 0 $n$ $n$ $n$ 0 0

 $0.13021597 E-05$
$-0.22554600 t-05$



$0.96058020 \mathrm{E}-10$
$-0.41222665 \mathrm{E}-06$
$0.50612459 \mathrm{E}-09$
$-0.16898400 \mathrm{E}-08$
 $0.28568593 E-06$
$0.40808939 E-08$
$0.13021597 E-05$
$0.10851800 \mathrm{E}-08$
$0.22249410 \mathrm{E}-08$



 -0.2434947 FE－0



 8
0
1
0
0
0
7
0
0
0
0
0
0 onino -399520
-366562
-39695 C
-366602
$-36550 \varepsilon$
$0.69626199 \mathrm{E}-06$
$0.28332230 \mathrm{E}-09$
$0.46966666 \mathrm{E}-06$
$0.22805686 \mathrm{E}-05$
$0.49847219 E-09$
$0.28568593 E-06$
$0.51471379 E-07$
$0.43541569 E-08$





80
0
1
0
0
0
0
6
0
0
0
0
0
0
0 0
0
1
1
1
0
0
0
0
0
0
0
0
0
0
0
0
0
0







岂



| $\sim 0$ |
| :--- |
| 00 |
| 1 |
| $u$ | $0.16827480 \mathrm{E}-05$

$0.22459099 \mathrm{E}-08$
$0.46875809 \mathrm{E}-08$
$0.56091599 \mathrm{E}-06$
$-0.63728899 \mathrm{E}-07$







## $\circ$ 0 $\dot{1}$ $\stackrel{1}{2}$




E－09







| Non |
| :--- |
| 00 |
| 1 |
| 1 |
| 1 |


 $0.89626199 E-06$
$-0.48963699 E-08$
$-0.35664400 E-07$
$0.69626199 E-06$
$0.13523550 E-07$ $-0.60620979 E-06$
$-0.67989500 E-07$
$0.37237049 E-07$

0．
$-0.21554600 \mathrm{~F}-06$
$0.51919840 \mathrm{E}-06$
$0.76336866 \mathrm{~F}-06$
$-0.28181099 \mathrm{E}-06$
$0.19024180 \mathrm{E}-06$
$0.28515480 E-06$
$0.16354869 E-07$
$-0.38722299 E-06$
$-0.3872229 \mathrm{E}-06$
0.162900 HOE－06
$0.7633 \mathrm{ARG6E-O6}$
○드응




.8 A344999E－07
$.8+141248 E-07$

 ii

|  |  |  |  | $\begin{aligned} & -10.439636991-08 \\ & -0.39664400 \mathrm{~F}-07 \\ & 0.69626231 \mathrm{E}-06 \end{aligned}$ | $\begin{array}{r} 0.69626731 E-06 \\ -0.45534493 E-07 \\ 0.24015090 E-10 \end{array}$ | $\begin{array}{r} -0.17 H 06399 E-10 \\ 0.6^{\circ} 7626231 E-06 \\ -0.45534699 E-07 \end{array}$ | $\begin{array}{r} -0.14545799 E-06 \\ 0.69626231 E-06 \end{array}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | 0．13523550E－07 | 0. | 0．79243379E－07 |  |
|  |  |  |  |  |  |  |  |
| －0．20004613－04 | O． $62+32050106$ | （）． 245154 H0E－06 | －0．212214ヶ＊－0\％ |  |  |  | $06$ |
|  | （O．H1）644891－0H |  | O． 2372 （）0゙刀－05 |  |  | $0.77400032 E-06$ | $28305930 E-06$ |
| U．6131301 $11-106$ | O． 1 361J2401－06 |  | $0.17400032 L-06$ $-0.201900905-06$ |  | $0.11170846 E-06$ | $-0.32879199 E-06$ | $0.77400032 E-06$ |
| － $1.343142014-116$ | $1.774000321-06$ $-0.34314299 t-06$ | $1.26140340 \mathrm{E}-06$ $0.17400032 \mathrm{t}-06$ | $-0.26136099-06$ $0.433335891-06$ | $-0.57414297 t-06$ | $0.77400032 \mathrm{E}-06$ | $\begin{array}{r} -0.32879199 \mathrm{E}-06 \\ 0.67673310 \mathrm{E}-06 \end{array}$ |  |
| $\begin{gathered} -0.231079 n 61-06 \\ 13 \end{gathered}$ | －0．34314299t－06 | U．17400032t－06 | 0．4333358\％－06 |  | 0．77400032E－06 |  |  |
| U．1 $10348191-1) 5$ | O．33H06ヶ「0t－0゙ | －0．220696\％9E－06 | 0．130215976－05 | 8 | 340E－06 | 6 | 8 |
| －0．57949500t－07 | O．HA131248t－07 | －0．48963693t－08 | －0．60411939\％－07 | 0．56612439E－05 | －0．63171299E－08 | 0．51495060E－09 |  |
| 0.1 H231269t－08 | O．21782110t－0＇） | U．13470日131－05 | 0．210374日01－07 | 0．13498429E－07 | $0.18870813 \mathrm{E}-05$ |  |  |
| U． 1 H8POHL3：－0\％ | －0．29084900t－07 | －0．207114908－06 | －0．HET7H333t－06 | 0.7823230 ¢E－08 | 0．21783289E－09 | $0.18870813 \mathrm{E}-05$ | $0.18705180 \mathrm{E}-07$ |
|  | O．1月H7OH13t－05 | 0．21092220E－07 | （1．13044800）－07 | 0．1．8870813E－05 | 0．57944289E－07 | 0．14743999E－06 |  |
| 14 |  |  |  |  |  |  |  |
|  | （1．120t953が－05 | －10．364474990－09 |  | 0．11069536E－05 | －0．48273599E－09 |  |  |
| －0．4 19219\％－06 | U．207514601－07 | 60t96262315－06 | 0．473H20791－06 | －0．63171799E－08 | 0．33208609E－05 | ．18784999E－09 | 8 |
|  |  |  | 0．111665361－05 | 0．85331499E－08 | －0．60194399E－09 | 0．11069536E－05 | 7 |
| U． $103473791-01$ | $0.1106 .35365-06$ | －0．3291790）－07 |  | 0．11069536E－05 | $0.64348698 \mathrm{E}-10$ |  | 0．11069536E－05 |
| 0．3．44日 2 24t－07 | －0．10205399t－09 | 0．11069536F－05 | （0． $854220 \% 95-07$ | 56391 AAE－ | 0.12069536 | －0．24170600E－07 |  |
| 15 |  |  |  |  |  |  |  |
| 3．68066H2 DE－08 | －0．14043400L－09 | 0．283154日OE－06 | －0．492063301－08 | －0．13497800E－09 | 6 |  |  |
| U．3051334 18－06 | $0.219807508-0 \mathrm{C}$ | －0．17306039F－10 | $0.305139431-06$ | 9 | 9 |  |  |
| U．13ヶ5577301－08 | U． 136752404 －08 | －0．152122901－08 | 0．293274691－07 | 0．32351274E－07 | 0．43611539E－08 | 7 | 0．30513943E－06 |
| J．34341149t－07 | － $0.22400000 \mathrm{t}-0 \mathrm{r}$ | $0.30515343 \mathrm{~L}-06$ | $0.150874501-08$ | 67899E－08 | 0．11170846E－06 | 8 | ． $29327480 \mathrm{E}-07$ |
| $\begin{gathered} -0.25707566 E-06 \\ 16 \end{gathered}$ | 0．43608119t－0． | －0．21036000E－07 | 0．305139431－06 | 71 | d |  |  |
| －0．4i222．665t－00 | －0．10122．29－（0） | 155154 $40 \mathrm{CO}-06$ | －0．412226654－06 | 9 |  |  |  |
|  | $0.2130121231-06$ | －0．463863リリヒ－09 | －0．26156199－06 | －0．41222665E－06 |  |  |  |
| －0．1く485200E－06 |  | －1）．41222665t－06 | 0．27573ヶ971－06 | －0．75605665E－07 | －0．41222665E－06 | －0．26433200E－06 | $0.69014749 \mathrm{E}-07$ |
| －U．41222665t－06 | －0．41222665t－06 | －0．41222665E－06 | －0．412226から！－06 | 0．34402629E－08 | $24139709 E-08$ | 6 | 0．76306669E－08 |
| O． 6 ¢75H3H9E－08 | －0．4122266：51－06 |  | OM1 | 222665E－06 | 700E－07 | 6 |  |
| 1 |  |  |  |  |  |  |  |
| 0．46．2437194－108 | 0．2242365st－05 | －（1）630）24つ91－09 | $4 \mathrm{BEO} \mathrm{H}-0 \mathrm{~A}$ |  |  |  |  |
|  |  | （3．6\％626231E－06 | （0．633130194－06 | 0.7423126 ＇e－08 | 064536E－05 | 8 | $16$ |
| （0．1）2．）1839F－04 | －0．40284009F－UT | （）． 1 गOB1H20E－06 | $0.440261331-05$ | 7 | 68YYE－06 | $4026 \pm 33 E-05$ | $.46966666 E-06$ |
| （）．1，134270L－06 | 0．441261331－05 | －0．11746600t－0） | 0．344029．94t－08 | 0．42858366E－05 | O． $9+302172$ TE－08 | $0.53092379 E-08$ | 0．28337319E－05 |
| －3．219075651－06 | －0．21336．3000－08 | 0．44028133t－05 | －0．75850．7332E－ 196 | －0．17144299E－07 | 44026133E－05 | 0．67106859E－06 |  |
| 14 |  |  |  |  |  |  |  |
| 0．2．1075301－0．0） | O． 176859.819 －09 | O．525かつO日SE－07 | O． 10851685104 | 9 |  |  |  |
| J． 1 is．75， 240 L －06 | （1．206034196－0日 | （）．24023610t－10 | 0．13675．404－08 | O． 217 H2110t－09 | －0．64338599E－10 | O．13675240E－06 | 0．86697899E－07 |
| －3，40284（0）91－197 | 0．41025720世－06 | －0．203277リ）f－07 | 0．13630400f－06 | 0．33512529E－07 | 8978E－07 | 7 | 6 |
| O， 2 －4ら57\％01－0\％ |  |  | O． 2413 3 819E－08 | 8 |  |  |  |
| U． $216.59 .405-06$ | 0．41505\％9\％t－08 | 0．17125669t－07 |  | 3657200 E | 426499E－08 | 40 |  |
| 14 |  |  |  |  |  |  |  |
| J． 1034 ALO |  |  | 0．1＊021以い71－05 | －0． |  |  |  |
|  | $0.280321238-06$ | －0． 4 1月70799\％－07 | －0． 2 23717901－00 |  |  |  |  |
| 0.10041 82） 006 | －0，20：32「矿！－0r |  | O． 311462104.06 | －0．7560566＇8－07 |  | $16$ | 0．20727500E－06 |
| J．くっ5306961－0． | 0． $25.30046 t-105$ |  | －0．88718331－06 | $0.53097809 \mathrm{E}-08$ | 9 | 0．25530696E－05 | 50902169E－08 |
|  |  | $0.500146409-08$ | －0．37734400 |  |  |  |  |
| $\therefore$ i |  |  |  |  |  |  |  |
| 1．1．6783108－07 | 0．22りく36ら61－05 | －0．534 5 （1）491－09 | －0． $26.70 \times 6 \mathrm{~mL}$ |  |  |  |  |
| $0.1 \times 335860$－ 00 | －（1．11：04 7－9－－ 11 | O．6．6．262111－06 | O． $7740003 \mathrm{C}^{\text {a }}$－06 |  |  | 0．76057666E－05 |  |
|  | （0．136，104001－06t | O． $311462101-06$ |  | －0．75605665F－07 | －0．26447699E－06 | $0.76057666 E-05$ | $0.46966666 E-06$ |
|  |  |  | $0.1630600101-08$ |  |  |  |  |
| －）．S Motsort－06 |  |  | －11．550173？${ }^{\text {a }}$ |  |  |  |  |
| $\therefore 1$ |  |  |  | 7 | 89E－07 |  |  |
| 9．10ヶ6けがリサー－17 |  |  |  | 0.134018424 －0．7 |  | 0．32351279E－07 | $5 E-0$ |
|  |  | W．3＂った64400F－07 |  | －0．226日1700E－06 | －0．65645594F－07 | －0．75605665E－07 | 0.75605 |
| 1.7 bistbor－01 |  |  | $\rightarrow B(140)-() \mathrm{H}$ | － 1.7 7605665E－07 | －0．75605665E－07 | －0．56893099E－09 | 0．75605665E－07 |
| ）． 1 －6icstost－01 |  |  | $-0.156(0,6,0: 5-0) 7$ | －0．29648OOOE－07 | －0．75605665E－07 | －0．75605665E－07 |  |
|  |  |  |  |  |  |  |  |

$-0.41222665 \mathrm{E}-06$
$0.22236349 \mathrm{E}-06$
$-0.33912300 \mathrm{E}-08$ 08
06
05
06 $0.43611539 E-$
$-0.27643400 E-$ 0.18870813E-05-0.80194399E-09 $\begin{array}{rrrr}0.24032123 E-06 & -0.45539399 F-07 & -0.37422099 F-06 \\ 0.43928979 F-07 & 0.25530696 E-05 & -0.26447699 E-06 \\ -0.25461199 E-05 & -0.93643799 E-06 & -0.84778333 E-06 \\ 0.25336959 E-05 & -0.34491999 E-08 & 0.55922649 E-08\end{array}$ $3.3332569 .9 E-05$
$3.1535894 E-05$
$0.545 J 7165 E-05$
$0.3019994 J E-08$
23



 $-24.59543997 \mathrm{E}-03$

 0
0
1
0
0
0
$n$
0
0
$n$
$n$
$n$
$i$
$i$ $0.13034819 E-05$
 $J .39210239 E-04$
$-J .21952400 E-06$ 26 $279100 \mathrm{E}-07$ 6
0
0
1
1
1
0
0
0
0
0
0 0 10
00
10
1
1
1
$m$
0
0
 $0.34283199 E-06$
$J .1 S 27008 J E-05$
$J .1: 746600 E-05$

 $-0.88778333 E-06$ $-J .25181697 \mathrm{E}-06$
$\mathrm{~J} .34402359 \mathrm{E}-08$ $0.64843799 E-06$
$0.12931140 E-05$ $0.4524519 J E-O B$

 30 $J .2: 7 J 783 J E-07$
$0.11170845 E-06$
$0.78021729 E-08$ $3,8659780 J E-07$
$J .12170845 E-06$ 31


[^10]| $907556 E-0$ | -0. | -0.14545799E-06 | -0.23907566E-06 | 0.13498519E-07 | 0.32448729E-07 | 6 | 8 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| -J.2390755SE-0s | 0.13675240E-06 | -0.76554999E-09 | -0.23907566E-06 | -0.75605665E-07 | 0.30199990E-08 | $-0.23907566 \mathrm{E}-06$ | .23907566E-06 |
| -0.2.952400E-05 | -0.23907566E-06 | -0.23907565E-06 | 0.12931140E-06 | -0.23907566E-06 | $0.11170846 \mathrm{E}-06$ | -0.65644900E-07 | -0.23907566E-06 |
| -0.71722699E-06 | 0.20936080E-06 | -0.23907566E-06 | -0.23907566E-06 | -0.29646600E-07 | 0.23907566E-06 | -0.23907566E-06 |  |
|  |  |  |  |  |  |  |  |
| 0,12334819E-05 | -0.11405999E-05 | 0.28515480E-06 | 0.13021597E-05 | 0.11303770E-05 |  | 06 | -07 |
| -0.33325699E-06 | 0.28032123E-06 | -0.54935300E-07 | -0.34314299E-06 | 0.18870813E-05 | -0.80205399E-09 | $0.43608119 \mathrm{E}-08$ | -0.41222665E-06 |
| -0.2133690JE-09 | 0.41506549E-08 | -0.64844099E-06 | -0.33920799E-08 | $0.30195549 \mathrm{E}-08$ | 0.25336959E-05 | -0.34497699E-08 | 0.55914239E-08 |
| 0.25336959E-05 | 0.25336959E-05 | -0.11746600E-05 | -0.88778333E-06 | $0.15934420 \mathrm{E}-06$ | $0.23455829 \mathrm{E}-07$ | 0.25336959E-05 |  |
| -220936080E-06 | 0.76010879E-05 | 410E-06 | 12613799E-06 | 0.25336959E-05 | 0.92379580E-06 | 0.64818719E-06 |  |
| 35 |  |  |  |  |  |  |  |
| 0.10275120E-0 | . 22823656 | 0.13790870E-0 | $0.37392440 \mathrm{E}-08$ | 0.2 |  | 0.468781905-08 | E-06 |
| $0.76336866 \mathrm{E}-05$ | -0.71399899E-08 | $0.69626231 E-06$ | $0.77400032 \mathrm{E}-06$ | $0.21092220 \mathrm{E}-07$ | $0.11069536 \mathrm{E}-05$ | -0.21036000E-07 | -08 |
| 0.44026.33E-05 | 0.17125669E-07 | $0.50014640 \mathrm{E}-08$ | $0.76057606 \mathrm{E}-05$ | -0.75605665E-07 | -0.34491999E-0才 | 0.76108632E-05 |  |
| -0.20432999E-06 | 0.56504766E-05 | $-0.11746600 E-05$ | $0.326347805-06$ | $0.42858366 E-05$ | $0.11170846 \mathrm{E}-06$ | -0.27643299E-06 | 0.76108632E-05 |
| ${ }_{-0.239075}$ | 0.32011410E-06 | 04 | -0.75507332E-06 | -0.46385600E-07 | 0.76108632E-05 | 5 |  |
|  |  |  |  |  |  |  |  |
| -3,59545499E-08 | -0.75507332E-06 | $0.28515480 \mathrm{E}-06$ | 0.50595900E-09 | -0.75507332 | 0.19555150E-06 | 8 |  |
| 0.19024180 EE 05 | 0.15595290E-07 | $0.35669090 \mathrm{E}-06$ | 0.43333589E-06 | -0.13044800E-07 | -0.85422099E-07 | 0.30513943E-06 | $-0.16081200 \mathrm{E}-08$ |
| -0,75507332E-06 | $0.13675240 \mathrm{E}-06$ | -0.37735400E-10 | -0.75507332E-76 | -0.75605665E-07 | 0.55912649 E -08 | -0.75507332E-06 | 0.46966666E-06 |
| 0.59014049E-07 | -0.75507332E-06 | -0.71849699E-06 | -0.20727500E-06 | -0.75507332E-06 | $0.11170846 \mathrm{E}-06$ | 0.22236320E-06 | 0.75507332E-06 |
| -0.23907566E-06 | 13797E-06 | 507332E-06 | .22652200E-05 | $0.27174439 \mathrm{E}-07$ | -0.75507332E-06 | 0.71044578E-07 |  |
|  |  |  |  |  |  |  |  |
| 13034817 E | 0.19101020E-06 | 0.28515480E-06 | 0.13021597E-05 | -0.22065400E-06 | 0.28568593E-06 | 6091599E-06 | 0.13523550E-07 |
| -0.58131397E-05 | 0.28032123E-06 | $0.13523550 \mathrm{E}-07$ | -0.57414299E-06 | 0.18870813E-05 | 0.85639188E-08 | -0.37138499E-09 | -0.41222665E-06 |
| -0,17144299E-07 | -0.23657200E-08 | $0.25530696 \mathrm{E}-05$ | -0.44646200E-07 | -0.29648000E-07 | 0.54507166E-05 | -0.46388400E-07 | 7 |
| 3.13070080E-04 | 0,22933510E-05 | -0.11746600E-05 | -0.88778333E-06 | -0.17142700E-07 | -0.23655299E-OB | 0.36206033E-05 |  |
| -3.29646600E-07 | 0.25336959E-05 | 5600E-07 | .27174439E-07 | 0.22044060E-03 | $0.51221232 \mathrm{E}-04$ | $0.24036849 \mathrm{E}-05$ |  |
| 38 |  |  |  |  |  |  |  |
| 0.38598579E-07 | 0.22923655E-05 | -0.92514499E-09 | 0.22249410E-08 | 0.22805686E-05 | -0.67133299E-08 |  |  |
| 0,75336865E-05 | 0,10563870E-07 | 0.69626231E-06 | $0.77400032 \mathrm{E}-06$ | $0.57944289 \mathrm{E}-07$ | 0.11069536E-05 | $0.69026009 \mathrm{E}-08$ | -0.31826700E-07 |
| 0.44026133E-05 | -0.34926499E-08 | -0.63437199E-06 | 0.76057566E-05 | -0.75605665E-07 | -0.92379399E-06 | $0.78671032 \mathrm{E}-05$ | -.31826700E-07 |
| -0.32823200E-07 | 0.56304766E-05 | -0.11746600E-05 | -0.83369499E-07 | 0.42858366E-05 | -0.34927499E-08 | $0.89848609 \mathrm{E}-06$ | $0.78603866 \mathrm{E}-05$ |
| -0.23907566E-05 | 6 | 8632 E | -0.75507332E-06 | 0.51221232E-04 | $0.15366369 \mathrm{E}-$ | 0.22490969E-05 |  |
| 39 |  |  |  |  |  |  |  |
| -3.24000299E-0S | 0.59792120E-06 | 0.28515480E-06 | 0.24084199E-06 | 871589E-06 | 588593 E |  |  |
| $0.50828000 \mathrm{E}-05$ | -0.68342800E-07 | 0.99243379E-07 | 0.67673310E-06 | -0.14743999E | -0.24170600E-07 | $0.30513943 \mathrm{E}-06$ | $1222665 E-06$ |
| 0.57106859E-05 | 0.13675240E-06 | -0.64824199E-06 | 0.14380430E-05 | -0.75605665E-07 | -0.64818699E-06 | 0.13671820E-05 | $1044520 \mathrm{E}-07$ |
| $0.24036849 \mathrm{E}-05$ | 0.15531819E-05 | -0.11746600E-05 | -0.88778333E-06 | $0.67106859 \mathrm{E}-06$ | 0.11170846E-06 | -0.64824199E-06 | .14380419E-05 |
| . $23907566 \mathrm{E}-0$ | 0.64818719E-06 | 20 | 0.71044578E-07 | 0.24036849E-05 | 0.22490969E-05 | 0.72110549E-05 |  |

Premultiplying Eq. (5) by $\left(W^{j}\right)^{T}$, the transponse of $W^{j}$,

$$
-\omega_{\mathbf{j}}^{2}\left(W^{\mathrm{j}}\right)^{\mathrm{T}} \mathrm{~A} W^{\mathrm{j}}+\left(W^{\mathrm{j}}\right)^{\mathrm{T}} \mathrm{M} W^{\mathrm{j}}=0
$$

or

$$
\begin{equation*}
\left(W^{\mathrm{j}}\right)^{\mathrm{T}} \mathbb{M W}^{\mathbf{j}}=\omega_{\mathrm{j}}^{2}\left(W^{\mathrm{j}}\right)^{\mathrm{T}} \mathbb{A} W^{\mathbf{j}} \tag{6}
\end{equation*}
$$

The orthogonality relationship of the modes may be expressed as

$$
\begin{equation*}
\left(W^{j}\right)^{T} M W^{k}=0 \quad j \neq k \tag{7}
\end{equation*}
$$

Eq. (6) may be put into a more general form by letting W represent the modal matrix (a matrix whose $j$-th column represents the $j$-th mode) and letting $\Omega$ represent a diagonal matrix containing the squares of the eigenvalues. Thus,

$$
\begin{equation*}
W^{\mathrm{T}} \mathrm{MW}=\Omega W^{\mathrm{T}} \mathrm{AW} \tag{8}
\end{equation*}
$$

## FORCED VIBRATIONS

The solution to Eq. (2) may be assumed to be of the form

$$
\begin{equation*}
X=\sum_{j} \phi_{\mathrm{j}} W^{j}=W \Phi \tag{9}
\end{equation*}
$$

in which $\phi_{\mathrm{j}}$ represents a scalar function of time associated with the j --th mode and $\Phi$. represents a column vector containing all the $\phi^{j}{ }^{\prime}$ S. Substitution of Eq. (9) into Eq. (2) yields the following equation

$$
\begin{equation*}
A W \ddot{\Phi}+M W \Phi=M D Q \tag{10}
\end{equation*}
$$

Premultiplying Eq. (10) by

$$
\begin{equation*}
W^{\mathrm{T}} A W \ddot{\oplus}+W^{\mathrm{T}} M W \Phi=W^{\mathrm{T}} M D Q \tag{11}
\end{equation*}
$$

Let,

$$
\mathrm{K}=\mathrm{W}^{\mathrm{T}} \mathrm{~A} \mathrm{~W}
$$

It can easily be shown that $K$ is a diagonal matrix. Premultiplying Eq. (11) by $K^{-1}$, i.e., the inverse of $K$, the result is

$$
\begin{equation*}
\ddot{\Phi}+K^{-1} W^{T} M W \Phi=K^{-1} W^{T} M D Q \tag{13}
\end{equation*}
$$

Recalling that $A=M D M$, noting that $K^{-1}=W^{-1} M^{-1} D^{-1} M^{-1}\left(W^{T}\right)^{-1}$, and using Eq. (8), the above equation takes the form of

$$
\begin{equation*}
\ddot{\Phi}+\Omega \Phi=K^{-1} W^{T} M D Q \tag{14}
\end{equation*}
$$

or

$$
\begin{equation*}
\ddot{\Phi}+\Omega \Phi=W^{-1} M^{-1} Q \tag{15}
\end{equation*}
$$

The solution to Eq. (15) depends on the nature of the force vector Q. In what follows the solution to this equation will be obtained using operational methods.

## TRANSFER FUNCTIONS

Taking the Laplace transform of both sides of Eq. (15)

$$
\begin{equation*}
\ddot{\bar{\Phi}}+\Omega \bar{\Phi}=W^{-1} M^{-1} \bar{Q} \tag{16}
\end{equation*}
$$

Similarly,

$$
\bar{X}=W \bar{\Phi}
$$

or

$$
\begin{equation*}
\bar{\Phi}=W^{-1} \bar{X} \tag{17}
\end{equation*}
$$

and

$$
\begin{equation*}
\ddot{\bar{\Phi}}=W^{-1} \ddot{\bar{X}} \tag{18}
\end{equation*}
$$

Substituting Eqs. (17) and (18) into Eq. (16)

$$
\begin{equation*}
W^{-1} \ddot{\bar{X}}+\Omega W^{-1} \overline{\mathrm{X}}=\mathrm{W}^{-1} \mathrm{M}^{-1} \overline{\mathbf{Q}} \tag{19}
\end{equation*}
$$

Assuming zero initial conditions and using the properties of the Laplace transforms concerning higher derivatives, Eq. (19) may be rewritten as

$$
\begin{equation*}
W^{-1}\left[s^{2} \bar{X}\right]+\Omega W^{-1} X=W^{-1} M^{-1} \bar{Q} \tag{20}
\end{equation*}
$$

where $s$ represents a scalar function of a complex variable. Let a diagonal matrix $Z$ be defined as

$$
\mathrm{Z}=\left[\begin{array}{ccc}
\left(\mathrm{s}^{2}+\omega_{1}^{2}\right) & 0 & 0 \cdots \cdots \cdot  \tag{21}\\
0 & \left(s^{2}+\omega_{2}^{2}\right) & 0 \cdots \cdots \cdot 0 \\
\cdot & & \\
\cdot & & \\
\cdot & & \\
\cdot & 0 & 0 \cdots\left(s^{2}+\omega_{n}^{2}\right)
\end{array}\right]
$$

in which, as before, $\omega_{j}$ represents the $j$-th natural angular frequency of vibration. With the use of matrix Z Eq. (20) becomes

$$
Z W^{-1} \overline{\mathrm{X}}=\mathrm{W}^{-1} \mathrm{M}^{-1} \overline{\mathrm{Q}}
$$

or

$$
\begin{equation*}
\bar{X}=W Z^{-1} W^{-1} M^{-1} \bar{Q} \tag{22}
\end{equation*}
$$

For computer programming purposes it is desirable to avoid finding the inverse of the modal matrix W. Noting that

$$
W^{-1}=K^{-1} W^{T} M D M
$$

Eq. (22) may be rewritten as

$$
\begin{equation*}
\overline{\mathrm{X}}=\mathrm{W} \mathrm{Z}^{-1} \mathrm{~K}^{-1} \mathrm{~W}^{\mathrm{T}} \mathrm{MD} \overline{\mathrm{Q}} \tag{23}
\end{equation*}
$$

Matrices $Z$ and $K$ being diagonal, their inverses can be found easily.

The transfer functions for the system under study may be found from Eq. (22) or Eq. (23) by applying a series of impulses one at a time at each of the degrees of freedom. Since the Laplace transform of a unit impulse function is unity, for each application of a unit impulse the matrix $\bar{Q}$ represents a column matrix containing all zero elements except a single non-zero element equal to one, at the appropriate position. To obtain the entire set of transfer functions it can be shown that $\bar{Q}$ has to simply be replaced by the identity matrix, I. Therefore, denoting by $T$ the matrix containing the transfer functions

$$
\begin{equation*}
T=W Z^{-1} W^{-1} M^{-1} Q=W Z^{-1} K^{-1} W^{T} M D \tag{24}
\end{equation*}
$$

Substituting Eq. (24) into Eq. (22)

$$
\begin{equation*}
\overline{\mathrm{X}}=\mathrm{T} \overline{\mathrm{Q}} \tag{25}
\end{equation*}
$$

Eq. (25) represents the solution to the forced vibration problem (see Eqs. (2) and (16)). If the Laplace transforms of the forcing functions have been found, the operational displacements can be computed using the above equation. The physical displacements can then be determined by finding the inverse transform of $\overline{\mathrm{X}}$.

## NUMERICAL RESULTS

Based on the above analysis a computer program has been developed to calculate the natural frequencies and modes of vibration, and the transfer matrix of the system. The natural frequencies and modes of vibration obtained in this manner are listed on following pages.

The above eigenvalues represent the $\omega_{j}^{2}$, or, the squares of the angular frequencies. The corresponding linear frequencies, the $f_{j}{ }^{\prime} s$, in cycles per second may be found from

$$
\begin{equation*}
f_{j}=\frac{\omega_{j}}{2 \pi} \tag{26}
\end{equation*}
$$

A simple computation shows that the above eigenvalues correspond to a frequency range of about 1.03 cps to 43.5 cps . This frequency range in general, shows some agreement with the experimentally measured frequency range.

The transfer coefficients obtained by using the computer program are also shown on following pages.
EIGENVALUES AND EIGFNVFGTORS
$0.16255552 \mathrm{E}-02$
$-0.14555386 \mathrm{E}-02$
$0.1032442 \mathrm{E}-02$
$0.15702818 \mathrm{E}-01$
$0.15614480 \mathrm{E}-01$
$0.16239659 \mathrm{E}-02$
$0.88277671 \mathrm{E}-02$
$0.12902088 \mathrm{E}-00$

 $-0.99149724 E-02$
$0.17288411 \mathrm{E}-01$
$0.11539984 \mathrm{E}-00$
$0.71043544 \mathrm{E}-02$
$-0.14865999 \mathrm{E}-01$
0.75857990 e 02 EIGENVALUE NO.
E[GEVVECTUR EIGE J3071045E-02
$0.74522286 E-01$
$-3.57571523 E-01$
$-0.44525 H 83 E-02$

No, $1=0.41606855 \mathrm{E} \mathrm{O2}$

$\begin{array}{rrr}0.47156435 \mathrm{E}-02 & 0.51422128 \mathrm{E}-02 & 0.92711081 \mathrm{E}-03 \\ 0.1403816 \mathrm{E}-03 & 0.76204967 \mathrm{E}-02 & 0.29392918 \mathrm{E}-02 \\ 0.14703825 \mathrm{E}-01 & -0.23547269 \mathrm{E}-03 & 0.14301590 \mathrm{E}-01 \\ -0.33005985 \mathrm{E}-02 & 0.95339906 \mathrm{E}-02 & 0.22226173 \mathrm{E}-04 \\ -0.15176263 \mathrm{E}-02 & 0.57311071 \mathrm{E} 00 & 0.31528994 \mathrm{E}-00\end{array}$ -015176263

$0.18282567 \mathrm{E}-02$
$0.32052314 \mathrm{E}-04$
$0.14919756 \mathrm{E}-01$
$0.12445194 \mathrm{E}-01$
$0.97208808 \mathrm{E}-02$ $-J .50350978 \mathrm{E}-03$

$-0.39374733 \mathrm{E}-00 \quad 0.66847448 \mathrm{E}$ $0.27626527 E-01$ $\begin{array}{lll}0.16160738 \mathrm{E}-00 & 0.71429315 \mathrm{E}-03 & 0.27626527 \mathrm{E}-01 \\ 0.36 .658435 \mathrm{E}-02 & 0.38457571 \mathrm{E}-01 & 0.29661585 \mathrm{E}-01\end{array}$ EIGEVVALUE NO. 3
EIGENVECTUR
$3=0.11089526 \mathrm{E} \mathrm{O}$
$\begin{array}{rr}0.16165175 \mathrm{E}-00 & 0.43311813 \mathrm{E}-02 \\ 0.5837605 \mathrm{E}-01 & 0.83646980 \mathrm{E}-01 \\ -0.30675647 \mathrm{E}-02 & 0.41137783 \mathrm{E}-01 \\ 0.19609089 \mathrm{E}-00 & 0.1420540 \mathrm{E}-02 \\ 0.14950383 \mathrm{E}-00 & -0.44824563 \mathrm{E}-00\end{array}$
 $\begin{array}{ll}16160738 E-00 & 0.71429315 E-03 \\ 36558435 E-02 & 0.38457571 E-01 \\ 33272903 E-03 & 0.40915098 E-01 \\ 27864814 E-00 & 0.84193633 E-02 \\ 37711270 E-01 & 0.22273489 E-00\end{array}$ $4=0.22252107 E 03$ $0.27054084 E-01$ $0.27054045 \mathrm{E}-01$
$-0.2<040454 \mathrm{E}-01$ $-0.24040464 \mathrm{E}-01$
$0.14815842 \mathrm{E}-00$
$0.89353450 \mathrm{E}-01$ $0.89353450 \mathrm{E}-01$
$-J .99850146 \mathrm{E}-02$ EIGEVVALUE NO.
EIGENVECTOR J. $11298352 \mathrm{E}-00$ $-0.25451530 \mathrm{E}-02$
-0.2 -0 $-3,28079295 \mathrm{E}-01$
 -0.32705064E-01
 $0.16166644 \mathrm{E}-01$
$-0.19884668 \mathrm{E}-01$
$0.14683228 \mathrm{E}-00$
$-0.2479602 \mathrm{E}-04$
$0.16733088 \mathrm{E}-00$

 $-0.30619451 E-02-0.28682537 E-01$
$0.13262576 E-00-0.40913482 E-01$
 $0.22765797 E \quad 03$ $\begin{array}{rrrrr}\text { EIGETVECTUR } & & & & \\ -0.1 ? 155650 E-03 & -0.23036480 E-00 & 0.65801434 E-02 & -0.74368916 E-04 \\ 0.10749691 E-00 & 0.31909300 E-03 & -0.39800844 E-01 & -0.12757464 E-01 \\ 0.45215984 E-01 & 0.20758864 E-02 & 0.18118485 E-01 & 0.34554205 E-00 \\ 0.14790940 E-00 & 0.18950063 E-00 & -0.13885865 E-00 & 0.25651259 E-01 \\ -0.77132415 E-03 & 0.32598915 E-01 & 0.52069471 E 00 & 0.12938587 E-01\end{array}$ $6=0.27718224 \mathrm{E} \mathrm{O}$ EIGEIVALUE NO.
EIGE JVECTUR $\begin{array}{llll}\text { EICEJVECTUR } & & \\ 0.10722940 E-00 & 0.55209408 E-01 & 0.46 \\ -0.27720074 E-01 & 0.23093747 E-01 & 0.7 \\ J .16953521 E-01 & 0.91052740 E-03 & 0.1\end{array}$ -25才5月288E EIGENVALUE NO, 0.94846264 L -
$7=0.29869724 E 03$
 ELTEVGLUE NO $8=0.30124167 E 03$ LITEVALIJE NO.
EIGENVECTUR $\begin{array}{llll}\text { EIGENVECTUR } & & \\ -0.4 S J 4 S 6 S 8 E-01 & -0.69345076 E-01 & -0.1 \\ -J .19676154 E-02 & 0.25999267 E-02 & 0.6\end{array}$

EIGENVALUENO,
EIGENVECTUR
N

 -0.50141016E-01


 $-0.74529521 \mathrm{E}-01$
$0.48877160 \mathrm{E}-01$ $-0.12440248 \mathrm{E}=01$
$0.89997501 \mathrm{E}-03$
$0.4456058 \mathrm{E}-02$
$0.24434158 \mathrm{E}-01$ $0.4887160 \mathrm{E}=01$
$-0.10514450 \mathrm{E}-02$
0.89153658 O
$-0.13359457 \mathrm{E}-01$ $10.40941392 E \quad 03$


[^11]$$
13=0.57339258 \mathrm{E} 03
$$


$\begin{array}{lllllllll} \\ \text { EIGENVECTOR } & 0.10428378 E-00 & -0.18323995 E-01 & -0.12559033 E-00 & -0.94260582 E-01 & -0.29450447 E-01 & 0.18794305 E-02 & -0.10151454 E-01\end{array}$

 EISEIVALUE ND. $15=0.59831614 E 03$

[^12]





\[

$$
\begin{aligned}
& \text { EIUEVVALIK NO, } 18=0.75459960 \mathrm{E} 03 \\
& \text { EIGENVEGTOR }
\end{aligned}
$$
\]



 $0.12244460 \mathrm{E}-00$





$$
\begin{aligned}
& \text { ELGENVALUE NU. } 19=0.10362227 E 04 \\
& \text { ELGEGVECTOR }
\end{aligned}
$$

$0.40185889 E-01$
$-0.53948984 E-02$
$-0.15389075 E-01$
$0.28102938 E-01$




 $0.17599892 E-01-0.72319412 E$

$0.12350552 E-00$
$0.36327866 \in-01$
$0.360791206-01$
 $0.69656391 E-02-0.39520121 E-0$

 $\begin{array}{ll}-0.46229744 \mathrm{E}-02 & 0.95440241 \mathrm{E}-01 \\ -0.30493324 \mathrm{E}-02 & 0.40245710 \mathrm{E}-00\end{array}$
 EIGEIVALUE NU, $2 E=0.22694551 E 04$ $0.2317142 \mathrm{E}-01$
$-0.134433 \mathrm{E}=01$
$-0.26441645 \mathrm{E}=02$ $-0.26941645 \mathrm{E}-02$

 $-0.10821098 \mathrm{E}=1$
$0.42909829 \mathrm{E}-00$ 0.15811483E-00

 $\begin{array}{lll}-0.3004600 \mathrm{E}-01 & 0.42909829 \mathrm{E}-00 \\ 0.102467297 \mathrm{E}-00\end{array}$ \begin{tabular}{l}
0 <br>
0 <br>
1 <br>
1 <br>
$t$ <br>
0 <br>
4 <br>

+ <br>
0 <br>
\hline- <br>
0 <br>
-1 <br>
0
\end{tabular}

[^13]$$
0.14610052 E 04
$$

## 22


IGE'IVALUE NO. $23=0.15130768 \mathrm{E} 04$

[^14]Noog:
-0.50201227 00 -0.2

 $\begin{array}{r}0.49810310 \mathrm{E}-00 \\ 0.0792294 \mathrm{E}-00 \\ -0.49632599 \mathrm{E}-00 \\ \hline .\end{array}$ 0.10326908 E OL


 $\begin{array}{lllll}0.41903601 E-01 & -0.32125656 E-00 & -0.13945498 E & 01 & 0.40676901 E-00 \\ 0.67602818 E-01 & -0.34590391 E-02 & -0.17856361 E-01 & 0.78433491 E\end{array}$ $\begin{array}{llll}0.0 .11985794 \mathrm{E}-01 & -0.14553655 \mathrm{E}-00 & -0.27679902 \mathrm{E}-01 & -0.78200810 \mathrm{E}-01 \\ -0.110201\end{array}$ $\begin{array}{llll}0.25899158 \mathrm{E}-01 & -0.14530648 \mathrm{E}-01 & -0.49402028 \mathrm{E}-01 \\ 0.11649328 \mathrm{E}-01 & 0.35357731 \mathrm{E}-02 & 0.68875507 \mathrm{E} & 00\end{array}$ $0.29625203 \mathrm{E}-00$
 $-0.2230459 E-00$
$0.27425583 E-01$

 $0.26127951 E-02-0.44534670 E-02-0.79702678 E-01$
 $0.74958660 \mathrm{E}-01$ $0.74958660 \mathrm{E}-01$
$-0.16501444 \mathrm{E}=01$
$-0.18048210 \mathrm{E}-01$
$0.1163800 \mathrm{E}-01$
$-0.12243196 \mathrm{E}-01$
 8\% 4:8 $68 \mathrm{E} \quad 00$
$35 \mathrm{E}-02$
 Elgevalue no. $28=0.28413404 \mathrm{E} 04$

| Eigevectior |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.30127999E-02 | 0.17532927E-01 | -0.76802229E-01 | 31835306 |  | .72787738E-02 | $0.24652439 \mathrm{E}-01$ | $0.35797635 \mathrm{E}-00$ |
| .19672693E-00 | $0.286 i 4633 E-00$ | -0.22749359E-1 | -0.39229204E-01 | $0.91764983 \mathrm{E}-02$ | O. $0.13267909 E-00$ | $0.59636144 \mathrm{E}-02$ | 0.87782980E-01 |
| $-0.49615698 \mathrm{E}-01$ $-0.33752162 \mathrm{E}-02$ | $-0.11849906 E-01$ $0.17003412 \mathrm{E}-00$ | -0.44264691E-00 | -0.19696567E-01 | $0.17149278 \mathrm{E}-01$ | -0.19317350E-01 | $0.80514076 \mathrm{E}-02$ | 0.28962847E-02 |
|  |  | -0.26765389E-01 | 0.31947003E-01 | 0.26127951E-02 | -0.44534670E-02 | -0.79702678E-01 |  |

0.10424555 E 01
 $0.38888605 \mathrm{E}=00$ -0.70535390E-0: -0.32668936E-00
$0.56377241 E-01$
$-0.15774137 E$ O1
$-0.56138125 E 00$
$0.14189667 E-00$
 $0.21178371 \mathrm{E}-01$
$0.23066368 \mathrm{E}-01$ - $14727010 \mathrm{E}-00-0.70535390 \mathrm{E}-02$ EIGEVVALUE NO. $29=0.29499044 E 04$
EIGENVECTOR 0. $0.31578216 \mathrm{E} \quad 04$
eigenvalue no, $31=0.34575645 \mathrm{E} 04$



 $\begin{aligned} & .13406059 \mathrm{E}-00 \\ & 0.0 .26585685 \mathrm{E}-00 \\ & 0.47374447 \mathrm{E}-00 \\ & 0.0 .12332663 \mathrm{E}-01\end{aligned}$
 $\begin{array}{llll}-3.3=157861 E-0 J & 0.14947321 \mathrm{E} & 01 & 0.29636370 \mathrm{E} \\ -0.3833246 \mathrm{E} & 00 & -0.42301204 \mathrm{E}-00 & 0.31881044 \mathrm{E}-01\end{array}$

[^15]
RRANS.FER COEFFICIENTS
$0.18540 E-06$
$-0.16312 E-06$
$0.11775 E-06$
$0.17910 \mathrm{E}-05$ 응nnn


$\begin{array}{rr}0.26971 E-06 & 0.23980 E-06 \\ 0.47267 E-08 & -0.21105 E-06\end{array}$

$\sim$
0
0
1
$\sim$
$m$
0
0
0
$n$
0
0
0
1
$\sim$
0
1
1
4
$m$
$m$
$m$
$m$
$m$
0
 $0.22690 \mathrm{E}=06$
$0.17723 \mathrm{E}-06$



UEGR: OF FREEDOM VO, 2
$0,70000 E-0 S ~ 0,90547 E-05$




## 

- 



$\circ$
0
$i$
$\vdots$
0
0
0
0
0



## -






No
$i$
$i$
$i$
0
0
0
0
0
0
0

## 






$\begin{array}{rr}0,44550 E-05 & 0.20254 E-05 \\ -0,57438 E-07 & 0.10755=-05\end{array}$
DEGREE OF FREADOM VO, 5
$0.53550 E-06 \quad 0.75350 E-06$

0.46519E-05 J.22i19E-05 $-0.75533 \mathrm{E}-07$ 0.2129こー0
$0.80338 E-07 \quad 0.71430 E-07$



 on
0
0
0
0
0
0
0
0
 0
0
1
1
0
0
1
1
$n$
$n$
$n$
0
0


 N
0
0
1
0
0
0
0
0
0
0
0.
0.

 ro
00
00
0
0
0
0
0
0
0
0


-J.20:45E-J7 O.41700E-06







 degree of freedóm no. il DEGREE OF FREEDOM VO. IL

 OEGREE OF FREEDM NOO
 J.16011E-07 0.207, 0 EE-07 $\begin{array}{rl}-0.20022 E-08 & 0.29633 E-08 \\ 0,30093 E-07 & 0.2329 E-10\end{array}$ -0,20075E-08 $0.32029 \mathrm{E}-07$













 $0.193 う 7 E-05$
$0.29773 E-06$
$-0.26204 E-06$
$0.1890 E-06$
$0.28760 E-05$







| 1 |
| :--- |
| 4 |
| 0 |
| 0 |
| -7 |
| -1 |
| 0 |
| 0 |
| 0 |
| 0 |
| 1 |
| 4 |




$0.86337 \mathrm{E}-06$


$0.71421 \mathrm{E}-09$





## 





0
0
1
1
+1
-1
$m$
$m$
0
0
0


$\therefore$


$1 n$
0
1
4
0
0
0
0
0
0
0
0
0
0
0
0
4
0
0
0
0
0
0

 $0.26078 \mathrm{E}-05$
0.20838






$\begin{array}{cc}\text { DEGREE OF FREEDOM NOO } 18 \\ 3,79233 E-09 & 30.0235 E-08 \\ -3,10271 E-99 & 301465 E-09 \\ 3.14873 E-08 & 0.1154 E-11 \\ 0.50765 E-08 & 0.29843 E-03 \\ -0.99221 E-10 & 0.15825 E-03\end{array}$
$0.2836 I E-06$
$-0.24961 E-06$
$0.18013 E-06$
$0.27396 E-05$








## © 0 in 0 m <br> $n$ 0 0 1 0 <br>  <br> 

in owin in
0000
10

 DEUREE OF FREEDOM NO， 20
$0,15771 E-05 \quad 0021692 ミ-05$


 $\begin{array}{rr}0.133 J 2 E-04 & 0,63246 E-0 \\ -0,2328 E-06 & 0,33538 E-05\end{array}$
$\begin{array}{llr} \\ -0.52469 E-03 & -0.10347 E-07 & -0.91999 E-08 \\ -0.2069 E-07 & -0.18133 E-09 & 0.80970 E-08 \\ --0.80940 E-07 & -0.8443 \mathrm{E}-07 & -0.58432 \mathrm{E}-08 \\ -0.12575 E-09 & -0.70434 \mathrm{E}-07 & -0.88370 \mathrm{E}-07 \\ -0.0 .17844 \mathrm{E}-05 & -0.55016 \mathrm{E}-07 & \end{array}$





000
0000
10
1
1
0
1
$0.62845 E-06$
$0.21013 E-07$
$0.51284 E-05$
$0.42779 E-05$
$0.33414 E-05$


 $0.44628 E-05$
$0.34858 E-05$

 DEGREE OF FREEDOM NU． 22
$0.15312 E-05$ J． $21093=-05$

 $\begin{array}{rr}0.129385-04 & 0.61516 E-05 \\ -0.20453 E-06 & 0.32020 E-05\end{array}$


DEGREE OF FREEDOM NO， 23 3，17017E－05 0．22010E－0 －J，2：2：29E－06 0．31494E－0
 -522 ij37E－06 Jo34030こ－05
 $0.30881 E-06$
$0.24121 E-06$



DEGREE OF FREEDON VO， 24
$0,1.775 E-O 6$ O． $25230 E-06$
$-0.2313 E-07 \quad 0.27793 E-07$

## No 0 1 1 0 0 m m 4 t 0 <br> 0.45366 E－07 $0.40336 \mathrm{E}-07$

 -11
1
0
0
0
10
10
n
0
0
0
$0.38964 \mathrm{E}-06$






0．4 $4519 E-05$ $n$
0

100
0
1
4
0
0
0
0
1
4
0
0
0
0
0
1
4
4
4
0
0
0
0
0
0


0．5i846E－03 DEGREE DF FREEDON NO， 25
$0,4-9305-05 \quad 0.55526 E-05$ $\begin{array}{rr}0,95396 \mathrm{E}-06 & 0.44407 \mathrm{E}-05 \\ -0.14764 \mathrm{E}-07 & 0.23548 \mathrm{E}-06\end{array}$






 $\begin{array}{ll}0 & 0 \\ 0 & 0 \\ 1 & 1 \\ \text { t } \\ 0 & t \\ N & 0 \\ 0 & 0 \\ \sim & 0 \\ m & 0 \\ 0 & 0\end{array}$ 0
0
1
1
1
1
1
0
0
$\sim$
$M$
$m$
0
0
0


DESREE OF FREEJOM NO 26
$0.204125-05$ J． $2640 \pm E-05$
$\begin{array}{rr}0.20412 E-05 & 0.26401 E-0 \\ -0.26544 E-05 & 0.37777 E-06\end{array}$

$\begin{array}{rr}0,3 E 363 E-05 & 0.29700 E-08 \\ 0,15170 E-04 & 0.76978 E-05 \\ -0,25594 E-06 & 0.40819 E-05\end{array}$

DEGREE OF FREEDOM NO． 27
$-0.56274^{\circ}-06-0.72786 E-06$

 －OON NO 28
$\begin{array}{ll}-0.26349 E-08 & -0.14758 E-05 \\ -0.37388 E-04 & -0.11527 E-05\end{array}$

$\begin{array}{rlr}-0.73545 \mathrm{E}-07 & -0.14504 \mathrm{E}-06 & -0.12895 \mathrm{E}-06 \\ -0.22524 \mathrm{E}-06 & -0.25418 \mathrm{E}-08 & 0.11349 \mathrm{E}-06\end{array}$
 $90-\exists \varepsilon 6 \angle 0+^{\circ} 0-90-360+L \varepsilon^{\circ} 0-\angle 0-\exists L L I 0^{\circ} 0-$ JEGREE DF FREEDOM ND． 28

$-0.12457 E-05$
$-0.98727 E-06$
$-0.77115 E-06$
$\begin{array}{llll}0.22577 E-06 & -0.75632 E-06 & -0.4526 E-08 \\ 0.12436 E-06 & -0.45464 E-04 & -0.25012 E-04\end{array}$

## 

un $n$
0
1
1
1
1
$n$
$\infty$

$m$
$m$
-1
0
0
 OESV E JF FREEUUN NAS 29

00
00
0
1
1 0

 $0.22275 \mathrm{E}-05$


## 

$0 \infty$
000
0
1 $0_{1} 0$


 $900 \infty$

 $D E O R=E$ OF FREEJUM NO， $3!$
$0.48623 \mathrm{E}-06$
$-0.4279 \mathrm{E}-06$
$0.30882 \mathrm{E}-06$



## 

12
0
1
1
0
0
0
0
0
0
0
0 77E－0
 －



## DESR EE OF FREEDOM NOS 32 $J .27910 E-05 \quad 0.23255 E-0 J$




 00
0
0
0
1
1
1
1





DEGR：E OF FREEOU NJ， 34
$D E S Z_{L E} J F F K E E O O M$ NO， 33
$-0.53834 E-07-0.89032 E-07$
$0.89524 E-J 8-0.22739 E-07$
$-0.1 \angle 937 E-05-0.20025 E-03$
$-5.54575 E-05-0025459 E-05$
$0.36307 E-08-0.23755 E-05$













うミうて＿ニ JF FREEJU＊dO， 35


Mơ険令

（1）
コン2



For the above transfer coefficients, the point of application of the unit impulse is indicated by referring to the "degree of freedom number." Following this designation, the coefficients associated with each degree of freedom are given sequentially (by row). Thus, for example, the 39 numerical values listed after the "Degree of Freedom No. 1" are the transfer coefficients for each degree of freedom due to a unit impulse applied at degree of freedom No. 1. If the transfer coefficients are divided by $s^{2}+\omega_{1}^{2}$ they will then yield the desired transfer functions. However, the transfer functions obtained in this manner represent only approximations to the actual values, since they are based only on the contribution due to the fundamental mode. However, the contributions of all 39 modes have been obtained by means of the computer program. For the purpose of brevity, they are not listed in this report, but may be obtained through the Antenna Systems branch, at Goddard Space Flight Center.

It may be observed that, as expected, the transfer coefficients possess symmetrical properties since the transfer matrix of a linear system is symmetrical.

## CONCLUSIONS

A simplified model has been used for determining the dynamical characteristics of the Rosman I Antenna. The model consisting of 39 degrees of freedom yields natural frequencies encompassing the experimentally observed frequency range. The transfer functions obtained by means of the model may be used in a control system simulation of the antenna. More details concerning this study and the computer results are furnished in a paper by the same authors, reported in the 1967 Goddard Summer Workshop Report.

# SPATIAL LIGHT MODULATION EMPLOYING A FERRO-MAGNETIC FLUID 

by

J. P. Uldrick* and P. A. Lux**


#### Abstract

This report describes some work done during the summer of 1967 concerning the feasibility of synthesizing a fluid light modulator by using a ferro-magnetic fluid for transducing an electrical signal into a corresponding spatial transparency distribution. Several cells using $3^{\prime \prime} \times 5^{\prime \prime}$ glass plates with $0.01^{\prime \prime}$ to $0.10^{\prime \prime}$ spacings were constructed. A clear fluid was passed through the region between the plates at about 1 foot per second. In the middle of the stream a small stream of opaque fluid was inserted in the clear carrier fluid to form a pattern between the plates. Since the flow is laminar the pattern will hold its shape as it traverses the cell. The storage time is adjusted by regulating the flow rate of the clear carrier fluid.


A ferro-magnetic fluid was manufactured for use as the opaque fluid. Since this fluid is highly responsive to magnetic fields, then it is possible to deflect it with an electromagnet. This gives a method for transducing an electric signal into an optical pattern. By properly shaping the poles of the electromagnet it is possible to exert a constant body force on the ferrofluid in some region between the poles.

It was necessary to design and construct an attritor for manufacturing the ferrofluid. The ferrofluid consists of $\mathrm{Fe}_{3} \mathrm{O}_{4}$ particles of about 100 antroms in diameter suspended in JP-4 jet fuel. A normal ball mill is not capable of grinding the iron powder small enough.

## INTRODUCTION

The object of the work reported herein was to investigate the feasibility of storing information in a moving fluid. Ultimately the information was to be read out using a laser.

[^16]A cell was constructed which consisted of two $3^{\prime \prime} \times 5^{\prime \prime}$ parallel glass plates spaced close together and a fluid was passed between the plates at relatively low velocity. Because of the close spacing and the low velocity the fluid flow is laminar and has uniform velocities in planes parallel to and between the glass plates. One such cell is shown in Figure 1.1.

If a clear carrier fluid is used between the plates and a pattern of dark fluid is immersed in the carrier fluid between the glass plates near the upstream end of the cell its pattern is carried downstream through the cell with very little distortion. The velocity with which the pattern traverses the cell can be adjusted by regulating the flow rate of the clear carrier fluid. In this way the storage time can be controlled.

Although this investigation did not progress that far, the ultimate use for this cell would be in an optical data processing system. In this application a laser beam would pass through the cell normal to the glass plates and the pattern of dark and clear fluid in the cell would be detected with an optical filter.

The fluids used in the study were JP-4 jet fuel for the clear fluid and a ferrofluid for the dark fluid. The manufacture of the ferrofluid will be discussed below.

## THE FLUID FLOW PROBLEM

The principle that will be discussed to transport the desired transparency across the aperture of the optical system will be the steady laminar flow between parallel boundaries. In order to use such a system one must be able to control precisely the trajectories of any given fluid particle. This requirement precludes any possibility of local fluctuation in the velocity or mixing of the fluid as it passes between the boundaries. Such will be the case if the characteristic flow parameter composed of the product of the mean velocity V , density $\rho$, and the spacing between the boundaries L, divided by the dynamic viscosity $\mu$ of the fluid (i.e., a Reynolds' number), has a magnitude not greater than a certain limiting value which is roughly of the order of $10^{3}$. For a fluid such as kerosene with a kinematic viscosity $\mu / \rho=2 \times 10^{-4} \mathrm{ft}^{2} / \mathrm{sec}$ requires that

$$
\begin{equation*}
\mathrm{V} \mathrm{~L} \leq 0.2 \tag{2.1}
\end{equation*}
$$

where V and L are defined above.
Let us now consider the laminar flow of an incompressible fluid between infinite parallel plates as shown in Figure 2.1.


Figure 1.1(a)


Figure 1.1(b)


Figure 2.1

The equation of motion for an incompressible viscous fluid is

$$
\begin{align*}
& \rho \frac{\partial \mathrm{V}_{\mathrm{x}}}{\partial \mathrm{t}}+\mathrm{V}_{\mathrm{x}} \frac{\partial \mathrm{~V}_{\mathrm{x}}}{\partial \mathrm{x}}+\mathrm{V}_{\mathrm{y}} \frac{\partial \mathrm{~V}_{\mathrm{x}}}{\partial \mathrm{y}}+\mathrm{V}_{\mathrm{z}} \frac{\partial \mathrm{~V}_{\mathrm{x}}}{\partial \mathrm{z}}=\rho \mathrm{B}_{\mathrm{x}}+\frac{\partial \mathrm{p}}{\partial \mathrm{x}}+\mu \nabla^{2} \mathrm{~V}_{\mathrm{x}} \\
& \rho \frac{\partial \mathrm{~V}_{\mathrm{y}}}{\partial \mathrm{t}}+\mathrm{V}_{\mathrm{x}} \frac{\partial \mathrm{~V}_{\mathrm{y}}}{\partial \mathrm{x}}+\mathrm{V}_{\mathrm{y}} \frac{\partial \mathrm{~V}_{\mathrm{y}}}{\partial \mathrm{y}}+\mathrm{V}_{\mathrm{z}} \frac{\partial \mathrm{~V}_{\mathrm{y}}}{\partial \mathrm{z}}=\rho \mathrm{B}_{\mathrm{y}}+\frac{\partial \mathrm{p}}{\partial \mathrm{y}}+\mu \nabla^{2} \mathrm{~V}_{\mathrm{y}} \\
& \rho \frac{\partial \mathrm{~V}_{\mathrm{z}}}{\partial \mathrm{t}}+\mathrm{V}_{\mathrm{x}} \frac{\partial \mathrm{~V}_{\mathrm{z}}}{\partial \mathrm{x}}+\mathrm{V}_{\mathrm{y}} \frac{\partial \mathrm{~V}_{\mathrm{z}}}{\partial \mathrm{y}}+\mathrm{V}_{\mathrm{z}} \frac{\partial \mathrm{~V}_{\mathrm{z}}}{\partial \mathrm{z}}=\rho \mathrm{B}_{\mathrm{z}}+\frac{\partial \mathrm{p}}{\partial \mathrm{z}}+\mu \nabla^{2} \mathrm{~V}_{\mathrm{z}} \tag{2.2}
\end{align*}
$$

where

$$
\begin{aligned}
\rho & - \text { fluid density } \\
\left\{\mathrm{V}_{\mathrm{x}}, \mathrm{~V}_{\mathrm{y}}, \mathrm{~V}_{\mathrm{z}}\right\} & - \text { velocity components } \\
\mathrm{p} & \text { - pressure } \\
\left\{\mathrm{B}_{\mathrm{x}}, \mathrm{~B}_{\mathrm{y}}, \mathrm{~B}_{\mathrm{z}}\right\} & - \text { body force components } \\
\mu & \text { - dynamic viscosity }
\end{aligned}
$$

For the case under study

$$
V_{x}=V_{y}=\frac{\partial V_{x}}{\partial t}=\frac{\partial V_{y}}{\partial t}=\frac{\partial V_{z}}{\partial t}=0
$$

If the body force is conservative, i.e., $\vec{B}=\nabla B$ where $B$ is the body force potential, then the equations given in (2.2) simplify to

$$
\begin{gather*}
\frac{\partial}{\partial \mathbf{x}}(\rho \mathbf{B}+\mathbf{p})=0  \tag{2.3a}\\
\frac{\partial}{\partial \mathbf{y}}(\rho \mathbf{B}+\mathbf{p})=0  \tag{2.3b}\\
-\frac{\partial}{\partial \mathbf{z}}(\rho \mathbf{B}+\mathbf{p})=\mu \frac{\partial^{2} \mathbf{V}_{\mathbf{z}}}{\partial \mathbf{y}^{2}}=\beta \tag{2.3c}
\end{gather*}
$$

Equation (2.3c) can be integrated to give

$$
\begin{equation*}
\mathbf{y}_{z}=\frac{\beta}{\mu}\left(\frac{\mathbf{y}^{2}}{2}+\mathrm{c}_{1} \mathbf{y}+\mathrm{c}_{2}\right) \tag{2.4}
\end{equation*}
$$

where $C_{1}$ and $C_{2}$ are constants of integration which can be evaluated by imposing the no slip condition at the boundary $\mathrm{y}= \pm \mathrm{h}$. By substituting these boundary conditions into Equation (2.4) yield

$$
\begin{equation*}
\mathrm{V}_{\mathrm{z}}=\frac{\beta}{2 \mu}\left(\mathrm{~h}^{2}-\mathrm{y}^{2}\right) \tag{2.5}
\end{equation*}
$$

which describes the familiar parabolic profile depicted in Figure 2.1. If we orient the plates such that the $x z$ plane is horizontal, Equations (2.3a) and (2.3b) specify that the pressure is independent of $x$ and $y$ giving no particular useful information here.

Now if one introduces a small quantity, say a cubical volume of side dimensions $k h(k<1)$ of a dyed fluid at the point $x=y=z=0$, then at any subsequent time $t$ later the center of the dye will be located at

$$
\begin{equation*}
\mathrm{z}=\frac{\beta \mathrm{h}^{2}}{2 \mu} \mathrm{t} \tag{2.6}
\end{equation*}
$$

according to Equation (2.5). The particles on the surface of the cube which was originally located at $z=0, x=0$, and $y=k h$ will be located at

$$
\begin{equation*}
\mathrm{z}=\frac{\beta \mathrm{h}^{2}}{2 \mu}\left(1-\mathrm{k}^{2}\right) \mathrm{t} \tag{2.7}
\end{equation*}
$$

any time $t$ later. Thus we see that the size of the dyed particle relative to the plate spacing is a measure of the smear along the $z$ axis. In the present application k should be made as small as possible. In principle then one should be able to calculate the dye distribution and hence the transparency distribution of the introduced dye as a function of space and time.

In order to use this principle for spatially modulating light a method must be devised to introduce the time dependent electric signal into the fluid. For this we have synthesized a magnetic responsive fluid similar to that employed by Papell and Faber of the Lewis Research Center and described by Neuringer and Rosensweig. The production and properties of this unique fluid will be discussed in Section IV below.

Basically the ferrofluid is a colloidal suspension of very fine (sub-micron) sized particles of ferrite such as $\mathrm{Fe}_{3} \mathrm{O}_{4}$ immersed in a carrier fluid such as JP-4 jet fuel. These fluids are highly responsive to magnetic fields. According to Neuringer and Rosensweig the force per unit volume induced into the fluid in the present of a magnetic field is given by

$$
\begin{equation*}
\overrightarrow{\mathbf{F}} / \Delta \mathrm{V}=\mu_{0} \mathrm{M} \nabla \mathrm{H} \tag{2.8}
\end{equation*}
$$

where is the free space permeability, $M$ is the magnetic moment per unit volume of the magnetic field, and H is the magnitude of the magnetic field. For magnetic fluids which are perfectly soft in the magnetic sense, the magnetization $M$ vanishes when the applied field $H$ is zero while at very high applied fields the magnetization must approach the constant, saturation value. For low to moderate applied fields it is reasonable to assume an empirical relationship of the form

$$
\begin{equation*}
\mathrm{M}=\mathrm{X} \mathrm{H} \tag{2.9}
\end{equation*}
$$

to hold, where the susceptibility is taken to be of the form

$$
\begin{equation*}
\mathrm{X}=\sigma(\mathrm{T}) \mathrm{H}^{\mathrm{n}-2} \tag{2.10}
\end{equation*}
$$

in which $\sigma(\mathrm{T})$ depends upon temperature only. Thus $\sigma$ and $n$ are magnetic material properties which must be determined by experiment.

By substituting Equations (2.9) and (2.10) into Equation (2.8) and by assuming an isothermal exchange of energy the force per unit volume can be written as

$$
\begin{equation*}
\overrightarrow{\mathrm{F}} / \Delta \mathrm{V}=\nabla \frac{\mu_{0} \sigma \mathrm{H}^{\mathrm{n}}}{\mathrm{n}} \tag{2.11}
\end{equation*}
$$

This indicates that the magnetically induced body force is conservative where the potential is proportional to the nth power of the magnitude of the magnetic field.

With the above results in mond let us study the following problem. Suppose a particle of opaque ferrofluid synthesized with a kerosene fluid as carrier be introduced into a transparent kerosene fluid which is flowing in the z direction between parallel glass plate as shown in Figure 2.2 and as discussed above.


Figure 2.2

As the particle moves along the $z$-axis with constant velocity $V_{z}$, it moves through a time dependent magnetic field which is generated by an electromagnetic external to the flow region. The potential for one convenient magnetic field can be written as

$$
\begin{equation*}
\phi=K_{1}[A+B \cos (\omega t+\psi)] y(x+c)\left[U_{-1}(z)-U_{-1}(z-a)\right] \tag{2.12}
\end{equation*}
$$

in which $K_{1}$ is a measure of the strength of the field; B $<A$ are electrical modulation constants, $C$ is a constant which depends on the shape of the poles, and $\mathrm{U}_{-1}(z)$ is the unit step function and $\psi$ is a phase angle. See Section III below for vari types of field shapes. In terms of the potential $\phi$ the magnitude of the magnetic field is given by

$$
\begin{equation*}
H=\sqrt{H_{x}^{2}+H_{y}^{2}}=\sqrt{\left(\frac{\partial \phi}{\partial x}\right)^{2}+\left(\frac{\partial \phi}{\partial y}\right)^{2}} \tag{2.13}
\end{equation*}
$$

By substituting Equations (2.12) and (2.13) into Equation (2.11) it can be shown that

$$
\begin{gather*}
\mathrm{F}_{\mathrm{x}} / \Delta \mathrm{V}=\mu_{0} \sigma \mathrm{n}\left\{\mathrm{~K}_{1}^{\mathrm{n}}[\mathrm{~A}+\mathrm{B} \cos (\omega \mathrm{t}+\psi)]^{\mathrm{n}}\left[\mathrm{U}_{-1}(\mathrm{z})-\mathrm{U}_{-1}(\mathrm{z}-\mathrm{a})\right]\left\{\mathrm{y}^{2}+(\mathrm{x}+\mathrm{c})^{2}\right\}^{\mathrm{n} / 2-1}(\mathrm{x}+\mathrm{c})\right\} \\
\mathrm{F}_{\mathrm{y}} / \Delta \mathrm{V}=\mu_{0} \sigma \mathrm{n}\left\{\mathrm{~K}_{1}^{\mathrm{n}}[\mathrm{~A}+\mathrm{B} \cos (\omega \mathrm{t}+\psi)]^{\mathrm{n}}\left[\mathrm{U}_{-1}(\mathrm{z})-\mathrm{U}_{-1}(\mathrm{z}-\mathrm{a})\right] \mathrm{y}\right.  \tag{2.14}\\
\mathrm{F}_{\mathrm{z}}=0
\end{gather*}
$$

If the particle entering the magnetic field is traveling in the $x-y$ plane then according to Equation (2.14) it will experience a force

$$
\begin{gather*}
\mathrm{F}_{\mathrm{x}}=\mu_{0} \sigma \Delta \mathrm{Vn} \mathrm{~K}_{1}^{\mathrm{n}}(\mathrm{~A}+\mathrm{B} \cos (\omega \mathrm{t}+\psi))^{\mathrm{n}}(\mathrm{X}+\mathrm{C})^{\mathrm{n}-1}\left[\mathrm{U}_{-1}(\mathrm{z})-\mathrm{U}_{-1}(\mathrm{z}-\mathrm{a})\right] \\
\mathrm{F}_{\mathrm{y}}=0 \\
\mathrm{~F}_{\mathrm{z}}=0 \tag{2.15}
\end{gather*}
$$

while in the magnetic field region.

Now it is desirable to be able to predict the location of the particle at any subsequent time after it is past the region of the magnetic field. By assuming laminar flow of the transparent fluid around the magnetically energized fluid particle, the drag experienced will be directly proportional to the relative velocity between the ferrofluid particle and the surrounding transparent fluid, i.e.,

$$
\mathrm{drag}=\mathrm{D} \dot{\mathrm{x}}
$$

where $D$ is a constant depending upon the shape and size of the ferrofluid particle relative to the plate spacing and the viscosity of the transparent fluid, and $\dot{x}$ its velocity in the $x$ direction. Then, to a very good approximation, the equations describing the motion of the ferrofluid particle will be

$$
\begin{gather*}
\rho_{\mathrm{f}} \Delta \mathrm{~V} \ddot{\mathrm{x}}=\Delta \mathrm{V} \mu_{0} \sigma \mathrm{n} \mathrm{~K}_{1}^{\mathrm{n}}(\mathrm{~A}+\mathrm{B} \cos (\omega \mathrm{t}+\psi))^{\mathrm{n}}(\mathrm{x}+\mathrm{c})^{\mathrm{n}-1}\left[\mathrm{U}_{-1}(\mathrm{z})-\mathrm{U}_{-1}(\mathrm{z}-\mathrm{a})\right]-\mathrm{D} \dot{\mathrm{x}} \\
\ddot{\mathrm{y}}=0, \quad \ddot{\mathrm{z}}=0 \tag{2.16}
\end{gather*}
$$

where $\rho_{\mathrm{f}}$ is the density of the ferrofluid. The initial condition are at $\mathrm{t}=0$ : $\mathrm{x}=\mathrm{y}=\mathrm{z}=\dot{\mathrm{x}}=\dot{\mathrm{y}}=0$, and $\dot{\mathrm{z}}=\mathrm{V}_{\mathrm{z}}=$ const.

As pointed out above the exponent $n$ is a property of the magnetic fluid. Neuringer and Rosensweig reports a value of $n=4 / 3$ for their fluid. It is reasonable to expect that the ferrofluid we synthesized at Goddard will possess magnetic properties of the same order of magnitude. If this is the case, then the forcing function given in Equation (2.16) depends rather weakly upon $x$, i.e., $x^{1 / 3}$. Furthermore, by assuming the induced displacement of the particle to be small while moving through the region of the magnetic field we may simplify our analysis considerably by saying that

$$
\begin{equation*}
c_{1}=(x+c)^{n-1} \tag{2.17}
\end{equation*}
$$

With this simplification Equation (2.16a) can be written in the form

$$
\begin{equation*}
\ddot{x}+\alpha \dot{x}=\beta[A+B \cos (\omega t+\psi)]^{n}\left[U_{-1}(z)-U_{-1}(z-a)\right] \tag{2.18}
\end{equation*}
$$

where

$$
\alpha=\frac{\mathrm{D}}{\rho_{\mathrm{f}} \Delta \mathrm{~V}}, \quad \beta=\frac{\mu_{0} \sigma \mathrm{nK}_{1}^{\mathrm{n}} \mathrm{C}_{1}}{\rho_{\mathrm{f}}}
$$

To obtain the general solution of Equation (2.18) we first expand the forcing function in a Fourier series as

$$
\begin{equation*}
\beta[\mathbf{A}+\mathbf{B} \cos (\omega t+\psi)]^{n}=\beta\left[\mathbf{A}_{0}+2 \sum_{n=1}^{\infty} \mathbf{A}_{m} \cos m(\omega t+\psi)\right] \tag{2.19}
\end{equation*}
$$

in which the $A_{m}$ can be determined from the Fourier inversion formula

$$
A_{m}=\frac{1}{\pi} \int_{0}^{\pi}[A+B \cos (\omega t+\psi)]^{n} \cos m(\omega t+\psi) d(\omega t+\psi)
$$

By employing Laplace transform techniques, Equation (2.16) may be integrated to yield

$$
\begin{align*}
x= & A_{0}\left(\frac{t}{\alpha}+\frac{1}{\alpha^{2}} e^{-\alpha t}-\frac{1}{\alpha^{2}}\right)-Z \sum_{m=1}^{\infty} \frac{A_{m} \cos m \psi}{(m \omega)^{2}+\alpha^{2}}\left[\cos m \omega t-\frac{\alpha}{m \omega} \sin m \omega t\right] \\
& +Z \sum_{m=1}^{\infty} A_{n} m \omega \sin m \psi\left[\frac{1}{m^{2} \omega^{2}\left(\alpha^{2}+m^{2} \omega^{2}\right)}(\alpha \cos m \omega t-m \omega \sin m \omega t)\right. \\
& \left.-\frac{1}{\alpha m^{2} \omega^{2}}+\frac{1}{\alpha\left(\alpha^{2}+m^{2} \omega^{2}\right)} e^{-\alpha t}\right] \tag{2.20}
\end{align*}
$$

for the transverse displacement of the particle, which is only applicable while the particle is in the region of the magnetic field.

Thus, as the particle leaves the magnetic field, $\mathrm{z}=\mathrm{a}$, its displacement will be

$$
\begin{align*}
x_{a}= & A_{0}\left(\frac{a}{a V_{z}}+\frac{1}{a^{2}} e^{-\alpha a / V_{z}}-\frac{1}{a^{2}}\right)-2 \sum_{m=1}^{\infty} \frac{A_{m} \cos m \psi}{(m \omega)^{2}+a^{2}}\left[\cos m \omega \frac{a}{V_{z}}-\frac{\alpha}{m \omega} \sin m \omega \frac{a}{V_{z}}\right] \\
& +2 \sum_{m=1}^{\infty} A_{n} m \omega \sin \operatorname{mi} \psi\left[\frac{1}{m^{2} \omega^{2}\left(\alpha^{2}+m \omega^{2}\right)}\left(\alpha \cos m \omega \frac{a}{V_{z}}-m \omega \sin m \omega \frac{a}{V_{z}}\right)\right. \\
& \left.-\frac{1}{\alpha m^{2} \omega^{2}}+\frac{1}{\alpha\left(\alpha^{2}+m^{2} \omega^{2}\right)} e^{-\alpha a / V_{z}}\right] \tag{2.21}
\end{align*}
$$

At this instant its velocity will be

$$
\begin{align*}
\dot{\mathrm{x}}_{\mathrm{a}}= & \frac{\mathrm{A}_{0}}{\alpha}-\frac{1}{\alpha} \mathrm{e}^{-\alpha a / \mathrm{V}_{z}}+2 \sum_{\mathrm{m}=1}^{\infty} \frac{\mathrm{A}_{\mathrm{m}} \cos \mathrm{~m} \psi}{\mathrm{~m}^{2} \omega^{2}+\alpha^{2}}\left[\mathrm{~m} \omega \sin \mathrm{~m} \omega \frac{\mathrm{a}}{\mathrm{~V}_{z}}+\alpha \cos m \omega \frac{\mathrm{a}}{\mathrm{~V}_{\mathrm{z}}}\right] \\
& -2 \sum_{\mathrm{m}=1}^{\infty} \mathrm{A}_{\mathrm{m}} \mathrm{~m} \omega \sin \mathrm{~m} \psi\left[\frac{1}{\mathrm{~m}^{2} \omega^{2}\left(\alpha^{2}+\mathrm{m}^{2} \omega^{2}\right)}\left(\alpha \mathrm{m} \omega \sin \mathrm{~m} \omega \frac{\mathrm{a}}{\mathrm{~V}_{z}}+\mathrm{m}^{2} \omega^{2} \cos m \omega \frac{a}{V_{z}}\right)\right. \\
& \left.+\frac{1}{\alpha^{2}+\mathrm{m}^{2} \omega^{2}} \mathrm{e}^{-\alpha a / v_{z}}\right] \tag{2.22}
\end{align*}
$$

To determine the transverse location of the particle at any subsequent time after leaving the magnetic field, we have for the motion equation

$$
\begin{equation*}
\ddot{x}=-a \dot{x} . \tag{2.23}
\end{equation*}
$$

This can be integrated simply to give

$$
\begin{equation*}
\dot{\mathbf{x}}=\dot{\mathbf{x}}_{\mathrm{a}} \mathrm{e}^{-a\left(\mathrm{t}-\mathrm{a} / \mathrm{v}_{\mathrm{z}}\right)} \tag{2.24}
\end{equation*}
$$

and

$$
\begin{equation*}
x=x_{a}-\frac{\dot{x}_{a}}{a}\left[1-e^{-\alpha\left(t-a / v_{z}\right)}\right] \tag{2.25}
\end{equation*}
$$

The $z$ displacement of this particle will be

$$
\begin{equation*}
\mathrm{z}=\mathrm{V}_{\mathrm{z}} \mathrm{t} \tag{2.26}
\end{equation*}
$$

Therefore, a particle location at $z=V_{z} t$ will have a lateral displacement

$$
\begin{equation*}
\mathrm{x}=\mathrm{x}_{\mathrm{a}}-\frac{\dot{\mathrm{x}}_{\mathrm{a}}}{\alpha}\left[1-\mathrm{e}^{-\alpha\left(\mathrm{z} / \mathrm{v}_{\mathrm{z}}-\mathrm{a} / \mathrm{v}_{\mathrm{z}}\right)}\right] \tag{2.27}
\end{equation*}
$$

In order to calculate the shape formed by a continuous line of streaming particles, it is necessary to relate the phase angle $\psi$ of the magnetic field with the particular particle under consideration. Assuming the phase angle of the magnetic field was $\psi_{0}$ when the particle which is located at $z=a$ entered the magnetic field at $z=0$, then the phase angle for the particle located at any $z>a$ would be

$$
\begin{equation*}
\psi=\psi_{0}+\omega\left(\frac{z-\mathrm{a}}{\mathrm{v}_{\mathrm{z}}}\right) \tag{2.28}
\end{equation*}
$$

By substituting Equation (2.28) into Equations (2.21) and (2.22) and then substituting the results into Equation (2.27), we can derive the shape of the resulting pattern of streaming particles at any instant. Due to lack of time, we were unable to pursue the analytic solution further.

In order to see if this type modulator was feasible, several cells were constructed and several different modulation configurations were tried. Descriptions of some of these will be found in the next section of this report.

## DEFLECTION SYSTEMS

In this section we discuss several possible methods for deflecting the stream of ferrofluid which is being transported in the nonmagnetic clear carrier fluid. Each of these methods are discussed on the premise that the ferrofluid can be displaced in the carrier fluid by inducing an electromagnetic body force in the ferrofluid.

The basic equation for governing the design of possible deflection systems is that given by Equation (2.11). Two of the systems studied are analyzed on the next page.

## Single Wire Normal to the Fluid Flow

As demonstrated by Rosensweig, a ferrofluid with a high magnetization $M$ can have impressed upon it a measurable force by passing a current through a wire located near the fluid. Thus, let us consider the force induced in a small volume $\Delta V$ of ferrofluid which is located near a wire carrying a current I as shown in Figure 3.1.


Figure 3.1

It is well known that the magnetic field in the region exterior to a current carrying straight wire whose center is located along the y axis is

$$
\begin{equation*}
\overrightarrow{\mathrm{H}}=\frac{\mathrm{I}}{2 \pi \mathrm{r}} \vec{\tau} \tag{3.1}
\end{equation*}
$$

where $I$ is the current, $r$ is the radial distance from the center of the wire to the point in question, and $\vec{\tau}$ is a unit vector normal to the radius vector $\vec{r}$ and the unit vector $\vec{j}$ along the y axis. By substituting Equation (3.1) into Equation (2.11), the force acting on a unit volume located at ( $x, o, z$ ) is

$$
\begin{equation*}
\frac{\overrightarrow{\mathrm{F}}}{\Delta \mathrm{~V}}=-\mu_{0} \sigma\left(\frac{\mathrm{I}}{2 \pi \mathrm{r}}\right)^{\mathrm{n}-1} \frac{\mathrm{I}}{2 \pi \mathrm{r}^{2}}\left[\frac{\partial \mathrm{r}}{\partial \mathrm{x}} \overrightarrow{\mathrm{i}}+\frac{\partial \mathrm{r}}{\partial \mathrm{y}} \overrightarrow{\mathrm{j}}+\frac{\partial \mathrm{r}}{\partial \mathrm{z}} \overrightarrow{\mathrm{k}}\right] \tag{3.2}
\end{equation*}
$$

But, in the present case $r=\sqrt{x^{2}+y^{2}}$. Thus, equation (3.2) reduces to

$$
\begin{equation*}
\overrightarrow{\mathbf{F}} / \Delta \mathbf{V}=\mu_{0} \sigma\left(\frac{\mathrm{I}}{2 \pi \mathbf{r}}\right)^{\mathrm{n}} \frac{1}{\mathbf{r}}\left[\frac{2 \mathbf{x}}{\mathrm{r}^{3 / 2}} \overrightarrow{\mathbf{i}}+\frac{2 \mathrm{z}}{\mathbf{r}^{3 / 2}} \overrightarrow{\mathbf{k}}\right] \tag{3.3}
\end{equation*}
$$

Assuming a value of $n=4 / 3$ this last expression can be written as

$$
\begin{equation*}
\overrightarrow{\mathrm{F}}=-2 \sigma \Delta \mathrm{~V}\left(\frac{\mathrm{I}}{2 \pi}\right)^{4 / 3} \frac{\mathrm{x} \overrightarrow{\mathrm{i}}+\mathrm{z} \overrightarrow{\mathrm{k}}}{\left(\mathrm{x}^{2}+\mathrm{y}^{2}\right)^{17 / 12}} \tag{3.4}
\end{equation*}
$$

This equation reveals some interesting results concerning the force on the small particle as it passes the wire. For a small volume of ferrofluid, the maximum force exerted will occur when the particle is closest to the wire. For example, if the particle is in contact with the wire as it crosses the $x$ axis, the force will be

$$
\begin{equation*}
\mathrm{F}=-2 \sigma \Delta \mathrm{~V}\left(\frac{\mathrm{I}}{2 \pi}\right)^{4 / 3} \mathrm{R}^{-11 / 6} \overrightarrow{\mathrm{i}} \tag{3.5}
\end{equation*}
$$

where $R$ is the radius of the wire. If $\Delta V$ is very small, either the wire must be very small or the current through the wire must be very high to expect any measurable force. However, the maximum current is limited by the burnout of the wire. The burnout current can be approximated by $I_{\max }=K R^{4}$. Thus we see that

$$
\begin{equation*}
\mathrm{F}_{\max }=2 \sigma \Delta \mathrm{~V} \frac{\mathrm{KR}^{16 / 3}}{(2 \pi)^{4 / 3}} \frac{1}{\mathrm{R}^{11 / 6}} \text { e } \mathrm{R}^{3.5} \Delta \mathrm{~V} \tag{3.6}
\end{equation*}
$$

where e is a bounded constant. This last equation indicates that this system is impractical in the present application.

There are several ways to shape the poles of an electromagnet for use in deflecting the stream of ferrofluid. Since we are interested in a system where the pole faces are exterior to the fluid region, we use the following analysis. To find the shape of the poles it will be assumed that the permeability of the core material is of sufficient quality such that the pole faces form magnetic equipotential surface. Since $\mu$ of iron is greater than 1000 , this is a reasonable assumption. The problem now at hand is to determine the shape of the equipotential surfaces to give the desired field along the x axis.

Let the fluid be flowing along the $z$ direction. It is desired to induce into the ferrofluid a force along the x axis with a magnitude which is approximately independent of x . Thus, according to Equation (2.11) we set

$$
\begin{equation*}
\vec{F} / \Delta V=\mu_{0} \sigma H^{n-1}\left(\frac{\partial A}{\partial x} \vec{i}+\frac{\partial H}{\partial y} \vec{j}+\frac{\partial H}{\partial z} \vec{k}\right)=-K \vec{i} \tag{3.7}
\end{equation*}
$$

where $K$ is weakly dependent upon x . Also, the H vector field must satisfy

$$
\begin{equation*}
\operatorname{Cur1} \overrightarrow{\mathrm{H}}=0, \quad \operatorname{div} \overrightarrow{\mathrm{H}}=0 \tag{3.8}
\end{equation*}
$$

In order to solve Equation (3.7) along with (3.8), we assume that $\eta$ is close enough to unity that $\mu_{0} \sigma \mathrm{H}^{\mathrm{n}-1}=$ constant. Also, in order that the pole faces be exterior to the flow we require that the $H$ field be normal to the x -axis, i.e., $\mathrm{H}_{\mathrm{x}}=\mathrm{H}_{\mathrm{z}}=0$ here. Thus we write

$$
\begin{equation*}
\frac{\partial H}{\partial x}=\frac{\partial H y}{\partial x}=-K_{1} \tag{3.9}
\end{equation*}
$$

where

$$
\mathrm{K}_{1}=\frac{\mathrm{K}}{\mu_{0} \sigma \mathrm{H}^{\mathrm{n}-1}}
$$

Now according to Equation (3.8)

$$
\begin{equation*}
\overrightarrow{\mathrm{H}}=-\nabla \phi \text { and } \nabla^{2} \phi=0 \tag{3.10}
\end{equation*}
$$

where $\phi$ is the magnetic potential.

Therefore, along the x axis

$$
\begin{equation*}
\frac{\partial^{2} \phi}{d x d y}=K_{1} \tag{3.11}
\end{equation*}
$$

One solution that will satisfy both Equations (3.10) and (3.11) is

$$
\begin{equation*}
\phi=\mathrm{K}_{1} \mathrm{y}\left(\mathrm{x}+\mathrm{C}_{2}\right) \tag{3.12}
\end{equation*}
$$

Now to find the shape of the pole faces, it is only necessary to set $\phi$ equal to a constant and to choose a suitable value for $C_{2}$. For example, the shape of the pole faces used in the experimental part of this program it was assumed that

$$
\begin{align*}
\phi / \mathrm{K}_{1} & = \pm 0.06 \\
\mathrm{C}_{2} & =-0.6 \tag{3.13}
\end{align*}
$$

Substituting these values into Equation (3.12) the shape of the poles becomes

$$
\begin{equation*}
y=\frac{ \pm 0.06}{(x-0.6)} \tag{3.14}
\end{equation*}
$$

The pole shapes as given by Equation (3.14) are shown in Figure 3.2.

## DISCUSSION OF PROPERTIES AND PRODUCTION OF FERROMAGNETIC FLUIDS

Ferromagnetic fluids are colloidal suspensions of very small (sub-micron) size ferromagnetic particles immersed in an organic carrier fluid such as kerosene. A dispersing agent is added to prevent flocculation of the iron particles. These ferrofluids are not commercially available. To synthesize a ferrofluid, a mixture of magnetic iron powder, a surfactant, and the carrier fluid is prepared and ground for several days in a ball mill.

Papell and Faber of the Lewis Research Center synthesized a magnetic colloid which they employed to simulate zero- and reduced gravity to study the effect of gravity on pool-boiling systems.


Figure 3.2

Neuringer and Rosensweig discuss the fluid dynamics and thermodynamics of magnetic responsive fluids. They treat the fluid as a continuum and develop what they refer to as the Ferrohydrodynamics equations of motion.

These ferrofluids are quite different from the magnetic clutch materials used in the 1940's. The magnetic fluids discussed here are truly colloidal in that the sufficiently fine particles can be suspended indefinitely in a liquid even though the density of the magnetic particles is greatly different from that of the liquid. A qualitative explanation of the mechanics of interaction of the particles will be helpful in understanding its manufacture and its response to magnetic fields. If we consider a distribution of small spheres in a liquid we can envision that the thermal agitation of the molecules striking the individual spheres will prevent them from flocculating and settling out. Now should these spheres be magnetized the magnetic energy of attraction is proportional to the square of the radius of the particle if they touch and inversely proportional to the cube of the radius. Therefore, if the radii of the spheres are made small enough, it is conceivable that the random impact of the molecules would be sufficient to prevent flocculation of the magnetized particles. For normal temperatures and for the magnetic
materials available, it can be shown that particles of 25-100 Anstroms in diameter would be stable under these conditions. However, when two particles of such small diameter are brought into contact there will be an attraction present due to the Van der Waal forces. To counteract this latter type force, we can imagine coating each sphere with an elastic material whose force originates from the deformation of the surface. Therefore, it can be seen that the secret for synthesizing a ferrofluid is to grind the ferromagnetic particles to a very fine size and at the same time use a surfactant to coat the individual particles.
2. Heptane Ferrofluid in Water Carrier. At our request, Stephen Papell of the Lewis Research Center, sent us a very small quantity of the ferrofluid they were using in some of their work. This fluid was synthesized with a Heptane carrier base. We tried to use this ferrofluid in place of the dye in the water cell. This proved to be fruitless because the heptane ferrofluid immediately balled up and diffused against the glass plates.
3. JP-4 Jet Fuel and Heptane. By using clear JP-4 jet fuel for the main flow and heptane ferrofluid for the dye, a very fine line could be maintained. This line could be deflected quite readily with a permanent magnetic.
4. JP-4 and JP-4 Ferrofluid. After some ferrofluid of our own manufacture was obtained, several types of deflection methods were tried. These methods are described below with comments about their operation. Again, the order is chronological.
a. Single wire transverse to the flow. A small hole was cut through both glass plates and a wire was passed through the holes. This wire was then transverse to the fluid flow and was small enough so that it did not have much effect on the flow. A small stream of ferrofluid was inserted close to the wire. A current was then passed through the wire and the effect on the stream noted.

The effect on the stream was zero. The measurement techniques were not extremely accurate, and this method was abandoned for the present.

It is possible that this method may still work. Referring to Rosenswig's paper, it is noted that he was able to exert considerable force on a ferrofluid using a current through a wire. For example, he caused a ferrofluid to climb up a wire with a current in it.

This method has an inherent non-linearity which makes it potentially difficult to control. For a theoretical discussion see Section III.
b. A single $0.006^{\prime \prime}$ thick lamination. A transformer lamination $0.006^{\prime \prime}$ thick and $1 / 2^{\prime \prime}$ wide and $1-1 / 2^{\prime \prime}$ long was inserted through the side of the cell parallel to the glass plates. Part of this lamination was left sticking outside the fluid and around this part a 100 turn coil was wound. A ferrofluid stream was passed close (approximately $0.1^{\prime \prime}$ ) to the end of the lamination which was in the fluid.

Since we were unable to obtain a sufficient quantity of existing magnetic fluid for our purposes, it was necessary to produce a quantity of fluid here at Goddard.

First, an attempt was made to grind the mixture using a standard ball mill. After two weeks of continuous grinding with a ball mill located in the Material Research and Development Branch, no appreciable colloid had formed. After a telephone conversation with Stephen Papell of the Lewis Research Center, we decided to design and construct a fast-grinding machine or attritor. The details of the attritor are shown in Figure 4.1.


Figure 4.1


Figure 4.2


Figure 4.3

Figure 4.4

The attritor was mounted on a drill press (as shown in Figures 4.2(a) and 4.3 ) which can be driven at four different speeds. To make the ferrofluid, the attritor was charged with 450 cc of JP-4 jet fuel, 25 cc or oleic acid, and 60 grams of $\mathrm{Fe}_{3} \mathrm{O}_{4}$ powder along with $50003 / 16^{\prime \prime}$ diameter chrome alloy steel ball bearings. The attritor was run at approximately 490 RPM continuously for one week. Cooling water was circulated through the jacket for temperature control of the mixture. To determine if any colloid was being formed, we extracted a small sample of fluid from the attritor and centrifuged it. If after one day of grinding the sample appeared opaque and responsive to a small permanent magnet, we were confident that a ferrofluid was forming and the run was continued. The response of a sample to a permanent magnetic is shown in Figure 4.4. This particular sample was extracted at the end of three days' grinding time.

## Experimentation

In order to ascertain whether this method of modulation has merit, we constructed several cells and experimented with different techniques for converting the input signal into some type of corresponding transparency pattern. These experiments are discussed in chronological order.

1. Water Cell. To demonstrate that the laminar flow could possibly be used for signal storage, a cell was constructed using two $3^{\prime \prime} \times 5^{\prime \prime}$ photographic glass plates with a spacing of $0.012^{\prime \prime}$. Strips of adhesive tape was used for spacing the plates and two $1 / 4^{\prime \prime}$ rubber hose were split and mounted on the ends for channeling the fluid through the plates. The system was sealed with epoxy. Water was syphoned through the assembly and a food coloring dye stream was inserted into the upstream portion between the plates with a hypodermic needle. The system is a type of Hele-Shaw apparatus. It was demonstrated that a very fine steady line could be maintained. A photograph of this cell is shown in Figure 5.1.

When a current was applied to the coil the fluid was deflected. The problem was that the fluid tended to stick to the lamination. This no doubt is due to the large magnetic gradient at the surface of the lamination.
c. Pair of laminations outside the fluid. Two laminations of the size of the previous case were placed normal to the glass plates, one on either side of the cell, as shown in Figure 5.3. Coils were wound on them. They were energized to give first two north poles against the glass and then a north and a south against the glass.

The ferrofluid stream tended to stick to the inside of the glass plates and it was not possible to deflect the stream in a useful way.


Figure 5.1

Figure 5.2


Figure 5.3
d. A tape wound core having a cross section of $1 / 4^{\prime \prime} \times 1 / 4^{\prime \prime}$ was cut to form pole shapes as derived in method a. above. The gap was about $0.2^{\prime \prime}$ on one edge and about $0.4^{\prime \prime}$ on the other. The shape was approximately that shown in Figure 3.3 and mounted as shown in Figure 5.4.

This electromagnet was used in conjunction with a cell with $0.1^{\prime \prime}$ spacing between the plates. A 300 turn coil was wound on the torricd and a d.c. current of $0-1$ ampere was passed through the coil.

The stream was deflected, but due to the mechanical instability of the system it was difficult to make any accurate measurements. For example, slight heating of the glass plates by the magnet caused large deflections. Another problem with this type of deflection is the large area over which the magnetic field acts. In some cases, it appeared that the magnetic field might cause a longitudinal instability of the ferrofluid stream.

To continue in this direction of investigation, it would be necessary to construct a second generation of fluid cells where the various parameters are under better control.

## CONCLUSION

It has been demonstrated that the laminar flow of a fluid between parallel plates can be used to store information in the form of an opaque fluid immersed in a moving transparent carrier fluid. This is possible because in laminar flow a signal of dark fluid inserted in the clear carrier fluid holds its shape as it flows through the space between the parallel plates. The plates are made of glass and a laser light beam could be passed through the storage cell to read out the stored information.

Several cells of different configurations were constructed to determine the problems associated with immersing information into the fluid stream between the plates. The method which appeared most promising was to use a small stream of dark ferrofluid in a clear carrier fluid. This small stream can be deflected by an electromagnetic. However, due to the fact that this method involves the mechanical movement of the dark fluid relative to the clear carrier, and also because of difficulties in containing a desirable magnetic field within a small region, the frequency response is poor. Due to these difficulties, it is recommended that other techniques be used in generating the pattern.


Figure 5.4

Other methods that warrant consideration are the following:

1. Use an electromagnetic ferrofluid transducer to control the flow rate of the opaque fluid before immersing it in the clear carrier fluid. For example, wind a coil around the hypodermic needle.
2. Design a surface wave generator employing the ferrofluid and read out the stored information out by optical techniques. The advantages of this method lie in the fact that a simulated gravity field can be generated with a permanent magnetic field thus allowing control of the wave propagation speed and at the same time it appears quite easy to generate the desired surface wave pattern with another independent electromagnet.
3. Use a cell as described herein but generate the transparency pattern with techniques that will not disturb the main flow of the clear carrier fluid. For example, generate the dye pattern by a flash Photolysis process as described by Popovich and Hummel.

## ACKNOWLEDGMENTS

The authors wish to express their appreciation to S. S. Papell and O. C. Faber of Lewis Research Center for their advice on the manufacture of the ferrofluid. Thanks are due to A. Shulman and his group for their help and to Ron Hunkeler for the use of his laboratory.

## REFERENCES

1. Rosensweig, R. E., "Magnetic Fluids." International Science and Technology, July 1966, pp. 48-56.
2. Papell, S. S., and O. C. Faber, Jr., "Zero- and Reduced-Gravity Simulation on a Magnetic-Colloid Pool-Boiling System." NASA TN D-3288, February 1966.
3. Neuringer, J. L., and R. E. Rosensweig, "Ferrohydrodynamics." Physics of Fluids, Vol. 7, No. 12. December 1964. pp. 1927-32.
4. Popovich, A. T., and R. L. Hummel, "A New Method for Nonturbing Turbulent Flow Measurements Very Close to a Wall." Chemical Engineering Science, Vol. 22, 1967, pp. 21-25.

ABSTRACTS

# PRECEDING PAGE BLANK NOT FILMED. 

## AUTOMATIC PROCESSING OF IONOGRAMS

by<br>William A. Barrett ${ }^{\dagger}$

The project that I undertook this summer as an ASEE-NASA Faculty Fellow had as its purpose the automatic processing of data from the topside sounder experiment aboard the Defense Research Telecommunications Establishment's Alouette satellite series. This data, commonly referred to as an 'Ionogram', is used to study the ion density in the upper layers of the atmosphere. This is accomplished by transmitting a swept-frequency signal and recording the time required for the echo to return and also the magnitude of the echo.

In my work I was involved with Dr. F. P. Kuhl in the formation of the logic structure that would serve to automatically classify the signal. This phase of the project consisted of modifying the procedure used in a digital computer character recognition program. The modification was necessary to overcome the noisy background of the ionogram. We accomplished this phase by using the main characteristics of the signal as a key.

While Dr. Kuhl proceeded to write the precise logic, I turned to enhancing the analog signal. My method called for the signal to be repeatedly integrated and squelched. I hoped that this would decrease the noise by essentially serving as a sharp cutoff filter. After assembling the necessary equipment and designing and building a squelch circuit, I found that this method gave a signal-to-noise voltage gain of approximately two.

This work is published in X-565-67-380.

[^17]
# EXPERIMENTAL AND THEORETICAL STUDIES OF THE CARBON DIOXIDE LASER 

by

John R. Bolte ${ }^{\dagger}$

The Quantum Optics Section at the Goddard Space Flight Center is concerned with the problem of developing a deep space communications capability using a high powered carbon dioxide laser operating at 10.6 microns. Part of this workinvolves a carefulstudy of the characteristics of the laser and is under the direction of Mr. Nelson McAvoy. Specific studies during the summer have included:

1. Lifetime Studies of Sealed Off Laser Tubes. Until very recently $\mathrm{CO}_{2}$ lasers had been operated only in gas flow systems because the lifetime of sealed off tubes was very short. Tube lifetimes of 2000 hours now appear to be possible using $\mathrm{CO}_{2}$ and helium mixtures and careful tube cleaning techniques.
2. Improvement of Techniques for RF Excitation of Plasma Tubes. The problem is to find efficient means of coupling rf power into the laser tube. New rf electrode designs are now in use which confine the electric field to the plasma region and additional studies are to be made in which the electrodes are placed inside the glass cooling jacket.
3. Experimental Work with Detectors for 10.6 Micron Laser Radiation. Thermal detectors in the form of thermopiles and bolometers as well as semiconductor detectors are available in the 10.6 micron infrared region. Thermal detectors have the disadvantage of having a slow response time while semiconductor detectors have to be cooled to liquid nitrogen or liquid helium temperatures. All of the devices have good sensitivity, however. Thermistor bridge detectors were constructed during the summer which could easily detectinfrared radiation from the hand held several feet from the detector.
4. Theoretical Studies of the Resonant Modes in a Fabry-Perot Optical Cavity. The journal articles by Fox and Li, Boyd and Gordon, and by Kogelnik have given abbreviated theoretical treatments of optical resonant cavities. A thorough understanding of these articles is necessary so that meaningful experimental work can be done in the design of mode stable lasers. Further theoretical study as well as experimental work is presently underway which should lead to the development of highly stable lasers which will be suitable for communications work.
[^18]
# A PERTURBATION THEORETIC TREATMENT OF THE FERMI FUNCTION 

by
B. $\mathrm{Chern}^{\dagger}$

During my tenure as an ASEE-NASA Fellow at the Laboratory for Theoretical Studies of the Goddard Space Flight Center, I launched a Field-theoretic perturbative treatment of the Fermifunction which is used in nuclear beta decay. A purpose of the investigation was to see whether the logarithmic divergence which occurs in the Fermi function when the nuclear radius goes to zero also occurs in the perturbative treatment. The larger overall objective is to obtain precise values of the vector coupling constant and to see whether the radiative corrections to beta decay can be made consistent with a Universal Fermi interation.

This work has direct application to nuclear beta decay processes in stellar interiors

[^19]
# BIT ERROR PROBABILITY TESTS FOR MEASURING THE PERFORMANCE OF THE UNIFIED S-BAND GROUND STATIONS 

by

David A. Dalgleish ${ }^{\dagger}$

## Project:

This summer project consisted of learning more about modern communications theory and such basic concepts as probability of bit error, noise, digital codes and formats, bi-phase and phase modulation, phase detection and phaselocked loops, and system signal-to-noise ratios.

Specifically, the study was directed at a Bit Error Probability test that is part of the Station Readiness Tests of the single 30 foot, split 30 foot, and 85 foot unified S-band (USB) stations that are part of the Manned Space Flight Network found around the world. It was desired to theoretically describe the test. The results are partly descriptive and partly theoretical.

## Results:

The test uses a simulator to put out a pulse code modulated (PCM) telemetry signal similar to that put out by the Apollo Spacecraft. This signal has a nonreturn to zero-level (NRZ-L) code format for the PCM which is bi-phased modulated on a 1024 KHZ sub-carrier. The sub-carrier is in turn used to phase modulate the main 2287.5 MHZ spacecraft down-link channel main carrier, which in this case is a simulated main carrier.

The simulated signal is then the input to the USB system either through the collimation tower and antenna or directly into the parametric amplifiers with the antenna held at a "quiet sky" or zenith position. This input signal goes through the processing of the double super-heterodyne S-band receiver. The 1024 KHZ sub-carrier is stripped off and, in turn the PCM data is removed in the Signal Data Demodulator System. The noisy serial bit train is processed through a signal conditioner and compared with the original signal in a comparator. The bit error count, actually a not-error count, is made for inputsignals having a variety of signal to noise (carrier to noise) levels.

[^20]In the test of the USB and PCM equipment, it was felt that the theoretically best result was

$$
\mathrm{P}_{\mathrm{e}}=\frac{1}{2} \mathrm{P}(1 / \mathrm{D})+\frac{1}{2} \mathrm{P}(\mathrm{D} / 1)
$$

where

$$
P(1 / D)=P(D / 1) \quad P_{e}=\frac{1}{2}\left[1-\operatorname{erf} \frac{1}{2} \sqrt{\frac{S}{N}}\right]
$$

where $\mathrm{S} / \mathrm{N}$ is a power ratio and erf

$$
\mathrm{x}=\frac{2}{\sqrt{\pi}} \int_{0}^{\mathrm{x}} \mathrm{e}^{-\mathrm{y}^{2}} \mathrm{dy}
$$

The actual signal to noise ratio for the various parts of the system can then be determined using effective bandwidths and temperatures. With this KTB concept, a final adjustment in the $S / N$ can be made so that a final curve of probability of error vs input carrier to noise ratio can be made. This curve shows the best possible condition in the system. Somewhat arbitrarily, an acceptable margin of +5 db was selected; i.e., the carrier-to-noise vs probability of error curve can be shifted +5 db to give an upper acceptable bound for the test.

Data which has been collected from many stations in the MSFN shows this test to be very useful in determining system performance and finding problems in the loop.

## Future Work:

Future work should be spent on diminishing the 5 db acceptable margin. Since the theoretical curve is obtained assuming an equal probability of "ones" and "zeros," some work will have to be done on simulation signals that do not have equal numbers of ones and zeros. Another aspect to consider is the transition density percentage. This means that alternate 1-0 formats will give a maximum number of transitions ( $100 \%$ ) and a long series of one's followed by a long series of zeros (for example 64 ones then 64 zeros) will give a very small number of transitions (about 2\%). The effects of transition density on probability of error then remains to be determined. Other probable causes of error would
be non-gaussian noise, a quiet sky not being quiet, inaccuracies in setting in the input signal to noise levels for the test, and long time temperature change effects since the overall test takes 3-4 hours.

## Acknowledgments:

I am indebted to my colleague, David J. Graham, Operations Analysis Section, for keeping me on a correct course, and to Wallace Goode and many others of the Manned Flight Operations Division who made this summer so meaningful.

# AN OPTICALLY INDUCED FREE CARRIER LIGHT MODULATOR 

by

Dr. Carl L. Gruber ${ }^{\dagger}$

My work as an ASEE-NASA Summer Faculty Fellow involved theoretical and experimental studies of a wideband optical modulator to be used in infra-red communication systems. The basic concept evolved from prior optical modulator development by my NASA colleague, Mr. William Richards and the theory from my investigation of problems related to other free carrier optical modulators.

A theoretical model for the modulator using a semiclassical approach was developed during the initial investigations. Using this model the optical transmission characteristics, depending upon several material properties and modulation power level, were determined. The final phase of development involved experimental verification of the model. Since limited time was available, complete experimental definition of the modulator was not possible. However, work will continue in this direction both at NASA and at South Dakota School of Mines.

The basic modulator is described by two components. A slab of semiconducting material possessing good transmission properties at the wavelength to be modulated and a source of readily modulated optical radiation at the band gap energy of the semiconductor material (a current modulated diode laser is ideal). The source illuminates the surface of the slab generating an inhomogeneous distribution of free carriers in the semi-conductor by electron-hole pair production. The generated free carriers in turn alter the complex optical index of refraction of the material in proportion to the incident source power level. Either phase of amplitude modulation may be effected depending upon material properties and incident source power.

This device exhibits characteristics superior to those of other modulators presently in use. The bandwidth should approach 1 GHz with $10 \%$ modulation power efficiency (three orders of magnitude better than previous devices) and further requires no direct electrical connections to the modulation element.

This work has resulted in a patent disclosure and NASA Tech Note coauthored by Mr. Richards and myself. A paper is to be submitted for formal publication in Proceedings of the IEEE pending further experimental definition of the device.

[^21]
# AUTOMATIC PROCESSING OF IONOGRAMS 

by

Frank P. Kuhl ${ }^{\dagger}$

The ionograms obtained from the ALOUETTE topside sounder are analyzed to show the variation with height of the electron density in the ionosphere. Ionograms can be collected over all regions of the earth for any time of day since the orbit of the ALOUETTE satellite traverses each region periodically every three months. The typical appearance of an ionogram is that of a pair of interwoven lines shaped like an " S " rotated 90 degrees counterclockwise. Figure 1 shows good ionograms.

An automatic system for machine encoding and identification of ionograms was proposed by the author in the document X-565-67-380, August 1966, Goddard Space Flight Center, as a result of last summer's work. This summer work toward implementing this system will be published in a new X-document entitled "Automatic Processing of Ionograms." Briefly the system consists of the functional blocks shown in Figure 2. Although complete specification of the operations of the classification blocks ( $3,4,5$, and 6 ) is not yet possible, such classification can now be shown to be feasible as a result of the development of FORTRAN flow charts for noise removal, continuity insertion and image encoding (blocks 1 and 2). The types of classification that can be developed are suggested in $X$-document X-565-67-380.

[^22]
(a)

(b)


Figure 1-Good lonograms
IDENTIFICATIONS AND

|  |  |
| :---: | :---: |



# PCM TELEMETRY ANALYSIS 

by

Nicholas Kyriakopoulos ${ }^{\dagger}$

The study was concerned with signal degradation in telemetry systems. The ultimate goal of such study would be the development of a mathematical model for a PCM telemetry system consisting of: antenna, preamplifier, multicoupler, receiver and data handling equipment. This model would be used for computer simulation and evaluation of telemetry systems.

The first phase of the program involved the investigation of the behavior of phase-locked loops, used in the detection and bit synchronization stages, in the presence of noise. The system considered was a second order loop whose phase modulated input was corrupted by additive, Gaussian noise. An analytical solution of the differential equation involving the probability distribution of the output signal was sought. It has, however, been shown that the output of the phaselock loop under the given conditions forms a Markov process and the differential equation involving the probability distribution of such process has as of today, no solution. It was, therefore, decided to abandon seeking analytical solution to these problems and proceed along the road of computer simulation.

The second problem investigated was the operation of a filter in the S-band transponder aboard a spacecraft which is part of the Range and Range Rate system. A question had been raised as to how closely the actual filter approached the claimed maximally flat time delay filter.

A computer program to analyze the circuit was used and the computed response did not approach the claimed ideal one. The investigation will continue.

[^23]Richard L. Martin ${ }^{\dagger}$

## Problem Statement

The problem considered here is:

1. Given a pulse frequency modulated signal (PFM) consisting of finite duration bursts of constant but unknown frequency distorted by clipping and corrupted by Gaussian noise.
2. Develop a computer algorithm which will extract the fundamental frequency component of each individual burst and give an indication of the signal-to-noise ratio.

Hardware Implementation
There is an existing hardware implementation which performs the stated function. It consists of a comb filter bank, switching logic and a zero-crossing counter.

The comb filter bank consists of 128 narrow band filter, each having a 100 Hz bandwidth, with each center frequency separated by 100 Hz covering the entire range of possible received signals ( 3.6 kHz to 16.3 kHz ). The received signal plus noise is fed into all elements of the comb filter simultaneously. As the signal progresses in time, the outputs of the various elements built up until one element reaches a predetermined nominal level. At this point a low resolution estimate of the frequency is obtained, i.e., the center frequency of that particular element.

The output of that element is then used as the input to the zero-crossing counter. The total noise power contained in the signal has been reduced by the ratio of the filter element equivalent noise bandwidth to the input noise bandwidth. Also the frequency of the signal is known within $1 \%$. These two results, a relatively clean signal and a fairly good estimate of the frequency, are necessary for the operation of the zero-crossing counter.

[^24]Several methods were considered for the digital computer frequency measurement, all of which use a zero-crossing counting scheme for the final high accuracy measurement.

The first method, a digital simulation of the comb filters, was rejected immediately as requiring a prohibitive member of computer manipulations. Assuming a 10 ms burst with a 20 kHz bandwidth, sampling of the signal must be done at a minimum 40 kHz rate resulting in at least 400 samples. Processing these 400 samples through 128 different filters was considered an excessive amount of work.

A second method consists of cross-correlating the input signal with 128 different square waves, the frequency of each being the center frequency of the filters stated previously. By using unit magnitudes for the square waves each cross-correlation is reduced to a summation rather than an actual multiplication plus summation procedure. If the input signal frequencies were from a discrete spectrum this would probably be the best technique to use. However, since the spectrum is continuous the possibility exists that a frequency lying midway between two correlation frequencies would be missed. Therefore, a technique which would give a smooth approximation to the total input spectrum was sought. One such technique has been derived from adaptive filter theory.

This project has been mainly concerned with the application of the adaptive filter technique. Briefly stated, this technique is as follows. Given a set of samples representing a constant amplitude signal plus noise, construct a sample data filter (composed of denominator terms only). This filter is to be such that if it were driven by white Gaussian noise, the output spectrum of the filter would give an approximation to the spectrum represented by the given samples. The sharpest pole of the filter should then represent the signal which was corrupted by the noise. Although the filter is derived mathematically by minimizing the mean square error of the output of the inverse filter when driven by the given samples, the implementation of the actual method is straightforward.

The derivation of this method plus results from computer runs have been submitted to GSFC.

# AN EXPLORATION ON HANSEN'S METHOD OF PARTIAL ANOMALIES 

by
George E. McCluskey, Jr. ${ }^{\dagger}$

Hansen's method of partial anomalies was investigated in order to determine if it would be suitable for use in the calculation of the perturbations of artificial satellites and space probes having highly eccentric orbits. This method has received very little attention since its publication in 1856. The main reason for this appears to be the lack of celestial bodies having highly eccentric orbits with the orbital parameters determined with sufficient accuracy to warrant a perturbation calculation. With the advent of the space age, this is no longer the case.

The commonly used methods of celestial mechanics express the perturbations of the perturbed body in trigonometric series involving multiples of the mean anomaly, eccentric anomaly, or true anomaly. If the eccentricity of the perturbed body is less than about 0.67 , the series will converge with sufficient rapidity, although for $\mathrm{e}>0.5$, the calculations become very tedious. If e $>0.67$, the convergence is extremely slow or may even fail completely. Consequently, for large eccentricities purely numerical techniques must be used.

Hansen's method of partial anomalies introduces quantities known as partial anomalies. If the perturbations are expressed as trigonometric series involving the partial anomalies, the convergence may be made very rapid for any value of the eccentricity. The method also has the great advantage over the numerical methods of being semi-analytic.

In consequence, Hansen's method of partial anomalies is well suited for the computation of perturbations of satellites, such as the AIMP and IMP satellites, and interplanetary probes with highly eccentric orbits.

This work was published in X-552-67-381.

[^25]by<br>Dr. John B. Peatman ${ }^{\dagger}$

For special purpose digital systems design a delay line can be used as a low cost memory, where the storage of $1000-20,000$ bits are required. This upper limit of 20,000 bits represents the present state of the art and is determined by the dispersion of a signal between input and output of the delay line. This limitation is dictated by the step response of a delay line. In essence, the input is differentiated, delayed by $\mathrm{T}_{\mathrm{d}}$ seconds and differentiated again at the output. Thus a step input should come out as a doublet. Dispersion spreads this doublet out so that $T$ seconds fall between its peaks. The maximum storage capacity is determined by the rate at which step inputs can be put into the delay line and still be distinguished at the output. This limits the storage capacity to $\mathrm{T}_{\mathrm{d}} / \mathrm{T}$ bits.

Manufacturers of delay lines expend great effort to reduce this dispersion so as to maximize storage capacity. In this way 20,000 bits of storage have been obtained.

The purpose of this study is to determine the problems associated with increasing this storage capacity through signal processing. We desire to apply step inputs to the delay line at such a rate that the output waveforms overlap. Then we wish to put the output through a filter such that the effect of each input step is apparent at the output of the filter. Several possibilities exist:

1. Inverse filter
2. Matched filter
3. Waveform compression filter

The inverse filter would recreate a step at the filter output for each step input to the delay line. Consequently if $F(s) e^{-T} d s$ is the transfer function of the delay line, then the transfer function of the inverse filter is $G(s)=1 / F(s)$. There are several drawbacks to the use of this filter. First, it is physically unrealizable since it must approach infinite gain as frequency approaches infinity in order to compensate for the attenuation of high frequencies inherent in $F(s)$. In addition, the inverse filter will include two integrations, which will cause baseline drift in the output in any real implementation of the filter. Finally in

[^26]the process of working toward an unattainable step output in response to a step input, any mismatch between $\mathrm{G}(\mathrm{s})$ and $1 / \mathrm{F}(\mathrm{s})$ will be accentuated in an uncontrolled fashion. That is, we have not designed the filter so as to minimize the sensitivity to such a mismatch.

The term "matched filter" comes from radar signal processing and denotes a filter whose impulse response is

$$
g(t)=K f(D-t)
$$

where $K$ and D are arbitrary constants, while $f()$ is the response of the delay line previously defined. This filter maximizes the height of the signal component output pulse with respect to the RMS noise output due to white noise in the circuit. For this problem this approach does not seem appropriate since noise is not the fundamental problem. The delay line plus output circuitry is relatively noise-free and predictable as compared with a radar system. Consequently it seems that a matched filter is based upon optimizing the wrong criterion.

A waveform compression filter generated $f(2 t)$ in response to $f(t)$, assuming it is set up to compress the step response by a factor of two. Such an approach seems as if it is putting emphasis where it really counts for this problem. However, if it is assumed that $f(t)$ consists of one period of a sinusoid (a usefully simple idealization) then $g(t)$, the impulse response of the filter, consists of a periodic waveform with complex shape. Such a waveform sounds unrealizable without getting into sensitivity problems. However since it is not necessary that the output be exactly equal to $f(2 t)$, and since $f(t)$ itself is not really a finite waveform, the required $g(t)$ may really be a decaying waveform. This seems as if it would really be necessary if the sensitivity to mismatch between $f(t)$ and $g(t)$ is not to be infinite.

# OPTICAL PROPERTIES OF LIGHT BEAMS 

by

Harvey N. Rexroad ${ }^{\dagger}$

## OPTICAL RESONATORS (THE CONFINEMENT OF LIGHT BEAMS)

My NASA colleague, Nelson McAvoy, and I had agreed prior to my coming to Goddard that a good way to begin my tenure as a ASEE-NASA Fellow would be to review the available literature on the subject of optical resonators. Besides affording a means of introducing the author to the word of laser technology, there were two additional motives for the project. The first was education. Several ideas from physics and mathematics at the advanced graduate level are required for an understanding of this topic, which is of fundamental importance for laser research. Consequently, there is a need for an organized collection of all of this material. The second motive was directly related to the work in progress in the laboratory at Goddard. It was hoped that this study would provide an answer to questions having practical significance (the optimum geometry for maximum degree of coherence, the field pattern radiated from a coupling hole, the interpretation of data from a confocal scanning interferometer, etc.).

A series of lectures and discussions, moderated by the author, did serve for the mutual educational benefit of everyone in the group. Notes have been completed and distributed covering the following topics:
I. Introduction
II. Basic Considerations
A. Uncertainty Principal and Confinement
B. Quality Factor
III. Introduction to Methods of Computing Properties of Optical Resonators
A. Beam Iteration; Phase Transformers
B. Self Consistent Field Condition
C. Resonator Mode Condition; Diffraction Loss

[^27]IV. The Self Consistent Field Equations of Optical Resonators
A. The Basic Diffraction Integral:

Fresnel-Huggen's Diffraction Integral for Optical Resonators
B. Derivation of the Basic Diffraction Integral Using the Integral Representation of 15 D .
C. Simplified Basic Diffraction Integral.
D. Identical Mirrors; Single Transit Field Consistent Equations
E. Complete Pass Self Consistent Equations
F. Symmetry Properties and Equivalence Relations for Optical Resonators

Appendix A. Review of Electromagnetic Theory
Appendix B. Gauss' and Green's Theorems: Delta Function Properties of

$$
\nabla^{2}(1 / \mathrm{r}) \text { and } \square^{2}\left(1 / \mathrm{r}^{2}\right)
$$

Appendix C. Determination of Electromagnetic Field From Sources
Appendix D. Kirchoff's Surface Integral Representation

Appendix E. Introduction to Integral Equations; Example of Solution of a Homogeneous Fredholm Equation of the Second Kind.

Appendix F. Spheroidal Wave Functions
The following additional notes will complete the coverage of lecture material and will be added in the near future:
V. Solution of the Confocal Square Mirror Problem
A. The Exact Solution
B. Gauss - Hermite Functions
C. Field Pattern of Various Modes
D. Propagation of Gaussian Waves
VI. The Waveguide Approach
VII. Image and Propagation Laws
A. Image Rules of Fresnel Diffraction
B. Propagation Laws
C. Ray Optics
D. The ABCD Law of Gaussian Imaging
E. Circle Diagrams; Slide Rule Solution of Gaussian Beam Propagation Problems
VIII. Practical Design of Optical Resonators
A. Stability Condition
B. Frequency Condition
C. Mirror Spot Size
D. Filling Factor

## IX. Perturbations

Appendix G. Integration of Diffraction Integral for Beam Propagation
Appendix H. Solution of Vector Wave Equation in Cylindrical Coordinates.
Appendix I. Fresnel Diffraction Image Rule
Appendix J. The Ray Transfer Matrix
In Progress: A computer program is being written to determine the radiation zone field patterns from a non-confocal optical resonator with a hole for output coupling aperture.

## SPREADING OF A LIGHT BEAM (Possible Publication)

A somewhat natural extension of some aspects of the main problem led to the investigation of the spreading of a light beam as it propagates in space, which is analogous to the spreading of a quantum mechanical wave packet in time. The exact problem considered utilized classical electrodynamics to determine the Fraunhofer (far-field) diffraction due to a wave launched from a rectangular aperture at 0 in Figure 1. A gain for the finite surface, $S$, at $F$ in the far field was defined as:
$\mathrm{G}=$ (time average power through S per unit solid angle) divided by (time average power through A over $4 \pi$ ).


Figure 1

I believe I have proved that a field distribution:

$$
E_{x_{n m}}(x, y) \propto S_{o n}(c, x / a) S_{o m}(c, y / b) \exp \left(-\frac{2 \pi}{i \lambda}\left(\frac{x^{2}+y^{2}}{2 d}\right)\right)
$$

with $n=m=0$, over the aperture at 0 produces the maximum possible gain, G.
In this equation $S_{\text {on }}(c, t)$ is a prolate spheroidal angle function, and the solid angle subtended by $S$ to 0 is $4 c^{2} \lambda^{2} /\left(\pi^{2} \mathrm{~A}\right)$. This maximum gain turns out to be:

$$
\mathrm{G}_{\max }=\mathrm{G}_{0}\left[\mathrm{R}_{00}^{(\mathrm{i})}(\mathrm{c}, 1)\right]^{4},
$$

where $R_{\text {on }}{ }^{\text {(i) }}$ is a prolate radial function of the first kind. This result agrees with a previously known one that $\mathrm{G}_{0}=4 \pi \mathrm{~A} / \lambda^{2}$ is the maximum gain when $\mathrm{c}=0$.

There is, of course, a direct application of this maximum gain theorem to space communications and satellite tracking. It is also a fortunate circumstance that the lowest mode of a laser automatically produces this desired field distribution.

Present Stage: A rough draft of a paper has been written. I am checking the literature to see if I have missed anything. Also, since the difference between this and previously known results is not large, there is some problem in assessing the value of this discovery.

# TOWARDS OPTICAL COMMUNICATIONS SYSTEMS 

by<br>Dr. Frederick Rojak ${ }^{\dagger}$

Work during this summer was divided in roughly four categories.
(1) Continuation of a study of three-dimensional optical data processing using holography. The advantages of holography over conventional photography became apparent for the case where the third dimension is traded for the acquisition of a large amount of two-dimensional information.

Attention was given to the problem of recording the parameters of spatially modulated coherent wavefronts.
(2) An explanation of the theory of variable scale optical correlators.
(3) An attempt at contributing to the development of an optical parity generator for a convolutional encoder.
(4) Giving two lectures on coherence theory and working toward a meaningful definition of "coherence length." Familiarized myself somewhat with Glanber's work on this.

[^28]
## REAL TIME CONTROL SYSTEMS

by

Linus Schrage ${ }^{\dagger}$

As part of the development program for an automated, computer controlled, satellite tracking station a study was made in order to outline the functional requirements of the various components of the system. A substantial effort was devoted to the specification of a language for programming a real-time control computer. The proposed language has a standard alogrithmic language such as ALGOL or FORTRAN as a proper subset but has additional statement types particularly useful in real-time control programming. Examples of these statement types are the CONNECT statement which allows one to specify that a given external interrupt will cause a specified sub-routine to be entered when the interrupt is triggered. There are also statements to assign priorities to subroutines and statements to disarm and disable interrupt channels. The execution of subroutine at a future point in time can be scheduled through the use of the CAUSE statement. A measure of multi-programming is possible through use of group sub-routine calls and the DELAY statement. Utility subroutines can be compiled into re-entrant or pure code by the use of the PURE PROCEDURE declaration.

[^29]
[^0]:    $\dagger$ University of Missouri, Columbia, Missouri

[^1]:    *Carnegie-Mellon University, Pittsburgh, Pa.

[^2]:    *U. S. Naval Academy, Annapolis, Md.

[^3]:    *Catholic University of America, Washington, D. C.

[^4]:    *Electrical Engineering Department University of Missouri at Rolla.

[^5]:    *Miami University, Oxford, Ohio

[^6]:    *Cycle time is the time to retrieve an instruction from core memory, bring it to the instruction register in the CPU, and rewrite the instruction in memory. Cycle time on the SIGMA 5 computer is .85 microseconds.

[^7]:    *Rutgers, The State University, New Brunswick, N. J.

[^8]:    * Norwich University, Northfield, Vt.

[^9]:    *The George Washington University, Washington, D. C.
    **Kutgers, The State University, New Brunswick, New Jersey.

[^10]:    $0.15678730 E-07$
    $0.75335835 E-06$
    $0.26337319 E-05$
    $-0.27573499 E-06$
    $-0.23957566 E-06$
    33

[^11]:    EIGENVALUENO.
    EGENVECTUR $-0.77805407 E-01$
    $0.29290821 E-02$
    $-0.48790627 E-02$ $\begin{array}{lr}0.127480809 \mathrm{E}-02 & 0.43895976 \mathrm{E}-01 \\ 0.214148816 \mathrm{E}-01 & -0.17231324 \mathrm{E}-00\end{array}$ $-0.81168407 \mathrm{E}-02$
    $-0.40250877 \mathrm{E}-01$ $0.33242125 \mathrm{E}-02$
    $0.2380521 \mathrm{E}-01$
    $-0.42134225 \mathrm{E}-02$ $-0.42134225 \mathrm{E}-02$ $-0.72402266 \mathrm{E}-01$
     $0.46050031 \mathrm{E}-01 \quad 0.93884882 \mathrm{E}-01$
     $\begin{array}{rrr}-0.29335537 E-00 & -0.15024813 E-00 \\ 0.17430697 E-01 & 0.70877004 \mathrm{E}-01\end{array}$

[^12]:    $-0.40586204 E-01-0.13329772 E-01-0.75060794 E-02$
    
    

[^13]:    
    $\begin{array}{rrrrr}J .44535203 E-J i & -0.12323527 E-00 & -0.55855289 E-01 & 0.93028909 E-01 \\ J .1221343 J E-J J & -0.41125065 E-01 & 0.37820406 E-01 & -0.13065866 E-01 \\ -J .17734934 E-01 & 0.60891882 E-02 & 0.48947428 E-00 & 0.32373012 E-00 \\ 0.3643393 E-0 J & -0.32362305 E-00 & 0.83964576 E 00 & -0.26653815 E-01 \\ -J .82635792 E-01 & -0.15237873 E ~ 01 & 0.19990189 E-00 & 0.97991502 E-01\end{array}$
    $\begin{array}{rrrrr}J .44535203 E-J i & -0.12323527 E-00 & -0.55855289 E-01 & 0.93028909 E-01 \\ J .1221343 J E-J J & -0.41125065 E-01 & 0.37820406 E-01 & -0.13065866 E-01 \\ -J .17734934 E-01 & 0.60891882 E-02 & 0.48947428 E-00 & 0.32373012 E-00 \\ 0.3643393 E-0 J & -0.32362305 E-00 & 0.83964576 E 00 & -0.26653815 E-01 \\ -J .82635792 E-01 & -0.15237873 E ~ 01 & 0.19990189 E-00 & 0.97991502 E-01\end{array}$ EIGENVALUE NO, $20=0.10902555 E 04$
    EIGENVECTOR
     J. $45712175 E-01-0.20010503 E-00-0.21606297 E-01$ - J. 1 S508735E-00 -0.47472778E-01 0.12156307E-00
    
    

[^14]:    0.78629863 E 00
    $0.06362355 \mathrm{E}-02$
    $0.1381873 \mathrm{E}-01$
    $-0.28905485 \mathrm{E}-01$
    
    
    (10)
    $0.34312701 E-01$
    $-0.43076939 E 00$
    
    
    
    

[^15]:    $0.42399381 \mathrm{E} \quad 04$
    
    
    
    
    
     eigenvalue no. $32=$ EIGEVVALURTOR
    EISENVE
    -3. 247216232
    

[^16]:    *United States Naval Academy, Annapolis, Maryland
    **Sacramento State College, Sacramento, California

[^17]:    ${ }^{\dagger}$ Union College, Schenectady, New York

[^18]:    ${ }^{\dagger}$ San Diego State College, San Diego, California

[^19]:    †University of Florida, Gainesville, Florida

[^20]:    ${ }^{\dagger}$ Spring Garden Institute, Philadelphia, Pennsylvania

[^21]:    ${ }^{\dagger}$ South Dakota School of Mines and Technology, Rapid City, S.D.

[^22]:    ${ }^{\dagger}$ Union College, Schenectady, New York

[^23]:    ${ }^{\dagger}$ The George Washington University, Washington, D.C.

[^24]:    ${ }^{\dagger}$ University of Maryland, College Park, Maryland

[^25]:    ${ }^{\dagger}$ Lehigh University, Bethlehem, Pennsylvania

[^26]:    ${ }^{\dagger}$ Georgia Institute of Technology, Atlanta, Georgia

[^27]:    ${ }^{\dagger}$ West Virginia University, Morgantown, West Virginia

[^28]:    ${ }^{\dagger}$ Lowell Technological Institute, Lowell, Massachusetts

[^29]:    ${ }^{\dagger}$ Stanford University, Stanford, California

