NASA SP-141

FIRST COMPILATION OF PAPERS ON TRAJECTORY ANALYSIS AND GUIDANCE THEORY

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NASA SP-141

FIRST COMPILATION OF PAPERS ON TRAJECTORY ANALYSIS AND GUIDANCE THEORY

Prepared by contractors for
the NASA Electronics Research Center
Guidance Laboratory, Cambridge, Massachusetts

Scientific and Technical Information Division
OFFICE OF TECHNOLOGY UTILIZATION 1967
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C.

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – Price $3.00

PRECEDING PAGE BLANK NOT FILMED.

Foreword

Compiled in this volume are seven papers from agencies working with the Guidance Laboratory of NASA/ERC. These papers are of special studies in the disciplines of trajectory analysis, astrodynamics, and celestial mechanics.

They include:

(1) An extension of variational theory to cover problems involving functions that can be represented approximately only, through as closely as desired;

(2) A presentation of an orthonormalization procedure for achieving a least squares approximation of a multivariate function;

(3) A development of an analytic solution for minimum fuel impulsive transfer between low eccentricity orbits;

(4) A development of a method for calculating "geometric" bounds for conditions where, if the total energy of the three-body problem is negative, then one body will recede to infinity from the other two bodies;

(5) A development of a power series for the problem of three bodies where the coefficients in the series are generated by reversive operations;

(6) A development of rigorous error bounds for approximate satellite orbit theories;

(7) A development of a general perturbation theory for the long-term behavior of high eccentricity orbits about Mars.

The first paper, along with extensions to this work, will be useful in trajectory analysis and guidance theory. The second paper should support trajectory analysis and guidance theory in supplying approximations to functions where only numerical values are available. The third paper will contribute to mission design and, in general, to astrodynamics studies. The fourth paper will be useful for trajectory analysis on mission design, along with contributing to studies in celestial mechanics. The last three papers will be useful in celestial mechanics, specifically orbit determination or prediction.

R. J. Hayes, Chief
Guidance Laboratory

W. E. Miner, Chief
Guidance Theory and
Trajectory Analysis Branch

SUMMARY

This volume contains technical papers on NASA-sponsored studies in the areas of trajectory analysis and guidance theory. The studies are being carried on by several universities and industrial companies. These papers cover a period ending October 1, 1966. The technical supervision of the contracts is under the personnel of the Guidance Laboratory.

Contents

Page

INTRODUCTION 1
W. E. Miner, GST, ERC

FUNCTIONS OF RELAXED CONTROLS 7
Dr. J. Warga, Northeastern University

DEVELOPMENT OF MULTIVARIATE FUNCTIONAL MODELS
BY LEAST SQUARES 21
Dr. D. E. Dupree and Dr. R. L. Truax,
Northeast Louisiana State College

A GENERAL SOLUTION FOR MINIMUM IMPULSE TRANSFERS
IN THE NEAR VICINITY OF A CIRCULAR ORBIT 37
Mr. T. N. Edelbaum, Analytical Mechanics
Associates, Inc.

REJECTION TO INFINITY IN THE PROBLEM OF THREE
BODIES WHEN THE TOTAL ENERGY IS NEGATIVE 73
Dr. D. C. Lewis, Control Research Associates

AN ALGORITHM TO OBTAIN SERIES EXPANSIONS IN THE
THREE-BODY PROBLEM 97
Dr. P. Sconzo, IBM

RIGOROUS ERROR BOUNDS ON POSITION AND VELOCITY
IN SATELLITE ORBIT THEORIES 121
Dr. J. V. Breakwell and Dr. J. Vagners,
Stanford University

AN INVESTIGATION OF HIGH ECCENTRICITY ORBITS
ABOUT MARS 175
Dr. J. V. Breakwell and Mr. R. D. Hensley,
Stanford University

Introduction

Compiled in this volume are seven technical papers from agencies working under contract, or grant, to NASA's Electronics Research Center (ERC) in the fields of guidance theory and trajectory analysis. This work was sponsored by the Guidance Laboratory at the NASA Electronics Research Center.

The following table presents the contributing institution, the section of greatest relationship, and the discipline of the paper.

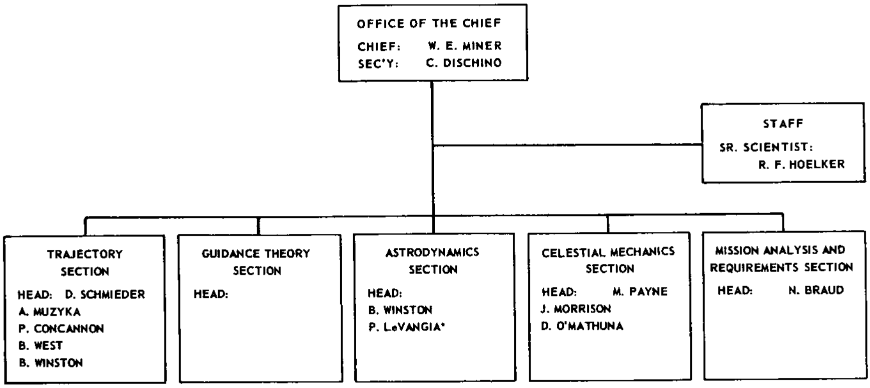
SECTION	INSTITUTION/COMPANY	DISCIPLINE
Trajectory	Northeastern Univ.	Calculus of Variation
Trajectory	Northeast La. State	Functional Models
Astrodynamics	AMA	Impulse Transfers
Celestial Mechanics	CRA	Celestial Mechanics
Celestial Mechanics	IBM	Celestial Mechanics
Celestial Mechanics	Stanford Univ.	Celestial Mechanics
Celestial Mechanics	Stanford Univ.	Celestial Mechanics

The present organization for the laboratory effort is shown in Figure 1.

The following are reviews of the individual papers.

Paper No. 1

The first paper by J. Warga of Northeastern University is a contribution toward the development of variational theory by the methods of modern analysis. Applications are expected to be made on practical problems that cannot be handled satisfactorily by the standard variational theory. The basic technique is to imbed the set of admissible control functions in a larger



*COLLEGE WORK STUDY PROGRAM

FIGURE 1

set that possesses an extremizing solution for the problem, and then to approximate that solution as closely as is desired with one of the original control functions. The present paper presents and proves the required existence and approximation theorems, and extends previous work on this theory done by Prof. Warga. Future work will include the development of the corresponding necessary conditions for an extremum and the consideration of constructive methods by which the solutions may be determined in practice.

Paper No. 2

The second paper is a technical progress report from Northeast La. State by D. E. Dupree and R. T. Truax. This paper first reviews the development of an orthonormalization procedure for achieving a least squares approximation of a multivariate function. The authors then go through the development of a weighting function to augment the least squares approximation. This is done in an attempt to require the resultant error to be less than a specified error tolerance. As yet the work is incomplete in that the proof only succeeds in showing that the weighted error will not exceed that of the unweighted function.

Multivariate functional models are useful whenever functional approximation is needed for numerical data. Examples are cut-off requirements, time-of-ignition, and steering angles which can often be numerically, but not analytically, generated. The objective here was to control the errors of such approximations. Work will continue in this area. The basic orthonormalization procedure has been programmed at ERC and will be used in application. To date no work has been done in implementing the weighting function.

Paper No. 3

The third technical paper, prepared by T. N. Edelbaum of AMA, entitled "A General Solution for Minimum Impulse Transfers in the Near Vicinity of a Circular Orbit," develops an analytic solution for the minimum fuel transfer between low eccentricity orbits. The author shows that in all cases a two-impulse transfer suffices. The results are derived by applying Lawden's primer concept. The range of validity of the results is determined by the applicability of the linearized equations of motion.

The fact that analytic solutions of the two-point boundary value problem were obtained makes this paper a valuable contribution to the state of the art in trajectory optimization and guidance theory. The impulsive solutions are ready approximations for finite time trajectory solutions and may be used as such in guidance modes for on-board as well as ground-based applications.

Paper No. 4

The fourth technical paper, on Rejection to Infinity by D. C. Lewis of Control Research Associates, sharpens a result for the three-body problem due to Birkhoff. Roughly speaking, the result is that if, for given total energy (negative) and total angular momentum, the three bodies are initially sufficiently close together, then one of the three will ultimately recede infinitely far from the other two. The result is, however, a very precise one, and a rigorous proof is given. The condition on the initial configuration is stated in terms of the Lagrange inertial radius R which is a measure of "closeness together" of the three bodies. A method is developed for calculating three quantities R_o, R_1, and t_1 such that if initially $R \geq R_o$, then, at t_1, $R > R_1$. Further, the distance between the two bodies closest together at t_1 will remain bounded (with an estimate given for the bound) thereafter, and the third body will recede to infinity from these two.

This study was done to get a better understanding of the problem of three bodies in celestial mechanics. Establishing the bounds is an advancement which allows us in-house to attempt to develop methods of studying analytically swing-by trajectories to distant places.

Paper No. 5

The fifth technical paper, entitled "An Algorithm to Obtain Series Expansions for the Three-Body Problem," by P. Sconzo, IBM, deals with power series solution, both in the time variable itself and also in terms of a regularizing variable. Outlines are presented for the recursive generation of the coefficients in the two series. A lower bound is given for the radius of convergence of the series in time and the relation of this series to the "f" and "g" series representation of the solution of the two-body problem is briefly discussed. (It should be noted that the method is of wider applicability and, in particular, the relation with the "f" and "g" series might well be exploited for artificial satellite theory.) Next the properties of a general class of regularizing transformations are presented with special reference to the transformation of Levi-Civita. Subsequent reports will deal with convergence of the series in the regularizing variable and application of a machine-generated series based on the recursive formulas mentioned above to a problem of Bohlen for which numerical calculations by Zumkley are available for comparison. Also included are some comments on a solution of the three-body problem recently obtained by Bazayevesky in terms of a power series in time.

While the method used in this report has been previously applied (Steffenson, Fehlberg, Rabe et al), there are some novel features in its development and in the symbolic program for the generation of the series. The theory developed in this report is directly applicable to problems for which the three-body problem forms a good model, for example, to ephemeris calculations for motion about the sun of two planets or one planet and its satellite.

Paper No. 6

The sixth technical paper by J. V. Breakwell and J. Vagners discusses rigorous error bounds for approximate satellite orbit theories. The paper is concerned with the error in prediction from initial conditions over a time interval of order $1/\varepsilon$ for a general perturbation theory developed in powers of ε. The problem is that if one, or more, of the variables possesses a linear dependence on time, the coefficients of time in these variables must be calculated to second order in ε in order to obtain a first-order approximation for time intervals of order $1/\varepsilon$. The development is carried out for earth satellites ($\varepsilon = J_2$) using the Brouwer-von Zeipel technique with Poincaré variables. The only one of these variables with a linear second-order (in ε) time dependence is an angle closely related to the mean anomaly. The authors make use of the energy integral to obtain this second-order coefficient, thus bypassing the necessity of a full second-order formulation and integration of the perturbation equation. The theory includes tesseral harmonics, but omits consideration of resonance effects arising from the tesseral harmonics and also from critical angles of inclination. These will be the object of future investigations. The theory is applied to a circular 2000-mi satellite for 7 days, and the resulting errors are consistent with the theory.

Paper No. 7

The seventh technical paper, by J. V. Breakwell and R. D. Hensley of Stanford University, develops a general perturbation theory for the long-term behavior of high eccentricity orbits about Mars. Mission requirements on a planetary orbiter would probably require a small pericenter distance (for observational purposes) and high eccentricity for saving of fuel. In this study, a pericenter requirement of 4000 to 6000 km was imposed and eccentricity larger than 0.5 was taken. For such an orbit about Mars, the orbital period is small compared to the Martian year which, in turn, is small compared to the rates of change in the orbital parameters due to Mars' oblateness and the sun. These facts make a double averaging procedure, first, over the orbital period and then over the Martian year useful for the study of long period effects. The dominant perturbation is due to Mars' polar oblateness and results in secular rates for the argument of the node. The perturbations caused by the sun result primarily in long period fluctuations in inclination and eccentricity and hence in pericenter distance. In addition to the oblateness critical angle of 63.4 degrees, a number of other critical angles occurs. It is in the neighborhood of these critical angles that the amplitude of the fluctuations in eccentricity are most pronounced. A detailed analysis of these resonance effects is given. In addition, an appendix contains an analysis for small eccentricity (J_2 for Mars). As a first step in the analysis of planetary orbiters, this is a very useful study and extension to include short period effects would be desirable. The applicability of the analysis to orbiters of other planets would require consideration of the relative magnitudes of the orbital period, the planet's period about the sun, and the rates of change of the orbital elements, as well as a study of which perturbations are dominant. Development of theories for planetary orbiters is, of course, essential to any extensive program of planetary exploration.

PRECEDING PAGE BLANK NOT FILMED.

Functions of Relaxed Controls*

by J. Warga
Northeastern University
Boston, Massachusetts

N67-29371

SUMMARY

Problems of the calculus of variations do not admit, in general, ordinary minimizing solutions. General existence theorems in the calculus of variations have been established only with the introduction of generalized, or relaxed, solutions of the given differential equations. These relaxed solutions represent the limits of ordinary solutions with rapidly oscillating derivatives.

In the present paper we consider a class of problems of the mathematical control theory that are not necessarily defined by systems of ordinary differential equations; they may involve solutions of certain partial differential equations, nonadditive set functions, or other functionals. The controls are functions ρ from a metric space T, with a given measure, to a metric space R, and are subject to restrictions of the form $\rho(t) \in R^{\#}(t)$ a.e. in T, where $R^{\#}(t)$ is, for every t in T, a given nonempty subset of T. We consider functionals $x(\rho, b) = (x^1(\rho, b), \ldots, x^n(\rho, b))$ depending on controls ρ and on parameters b restricted to a compact set B_0. The variational problem consists in determining the minimum of $x^1(\rho, b)$, subject to the previously mentioned restrictions on ρ and b and the condition that $x(\rho, b) \in B_1$, where B_1 is a given set in the euclidean n-space.

We establish an existence theorem asserting that, in a large class of such problems, the restricted minimum is achieved by "relaxed" controls. We also establish an approximation theorem stating that each such relaxed control can be constructively approximated by original, or ordinary, controls.

Necessary conditions for minimum will be discussed in a forthcoming report.

* This research was initiated at Avco Corporation, Wilmington, Massachusetts, under N.A.S.A.-ERC contract NAS 12-112, and has been continued at Northeastern University under grant NGR 22-011-020.

1. Introduction. The mathematical control theory deals primarily with functionals defined in terms of a system of ordinary differential equations. It is our purpose here, and in a paper to follow, to extend certain methods and results of the control theory to a more general setting. In particular, we wish to generalize certain results of references [1] and [2]. In this endeavor, and especially in arguments pertaining to problems of existence, we continue to apply certain concepts first introduced by L. C. Young in his study of "generalized curves" [3], [4].

Let $\mathfrak{R}$ be a class of mappings from a set T to a set R, and let $x^i (i=1,.., n)$ be real-valued functionals on $\mathfrak{R}$, that is, functions from $\mathfrak{R}$ to the real line. In many problems, the sets T and R are metric and the vector x of functionals is characterized by a "weak" continuity, in the sense that $x(\rho)$ and $x(\bar{\rho})$ differ little if $\rho \in \mathfrak{R}$, $\bar{\rho} \in \mathfrak{R}$ and $\rho(t)$ and $\bar{\rho}(t)$ are at a small distance from one another for all (or almost all) values of t in T. This type of continuity (with respect to uniform convergence) has been most frequently investigated in connection with differential equations (e.g. in the study of perturbation methods), in the calculus of variations (weak variations), etc.

In a large class of problems, the functionals x^i are also continuous in a different sense. Let us consider, as an example, the problem of determining the temperature $\theta(\bar{t}, \bar{\xi})$ at a time $\bar{t}$ and at a fixed point $\bar{\xi}$ in the interior of a conducting body whose surface is subjected to a heat flux that varies with time and position. If the heat flux is only slightly changed at all times and over the entire surface, the value $\theta(\bar{t}, \bar{\xi})$ will change but little. This is so because $\theta(\bar{t}, \bar{\xi})$ is a continuous functional of the heat flux in the previously described sense. Assume now that the interval of time $[0, \bar{t}]$ is subdivided into 2k equal subintervals, and that at every point z of the surface, the flux $h(t, z)$ during the interval $[\frac{2i}{2k}\bar{t}, \frac{2i+1}{2k}\bar{t}]$ is replaced by $h(t+\frac{\bar{t}}{2k}, z)$ and the flux $h(t, z)$ during the interval $[\frac{2i+1}{2k}\bar{t}, \frac{2i+2}{2k}\bar{t}]$ is replaced by $h(t-\frac{\bar{t}}{2k}, z)$, for $i = 0, 1, \ldots, k-1$. Then for large values of k, we would expect $\theta(\bar{t}, \bar{\xi})$ to be affected little by this "permutation", even if the flux $h(t, z)$ is a rapidly varying, or even a discontinuous, function of t. We would also expect $\theta(\bar{t}, \bar{\xi})$ to be little affected by "permuting" the heat flux between many adjacent small areas of the surface of comparable measures. This second type of continuity is relative to a mode of convergence resembling the weak convergence of measures. It is of fundamental importance in control theory and in our further considerations.

The continuity of functionals in this second sense makes it possible to simulate mathematically the limits of rapidly oscillating functions in $\mathfrak{R}$. We shall refer to elements of $\mathfrak{R}$ (functions from T to R) as "original controls", and we shall imbed $\mathfrak{R}$ in a larger space $\mathcal{S}$ of "relaxed controls." If we assume that R is a compact Hausdorff space, then we can define the class S of regular probability measures on Borel subsets of R. A "relaxed control" σ is a function from T to S. The relaxed control σ will simulate the limit of rapidly oscillating ordinary controls $\rho_1, \rho_2, \ldots$ when for all (or almost all) t and all Borel subsets R_1 of R, the $\sigma(t)$-measure of R_1 represents, in some sense, the limit, as $j \to \infty$, of the relative frequencies of occurrence inside R_1 of the points $\rho_j(\tau)$ in the neighborhood of t. A relaxed control σ with the property that the measure $\sigma(t)$ is, for every t, concentrated at a single point $\rho(t)$, can be identified with the original control ρ.

Relaxed controls, patterned after L. C. Young's definition of "generalized curves", provide a means of completing the space $\mathfrak{R}$ of original controls. Their use paves the way, in complete analogy with the calculus of variations, to existence and approximation theorems which we state in section 2 and prove in section 3. The next task is the derivation of necessary conditions for minimum. We propose to discuss this subject in a paper to follow.

2. Existence and approximation theorems. We shall assume in this section, and in section 3, that T and R are compact metric spaces, that B_0 is a compact Hausdorff space, and B_1 is a closed set in the euclidean n-space E_n. We shall assume, further, that a nonnegative, finite, regular, complete, and nonatomic measure is defined on T. We represent by ρ, and sometimes by $\rho(\ )$, a mapping from T to R, and by $\rho(t)$ the image of a point t under the mapping. A similar distinction is consistently made between a function (mapping) and the image of a particular point under the mapping. A mapping from T to R is "continuous" at t (on a set T_1) if $|\rho(t), \rho(t')| \to 0$ as $t' \to t (t' \to t$ for all $t \in T_1)$. Here $| r_1, r_2 |$ designates the distance of points r_1 and r_2 in R, and similarly $| t', t |$ will designate the distance in T. A mapping from T to R is "measurable" if, for every $\epsilon > 0$, there exists a closed set F_ε in T, whose measure $| F_\varepsilon |$ is at least $|T| - \varepsilon$, and such that the function ρ is continuous on F_ε when restricted to F_ε.

These definitions of continuity and measurability of ρ can be easily seen to be equivalent to the following statement: ρ is continuous at t (on a set T_1), respectively measurable on T_1, if the function ψ, defined by $\psi(t) = \phi(\rho(t))$ on T, is continuous at t(on T_1), respectively measurable on T_1, for every choice of a continuous function ϕ from R to E_1.

Let $R^\#$ be a mapping from T to the class of nonempty subsets of R, and let $\mathfrak{R}$ be the space of measurable functions ρ from T to R. We are given a function $x: \mathfrak{R} \times B_0 \to E_n$. We wish to investigate the original problem of determining the minimum of $x^1(\rho, b)$, subject to the restrictions that

$\rho(t) \epsilon R^{\#}(t)$ a. e. in T and $x(\rho, b) \epsilon B_1$. Alternately, we wish to consider "approximate minimizing solutions" to this problem. An "approximate solution" $\propto$ is a sequence $\{\rho_j, b_j\}_{j=1}^{\infty}$ such that the ρ_j are measurable functions from T to R, $b_j \epsilon B_0$, $\rho_j(t) \epsilon R^{\#}(t)$ a. e. in T, $x(\rho_j, b_j)$ converges, as $j \to \infty$, to a point $x_\propto$, and $x_\propto \epsilon B_1$. An "approximate minimizing solution" is an approximate solution that minimizes $x^1_\propto$.

Let now S be the class of regular probability measures defined on the Borel subsets of R. We shall refer to a function ρ from T to R as an "original control", and to a function σ from T to S as a "relaxed control". A relaxed control σ is "continuous", respectively "measurable", on a set T_1 if $\int_R \phi(r)\sigma(dr;t)$ is a continuous, respectively measurable, function on T_1 for every choice of a continuous function $\phi : R \to E_1$. Here $\sigma(R_1;t)$ represents the $\sigma(t)$-measure of a Borel set R_1. We can easily verify that if σ is a measurable control and R_1 is a Borel subset of R, then $\sigma(R_1;t)$ is measurable. We shall denote by $\mathcal{S}$ the set of measurable relaxed controls.

If a relaxed control σ_ρ has the property that $\sigma_\rho(t)$ is a measure consisting of a single mass point $\rho(t)$ a. e. in T, then we refer to it, somewhat loosely but without any fear of confusion, as the original control ρ. In this sense we consider $\mathfrak{R}$ to be a subset of $\mathcal{S}$. We shall also treat original, respectively relaxed, controls as identical if they differ only on a set of measure 0 in T.

Definition 2.1. The Young representation. We shall say that a function $y: \mathcal{S} \times B_0 \to E_n$ is a "Young representation of x" if y coincides with x on $\mathfrak{R} \times B_0$; that is, if $y(\sigma_\rho, b) = x(\rho, b)$ for every b in B_0 and every relaxed control σ_ρ such that $\sigma_\rho(t)$ is concentrated at the single point $\rho(t)$ a. e. in T.

(We observe that the relaxed control σ_ρ is measurable if, and only if, the original control ρ is measurable).

Let now $L_1(T)$ be the Banach space of real-valued integrable functions on T, and let C(R) be the Banach space of real-valued continuous functions on R, both with their conventional norms. ($|\phi|_{C(R)} = \max_{r \in R} |\phi(r)|$, $|f|_{L_1(T)} = \int_T |f|$). We shall define functionals $k(\phi, f, \sigma)$ on $C(R) \times L_1(T) \times \mathcal{S}$ by

$$k(\phi, f, \sigma) = \int_T f(t) \int_R \phi(r) \sigma(dr;t).$$

If σ_ρ is an original control ρ, we shall write

$$k(\phi, f, \sigma_\rho) = k(\phi, f, \rho) = \int_T f(t)\, \phi(\rho(t)).$$

Definition 2.2. The Young topology on $\mathcal{S}$. We shall say that a sequence $\sigma_1, \sigma_2, \ldots$ in $\mathcal{S}$ is convergent if the sequence of real numbers $\{k(\phi, f, \sigma_j)\}_{j=1}^{\infty}$ is convergent for every choice of (ϕ, f) in $C(R) \times L_1(T)$. We shall say that σ in $\mathcal{S}$ is a limit of the sequence $\sigma_1, \sigma_2, \ldots$ if

$$k(\phi, f, \sigma) = \lim_{j \to \infty} k(\phi, f, \sigma_j) \text{ on } C(R) \times L_1(T).$$

We now consider a mapping $R^{\#}$ satisfying

Assumption 2.3. For every $\varepsilon > 0$ there exists a closed subset T_ε of T, of measure at least $|T| - \varepsilon$, with the property that

(2.3.1) for every $\bar{t} \in T_\varepsilon$ and every $r \in \bar{R}^{\#}(\bar{t})$ (the closure of $R^{\#}(t)$) there exists a measurable original control ρ, continuous at $\bar{t}$ when restricted to T_ε, and such that $|\rho(\bar{t}), r| < \varepsilon$ and $\rho(t) \in R^{\#}(t)$ on T;

(2.3.2) the mapping $R^{\#}$, when restricted to T_ε, is continuous with respect to inclusion, i.e. for every $\bar{t}$ in T_ε and every $h > 0$, there exists $\delta = \delta(h, \bar{t})$ such that $R^{\#}(t) \subset U(R^{\#}(\bar{t}), h)$ and $R^{\#}(\bar{t}) \subset U(R^{\#}(t), h)$

if $t \in T_\varepsilon$ and $|t, \bar{t}| < \delta$. Here

$$U(R_1, h) = \{r \in R \mid |r, r_1| < h \text{ for some } r_1 \text{ in } R_1\}.$$

We can now state our basic approximation and existence theorems. The symbol $\overline{R}^{\#}(t)$ will denote the closure of $R^{\#}(t)$.

Theorem 2.4. Let the mapping $R^{\#}$ satisfy condition (2.3.1). Then every measurable relaxed control σ can be approximated (in the Young topology) by measurable original controls $\rho_1, \rho_2, \ldots$. If $\sigma(\overline{R}^{\#}(t);t) = 1$ a.e. in T then the controls $\rho_1, \rho_2, \ldots$ can be chosen so that $\rho_j(t) \in R^{\#}(t)$ a.e. in T ($j = 1, 2, \ldots$).

Theorem 2.5. Let $\mathcal{S}^{\#} = \{\sigma \in \mathcal{S} \mid \sigma(\overline{R}^{\#}(t);t) = 1 \text{ a.e. in } T\}$, and assume that the mapping $R^{\#}$ satisfies Assumption 2.3. Then the set $\mathcal{S}^{\#}$ is sequentially compact.

Let y be a Young representation of x, and assume that y is continuous on $\mathcal{S}^{\#} \times B_0$ (with respect to the product topology on $\mathcal{S}^{\#} \times B_0$). Let $\mathfrak{R}^{\#} = \{\rho \in \mathfrak{R} \mid \rho(t) \in R^{\#}(t) \text{ a.e. in } T\}$, $X = \{x(\rho, b) \mid \rho \in \mathfrak{R}^{\#}, b \in B_0\}$ and $Y = \{y(\sigma, b) \mid \sigma \in \mathcal{S}^{\#}, b \in B_0\}$. Then Y is the closure of X.

As a corollary of Theorems 2.4 and 2.5, we derive

Theorem 2.6. Let $\mathfrak{R}^{\#}$ and X be defined as in Theorem 2.5, and let the assumptions be the same as in Theorem 2.5. Then either $Y \cap B_1$ is empty, or there exist $\bar{\sigma} \in \mathcal{S}^{\#}$ and $\bar{b} \in B_0$ that yield the minimum of $y^1(\sigma, b)$ subject to the condition $y(\sigma, b) \in B_1$. If the construction described in 3.3 is used to approximate $\bar{\sigma}$ with a sequence $\bar{\rho}_1, \bar{\rho}_2, \ldots$ in $\mathfrak{R}^{\#}$, then the sequence $\{\bar{\rho}_j, \bar{b}\}_{j=1}^{\infty}$ is a minimizing approximate solution of the original problem.

3. Proofs of the approximation and existence theorems.

Definition 3.1. A dense sequence of partitions of T [5, pp. 171-174]. We shall say that P_T is a dense sequence of partitions of T if $P_T = \{P_T^1, P_T^2, \ldots\}$; $P_T^i = \{T_1^i, T_2^i, \ldots, T_{j_i}^i\}$ $(i = 1, 2, \ldots)$; the sets T_j^i $(j = 1, \ldots, j_i)$ are, for each $i = 1, 2, \ldots,$ measurable and disjoint and $\bigcup_{j=1}^{j_i} T_j^i = T$; every element of P_T^{i+1} is contained in some element of P_T^i, for $i = 1, 2, \ldots$; and to every measurable subset E of T and every $\varepsilon > 0$ there correspond a positive integer $i(\varepsilon)$ and a subset $J(E, \varepsilon)$ of $\{1, 2, \ldots, j_{i(\varepsilon)}\}$ such that $|E - E_0| + |E_0 - E| < \varepsilon$, where $E_0 = \bigcup_{j \in J(E, \varepsilon)} T_j^{i(\varepsilon)}$.

It is well known [5, Th. C, p. 173] that there exists a dense sequence of partitions of T as a consequence of T being metric and compact, and the measure on T having the properties listed at the beginning of section 2.

We shall require a lemma.

Lemma 3.2. Let $\varepsilon > 0$, T_ε have the properties described in Assumption (2.3.1), F be a measurable subset of T_ε, $\{R_1, R_2, \ldots, R_m\}$ be a partition of R into disjoint nonempty Borel subsets, and $\sigma \in \mathcal{S}$, and assume that the support of $\sigma(t)$ is contained in $\bar{R}^{\#}(t)$ (the closure of $R^{\#}(t)$) for all t in F. Let $\alpha^k = \int_F \sigma(R_k; t)$ $(k = 1, \ldots, m)$. Then there exist a partition of F into disjoint measurable sets $F_1, F_2, \ldots, F_m$ and a measurable original control ρ

such that, for $k = 1, 2, \ldots, m$, $|F_k| = \alpha^k$, $\rho(t) \in R^{\#}(t)$ on T, and $\rho(t)$ is within a distance 2ε of R_k a.e. in F_k.

Proof. Let k represent integers from 1 to m, and let $G_k = \{t \in F \mid \sigma(R_k;t) \neq 0\}$. For every nonempty subset A of $\{1, 2, \ldots, m\}$, let $G_A = \bigcap_{k \in A} G_k$. Since $\sum_{k=1}^{m} \sigma(R_k;t) = \sigma(R;t) = 1$ in T, it follows that $F = \bigcup_A G_A$ (the union over all nonempty subsets A of $\{1, 2, \ldots, m\}$). For every set A, we can partition G_A into disjoint measurable subsets $G_A^k (k \in A)$ of measure $\int_{G_A} \sigma(R_k;t)$. If $k \notin A$ we define G_A^k to be the empty set. We now let $F_k = \bigcup_A G_A^k$, and verify that $|F_k| = \alpha^k$ $(k=1, \ldots, m)$.

Let now k be fixed. Since, by construction, $\sigma(R_k;t) \neq 0$ for $t \in F_k$, for every $\bar{t}$ in F_k there exists a point $r_{\bar{t}}^k$ in $\bar{R}^{\#}(t) \cap R_k$ and, by Assumption (2.3.1), there exists a measurable original control $\rho_{\bar{t}}^k$, continuous at $\bar{t}$ when restricted to F, and such that $|\rho_{\bar{t}}^k(t), r_{\bar{t}}^k| \leq \varepsilon$ and $\rho_{\bar{t}}^k(t) \in R^{\#}(t)$ on T. Because $\rho_{\bar{t}}^k$ is continuous at $\bar{t}$ when restricted to F, there exists a neighborhood (relative to F_k) $N_k(\bar{t})$ of $\bar{t}$ such that $|\rho_{\bar{t}}^k(t), r_{\bar{t}}^k| < 2\varepsilon$ in $N_k(\bar{t})$; hence $\rho_{\bar{t}}^k(t)$ is within 2ε of R_k for $t \in N_k(\bar{t})$. Since F_k is covered by open neighborhoods (relative to F_k) $N_k(\bar{t})$, it must be covered a.e. by a denumerable subfamily, say $N_k(\bar{t}_1), N_k(\bar{t}_2), \ldots$. We now let $\rho(t) = \rho_{\bar{t}_j}^k(t)$ for $t \in N_k(\bar{t}_j)$ $(k=1, \ldots, m; j=1, 2, \ldots)$ and $\rho(t) = \rho_{\bar{t}_1}^1(t)$ everywhere else on F. Since $\rho_{\bar{t}_j}^k(t) \in R^{\#}(t)$ for all k and j, it follows that $\rho(t) \in R^{\#}(t)$ on T. We also observe that ρ is measurable and $\rho(t)$ is within a distance 2ε from R_k a.e. in F_k $(k=1, 2, \ldots, m)$.

3.3. Proof of Theorem 2.4. Let the sets T_j^i $(j=1, \ldots, j_i; i=1, 2, \ldots)$ define a dense sequence of partitions of T as in Definition 3.1. Since R is metric and compact, for every positive integer i we can partition R into disjoint Borel subsets R_k^i $(k=1,\ldots,k_i; i=1,2,\ldots)$ of diameters not exceeding $\frac{1}{i}$. In every one of these sets R_k^i we may arbitrarily select a point r_k^i. Let $T_{1/i}$ be defined as in Assumption 2.3, and let $T_j^{*i} = T_j^i \cap T_{1/i}$ $(j=1,2,\ldots,j_i; i=1,2,\ldots)$. For every fixed positive integer i and for every $j=1,2,\ldots,j_i$, we may define sets $T_{j,k}^i (k=1,\ldots,k_i)$ and original controls ρ_j^i that have the properties described in the statement of Lemma 3.2, with $1/i$, T_j^{*i}, $T_{j,k}^i$, R_k^i, and ρ_j^i replacing ε, F, F_k, R_k, and ρ, respectively. Let now a measurable original control ρ_i be defined for $i=1,2,\ldots$ by the relations

$$\rho_i(t) = \rho_j^i(t) \text{ on } T_j^{*i} \quad (j=1,2,\ldots,j_i),$$

$$\rho_i(t) = \rho_1^1(t) \text{ on } T-T_{1/i}.$$

We observe that $\rho_i(t) \in R^{\#}(t)$ on T, $\rho_i(t)$ is within a distance $3/i$ of r_k^i a.e. in $T_{j,k}^i$, and $\left| T_{j,k}^i \right| = \int_{T_j^{*i}} \sigma(R_k^i;t)$ $(k=1,2,\ldots,k_i)$.

Let now $\varepsilon > 0$, $\phi \in C(R)$, and let E be a measurable subset of T. The symbol $|\phi|$ will denote the $C(R)$-norm of ϕ, i.e., $\underset{r \in R}{\operatorname{Max}} |\phi(r)|$. We may choose an integer i_0 sufficiently large so that, for every $i \geq i_0$, there exist a subset J_i of $\{1,2,\ldots,j_i\}$ and a measurable set E_i in T such that $E_i = \bigcup_{j \in J_i} T_j^{*i}$, $|E-E_i| + |E_i - E| < \frac{1}{4}\varepsilon / |\phi|$ and $|\phi(r) - \phi(r')| < \frac{1}{4}\varepsilon / |T|$ if $|r, r'| \leq 3/i$. Finally, let $0(\alpha)$ represent here a quantity not exceeding α in absolute value. Then, for all $i \geq i_0$,

$$\int_E \int_R \phi(r)\sigma(dr;t) = \int_E \sum_{k=1}^{k_i} \phi(r_k^i)\,\sigma(R_k^i;t) + 0\left(\tfrac{1}{4}\varepsilon\right)$$

$$= \sum_{j \in J_i} \int_{T_j^{*i}} \sum_{k=1}^{k_i} \phi(r_k^i)\sigma(R_k^i;t) + 0\left(\tfrac{1}{2}\varepsilon\right)$$

$$= \sum_{j\epsilon J_i} \sum_{k=1}^{k_i} \phi(r_k^i) \; |T^i_{j,k}| + 0(\tfrac{1}{2}\varepsilon)$$

$$= \sum_{j\epsilon J_i} \sum_{k=1}^{k_i} \int_{T^i_{j,k}} \phi(\rho_i(t)) + 0(\tfrac{3}{4}\varepsilon)$$

$$= \sum_{j\epsilon J_i} \int_{T^{*i}_j} \phi(\rho_i(t)) + 0(\tfrac{3}{4}\varepsilon)$$

$$= \int_E \phi(\rho_i(t)) + 0(\varepsilon).$$

Thus $\lim_{i\to\infty} \int_T f_E(t)\phi(\rho_i(t)) = \lim_{i\to\infty} k(\phi, f_E, \rho_i) = k(\phi, f_E, \sigma)$ for every $\phi\epsilon C(R)$ and every measurable characteristic function f_E. It follows that $\lim_{i\to\infty} k(\phi, f, \rho_i) = k(\phi, f, \sigma)$ for every $(\phi, f) \epsilon C(R)\times L_1(T)$. This completes the proof of the theorem.

The proof of Theorem 2.5 is largely a generalization of a construction of L. C. Young [3].

3.4 Proof of Theorem 2.5. We shall first prove that $\mathcal{S}$ is sequentially compact in the Young topology. Let $\sigma_1, \sigma_2, \ldots$ be a sequence in $\mathcal{S}$, and let $k_j(\phi, f) = k(\phi, f, \sigma_j) = \int_T f(t) \int_R \phi(r)\sigma_j(dr;t)$ $(\phi\epsilon C(R), f\epsilon L_1(T), j = 1, 2, \ldots)$. The bilinear functionals k_j clearly satisfy the relation

(3.4.1) $\quad |k_j(\phi, f)| < |f| \; |\phi| \quad (\phi\epsilon C(R), f\epsilon L_1(T), j=1, 2, \ldots)$, where $|f| = |f|_{L_1(T)} = \int_T |f(t)|$ and $|\phi| = |\phi|_{C(R)} = \operatorname{Max}_{r\epsilon R} |\phi(r)|$.

Let now $C'(R)$ be a dense denumerable subset of the separable space $C(R)$. For all $\phi\epsilon C'(R)$ such that $|\phi| \leq 1$, the sequence of linear functionals $\{k_j(\phi, \;)\}_{j=1}^{\infty}$ on $L_1(T)$ has norms bounded by 1, as a consequence of relation (3.4.1). It follows that there exists a subsequence of the functionals $k_j(\phi, \;)$, which we shall continue to designate by $\{k_j(\phi,\ )\}_{j=1}^{\infty}$, that converges to a bounded linear functional $k(\phi,\ )$ for all $\phi \in C'(R)$. Relation (3.4.1) clearly continues to apply to $k(\phi, f)$ for all $\phi \in C'(R)$ and $f \in L_1(T)$. We can easily extend, by continuity, the definition of $k(\phi,\ )$, as a bounded linear functional on $L_1(T)$, for all $\phi \in C(R)$ and we verify that k is bilinear and satisfies the inequality (3.4.1).

For every fixed ϕ in $C(R)$, $k(\phi,\ )$ is a bounded linear functional on $L_1(T)$ and, as such, can be represented by

$$\int_T \ell(\phi, t) f(t),$$

where $\ell(\phi, t)$ is a bounded measurable function on T. Since, for each ϕ, $\ell(\phi, t)$ can be arbitrarily changed on a subset of T of measure 0, we can determine a subset T' of T such that $|T'| = |T|$ and $\ell(\ , t)$ is a bounded linear functional on $C'(R)$ for all $t \in T'$. We then verify that $\ell(\ , t)$ can be extended, by continuity, for each $t \in T'$, to a bounded linear functional on $C(R)$, and that

$$k(\phi, f) = \int_T \ell(\phi, t) f(t) \quad (\phi \in C(R),\ f \in L_1(t)).$$

Furthermore, relation (3.4.1), applied to $k(\phi, f)$, implies that there exists a subset T^* of T', of measure $|T|$, such that

$$|\ell(\phi, t)| \leq |\phi| \text{ on } C(R) \times T^*.$$

We can prove this last relation first for all $\phi \in C'(R)$, and then, by continuity, for all $\phi \in C(R)$.

We can, therefore, conclude that there exists a signed regular measure $\sigma(t)$ for all $t \in T^*$ such that

$$(3.4.2) \quad \ell(\phi, t) = \int_R \phi(r)\sigma(dr;t) \text{ on } C(R) \times T^*.$$

Thus

$$\lim_{j\to\infty} k(\phi, f, \sigma_j) = \lim_{j\to\infty} k_j(\phi, f) = k(\phi, f) = \int_T f(t) \int_R \phi(r)\sigma(dr;t)$$

$$= k(\phi, f, \sigma) \quad (f \in L_1(T), \phi \in C(R)).$$

Since $k_j(\phi, f) \geq 0$ for all j if $\phi(r) \geq 0$ on R and $f(t) \geq 0$ on T, it follows that $k(\phi, f)$ has this property and, therefore, $\ell(\phi, t) \geq 0$ a.e. in T^*, say in $T^*(\phi)$, if $\phi(r) \geq 0$ on R. Furthermore, $k_j(\phi_1, f) = \int_T f(t)$ if $\phi_1(r) = 1$ on R, hence $k(\phi_1, f) = \int_T f(t)$, and it follows that $\ell(\phi_1, t) = 1$ a.e. in T^*. Since $C(R)$ is separable, we can prove, as in previous arguments, that there exists a subset of $T^{\#}$ of T of measure $|T|$ such that $\ell(\phi, t) \geq 0$ on $T^{\#}$ for every non-negative ϕ in $C(R)$; and $\ell(\phi_1, t) = 1$ on $T^{\#}$ if $\phi_1(r) = 1$ on R. It follows then from (3.4.2) that $\sigma(t)$ is a regular probability measure for every t in $T^{\#}$. The corresponding mapping σ is measurable on T since $\ell(\phi, t)$ is measurable for every $\phi \in C(R)$. This shows that $\sigma \in \mathcal{S}$ and completes the proof that $\mathcal{S}$ is sequentially compact in the Young topology.

We shall next show that if a sequence $\sigma_1, \sigma_2, \ldots$ converges to σ in the Young topology, and if $\sigma_j(\overline{R}^{\#}(t);t) = 1$ a.e. in T $(j=1, 2, \ldots)$ then $\sigma(\overline{R}^{\#}(t);t) = 1$ a.e. in T. Indeed, let $\eta > 0$ and let T_η of measure at least $|T| - \eta$ be a closed subset of T such that $R^{\#}$ is continuous when restricted to T_η. Let $\varepsilon > 0$, $\overline{t} \in T_\eta$, $S_\eta(\overline{t}, \delta) = \{t \in T_\eta \mid |t, \overline{t}| < \delta\}$, $U(R_1, h)$ be the open h-neighborhood of a set $R_1 \subset R$, $U_{1/2} = U(\overline{R}^{\#}(\overline{t}), \frac{\varepsilon}{2})$, $\overline{U}_{1/2}$ the closure of $U_{1/2}$, $U_1 = U(\overline{R}^{\#}(\overline{t}), \varepsilon)$, and let $\delta = \delta(\varepsilon)$ be such that $\overline{R}^{\#}(t) \subset U_{1/2}$ and $\overline{R}^{\#}(\overline{t}) \subset U(R^{\#}(t), \frac{\varepsilon}{2})$ for all t in $S_\eta(\overline{t}, \delta)$. Let $\overline{\Phi}$ in $C(R)$ be such that $\overline{\Phi}(r) = 0$ on $\overline{U}_{1/2}$, $0 \leq \overline{\Phi}(r) \leq 1$ on R, and $\overline{\Phi}(r) = 1$ on $R - U_1$. Then

$$0 = \lim_{j\to\infty} \int_{S_\eta(\bar{t},\delta)} \int_R \bar{\Phi}(r)\sigma_j(dr;t) = \int_{S_\eta(\bar{t},\delta)} \int_R \bar{\Phi}(r)\sigma(dr;t)$$
$$\geq \int_{S_\eta(\bar{t},\delta)} \sigma(R - U_1;t);$$

hence $\sigma(R - U_1;t) = 0$ a. e. in $S_\eta(\bar{t}, \delta(\varepsilon))$ or $\sigma(U_1;t) = 1$ a. e. in $S_\eta(\bar{t}, \delta(\varepsilon))$. We observe that, for all t in $S_\eta(\bar{t}, \delta(\varepsilon))$, $U_1 \subset U(\bar{R}^{\#}(t), \frac{3\varepsilon}{2})$. It follows that T_η can be covered by open (relative to T_η) neighborhoods in each of which $\sigma(t)$ is a. e. supported by $U(\bar{R}^{\#}(t), \frac{3\varepsilon}{2})$. Since ε is arbitrary, $\bar{R}^{\#}(t)$ is compact for all t, and the measure $\sigma(t)$ is regular, we conclude that $\sigma(\bar{R}^{\#}(t);t) = 1$ a. e. in T_η. Since η is arbitrary, it follows that $\sigma(\bar{R}^{\#}(t);t) = 1$ a. e. in T. Thus $\mathcal{S}^{\#}$ is sequentially compact.

It follows from the above conclusion and from the continuity of y on the sequentially compact space $\mathcal{S}^{\#} \times B_0$ that the set Y is closed. By Theorem 2. 4, Y is contained in the closure of X. Since X is obviously a subset of Y, we conclude that T is the closure of X. This completes the proof of the theorem.

References

1. J. Warga, Relaxed variational problems, J. Math. Anal. Appl. 4, (1962), pp. 111-128.

2. —, Necessary conditions for minimum in relaxed variational problems, J. Math. Anal. Appl. 4, (1962), pp. 129-145.

3. L. C. Young, Generalized curves and the existence of an attained absolute minimum in the calculus of variations, Compt. rend, soc. sci. et lettres Varsovie, CL III. 30 (1937), pp. 212-234.

4. —, Necessary conditions in the calculus of variations, Acta Math. 69 (1938), pp. 239-258.

5. P. R. Halmos, Measure Theory, New York; D. Van Nostrand, 1950.

Development of Multivariate Functional Models by Least Squares*

by D. E. Dupree and R. L. Truax
Northeast Louisiana State College
Monroe, Louisiana

SUMMARY

A technique for deriving an approximating function yielding an error, in the sense of least squares, less than a specified error tolerance is developed.

A COMPUTATIONAL TECHNIQUE FOR DERIVING THE LEAST SQUARES APPROXIMATING FUNCTION

Given: $n + 1$ tabular points $\{\beta_0, X(\beta_0)\}$, $\{\beta_1, X(\beta_1)\}$, ... , $\{\beta_n, X(\beta_n)\}$ for the function $X = X(\beta)$, where $\beta = (x_1, x_2, \ldots, x_t)$.

Problem: Choose $N + 1$ independent functions $\varphi_0(\beta)$, $\varphi_1(\beta), \ldots, \varphi_N(\beta)$ and determine the polynomial $\sum_{j=0}^{N} A_j\varphi_j(\beta)$ satisfying the property that

$$F(A_0, A_1, \ldots, A_N) = \sum_{i=0}^{n} \{X(\beta_i) - \sum_{j=0}^{N} A_j\varphi_j(\beta_i)\}^2$$

is minimum.

A necessary condition for F to be minimum is for $\frac{\partial F}{\partial A_0} = \frac{\partial F}{\partial A_1} = \ldots = \frac{\partial F}{\partial A_N} = 0$. This yields the following system of

*Performed under NASA Grant NGR-19-006-001 normal equations:

$$A_0(\overline{\varphi}_0,\overline{\varphi}_0) + A_1(\overline{\varphi}_1,\overline{\varphi}_0) + \cdots + A_N(\overline{\varphi}_N,\overline{\varphi}_0) = (\overline{X},\overline{\varphi}_0)$$

$$A_0(\overline{\varphi}_0,\overline{\varphi}_1) + A_1(\overline{\varphi}_1,\overline{\varphi}_1) + \cdots + A_N(\overline{\varphi}_N,\overline{\varphi}_1) = (\overline{X},\overline{\varphi}_1)$$

$$\vdots$$

$$A_0(\overline{\varphi}_0,\overline{\varphi}_N) + A_1(\overline{\varphi}_1,\overline{\varphi}_N) + \cdots + A_N(\overline{\varphi}_N,\overline{\varphi}_N) = (\overline{X},\overline{\varphi}_N)$$

where $\overline{X} = \{X(\beta_0), X(\beta_1), \ldots, X(\beta_N)\}$,

$\overline{\varphi}_j = \{\varphi_j(\beta_0), \varphi_j(\beta_1), \ldots, \varphi_j(\beta_n)\}$, $j = 0,1,\ldots,N$,
and $(\overline{\varphi}_k,\overline{\varphi}_j)$ denotes the inner product $\sum\limits_{i=0}^{n} \varphi_k(\beta_i)\varphi_j(\beta_i)$.

Using the vectors $\overline{\varphi}_0,\overline{\varphi}_1,\ldots,\overline{\varphi}_N$, define a set of vectors $\overline{e}_0,\overline{e}_1,\ldots,\overline{e}_N$ as follows:

$$\overline{e}_j = \frac{\{\overline{\varphi}_j - (\overline{\varphi}_j,\overline{e}_0)\overline{e}_0 - (\overline{\varphi}_j,\overline{e}_1)\overline{e}_1 - \cdots - (\overline{\varphi}_j,\overline{e}_{j-1})\overline{e}_{j-1}\}}{\{(\overline{\varphi}_j,\overline{\varphi}_j) - \sum\limits_{k=0}^{j-1} (\overline{\varphi}_j,\overline{e}_k)^2\}^{1/2}}$$

This is the orthonormal collection yielded by the Gram-Schmidt orthogonalization process; that is $(\overline{e}_k,\overline{e}_j) = 0$, $k \neq j$, and $(\overline{e}_j,\overline{e}_j) = 1$, $j = 0,1,\ldots,N$.

In addition, define the triangular array of coefficients

$$
\begin{array}{llllll}
A_0(-1) & & & & & \\
A_1(-1) & A_1(0) & & & & \\
A_2(-1) & A_2(0) & A_2(1) & & & \\
A_3(-1) & A_3(0) & A_3(1) & A_3(2) & & \\
\multicolumn{6}{c}{----------------------------} \\
A_N(-1) & A_N(0) & A_N(1) & A_N(2) & ----- & A_N(N-1)
\end{array}
$$

where

$$A_\gamma(-1) = \frac{1}{\{(\bar{\varphi}_\gamma, \bar{\varphi}_\gamma) - \sum_{j=0}^{\gamma-1} (\bar{\varphi}_\gamma, \bar{e}_j)^2\}^{1/2}}$$

$$A_\gamma(k) = \frac{(\bar{\varphi}_\gamma, \bar{e}_k)}{\{(\bar{\varphi}_\gamma, \bar{\varphi}_\gamma) - \sum_{j=0}^{\gamma-1} (\bar{\varphi}_\gamma, \bar{e}_j)^2\}^{1/2}}, \quad k = 0,1,\ldots,\gamma-1.$$

Then $\bar{e}_j$ and $A_j(k)$, $0 \leq j \leq N$ and $k = 0,1,2,\ldots,j-1$, can be written as follows:

$$A_j(k) = A_j(-1)A_k(-1)(\bar{\varphi}_j, \bar{\varphi}_k) - \sum_{i=0}^{k-1} A_j(i)A_k(i)$$

$$\bar{e}_j = A_j(-1)\bar{\varphi}_j - \sum_{i=0}^{j-1} A_j(i)\bar{e}_i \ .$$

Then the coefficients of the triangular array and the $\bar{e}_j$, $0 \leq j \leq N$, can be written recursively as follows:

$A_0(-1) = \dfrac{1}{(\bar{\varphi}_0, \bar{\varphi}_0)^{1/2}}$		
$\bar{e}_0 = A_0(-1)\bar{\varphi}_0$		
$A_1(-1) = \dfrac{1}{\{(\bar{\varphi}_1, \bar{\varphi}_1) - (\bar{\varphi}_1, \bar{e}_0)^2\}^{1/2}}$	$A_1(0) = A_1(-1)A_0(-1)(\bar{\varphi}_1, \bar{\varphi}_0)$	
$\bar{e}_1 = A_1(-1)\bar{\varphi}_1 - A_1(0)\bar{e}_0$		
$A_2(-1) = \dfrac{1}{\{(\bar{\varphi}_2, \bar{\varphi}_2) - \sum\limits_{j=0}^{1} (\bar{\varphi}_2, \bar{e}_j)^2\}^{1/2}}$	$A_2(0) = A_2(-1)A_0(-1)(\bar{\varphi}_2, \bar{\varphi}_0)$	$A_2(1) = A_2(-1)A_1(-1)(\bar{\varphi}_2, \bar{\varphi}_1)$ $- A_2(0)A_1(0)$
$\bar{e}_2 = A_2(-1)\bar{\varphi}_2 - \sum\limits_{i=0}^{1} A_2(i)\bar{e}_i = A_2(-1)\bar{\varphi}_2 - A_2(0)\bar{e}_0 - A_2(1)\bar{e}_1$		

.

.

.

$\cdot$
$\cdot$
$\cdot$

$A_\gamma(-1) = \dfrac{1}{\{(\bar{\varphi}_\gamma,\bar{\varphi}_\gamma) - \sum_{j=0}^{\gamma-1} (\bar{\varphi}_\gamma,\bar{e}_j)^2\}^{1/2}}$	$A_\gamma(0) = A_\gamma(-1)A_0(-1)(\bar{\varphi}_\gamma,\bar{\varphi}_0)$	$\cdots$	$A_\gamma(\gamma-1) = A_\gamma(-1)A_{\gamma-1}(-1)(\bar{\varphi}_\gamma,\bar{\varphi}_{\gamma-1}) - \sum_{i=0}^{\gamma-2} A_{\gamma-1}(i)$

$$\bar{e}_\gamma = A_\gamma(-1)\bar{\varphi}_\gamma - \sum_{i=0}^{\gamma-1} A_\gamma(i)\bar{e}_i = A_\gamma(-1)\bar{\varphi}_\gamma - A_\gamma(0)\bar{e}_0 - A_\gamma(1)\bar{e}_1 - A_\gamma(2)\bar{e}_2 - \ldots - A_\gamma(\gamma-1)\bar{e}_{\gamma-1}$$

To see how the coefficients of the triangular array and the $\bar{e}_j$, $0 \le j \le N$, are used, define N + 1 functions as follows:

$$f_0(\beta) = A_0(-1)\varphi_0(\beta)$$

$$f_j(\beta) = A_j(-1)\varphi_j(\beta) - \sum_{i=0}^{j-1} A_j(i)f_i(\beta), \quad 1 \le j \le N.$$

Then each $f_j(\beta)$ is a linear combination of the N + 1 functions

$\varphi_0(\beta)$, $\varphi_1(\beta)$, ... , $\varphi_N(\beta)$, and

$$\bar{f}_j = (f_j(\beta_0),\ f_j(\beta_1),\ \ldots\ ,\ f_j(\beta_n)) = \bar{e}_j.$$

Thus,
$$F(A_0, A_1, \ldots, A_N) = \sum_{i=0}^{n} \{X(\beta_i) - \sum_{j=0}^{N} A_j\varphi_j(\beta_i)\}^2$$

$$= \sum_{i=0}^{n} \{X(\beta_i) - \sum_{j=0}^{N} A'_j f_j(\beta_i)\}^2$$

$$= F'(A'_0, A'_1, \ldots, A'_N)$$

and the necessary condition that F' be minimum yields the normal equations

$$A'_0 = (\bar{X}, \bar{f}_0)$$

$$A'_1 = (\bar{X}, \bar{f}_1)$$

$$\vdots$$

$$A'_N = (\bar{X}, \bar{f}_N).$$

This yields the function $\sum_{j=0}^{N} A_j' f_j(\beta) = \sum_{j=0}^{N} A_j \varphi_j(\beta)$ such that

$\sum_{i=0}^{n} \{X(\beta_i) - \sum_{j=0}^{N} A_j \varphi_j(\beta_i)\}^2$ is minimum.

DERIVATION OF A FUNCTION SATISFYING A GIVEN ERROR TOLERANCE IN THE SENSE OF LEAST SQUARES

Although $E = \sum_{i=0}^{n} \{X(\beta_i) - \sum_{j=0}^{N} A_j \varphi_j(\beta_i)\}^2$ is minimized by the least squares procedure, there is no assurance of the relative size of this error. Thus, we need to be able to determine an approximating function in such a fashion that the error, in the sense of least squares, will not exceed a given tolerance. Before doing this, notice that

$$E = \sum_{i=0}^{n} [X(\beta_i) - \sum_{j=0}^{N} A_j \varphi_j(\beta_i)]^2 = \| \overline{X} - A_0 \overline{\varphi}_0 - \cdots - A_N \overline{\varphi}_N \|^2$$

$$= \| \overline{X} - \sum_{j=0}^{N} (\overline{X}, \overline{e}_j) \overline{e}_j \|^2 = \| \overline{X} \|^2 - [\overline{X}, \sum_{j=0}^{N} (\overline{X}, \overline{e}_j) \overline{e}_j]$$

$$- [\overline{X}, \sum_{j=0}^{N} (\overline{X}, \overline{e}_j) \overline{e}_j] + \| \sum_{j=0}^{N} (\overline{X}, \overline{e}_j) \overline{e}_j \|^2$$

$$= \| \overline{X} \|^2 - \sum_{j=0}^{N} (\overline{X}, \overline{e}_j)^2 .$$

From this we note the following points:

1. $\| \overline{X} \|^2$ is an upper bound of E.
2. A sum of any k of the N + 1 terms $(\overline{X}, \overline{e}_j)^2$, $0 < k < N$, will yield an error $E' > E$.
3. If $\overline{e}_{N+1}$ is any other non-zero vector orthonormal to each of $\overline{e}_0, \overline{e}_1, \ldots, \overline{e}_N$, then $\| \overline{X} \|^2 - \sum_{j=0}^{N+1} (\overline{X}, \overline{e}_j)^2 < E$.

Thus, we may define the problem stated above as follows:

Problem: After evaluating $\| \overline{X} \|^2 - \sum_{j=0}^{N} (\overline{X}, \overline{e}_j)^2$, we find that this value still exceeds a given error tolerance δ. Then we need to find $\overline{\varphi}_{N+1}$ such that

$$\| \overline{X} \|^2 - \sum_{j=0}^{N} (\overline{X}, \overline{e}_j)^2 - (\overline{X}, \overline{e}_{N+1})^2 \leq \delta, \text{ or}$$

$$(\overline{X}, \overline{e}_{N+1})^2 \geq \| \overline{X} \|^2 - \sum_{j=0}^{N} (\overline{X}, \overline{e}_j)^2 - \delta,$$

where $\overline{e}_{N+1}$ is the vector associated with $\overline{\varphi}_{N+1}$ that is orthonormal to $\overline{e}_0, \overline{e}_1, \ldots, \overline{e}_N$.

Solution: Let $\overline{\varphi}_{N+1} = (\lambda_0, \lambda_1, \ldots, \lambda_n)$. Then

$$\overline{\varphi}'_{N+1} = \overline{\varphi}_{N+1} - \sum_{j=0}^{N} (\overline{\varphi}_{N+1}, \overline{e}_j)\overline{e}_j$$

$$= (\lambda_0, \lambda_1, \ldots, \lambda_n) - \sum_{j=0}^{N} \left(\sum_{i=0}^{n} \lambda_i e_{ji} \right) \overline{e}_j,$$

where $\overline{e}_j = (e_{j0}, e_{j1}, \ldots, e_{jn}), \; 0 \leq j \leq N.$

Then

$$\overline{e}_{N+1} = \frac{(\lambda_0, \lambda_1, \ldots, \lambda_n) - \sum_{j=0}^{N} \left(\sum_{i=0}^{n} \lambda_i e_{ji} \right) \overline{e}_j}{\left\{ \sum_{i=0}^{n} \lambda_i^2 - \sum_{j=0}^{N} \left(\sum_{i=0}^{n} \lambda_i e_{ji} \right)^2 \right\}^{1/2}},$$

and if $\overline{X} = (t_0, t_1, \ldots, t_n)$, then

$$(\overline{X}, \overline{e}_{N+1})^2 = \frac{[\sum_{i=0}^{n} \lambda_i t_i - \sum_{j=0}^{N} (\sum_{i=0}^{n} \lambda_i e_{ji})(\overline{X}, \overline{e}_j)]^2}{\sum_{i=0}^{n} \lambda_i^2 - \sum_{j=0}^{N} (\sum_{i=0}^{n} \lambda_i e_{ji})^2} .$$

Thus, to have $(\overline{X}, \overline{e}_{N+1})^2 \geq \| \overline{X} \|^2 - \sum_{j=0}^{N} (\overline{X}, \overline{e}_j)^2 - \delta$,

we must have

$$[\sum_{i=0}^{n} \lambda_i t_i - \sum_{j=0}^{N} (\sum_{i=0}^{n} \lambda_i e_{ji})(\overline{X}, \overline{e}_j)]^2 \geq$$

$$[\sum_{i=0}^{n} \lambda_i^2 - \sum_{j=0}^{N} (\sum_{i=0}^{n} \lambda_i e_{ji})^2] [\| \overline{X} \|^2 - \sum_{j=0}^{N} (\overline{X}, \overline{e}_j)^2 - \delta], \text{ or,}$$

equivalently, we must have

$$\sum_{i=0}^{n} [\lambda_i^2 \{t_i - (\overline{X}, \overline{e}_0)e_{0i} - \ldots - (\overline{X}, \overline{e}_N)e_{Ni}\}^2$$

$$+ \{ \| \overline{X} \|^2 - \sum_{j=0}^{N} (\overline{X}, \overline{e}_j)^2 - \delta\} \{e_{Ni}^2 + \ldots + e_{0i}^2 - 1\}]$$

$$+ 2\lambda_i \left[\sum_{\substack{k=0 \\ k>i}}^{n} \lambda_k \{t_i - (\overline{X},\overline{e}_0)e_{0i} - \dots - (\overline{X},\overline{e}_N)e_{Ni}\}\{t_k - (\overline{X},\overline{e}_0)e_{0k} - \dots - (\overline{X},\overline{e}_N)e_{Nk}\} \right.$$

$$+ \{ \| \overline{X} \|^2 - \sum_{j=0}^{N} (\overline{X},\overline{e}_j)^2 - \delta\}\{ \sum_{\substack{k=0 \\ k>i}}^{n} \lambda_k e_{0i}e_{0k} + \dots + \sum_{\substack{k=0 \\ k>i}}^{n} \lambda_k e_{Ni}e_{Nk}\}] \geq 0.$$

We can write this as

$$\sum_{i=0}^{n} (A_i\lambda_i^2 + B_i\lambda_i) \geq 0,$$

where

$$A_i = [\{t_i - (\overline{X},\overline{e}_0)e_{0i} - \dots - (\overline{X},\overline{e}_N)e_{Ni}\}^2 + \{ \| \overline{X} \|^2 - \sum_{j=0}^{N} (\overline{X},\overline{e}_j)^2 - \delta\}\{e_{Ni}^2 + \dots + e_{0i}^2 - 1\}]$$

and

$$B_i = 2[\sum_{\substack{k=0 \\ k>i}}^{n} \lambda_k \{t_i - (\overline{X},\overline{e}_0)e_{0i} - \dots - (\overline{X},\overline{e}_N)e_{Ni}\}\{t_k - (\overline{X},\overline{e}_0)e_{0k} - \dots - (\overline{X},\overline{e}_N)e_{Nk}\}$$

$$+ \{ \| \bar{X} \|^2 - \sum_{j=0}^{N} (\bar{X}, \bar{e}_j)^2 - \delta\}\{ \sum_{\substack{k=0 \\ k>i}}^{n} \lambda_k e_{0i} e_{0k} + \dots + \sum_{\substack{k=0 \\ k>i}}^{n} \lambda_k e_{Ni} e_{Nk}\}] .$$

The inequality $\sum_{i=0}^{n} (A_i\lambda_i^2 + B_i\lambda_i) \geq 0$ is satisfied if $A_i\lambda_i^2 + B_i\lambda_i \geq 0$, for $i = 0,1,\dots,n$.

Case 1: $A_i \geq 0$, for some i, $0 \leq i \leq n$. Assume $A_n < 0$, $A_{n-1} < 0$, ..., $A_{i+1} < 0$ and $A_i \geq 0$. Then we need to solve

$$(1) \quad A_i\lambda_i^2 + B_i\lambda_i + (A_n\lambda_n^2 + A_{n-1}\lambda_{n-1}^2 + B_{n-1}\lambda_{n-1} + A_{n-2}\lambda_{n-2}^2 + B_{n-2}\lambda_{n-2} + \dots + A_{i+1}\lambda_{i+1}^2 + B_{i+1}\lambda_{i+1}) \geq 0.$$

Thus, we must have

$$B_i^2 - 4A_iA_n\lambda_n^2 - 4A_iA_{n-1}\lambda_{n-1}^2 - 4A_iB_{n-1}\lambda_{n-1} - 4A_iA_{n-2}\lambda_{n-2}^2 - 4A_iB_{n-2}\lambda_{n-2} - \dots - 4A_iA_{i+1}\lambda_{i+1}^2 - 4A_iB_{i+1}\lambda_{i+1} \geq 0.$$

Then choose λ_n arbitrarily and $\lambda_{n-1}, \lambda_{n-2}, \dots, \lambda_{i+1}$ as follows:

sign $\lambda_j = -(\text{sign } B_j)$, $i+1 \leq j \leq n-1$.

In addition, choose λ_i to be any solution of (1) and λ_k to be any solution of $A_k\lambda_k^2 + B_k\lambda_k = \lambda_k(A_k\lambda_k + B_k) \geq 0$, $k = 0,1,\dots,i-1$.

Case 2: If $A_i < 0$ for all i, choose $\lambda_n = 1$ and examine

(1) $B^2_{n-1} - 4A_nA_{n-1}$.

If (1) is greater than or equal to zero, we are assured of a solution to the equation

(2) $A_{n-1}\lambda^2_{n-1} + B_{n-1}\lambda_{n-1} + A_n\lambda^2_n = 0$. Then let λ_{n-1} be either solution of (2) and let $\lambda_j = -\dfrac{B_j}{A_j}$, $j = 0,1,\ldots,n-2$.

If (1) is negative choose $\lambda_n = 0$, and let $\lambda_{n-1} = 1$ and examine

(3) $B^2_{n-2} - 4A_{n-1}A_{n-2}$.

If (3) is positive or zero, then

$A_{n-2}\lambda^2_{n-2} + B_{n-2}\lambda_{n-2} + A_{n-1}\lambda^2_{n-1} = 0$ has a solution.

Let λ_{n-2} be either of these and let $\lambda_j = -\dfrac{B_j}{A_j}$,

$j = 0,1,\ldots,n-3$, etc.

In order for $A_i \geq 0$ for some i; we must have

$$[t_i - \sum_{j=0}^{N} (\overline{X},\overline{e}_j)e_{ji}]^2 + [\,\|\,\overline{X}\,\|^2 - \sum_{j=0}^{N} (\overline{X},\overline{e}_j)^2 - \delta]$$

$$[(\sum_{j=0}^{N} e_{ji}{}^2) - 1] > 0, \text{ or}$$

$$[\,\|\bar{X}\|^2 - \sum_{j=0}^{N} (\bar{X},\bar{e}_j)^2 - \delta]\,[1 - \sum_{j=0}^{N} e_{ji}^{\,2}] < [t_i - \sum_{j=0}^{N} (\bar{X},\bar{e}_j)e_{ji}]^2, \text{ or}$$

$$[\,\|\bar{X}\|^2 - \sum_{j=0}^{N} (\bar{X},\bar{e}_j)^2 - \delta] < \frac{[t_i - \sum_{j=0}^{N} (\bar{X},\bar{e}_j)e_{ji}]^2}{1 - \sum_{j=0}^{N} e_{ji}^{\,2}}.$$ But we may write

this as $$\delta > \|\bar{X}\|^2 - \sum_{j=0}^{N} (\bar{X},\bar{e}_j)^2 + \frac{[t_i - \sum_{j=0}^{N} (\bar{X},\bar{e}_j)e_{ji}]^2}{\sum_{j=0}^{N} e_{ji}^{\,2} - 1}, \text{ or}$$

$$\delta > E + \frac{[t_i - \sum_{j=0}^{N} (\bar{X},\bar{e}_j)e_{ji}]^2}{\sum_{j=0}^{N} e_{ji}^{\,2} - 1}.$$

In the newly computed vector $\bar{\varphi}_{N+1} = (\lambda_0,\lambda_1,\ldots,\lambda_n)$ let λ_i be the value of some function $\varphi_{N+1}(\beta)$ at β_i, i.e., $\varphi_{N+1}(\beta_i) = \lambda_i$. Then

$$E = \sum_{i=0}^{n} [X(\beta_i) - \sum_{j=0}^{N} A_j\varphi_j(\beta_i) - A_{N+1}\varphi_{N+1}(\beta_i)]^2 < \delta.$$

<u>Problem</u>: Determine $\varphi_{N+1}(\beta')$ for some value $\beta' \neq \beta_i$, $0 \leq i \leq n$, such that the error obtained by using $\sum_{j=0}^{N+1} A_j\varphi_j(\beta')$ to approximate $X(\beta')$ in the sense of least squares, does not exceed the error obtained

by approximating $X(\beta')$ with $\sum_{j=0}^{N} A_j \varphi_j(\beta')$.

Solution: Compute $A_{N+1}(k)$, $k = -1, 0, 1, \ldots, N$, $\bar{e}_{N+1}$ and A'_{N+1} as follows:

$$A_{N+1}(-1) = \frac{1}{\left\| \bar{\varphi}_{N+1} - \sum_{j=0}^{N} (\bar{\varphi}_{N+1}, \bar{e}_j)\bar{e}_j \right\|}$$

$$A_{N+1}(0) = A_{N+1}(-1)A_0(-1)(\bar{\varphi}_{N+1}, \bar{\varphi}_0)$$

$$\vdots$$

$$A_{N+1}(N) = A_{N+1}(-1)A_N(-1)(\bar{\varphi}_{N+1}, \bar{\varphi}_N) - \sum_{j=0}^{N-1} A_{N+1}(j)A_N(j)$$

$$\bar{e}_{N+1} = A_{N+1}(-1)\bar{\varphi}_{N+1} - \sum_{j=0}^{N} A_{N+1}(j)\bar{e}_j.$$

$$A'_{N+1} = (\bar{X}, \bar{e}_{N+1}).$$

Finally, compute the $(N+2) A_j$'s, $j = 0, 1, \ldots, N+1$, as follows:

$$A_{N+1} = A'_{N+1}A_{N+1}(-1)$$

$$A_N = A_N(-1)[A'_N - A'_{N+1}A_{N+1}(N)]$$

$$A_{N-1} = A_{N-1}(-1)\{A'_{N-1} - A'_N A_N(N-1) + A'_{N+1}[-A_{N+1}(N-1) + A_{N+1}(N)A_N(N-1)]\}$$

$$\vdots$$

Now let $\beta_{i'}$ be a β_i such that $\| \beta_{i'} - \beta' \| = \min\limits_{0 \leq i \leq n} \{ \| \beta_i - \beta' \| \}$.

Define the following function

$$\sum_{j=0}^{N} A_j \varphi_j(\beta') + A_{N+1} M(\beta'),$$

where

$$M(\beta') = \lambda_{i'} \left[\frac{L(\beta_{i'}) - 2 \| \beta_{i'} - \beta' \|}{L(\beta_{i'})} \right]$$

for $2 \| \beta_{i'} - \beta' \| < L(\beta_{i'})$

$$= 0, \text{ otherwise,}$$

and $L(\beta_{i'}) = \min\limits_{\substack{0 \leq i \leq n \\ i \neq i'}} \{ \| \beta_i - \beta_{i'} \| \}$.

Thus, when β' is chosen, we are able to use the function above to approximate $X(\beta')$, being assured that the approximation obtained here is no worse than the value $\sum_{j=0}^{N} A_j \varphi_j(\beta')$ obtained by using the initial least squares approximating function.

Writing this multiple of λ_i as

$$\frac{\frac{1}{2} L(\beta_{i'}) - \| \beta_{i'} - \beta' \|}{\frac{1}{2} L(\beta_{i'})},$$

we see that we have a factor which varies from zero to one as β' varies from a position on the boundary to a position at the center of the ball

$$\left\{ \beta \mid \| \beta_{i'} - \beta' \| \le \frac{1}{2} L(\beta_{i'}) \right\}.$$

Thus, the factor λ_i, which was derived in association with the vector β_i, is weighted depending on the nearness of β' to $\beta_{i'}$.

Achieser, N. I., *Theory of Approximation*, Ungar Publishing Company, New York, 1956.

Approximation of Functions, Symposium held at General Motors Research Laboratories, Elsevier Publishing Company, 1965, edited by Henry L. Garabedian.

On Approximation Theory, Proceedings of the conference held in the Mathematical Research Institute at Oberwolfach, Black Forest, 1963, Birkhauser Verlag, 1964.

On Numerical Approximation, Proceedings of a Symposium conducted by the Mathematics Research Center, United States Army, at the University of Wisconsin, Madison, 1958, edited by Rodolph E. Langer.

A General Solution for Minimum Impulse Transfers in the Near Vicinity of a Circular Orbit*

by T. N. Edelbaum
Analytical Mechanics Associates, Inc.
Cambridge, Massachusetts

ABSTRACT

An analytic solution is obtained for minimum impulse transfer between two neighboring low-eccentricity orbits. The orientations of the arguments of perigee and the line of nodes are completely arbitrary. The solutions obtained are of three different classes. The first class consists of two-impulse nodal transfers with both impulses occurring on the line of nodes and having equal but opposite radial, circumferential and normal direction cosines for the impulses. The second class also consists of two-impulse solutions but with equal circumferential direction cosines, and equal and opposite radial and normal direction cosines. The location of the two impulses for this class is a function of the particular transfer. The third class consists of singular solutions with a well-defined thrust direction at every point but with an infinite number of solutions for the distribution of impulses (or even continuous finite thrusts) all having the same fuel consumption. The singular solution can be realized with two impulses so that two impulses suffice for all of these optimum transfers.

*Performed under contract NAS 12-26.

NOMENCLATURE

a	Semi-major axis
a_o	Semi-major axis of circular reference orbit
C	$\cos(\theta - \theta_o)$
e	Eccentricity
e_x	x component of eccentricity
e_y	y component of eccentricity
H	Variational Hamiltonian
i	Inclination of orbit planes
i_x	x component of inclination
i_y	y component of inclination
K	Gravitational constant
ℓ_N	Direction cosine of the thrust normal to the orbit plane
ℓ_R	Radial direction cosine of the thrust
ℓ_T	Circumferential direction cosine of the thrust
p	Magnitude of the primer vector
R	Defined by Eq. (50)
S	$\sin(\theta - \theta_o)$
T	Defined by Eq. (61)
u	Time integral of thrust acceleration
x_i	Defined by Eqs. (1)-(5)
δ	Defined by Eq. (31)
θ	Central angle
λ	Radial component of the primer vector
λ_i	Lagrange multipliers
μ	Circumferential component of the primer vector
ν	Normal component of the primer vector
ω	Defined by Eq. (60)
Ω	Defined by Eq. (60)

INTRODUCTION

The theory of optimal space vehicle guidance has much in common with classical applied mechanics and celestial mechanics. As might be expected, many of the techniques of these classical disciplines have been applied to problems of optimal space vehicle guidance. Such techniques include Taylor series expansions, linear perturbation theory, higher order perturbation theory, matched asymptotic expansions, the method of averaging, and Hamilton-Jacobi theory. Many of these techniques have been quite successful for some problems. However, optimal guidance problems introduce some special difficulties that are not present in typical problems of celestial mechanics.

One of the usual procedures for transforming an optimal guidance problem into a problem in Hamiltonian mechanics is to use the maximum principle, or its classical equivalents, to express the control in terms of the adjoint. The control is often a highly nonlinear function of the adjoint so that small changes in the adjoint may cause large changes in the control. This simple fact can cause many of the methods which are successful for problems in celestial mechanics to break down when applied to problems in optimal guidance.

This simple fact can even cause trouble with perturbation methods specifically developed for optimal guidance, such as the method of "neighboring optimal" or "second variation" guidance. This guidance method linearizes the state, the adjoint and the control, and can cause difficulty towards the terminal point where small changes in the state may require large nonlinear changes in the control. The same difficulty can occur with higher order schemes, such as n^{th} variation guidance, because the expansion of the solution in variations of various orders may not converge.

One method of studying this terminal accuracy problem is to linearize the problem about the final state. The resulting trajectory optimization and optimal guidance problems may then sometimes be solved analytically. For power-limited rockets, the control is a linear function of the adjoint and the above difficulties do not arise. Minimum-fuel power-limited solutions have been obtained for three-dimensional linearizations about circular orbits (Ref. 1) and elliptic orbits (Ref. 2). These solutions are for rendezvous in a fixed time.

For rockets with constant exhaust velocity, the control is a nonlinear function of the adjoint and the problem is far more difficult. Even if no bounds are placed on the control magnitude (so that impulses are allowed) and the transfer time and terminal positions are left open, only limited (but significant) success has been achieved. The coplanar problem for elliptic terminal orbits has been partially solved in Refs. 3 and 4. A more complete solution has been obtained for three-dimensional transfers in the vicinity of a circular orbit (Refs. 4 and 5). The present paper contains additional details on the solution of Ref. 4, including the synthesis of the optimal control and the determination of the minimum number of impulses required for the singular case. Ref. 5 represents an independent derivation of the results contained herein and in Ref. 4. It includes some consideration of the effects of a slightly eccentric reference orbit.

For a more complete review of the existing literature on optimal orbital transfer, the two survey papers prepared under this contract should be consulted (Refs. 6 and 7).

ANALYSIS

The solution of this problem is conveniently carried out in terms of Lagrange's planetary variables. By linearizing the variation of parameter equations about a circular reference orbit, the equations of motion (1)-(5) are obtained.

$$\frac{dx_1}{du} = \frac{da/a_o}{du} = 2\sqrt{\frac{a_o}{K}}\, \ell_T \tag{1}$$

$$\frac{dx_2}{du} = \frac{de_y}{du} = \sqrt{\frac{a_o}{K}}\, (2\,\ell_T \sin\theta - \ell_R \cos\theta) \tag{2}$$

$$\frac{dx_3}{du} = \frac{de_x}{du} = \sqrt{\frac{a_o}{K}}\, (2\,\ell_T \cos\theta + \ell_R \sin\theta) \tag{3}$$

$$\frac{dx_4}{du} = \frac{di_y}{du} = \sqrt{\frac{a_o}{K}}\, \ell_N \sin\theta \tag{4}$$

$$\frac{dx_5}{du} = \frac{di_x}{du} = \sqrt{\frac{a_o}{K}}\, \ell_N \cos\theta \tag{5}$$

Following Contensou (Ref. 8) and Breakwell (Ref. 9), the independent variable is taken as the time integral of the thrust acceleration. If the thrust is impulsive, this integral is equal to the sum of the absolute magnitudes of the impulses. The variable θ is the angular position in the reference orbit measured from the x axis. Both eccentricity and inclination are treated as vectors having components along the x and y axes which lie in the plane of the reference orbit. The ℓ's are the direction cosines of the thrust in the radial, circumferential and normal directions.

The variational problem is formulated in terms of the Hamiltonian defined by Eq. (6).

$$H = \sum_{i=1}^{5} \lambda_i \frac{dx_i}{du} \tag{6}$$

Following Lawden (Ref. 10), we shall introduce a primer vector which is the adjoint of the velocity vector. The magnitude of this vector will be denoted by p and its components in the radial, circumferential and normal directions will be denoted by λ, μ and ν, respectively. The following equations for the optimum thrust direction are derived by means of the maximum principle.

$$\ell_R = \frac{\lambda}{p} \qquad \ell_T = \frac{\mu}{p} \qquad \ell_N = \frac{\nu}{p} \tag{7}$$

$$\lambda = \sqrt{\frac{a_o}{K}} \, (-\lambda_2 \cos\theta + \lambda_3 \sin\theta) \tag{8}$$

$$\mu = \sqrt{\frac{a_o}{K}} \, (2\lambda_1 + 2\lambda_2 \sin\theta + 2\lambda_3 \cos\theta) \tag{9}$$

$$\nu = \sqrt{\frac{a_o}{K}} \, (\lambda_4 \sin\theta + \lambda_5 \cos\theta) \tag{10}$$

The location of the impulse is given by the value of θ where p takes on its absolute maximum. If there is to be more than one impulse, all of these maxima must be equal in magnitude.

Following Lawden (Ref. 10), Eqs. (8)-(10) will be rewritten in a different form.

$$\lambda = \sqrt{\frac{a_o}{K}} \sqrt{\lambda_2^2 + \lambda_3^2} \, \sin(\theta - \theta_o) \tag{11}$$

$$\mu = \sqrt{\frac{a_o}{K}} \left[2\lambda_1 + 2\sqrt{\lambda_2^2 + \lambda_3^2} \cos(\theta - \theta_o) \right] \tag{12}$$

$$\nu = \sqrt{\frac{a_o}{K}} \left[\frac{\lambda_3\lambda_4 - \lambda_2\lambda_5}{\sqrt{\lambda_2^2 + \lambda_3^2}} \sin(\theta - \theta_o) + \frac{\lambda_2\lambda_4 + \lambda_3\lambda_5}{\sqrt{\lambda_2^2 + \lambda_3^2}} \cos(\theta - \theta_o) \right] \tag{13}$$

$$\tan \theta_o = \frac{\lambda_2}{\lambda_3} \tag{14}$$

Equations (11)-(13) are the equations of an ellipse in three-space. This ellipse is formed by the intersection of a two-to-one elliptic cylinder parallel to the ν axis and a plane which passes through the intersection of the cylinder axis with the λ, μ plane. A typical case is illustrated in Fig. 1 which also shows the projection of the ellipse on the λ, μ plane.

There are only three configurations of this elliptical primer locus which allow the primer vector to have more than one maximum. The first configuration is a family of solutions where the center of the ellipse is located at the origin (Fig. 2). In this case the two equal maxima occur on the major axis of the ellipse and are separated by 180^o. The following equations characterize this case, which will be referred to as the nodal case.

$$\lambda_1 = 0 \tag{15}$$

$$\theta_2 = \theta_1 + \pi \tag{16}$$

$$\left.\begin{aligned} \lambda(\theta_1) &= -\lambda(\theta_2) \\ \mu(\theta_1) &= -\mu(\theta_2) \\ \nu(\theta_1) &= -\nu(\theta_2) \end{aligned}\right\} \tag{17}$$

The second configuration, which allows two equal maxima of the primer vector, corresponds to cases where the ellipse passes through the μ axis and the primer vectors again lie in a single plane (Fig. 3). This case, which will be referred to as the nondegenerate case, is characterized by the following equations and inequalities

$$\lambda_2 \lambda_4 = -\lambda_3 \lambda_5 \tag{18}$$

$$\theta_2 - \theta_o = \theta_o - \theta_1 \tag{19}$$

$$\left.\begin{aligned} \lambda(\theta_1) &= -\lambda(\theta_2) \\ \mu(\theta_1) &= \mu(\theta_2) \\ \nu(\theta_1) &= -\nu(\theta_2) \end{aligned}\right\} \tag{20}$$

$$\cos(\theta_1 - \theta_o) = \frac{4\lambda_1\left(\sqrt{\lambda_2^2 + \lambda_3^2}\right)^3}{(\lambda_3\lambda_4 - \lambda_2\lambda_5)^2 - 3(\lambda_2^2 + \lambda_3^2)^2} \tag{21}$$

$$3(\lambda_2^2 + \lambda_3^2)^2 < (\lambda_3\lambda_4 - \lambda_2\lambda_5)^2 \tag{22}$$

$$4\lambda_1\left(\sqrt{\lambda_2^2 + \lambda_3^2}\right)^3 \leq (\lambda_3\lambda_4 - \lambda_2\lambda_5)^2 - 3(\lambda_2^2 + \lambda_3^2)^2 \tag{23}$$

The third configuration is a combination of the two previous cases where the primer locus forms a circle and the primer vector has the same magnitude at all points on the orbit (Fig. 4). This singular case is characterized by the following equations where the magnitude of p has been taken to be unity.

$$\lambda_1 = 0 \tag{24}$$

$$\lambda_2 \lambda_4 = - \lambda_3 \lambda_5 \tag{25}$$

$$3(\lambda_2^2 + \lambda_3^2)^2 = (\lambda_3 \lambda_4 - \lambda_2 \lambda_5)^2 = \frac{3}{4} \frac{K}{a_o} \tag{26}$$

$$\lambda = \frac{1}{2} \sin (\theta - \theta_o) \tag{27}$$

$$\mu = \cos (\theta - \theta_o) \tag{28}$$

$$\nu = \frac{\sqrt{3}}{2} \sin (\theta - \theta_o) \tag{29}$$

The above results for the equal maxima of the primer vector were first obtained by purely geometrical reasoning. They have been analytically verified by M. Washington in Ref. 4.

The admissible adjoint solutions have now been determined. The next problem is the solution of the two-point boundary value problem to determine the optimum transfers corresponding to each adjoint solution. The simplest case is that of nodal transfer. The x axis may be aligned with the line of nodes between the initial and final orbits. Both impulses must occur on this line of nodes because of the 180^o central angle separating the impulse locations. The total change in the orbital elements is given by Eqs. (30)-(36) using Eq. (17).

$$u \equiv u_1 + u_2 \tag{30}$$

$$\delta \equiv \frac{u_2}{u} \tag{31}$$

$$\frac{\Delta x_1}{u} = \frac{\Delta a/a_o}{u} = 2\sqrt{\frac{a_o}{K}}\ \ell_{T_1}(1-2\delta) \tag{32}$$

$$\frac{\Delta x_2}{u} = \frac{\Delta e_y}{u} = -\sqrt{\frac{a_o}{K}}\ \ell_{R_1} \tag{33}$$

$$\frac{\Delta x_3}{u} = \frac{\Delta e_x}{u} = 2\sqrt{\frac{a_o}{K}}\ \ell_{T_1} \tag{34}$$

$$\frac{\Delta x_4}{u} = 0 \tag{35}$$

$$\frac{\Delta x_5}{u} = \frac{\Delta i}{u} = \sqrt{\frac{a_o}{K}}\ \ell_{N_1} \tag{36}$$

These equations may be solved simultaneously to determine the total required impulse [Eq. (37)] and the magnitude of the first impulse [Eq. (38)].

$$u = \sqrt{\frac{K}{a_o}}\sqrt{\Delta i^2 + \frac{\Delta e_x^2}{4} + \Delta e_y^2} \tag{37}$$

$$u_1 = \frac{\Delta e_x + \Delta a/a_o}{2\Delta e_x}\sqrt{\frac{K}{a_o}}\sqrt{\Delta i^2 + \frac{\Delta e_x^2}{4} + \Delta e_y^2} \tag{38}$$

There are two bounds on the region of applicability of this solution. The first is given by the requirement that each impulse must point in the positive direction of the primer vector.

$$\left(\frac{\Delta a}{a_o}\right)^2 \le \Delta e_x^2 \tag{39}$$

The second inequality is given by the requirement that the primer vector must have a maximum at the impulse locations.

$$\Delta i^2 \geq 3 \Delta e_y^2 \tag{40}$$

The optimal directions of the impulses must still be determined. An instructive way to do this is to differentiate the payoff with respect to the state. If the maximum magnitude of the primer vector (which is equal to the magnitude of the Hamiltonian) is taken as unity, the Lagrange multiplier for each component of the state is the partial derivative of the payoff with respect to that component of the state.

$$\lambda_1 = 0 \tag{41}$$

$$\lambda_2 = \sqrt{\frac{K}{a_o}} \frac{\Delta e_y}{\sqrt{\Delta i^2 + \frac{\Delta e_x^2}{4} + \Delta e_y^2}} \tag{42}$$

$$\lambda_3 = \sqrt{\frac{K}{a_o}} \frac{\Delta e_x}{4\sqrt{\Delta i^2 + \frac{\Delta e_x^2}{4} + \Delta e_y^2}} \tag{43}$$

$$\lambda_4 = -\frac{3}{4}\sqrt{\frac{K}{a_o}} \frac{\Delta e_x \Delta e_y}{\Delta i_x \sqrt{\Delta i^2 + \frac{\Delta e_x^2}{4} + \Delta e_y^2}} \tag{44}$$

$$\lambda_5 = \sqrt{\frac{K}{a_o}} \frac{\Delta i}{\sqrt{\Delta i^2 + \frac{\Delta e_x^2}{4} + \Delta e_y^2}} \tag{45}$$

The direction cosines of the impulses are given by the components of the primer vector at the impulses.

$$\lambda_1 = \frac{-\Delta e_y}{\sqrt{\Delta i^2 + \frac{\Delta e_x^2}{4} + \Delta e_y^2}} \tag{46}$$

$$\mu_1 = \frac{\Delta e_x}{2\sqrt{\Delta i^2 + \frac{\Delta e_x^2}{4} + \Delta e_y^2}} \tag{47}$$

$$\nu_1 = \frac{\Delta i}{\sqrt{\Delta i^2 + \frac{\Delta e_x^2}{4} + \Delta e_y^2}} \tag{48}$$

The solution of the two-point boundary value problem for the nondegenerate case is more complicated and involves considerable algebra. It is convenient to use a different normalization of the Lagrange multipliers. This normalization will be retained only to Eq. (67), where the conventional normalization with the primer vector equal to unity will be reintroduced.

$$\sqrt{\frac{a_o}{K}} \; \frac{\lambda_3 \lambda_4 - \lambda_2 \lambda_5}{\sqrt{\lambda_2^2 + \lambda_3^2}} \equiv 1 \tag{49}$$

$$R \equiv \sqrt{\frac{a_o}{K}} \sqrt{\lambda_2^2 + \lambda_3^2} \tag{50}$$

$$\left.\begin{aligned} C &\equiv \cos(\theta - \theta_o) \\ S &\equiv \sin(\theta - \theta_o) \end{aligned}\right\} \tag{51}$$

With these definitions, the direction cosines of the thrust at the impulse locations may be written:

$$\ell_R = \frac{\lambda}{p} = \frac{2R^2 S}{\sqrt{1+R^2}\,\sqrt{4R^2 + (1-3R^2)C^2}} \tag{52}$$

$$\ell_T = \frac{\mu}{p} = \frac{\sqrt{1+R^2}\quad C}{\sqrt{4R^2 + (1-3R^2)C^2}} \tag{53}$$

$$\ell_N = \frac{\nu}{p} = \frac{2RS}{\sqrt{1+R^2}\,\sqrt{4R^2 + (1-3R^2)C^2}} \tag{54}$$

The Lagrange multiplier λ_1 has been eliminated from these equations by use of Eq. (21). The total change in each element of the orbit can now be given by Eqs. (55)-(59).

$$\frac{\Delta x_1}{u} = 2C\sqrt{\frac{1+R^2}{4R^2 + (1-3R^2)C^2}} \tag{55}$$

$$\frac{\Delta x_2}{u} = \frac{2(C^2 + R^2)}{\sqrt{1+R^2}\,\sqrt{4R^2 + (1-3R^2)C^2}} \tag{56}$$

$$\frac{\Delta x_3}{u} = \frac{2CS(1-2\delta)}{\sqrt{1+R^2}\ \sqrt{4R^2+(1-3R^2)C^2}} \tag{57}$$

$$\frac{\Delta x_4}{u} = \frac{2RS^2}{\sqrt{1+R^2}\ \sqrt{4R^2+(1-3R^2)C^2}} \tag{58}$$

$$\frac{\Delta x_5}{u} = \frac{2RCS(1-2\delta)}{\sqrt{1+R^2}\ \sqrt{4R^2+(1-3R^2)C^2}} \tag{59}$$

The x axis has been aligned with θ_o for these equations. This orientation will also be abandoned after Eq. (67), when the orientation along the line of nodes will be reintroduced.

The two-point boundary value problem will now be solved in terms of the angle between the vectors describing the eccentricity change and the inclination change. The following two angles are first introduced.

$$\tan\omega \equiv \frac{\Delta x_2}{\Delta x_3} \qquad \tan\Omega \equiv \frac{\Delta x_4}{\Delta x_5} \tag{60}$$

The variable T defined in Eq. (61) can be determined from Eqs. (56)-(59).

$$T \equiv \tan(\omega-\Omega) = \frac{(1-2\delta)^2 C^2 S - (C^2+R^2)S}{(1-2\delta)(1+R^2)C} \tag{61}$$

Eq. (61) can now be solved for the impulse split in terms of T.

$$1-2\delta = \frac{T(1+R^2) - \sqrt{T^2(1+R^2) + 4S^2(C^2+R^2)}}{2SC} \tag{62}$$

The total change in eccentricity and inclination can now be calculated in terms of C, R, S, and T.

$$\frac{\Delta e}{u} = \sqrt{\frac{\Delta x_2^2}{u^2} + \frac{\Delta x_3^2}{u^2}}$$

$$= \sqrt{\frac{4(C^2 + R^2) + 2T^2(1+R^2) - 2T\sqrt{T^2(1+R^2)^2 + 4S^2(R^2 + C^2)}}{4R^2(1-3R^2)C^2}} \tag{63}$$

$$\frac{\Delta i}{u} = \sqrt{\frac{\Delta x_4^2}{u^2} + \frac{\Delta x_5^2}{u^2}}$$

$$= R\sqrt{\frac{4S^2 + 2T^2(1+R^2) - 2T\sqrt{T^2(1+R^2)^2 + 4S^2(R^2 + C^2)}}{4R^2(1-3R^2)C^2}} \tag{64}$$

By multiplying Eq. (63) by R and dividing it by Eq. (55), and dividing Eq. (64) by Eq. (55), and subtracting the squares of these two quantities, the following result may be obtained.

$$C^2 = \frac{R^2(1-R^2)\dfrac{\Delta a^2}{a_o^2}}{2R^2\dfrac{\Delta a^2}{a_o^2} - (1+R^2)(R^2\Delta e^2 - \Delta i^2)} \tag{65}$$

This equation allows C and S to be eliminated from the previous equations and allows R to be determined in terms of the changes in the orbital elements.

$$R^2 = 1 + \frac{1}{2}\left[\frac{\frac{\Delta a^2}{a_o^2} + \Delta i^2 - \Delta e^2}{\Delta i\ \Delta e \sin(\omega - \Omega)}\right]^2 - \sqrt{\left[\frac{\frac{\Delta a^2}{a_o^2} + \Delta i^2 - \Delta e^2}{\Delta i\ \Delta e \sin(\omega - \Omega)}\right]^2 + \frac{1}{4}\left[\frac{\frac{\Delta a^2}{a_o^2} + \Delta i^2 - \Delta e^2}{\Delta i\ \Delta e \sin(\omega - \Omega)}\right]^4} \tag{66}$$

The payoff now can be calculated in terms of the orbital elements.

$$u = \sqrt{\frac{K}{2a_o}}\sqrt{\Delta i^2 + \Delta e^2 - \frac{\Delta a^2}{2a_o^2} + \sqrt{\left(\Delta i^2 - \Delta e^2 + \frac{\Delta a^2}{a_o^2}\right)^2 + 4\Delta i^2 \Delta e^2 \sin^2(\omega - \Omega)}} \tag{67}$$

At this point the orientation of the x axis along the line of nodes and the normalization of the Lagrange multipliers with $p = 1$, used in the rest of this report, are reintroduced. The equation for the total required impulse now becomes Eq. (68).

$$u = \sqrt{\frac{K}{2a_o}}\sqrt{\Delta i^2 + \Delta e_x^2 + \Delta e_y^2 - \frac{\Delta a^2}{2a_o^2} + \sqrt{\left(\Delta i^2 - \Delta e_x^2 - \Delta e_y^2 + \frac{\Delta a^2}{a_o^2}\right)^2 + 4\Delta i^2 \Delta e_y^2}} \tag{68}$$

The Lagrange multipliers are now most easily obtained by differentiating the payoff, as was done for the nodal case.

$$\lambda_1 = \frac{K\Delta a}{2a_o^2 u}\left[-\frac{1}{2} + \frac{\Delta i^2 - \Delta e_x^2 - \Delta e_y^2 + \frac{\Delta a^2}{a_o^2}}{\sqrt{\left(\Delta i^2 - \Delta e_x^2 - \Delta e_y^2 + \frac{\Delta a^2}{a_o^2}\right)^2 + 4\Delta i^2 \Delta e_y^2}}\right] \tag{69}$$

$$\lambda_2 = \frac{K\,\Delta e_y}{2 a_o u}\left[1 - \frac{-\Delta i^2 - \Delta e_x^2 - \Delta e_y^2 + \dfrac{\Delta a^2}{a_o^2}}{\sqrt{\left(\Delta i^2 - \Delta e_x^2 - \Delta e_y^2 + \dfrac{\Delta a^2}{a_o^2}\right)^2 + 4\,\Delta i^2\,\Delta e_y^2}}\right] \tag{70}$$

$$\lambda_3 = \frac{K\,\Delta e_x}{2 a_o u}\left[1 - \frac{\Delta i^2 - \Delta e_x^2 - \Delta e_y^2 + \dfrac{\Delta a^2}{a_o^2}}{\sqrt{\left(\Delta i^2 - \Delta e_x^2 - \Delta e_y^2 + \dfrac{\Delta a^2}{a_o^2}\right)^2 + 4\,\Delta i^2\,\Delta e_y^2}}\right] \tag{71}$$

$$\lambda_4 = -\frac{K\,\Delta e_y}{2 a_o u}\left[\frac{2\,\Delta e_x\,\Delta i}{\sqrt{\left(\Delta i^2 - \Delta e_x^2 - \Delta e_y^2 + \dfrac{\Delta a^2}{a_o^2}\right)^2 + 4\,\Delta i^2\,\Delta e_y^2}}\right] \tag{72}$$

$$\lambda_5 = \frac{K\,\Delta i}{2 a_o u}\left[1 + \frac{\Delta i^2 - \Delta e_x^2 + \Delta e_y^2 + \dfrac{\Delta a^2}{a_o^2}}{\sqrt{\left(\Delta i^2 - \Delta e_x^2 - \Delta e_y^2 + \dfrac{\Delta a^2}{a_o^2}\right)^2 + 4\,\Delta i^2\,\Delta e_y^2}}\right] \tag{73}$$

The orientation of the primer locus relative to the line of nodes can be found by use of Eq. (14). The location of the impulses can be found from Eq. (21) and the direction of the impulses from Eqs. (8)-(10). The impulse split between the two impulses can be found from Eq. (62) or from one of the boundary conditions.

The payoff for the singular case is independent of the semi-major axis, as it is for the nodal case, because the partial derivative of the payoff with respect to the semi-major axis, λ_1, is equal to zero. The required value of change in the semi-major axis is easily determined by setting λ_1 equal to zero in the nondegenerate solution for λ_1, Eq. (69).

$$\Delta a^2 = \Delta e_x^2 + \Delta e_y^2 + \frac{2}{\sqrt{3}} \Delta i \, \Delta e_y - \Delta i^2 \tag{74}$$

For changes in the semi-major axis smaller than those given by Eq. (74), the solution will be singular (or nodal) and will have the same payoff as the nondegenerate solution with the semi-major axis given by Eq. (74).

$$u = \sqrt{\frac{K}{4 a_o}} \sqrt{\left(\sqrt{3} \; \Delta i + \Delta e_y\right)^2 + \Delta e_x^2} \tag{75}$$

The orientation of the singular locus may also be found as a special case of the nondegenerate case.

$$\tan \theta_o = \frac{\Delta e_y + \sqrt{3} \; \Delta i}{\Delta e_x} \tag{76}$$

The optimum thrust directions are given by Eqs. (27)-(29).

In this singular case, the solution space collapses from five-dimensional to three-dimensional and may be visualized in three-space. The simplest way to show this is to once again align the x axis with the θ_o direction of the primer locus. The rates of change of the elements are then given by Eqs. (77)-(81).

$$\frac{dx_1}{du} = 2C \tag{77}$$

$$\frac{dx_2}{du} = \frac{3}{2} S C \tag{78}$$

$$\frac{dx_3}{du} = 2 - \frac{3}{2} S^2 \tag{79}$$

$$\frac{dx_4}{du} = \frac{\sqrt{3}}{2} S^2 \tag{80}$$

$$\frac{dx_5}{du} = \frac{\sqrt{3}}{2} S C \tag{81}$$

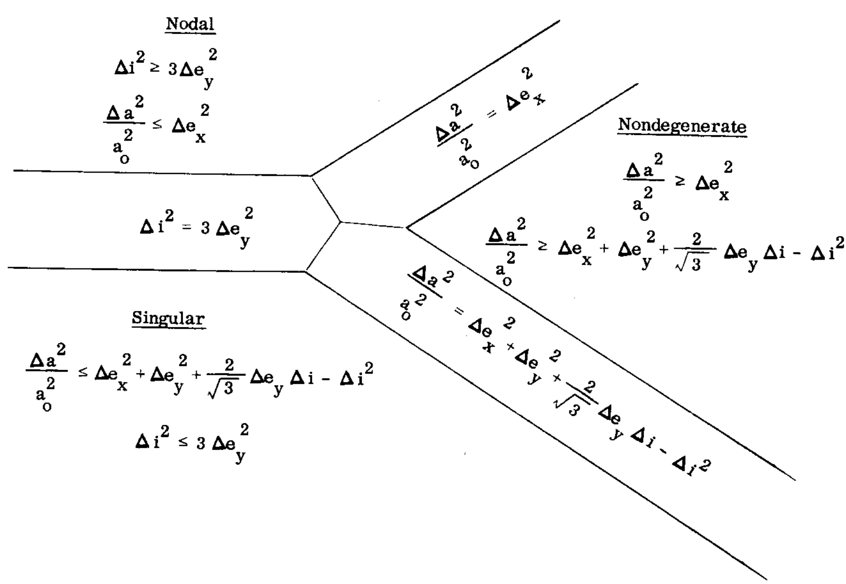
It is seen that x_2 is linearly dependent on x_5 and that x_3 is linearly dependent on x_4. Figure 5 is a plot of Eqs. (77), (80) and (81) showing the possible changes in the elements as a function of the position of the impulse. If only one impulse is allowed, the reachable states lie on the line of intersection of a circular cylinder and a parabolic cylinder which form the convex hull of the intersection. It is possible to reach points on the convex hull by using two impulses (see e.g. Ref. 8). The interesting question is whether points in the interior of the volume can also be reached with two impulses. If they can, then all minimum impulse transfers between orbits in the near vicinity of a circular orbit can be realized with only two impulses.

A proof that every interior point of the convex hull is reachable with two impulses has been suggested by W. D. Hayes and S. H. Lam of Princeton University. The geometric interpretation is that a straight line which touches the space curve at two points can be passed through every interior point of the volume. The proof is constructed by drawing a unit sphere about an arbitrary interior point and projecting the space curve onto the sphere. The antipodal curve to this curve on the sphere is also drawn. The geometry of the problem is such that the curve and its antipodal curve will always intersect. As the intersection is sufficient for a straight line which touches the curve at two points to pass through the center of the sphere, two-impulse solutions are always possible. Even these two-impulse solutions are not unique, as there are usually at least two intersections of the two projected curves.

The explicit calculation of the two-impulse solutions is difficult and has not been accomplished. However, three-impulse solutions are easily calculated explicitly. For example, the same value of $\sin^2(\theta - \theta_o)$ may be used for each impulse and the impulse splits adjusted to meet the other boundary conditions.

Because these solutions are singular, they are also realizable with finite thrusts. Any transfer with the optimum thrust directions that meets the boundary conditions will realize the minimum fuel consumption.

The type of solution corresponding to any particular transfer can be found from the following diagram.



RESULTS

The total impulse required for any transfer as well as the transfer regions are illustrated in Figs. 6-11. In each figure, the contours are of constant values of $\Delta i/u$ going from zero at the outer boundary to unity at the origin in increments of 0.05. Each figure is drawn for a different value of the angle between the eccentricity change vector and the inclination change vector. Figure 6 corresponds to co-axial transfers where the solution is known to consist of inclined Hohmann transfers. There is no singular solution in this case. One set of Hohmann transfers is a special case of the nodal transfers while the other is a special case of the nondegenerate transfers. This becomes clearer in Fig. 7 where the singular solution appears near the outer boundary of the figure. The nodal case is the triangular region adjoining the singular region. The nondegenerate case occupies the rest of the figure.

As the angle between the eccentricity change and the inclination change increases, the nodal region rapidly decreases in size while the singular region grows. Finally, in Fig. 11 the nodal region has completely disappeared.

SUGGESTIONS FOR FURTHER RESEARCH

The solutions obtained herein are quite general and contain a number of new results. They naturally suggest a number of possibilities for further research, particularly for nonlinear transfers. The following paragraphs will briefly describe some of these research areas.

1. The existence of more general nodal transfers than the Hohmann transfer in this problem suggests a general investigation of nodal transfers.

2. The singular solution obtained herein may be the limit point of a nonlinear singular solution like the Lawden spiral. Unlike the Lawden spiral, half of this curve satisfies the necessary condition derived by Kelley, Kopp and Moyer (Ref. 11) and by Robbins (Ref. 12).

3. Either or both of the infinitesimal impulses of the present solutions may be allowed to become finite. The resulting nonlinear transfers may be investigated by the methods of Moyer (Ref. 13) or Winn (Ref. 14).

4. The consideration of higher order terms will remove the degeneracy of the singular solution and will introduce three-impulse solutions (if not even more complex ones). The next step is the consideration of quadratic terms.

5. The terminal guidance problem can be considered since the sensitivity of the transfer to various errors can be analytically determined. Finite thrust guidance might be approached as Robbins does in Ref. 15.

REFERENCES

1. Gobetz, F.W.: "A Linear Theory of Optimum Low-Thrust Rendezvous Trajectories," Journal of the Astronautical Sciences, Vol. 12, No. 3, Fall 1965, pp. 69-76.

2. Edelbaum, T.N.: "Optimum Low-Thrust Rendezvous and Stationkeeping," AIAA Journal, Vol. 2, No. 7, July 1964, pp. 1196-1201.

3. Marchal, C.: "Transferts Optimaux Entre Orbites Elliptiques Coplanaires (Duree Indifferente)," Astronautica Acta, Vol. 11, No. 6, Nov.-Dec. 1965.

4. Gobetz, F.W., Washington, M. and Edelbaum, T.N.: "Minimum-Impulse Time-Free Transfer Between Elliptic Orbits," NASA CR-636, Nov. 1966.

5. Marec, J.P.: "Transferts Infinitesimaux Impulsionnels Economiques entre Orbites Quasi-Circulaires non Coplanaires," presented at 17th IAF Congress, Madrid, October 1966.

6. Edelbaum, T.N.: "Optimization Problems in Powered Space Flight," Recent Developments in Space Flight Mechanics, Vol. 9, Science and Technology Series, American Astronautical Society, Tarzana, Calif., 1966.

7. Edelbaum, T.N.: "How Many Impulses?" AIAA Paper No. 66-7, January 1966. Revised and updated version to appear in Astronautics and Aeronautics.

8. Contensou, P.: "Etude Theorique des Trajectories Optimaler dans un Champ de Gravitation. Application au Cas d'un Centre d'Attraction Unique," Astronautica Acta, Vol. 8, No. 2 & 3, 1962.

9. Breakwell, J. V.: "Minimum Impulse Transfer," AIAA Progress in Astronautics Series, Vol. 14, 1963.

10. Lawden, D. F.: Optimal Trajectories for Space Navigation, Butterworths, London, 1963.

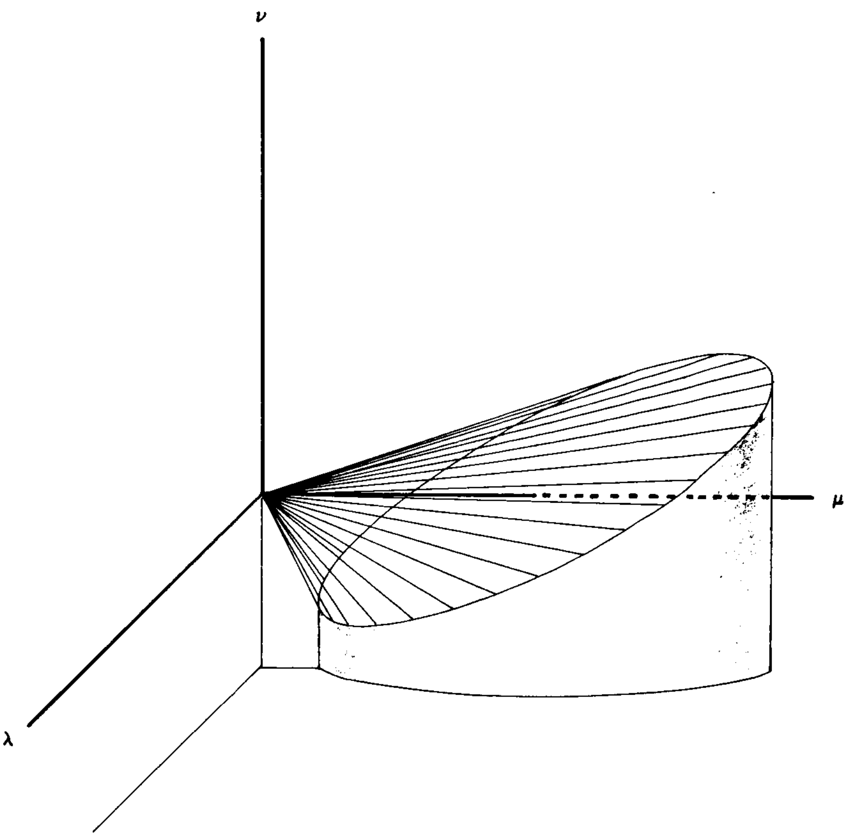
11. Kelley, H. J., Kopp, R. E. and Moyer, H. G.: "Singular Extremals," Chapter 3 of Optimization Theory and Applications, G. Leitmann, Editor, Academic Press, New York, to appear.

12. Robbins, H. M.: "Optimality of Intermediate Thrust Arcs of Rocket Trajectories," AIAA Journal, Vol. 3, pp. 1094-1098, 1965.

13. Moyer, H. G.: "Necessary Conditions for Optimal Single Impulse Transfer," AIAA Journal, Vol. 4, pp. 1405-1410, 1966.

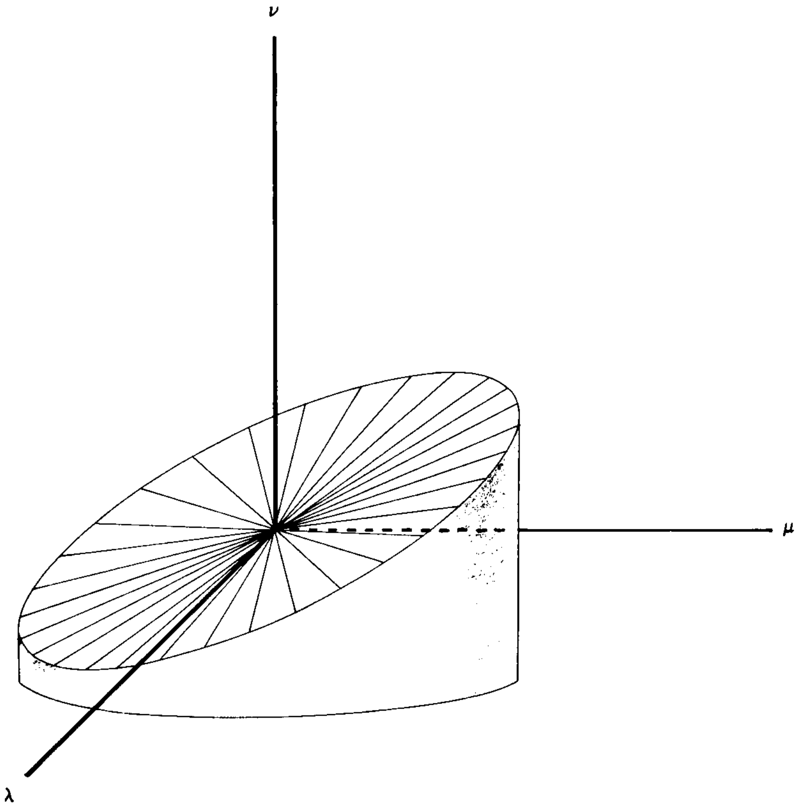
14. Winn, C. B.: "Minimum Fuel Transfer Between Co-Axial Orbits, Both Coplanar and Noncoplanar," AAS Preprint 66-119, July 1966.

15. Robbins, H. M.: "An Analytical Study of the Impulsive Approximation," AIAA Journal, Vol. 4, pp. 1417-1423, 1966.



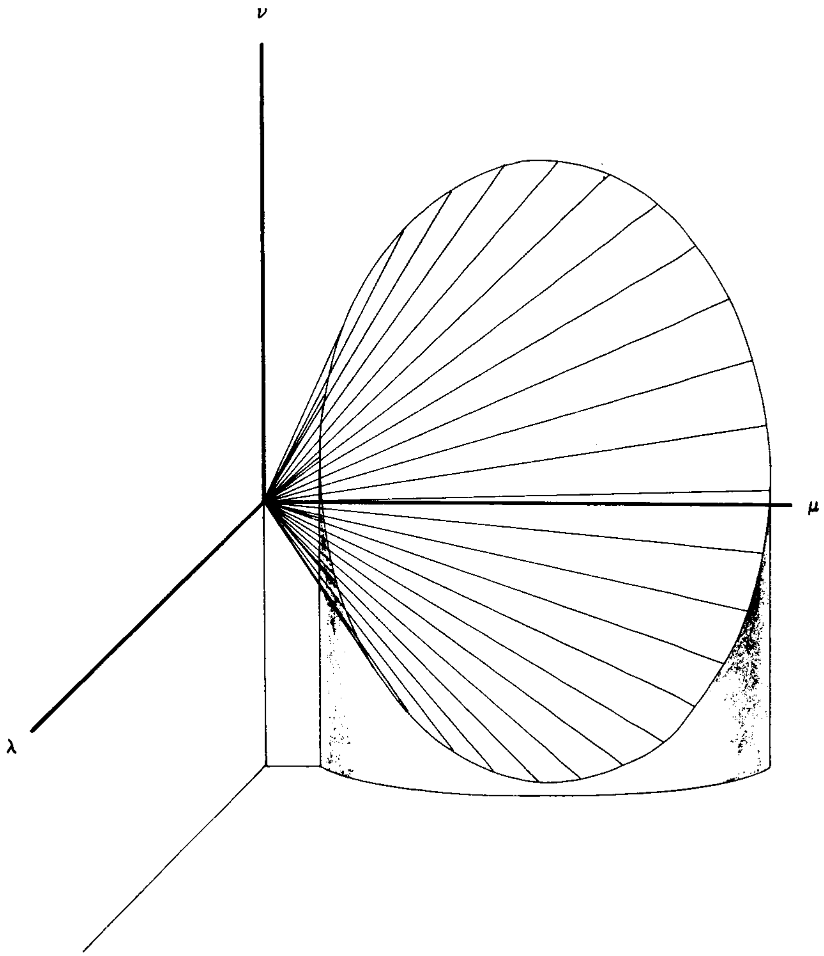
PRIMER LOCUS DIAGRAM

Figure 1



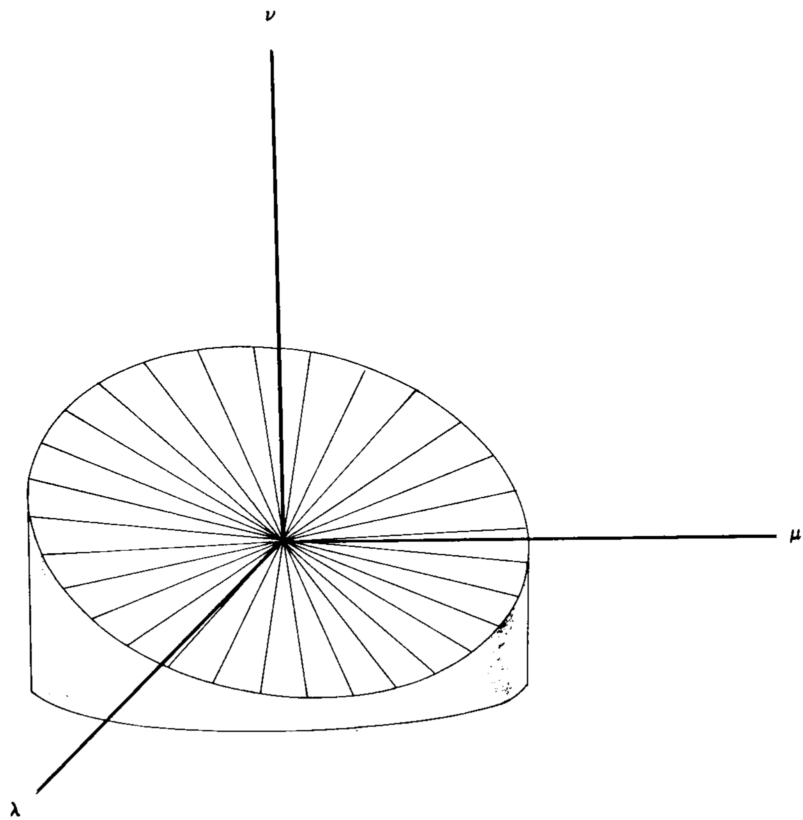
PRIMER LOCUS DIAGRAM

Figure 2



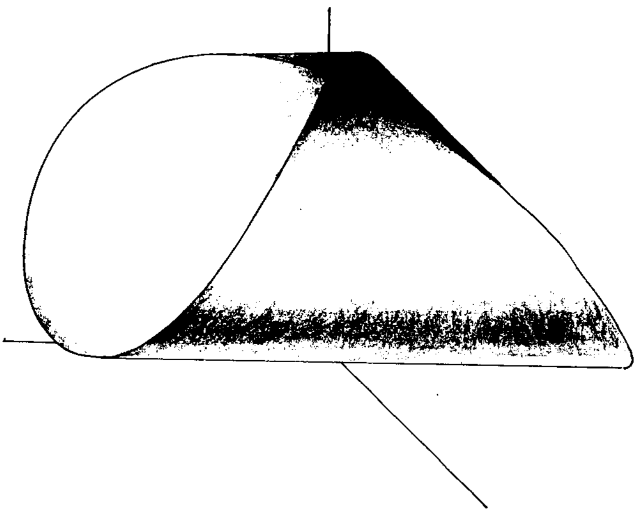
PRIMER LOCUS DIAGRAM

Figure 3



PRIMER LOCUS DIAGRAM

Figure 4



REACHABLE STATES - - - SINGULAR CASE

Figure 5

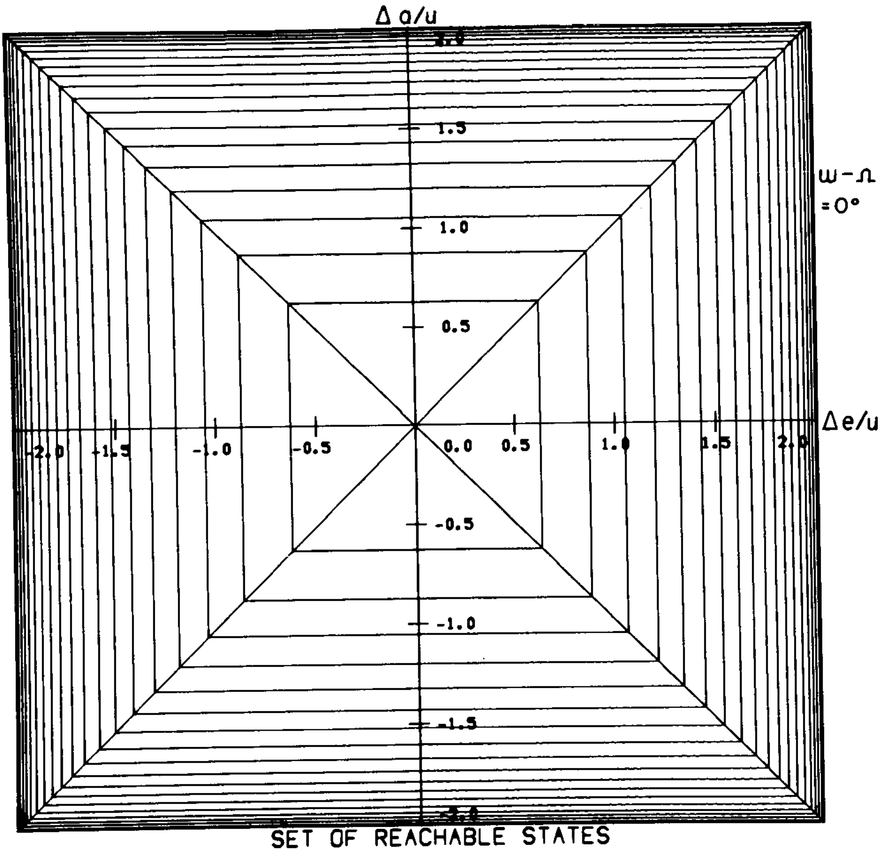

Δa/u
1.5
1.0
0.5
-0.5
-1.0
-1.5
ω-Ω
=0°
Δe/u
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
SET OF REACHABLE STATES

Figure 6

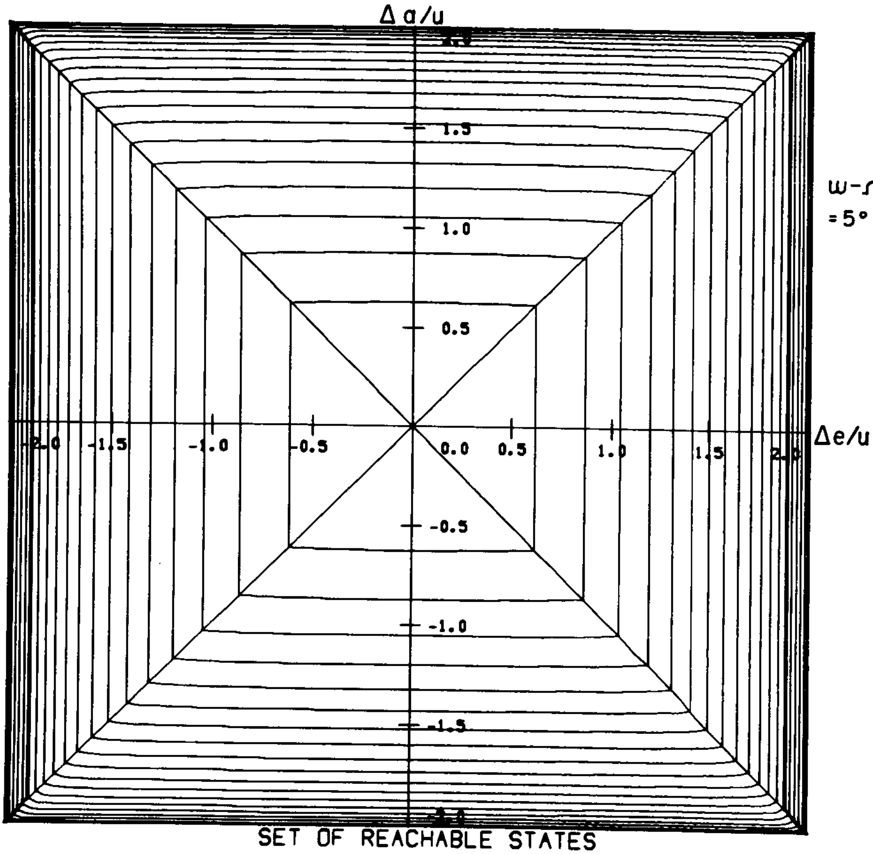

Δa/u
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
Δe/u
-2.0
-1.5
-1.0
-0.5
0.5
1.0
1.5
2.0
ω−Ω
=5°
SET OF REACHABLE STATES

Figure 7

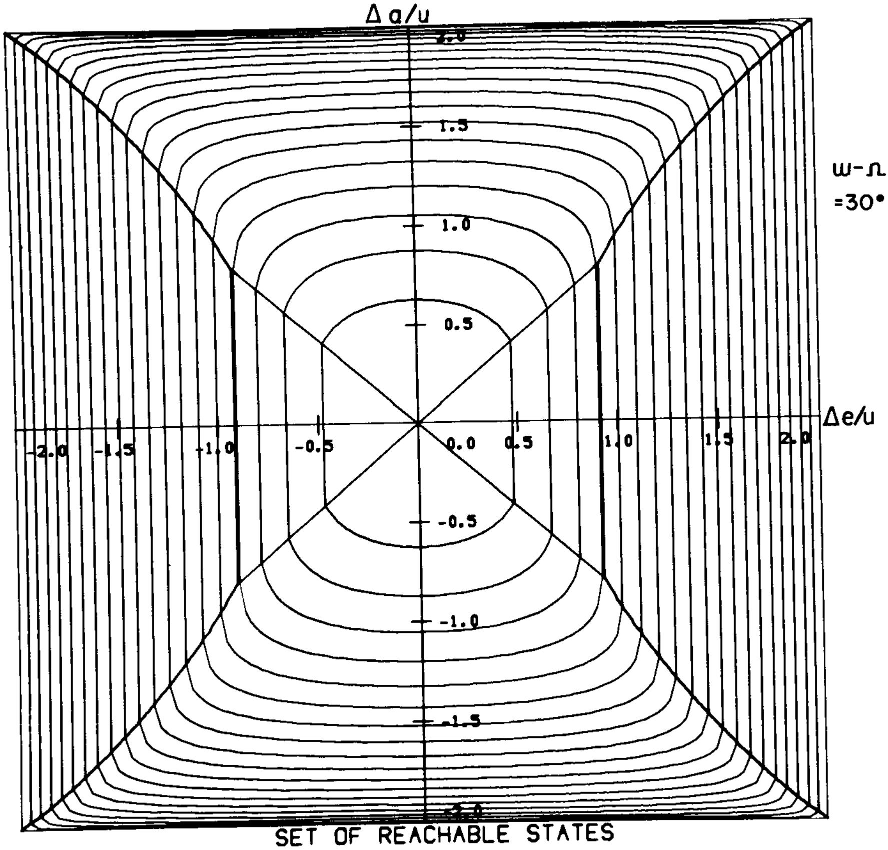

Δ a/u
2.0
1.5
1.0
0.5
-0.5
-1.0
-1.5
-2.0
Δe/u
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
ω-Ω
=30°
SET OF REACHABLE STATES

Figure 8

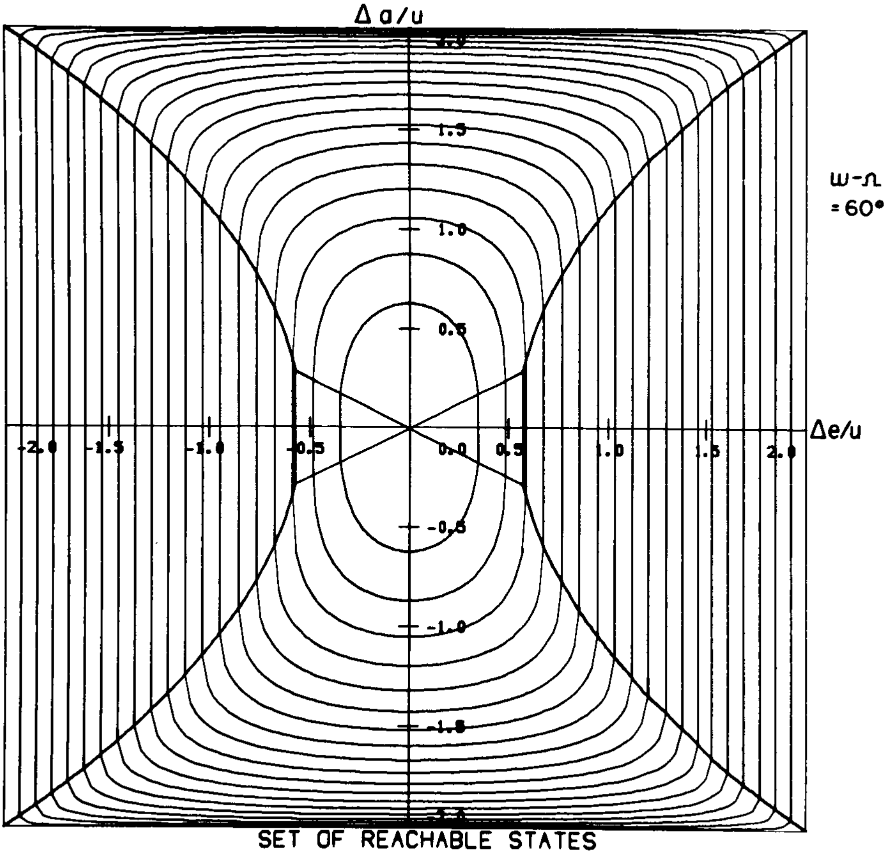
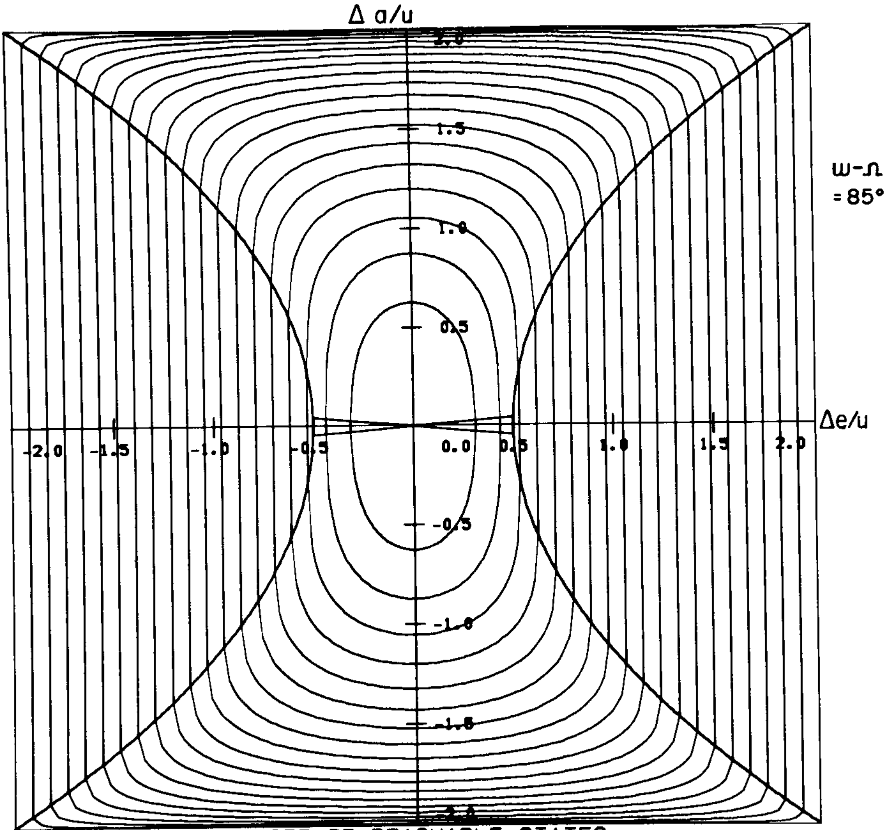

Δ a/u
1.5
1.0
0.5
-0.5
-1.0
-1.5
Δe/u
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
ω-Ω
=60°
SET OF REACHABLE STATES

Figure 9



SET OF REACHABLE STATES

Figure 10

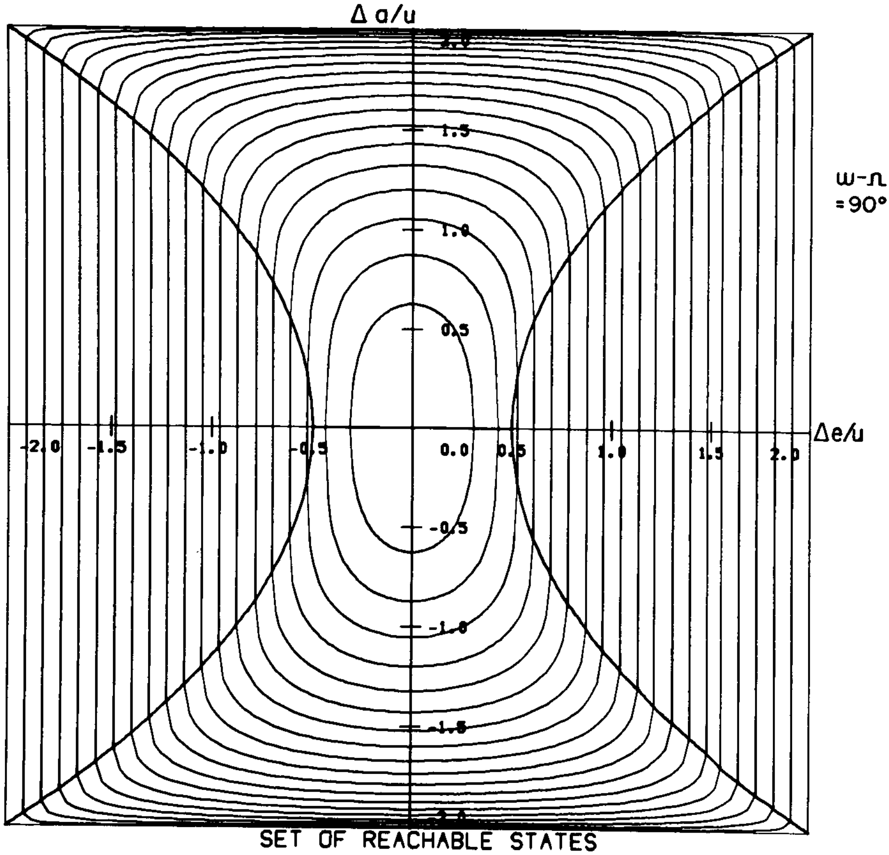

Δ a/u
2.0
1.5
1.0
0.5
-0.5
-1.0
-1.5
-2.0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Δe/u
ω-Ω
=90°
SET OF REACHABLE STATES

Figure 11

PRECEDING PAGE BLANK NOT FILMED.

Rejection to Infinity in the Problem of Three Bodies When the Total Energy Is Negative*

by D. C. Lewis
Control Research Associates
Baltimore, Maryland

1. Introduction

N67-29374

In the problem of three bodies, we take the kinetic energy relative to the center of mass of the three bodies and we take the zero level of potential energy as occurring when the three bodies are infinitely far from each other. This makes the potential energy always negative for any actual configuration, and the total energy must then be negative also provided that the velocities of the three bodies relative to the mass-center are sufficiently small for the given configuration. In this context we consider motions for which the total energy has a preassigned negative value $-K (K > 0)$ and for which the length of the angular momentum vector (taken relative to the mass center) also has a preassigned value $f > 0$.

---

*Performed under contract NAS 12-93.

A useful quantity for measuring the dispersion of the three bodies is the so-called Lagrangian inertial radius R relative to the mass center. Its definition will be deferred to the next section. We merely note here that R becomes infinitely large if and only if the perimeter of the triangle formed by the three masses becomes infinitely large, and R is small if and only if the perimeter is small.

G. D. Birkhoff has shown that if R is at any time t_0 less than a sufficiently small positive quantity R_0, then one of the three bodies will recede infinitely far from the other two as $(t - t_0) \to \infty$, while the distance between the latter two will remain bounded.

It is clear from his work that R_0 depends on the three masses m_0, m_1, m_2 as well as on the assigned values of K and f, so that

$$R_0 = R_0(K, f, m_0, m_1, m_2) \; .$$

But he did not give an explicit estimate of R_0. The main purpose of this paper is to calculate such an estimate. The final result is formulated with some precision in Theorem 10.

We also wish to clear up some obscurities, if not actual errors, in Birkhoff's treatment. For instance, Birkhoff never proved the statement about upward concavity of the part of the curve $R = R(t)$ for which $R < f/(2^{1/2} K^{1/2})$, cf. Reference [1] p. 278, ℓ. 1. He only proved that the second derivative was positive at points where the curve was horizontal. For such reasons we have discussed this theorem in much greater detail than was done by Birkhoff. Our version of this part of his work appears as Theorem 3.

We have ignored the possibility of collisions in this paper. Since $f > 0$, the case of a simultaneous collision of all three bodies is ruled out by the conjecture of Weierstrass, which was rigorously proved by Sundman. The remaining difficulties afforded by the possible occurrences of binary collisions are readily eliminated with the help of a regularizing parametrization due also to Sundman.

2. Notational Prelude.

The nine second order differential equations for the three body problem may be written in the form of three vector equations

$$m_i \frac{d^2 q_i}{dt^2} = \frac{\partial U}{\partial q_i}, \qquad i = 0, 1, 2, \tag{2.1}$$

where q_i is the vector with components x_i, y_i, z_i. The mass m_i is thus regarded as having rectangular coordinates x_i, y_i, z_i in an inertial frame of reference. The (scalar) force function is

$$U = \frac{m_1 m_2}{r_0} + \frac{m_2 m_0}{r_1} + \frac{m_0 m_1}{r_2}$$

where r_i is the distance between the masses m_j and m_k. Here (i, j, k) is a permutation of $(0, 1, 2)$, and $r_i = |q_j - q_k| = [(x_j - x_k)^2 + (y_j - y_k)^2 + (z_j - z_k)^2]^{\frac{1}{2}}$.

The equations (2.1) are well known to admit a total of ten elementary first integrals, the first six of which express the conservation of linear momentum and permit us to use a frame of reference whose origin is at the center of gravity of the three bodies. We may thus assume that

$$m_0 q_0 + m_1 q_1 + m_2 q_2 \equiv 0. \tag{2.2}$$

The energy integral is written in the form

$$\sum_i \frac{1}{2} m_i \, |\dot{q}_i|^2 = U - K ,$$

where the dot is used to denote differentiation with respect to t and where the integration constant K is the <u>negative</u> of the total energy.

The three integrals of angular momentum appear as a single vector equation

$$\sum_i m_i (q_i \times \dot{q}_i) = c .$$

The length of the vector c of total angular momentum will be denoted in the sequel by f.

The Lagrangian inertial radius R may be defined by the equation

$$R^2 = \sum_i m_i \, |q_i|^2 ,$$

but, if (2.2) is assumed, it is not hard to see that we also have

$$R^2 = M^{-1}(m_0 m_1 r_2^2 + m_1 m_2 r_0^2 + m_2 m_0 r_1^2) ,$$

where M, as in (2.12), is the sum of the masses.

Lagrange's well known identity is to the effect that

$$\frac{d^2R^2}{dt^2} = 2(U - 2K) . \tag{2.3}$$

We refer to the following as Sundman's inequality

$$\dot{R}^2 + 2R\ddot{R} + 2K \geq f^2 R^{-2} \tag{2.4}$$

which is well known. The auxiliary function of Sundman is defined by the equation

(2.5) $$H = R\,\dot{R}^2 + 2\,KR + f^2\,R^{-1}\;.$$

Its derivative with respect to t is seen to be of the form $F\,\dot{R}$, where $F = \dot{R}^2 + 2R\,\ddot{R} + 2K - f^2\,R^{-2} \geq 0$ because of (2.4). We thus note the important fact that, on any interval for t on which R is monotonic, H is also monotonic in the same sense.

The formulas thus far used have the advantage of being symmetric in m_0, m_1, m_2 and q_0, q_1, q_2. In other words they are invariant under any permutation of the subscripts $(0, 1, 2)$. On the other hand they have the disadvantage of not reflecting the full potentiality of (2.2) for reducing the number of unknowns. One way of accomplishing this is due to Lagrange, but a single reduction of this sort necessarily sacrifices the desirable symmetry noted above. Consequently we contemplate the whole class of such reductions in the following way.

Let (i, j, k) be any permutation of $(0, 1, 2)$. Let the vector q, with components (x, y, z), determine the position of the mass m_j relative to m_i, so that

(2.6) $$q = q_j - q_i\;.$$

The center of gravity of m_i and m_j, relative to the origin of the original frame of reference, is evidently at the point corresponding to the vector $\alpha_i q_i + \alpha_j q_j$, where $\alpha_i = m_i(m_i + m_j)^{-1}$ and $\alpha_j = m_j(m_i + m_j)^{-1}$.

We now use a vector s with components (ξ, η, ζ) to determine the position of the third body m_k relative to the center of gravity of m_i and m_j, so that

$$s = q_k - \alpha_i q_i - \alpha_j q_j \tag{2.7}$$

Assuming that the origin of the original frame is at the center of gravity of all three bodies, so that $m_i q_i + m_j q_j + m_k q_k = 0$, which is simply another way of writing (2.2), we may easily find q_i, q_j, and q_k in terms of q and s. In fact we have

$$\begin{aligned} q_i &= -\alpha_j q - m_k M^{-1} s \\ q_j &= +\alpha_i q - m_k M^{-1} s \\ q_k &= (m_i + m_j) M^{-1} s \end{aligned} \tag{2.8}$$

where M, as previously, denotes the sum of the three masses.

If we set

$$m = \frac{m_i m_j}{m_i + m_j} \quad \text{and} \quad \mu = \frac{(m_i + m_j)\, m_k}{M} \tag{2.9}$$

it may be shown that the equations (2.1) are equivalent, with due regard to (2.2), to the equations

$$m\,\ddot{q} = \frac{\partial U}{\partial q} \quad , \quad \mu\,\ddot{s} = \frac{\partial U}{\partial s} \quad .$$

The energy integral may be written in the form

$$\frac{1}{2} m|\dot{q}|^2 + \frac{1}{2}\mu|\dot{s}|^2 = U - K \tag{2.10}$$

and the Lagrangian inertial radius R, previously introduced, satisfies the equation

$$R^2 = m r^2 + \mu \rho^2 , \tag{2.11}$$

where $r = |q| = \sqrt{x^2 + y^2 + z^2}$ and $\rho = |s| = \sqrt{\xi^2 + \eta^2 + \zeta^2}$.

Most of the above formulas, beginning with (2.6), depend heavily upon the particular choice of the permutation (i, j, k) of $(0, 1, 2)$. It will be essential to have upper and lower bounds for both m and μ which are independent of this permutation. We will, in fact, calculate such bounds in terms of the following four symmetric functions of m_0, m_1, and m_2.

$$\begin{aligned} M &= m_0 + m_1 + m_2 \\ P &= m_0 m_1 m_2 \\ m^* &= \text{minimum of the three masses .} \\ \overline{m} &= \text{greater of the two smallest masses .} \end{aligned} \tag{2.12}$$

From (2.9) we find that $dm/dm_i = m_j^2(m_i + m_j)^{-2} > 0$. Hence m is an increasing function of m_i when m_j is held fixed. On the interval $m_j \leq m_i < +\infty$, it takes on its minimum at $m_i = m_j$, at which point $m = \frac{1}{2} m_j \geq \frac{1}{2} m^*$; and as $m_i \to \infty$, m tends to its least upper bound m_j, and, since under present conditions $m_i \geq m_j$, m_j cannot exceed the greater of the two smallest masses, which we denote by $\overline{m}$. We conclude therefore that

(2.13) $$\frac{1}{2} m^* \leq m < \bar{m}$$

if $m_i \geq m_j$. But, since m is symmetric in m_i and m_j (and so are m^* and $\bar{m}$), we see that (2.13) holds also even if $m_i < m_j$. It is seen from (2.9) and (2.12) that $\mu = P\, m^{-1} M^{-1}$. Hence it follows at once from (2.13), that

(2.14) $$\underline{\mu} < \mu \leq \bar{\mu} \, ,$$

where

(2.15) $$\bar{\mu} = \frac{2P}{m^* M} \quad \text{and} \quad \underline{\mu} = \frac{P}{\bar{m} M} \, .$$

3. Fundamental Results.

Theorem 1. In case $K > 0$, the least of the three mutual distances, r_0, r_1, r_2 can not exceed $M^2/(3K)$.

Proof. From the energy integral it is clear that

$$K \leq U = \frac{m_0 m_1}{r_2} + \frac{m_0 m_2}{r_1} + \frac{m_1 m_2}{r_0} \leq \frac{m_0 m_1 + m_0 m_2 + m_1 m_2}{r} \, ,$$

where $r = \min(r_0, r_1, r_2)$. But $M^2 = (m_0 + m_1 + m_2)^2$ $= \frac{1}{2}(m_0^2 + m_1^2) + \frac{1}{2}(m_0^2 + m_2^2) + \frac{1}{2}(m_1^2 + m_2^2) + 2(m_0 m_1 + m_0 m_2 + m_1 m_2)$ $\geq 3(m_0 m_1 + m_0 m_2 + m_1 m_2)$. Thus we have

$$K \leq \frac{m_0 m_1 + m_0 m_2 + m_1 m_2}{r} \leq \frac{M^2}{3r} \, ,$$

so that

$$r \leq \frac{M^2}{3K}$$

as desired.

We continue throughout the rest of the paper to assume that $K > 0$. Moreover, as in the proof of the next theorem, we shall always use r to denote the least of the three mutual distances and we shall use ρ to denote the distance from the center of gravity of the two mutually closest bodies to the third body. This notation is consistent with the notation of (2.9) and (2.11), since we merely have to take m_i and m_j as the two closest bodies.

<u>Theorem 2.</u> If

$$R \geqq \frac{M^2}{3K} \sqrt{\overline{m} + 4 \overline{\mu}} ,$$

then $\rho \geqq \frac{2}{3} M^2 K^{-1}$.

<u>Proof.</u> Evidently, by (2.13) and (2.14),

$$R \geqq \frac{M^2}{3K} \sqrt{m + 4 \mu} .$$

Since $R^2 = m r^2 + \mu \rho^2$, we have, from Theorem 1,

$$\rho^2 = \frac{1}{\mu} R^2 - \frac{m}{\mu} r^2 \geqq \frac{M^4}{9 K^2 \mu} (m + 4 \mu) - \frac{m}{\mu} \frac{M^4}{9 K^2} = \frac{4 M^4}{9 K^2} ,$$

so that

$$\rho \geqq \frac{2 M^2}{3 K}$$

as stated.

In the sequel it is convenient to refer several times to the following elementary lemmas. Although they are essentially well known, it is more convenient to give the proofs here than to refer to the pertinent literature.

Lemma 1. Suppose that $F(t)$ is differentiable for $t > a$ and that $\lim_{t\to\infty} F(t) = b$, where a and b are constants. Then there exists a sequence $t_1, t_2, \ldots$ such that $\lim_{n\to\infty} t_n = \infty$ and $\lim_{n\to\infty} F'(t_n) = 0$.

Proof. Assuming $n > a$, we may define t_n, by the mean value theorem, to be such that $n < t_n < n + 1$ and $F'(t_n) = F(n + 1) - F(n)$, and the lemma follows at once since $\lim_{n\to\infty} [F(n + 1) - F(n)] = b - b = 0$.

Lemma 2. Let $F(t, P)$ be a continuous real valued function of its two real arguments t, P; and let it satisfy a local Lipschitz condition in P for $t_0 \leq t$ and $P > 0$. Let $R(t)$ be a positive differentiable function of t, defined for $t \geq t_0$, such that

$$\text{(I)} \qquad R'(t) \leq F[t, R(t)]$$

Let $Q(t)$ be positive and satisfy the differential equation

$$Q'(t) = F[t, Q(t)]$$

on the interval $t_0 \leq t \leq t_1$ and suppose also that $Q(t_0) = R(t_0)$. Then $Q(t) \geq R(t)$ for $t_0 \leq t \leq t_1$.

Proof. Let t_2 be any real number between t_0 and t_1. Then by the existence and continuity theorems for differential equations, we may define a function $P(t) = P(t, \varepsilon)$ which satisfies the differential equation

$$\text{(II)} \qquad P'(t) = F[t, P(t)] + \varepsilon ,$$

and the initial condition

$$P(t_0) = R(t_0) ,$$

this definition being valid for $t_0 \leq t \leq t_2$, at least, if $|\varepsilon|$ is sufficiently small, depending perhaps on the choice of t_2. Moreover

(III) $$\lim_{\varepsilon \to 0} P(t, \varepsilon) = Q(t)$$

on the interval $t_0 \leq t \leq t_2$.

We first prove that, if ε is positive, then

(IV) $$P(t, \varepsilon) > R(t) \qquad \text{for } t_0 < t \leq t_2 .$$

Suppose that this were not so. Then there would exist a point t^*, such that $t_0 < t^* \leq t_2$, where the function $w(t) = P(t, \varepsilon) - R(t)$ is not positive. Also, since $w(t_0) = P(t_0, \varepsilon) - R(t_0) = 0$ and

(V) $$w'(t_0) = P'(t_0, \varepsilon) - R'(t_0) \geq [F(t_0, P(t_0)) + \varepsilon] - F(t_0, R(t_0)) = \varepsilon > 0$$

there is a point $t^{**} > t_0$, but not as great as t^*, such that $w(t) > 0$ on the interval $t_0 < t \leq t^{**}$. For otherwise we would have $w(t)$ non-positive at an infinite number of points in any neighborhood of t_0, necessitating $w'(t_0) \leq 0$, contrary to (V). Thus, since $w(t^{**}) > 0$ and $w(t^*) \leq 0$, there is a non-empty set S of points on the interval $t^{**} < t \leq t_2$ where w vanishes. Since w is continuous, S is closed. Let $\bar{t}$ = the greatest lower bound of S. Then $t_0 < t^{**} < \bar{t} \leq t_2$ and $w(\bar{t}) = 0$. That is,

(VI) $$P(\bar{t}, \varepsilon) = R(\bar{t}) .$$

But $w(t) > 0$ for $t_0 < t < \bar{t}$. Since w is differentiable at $\bar{t}$, it follows that $w'(\overline{t}) \leqq 0$, which contracts the fact, following from (VI), that

$w'(\overline{t}) = P'(\overline{t},\epsilon) - R'(\overline{t}) \geqq [F(\overline{t}, P(\overline{t})) + \varepsilon] - F(\overline{t}, R(\overline{t})) = \varepsilon > 0$.

This finishes the proof of (IV).

By letting $\varepsilon \to 0$ in (IV), we find from (III), that $Q(t) \geqq R(t)$ for $r_0 \leqq t \leqq t_2$. But, since t_2 is arbitrary on the open interval from t_0 to t_1 and since both $Q(t)$ and $R(t)$ are continuous, the above inequality remains valid on the whole closed interval from t_0 to t_1, as we wished to prove.

Theorem 3. In case $f > 0$, $K > 0$, any connected part C of the curve $R = R(t)$, for which $R < f(2K)^{-1/2}$, consists of an arc with a single minimum. If $R = R_0$ at this minimum and if θ is any number between 0 and 1, the curve rises on either side until $R > \theta f^2/(2K R_0)$ with slope $\dot{R} = (d R/dt)$ at least as great in absolute value as demanded by the inequality

$$\dot{R}^2 \geqq \frac{R - R_0}{R} \left(\frac{f^2}{R_0 R} - 2K\right)$$

at every intermediate stage. It is hereby implied that $R_0 > 0$.

Proof. If $\ddot{R}(t) \leqq 0$ and $\dot{R}(t) = 0$, the Sundman inequality, (2.4), $\dot{R}^2 + 2R\ddot{R} + 2K \geqq f^2 R^{-2}$, shows that $R(t) \geqq f(2K)^{-1/2}$. We thus see that, as long as we confine attention to C (for which $R(t) < f(2K)^{-1/2}$), we must have $\ddot{R} > 0$ at all points where $\dot{R} = 0$. This means in particular that R cannot be a constant on C or any subarc of C. It also means that $\dot{R}$ can vanish at not more than one point of C. We proceed to show that such a point actually exists.

Otherwise, since $\dot{R}(t)$ is continuous and never zero, $R(t)$ would be strictly monotonic (in one sense or the other) on C, and hence the Sundman function

$$H = R\,\dot{R}^2 + 2K\,R + f^2\,R^{-1}$$

must be monotonic in the same sense. This shows that R cannot decrease monotonically to 0 for either increasing or decreasing t, tending to either a finite or an infinite limit, for this would make H tend to $+\infty$, so that H could not possibly be monotonic in the same sense as R. Hence let $R_0 > 0$ be the greatest lower bound of $R(t)$ on C. Then, since $R(t)$ is presently assumed to be strictly monotonic on C, we must have either

$$\lim_{t\to\infty} R(t) = R_0 > 0 \tag{3.1}$$

or else

$$\lim_{t\to-\infty} R(t) = R_0 > 0\,, \tag{3.1*}$$

and since the differential equations are invariant under change of sign of t, it will be sufficient to confine attention to (3.1). We wish to show that (3.1) leads to a contradiction. By Lemma 1 there is an infinite sequence $t_1, t_2, \ldots$ tending to ∞, such that $\lim_{n\to\infty} \dot{R}(t_n) = 0$. Hence, from (2.5), we have $\lim_{n\to\infty} H(t_n) = 2K\,R_0 + f^2\,R_0^{-1}$. But, since H is monotonic in the same sense as R and is never negative, we know that $\lim_{t\to\infty} H(t)$ exists.

It follows therefore that $\lim_{t\to\infty} H(t) = 2K R_0 + f^2 R_0^{-1}$. Moreover $H(t) \geq 2K R_0 + f^2 R_0^{-1}$ for every finite t, since this corresponds to the monotonic sense in which H tends to its limit. From the definition (2.5) of $H(t)$, we find therefore that

$$R \dot{R}^2 + 2K R + f^2 R^{-1} \geq 2K R_0 + f^2 R_0^{-1}$$

which is equivalent to

$$\dot{R}^2 \geq \left(\frac{R - R_0}{R}\right) \left(\frac{f^2}{R_0 R} - 2K\right) . \tag{3.2}$$

And, since $R(t)$ is presently assumed to decrease monotonically to R_0 as $t \to \infty$, we also have

$$\dot{R}(t) < 0 \tag{3.3}$$

on C.

We wish to show that (3.1), (3.2), and (3.3) are mutually inconsistant. To do this we take any fixed t^* corresponding to a point $(t^*, R(t^*))$ on C and define a function $Q(t)$ by means of the differential equation,

$$\dot{Q}(t) = - \left[\frac{Q(t) - R_0}{Q(t)}\right]^{\frac{1}{2}} \left[\frac{f^2}{R_0 Q(t)} - 2K\right]^{\frac{1}{2}} \tag{3.4}$$

and the initial condition $Q(t^*) = R(t^*)$. Such a definition can be effected by the usual existence theorems for differential equations at least for some sufficiently short interval $t^* \leq t < t^{**}$. It follows from (3.2), (3.3), and (3.4) that $0 > \dot{Q}(t) \geq \dot{R}(t)$ whenever $Q(t) = R(t) > R_0$, in particular when $t = t^*$. Hence by Lemma 2 the two curves can never cross for $t > t^*$ and moreover we must have $Q(t) \geq R(t) > R_0$. Since $R(t)$ is presently assumed to be defined for $t^* \leq t < \infty$, and since, from (3.4), $\dot{Q}(t) \leq 0$, it is clear that $\lim_{t \to t^{**}} Q(t) = Q^{**}$, say, must exist and must exceed R_0. Furthermore $Q^{**} \leq Q(t^*) = R(t^*) < f(2K)^{-1/2}$, since $(t^*, R(t^*))$ is on C. Thus $R_0 < Q^{**} < f(2K)^{-1/2}$, so that both $Q - R_0$ and $f^2/R_0Q - 2K$ are positive when $Q = Q^{**}$. It follows by repeated application of the existence theorems for differential equations that the definition of $Q(t)$ can be extended over the whole infinite interval $t^* \leq t < \infty$, and that over this whole interval it is monotonic non-increasing and bounded from below by R_0. Hence

$$(3.5) \qquad Q_0 = \lim_{t \to \infty} Q(t)$$

exists. We also see from Lemma 1 that there exists an infinite sequence of points $t_1, t_2, \ldots$ tending to $+\infty$, such that $\lim_{n \to \infty} \dot{Q}(t_n) = 0$. But from (3.4) and (3.5) we know that $\lim_{t \to \infty} \dot{Q}(t)$ exists (even when we do not restrict ourselves to such a sequence). We are thus justified in writing

$$(3.6) \qquad \lim_{t \to \infty} \dot{Q}(t) = -\left[\frac{Q_0 - R_0}{Q_0}\right]^{\frac{1}{2}} \left[\frac{f^2}{R_0 Q_0} - 2K\right]^{\frac{1}{2}} = 0 .$$

Since the point $(t^*, R(t^*))$ is on C and since both Q_0 and R_0 do not exceed $R(t^*)$, as follows from the monotonicity of $Q(t)$ and $R(t)$, we have

$$\frac{f^2}{R_0 Q_0} > \frac{f^2}{R(t^*)^2} > 2K .$$

From this result and (3.6), it follows that $Q_0 = R_0$. We have thus established for the function $Q(t)$ that

$$\lim_{t\to\infty} Q(t) = R_0 \tag{3.7}$$

and that

$$\lim_{t\to\infty} \dot{Q}(t) = 0 \,. \tag{3.8}$$

An elementary calculation based on (3.4), (3.7), and (3.8) shows that

$$\lim_{t\to\infty} \ddot{Q}(t) = \frac{1}{2R_0}\left(\frac{f^2}{R_0^2} - 2K\right) > \frac{1}{2R_0}\left(\frac{f^2}{R(t^*)^2} - 2K\right) > 0 \tag{3.9}$$

But from (3.8) and Lemma 1 there exists a sequence of points tending to ∞ such that $\ddot{Q}(t)$ tends to 0. This contradicts (3.9), and hence we have finally shown the inconsistency of (3.1), (3.2), and (3.3).

This finishes the proof of the fact that C contains just one point $(t_0, R(t_0))$ where R takes on a minimum value $R_0 = R(t_0)$. It has also been established that $R_0 > 0$.

Evidently the curve $R = R(t)$ must rise on either side of the minimum point (t_0, R_0) on C either indefinitely or until it reaches a relative maximum point (t_1, R_1), where, of course, $\dot{R}(t_1) = 0$, $\ddot{R}(t_1) \leqq 0$, so that this point (t_1, R_1) is beyond the upper bound for points on C. Thus $R_1 > f(2K)^{-1/2} > R_0$. From the Sundman inequality, $H(R(t_0), \dot{R}(t_0)) \leqq H(R(t_1), \dot{R}(t_1))$ together with the fact that $\dot{R}(t_0)$ and $\dot{R}(t_1)$ are both zero, we find that

$$(3.10) \qquad 2K\ R_0 + \frac{f^2}{R_0} \leq 2K\ R_1 + \frac{f^2}{R_1} \qquad (R_1 = R(t_1)) .$$

This can be written in the form $2K\ R_1^2\ R_0 + f^2\ R_0 - 2K\ R_0^2\ R_1 - f^2\ R_1 \geq 0$ or $2K\ R_1\ R_0\ (R_1 - R_0) - f^2(R_1 - R_0) \geq 0$. Since $(R_1 - R_0) > 0$, we find that $2K\ R_1\ R_0 \geq f^2$. In other words $R_1 \geq f^2/2K\ R_0$. Thus, R must rise to at least this value, if there is a relative maximum.

If there is no relative maximum and if R still does not attain the value $f^2(2K\ R_0)^{-1}$, it is clear that $R(t)$ is monotonic and bounded by $f^2(2K\ R_0)^{-1}$, so that it tends to a limit as $t \to \pm\infty$. This limit is greater than R_0 but not greater than $f^2(2K\ R_0)^{-1}$. So we may write

$$(3.11) \qquad R_0 < \lim_{t\to\infty} R(t) = \overline{R} \leq f^2(2K\ R_0)^{-1} .$$

We hereby restrict attention to the case $t \to \infty$, as we are entitled to do because of the invariance of (2.1) under change of sign of t. Let R_2 be any number such that $R_0 < R_2 < \overline{R}$, and let θ be any number between 0 and 1; then set

$$(3.12) \qquad \varepsilon = f\ (R_2 - R_0)^{\frac{1}{2}}\ (1 - \theta)^{\frac{1}{2}}\ R_0^{-\frac{1}{2}}\ \overline{R}^{\,-1} > 0 .$$

From (3.11) and Lemma 1 we can find $t_1 > t_0$ such that $0 \leq \dot{R}(t_1) < \varepsilon$ and such that $R_2 < R_1 \leq \overline{R}$, where $R_1 = R(t_1)$. Using the Sundman inequality as before we now get a slightly modified form of (3.10), namely $R_1\ \varepsilon^2 + 2K\ R_1 + f^2\ R_1^{-1} \geq 2K\ R_0 + f^2\ R_0^{-1}$. This leads to the inequality

$$R_1 \geq \frac{f^2}{2K\ R_0} - \frac{R_1^2\ \varepsilon^2}{2K(R_1 - R_0)} \geq \frac{f^2}{2KR_0} - \frac{\overline{R}^2\ \varepsilon^2}{2K(R_2 - R_0)}$$

Using the expression for ε given in (3.12), we find that $R_1 \geq f^2\theta/2K\ R_0$, as we desired to prove.

For any t between t_0 and t_1, either in the case just treated or in the previous case where $R(t_1) \geq f^2(2K\ R_0)^{-1}$, we of course have $R(t) > R_0$; and hence, the Sundman inequality, with the monotonicity of $R(t)$, leads to the result that $R(t)\ \dot{R}(t)^2 + 2KR(t) + f^2R(t)^{-1} \geq 2KR_0 + f^2R_0^{-1}$. From this, we get after an elementary calculation the inequality given in the statement of the theorem, the proof of which is now complete.

Theorem 4. If $R_0 < \min[f(2K)^{-1/2},\ 3f^2\theta\ 2^{-1}\ M^{-2}(\overline{m} + 4\overline{\mu})^{-1/2}]$, then ρ attains a value at least as great as $2M^2/(3K)$.

Proof. By Theorem 3, R attains a value $> \theta f^2/(2K\ R_0)$ which cannot be less than

$$\frac{\theta\ f^2}{2k\left[\dfrac{3f^2\theta}{2M^2(\overline{m} + 4\overline{\mu})^{1/2}}\right]} = \frac{M^2}{3K}(\overline{m} + 4\overline{\mu})^{\frac{1}{2}}.$$

Hence, by Theorem 2, the corresponding value of $\rho \geq 2M^2/(3K)$.

Theorem 5. Let $\lambda = 2M^2/[3(K\ m^*)^{1/2}]$, then $|r\ \dot{r}| \leq \lambda$.

Proof. From $r^2 = x^2 + y^2 + z^2$ and $\dot{r} = \dot{x}(\frac{x}{r}) + \dot{y}(\frac{y}{r}) + \dot{z}(\frac{z}{r})$ we find from the Cauchy-Schwarz inequality that $\dot{r}^2 \leq \dot{x}^2 + \dot{y}^2 + \dot{z}^2$, which by the energy integral (2.10) is less than $2m^{-1}\ U$. Since r is the least of the three mutual distances, we have

$$\dot{r}^2 < \frac{2(m_0m_1 + m_0m_2 + m_1m_2)}{m\ r}.$$

As shown in the proof of Theorem 1, $m_0m_1 + m_0m_2 + m_1m_2 \leq \frac{1}{3} M^2$. Hence

$$r^2\ \dot{r}^2 < \frac{2\ M^2\ r}{3m} \leq \frac{4\ M^2\ r}{3\ m^*}$$

by (2.13). Now, by Theorem 1, $r \leq M^2(3K)^{-1}$. Hence

$$r^2\ \dot{r}^2 \leq \frac{4\ M^4}{9\ m^*\ K},$$

so that

$$|r\ \dot{r}| \leq \frac{2\ M^2}{3(m^*\ K)^{1/2}} = \lambda,$$

as desired.

<u>Theorem 6</u>. If $R\ \dot{R} > B(R)^{1/2}$, where $B(R) = (4\ M^{1/2}\ \bar{m}^{3/4}\ R^{1/2} + \lambda\ \bar{m})^2$, then $\rho\ \dot{\rho} > 4\ M^{1/2}\ \rho^{1/2}$.

<u>Proof</u>. By (2.11) $\mu\ \rho^2 < R^2$, so that

$$R^{\frac{1}{2}} > \mu^{\frac{1}{4}}\ \rho^{\frac{1}{2}}. \tag{3.13}$$

Also by differentiation of (2.11), we have $\mu\rho\dot{\rho} = R\ \dot{R} - m\ r\ \dot{r}$. Applying Theorem 5, we see that $-\ m\ r\ \dot{r} \geq -\ \lambda m$, while $R\ \dot{R} > 4\ M^{1/2}\ \bar{\mu}^{3/4}\ R^{1/2} + \lambda\bar{m} \geq 4\ M^{1/2}\ \mu^{3/4}\ R^{1/2} + \lambda m$, in accordance with (2.13) and (2.14). Hence $\mu\rho\dot{\rho} > 4\ M^{1/2}\ \mu^{3/4}\ R^{1/2}$. Referring back to (3.13), we see that $\mu\rho\dot{\rho} > 4\ M^{1/2}\ \mu\ \rho^{1/2}$, whence the stated result follows on division by μ.

Theorem 7. Let

$$\varepsilon = \min\left[f\left(\frac{\theta}{2K}\right)^{\frac{1}{2}} , \; \frac{3\,\theta\, f^2}{2\, M^2 (\bar{m} + 4\,\bar{\mu})^{1/2}} \right]$$

and let

$$\sigma = \frac{\theta f^2}{2\,K\,\varepsilon} \quad .$$

If

$$R_0 \leq \min\left[\varepsilon \; , \; \frac{\sigma}{2} \; , \; \frac{f^2\,\sigma}{2[B(\sigma) + K\,\sigma^2]} \right] ,$$

then R takes on the value σ at some later time $t = t_{R_0}$ and the corresponding value of $\dot{R}$ is such that $\sigma^2\,\dot{R}^2 > B(\sigma)$.

Proof. If

$$R_0 \leq \varepsilon \leq \frac{\theta^{\frac{1}{2}}\, f}{\sqrt{2K}} < \frac{f}{\sqrt{2K}} ,$$

we know from Theorem 3 that R will attain a value $\geq \theta\, f^2/(2K\, R_0) \geq \theta\, f^2/(2K\varepsilon) = \sigma$. R also takes on the value

$$R_o \leq \varepsilon \leq \frac{\theta^{1/2}\, f}{\sqrt{2K}} = \eta ,$$

say. But then

$$\eta^2 = \frac{\theta\, f^2}{2K} ,$$

so that

$$\eta = \frac{\theta\, f^2}{2K\,\eta} \leq \frac{\theta f^2}{2K\varepsilon} = \sigma \quad .$$

Hence R takes on values both less than σ and greater than (or equal to) σ . Hence, since $R(t)$ is continuous, it takes on the value σ at some $t = t_{R_0}$.

By the second part of Theorem 3, we see that at $t = t_{R_0}$, where $R = \sigma$, we must have

$$\sigma^2 \dot{R}^2 \geq (\sigma - R_0)\left(\frac{f^2}{R_0} - 2K\sigma\right) \geq \frac{\sigma}{2}\left[\frac{f^2}{\left(\frac{f^2\sigma}{2[B(\sigma) + K\sigma^2]}\right)} - 2K\sigma\right] = B(\sigma)$$

as we wished to prove.

Theorem 8. Under the hypotheses of Theorem 7, the value of ρ at $t = t_{R_0}$ is $> \frac{2}{3}\frac{M^2}{K}$, while that of R is $> M^2(3K)^{-1}(\bar{m} + 4\bar{\mu})^{1/2}$.

Proof.

$$\sigma = \frac{\theta f^2}{2K\varepsilon} \quad \text{and} \quad \varepsilon < \frac{3\theta f^2}{2M^2(\bar{m} + 4\bar{\mu})^{1/2}}. \quad \text{Hence}$$

$$\sigma > \frac{\theta f^2}{2K\left[\frac{3\theta f^2}{2M^2(\bar{m} + 4\bar{\mu})^{1/2}}\right]} = \frac{M^2(\bar{m} + 4\bar{\mu})^{\frac{1}{2}}}{3K}.$$

Moreover $R = \sigma$ when $t = t_{R_0}$. Hence at $t = t_{R_0}$ we have

$$R > \frac{M^2(\bar{m} + 4\bar{\mu})^{1/2}}{3K}.$$

Hence, by Theorem 2, we have $\rho \geq \frac{2}{3}\frac{M^2}{K}$ when $t = t_{R_0}$.

Theorem 9. If $\rho > \frac{2}{3}\frac{M^2}{K}$, then $\ddot{\rho} > -8M\rho^{-2}$. If for any such value of ρ, we have $\dot{\rho} \geq 4M^{1/2}\rho^{-1/2}$, ρ will constantly increase without bound.

We omit the proof of Theorem 9 because the proof in [1] is entirely satisfactory.

Theorem 10. In the three body problem with masses m_0, m_1, m_2, suppose the total energy $-K$ to be negative; so that $K > 0$. Suppose also that the angular momentum c of the system about its center of mass is not zero; so that $f = |c| > 0$. Let R be defined by the formula

$$R^2 = M^{-1}(m_0 m_1 r_2^2 + m_1 m_2 r_0^2 + m_2 m_0 r_1^2) \quad ,$$

where $M = m_0 + m_1 + m_2$ and where r_i is the distance between m_j and m_k, $j \neq i$, $k \neq i$, $j \neq k$.

I. Then the minimum of these three mutual distances is under all circumstances not greater than $M^2(3K)^{-1}$ and there exists a positive number R_0 such that, if $R \leq R_0$ at any time $t = t_0$, there exists a later time t_1 such that $R(t_1) > \frac{M^2}{3K}(\overline{m} + 4\,\overline{\mu})^{1/2}$ where $\overline{m}$ is the greater of the two smaller masses and $\overline{\mu} = 2\, m_0 m_1 m_2\, M^{-1}/m^*$, m^* being the least of the three masses.

II. For $t \geq t_1$, the pair of masses closest together retain their identities and $\lim_{t\to\infty} R(t) = \infty$, so that two of the r_i become infinite while the third one remains bounded (by $M^2(3K)^{-1}$) as $t \to \infty$.

III. A number R_0 for which the above is true may be calculated as follows: Let $\lambda = 2\, M^2/[3(K\, m^*)^{1/2}]$ and $B(R) = (4\, M^{1/2}\, \overline{\mu}^{-3/4}\, R^{1/2} + \lambda \overline{m})^2$. Choose any positive number $\theta < 1$. Let

$$\varepsilon = \min\left[f\left(\frac{\theta}{2K}\right)^{\frac{1}{2}} ,\ \frac{3\,\theta\, f^2}{2\, M^2(\overline{m} + 4\,\overline{\mu})^{1/2}} \right]$$

and let $\sigma = \frac{\theta\, f^2}{2\, K\, \varepsilon}$. Then finally take

$$R_0 = \min\left[\varepsilon\ ,\ \frac{\sigma}{2}\ ,\ \frac{f^2\,\sigma}{2[B(\sigma)\ +\ K\,\sigma^2]}\right].$$

Proof. The assertion under I to the effect that the least of the three mutual distances never exceeds $M^2(3K)^{-1}$ follows from Theorem 1. The other result under I follows from Theorem 8, assuming that ε and R_0 are determined as described under III, so that the hypotheses of Theorems 7 and 8 are both fulfilled. The t_1 of the present theorem can, of course, be taken as the t_{R_0} of Theorems 7 and 8.

From Theorem 7 we also have

$$R^2\,\dot{R}^2 \geq B(R) \tag{3.14}$$

when $t = t_1$, since $R(t_1) = \sigma$.

Let m_i and m_j be the two bodies closest together at $t = t_1$ and let their distance apart at any time t be $r = r(t)$. Let the third body m_k be distant $\rho(t)$ from the center of gravity of m_i and m_j. Then from Theorem 8 we know that at $t = t_1$ we must have

$$\rho > 2\,M^2/(3K) \geq 2r\ . \tag{3.15}$$

Now if, for $t > t_1$, ρ never decreases, the inequalities (3.15) persist, since, as long as $r(t)$ remains the least mutual distance, $2r$ can never exceed $2M^2/(3K)$ whereas both r_i and r_j exceed $\rho - r > 2r - r = r$, and for r_i (for instance) ever to become less than r, it would be necessary (by continuity) for r_i and r to be equal at some $t^* > t_1$. But this is impossible, since the posibility of changing the first inequality sign in (3.15) to an equality sign is ruled out. This means that the pair of masses closest together retain their identities for $t > t_1$, at least if ρ never decreases.

That ρ does indeed never decrease for $t > t_1$ but actually increases without limit may be seen as follows: From (3.14) and Theorem 6 we see that at time $t = t_1$ we must have

$$\dot{\rho} > 4 M^{\frac{1}{2}} \rho^{-\frac{1}{2}} \tag{3.16}$$

We now combine (3.15), (3.16) and Theorem 9 to show that

$$\lim_{t\to\infty} \rho(t) = \infty . \tag{3.17}$$

Finally, the proof of Theorem 10 is completed by the remark that the result

$$\lim_{t\to\infty} R(t) = \infty$$

follows from (2.12), (3.17), and Theorem 1.

Reference

[1] George D. Birkhoff. Dynamical Systems, American Mathematical Society Colloquium Publications, vol. IX, Chapter IX. Especially pp. 275-282.

An Algorithm To Obtain Series Expansions in the Three-Body Problem

by P. Sconzo
IBM, Cambridge Advanced Space Systems
Cambridge, Massachusetts

N67-29375

SUMMARY

This report on Contract NAS 12-98, "Modern Celestial Mechanics," awarded by NASA-ERC to IBM-Cambridge Advanced Space Systems, deals with the analytical solution of the three-body problem by means of convergent power series expansions. The method followed uses only simple algebraic tools and fully exploits the properties of Levi-Civita's regularizing transformation. Stimulated by the suggestion made some years ago by Vernić (1955) to take advantage of this transformation for purposes of practical computation, we present here, for the first time in the long history of the three-body problem, a workable algorithm for constructing recursively its power series solution in terms of Levi-Civita's regularizing variable. This method solves the three-body problem formulated in its utmost generality, since no restrictions at all are made on the order of magnitude of masses and distances and none of the three bodies is restricted to move along a conic section orbit. Besides, the reference system used is a tridimensional inertial one.

After an introductory section dedicated to the historical background of the problem, the second section deals with the discussion of the algebra involved in the derivation of the time series solution. In a third section, the radius of convergence of this solution in the neighborhood of a non-collision point is determined according to Sundman. Section number four describes the property of regularizing variables with particular emphasis on Levi-Civita's variable u. Finally, in the fifth and last section the power series solution in the new variable u is constructed by a procedure of successive approximations in which only elementary algebraic operations are to be performed on polynomials.

In the conclusion, a comment is made about the merits of the algorithm described in Section V, and the items now being developed to implement this investigation are listed.

I. INTRODUCTORY BACKGROUND

The first remarkable contribution to the analytical solution of the general problem of the motion of three bodies subject to mutual Newtonian attractions was made by Painlevé (1896). Here, the designation "general problem" is used in contrast to "restricted problem," which characterizes a special case of the three-body problem, namely, that when one of the three masses is negligible and the two bodies of finite mass revolve around one another in circular orbits.

In various papers and in his masterly lectures at Stockholm (1897), Painlevé[1] demonstrated that: a) the motion is regular and the coordinates may be represented by series of polynomials in t (time) in any interval in which no collisions occur; and b) the motion ceases to be regular in either of the following cases, when the three bodies collide at the same point or when only two of them collide, their distance from the third body remaining a finite quantity. The analytical characterization of the occurrence of collisions was later examined by Levi-Civita[2] (1903-04) and his disciple Bisconcini[3] (1904). They also investigated qualitatively the holomorphic representation of all possible trajectories starting at a collision point within a small domain around this point.

Subsequently, Sundman[4] (1907-1912) made the second and decisive contribution to the analytical solution of the general three-body problem. He demonstrated the possibility of representing the coordinates and the velocities of the three bodies by means of convergent series and determined their radius of convergence. The key points of his proof are: a) the existence theorem by Cauchy-Picard for the solution of a system of differential equations of the 1st order; and b) the transformation of the independent variable t into a new variable u by means of

$$du = \frac{1}{r}\,dt \qquad \text{(I. 1)}$$

where r is the distance between two colliding bodies, that is the distance which will vanish at a given value $\bar{t}$ of t. After showing that $\lim_{t \to \bar{t}} u$ exists and is finite, Sundman demonstrated that the coordinates and velocities, even in the neighborhood of a binary collision point $t = \bar{t}$, can be expanded in convergent power series of $(t - \bar{t})^{1/3}$. This helps the understanding of the analytical nature of the singularity point. In fact, three branches of the same function can be permuted one into another around the point $t = \bar{t}$. It is possible, therefore, to continue analytically the representation of u, the coordinates and the velocities after, or before, the collision point by always taking the real value of the cubic root of $t - \bar{t}$.

It is of historical interest, however, to note that in order to remove the singularity at a binary collision point $t = \bar{t}$, the appropriate choice of the variable u, according to the transformation (I. 1) had been pointed out earlier, but not exploited, by Bruns[5] (1884).

After Sundman's work, the regularizing variable concept was generalized in various directions (successive binary collisions, triple collisions, imaginary collisions, other regularizing transformations, interpretation in the complex domain, etc.) and also gained acceptance among authors of classical books on Celestial Mechanics, notably, Charlier[6], Wintner[7], Siegel[8] and others. Recently, Arenstorff[9] has published a series of papers on this subject using methods taken from the theory of the functions of complex variables. The goal of this paper is to describe an algorithm leading to the construction of convergent series expansions of the coordinates when a regularizing variable of Sundman's type is used as the independent variable. Our choice for this variable is given by the transformation

$$du = V dt \tag{I. 2}$$

indicated by Levi-Civita at the end of his celebrated memoir of the year 1917.[10] In (I. 2) V is the total potential function defined as the negative of the total potential energy. Vernić[11] advocated also the application of the transformation (I. 2) to numerical computations. A goal similar to ours is that pursued by Steffensen[12] to obtain the time series expansion in the case of the planar restricted three-body problem by means of recursion formulas. It turns out that replacing t by u, recursion formulas can also be found for the case of the general three-body problem. In fact, explicit expressions for the derivative of the coordinates $\frac{d^\nu x_i}{du^\nu}$, for sufficiently large values of ν will be established later in this report. The application of the transformation (I. 1) or (I. 2) to the 2-body problem leads to a unified set of formulas valid for any kind of conic sections. A heuristic approach to the derivation of such universal formulas has been illustrated elsewhere.[13]

II. THE EQUATIONS OF MOTION AND THEIR SOLUTION BY TIME POWER SERIES

Let m_0, m_1, m_2 be three non-zero masses and write the equations of motion of all three bodies in an inertial Cartesian reference frame

$$m_i \ddot{x}_i = \frac{\partial V}{\partial x_i}, \qquad \begin{matrix}(i = 0, 1, 2)\\ (x \rightarrow y \rightarrow z)\end{matrix} \tag{II. 1}$$

Here, the dots represent derivatives with respect to the variable τ

(II. 2) $$\tau = kt,$$

k = gravitational constant,
t = time.

The total potential function V in (II. 1) is defined by

(II. 3) $$V = \sum_{ij} m_i m_j \frac{1}{r_{ij}} , \quad (i \neq j)$$

where the mutual distances r_{ij} between bodies are given by

(II. 4) $$r_{ij} = \left[(x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2 \right]^{\frac{1}{2}}$$

Then, the explicit expression for equation (II. 1) is

(II. 5) $$\ddot{x}_i = m_j \mu_{ij} (x_j - x_i) + m_k \mu_{ik} (x_k - x_i) ,$$

where the subscripts i, j, k are permuted cyclically according to the following rule

(II. 6)

i	j	k
0	1	2
1	2	0
2	0	1

and, in general, $\mu_{ij} = \mu_{ji}$ is the inverse cube of the distance r_{ij}

(II. 7) $$\mu_{ij} = \frac{1}{r_{ij}^3} .$$

Equations similar to (II. 5) can be written for the other coordinates y_i and z_i.

The formal power series solution of the equations of type (II. 5) can be written as follows

(II. 8) $$x_i(\tau) = \sum_{\nu=0}^{\infty} x_{i\nu} \tau^{\nu} , \qquad \begin{array}{l} (i = 0, 1, 2) \\ (x_i \rightarrow y_i \rightarrow z_i) \end{array}$$

where

(II. 9) $$x_{i\nu} = \frac{1}{\nu!}\left[\frac{d^\nu x_i}{dt^\nu}\right]_{\tau=0} ,$$

and $\tau = 0$ is an initial (origin) value of τ. The first two coefficients of the expansion (II. 8) are

(II. 10) $$x_{i0} = x_i(0) , \quad x_{i1} = \dot{x}_i(0) ,$$

that is, position and velocity components at origin $\tau = 0$. Initial position and velocities components of all three bodies are assumed to be known.

The next coefficients x_{i2} of the Maclaurin expansion (II. 8) is $\frac{1}{2}\ddot{x}_i(0)$ which can also be obtained by directly evaluating the RHS* of equation (II.5) at $\tau = 0$. The other coefficients of the higher order terms can be obtained by repeated differentiation with respect to τ of equation (II. 5).

This repeated differentiation requires, however, a considerable effort in algebraic manipulations. In fact, there are nine differential equations of 2nd order, coupled together by the mutual distances; hence, there are nine Maclaurin series expansions in which the coefficients depend on several intermixed parameters (mutual distances, masses, positions and velocities). The expressions for these coefficients become extremely complicated, and their complexity increases with the order of the differentiation. This explains why in the past the computation of these coefficients was considered to be impractical, until Steffensen[12] found a relatively easy way of computing them in a special case. This author showed that it is possible in the planar restricted case of the three-body problem, which reduces to only two coupled 2nd order differential equations, to establish recursion formulas for the computation of the coefficients of the two corresponding time series. Earlier attempts on this subject were those for the asteroidal case of the three-body problem by Stumpff[14] and Sconzo[15]. Rabe[16], Deprit[17] et al. have recently made use of Steffensen's formulation in numerical studies of planar restricted problems. Another interesting attempt to get the high order terms of the time series has been described by Gröbner[18] who applied Lie's concept of infinitesimal transformation[19] to the study of the n-body problem. In the application of this concept to the planar case of the three-body system formed by the Earth, Moon and spaceship, however, only the 4th order coefficient of the time series has been computed explicitly. In another paper by Bogayevskiy[20], where a lengthy recursion formula for Gröbner coefficients is presented, only the 3rd order coefficient is given explicitly in terms of the initial conditions. Bogayevskiy's procedure is complicated not only because the notation used is very complex but also because the number of terms which

*Right-hand side

constitute the desired expression is to be determined solving Diophantine equations.

Now, we deem it noteworthy to present the expressions for x_{03} and x_{04} arrived at by two successive differentiations of equation (II. 5) written for $i = 0$. The following concise expressions can, in fact, be derived

(II. 11)
$$x_{03} = \frac{1}{3!}\left[\frac{d^3x_0}{d\tau^3}\right]_{(0)} = \frac{1}{2}\left[m_1\mu_{01}\sigma_{01} + m_2\mu_{02}\sigma_{02}\right]_{(0)} x_0(0) - \frac{1}{6}\left[m_1\mu_{01} + m_2\mu_{02}\right]_{(0)} \dot{x}_0(0) - \frac{1}{2}\left[m_1\mu_{01}\sigma_{01}\right]_{(0)} x_1(0) + \frac{1}{6}\left[m_1\mu_{01}\right]_{(0)} \dot{x}_1(0) - \frac{1}{2}\left[m_2\mu_{02}\sigma_{02}\right]_{(0)} x_2(0) + \frac{1}{6}\left[m_1\mu_{02}\right]_{(0)} \dot{x}_2(0) \ ,$$

(II. 12)
$$x_{04} = \frac{1}{4!}\left[\frac{d^4x_0}{d\tau^4}\right]_{(0)} = \left[\frac{1}{8}\left\{m_1\mu_{01}(\epsilon_{01} - 5\sigma_{01}^2) + m_2\mu_{02}(\epsilon_{02} - 5\sigma_{02}^2)\right\} + \frac{1}{24}\left\{(m_1\mu_{01} + m_2\mu_{02})^2 + m_0(m_1\mu_{01}^2 + m_2\mu_{02}^2)\right\}\right]_{(0)} x_0(0) + \frac{1}{4}\left[m_1\mu_{01}\sigma_{01} + m_2\mu_{02}\sigma_{02}\right]_{(0)} \dot{x}_0(0) + m_1\left[-\frac{1}{8}\mu_{01}(\epsilon_{01} - 5\sigma_{01}^2) + \frac{1}{24}\left\{m_2\mu_{02}\mu_{12} - (m_0 + m_1)\mu_{01}^2 - m_2\mu_{01}(\mu_{02} + \mu_{12})\right\}\right]_{(0)} x_1(0) - \frac{1}{4}\left[m_1\mu_{01}\sigma_{01}\right]_{(0)} \dot{x}_1(0) + m_2\left[-\frac{1}{8}\mu_{02}(\epsilon_{02} - 5\sigma_{02}^2) + \frac{1}{24}\left\{m_1\mu_{01}\mu_{12} - (m_0 + m_2)\mu_{02}^2 - m_1\mu_{02}(\mu_{01} + \mu_{12})\right\}\right]_{(0)} x_2(0) - \frac{1}{4}\left[m_2\mu_{02}\sigma_{02}\right]\dot{x}_2(0) \ ,$$

if the following notations are adopted

(II. 13) $$\sigma_{01} = \frac{1}{r_{01}^2} s_{01} ,$$

(II. 14) $$\epsilon_{01} = \frac{1}{r_{01}^2} \dot{s}_{01} ,$$

where

$$s_{01} = \underset{(y,z)}{S} \left[(x_1 - x_0)(\dot{x}_1 - \dot{x}_0) \right] ,$$

$$\dot{s}_{01} = \underset{(y,z)}{S} \left[(\dot{x}_1 - \dot{x}_0)^2 - \frac{m_0 + m_1}{r_{01}} - m_2(x_1 - x_0) \left\{ \mu_{02}(x_2 - x_0) - \mu_{12}(x_2 - x_1) \right\} \right] .$$

The symbol $\underset{(y,z)}{S}\left[\ldots\right]$ means a sum extended to terms in the variables y and z similar to that in x.

Inspecting the structure of the RHS of both equations (II. 11) and (II. 12) we see that up to the third order term the perturbation effect upon the coordinates of the three bodies is of the first order with respect to the masses, while starting from the 4th order term this effect becomes of 2nd order or higher.

Applying the cyclic permutation of the indices indicated by (II. 6) to the equations (II. 13) and (II. 14), the expressions for x_{13}, x_{14} and x_{23}, x_{24} can easily be derived from (II. 11) and (II. 12), respectively. Similar expressions for the y and z coordinates can also be easily derived. We observe that Bogayevskiy's expression for x_{03}, presented in ten strings of terms, can be simplified to only 8 non-vanishing terms and in its simplified version it coincides with our polynomial expression (II. 11) which contains precisely 8 terms.

We want now to call the attention of the reader to the fact that both equations (II. 11) and (II. 12) have been brought to a form analogous to that of the coefficients of the Lagrangian f and g series in the two body problem.[21] The symbols μ_{ij}, σ_{ij}, ϵ_{ij} are also an extension of the symbols μ, σ, ϵ adopted for the two-body problem in the referenced paper. Although the said analogy can be extended further and formulas similar to Cipolletti's[22] recursion formulas could be established (this will be the object of a later IBM funded investigation) here, we will not try to derive explicitly the expressions for the derivatives of the coordinates with respect to the variable τ corresponding to an order $\nu \geqslant 5$. Our aim as it has been stated in the introduction is to obtain these high order derivatives with respect to the new variable u.

It is important, however, to determine the radius of convergence of the series expansion (II. 8). We dedicate the next section to this problem.

III. CONVERGENCE OF THE TIME SERIES EXPANSION

It can be said, in general, that the convergence radius of the series expansion (II. 8) is determined by the distance of the nearest singularity point (real or imaginary) to the origin. It can be demonstrated that the series (II. 8) is convergent in the neighborhood of a non-collision point. An upper positive bound T of the independent variable τ reckoned from the time of a non-collision point $\bar{\tau}$ can, in fact, be found such that within the interval (-T, T), centered at $\bar{\tau}$, the series (II. 8) is convergent. To demonstrate this we will use Cauchy's theorem as implemented by Picard[23] on the existence and uniqueness of the solution of a system of first order differential equations in the real domain.

Let

(III. 1) $$f_i\left[x_1(\tau),\ x_2(\tau),\ \dots,\ x_n(\tau)\right],\quad (i = 1, 2, \dots, n)$$

be real functions of the real independent variable τ, but not containing τ explicitly, satisfying the following conditions

a) they can be expanded in convergent power series of $x_i - \bar{x}_i$

when

(III. 2) $$| x_i - \bar{x}_i | < \eta_i \ ,$$

where $\bar{x}_i = x_i(\bar{\tau})$, $\bar{\tau}$ is an arbitrary value of τ, which can be taken $\bar{\tau} = 0$, and all η_i are positive quantities;

b) it is

(III. 3) $$| f_i | < F_i \ ,$$

when x_i satisfies the inequality (III. 2) and where all F_i are positive quantities. Then, the system of differential equations

(III. 4) $$\frac{dx_i}{d\tau} = f_i$$

admits one and only one solution such that

(III. 5) $$\lim_{\tau\to\bar{\tau}} x_i = \bar{x}_i \quad .$$

Under the conditions specified above it can be added that it is possible to expand the unknown functions $x_i(\tau)$ into a convergent power series of $\tau - \bar{\tau}$ for any value of $\bar{\tau}$ such that

(III. 6) $$|\tau - \bar{\tau}| < T ,$$

where

(III. 7) $$T = \min\left\{\frac{\eta_i}{F_i}\right\} .$$

Furthermore, the inequality (III.2) is verified when τ is chosen according to (III. 6).

In order to apply the Cauchy-Picard theorem stated above to our system of nine differential equations of 2nd order (II. 5), we transform this system into another system of 18 first order equations as follows

(III. 8) $$\frac{dx_i}{d\tau} = \dot{x}_i \quad ,$$

(III. 9) $$\frac{d\dot{x}_i}{d\tau} = m_j\mu_{ij}(x_j - x_i) + m_k\mu_{ik}(x_k - x_i). \qquad (x_i \to y_i \to z_i)$$

Now, let η_0 and $\dot{\eta}_0$ be two positive constants such that

(III. 10) $$|x_i - \bar{x}_i|,\ |y_i - \bar{y}_i|,\ |z_i - \bar{z}_i| < \eta_0 ,$$

(III. 11) $$|\dot{x}_i - \bar{\dot{x}}_i|,\ |\dot{y}_i - \bar{\dot{y}}_i|,\ |\dot{z}_i - \bar{\dot{z}}_i| < \dot{\eta}_0 , \qquad (i = 0, 1, 2)$$

and we will demonstrate that condition a) of Cauchy's theorem is satisfied, that is the RHS of equations (III. 8) and (III. 9) can be expanded in convergent power series of the differences $x_i - \bar{x}_i$, etc. It suffices to demonstrate that under the conditions (III. 10) the inverse cube of any of the three distances r_{ij} can be expanded into a convergent series.

We take into consideration, for example, the distance r_{01} and we write

(III. 12) $$r_{01}^2 = \bar{r}_{01}^2 + P_{01} \quad ,$$

where

(III. 13) $$\bar{r}_{01}^2 = r_{01}^2(\bar{T}) \quad ,$$

and P_{01} is a second degree polynomial in the differences $x_0 - \bar{x}_0$, etc.

The explicit expression of P_{01} is

(III. 14) $$P_{01} = \underset{(y,z)}{S} \left[2(\bar{x}_1 - \bar{x}_0)\left\{(x_1 - \bar{x}_1) - (x_0 - \bar{x}_0)\right\} + \left\{(x_1 - \bar{x}_1) - (x_0 - \bar{x}_0)\right\}^2 \right].$$

It follows that

(III. 15) $$|P_{01}| \leqslant 12\,\bar{r}_{01}\eta_0 + 12\eta_0^2$$

because

(III. 16) $$|\bar{x}_1 - \bar{x}_0|,\ |\bar{y}_1 - \bar{y}_0|,\ |\bar{z}_1 - \bar{z}_0| \leqslant \bar{r}_{01} .$$

Consequently, the expansion of $\frac{1}{r_{01}}$ as well as that of $\mu_{01} = \frac{1}{r_{01}^3}$

(III. 17) $$\mu_{01} = \frac{1}{\bar{r}_{01}^3}\left(1 + \frac{P_{01}}{\bar{r}_{01}^2}\right)^{-\frac{3}{2}}$$

is convergent if we choose η_0 such that

(III. 18) $$12\,\bar{r}_{01}\eta_0^2 + 12\eta_0^2 < r_{01}^2$$

or

(III. 19) $$\eta_0 < \frac{\bar{r}_{01}}{6 + 4\sqrt{3}} \quad .$$

Condition (III. 19) is satisfied <u>a fortiori</u> if we take

(III. 20) $$\eta_0 = \frac{\bar{r}_{01}}{14} .$$

We determine now two upper bounds for the RHS of both equations (III. 8) and (III. 9). We begin with the equations of type (III. 9) and we observe that with the choice (III. 20) we can first deduce

(III. 21) $$r_{01} > \sqrt{\bar{r}_{01}^2 - 12\bar{r}_{01}\eta_0 - 12\eta_0^2} = \frac{2}{7}\,\bar{r}_{01}$$

and since

$$|x_1 - x_0| \leqslant |\bar{x}_1 - \bar{x}_0| + |x_0 - \bar{x}_0| + |x_1 - \bar{x}_1| ,$$

we deduce also

$$|x_1 - x_0| < \bar{r}_{01} + 2\eta_0 = \frac{8}{7}\bar{r}_{01}. \tag{III. 22}$$

Then

$$|\mu_{01}(x_1 - x_0)| < \left(\frac{7}{2\bar{r}_{01}}\right)^3 \frac{8}{7}\bar{r}_{01} = \frac{1}{4\eta_0^2} \tag{III. 23}$$

and other inequalities similar to (III. 23) can be established for $|\mu_{02}(x_2 - x_0)|$, etc.

Now, if $\tau = \bar{\tau}$ is a non-collision point we can find a lower bound of the three distances at time $\tau = \bar{\tau}$, that is a positive number $\eta \leqslant \eta_0$ such that

$$\bar{r}_{01}, \bar{r}_{02}, \bar{r}_{12} > 14\eta . \tag{III. 24}$$

Thus, the inequalities (III. 10) can be replaced by

$$|x_i - \bar{x}_i|, |y_i - \bar{y}_i|, |z_i - \bar{z}_i| < \eta \tag{III. 25}$$

and the inequality (III. 23) by

$$|\mu_{01}(x_1 - x_0)| < \frac{1}{4\eta^2} . \tag{III. 26}$$

Considering the equation (III. 9) we then obtain

$$\left| \frac{d\dot{x}_i}{d\tau} \right| < \frac{1}{4\eta^2} (m_j + m_k) < \frac{M}{4\eta^2} , \tag{III. 27}$$

where

$$M = m_0 + m_1 + m_2 . \tag{III. 28}$$

Now, we deal with the equations of type (III. 8). We first deduce the existence of upper bounds for the velocities at time $\tau = \bar{\tau}$. We prove this by considering the energy integral

(III. 29) $$\frac{1}{2}\sum_{i=0}^{2} m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) - V = h$$

rewritten as follows

(III. 30) $$\sum_{i=0}^{2} m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) - 2\,\frac{m_0 m_1 m_2}{M}\,U = \frac{m_0 m_1 m_2}{M}\,h ,$$

where

(III. 31) $$\frac{m_0 m_1 m_2}{M}\,U = V = \frac{m_0 m_1}{r_{01}} + \frac{m_0 m_2}{r_{02}} + \frac{m_1 m_2}{r_{12}}$$

and h is a constant.

When $\tau \to \bar{\tau}$, $U \to \bar{U}$, where $\bar{U}$ according to (III. 31) and (III. 24) satisfies the inequality

(III. 32) $$\frac{m_0 m_1 m_2}{M}\,\bar{U} < (m_0 m_1 + m_0 m_2 + m_1 m_2)\,\frac{1}{14\eta} .$$

Now, we can apply to the RHS of (III. 32) the following obvious result

$$M^2 = (m_0 + m_1 + m_2)^2 \geqslant 3(m_0 m_1 + m_0 m_2 + m_1 m_2).$$

Doing so, the inequality (III. 32) becomes

(III. 33) $$\frac{m_0 m_1 m_2}{M}\,\bar{U} \leqslant \frac{M^2}{42\eta} .$$

It can also be proved that,

$$M^2 > 4m_0 m_1 ,\ 4m_0 m_2 ,\ 4m_1 m_2$$

then, we have in general

(III. 34) $$\frac{M}{4} > \frac{m_j m_k}{M} . \qquad (j \neq k)$$

If we now divide by m all terms of equation (III. 30) written for $\tau = \bar{\tau}$, where

$$m = \min\{m_0, m_1, m_2\} ,$$

we deduce

$$\dot{\bar{x}}_i^2\,,\ \dot{\bar{y}}_i^2\,,\ \dot{\bar{z}}_i^2 < \frac{M^2}{21m\eta} + \frac{M}{4}|h|.$$

Letting

(III. 35) $$\dot{\eta}_0 = \sqrt{\frac{M^2}{21m\eta} + \frac{M}{4}|h|}$$

we have

$$|\dot{\bar{x}}_i|\,,\ |\dot{\bar{y}}_i|,\ |\dot{\bar{z}}_i| < \dot{\eta}_0$$

and we obtain from (III. 11)

(III. 36) $$|\dot{x}_i|,\ |\dot{y}_i|,\ |\dot{z}_i| < 2\dot{\eta}_0\,.$$

We may conclude that selecting T as follows

(III. 37) $$T = \min\left\{\frac{\eta}{2\dot{\eta}_0}, \frac{\dot{\eta}_0}{(M/4\eta^2)}\right\},$$

the series expansion (II. 8) is convergent for any τ such that

(III. 38) $$|\tau - \bar{\tau}| < T.$$

The entire proof given above has been borrowed, with only slight modifications, from the original memoir by Sundman [4c]. It is worth noting that, of both ratios contained in the RHS of equation (III. 37), the first is smaller than the second. In fact

$$-\frac{\eta}{2\dot{\eta}_0} + \frac{\dot{\eta}_0}{(M/4\eta^2)} = \frac{\eta}{2\dot{\eta}_0}\left(\frac{8\eta\dot{\eta}_0^2}{M} - 1\right) = \frac{\eta}{2\dot{\eta}_0}\left(\frac{8M}{21m} + 2\eta|h| - 1\right) > 0$$

because $3m \leq M$, and consequently $\frac{8M}{24m} \geqslant 1$, $\frac{8M}{21m} > 1$.

Hence, all the series expansions of type (II. 8) are convergent for any τ such that

(III. 39) $$|\tau - \bar{\tau}| < \frac{\eta}{\sqrt{\frac{4M^2}{21m\eta} + M|h|}}\,,$$

and these series represent the analytical solution of the three-body problem which satisfies the preset initial conditions for position and velocity at time $\tau = \bar{\tau}$.

IV. ON REGULARIZING VARIABLES

We say that a variable is a regularizing variable if it has the property of removing that singularity of the differential equations of the motion which corresponds to a binary collision.

We will show that the variable u introduced by the Levi-Civita transformation (I. 2) is a regularizing variable. We demonstrate first the following lemma: if $S(r_{01}, r_{02}, r_{12})$ is a symmetric and homogeneous function of first degree of the three distances r_{01}, r_{02}, r_{12}, then the variable u defined by means of the differential operator

$$\frac{d}{d\tau} = \frac{1}{S}\frac{d}{du} \tag{IV. 1}$$

is a regularizing variable.

In fact, applying twice the operator (IV. 1) to x_i and denoting by primes the derivatives with respect to u, we obtain

$$\dot{x}_i = \frac{1}{S} x_i' \tag{IV. 2}$$

$$\ddot{x}_i = \frac{1}{S^2} x_i'' - \frac{S'}{S^3} x_i' , \tag{IV. 3}$$

or vice versa

$$x_i' = S\dot{x}_i , \tag{IV. 4}$$

$$x_i'' = S\dot{S}\dot{x}_i + S^2\ddot{x}_i , \tag{IV. 5}$$

if we observe that $S\dot{S} = S'$.

By virtue of (IV. 3) and (IV. 5), the equation of motion (II. 1) acquires the form

$$x_i'' = \frac{S'}{S} x_i' + \frac{S^2}{m_i}\frac{\partial V}{\partial x_i} \tag{IV. 6}$$

or the equivalent form

(IV. 7)
$$x_i'' = S\dot{S}\dot{x}_i + \frac{S^2}{m_i}\frac{\partial V}{\partial x_i} .$$

The latter can also be rewritten as follows

(IV. 8)
$$x_i'' = S\dot{x}_i \sum_{ij} \frac{\partial S}{\partial r_{ij}} \dot{r}_{ij} + \frac{S^2}{m_i} \sum_{ij} \frac{\partial V}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial x_i} .$$

We examine now the order of the infinitesimal terms which constitute the RHS of equation (IV. 8) written for $i = 0, 1$ when the two corresponding bodies collide at a certain time $\tau = \bar{\tau}$, that is when $\lim_{\tau \to \bar{\tau}} r_{01} = 0$. Considering r_{01} as an infinitesimal quantity of first order, the order of all functions depending on r_{01} can be listed as follows

Function	S	$\dot{x}_i$ (and also $\dot{y}_i$, $\dot{z}_i$)	$\frac{\partial S}{\partial r_{01}}$	$\dot{r}_{01}$	S^2	$\frac{\partial V}{\partial r_{01}}$	$\frac{\partial r_{01}}{\partial x_0, \partial x_1}$
Order	1	$-\frac{1}{2}$	0	$-\frac{1}{2}$	2	-2	0

All these results are evident except those for $\dot{x}_i$ and $\dot{r}_{01}$ which may be deduced as a consequence of the following four limits known as Bisconcini-Sundman's [4c, 3] relationships

(IV. 9)
$$\lim_{\tau \to \bar{\tau}} \left(\sqrt{r_{01}}\ \dot{r}_{01}\right) = H = -\sqrt{2(m_0 + m_1)} ,$$
$$\lim_{\tau \to \bar{\tau}} \left(\sqrt{r_{01}}\ \dot{x}_i\right) = C_x H ,$$
$$\lim_{\tau \to \bar{\tau}} \left(\sqrt{r_{01}}\ \dot{y}_i\right) = C_y H , \qquad (i = 0, 1)$$
$$\lim_{\tau \to \bar{\tau}} \left(\sqrt{r_{01}}\ \dot{z}_i\right) = C_z H ,$$

where C_x, C_y, C_z are three constants such that

$$C_x^2 + C_y^2 + C_z^2 = 1 .$$

Then, it follows from (IV. 8)

$$O\left(\lim_{u \to \bar{u}} x_i'' \right) = O\left(r_{01}^{0} \right), \quad (i = 0, 1)$$

where

$$\bar{u} = \lim_{\tau \to \bar{\tau}} u \quad .$$

We conclude that

$$\lim_{u \to \bar{u}} x_i^{*} = C \neq 0, \quad (i, = 0, 1) , \tag{IV.10}$$

that is, the RHS of equation (IV.8) is a constant C different from zero when the two bodies (i = 0, 1) collide.

As a consequence of this lemma we conclude that the variable u defined by Levi-Civita differential operator

$$\frac{d}{d\tau} = V \frac{d}{du} \tag{IV.11}$$

is also a regularizing variable.

V. USE OF LEVI CIVITA'S DIFFERENTIAL OPERATOR AND RECURSION FORMULAS FOR THE SOLUTION

Applying the differential operator (IV.11) to the equation of motion (II.1), we obtain

$$V x_i^{*} + V' x_i' = \frac{1}{m_i V} \frac{\partial V}{\partial x_i} \quad . \qquad \begin{matrix} (i = 0, 1, 2) \\ (x_i \to y_i \to z_i) \end{matrix} \tag{V.1}$$

In this equation the function V is to be considered as a function of the variable u. Since V is a known function of τ, in order to obtain its dependence on u we proceed as follows. First, we calculate

$$u(\tau) = \int_0^{\tau} V(\xi)\, d\xi \quad . \tag{V.2}$$

The RHS of equation (V.2) is a power series in τ , which for practical purposes can be truncated to a polynomial of certain degree.

Now, reversing the series given by (V.2), we get

$$\tau = \tau(u) \; ; \tag{V. 3}$$

then we insert (V. 3) into the expression of $V(\tau)$. Denoting by $\overset{*}{V}(u)$ the result of this operation we can rewrite equation (V. 1) in the new form

$$\overset{*}{V} x_i'' + \overset{*}{V}' x_i' = \overset{*}{X}_i \; , \tag{V. 4}$$

where

$$\overset{*}{X}_i = \frac{1}{m_i} \overset{*}{V}^{-1} \frac{\partial \overset{*}{V}}{\partial x_i} \; . \tag{V. 5}$$

Next, let

$$x_i(u) = \sum_{\nu=0}^{\infty} \overset{*}{x}_{i\nu} u^{\nu} \tag{V. 6}$$

be the formal series expansion of the solution of equation (V. 4). Letting also

$$\overset{*}{X}_i(u) = \sum_{\nu=0}^{\infty} \overset{*}{X}_{i\nu} u^{\nu} \; , \tag{V. 7}$$

$$\overset{*}{V} = \sum_{\nu=0}^{\infty} \overset{*}{V}_{\nu} u^{\nu} \; , \tag{V. 8}$$

and inserting (V. 6), (V. 7) and (V. 8) into equation (V. 4) it will not be difficult to arrive at the following recursion formula

$$\overset{*}{x}_{i\nu} = \frac{1}{\nu(\nu-1)\overset{*}{V}_0} \overset{*}{X}_{i\nu-2} - \frac{1}{\nu \overset{*}{V}_0} \sum_{k=1}^{\nu-1} (\nu-k) \overset{*}{V}_k \overset{*}{x}_{i\nu-k} , \quad (\nu \geqslant 2) \; . \tag{V. 9}$$

Formulas similar to (V. 9) can also be established for $\overset{*}{y}_{i\nu}$, $\overset{*}{z}_{i\nu}$ after defining

$$\overset{*}{Y}_i = \frac{1}{m_i} \overset{*}{V}^{-1} \frac{\partial \overset{*}{V}}{\partial y_i} \; , \tag{V. 10}$$

$$\overset{*}{Z}_i = \frac{1}{m_i} \overset{*}{V}^{-1} \frac{\partial \overset{*}{V}}{\partial z_i} \; . \tag{V. 11}$$

Since

(V.12) $$\overset{*}{x}_{i0} = x_i(0)$$

(V.13) $$\overset{*}{x}_{i1} = \frac{1}{\overset{*}{V}_0} \dot{x}_i(0)$$

are known, the use of (V.9) for $\nu = 2, 3, \ldots$ provides all the coefficients needed to write the formal expansion of $\overset{*}{x}_i$ in a power series of u. We need to start from known values of $\overset{*}{X}_{i0}$, $\overset{*}{V}_0$ and $\overset{*}{V}_1$ as functions of the initial conditions. For this purpose we observe that starting from the given initial conditions (II.10) we can write the following linear approximation

(V.14) $$x_i = x_{i0} + x_{i1}\tau, \quad \begin{matrix} (i = 0, 1, 2) \\ (x_i \to y_i \to z_i) \end{matrix}.$$

Then, we have

(V.15) $$r_{ij}^2 = \bar{r}_{ij}^2 + p_{ij1}\tau,$$

where

$$\bar{r}_{ij}^2 = \underset{(y,z)}{S}\left[x_{j0} - x_{i0}\right]^2,$$

$$p_{ij1} = 2 \underset{(y,z)}{S}\left[(x_{j0} - x_{i0})(x_{j1} - x_{i1})\right].$$

It follows from (V.15) that

(V.16) $$\frac{m_i m_j}{r_{ij}} = g_{ij0} + g_{ij1}\tau,$$

where

$$g_{ij0} = \frac{m_i m_j}{\bar{r}_{ij}},$$

$$g_{ij1} = -\frac{1}{2\bar{r}_{ij}^2} g_{ij0}\, p_{ij1}.$$

Hence

(V.17) $$V = V_0 + V_1\tau,$$

where

$$V_0 = \sum_{ij} g_{ij0} , \quad V_1 = \sum_{ij} g_{ij1} ,$$

and according to (V.2) we obtain

(V.18) $$u(\tau) = V_0\tau + \frac{1}{2} V_1\tau^2 .$$

Reversing (V.18) we get

(V.19) $$\tau = \frac{1}{V_0} u - \frac{1}{2} \frac{V_1}{V_0^3} u^2 .$$

Inserting (V.19) into (V.17) and truncating the result to the linear approximation we obtain

$$\overset{*}{V}(u) = \overset{*}{V}_0 + \overset{*}{V}_1 u ,$$

$$\overset{*}{V}{}^{-1}(u) = \frac{1}{\overset{*}{V}_0} - \frac{\overset{*}{V}_1}{\overset{*}{V}{}_0^{2}} u ,$$

where

(V.20) $$\overset{*}{V}_0 = V_0 , \qquad \overset{*}{V}_1 = \frac{V_1}{V_0} .$$

We have also

(V.21) $$\mu_{ij} = \frac{1}{\bar{r}_{ij}^3} - \frac{3}{2\bar{r}_{ij}^5} P_{ij1} \tau ,$$

(V.22) $$x_j - x_i = (x_{j0} - x_{i0}) + (x_{ji} - x_{i1}) \tau$$

Finally, inserting (V.19) into (V.21) and (V.22) and performing the products $m_j\mu_{ij}(x_j - x_i)V^{-1}$, $m_k\mu_{ik}(x_k - x_i)V^{-1}$ and their sum we obtain

$$\overset{*}{X}_i = \overset{*}{X}_{i0} + \overset{*}{X}_{i1} u ,$$

where $\overset{*}{X}_{i0}$ and $\overset{*}{X}_{i1}$ are now known expressions of the initial conditions (II. 10). As a matter of fact we can find $\overset{*}{x}_{i2}$ directly from its definition

$$\overset{*}{x}_{i2} = \frac{1}{2}\left[\frac{d^2 x_i}{du^2}\right]_{u=0} .$$

Applying twice the inverse operator of (IV. 11) to x_i we have in fact

$$\frac{1}{2}\frac{d^2x_i}{du^2} = \frac{1}{2}\frac{1}{V}\frac{d}{d\tau}\left(\frac{1}{V}\frac{dx_i}{d\tau}\right) = \frac{1}{2}\left(\frac{1}{V^2}\frac{d^2x_i}{d\tau^2} - \frac{1}{V^3}\frac{dV}{d\tau}\frac{dx_i}{d\tau}\right)$$

and the evaluation of both sides at the origin (u = 0 and τ = 0, respectively) provides

$$\overset{*}{x}_{i2} = \frac{1}{2V_0^2}[\xi_i]_{t=0} - \frac{1}{2V_0^3} V_1 x_{i1} , \tag{V. 23}$$

where

$$[\xi_i]_{t=0} = x_{i2} = \frac{1}{m_i}\left[\frac{\partial V}{\partial x_i}\right]_{t=0} .$$

The result expressed by (V. 23) coincides with that obtained from (V. 9) if we put $\nu = 2$. We take into consideration (V. 13) and (V. 20), and we observe that by virtue of (V. 5) $\overset{*}{X}_{i0} = \frac{1}{V_0}[\xi_i]_{t=0}$.

Next, in order to obtain $\overset{*}{x}_{i3}$, we need to know $\overset{*}{V}_2$. We arrive at $\overset{*}{V}_2$ by replacing (V. 14) by a second degree polynomial in u which is known because $\overset{*}{x}_{i2}$ is now a known quantity. The computation of the 2nd order expressions in the variable u for r_{ij}, $\frac{m_i m_j}{r_{ij}}$ and $\overset{*}{V}$ will provide the desired $\overset{*}{V}_2$. The same procedure is then applied to compute $\overset{*}{V}_3, \ldots$, and, therefore, all other coefficients of the series expansion (V. 6).

The algebra involved in this computation, in summary, consists of the following operations to be performed on polynomials

- subtraction
- multiplication
- integer and $-\frac{1}{2}$ and $-\frac{3}{2}$ power
- reversion and inversion

- substitution of the variable of a polynomial by another polynomial
- integration

This algebra is being programmed for polynomials of arbitrarily high degree.

We observe that the polynomials in the variable τ (polynomials of first degree) of type (V.14) are needed only to initialize the computation. The computation is then continued by using polynomials in the variable u. At each successive approximation, a new term will be added to the polynomials expressing the nine coordinates.

The relationship between τ and u should also be computed at each approximation replacing equation (V.2) by the equivalent equation

$$\tau = \int_0^u \frac{1}{\overset{*}{V}(\sigma)} \, d\sigma \quad . \tag{V.24}$$

CONCLUSION

The series expansion of the solution of the three-body problem has been found in terms of Levi-Civita's regularizing variable u. The recursion formulas of type (V.9) are the key points of the computation. By means of these formulas the coefficients of the series expansions for all nine coordinates are expressed as functions of the masses and the initial conditions.

That these series are convergent can be demonstrated using a well known theorem applied to the solution of linear differential equations of type (V.1) when they are solved by the power series method. This proof is being adapted to our type of equations and it will be presented in the next report.

It is expected that by using power series in the variable u, the analytical representation of the motion can be extended to an interval of time much larger than the step-size used in numerical integration procedures. If so, the method could advantageously be utilized in place of numerical integration. It would, in fact, lend itself to being exploited as a numerical tool for the analytical continuation of the solution. A previous effort in this direction, using time series expansions, was made in 1955 by the author of this investigation.[15]

In the next report we also intend to present a computed numerical case of the three-body problem using the method illustrated above. Specifically, we will choose an example considered as a "zweckmässig" one by Bohlin [24] and for which results, both in tabular and graphical form, have been obtain by Zumkley [25], who used a numerical integration procedure. In this examp the masses and the distances have the same order of magnitude, and the example is, therefore, appropriate to be used as a test case of the method described above. The only simplification used in this example is that it deals with a planar case instead of a tridimensional one, and this has been done in order to reduce the burden of the computation without prejudicing th general validity of the formulation. Some preliminary computed results we have obtained by hand computation, carried out up to 5th order terms in u, confirm the expectation that large intervals of time can be covered by this kind of truncated series representation of the motion of all three bodies.

REFERENCES

1. Painlevé, P.: a) Sur les singularités des équations de la dynamique et sur le problème des trois corps. C.R. t. 123, 871(1896); ibid. t. 125, 1075 (1897). b) Leçons sur la théorie analytique des équations différentiel, professées à Stockholm. Paris, Hermann (1897).

2. Levi-Civita, T.: a) Traiettorie singolari ed urti nel problema ristret dei tre corpi. Ann. Mat. S. 3, t. 9, 1 (1903); b) Sur la résolution qualitative du problème restreint des trois corps. Acta Math. t. 30, 305 (1904).

3. Bisconcini, G.: Sur le problème des trois corps. Acta Math. t. 30, 49 (1904).

4. Sundman, K. F.: a) Recherches sur le problème des trois corps. Acta Soc. Scientiarum Fennicae. t. 34, No. 6 (1907). b) Nouvelles recherc etc. ibid. t. 35, No. 9 (1909). c) Mémoire sur le problème des trois corps. Acta Math. t. 36, 105 (1912).

5. Bruns, H.: Bemerkungen zur Theorie der allgemeinen Störungen. Ast Nach. Bd. 109, 215 (1884).

6. Charlier, C. V. L.: Die Mechanik des Himmels, two volumes. Walter Gruyter, Berlin (1927).

7. Wintner, A., Analytical Foundations of Celestial Mechanics. Princeto University Press, Princeton, N. J. (1947).

8. Siegel, C. L.: Vorlesungen über Himmelsmechanik. Springer Verlag. Berlin (1956).

9. Arenstorf, R.: On the representation of the solution of the restricted problem of three bodies by power-series. Ballistic Missile Agency. Redstone Arsenal, Alabama. Rep. No. DC-TR-1-59 (1959).

10. Levi-Civita, T.: Sur la régularization du problème des trois corps. Acta Math. 42, 99 (1918).

11. Vernić, R.: Numerisko rjesavanje orseg problema triju tijela. Jugoslavenske Akademije znanosti i umjet nosti. Zagreb, 302, 47 (1955).

12. Steffensen, J. F.: Kgl. Danske Videnskab. Selskab. Mat. fys. Medd. 30, No. 18. Kopenhagen (1956),

13. Sconzo, P.: A heuristic approach to the introduction of a universal variable (to be printed in Mem. Soc. Astr. Italiana).

14. Stunpff, K., Untersuchungen über das Problem der speziellen Störungen in der rechtwinkligen Koordinaten. Astr. Nach. Bd. 273, 105 (1942).

15. Sconzo, P.: Las perturbaciones en las coordenadas heliocéntricas calculadas con un procedimiento numerico de prolongación analítica. Rev. Unión Mat. Argentina. Vol. 17, 223 (1955).

16. Rabe, E.: Determination and survey of periodic Trojan orbits in the restricted problem of three bodies. The Astr. Journal, 66, 500 (1961).

17. Deprit, A. and Price, J. F.: The computation of characteristic exponents in the planar restricted problem of three bodies. Boeing Scientific Research Lab. D1-82-0450 (1965).

18. Gröbner, W. and Cap, F.: The three-body problem Earth - Moon - Spaceship. Astronautica Acta. 5, 287 (1959).

19. Lie, S.: Forhand. Vid. -Selsk, 243. Christiania (1871).

20. Bogayevskiy, A. N.: Solution of the three-body problem in series form (1964). (Unknown source, only an English translation kindly supplied by NASA/ERC is at hand.)

21. Sconzo, P., LeSchack, A. R., Tobey, R.: Symbolic computation of f and g series by computer. The Astr. Journal, 70, 269 (1965).

22. Cipolletti, D.: Espressioni generali dello sviluppo in serie delle coordinate di un corpo celeste. Publ. Oss. Astr. Firenze, Riv. Sc.-Ind., Luglio (1872).

23. Picard, E.: a) Sur la convergence des séries représentant les intégrales des équations différentielles. Bull. Sc. Math. (1888).
b) Traité d'Analyse, t. 2, 301. Gauthier-Villars, Paris (1893).

24. Bohlin, K.: Über ein zweckmässiges Beispiel der Bewegung im allgemeinen Dreikörper problem. Astr. Iaktt. Unders. Stockholms Obs. 10, No. 11 (1923).

25. Zumkley, J.: Ein numerisch gerechneter Spezialfall des allgemeinen Dreikörper problems in vereinfachter Behandlung. Astr. Nach. Bd. 272, 66 (1941).

Rigorous Error Bounds on Position and Velocity in Satellite Orbit Theories*

by J. V. Breakwell and J. Vagners
Stanford University
Stanford, California

ABSTRACT

By utilizing results of Hamiltonian theory and the von Zeipel method for treating artificial satellite orbits, error bounds are derived for a general class of orbits with eccentricity less than one. In order to extend the error bounds for the general axisymmetric problem to time intervals of the order $1/J_2$, the known integral of energy is utilized to calibrate the governing differential equations for the rapidly rotating phase. The non-singular rapid phase in this analysis is taken to be the sum of the mean anomaly, argument of periapsis and the right ascension of the ascending node. A corresponding analysis for the general asymmetric problem (including the tesseral harmonics) is also given. From the general error analysis an algorithm is derived for the computation of the correct initial conditions consistent with the expected accuracy of the theory. Numerical results verifying the conclusions of the theory presented in this paper are also given.

*This work was performed in association with research sponsored by the National Aeronautics and Space Administration under Research Grant NsG-133-61

I. INTRODUCTION

The analytical theory of artificial satellite motion has been the subject of very intensive study since the launching of the first artificial satellite in 1957. In fact, many aspects of the problem had been studied before that time in connection with the theories of celestial mechanics. The result of the study has been a very extensive list of papers offering solutions of many differing forms and techniques of achieving them. However, with the exception of the work of Kyner (Ref. 1), no other solution is known to the authors that offers rigorous error bounds on the position and velocity for a general class of orbits, e.g., inclined orbits of any eccentricity less than one. Naturally, the orbits at critical inclination and orbits in resonance with the tesseral harmonics must be excepted from the general class. It is then a matter of general interest to derive such error bounds.

From a fundamental point of view, the problem of artificial satellite motion can be classified as a special case of a general class of non-linear oscillation problems. Non-linear oscillation problems can be treated with varying degrees of success by the general averaging methods developed by Krylov, Bogoliubov and Mitropolskii (i.e., Ref. 2). For these methods of averaging there exists an associated technique for establishing bounds on the error build-up in a specified time between the exact and the approximate solutions (first order or higher order). Now, the method of application of the technique of averaging to the problem of artificial satellite motion depends rather heavily on the particular choice of variables employed. In the case of Kyner's work, averaging could be applied directly; in most other approaches to the problem the use of averaging is more or less disguised.

One of the most widely used perturbation methods in treating artificial satellite orbits has been the method of von Zeipel as adopted by Brouwer (Ref. 3) and Kozai (Ref. 4). This method is one of successive canonical transformations and is necessarily carried out in the variables of Delaunay (L,G,H,ℓ,g,h). With a slight change of variables and a choice of a different intermediary orbit, the same method was applied by Garfinkel (Refs. 5,6). Furthermore, it has been shown (Refs. 7,8) that the von Zeipel method of canonical transformations is a particular form of the method of averaging. Hence, by drawing on the equivalence to averaging, rigorous error bounds could be established for the Delaunay variables directly. Unfortunately, bounds obtainable in this way for the Delaunay variables ℓ and g are unsatisfactory for very small eccentricity (i.e., $e' < J_2$ where J_2 is the oblateness parameter of $O(10^{-3})$) due to a singularity at zero eccentricity in the short period terms. A further drawback is the singularity at zero inclination. Since no singularities exist in the coordinates for zero eccentricity and/or inclination, one would expect that these objections to the bounds would not exist for a suitable choice of variables. The error bounds derived by such direct application of differential equation theory turn out to be unsatisfactory for large time intervals i.e., time intervals of the order $1/J_2$. Since one of the problems of interest in applying closed-form orbit theories is orbit prediction over long periods of time, the error theory must be modified. The modification is a more involved problem and a separate treatment is presented here.

In this report, the problem is analyzed in canonical variables;

the three sets of interest are those due to Delaunay, Hill and Poincaré. Of these variables, the Poincaré set is non-singular for both zero eccentricity and inclination, the Hill set singular for zero inclination and the Delaunay set singular for both zero inclination and eccentricity. The advantages of the Hill set are the simple forms of the in-plane coordinate perturbations which are obtained directly from known generating functions. It was shown by Izsak (Ref. 9) that, to first order in the oblateness coefficient J_2, the in-plane position and velocity components of a satellite are obtainable by converting via Keplerian formulae from Brouwer's averaged Delaunay variables (L', G', H', ℓ', g') to corresponding "averaged" position and velocity and then superimposing the short-period fluctuations. These short-period fluctuations were shown to be obtainable by rewriting Brouwer's short-period generating function S_1 in terms of the Hill variables and taking appropriate partial derivatives. These short-period fluctuations are well-behaved (unlike those in ℓ, g) when eccentricity goes to zero. Recent investigations by Vagners (Ref. 10) have obtained in the same manner first order long-period fluctuations in the Hill variables by rewriting Brouwer's long-period generating function S_1^* relating $(L'', G'', H'', \ell'', g'', h'')$ to $(L', G', H', \ell', g', h')$, including general formulas for the effects of any zonal harmonic. Analogous "medium-period" (i.e., daily) fluctuations in the Hill variables were obtained in a general form for the effects of the tesseral and sectorial harmonics. Since the analysis given by Vagners was applicable to any set of canonical variables, then similar results could readily be obtained for the Poincaré variables.

Utilizing the results of Izsak and Vagners, an analysis is carried out in this report which parallels every canonical tranformation of the Delaunay variables by an appropriate canonical transformation of some general set of canonical variables including the removal of second order short-period terms from the Hamiltonian. In this way, rigorous error bounds on the first-order solution are established which are independent of the eccentricity for Hill variables and independent of eccentricity and inclination for the Poincaré variables (as long as e is not too close to one). As is shown, these bounds are unsatisfactory for long time intervals and another method is offered.

A discussion is presented of the various terms arising in the error bound. Particular attention is focused on the question of initial condition errors; this question is of interest when computing by means of a "closed-form" satellite theory a satellite's ephemeris from some given initial position and velocity vectors. In view of the extensive comparison studies of different orbit theories conducted by Arsenault, Enright and Purcell (Ref. 11), wherein the problem of initialization plays such an important role, this question assumes considerable importance. An energy method is then given for greatly decreasing the primary in-track position error build-up due to initial conditions and some typical results are quoted. The algorithm of computing the correct initial conditions arises directly from the extended error bound theory.

The authors wish to acknowledge the contribution of Small (Ref. 12), who first utilized the energy method in reducing initialization errors in his solution to the problem of satellite motion about an oblate planet.

II. GENERAL BOUNDS ON SATELLITE MOTION

Before proceeding to more specific treatment of the error problem, some general statements concerning the a priori bounds on the motion may be made. First, one can consider the motion of a satellite in a general axi-symmetric gravitational field for which two integrals of the motion are known. If the potential field is represented by

$$V = -\frac{\mu}{r}\left\{1 - \sum_{N=2}^{\infty} J_N \left(\frac{R_\oplus}{r}\right)^N P_N(\sin\beta)\right\} = -\left[U_o(r) + U_1(r,\beta)\right] \quad \text{(II-1)}$$

where β is the latitude, $R_\oplus$ the equatorial radius, μ the gravitational constant, r the radius and J_N numerical coefficients, then it can readily be shown that the total energy and the polar component of the angular momentum are constants of the motion. The two exact integrals may be written in the form

$$\frac{1}{a} + \frac{2}{\mu} U_1(r,\beta) = k_1 \quad \text{(II-2)}$$

$$\text{and} \quad H = \sqrt{\mu a(1 - e^2)} \cos i = k_2 \quad \text{(II-3)}$$

where a is the semi-major axis of the orbit, e the eccentricity, i the orbital inclination and k_1 k_2 are constants.

The two integrals (II-2) and (II-3) imply that if k_1 and k_2 are given, then the motion of the satellite is confined to a region bounded by a "zero velocity" surface (Ref. 13). With initial conditions specifying k_1 and k_2 one can write the a priori bounds in the form

$$0 < \delta_1(k_1, k_2, \epsilon) < r < \delta_2(k_1, k_2, \epsilon) \quad \text{(II-4)}$$

where $\epsilon \triangleq J_2$, $k_1 > 0$, $k_2 > 0$ and J_N/J_2^2 are assumed values of $0(1)$.

General bounds of this type are developed by Poritsky (Ref. 14) and given for $\epsilon = 0$ by Kyner (Ref. 1). Here the explicit forms of δ_1 and δ_2 are not of direct interest.

In the more general problem of a longitude dependent potential one no longer has the two integrals (II-2) and (II-3). Such a potential arises when one includes the tesseral harmonics of the Earth's field in the general satellite problem. However, by considering a rotating coordinate system fixed in the primary, one can readily determine the Jacobi integral of the system. In this case one specifies only the upper bound by the zero-velocity surface.

One assumes then that a priori bounds on the state vector x are known; namely, if the initial state vector $x(0)$ is in a set D, then

$$|x| \leq C(x(0),\ \epsilon) \tag{II-5}$$

where the solution depends on a small parameter ϵ. Since for near-earth satellites one is concerned with elliptical orbits, the set D will be specified by the requirement of negative energy and a non-zero initial value of the angular momentum. If the state vector chosen for the description of the motion is some canonical set (q,p), then the equations of motion take the form

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \dot{x} = \Phi_o \mathcal{H}_{\tilde{x}}\ (x,\epsilon) \tag{II-6}$$

where $\mathcal{H}$ is the Hamiltonian of the problem

$\Phi_o = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ the canonical matrix

$\tilde{x}$ denotes the partials of $\mathcal{H}$ with respect to x and the super tilda denotes the transpose of the vector x.

Then, since $\mathcal{H}_{\tilde{x}}$ is continuous and satisfies a Lipschitz condition locally in x in some bounded region $\mathcal{R}$ (then) a solution for all t exists as a consequence of (II-5).

Note that implicit in (II-5) is also a restriction on how close the energy and the angular momentum may be to zero. For the general bounds to hold, these initial values must be sufficiently different from zero so that the perturbations, of order ϵ in the satellite problem, do not cause the state vector x to become arbitrarily large.

III. THE SECOND-ORDER HAMILTONIAN

Inherent in a specific discussion of error bounds is a knowledge of the characteristics of the analytical method used in the fundamental solution and a knowledge of the behavior of various functions arising therein. The method utilized in the following analysis is the von Zeipel method and the system analyzed is a Hamiltonian system. For a brief review of the von Zeipel procedure, the reader is referred to Ref. 10; the specific details of the orbit problem solution may be found in Refs. 3 and 4.

It turns out to be convenient to introduce the three sets of canonical variables due to Delaunay, Hill and Poincaré. (Recall that the original solution of Brouwer was carried out in Delaunay variables.) These sets of variables are defined in the following manner: The Delaunay variables, denoted by y, are given as

$$y = \begin{bmatrix} Q \\ P \end{bmatrix} = \begin{bmatrix} \ell \\ g \\ h \\ L \\ G \\ H \end{bmatrix} \tag{III-1}$$

where ℓ is the mean anomaly

$g = \omega$ the argument of pericenter

$h = \Omega$ the right ascension of the ascending node

$L = \sqrt{\mu a}$

$G = \sqrt{\mu a(1 - e^2)}$

$H = G \cos i$

Denoting the Hill variables by z, one finds

$$z = \begin{bmatrix} r \\ u \\ h \\ R \\ G \\ H \end{bmatrix} \tag{III-2}$$

where u the central angle or argument of latitude

$R = \dot{r}$ the radial velocity

And, finally, the Poincaré variables, denoted by x, are

$$x = \begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} \lambda \\ \eta_1 \\ \eta_2 \\ L \\ \xi_1 \\ \xi_2 \end{bmatrix} \tag{III-3}$$

with

$$\begin{aligned} \lambda &= \ell + g + h & L &= L \\ \eta_1 &= [2(L - G)]^{\frac{1}{2}} \cos (g + h) & \xi_1 &= [2(L - G)]^{\frac{1}{2}} \sin (g + h) \\ \eta_2 &= [2(G - H)]^{\frac{1}{2}} \cos h & \xi_2 &= [2(G - H)]^{\frac{1}{2}} \sin h \end{aligned} \tag{III-4}$$

Note that equations (III-4) give the transformation from Delaunay to Poincaré, and that no singularities are introduced in this transformation. The inverse transformation is given by

$$\begin{aligned} \ell &= \lambda - \tan^{-1} \frac{\xi_1}{\eta_1} & L &= L \\ g &= \tan^{-1} \frac{\xi_1}{\eta_1} - \tan^{-1} \frac{\xi_2}{\eta_2} & G &= L - \frac{\xi_1^2 + \eta_1^2}{2} \\ h &= \tan^{-1} \frac{\xi_2}{\eta_2} & H &= G - \frac{\xi_2^2 + \eta_2^2}{2} \end{aligned} \tag{III-5}$$

In the transformation (III-5), the equations for the momenta L,G,H exhibit no singularities, whereas in the coordinates ℓ,g,h singularities will arise for zero eccentricity ($\eta_1 = 0$) and for zero inclination ($\eta_2 = 0$). This feature of the transformation will be important in later analysis.

If one denotes a general canonical set of variables by w, then the equations of motion take the form (see Eq. (II-6)):

$$\dot{w} = \Phi_o \frac{\partial}{\partial \tilde{w}} \mathcal{H}(w,\epsilon) \tag{III-6}$$

with $w = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ α the generalized coordinates

β the associated momenta

where for the artificial Earth satellite problem the Hamiltonian is written as

$$\mathcal{H}(w,\epsilon) = -\frac{\mu^2}{2L^2(w)} + \epsilon\, \mathcal{H}^{(1)}(w) + \epsilon^2 \mathcal{H}^{(2)}(w) \tag{III-7}$$

The oblateness coefficient J_2 has been taken as the small parameter ϵ for convenience. Since all of the higher harmonics in the expansion for the Earth's field are of at least $0(J_2^2)$, one can represent their contribution as an ϵ^2 term (Eq. (II-4)).

Now, apply a stationary canonical transformation to define a new set of variables w':

$$\begin{aligned} \alpha &= \alpha' - \epsilon\, \mathcal{D}^{(1)}_{\tilde{\beta}'}(\beta',\alpha) - \epsilon^2 \mathcal{D}^{(2)}_{\tilde{\beta}'}(\beta',\alpha) \\ \beta &= \beta' + \epsilon\, \mathcal{D}^{(1)}_{\tilde{\alpha}}(\beta',\alpha) + \epsilon^2 \mathcal{D}^{(2)}_{\tilde{\alpha}}(\beta',\alpha) \end{aligned} \tag{III-8}$$

which has been truncated with the second order terms. The $\mathcal{D}^{(i)}$ are the "generating functions" of the canonical transformation. The new

Hamiltonian is then

$$\mathcal{H}'(w',\epsilon) = \mathcal{H}(w,\epsilon) = -\frac{\mu^2}{2L^2} + \epsilon\left\{\mathcal{H}^{(1)} + \frac{\mu^2}{L^3}\left(L_{\beta'}\mathcal{D}^{(1)}_{\tilde{\alpha}'} - L_{\alpha'}\mathcal{D}^{(1)}_{\tilde{\beta}'}\right)\right\} +$$

$$+\epsilon^2\left\{\mathcal{H}^{(2)} + \left(\mathcal{H}^{(1)}_{\beta'}\mathcal{D}^{(1)}_{\tilde{\alpha}'} - \mathcal{H}^{(1)}_{\alpha'}\mathcal{D}^{(1)}_{\tilde{\beta}'}\right) + \frac{\mu^2}{L^3}\left(L_{\beta'}\mathcal{D}^{(2)}_{\tilde{\alpha}'} - L_{\alpha'}\mathcal{D}^{(2)}_{\tilde{\beta}'}\right)\right.$$

$$-\frac{3\mu^2}{2L^4}\left(L_{\beta'}\mathcal{D}^{(1)}_{\tilde{\alpha}'} - L_{\alpha'}\mathcal{D}^{(1)}_{\tilde{\beta}'}\right)^2 + \frac{\mu^2}{L^3}\left(-L_{\beta'}\mathcal{D}^{(1)}_{\tilde{\alpha}'\alpha'}\mathcal{D}^{(1)}_{\tilde{\beta}'} + L_{\alpha'}\mathcal{D}^{(1)}_{\tilde{\beta}'\alpha'}\mathcal{D}^{(1)}_{\tilde{\beta}'}\right)$$

$$\left.+\frac{1}{2}\frac{\mu^2}{L^3}\left(\mathcal{D}^{(1)}_{\alpha'}, -\mathcal{D}^{(1)}_{\beta'}\right)\begin{pmatrix} L_{\tilde{\beta}'\beta'} & L_{\tilde{\beta}'\alpha'} \\ L_{\tilde{\alpha}'\beta'} & L_{\tilde{\alpha}'\alpha'}\end{pmatrix}\begin{pmatrix}\mathcal{D}^{(1)}_{\tilde{\alpha}'} \\ -\mathcal{D}^{(1)}_{\tilde{\beta}'}\end{pmatrix}\right\} + \epsilon^3 f(w',\epsilon) \qquad \text{(III-9)}$$

All functions in Eq. (III-9) are to be evaluated at w'. Choose $\mathcal{D}^{(1)}(\alpha,\beta')$ and $\mathcal{D}^{(2)}(\alpha,\beta')$ so that $\mathcal{H}'(w')$ contains no short period terms except in $f(w',\epsilon)$. This requirement is defined by

$$\frac{\partial}{\partial \ell'}\left[\mathcal{H}'(w',\epsilon) - \epsilon^3 f(w',\epsilon)\right] = 0 \qquad \text{(III-10)}$$

with ℓ' the Delaunay variable conjugate to $L' = L(w')$. The Poisson bracket

$$[A,B] = A_{\beta'}B_{\tilde{\alpha}'} - A_{\alpha'}B_{\tilde{\beta}'} = A_{w'}\Phi_o B_{\tilde{w}'} \qquad \text{(III-11)}$$

is easily shown to be invariant under a canonical transformation. In particular

$$[L', \mathcal{D}^{(1)}] = \frac{\partial \mathcal{D}^{(1)}}{\partial \ell'} \qquad \text{(III-12)}$$

then if one writes

$$\mathcal{H}^{(i)}(w') = \bar{\mathcal{H}}^{(i)}(w') + \hat{\mathcal{H}}^{(i)}(w')$$

with $\bar{\mathcal{H}}^{(i)}(w') = \underset{\ell'}{\text{av}}\ \mathcal{H}^{(i)}(w')$

one chooses $\mathcal{D}^{(1)}(w')$ such that

$$\frac{\mu^2}{L^3}\frac{\partial \mathcal{D}^{(1)}}{\partial \ell'} = - \hat{\mathcal{H}}^{(1)} \tag{III-13}$$

This defines $\mathcal{D}^{(1)}(w')$ uniquely up to an additive function of the Delaunay variables other than ℓ'. It is then convenient to choose $\mathcal{D}^{(1)}$ to be identical with Brouwer's $S^{(1)}(L',G',H,\ell,g,\text{--})$ expressed as $S^{(1)}(\alpha,\beta')$. Note that the function $S^{(1)}$ is non-singular for zero eccentricity and/or inclination and is a function (as Brouwer writes it) of both L,G explicitly and implicitly through e and f, the true anomaly. When computing the required partial derivatives for ℓ and g short period variations, the singularity for zero eccentricity, for example, arises in the following way

$$\begin{aligned}
\frac{\partial S^{(1)}}{\partial L'} &= \left(\frac{\partial S^{(1)}}{\partial L'}\right)_{\text{expl.}} + \left(\frac{G'^2}{e'L'^3}\right)\frac{\partial S^{(1)}}{\partial e'} \\
\frac{\partial S^{(1)}}{\partial G'} &= \left(\frac{\partial S^{(1)}}{\partial G'}\right)_{\text{expl.}} - \left(\frac{G'}{e'L'^2}\right)\frac{\partial S^{(1)}}{\partial e'}
\end{aligned} \tag{III-14}$$

As shown in Ref. 10, no $\frac{1}{e'}$ terms arise in the case of the Hill variables; however, zero inclination singularities still exist. That no singularities occur for the Poincaré variables can readily be demonstrated. The argument is given for the variable λ; similar arguments apply to the other variables. The function $S^{(1)}$ is given explicitly as $S^{(1)}(e',f',g',G',H')$ so $S^{(1)}$ depends on L' also through e' and f'. According to the von Zeipel procedure the first order short-period variations of λ are given by

$$\delta\lambda_1 = \frac{\partial S^{(1)}}{\partial L'} \qquad \text{(III-15)}$$

which then can be written as (dropping the primes for convenience)

$$\delta\lambda_1 = \frac{\partial S^{(1)}}{\partial L} = \left(\frac{\partial S^{(1)}}{\partial L}\right)_{\text{expl.}} + \frac{\partial e}{\partial L}\frac{\partial S^{(1)}}{\partial e} + \frac{\partial G}{\partial L}\frac{\partial S^{(1)}}{\partial G} + \frac{\partial H}{\partial L}\frac{\partial S^{(1)}}{\partial H} \qquad \text{(III-16)}$$

with

$$\frac{\partial S^{(1)}}{\partial e} = \left(\frac{\partial S^{(1)}}{\partial e}\right)_{\text{expl.}} + \left(\frac{a}{r} - \eta^{-2}\right)\frac{\partial S^{(1)}}{\partial f}\sin f$$

$$\eta = (1 - e^2)^{\frac{1}{2}}$$

So

$$\delta\lambda_1 = \frac{\partial e}{\partial L}\left[\frac{\partial S^{(1)}}{\partial e} + \left(\frac{a}{r} - \eta^{-2}\right)\frac{\partial S^{(1)}}{\partial f}\sin f\right] + \left(\frac{\partial S^{(1)}}{\partial G}\right)_{\text{expl.}} + \left(\frac{\partial S^{(1)}}{\partial H}\right)_{\text{expl.}}$$

where

$$e = \left\{1 - \frac{1}{L^2}\left[L - \frac{\xi_1^2 + \eta_1^2}{2}\right]^2\right\}^{\frac{1}{2}} \qquad \text{(III-17a)}$$

then

$$\frac{\partial e}{\partial L} = \frac{\sqrt{1 - e^2}}{eL}\left[\sqrt{1 - e^2} - 1\right] = \frac{\eta(\eta - 1)}{eL}$$

$$= \frac{\eta(\eta - 1)(\eta + 1)}{eL(\eta + 1)} = -\frac{e\eta}{L(\eta + 1)} \qquad \text{(III-17b)}$$

The derivatives with respect to e and f explicitly introduce no singularities and neither do the last two terms of Eq. (III-17a). Thus $\delta\lambda_1$ is well-behaved as e(and i) $\to 0$.

Next, one can show that the second order long-period Hamiltonian is independent of the particular canonical variables used and, furthermore, that the second order generating function $\mathfrak{D}^{(2)}$ is non-singular for zero eccentricity and/or inclination. Recall that the function $\mathfrak{D}^{(2)}$ is chosen so as to cancel all second order short-period terms of the

Hamiltonian. In order to obtain the desired results, note that

$$\mathcal{H}^{(1)}_{\beta'} S^{(1)}_{\tilde{\alpha}'} - \mathcal{H}^{(1)}_{\alpha'} S^{(1)}_{\tilde{\beta}'} = \left[\mathcal{H}^{(1)}, S^{(1)} \right] \tag{III-18}$$

is an invariant for canonical variables. Furthermore,

$$L_{\beta'} \mathcal{D}^{(2)}_{\tilde{\alpha}'} - L_{\alpha'} \mathcal{D}^{(2)}_{\tilde{\beta}'} = \frac{\partial \mathcal{D}^{(2)}}{\partial \ell'} \tag{III-19}$$

and

$$\left(L_{\beta'} S^{(1)}_{\tilde{\alpha}'} - L_{\alpha'} S^{(1)}_{\tilde{\beta}'} \right) = \left(\hat{\mathcal{H}}^{(1)} \right)^2 \tag{III-20}$$

One can rewrite (from Eq. (III-9))

$$- L_{\beta'} S^{(1)}_{\tilde{\alpha}'\alpha'} S^{(1)}_{\tilde{\beta}'} + L_{\alpha'} S^{(1)}_{\tilde{\beta}'\alpha'} S^{(1)}_{\tilde{\beta}'}$$

as follows (where α, β are understood to be the primed variables)

$$\begin{aligned}
- L_\beta S^{(1)}_{\tilde{\alpha}\alpha} S^{(1)}_{\tilde{\beta}} + L_\alpha S^{(1)}_{\tilde{\beta}\alpha} S^{(1)}_{\tilde{\beta}} &= -\frac{1}{2} L_\beta S^{(1)}_{\tilde{\alpha}\alpha} S^{(1)}_{\tilde{\beta}} + \frac{1}{2} L_\alpha S^{(1)}_{\tilde{\beta}\alpha} S^{(1)}_{\tilde{\beta}} \\
&- \frac{1}{2} \left(L_\beta \frac{\partial}{\partial \tilde{\alpha}} - L_\alpha \frac{\partial}{\partial \tilde{\beta}} \right) \left(S^{(1)}_\alpha S^{(1)}_\beta \right) + \frac{1}{2} L_\beta S^{(1)}_{\tilde{\alpha}\beta} S^{(1)}_{\tilde{\alpha}} - \frac{1}{2} L_\alpha S^{(1)}_{\tilde{\beta}\beta} S^{(1)}_{\tilde{\alpha}} \\
&= -\frac{1}{2} \left\{ \frac{\partial}{\partial \alpha} \left(L_\beta S^{(1)}_{\tilde{\alpha}} - L_\alpha S^{(1)}_{\tilde{\beta}} \right) \right\} S^{(1)}_{\tilde{\beta}} + \frac{1}{2} \left\{ \frac{\partial}{\partial \beta} \left(L_\beta S^{(1)}_{\tilde{\alpha}} - L_\alpha S^{(1)}_{\tilde{\beta}} \right) \right\} S^{(1)}_{\tilde{\alpha}} \\
&+ \frac{1}{2} \left(S^{(1)}_{\tilde{\alpha}} L_{\tilde{\beta}\alpha} - S^{(1)}_\beta L_{\tilde{\alpha}\alpha} \right) S^{(1)}_{\tilde{\beta}} - \frac{1}{2} \left(S^{(1)}_{\tilde{\alpha}} L_{\tilde{\beta}\beta} - S^{(1)}_{\tilde{\beta}} L_{\tilde{\alpha}\beta} \right) S^{(1)}_{\tilde{\alpha}} \\
&- \frac{1}{2} \frac{\partial}{\partial \ell} \left(S^{(1)}_\alpha S^{(1)}_{\tilde{\beta}} \right) = -\frac{1}{2} \left[\hat{\mathcal{H}}^{(1)}, S^{(1)} \right] - \frac{1}{2} \frac{\partial}{\partial \ell} \left(S^{(1)}_\alpha S^{(1)}_{\tilde{\beta}} \right) \\
&+ \frac{1}{2} \left(S^{(1)}_{\tilde{\alpha}} L_{\tilde{\beta}\alpha} - S^{(1)}_\beta L_{\tilde{\alpha}\alpha} \right) S^{(1)}_{\tilde{\beta}} - \frac{1}{2} \left(S^{(1)}_{\tilde{\alpha}} L_{\tilde{\beta}\beta} - S^{(1)}_{\tilde{\beta}} L_{\tilde{\alpha}\beta} \right) S^{(1)}_{\tilde{\alpha}}
\end{aligned} \tag{III-21}$$

then also

$$\frac{1}{2}\left(S_{\alpha}^{(1)}, -S_{\beta}^{(1)}\right)\begin{pmatrix} L_{\tilde{\beta}\beta} & L_{\tilde{\beta}\alpha} \\ L_{\tilde{\alpha}\beta} & L_{\tilde{\alpha}\alpha} \end{pmatrix}\begin{pmatrix} S_{\tilde{\alpha}}^{(1)} \\ -S_{\tilde{\beta}}^{(1)} \end{pmatrix}$$

$$= \frac{1}{2}\left(S_{\alpha}^{(1)}, -S_{\beta}^{(1)}\right)\begin{pmatrix} S_{\tilde{\alpha}}^{(1)} L_{\tilde{\beta}\beta} - S_{\tilde{\beta}}^{(1)} L_{\tilde{\beta}\alpha} \\ S_{\tilde{\alpha}}^{(1)} L_{\tilde{\alpha}\beta} - S_{\tilde{\beta}}^{(1)} L_{\tilde{\alpha}\alpha} \end{pmatrix} \tag{III-22}$$

Thus it can be seen that Eq. (III-22) cancels the last two terms of Eq. (III-21). The second order part of the Hamiltonian (III-9) consequently is given by

$$\epsilon^2\left\{\mathcal{H}^{(2)} + \left[\mathcal{H}^{(1)}, S^{(1)}\right] + \frac{\mu^2}{L'^3}\mathcal{D}_{\ell'}^{(2)} - \frac{3}{2}\frac{L'^2}{\mu^2}\left(\hat{\mathcal{H}}^{(1)}\right)^2 - \frac{\mu^2}{2L'^3}\left[\hat{\mathcal{H}}^{(1)}, S^{(1)}\right] - \frac{1}{2}\frac{\partial}{\partial \ell'}\left(S_{\alpha'}^{(1)} S_{\tilde{\beta}'}^{(1)}\right)\right\} \tag{III-23}$$

The only term of Eq. (III-23) apart from $\frac{\mu^2}{L'^3}\mathcal{D}_{\ell'}^{(2)}$ that depends on the particular canonical variables used is $-\frac{1}{2}\frac{\partial}{\partial \ell'}\left(S_{\alpha'}^{(1)} S_{\tilde{\beta}'}^{(1)}\right)$, which is necessarily short period, and $\mathcal{D}^{(2)}$, of course, is chosen so that the term $\frac{\mu^2}{L'^3}\mathcal{D}_{\ell'}^{(2)}$ cancels all short-period terms.

From this invariance property of the terms in Eq. (III-23) one deduces that the difference between the second order generating functions $\mathcal{D}^{(2)}$ of two different sets of canonical variables, such as any arbitrary set w and the Delaunay set y for example, will be given by

$$\mathcal{D}^{(2)} - S^{(2)} = \frac{L'^3}{2\mu^2}\left(S_{\alpha'}^{(1)} S_{\tilde{\beta}'}^{(1)}\right) - \frac{L'^3}{2\mu^2}\left(S_{Q'}^{(1)} S_{\tilde{P}}^{(1)}\right) + \text{arbitrary long period term} \tag{III-24}$$

Equation (III-24) gives a convenient algorithm for computing the generating function $\mathcal{D}^{(2)}$ for any set of canonical variables. In order to assure that all of the functions arising in the error bound determination remain bounded, one must establish that $\mathcal{D}^{(2)}$ contains no singularities. This may be done utilizing the known results of Izsak (Ref. 9), Brouwer (Ref. 3) and Kozai (Ref. 4).

In Kozai's paper, the expression given for the J_2^2 component of the function $S^{(2)}$ shows the factor $1/e$ for the trigonometric arguments $\sin f$, $\sin (f + 2g)$, $\sin (3f + 2g)$, $\sin (3f + 4g)$ and $\sin (5f + 4g)$. The appearance of this factor is unnecessary and a suitable rearrangement of terms eliminates it. Such rearrangement will be shown explicitly here for the coefficient of $\sin f$; the other terms can be treated similarly. The coefficient of $\sin f$ as given by Kozai is (omitting a nonsingular multiplying factor):

$$\frac{1}{e}\Big[9(11 - 30\,\theta^2 + 27\,\theta^4) - 8\eta^2(17 - 38\,\theta^2 + 11\,\theta^4) - 4\,\eta^3(1 - 3\,\theta^2)^2 \times (3 + \eta^2) + \eta^4(53 - 130\,\theta^2 - 11\,\theta^4)\Big] \tag{III-25}$$

$$\text{where} \quad \theta = \frac{H}{G} = \cos i$$

Equation (III-25) can be rewritten as

$$\frac{1}{e}\Big[99 - 270\,\theta^2 + 243\,\theta^4 - 136 + 304\,\theta^2 - 88\,\theta^4 - 4\eta^3(1 - 3\,\theta^2)^2(3 + \eta^2) + 53 - 130\,\theta^2 - 11\,\theta^4 - 2e^2(121 - 282\,\theta^2 + 33\,\theta^4) + e^4(53 - 13\,\theta^2 - 11\,\theta^4)\Big]$$

or dropping the e^2 and e^4 terms and combining:

$$\frac{4(1 - 3\,\theta^2)^2}{e}\,[4 - 3\eta^3 - \eta^5]$$

Now, this can be rewritten as follows:

$$\frac{4 - 3\eta^2 - \eta^5}{e} = \frac{(1 - \eta)}{e} [\eta^4 + \eta^3 + 4\eta^2 + 4\eta + 4]$$

$$= \frac{(1 - \eta)(1 + \eta)}{e(1 + \eta)} [\eta^4 + \eta^3 + 4\eta^2 + 4\eta + 4]$$

$$= \frac{e}{1 + \eta} [\eta^4 + \eta^3 + 4\eta^2 + 4\eta + 4] \qquad \text{(III-26)}$$

which remains bounded as $e \to 0$. In a similar manner the other expressions given by Kozai (for higher order harmonics J_N) may be rearranged and thus it can be shown that $S^{(2)}$ contains no $1/e$ factors. Of the terms on the right-hand side of Eq. (III-24), the first is known to be bounded (III-17 a,b); the second can be shown to be bounded by the above technique of rearranging.

The (new) canonical variables w' satisfy the differential equations

$$\dot{w}' = \Phi_o \mathcal{H}'_{\tilde{w}'}(w', \epsilon) \qquad \text{(III-27)}$$

where one can write the Hamiltonian in the form

$$\mathcal{H}'(w', \epsilon) = -\frac{\mu^2}{2L'^2(w')} + \epsilon \, \overline{\mathcal{H}}^{(1)}(w') + \epsilon^2 K^{(2)}(w') + \epsilon^3 s(w', \epsilon) \qquad \text{(III-28)}$$

an analytic function of the variables w' and the small parameter ϵ. Define next a transformation (canonical) to the "secular" variables w'' by the truncated expressions

$$\begin{aligned} \alpha'' &= \alpha' - \epsilon \, S^*_{\tilde{\beta}''}(\beta'', \alpha') \\ \beta'' &= \beta' + \epsilon \, S^*_{\tilde{\alpha}'}(\beta'', \alpha') \end{aligned} \qquad \text{(III-29)}$$

where[†] S^* is chosen so as to cancel the long-period part of the Hamiltonian (except near critical inclination). Then $S^*_{\tilde{\beta}''}$ and $S^*_{\tilde{\alpha}'}$ give the first order long-period variations of α and β i.e.:

$$\frac{1}{2\pi}\int_0^{2\pi} S^*_{\tilde{\beta}''}\, dg' = \frac{1}{2\pi}\int_0^{2\pi} S^*_{\tilde{\alpha}'}\, dg' = 0 \qquad \text{(III-30)}$$

The governing differential equations for w'' become then

$$\dot{w}'' = \Phi_0 \left[\mathcal{H}''_{\tilde{w}''}\ (w'', \epsilon) + \epsilon^3\ \psi_{\tilde{w}''}(w'', \epsilon) \right] \qquad \text{(III-31)}$$

and the solution can be written in the form

$$\alpha = \alpha'' - \epsilon S^*_{\tilde{\beta}''}\ (\beta'', \alpha') - \epsilon \overset{(1)}{S}_{\tilde{\beta}'}\ (\beta', \alpha) - \epsilon^2 \overset{(2)}{\mathfrak{D}}_{\tilde{\beta}'}\ (\beta', \alpha)$$

$$\beta = \beta'' + \epsilon S^*_{\tilde{\alpha}'}\ (\beta'', \alpha') + \epsilon \overset{(1)}{S}_{\tilde{\alpha}}\ (\beta', \alpha) + \epsilon^2 \overset{(2)}{\mathfrak{D}}_{\tilde{\alpha}}\ (\beta', \alpha)$$

(III-32)

where

$$\alpha' = \alpha'' - \epsilon S^*_{\tilde{\beta}''}\ (\beta'', \alpha')$$

$$\beta' = \beta'' + \epsilon S^*_{\tilde{\alpha}'}\ (\beta'', \alpha')$$

In the definitions of what constituted long-period and/or short-period variations, the Delaunay variables were used explicitly (see Eqs. (III-30) and (III-10)). If the von Zeipel technique is carried out for the Delaunay variables, then it is found that $P''(\equiv \beta'')$ are constants, the "secular" Hamiltonian is a function of P'' only and the coordinates $Q''(\equiv \alpha'')$ have constant rates. If one is dealing in any other canonical variables, for example the Poincaré set x, then x'' are defined to be the same functions of y'' as x are of y.

[†] The function S^* can be chosen to be identical to Brouwer's long-period generating function considered as a function of w (see Ref. 10).

In this section it has been established that the generating functions of transformations (III-8) and (III-29) and their partial derivatives are bounded. Note that Eqs. (III-8) constitute transcendental equations for w' which may be written in the form

$$\alpha' = \alpha + \epsilon \mathcal{L}_1(\beta',\alpha)$$
$$\beta' = \beta + \epsilon \mathcal{L}_2(\beta',\alpha) \tag{III-33}$$

From the general bounds on w one has

$$|\beta| \leq B \tag{III-34}$$

The functions $\mathcal{L}_i(\beta',\alpha)$ depend on trigonometric functions of α. Suppose now that β is bounded within some region R, and specifically, that β is bounded away from the boundary of R by at least $\epsilon A/1-\epsilon K$ where A and K are defined by

$$|\mathcal{L}_2(\beta,\alpha)| \leq A$$
$$\text{and} \quad |\mathcal{L}_2(\beta_i,\alpha) - \mathcal{L}_2(\beta_j,\alpha)| \leq K\,|\beta_i - \beta_j| \text{ for all } \beta_i,\beta_j \text{ in } R \tag{III-35}$$

Assume further that $\epsilon K < 1$; this in effect imposes a restriction on how close the energy and angular momentum may be to zero. Now assume the following iterative algorithm for computing the primed variables β':

$$\beta'_{n+1} = \beta + \epsilon \mathcal{L}_2(\beta'_n,\alpha)$$
$$\text{with} \quad \beta'_o \triangleq \beta \tag{III-36}$$

then

$$|\beta'_1 - \beta| \leq \epsilon A$$
$$|\beta'_2 - \beta'_1| \leq \epsilon K|\beta'_1 - \beta|, \quad \text{since } \beta'_1 \text{ is also in } R$$
$$|\beta'_3 - \beta'_2| \leq \epsilon K|\beta'_2 - \beta'_1|, \quad \text{since } \beta'_2 \text{ is also in } R \tag{III-37}$$
$$\vdots$$
$$|\beta'_n - \beta'_{n-1}| \leq \epsilon K|\beta'_{n-1} - \beta'_{n-2}|$$

and $$|\beta_2' - \beta| \leq |\beta_2' - \beta_1'| + |\beta_1' - \beta| \leq |\beta_1' - \beta|(1 + \epsilon K)$$

$$|\beta_3' - \beta| \leq |\beta_3' - \beta_2'| + |\beta_2' - \beta| \leq |\beta_1' - \beta|(1 + \epsilon K + (\epsilon K)^2)$$

$$\vdots$$

$$|\beta_n' - \beta| \leq |\beta_1' - \beta| \sum_{j=0}^{n-1} (\epsilon K)^j \tag{III-38}$$

or $$|\beta_n' - \beta| \leq \epsilon A \sum_{j=0}^{n-1} (\epsilon K)^j$$

Taking the limit

$$\lim_{n \to \infty} |\beta_n' - \beta| \leq \epsilon A \lim_{n \to \infty} \sum_{j=0}^{n-1} (\epsilon K)^j$$

$$\lim_{n \to \infty} |\beta_n' - \beta| \leq \epsilon A/1-\epsilon K \tag{III-39}$$

A similar argument can be applied to the long-period transformation (III-29) to deduce that β'' will remain (sufficiently) close to β' . As a consequence of the above, one has

$$|\alpha' - \alpha| = \epsilon|\mathcal{L}_1(\beta',\alpha)| \leq \epsilon A_1 \tag{III-40}$$

Also, note that the iterative procedure converges, i.e.

$$|\beta_{n+1}' - \beta_n'| \leq \epsilon A(\epsilon K)^n \tag{III-41}$$

so that as $n \to \infty$, $|\beta_{n+1}' - \beta_n'| \to 0$ if $\epsilon K < 1$.

IV. ERROR BOUNDS FOR THE AXISYMMETRIC PROBLEM

All of the information necessary to derive formal error bounds has been given in Sections II and III. One can proceed then in a straightforward manner to derive the bounds utilizing known theorems from the theory of differential equations. However, it turns out that because of the nature of the differential equations, the bounds obtainable in this manner prove to be unsatisfactory for time intervals of the order of $1/\epsilon$. This fact is a natural consequence of the existence of a rapidly rotating phase in the governing system of differential equations; however, since only one such phase appears in the case of satellite motion, one can circumvent the difficulty by appealing to a known integral of the motion. In this section the conventional method of error analysis will be presented first and then the extension to the large time intervals will be given for the problem with an axisymmetric potential.

A. BOUNDS FOR SMALL TIME INTERVALS

In order to simplify the following presentation, some new notation will be introduced at this point. If A and B denote n-dimensional vectors then $|A-B|$ will denote the matrix of absolute values of the component differences of A and B i.e.:

$$|A-B| \triangleq \begin{bmatrix} |A_1-B_1| \\ |A_2-B_2| \\ \vdots \\ |A_n-B_n| \end{bmatrix} \tag{IV-1}$$

The governing differential equations for the "secular" variables were given as (Eq. (III-31))

$$\dot{w}'' = \Phi_o \frac{\partial}{\partial \tilde{w}'} [\mathcal{K}''(w'',\epsilon) + \epsilon^3 \psi(w'',\epsilon)] \tag{IV-2}$$

from which the approximate state vector w''_A is defined by

$$\begin{aligned} \dot{w}''_A &= \Phi_o \frac{\partial}{\partial \tilde{w}''_A} \mathcal{K}''(w''_A,\epsilon) \\ w''_A(0) &= w''(0) \end{aligned} \tag{IV-3}$$

For convenience, Eqs. (IV-2) and (IV-3) can be rewritten in the form

$$\begin{aligned} \dot{w}'' &= \Lambda(w'',\epsilon) + \epsilon^3 \Psi(w'',\epsilon) \\ \dot{w}''_A &= \Lambda(w''_A,\epsilon) \end{aligned} \tag{IV-4}$$

Since the functions $\frac{\partial}{\partial \tilde{w}''} \mathcal{K}''$ and $\frac{\partial}{\partial \tilde{w}''_A} \mathcal{K}''$ satisfy a Lipschitz condition on the domain of definition of $w(t)$, it follows that

$$| \Lambda (w'',\epsilon) - \Lambda(w''_A,\epsilon) | \leq \underline{k} |w'' - w''_A| \triangleq \underline{k}m'' \tag{IV-5}$$

where $\underline{k}$ is an $n \times n$ matrix if m'', the matrix of absolute values of the component differences of $w'' - w''_A$, is $n \times 1$. The particular form of Eq. (IV-5) was chosen since a vector function, say $\Lambda(w'')$, satisfies a Lipschitz condition on w'' if and only if each of its components $\Lambda_i(w'',t)$ does. Since the constants may be different, the use of the matrix of Lipschitz constants $\underline{k}$ can afford a more precise bound than that usually provided by the norm $\|w'' - w''_A\|$.

As a consequence of (IV-5) one can immediately write

$$\frac{d}{dt} m'' \leq \underline{k}m'' + \epsilon^3 \Psi(w'',\epsilon) \tag{IV-6}$$

where from a priori bounds on $w(t)$

$$|\Psi(w'',\epsilon)| \leq W(C) \tag{IV-7}$$

Hence

$$\frac{d}{dt} m'' - \underline{k}m'' \leq \epsilon^3 W \tag{IV-8}$$

which is readily integrated to give

$$\begin{gathered} m'' \leq m''(0) \exp \underline{k}t + \epsilon^3 W\underline{k}^{-1} [\exp \underline{k}t - I] \\ 0 \leq t \leq T \ , \ t_o \overset{\Delta}{=} 0 \end{gathered} \tag{IV-9}$$

However, since it was assumed that $w_A''(0) = w''(0)$, the initial error $m''(0) = 0$ and

$$m'' = |w'' - w_A''| \leq \epsilon^3 W\underline{k}^{-1} [\exp \underline{k}t - I] \tag{IV-10}$$

At this point, several difficulties of (IV-10) can be pointed out. The bounds (IV-10) prove to be unsatisfactory for Delaunay variables for small eccentricity and/or inclination since $\underline{k}$ contains the factors $\frac{1}{e}$ and $\frac{1}{\sin i}$. If w is taken to be the Hill set z, the zero eccentricity difficulty is removed. Although the zero inclination singularity remains, for many purposes the Hill variables are a convenient set to use due to the relatively simple expressions for the periodic variations of the in-plane coordinates (see Vagners, Ref. 10). Taking w to be the Poincaré set x, satisfactory behavior is assured for both zero eccentricity and inclination. A much more serious difficulty occurs if one wishes to examine the bounds for time intervals of the order $1/\epsilon$. Expansion of (IV-10) yields, for "small" time intervals

$$\epsilon^3 W\underline{k}^{-1} [\exp \underline{k}t - I] = \epsilon^3 Wt + \epsilon^3 W \sum_{j=2}^{\infty} \frac{\underline{k}^{j-1} t^j}{j!} \tag{IV-11}$$

However, for time intervals of order $1/\epsilon$, the bound (IV-10) becomes very large i.e., behaves like $\exp 1/\epsilon$.

Assuming now that m'' is at most $O(\epsilon)$ (i.e., bound (IV-10) is satisfactory) then one can complete the analysis by including the periodic terms. If this is done, the total approximate solution of interest here is written as

$$w_c = w''_A + \epsilon\gamma(w''_A) \qquad \text{(IV-12)}$$

with $\gamma(w''_A)$ giving the first order periodic parts of w as defined by eqs. (III-8) and (III-29) with the generating functions considered as functions of the double primed variables. Equations (III-32) can be written in the form

$$w = w'' + \epsilon\gamma(w'') + \epsilon^2\zeta(w'',\epsilon) \qquad \text{(IV-13)}$$

then

$$\begin{aligned} |w - w_c| &= |w'' + \epsilon\gamma(w'') + \epsilon^2\zeta(w'',\epsilon) - w''_A - \epsilon\gamma(w''_A)| \leq \\ &|w'' - w''_A| + \epsilon|\gamma(w'') - \gamma(w''_A)| + \epsilon^2|\zeta(w'',\epsilon)| \end{aligned} \qquad \text{(IV-14)}$$

Since $\epsilon|\gamma(w'') - \gamma(w''_A)|$ gives the error in the first order periodic terms of the solution and m'' is $O(\epsilon)$, then the term contributes error of second order. Thus the effect of the last two terms of (IV-14) can be combined into one second order term $\epsilon^2 Z$ to account for all periodic errors of the solution. The error bound for "small" time intervals, assuming exact initial conditions, assumes the form

$$|w - w_c| \leq \epsilon^3 W\underline{k}^{-1}\,[\exp \underline{k}t - I] + \epsilon^2 Z \qquad \text{(IV-15)}$$

or, effectively,

$$\epsilon^3 Wt + \epsilon^2 Z \qquad \text{(IV-16)}$$

The difficulty of the above bounds for $t \sim 1/\epsilon$ is a direct consequence of the existence of a rapidly rotating phase in the dynamical system. In the treatment of systems with rapidly rotating phases by the method of averaging, the governing equations for these phases are considered separately. The general result obtained then is that the error is $O(\epsilon)$ for $t \sim 1/\epsilon$ rather than $O(\epsilon^2)$ as one would expect from the truncations performed i.e., truncation of $O(\epsilon^2)$ periodic and $O(\epsilon^3)$ secular terms. In the following, such separation will be effected and, by appealing to known integrals, the bounds will be derived for all variables to $O(\epsilon^2)$ for $t \sim 1/\epsilon$.

B. EXTENDED TIME ERROR BOUNDS

The following analysis will be carried out for the Poincaré variables explicitly utilizing known results for the Delaunay variables and their rates. The secular Hamiltonian was defined from the von Zeipel procedure as being a function of the Delaunay momenta P'' only (to second order), hence in the Poincaré variables one writes

$$\bar{\mathcal{H}}''(x'',\epsilon) = -\frac{\mu^2}{2L''^2} + \epsilon\,\bar{\mathcal{H}}^{(1)}(x'') + \epsilon^2\mathcal{R}^{(2)}(x'') - \epsilon^3\,\varphi(x'',\epsilon) \qquad \text{(IV-17)}$$

where $\bar{\mathcal{H}}^{(1)} = \dfrac{\mu^4 R_\oplus^2}{4L''^3G''^3}\left(3\,\dfrac{H^2}{G''^2} - 1\right)$, H and G functions of x'' (Eq. (III-5)) and $\mathcal{R}^{(2)}$ is F^{**} of Brouwer considered as a function of x'' . (Explicit expressions for $\mathcal{R}^{(2)}$ in terms of Hill variables for any J_n may be found in Ref. 10, which could then be transformed to Poincaré variables if necessary.) Equation (IV-17) can be rewritten more conveniently as

$$\bar{\mathcal{H}}''(x'',\epsilon) = -\frac{\mu^2}{2L''^2} + \epsilon\,\mathcal{J}(x'',\epsilon) - \epsilon^3\varphi(x'',\epsilon) \qquad \text{(IV-18)}$$

where

$$\mathcal{F}(x'',\epsilon) = \mathcal{F}\left(L'', \eta_1''^2 + \xi_1''^2, \eta_2''^2 + \xi_2''^2, \epsilon\right) \tag{IV-19}$$

The equations for the Poincaré variable rates become

$$\dot{\lambda}'' = \frac{\mu^2}{L''^3} + \epsilon \frac{\partial}{\partial L''} [\mathcal{F}(x'',\epsilon) - \epsilon^2 \varphi(x'',\epsilon)] \tag{IV-20}$$

$$\dot{L}'' = \epsilon^3 \frac{\partial \varphi}{\partial \lambda''} (x'',\epsilon)$$ since $\mathcal{H}^{(1)}$ and $\mathcal{R}^{(2)}$ do not contain λ'' and

$$\dot{x}_R'' = \epsilon \Phi_o \frac{\partial}{\partial \tilde{x}_R''} [\mathcal{F}(x'',\epsilon) - \epsilon^2 \varphi(x'',\epsilon)] \tag{IV-21}$$

with

$$x_R'' = \begin{bmatrix} \eta_1'' \\ \eta_2'' \\ \xi_1'' \\ \xi_2'' \end{bmatrix}, \quad \Phi_o \text{ a } 4 \times 4 \text{ matrix.}$$

The approximate variables x_A'' are defined by Eqs. (IV-20) and (IV-21) with $\varphi(x'',\epsilon)$ set equal to zero and $x_A''(0) = x''(0)$. Consider first the differential equations for η_1'' and ξ_1'':

$$\begin{aligned} \dot{\eta}_1'' &= \epsilon \frac{\partial}{\partial \xi_1''} \mathcal{F}(x'',\epsilon) - \epsilon^3 \frac{\partial \varphi}{\partial \xi_1''} \\ \dot{\xi}_1'' &= -\epsilon \frac{\partial}{\partial \eta_1''} \mathcal{F}(x'',\epsilon) + \epsilon^3 \frac{\partial \varphi}{\partial \eta_1''} \end{aligned} \tag{IV-22}$$

Recall that $\mathcal{F}(x'',\epsilon)$ is given by Eq. (IV-19) so that with

$$N_1 \triangleq 2 \frac{\partial \mathcal{F}}{\partial (\eta_1''^2 + \xi_1''^2)}$$

a bounded quantity, one obtains

$$\dot{\eta}_1'' = \epsilon N_1 \xi_1'' - \epsilon^3 \frac{\partial}{\partial \xi_1''} \varphi$$

$$\dot{\xi}_1'' = - \epsilon N_1 \eta_1'' + \epsilon^3 \frac{\partial}{\partial \eta_1''} \varphi \tag{IV-23}$$

Hence

$$\frac{1}{2} \frac{d}{dt} \left[\xi_1''^2 + \eta_1''^2 \right] = \epsilon^3 \left[\xi_1'' \frac{\partial \varphi}{\partial \eta_1''} - \eta_1'' \frac{\partial \varphi}{\partial \xi_1''} \right] \tag{IV-24}$$

The approximate solution x_A'' of interest is given by $\varphi = 0$; thus from the boundedness of φ (and its partial derivatives)

$$\left| \left(\eta_1''^2 - \eta_{1A}''^2 \right) + \left(\xi_1''^2 - \xi_{1A}''^2 \right) \right| = \left| \Delta \left(\xi_1''^2 + \eta_1''^2 \right) \right| \leq \epsilon^3 M_o t \tag{IV-25}$$

Here, as well as in the following discussion, the extended time interval will be taken as $t \sim 1/\epsilon$ so that

$$\left| \Delta \left(\xi_1''^2 + \eta_1''^2 \right) \right| \leq \epsilon^2 M_o \tag{IV-26}$$

The reader may prefer to think of the time interval as defined by $nt \sim 1/\epsilon$ where n is taken to be the (suitable) mean motion. For mathematical convenience, the definition $t \sim 1/\epsilon$ will be used.

Now, rewrite Eqs. (IV-23) as a single complex equation ($j \triangleq \sqrt{-1}$):

$$\dot{\xi}_1'' + j\dot{\eta}_1'' = j\epsilon N_1 \left(\xi_1'' + j\eta_1'' \right) + \epsilon^3 \left[\frac{\partial \varphi}{\partial \eta_1''} - j \frac{\partial \varphi}{\partial \xi_1''} \right] \tag{IV-27}$$

and the approximate equations as

$$\dot{\xi}_{1A}'' + j\, \dot{\eta}_{1A}'' = j\, \epsilon\, N_{1A} \left(\xi_{1A}'' + j\eta_{1A}'' \right) \tag{IV-28}$$

since

$$N_1 = N_1 \left(\xi_1''^2 + \eta_1''^2 , \xi_2''^2 + \eta_2''^2 , L'', \epsilon \right)$$

Difference Eqs. (IV-27) and (IV-28) to get

$$\Delta\dot{\xi}_1'' + j\Delta\dot{\eta}_1'' = j\epsilon\Delta N_1(\xi_1'' + j\eta_1'') + j\epsilon N_{1A}(\Delta\xi_1'' + j\Delta\eta_1'') + \epsilon^3\left[\frac{\partial\varphi}{\partial\eta_1''} - j\,\frac{\partial\varphi}{\partial\xi_1''}\right] \tag{IV-29}$$

From Eq. (IV-26):

$$\Delta N_1 = |N_1 - N_{1A}| \leq M_1\Delta\left(\xi_1''^2 + \eta_1''^2\right) \leq \epsilon^2 M_2$$

and (IV-29) thus becomes, with the aid of an integrating factor,

$$\left|\frac{d}{dt}\,[(\Delta\xi_1'' + j\Delta\eta_1'')e^{-j\epsilon N_{1A}t}]\right| \leq \epsilon^3\left|e^{-j\epsilon N_{1A}t}\left[\frac{\partial\varphi}{\partial\eta_1''} - j\,\frac{\partial\varphi}{\partial\xi_1''} + jM_2\xi_1'' - M_2\eta_1''\right]\right| \tag{IV-30}$$

Then, since the right hand side of (IV-30) is bounded, it follows that

$$\left|(\Delta\xi_1'' + j\Delta\eta_1'')e^{-j\epsilon N_{1A}t}\right| \leq \epsilon^3 M_3 t = \epsilon^2 M_3 \quad t \sim 1/\epsilon \tag{IV-31}$$

but $|e^{-j\epsilon N_{1A}t}| = 1$ so

$$|\Delta\xi_1'' + j\Delta\eta_1''| \leq \epsilon^2 M_3 \ . \tag{IV-32}$$

From similar arguments, it follows that for ξ_2'' and η_2''

$$|\Delta\xi_2'' + j\Delta\eta_2''| \leq \epsilon^2 M_4 \ , \quad t \sim 1/\epsilon \tag{IV-33}$$

Also, from the differential Eq. (IV-20)

$$|L'' - L_A''| \equiv \left|\epsilon^3 \int \frac{\partial\varphi}{\partial\lambda''}\,dt\right| \leq \epsilon^3 M_5 t = \epsilon^2 M_5 \tag{IV-34}$$

The remaining coordinate λ'' causes some difficulty, since with $|L'' - L_A''|$ known to $0(\epsilon^2)$, a straightforward analysis of the $\dot{\lambda}''$

149

equation gives $|\lambda'' - \lambda''_A|$ only to $O(\epsilon)$ for $t \sim 1/\epsilon$. In order to obtain bounds for λ'' consistent with those of the other coordinates, one must appeal to the knowledge of an exact integral for the axi-symmetric problem. In effect, one can re-define the mean motion as introduced by Brouwer (Ref. 3), who wrote

$$n_o \triangleq \frac{\mu^2}{L'^3} \tag{IV-35}$$

and hence, with ℓ_1 and ℓ_2 functions of L', G'', H only

$$\dot{\ell}'' = n_o[1 + \epsilon\ell_1 + \epsilon^2\ell_2] + O(\epsilon^3) \tag{IV-36}$$

Recall that $\dot{\lambda}''$ was defined by Eq. (IV-20), which in a more explicit form is given by

$$\dot{\lambda}'' = \frac{\mu^2}{L''^3}\left\{1 - \epsilon\,\frac{3}{2}\,k_1 L''^{-1} - \epsilon\,\frac{3\mu^2 R_\oplus^2}{4G''^4}\left[5\left(\frac{H''}{G''}\right)^2 - 1 - 2\,\frac{H''}{G''}\right] + \epsilon^2\delta_2\right\} - \epsilon^3\,\frac{\partial\varphi}{\partial L''} \tag{IV-37}$$

where

$$k_1 = \frac{\mu^2 R_\oplus^2}{2G''^3}\left(3\left(\frac{H''}{G''}\right)^2 - 1\right) \quad \text{with} \quad G'' = G''(x'') \quad \text{and} \quad H'' = H''(x'')$$

$$\delta_2 = \frac{L'^3}{\mu^2}\,\frac{\partial\mathcal{R}^{(2)}}{\partial L''}\,, \quad \text{a bounded quantity}$$

Define now a new constant $\hat{a}$ by $\left(\text{with } \mathcal{R}^* = \dfrac{2L''^2}{\mu^2}\,\mathcal{R}^{(2)}\right)$

$$\frac{\mu}{2\hat{a}} = -\,\mathcal{H} = -\,\mathcal{H}'' = \frac{\mu^2}{2L''^2}\,[1 - \epsilon k_1 L''^{-1} - \epsilon^2\mathcal{R}^* + \epsilon^3\varphi_1] \tag{IV-38}$$

and a new "mean motion" by

$$\hat{n} = \mu^{\frac{1}{2}}\left(\frac{1}{\hat{a}}\right)^{3/2} = \frac{\mu^2}{L''^3}\left[1 - \epsilon k_1 L''^{-1} - \epsilon^2\mathcal{R}^* + \epsilon^3\varphi_1\right]^{3/2} \tag{IV-39}$$

or expanding

$$\hat{n} = \frac{\mu^2}{L''^3}\left[1 - \epsilon\,\frac{3}{2}\,k_1 L''^{-1} - \epsilon^2\,\frac{3}{2}\,\mathcal{R}^* + \epsilon^2\,\frac{3}{8}\,k_1^2 L''^{-2} + \epsilon^3 \varphi_1^*(x'',\epsilon)\right] \quad \text{(IV-40)}$$

Thus

$$\dot{\lambda}'' = \hat{n} - \epsilon\,\frac{\mu^2}{L''^3}\left\{\frac{3\mu^2 R_\oplus}{4G''^4}\left[5\left(\frac{H''}{G''}\right)^2 - 1 - 2\,\frac{H''}{G''}\right] - \epsilon\left[\delta_2 + \frac{3}{2}\,\mathcal{R}^* - \frac{3}{8}\,k_1^2 L''^{-2}\right]\right\} - \epsilon^3\left[\frac{\mu^2}{L''^3}\,\varphi_1^*(x'',\epsilon) + \frac{\partial\varphi}{\partial L''}\right] \quad \text{(IV-41)}$$

Again, the approximate λ''_A is defined by (IV-41) with $\frac{\partial\varphi}{\partial L''} = 0$ and $\varphi_1^* = 0$. From the exact known integral, $H'' = H$ in Eq. (IV-41) and from the definition of G'':

$$G'' = L'' - \frac{\xi_1''^2 + \eta_1''^2}{2} \quad \text{(IV-42)}$$

then from Eqs. (IV-32) and (IV-34)

$$|G'' - G''_A| \le \epsilon^2 M_6 \qquad t \sim 1/\epsilon \quad \text{(IV-43)}$$

so that finally

$$|\lambda'' - \lambda''_A| \le \epsilon^2 M_7 \qquad t \sim 1/\epsilon \quad \text{(IV-44)}$$

If one is interested in orbits with non-zero eccentricity and/or inclination (i.e., $e \gg \epsilon$, $\sin i \gg \epsilon$) then the above analysis can be carried out analogously for the Delaunay and/or Hill variables. In particular, for the Delaunay variables, the rapid phase is ℓ'' and an equation similar to (IV-41) (but somewhat simpler in form) results for $\dot{\ell}''$. Due to choice of g and h as the other two coordinates, the ϵ term of (IV-41) is found to disappear. (Of course, the functions $\delta_2, \mathcal{R}^*$ and φ are different than for the Poincaré variables).

C. THE INITIALIZATION PROBLEM

At this point the relevance of the above results to the so-called initialization problem may be noted. The two primary uses of an analytic (artificial satellite) orbit theory are orbit determination by fitting to observational data and orbit prediction from some initial state vector. In the case of orbit determination, the mean (double primed) variables are obtained to high accuracy by fitting to observational data. This accuracy depends on the number and quality of the data points. In this application, the question of initial value errors does not arise.

The initialization problem may be defined as follows: given some initial radius and velocity vectors, compute a satellite ephemeris for some extended time interval via an analytic theory. The initial radius and velocity, and hence the instantaneous elements, are assumed to be known exactly. Analytic theories are usually formulated so that certain constants of the solution are mean elements, for example L', G'' and H in the Brouwer theory, instead of initial values. Thus from the known set of instantaneous elements, the mean elements must be formed by subtracting out the periodic variations. Since one is considering a first order theory, the mean elements thus defined will be in error by $O(\epsilon^2)$. It can be noted here that a numerical iteration procedure has been applied to the determining equations (Cain Ref. 15, Arsenault, et al Ref. 11) which are written as

$$\begin{aligned} Q &= Q'' - \epsilon\left[S^{(1)}_{\tilde{P}'}(P',Q) + S^{(*)}_{\tilde{P}''}(P'',Q')\right] \\ P &= P'' + \epsilon\left[S^{(1)}_{\tilde{Q}}(P',Q) + S^{(*)}_{\tilde{Q}'}(P'',Q')\right] \end{aligned} \qquad \text{(IV-45)}$$

Such a procedure can, of course, only remove that second order error that arises from considering the S functions to be functions of the instantaneous elements (P,Q), but still cannot account for the truncated second order terms. Thus from an accuracy point of view, such iteration procedures are of dubious value, since as shown by Eq. (IV-37) (with ϵ^3 terms truncated) the error in λ_A'', or equivalently ℓ_A'', will still grow as $\epsilon^2 t$ from the zero order term. The other variables of either x'' or y'' do not present any problem since their rates are either zero or multiples of ϵ, so that an initial value error of $O(\epsilon^2)$ will grow as $\epsilon^3 t$ giving results consistent with the expected accuracy of the truncated theory.

With the algorithm suggested by the analysis of subsection IV-B, the initialization difficulty can be resolved. As noted, for all variables except the rapidly rotating phase, the use of mean elements defined by instantaneous value minus the periodic terms (considered as functions of the instantaneous elements) will lead to no difficulty. The necessary initialization procedure for λ_A'' is given by Eqs. (IV-38), (IV-39) and (IV-41). The numerical value of $\hat{n}$ is known exactly from instantaneous $\mathcal{H}$ and the remaining terms of (IV-41) have at least an ϵ multiplier. For the Delaunay variables, one rewrites Eq. (IV-36) with n'' the mean motion defined by

$$n'' \triangleq \dot{\ell}'' = \frac{\mu^2}{L'^3}\left[1 - \epsilon\,\frac{3}{2}\,k_1 L'^{-1} + \epsilon^2\delta_2\right] + 0\ (\epsilon^3) \qquad \text{(IV-46)}$$

so that, with use of energy†

† The explicit expression for $\mathcal{R}^*$ is identical to $\frac{2L'^3}{\mu}\,\dot{F}^{**}$. [See Vagners (Ref. 10) where it is given as $\frac{2L'^3}{\mu^2}\,[\overline{\psi}_2 + \overline{F}_2]$].

$$n'' = \dot{\ell}'' = \hat{n} + \epsilon^2 \frac{\mu^2}{L'^3} \left[\frac{3}{2} \mathcal{R}^* + \delta_2 - \frac{3}{8} k_1^2 L'^{-2} \right]_{t = 0} + \epsilon^3 \text{ terms} \qquad \text{(IV-47)}$$

in which in the ϵ^2 terms one replaces the double primed variables with the instantaneous elements. The new mean motion $\hat{n}$ is again given by (IV-38) and (IV-39). All terms appearing in the brackets of (IV-47) are functions already known from the general theory.

V. THE ASYMMETRIC POTENTIAL FIELD

If one includes the longitude dependent terms (tesseral harmonics) in the gravitational potential, some modifications to the analysis of Section IV are necessary. The additional terms in the Hamiltonian are

$$\mathcal{H}_T = \frac{\mu}{r} \sum_{n=2}^{\infty} \sum_{m=0}^{n} J_{n,m} \left(\frac{R_\oplus}{r} \right)^n P_n^m(\sin \beta) \cos m(\lambda - \lambda_{n,m}) \qquad \text{(V-1)}$$

where $J_{n,m}$ $\lambda_{n,m}$ are constants with $J_{n,m} \sim 0(J_2^2)$

$$\lambda = h + \tan^{-1}(\cos i \tan u) - \omega_\oplus t$$

$\omega_\oplus$ the angular velocity of the Earth

(Time is measured from an instant when the right ascension of Greenwich is zero.)

In this discussion, $\mathcal{H}_T$ will be considered first as a function of the Delaunay variables $\mathcal{H}_T(L,G,H,\ell,g,h - \omega_\oplus t)$. To remove the explicit time dependence, define a new canonical variable as $h^* = h - \omega_\oplus t$ conjugate to H with the associated Hamiltonian given by

$$K = \mathcal{H} - \omega_\oplus H \qquad \text{(V-2)}$$

where $\mathcal{H}$ now is the original Hamiltonian including both zonal and tesseral harmonic effects. Since time is not present explicitly in K, it is a constant of the motion.

Following the von Zeipel procedure, "remove" all of the periodic parts of the extended Hamiltonian K via a suitable generating function, defined here up to second order (since $\mathcal{H}_T$ is second order in ϵ) so that the new variables become

$$L = L' + \epsilon \frac{\partial S^{(1)}}{\partial \ell} + \epsilon^2 \frac{\partial S^{(2)}}{\partial \ell} + \frac{\partial S_T}{\partial \ell}$$

$$\ell' = \ell + \epsilon \frac{\partial S^{(1)}}{\partial L'} + \epsilon^2 \frac{\partial S^{(2)}}{\partial L'} + \frac{\partial S_T}{\partial L'} \qquad \text{(V-3)}$$

$$H = H' + \frac{\partial S_T}{\partial h^*}$$

Equations similar to those above hold for the other variables. Note that H now contains fluctuations but that these are of second order. As before, $S^{(1)}$ and $S^{(2)}$ are chosen to cancel all zonal short-period terms up to, and including, second order. Thus one is left with (omitting the ϵ^3 function for the time being)

$$K' = -\frac{\mu^2}{2L'^2} + \epsilon\, \bar{\mathcal{K}}_1 + \epsilon^2 \mathcal{R}^{(2)} - \omega_\oplus H' + \frac{\mu^2}{L'^3} \frac{\partial S_T}{\partial \ell} - \omega_\oplus \frac{\partial S_T}{\partial h^*} - \mathcal{K}_T \qquad \text{(V-4)}$$

where

$$\mathcal{K}_T = \tilde{\mathcal{K}} + \bar{\mathcal{K}}_T = \sum_{k_2} \sum_{k_1} \sum_{m} A_{k_2 k_1 m}(a,e,i)\cos(k_1\ell + k_2 g + mh^* + \text{phase}) \qquad \text{(V-5)}$$

in which $\tilde{\mathcal{K}}_T$ includes all terms with $k_1 \neq 0$, the short-period part of $\mathcal{K}_T$, and $\bar{\mathcal{K}}_T$ gives the daily fluctuations from h^*, $k_1 = 0$. Choose S_T so that

$$-\frac{\mu^2}{L'^3} \frac{\partial S_T}{\partial \ell} + \omega_\oplus \frac{\partial S_T}{\partial h^*} = -\mathcal{K}_T \qquad \text{(V-6)}$$

Strictly speaking, (V-6) should be written as

$$-\frac{\mu^2}{L'^3} \frac{\partial S_T}{\partial \ell} + \omega_\oplus \frac{\partial S_T}{\partial h^*} + (\omega_\oplus - \dot{\Omega}'') \frac{\partial S_T^*}{\partial h^{*\prime}} =$$

$$-\tilde{\mathcal{K}}_T(L', G', H', \ell, g, h^*) - \bar{\mathcal{K}}_T(L'', G'', H'', g', h^{*\prime}) \qquad \text{(V-7)}$$

however, (V-6) is accurate to $O(\epsilon^2)$ since $\dot{\Omega}''$ is $O(\epsilon)$. Solving by inspection, with $n' \triangleq \dfrac{\mu^2}{L'^3}$,

$$S_T = \sum_{k_2} \sum_{k_1} \sum_{m} \frac{m A_{k_1 k_2 m}}{k_1 n' - m\omega_\oplus} \sin(k_1 \ell' + k_2 g' + m h^{*'} + \text{phase}) \qquad \text{(V-8)}$$

In Eq. (V-8) (as well as (V-6)) one can use the primed variables or instantaneous variables with $O(\epsilon^3)$ error. Clearly, there exist orbits for which n' is commensurable with $\omega_\oplus$ and hence a particular $(k_1 n' - m\omega_\oplus)$ goes to zero. These are the so-called tesseral resonance cases and will not be considered here. Introduction of S_T in the above manner leaves

$$K'' = -\frac{\mu^2}{2L'^2} + \epsilon\, \overline{\mathcal{H}}_1 + \epsilon^2 \mathcal{R}^{(2)} - \omega_\oplus H'' - \epsilon^3 \varphi^{\ddagger} \qquad \text{(V-9)}$$

where the ϵ^3 function is now included, and is different from the ϵ^3 function of $\mathcal{H}''$ of the axisymmetric problem.

The instantaneous elements may be written as a sum of the "secular" (double-primed) part and the periodic parts. In particular, with H'' a constant,

$$H = H'' + H_{S.P} + H_{M.P} \qquad \text{(V-10)}$$

From (V-3) and (V-6) it follows that

$$\omega_\oplus H'' = \omega_\oplus H + \mathcal{H}_T - n' \frac{\partial S_T}{\partial \ell} \qquad \text{(V-11)}$$

thus, from (V-8)

$$n' \frac{\partial S_T}{\partial \ell} = \sum_{k_2} \sum_{k_1} \sum_{m} \frac{k_1 n'}{k_1 n' - m\omega_\oplus} A_{k_1 k_2 m} \cos(k_1 \ell' + k_2 g' + m h^{*'} + \text{phase}) \qquad \text{(V-12)}$$

Combining (V-12) with (V-5) one obtains

$$\mathcal{H}_T - n' \frac{\partial S_T}{\partial \ell} = \sum_{k_2} \sum_{k_1} \sum_m \left(1 - \frac{k_1 n'}{k_1 n' - m\omega_\oplus}\right) A_{k_1 k_2 m} \cos(k_1 \ell' + k_2 g' + mh^{*\prime} + \text{phase}) \quad \text{(V-13)}$$

$$= - \sum_{k_2} \sum_{k_1} \sum_m \frac{m\omega_\oplus}{k_1 n' - m\omega_\oplus} A_{k_1 k_2 m} \cos(k_1 \ell' + k_2 g' + mh^{*\prime} + \text{phase})$$

The constancy of K and $\mathcal{H}''$ to 2nd order implies (because of (V-11)) that

$$- \mathcal{H}_T + n' \frac{\partial S_T}{\partial \ell} = \text{tesseral fluctuation of } \mathcal{H} = \int_A \frac{\partial \mathcal{H}}{\partial t} dt \quad \text{(V-14)}$$

where $\int_A \frac{\partial \mathcal{H}}{\partial t} dt$ denotes a specific second order approximation to the indefinite integral, considering only ℓ' and h^* as time varying. This may be checked by forming

$$\frac{\partial \mathcal{H}}{\partial t} = \frac{\partial \mathcal{H}_T}{\partial t} = \sum_{k_2} \sum_{k_1} \sum_m m\omega_\oplus A_{k_1 k_2 m} \sin(k_1 \ell' + k_2 g' + mh^{*\prime} + \text{phase}) \quad \text{(V-15)}$$

If one assumes that, to the order necessary here,

$$k_1 \ell' + mh^{*\prime} \approx [k_1 n' + m\omega_\oplus]t \quad \text{(V-16)}$$

then conclusion (V-14) follows.

For discussion of the error bounds, the formalism of Section IV can be retained to a large extent. Due to the absence of secular tesseral terms, only the φ function of the secular Hamiltonian will change and hence the bound on that term will be different. It follows then, that the error bounds on x'', with the exception of λ'', are derived in exactly the same manner as before with different values for the constants M_i . The derivation of an error bound on λ'' is not as simple as before, since for the asymmetrical problem the two separate integrals of energy and polar component of angular momentum no longer exist.

From the extended Hamiltonian (V-14) one finds

$$\dot{\lambda}'' = \frac{\mu^2}{L'^3}\left\{1 - \epsilon\frac{3}{2}k_1 L'^{-1} - \epsilon\frac{3\mu^2 R_\oplus}{4G''^4}\left[5\left(\frac{H''}{G''}\right)^2 - 1 - 2\frac{H''}{G''}\right]\right.$$

$$\left. + \epsilon^2\delta_2\right\} - \epsilon^3\frac{\partial\varphi^\ddagger}{\partial L'} \qquad \text{(V-17)}$$

The object again is to obtain an expression for $\frac{\mu^2}{L'^3}(1 - \epsilon\frac{3}{2}k_1 L'^{-1})$ accurate up to, and including, second order. Since K is a constant of the motion,

$$K = \mathcal{H} - \omega_\oplus H = K'' = \mathcal{H}'' - \omega_\oplus H'' \qquad \text{(V-18)}$$

so that

$$\mathcal{H}'' = \mathcal{H} - \omega_\oplus(H - H'') \qquad \text{(V-19)}$$

From general theory for time dependent Hamiltonians

$$\frac{d\mathcal{H}}{dt} = \frac{\partial\mathcal{H}}{\partial t} \qquad \text{(V-20)}$$

and

$$\mathcal{H} - \int\frac{\partial\mathcal{H}}{\partial t}\,dt = \text{constant} \qquad \text{(V-21)}$$

The constant is related to quantity $\mathcal{H}''$, which is also a constant to second order as defined by the von Zeipel procedure. In fact, if one chooses a particular approximate evaluation of the indefinite integral, as in (V-14), then

$$\mathcal{H}'' = \mathcal{H} - \int_A \frac{\partial\mathcal{H}}{\partial t}\,dt \qquad \text{(V-22)}$$

Note that as defined by (V-19), $\mathcal{H}''$ is known only to second order with third order secular terms (from H''). After times of order $1/\epsilon$ this constitutes an error in $\mathcal{H}''$ of second order secular and thus finally to first order in λ'' if L' is defined through $\mathcal{H}''$. Using the relation (V-22) avoids this difficulty, since it turns out that the third order evaluation of the specific integral $\int_A \partial\mathcal{H}/\partial t \; dt$ yields terms that are still third order after $t \sim 1/\epsilon$, with tesseral resonance situations still ruled out. This may be verified by considering the first order variations of the variables in $\partial\mathcal{H}/\partial t$ and noting that $m \neq 0$ in (V-15).

Next, define as before (with $\mathcal{H}''$ defined by (V-22))

$$\frac{\mu}{2\tilde{a}} = -\mathcal{H}'' \tag{V-23}$$

and the mean motion by

$$\tilde{n} = [\mu/\tilde{a}^3]^{\frac{1}{2}} \tag{V-24}$$

so that

$$\tilde{n} = \frac{\mu^2}{L'^3}\left[1 - \epsilon\,\frac{3}{2}\,k_1 L'^{-1} + \epsilon^2\,\frac{3}{8}\,k_1^2 L'^{-2} - \epsilon^2\,\frac{3}{2}\,\mathcal{R}^*\right] + \epsilon^3 \varphi_1^{\ddagger}(x'',\epsilon) \tag{V-25}$$

and finally, the $\dot{\lambda}''$ equation

$$\dot{\lambda}'' = \tilde{n} - \epsilon\,\frac{\mu^2}{L'^3}\left\{3\,\frac{\mu^2 R_\oplus^2}{4G''^4}\left[5\left(\frac{H''}{G''}\right)^2 - 1 - 2\,\frac{H''}{G''}\right] - \epsilon\left[\delta_2 + \frac{3}{2}\,\mathcal{R}^* - \frac{3}{8}\,k_1^2 L'^{-2}\right]\right\}$$

$$- \;\epsilon^3\left[\frac{\partial\varphi^{\ddagger}}{\partial L'} + \varphi_1^{\ddagger}\right] \tag{V-26}$$

The approximate solution λ_A'' is again defined by (V-26) with the $\varphi^{\ddagger}$ functions equal to zero. The argument for obtaining the error estimate follows as before.

The algorithm for computing the correct initial value of the mean motion $\tilde{n}$ now involves the evaluation of the integral $\int_A \partial\mathcal{H}/\partial t \; dt$. This may be done by a suitable expansion on eccentricity; one such evaluation is given in the Appendix.

VI. NUMERICAL VERIFICATION

Equation (V-26) then provides an algorithm for computing the correct initial conditions (to the order of accuracy demanded by the general solution) in the case when the total potential of the Earth is taken into account. The algorithm includes the (suitable) evaluation of the indefinite integral $\int_A \frac{\partial \mathcal{H}_T}{\partial t}\,dt$. It is of interest to obtain numerical verification of the general accuracy theory of Section V . The explicit expression for $\int_A \frac{\partial T_r}{\partial t}\,dt$ has been derived earlier by Vagners (Ref. 16), and was subsequently incorporated into the Lockheed Closed Form Orbit Determination Program (Ref. 17). This program utilizes a complete first order analytic solution that is equivalent to the extended Brouwer solution. (The extended Brouwer solution is taken to include J_2 short-period, J_2^2 and general J_N long-period, $J_{n,m}$ medium period (daily) effects and all second order secular effects not accounting for tesseral resonances (c.f. Giacaglia (Ref. 18), and Garfinkel (Ref. 19).) The Lockheed solution is due to Small (Ref. 12), and Vagners (Ref. 16).

Since the error in the mean anomaly (or equivalently λ'') is directly related to in-track position error, the simplest test of overall accuracy is to compare the in-track, cross-track and radial positions as predicted by the analytic solution and numerical integration of coordinates. Since the time intervals of interest are of the order $1/\epsilon$, the comparison was performed over a seven day interval. The test orbit was of 2000 nautical mile altitude and circular. Such an orbit, not including the results of Section V (roughly a 200 foot error in the semi-major axis), resulted in a 200 mile in-track error†. After "tuning" the mean motion with the

† The comparison study was carried out to determine the effects of inclusion (or omission) of tesseral harmonic short period terms in the semi-major axis. The energy had already been incorporated in the formulation of the axisymmetric problem.

energy, the secular error was decreased to 900 feet, which is an order ϵ decrease as demonstrated by the theory of Section V. The comparison is shown in Fig. 1, where it can be seen that the periodic errors and the secular error are now of the same order of magnitude i.e., $O(\epsilon^2)$. The cross-track and radial errors are periodic and have amplitudes of $\pm$ 120 feet and $\pm$ 350 feet, respectively, for the comparison orbit.

VII. CONCLUDING REMARKS

In the previous sections error bounds were derived specifically for the Brouwer procedure using the Poincaré variables. From the general theory, an algorithm was derived for the correct computation of the initial conditions for the Brouwer theory. It is then of interest to note the relevance of the results of this paper to other orbit theories, and also to present the computation of coordinates from the Poincaré elements.

Insofar as the first item is concerned, exactly equivalent errors are to be expected from any complete first order theory provided that care is taken in establishing the correct mean elements for that theory. A complete first order theory is defined as one that includes the first order periodic and second order secular influences of any harmonic. This distinction is necessary if one wishes to compare theories for prediction of orbits from a fit to observational data or for prediction from an initial state vector, i.e. the initial value problem. For example, the theories of Kyner (Ref. 1), Petty and Breakwell (Ref. 21), including a time equation carried only to first order secular terms, would give satisfactory results if applied to orbit prediction from a fit to data. However, for the initial value problem, these theories would prove unsatisfactory (giving ϵ errors for time $t \sim 1/\epsilon$). The latter difficulty could be remedied if the time equation (or its equivalent) would be carried out to include second order secular effects and an energy algorithm used to calibrate the mean motion. The theory of Small (Ref. 12, Ref. 16) is a complete first order theory and includes the correct algorithm for computation of initial conditions.

With more or less difficulty, any theory appearing in the literature may be analyzed in a manner analogous to that given in this paper and equivalent results obtained. In each case, the energy will have to be used to establish the mean motion (or the constant rate of the fast variable) to second order, unless complete second order periodic expressions for the semi-major axis are available. The questions of error bounds become more difficult if one admits orbits at critical inclination and/or orbits at resonance with the tesseral harmonics. Such orbits are excluded from the general class investigated in this report and remain the topic of future investigations.

The last point to consider is the computation of the coordinates from the Poincaré elements in which most of the theory of this paper was developed. In terms of conventional orbit elements,

$$\lambda = M + \omega + \Omega \qquad L = \sqrt{\mu a}$$

$$\eta_1 = [2\sqrt{\mu a}\,(1 - \sqrt{1-e^2})\,]^{\frac{1}{2}}\cos(\omega + \Omega) \qquad \xi_1 = [2\sqrt{\mu a}\,(1 - \sqrt{1-e^2})]^{\frac{1}{2}}\sin(\omega + \Omega)$$

$$\eta_2 = [2\sqrt{\mu p}\,(1 - \cos i)]^{\frac{1}{2}}\cos\Omega \qquad \xi_2 = [2\sqrt{\mu p}\,(1 - \cos i)]^{\frac{1}{2}}\sin\Omega$$

where M is the mean anomaly and $p = a(1 - e^2)$. The remaining elements were defined in Section III, Eq. (III-1). Known the time t one can find $\lambda, \xi_1, \eta_1, \xi_2, \eta_2$ and L, then compute

$$e\cos(\omega + \Omega) = \frac{\eta_1}{L^{\frac{1}{2}}}\left[1 - \frac{\xi_1^2 + \eta_1^2}{4L}\right]^{\frac{1}{2}}$$

$$e\sin(\omega + \Omega) = \frac{\xi_1}{L^{\frac{1}{2}}}\left[1 - \frac{\xi_1^2 + \eta_1^2}{4L}\right]^{\frac{1}{2}} \qquad \text{(VII-2)}$$

An iterative procedure yields Δ, $e\cos f$, $e\sin f$ defined by

$$\Delta_n = 2\,\frac{\tan^{-1}(e\sin f)_n}{1+\sqrt{1-e^2}+(e\cos f)_n} + \frac{\sqrt{1-e^2}\,(e\sin f)_n}{1+(e\cos f)_n} \qquad \text{(VII-3)}$$

$$(e\cos f)_{n+1} = (e\cos(\omega+\Omega))\cos(\lambda+\Delta_n) + (e\sin(\omega+\Omega))\sin(\lambda+\Delta_n)$$

$$(e\sin f)_{n+1} = (e\cos(\omega+\Omega))\sin(\lambda+\Delta_n) - (e\sin(\omega+\Omega))\cos(\lambda+\Delta_n)$$

where $\sqrt{1-e^2} = 1 - \dfrac{\xi_1^2+\eta_1^2}{2L}$

So that the radius is given by

$$r = \frac{L^{5/2}(1-e^2)}{D} \qquad \text{(VII-4)}$$

where

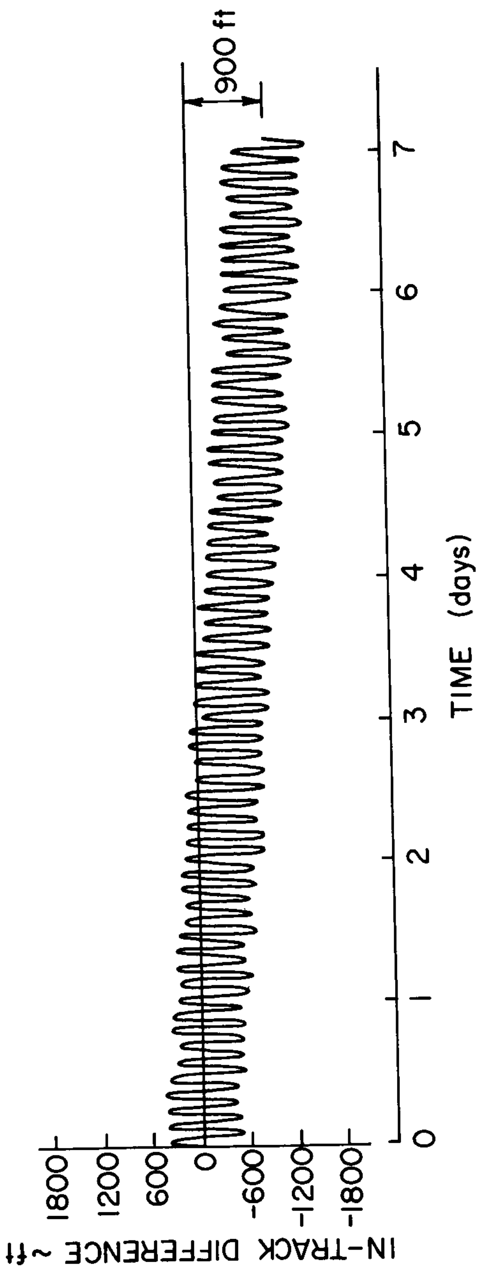
$$D = \mu\left\{ L^{\frac{1}{2}} + \left[1 - \frac{\xi_1^2+\eta_1^2}{4L}\right]^{\frac{1}{2}} [\eta_1\cos(\lambda+\Delta) + \xi_1\sin(\lambda+\Delta)]\right\}$$

and the cartesian coordinates x,y,z by

$$x = \frac{L^{3/2}(1-e^2)^{\frac{1}{2}}}{2D}\,[(2L-\xi_1^2-\eta_1^2-\xi_2^2)\cos(\lambda+\Delta) + \eta_2\xi_2\sin(\lambda+\Delta)]$$

$$y = \frac{L^{3/2}(1-e^2)^{\frac{1}{2}}}{2D}\,[\xi_2\eta_2\cos(\lambda+\Delta) + (2L-\xi_1^2-\eta_1^2-\eta_2^2)\sin(\lambda+\Delta)]$$

$$z = \frac{L^{3/2}(1-e^2)^{\frac{1}{2}}}{2D}\,(4L-2\xi_1^2-2\eta_1^2-\xi_2^2-\eta_2^2)^{\frac{1}{2}}\,[\eta_2\sin(\lambda+\Delta) - \xi_2\cos(\lambda+\Delta)]$$



COMPARISON OF NUMERICAL INTEGRATION
AND ANALYTIC THEORY ORBIT PREDICTION

FIGURE 1

(Reproduced from LMSC Document No. 579774
Courtesy of Lockheed Missiles and Space Company)

REFERENCES

1. Kyner, W. T.,"A Mathematical Theory of the Orbits About an Oblate Planet," Technical Report, Department of Mathematics, University of Southern California, Feb. 1963.

2. Bogoliubov, N. N. and Mitropolskii, Y. A.,"Asymptotic Methods In The Theory of Nonlinear Oscillations," Gordon and Breach, New York, 1961.

3. Brouwer, D., "Solution of the Problem of Artificial Satellite Theory Without Drag," A. J. 64, 1959.

4. Kozai, Y., "Second-Order Solution of Artificial Satellite Theory Without Air Drag," A. J. 67, 1962.

5. Garfinkel, B., "The Orbit of a Satellite of an Oblate Planet," A. J. 64, 1959.

6. Garfinkel, B. and McAllister, G. T. "The Zonal Harmonic Perturbations of an Artificial Satellite," A. J. 69, 1964.

7. Burshtein, E. L. and Solov'ev, L. T., "Hamiltonian of Averaged Motion," Soviet Physics-Doklady, 6, 1962.

8. Morrison, J. A., "The Generalized Method of Averaging and the von Zeipel Method," AIAA Preprint No. 65-687, Sept. 1965.

9. Izsak, I., "A Note on Perturbation Theory," A. J. 68, 1963.

10. Vagners, J., "Direct Determination of Satellite Coordinate Perturbations by the von Zeipel Method," Sudaer Report No. 252, Nov. 1965.

11. Arsenault, J. L., Enright, J. D., and Purcell, C., "General Perturbation Techniques for Satellite Orbit Prediction Study," Vols. I-II, Technical Documentary Report No. AL-TDR-64-70, April 1964 AD No. 437-475 and No. 437-476.

12. Small, H. W., "Satellite Motion Around an Oblate Planet," AIAA Preprint No. 63-393, Aug. 1963.

13. Macmillan, W. D., "Periodic Orbits About an Oblate Spheroid," Trans. AMS 11 55 (1910).

14. Poritsky, H., "Motion of a Satellite Around an Oblate Earth," A. J. 67, 1962.

15. Cain, B., "Determination of Mean Elements for Brouwer's Satellite Theory," A. J. 67, Aug. 1962.

16. Vagners, J., "Influences of the Tesseral and Sectorial Harmonics of Earth's Gravitational Potential on Satellite Orbits," Lockheed

Missiles and Space Company Document LMSC/577084, Sunnyvale, Calif. 22 June 1964.

17. Warmke, C. T., "Lockheed Closed Form Orbit Determination Program (LOCFOD)," Lockheed Missiles and Space Company Document LMSC/577615, Sunnyvale, Calif. 23 Sept. 1964.

18. Giacaglia, G. E. O., "The Influence of High-Order Zonal Harmonics on the Motion of an Artificial Satellite Without Drag," A. J. 69, 1964.

19. Garfinkel, B., "Tesseral Harmonic Perturbations of an Artificial Satellite," Private Communication from author, to be published.

20. Yoo, R. J. and Kulick, R. M., "Comparison of Ephemeris Generation Using Cowell Integration and Closed Form Solution of the Equations of Motion," Lockheed Missiles and Space Company Document LMSC/579774, 12 July 1965.

21. Petty, C. M. and Breakwell, J. V., "Satellite Orbits About a Planet With Rotational Symmetry," Journal of the Franklin Institute, Vol. No. 4, Oct. 1960.

22. Kaula, W. M., "Analysis of Gravitational and Geometric Aspects of Geodetic Utilization of Satellites," Geophysical Journal 5, 1961.

APPENDIX

Explicit Evaluation of $\int_A \frac{\partial \mathcal{H}_T}{\partial t}\, dt$

For the evaluation of the initial value problem, the indefinite integral $\int_A \frac{\partial \mathcal{H}_T}{\partial t}\, dt$ must be evaluated or, equivalently, the generating function S_T must be found. It will be assumed here that the integrand is given by (V-15) and the integration will be carried out in conventional variables.

The following expressions prove useful:

$$P_n^m(\sin\beta) = \frac{\cos^m\beta}{2^n n!} \sum_{\zeta=0}^{\nu} \frac{(2n-2\zeta)!}{(n-m-2\zeta)!} \binom{n}{\zeta} (-1)^{\zeta} \sin^{n-m-2\zeta} i \, \sin^{n-m-2\zeta} u \qquad \text{(A-1)}$$

where

$$\nu = \begin{cases} \dfrac{n-m}{2} & \text{for } n-m \text{ even} \\[2ex] \dfrac{n-m-1}{2} & \text{for } n-m \text{ odd} \end{cases}$$

$$\cos m(\lambda - \lambda_{n,m}) = \sum_{s=0}^{m} \binom{m}{s} \frac{\cos^{m-s} u \, \sin^{s} u \, \cos^{s} i}{\cos^m\beta} (-1)^{\xi} \times \qquad \text{(A-2)}$$

$$[\gamma_1 \cos m(h^* - \lambda^*) + \gamma_2 \sin m(h^* - \lambda^*)]$$

where

$$\xi = \begin{cases} s/2 & \text{if } s \text{ is even; } \gamma_1 = 1, \ \gamma_2 = 0 \\[2ex] \dfrac{s+1}{2} & \text{if } s \text{ is odd; } \gamma_1 = 0, \ \gamma_2 = 1 \end{cases}$$

$\lambda^* = \alpha_o + \lambda_{n,m}$ with α_o the right ascension of Greenwich at a base time t_o

$$\frac{1}{r^{n+1}} = \sum_{p=0}^{n+1} \binom{n+1}{p} \frac{e^p}{2^p(1 - e^2)^{p+1}} \frac{1}{a^{n+1}} \sum_{q=0}^{p} \binom{p}{q} \cos(p - 2q)(u - \omega) \quad \text{(A-3)}$$

and

$$\sin^j u \cos^k u = \sum_{c=0}^{j} \sum_{d=0}^{k} \binom{j}{c} \binom{k}{d} \frac{(-1)^{\ell}}{2^{j+k}} [\delta_1 \cos(j + k - 2c - 2d)u$$

$$+ \delta_2 \sin(j + k - 2c - 2d)u] \quad \text{(A-4)}$$

where

$$\ell = \begin{cases} j + c + j/2 & \text{if } j \text{ is even; } \delta_1 = 1, \ \delta_2 = 0 \\ j + c + \dfrac{j+1}{2} & \text{if } j \text{ is odd; } \delta_1 = 0, \ \delta_2 = 1 \end{cases}$$

At this point, the assumptions under which $\int_A \frac{\partial \mathcal{H}}{\partial t} \, dt$ will be integrated may be stated. The inclination angle i and the eccentricity e will be taken as constants. Since no appreciable difficulty is incurred thereby, the following will be adopted

$$\omega = \omega_o + \omega' u$$

$$h^* = \Omega_o + \Omega' u - \omega_{\oplus} t$$

with

$$\omega' = \frac{3}{2} \epsilon \left(\frac{R_{\oplus}}{p}\right)^2 (2 - 5/2 \sin^2 i) \quad \text{(A-5)}$$

$$\Omega' = -\frac{3}{2} \epsilon \left(\frac{R_{\oplus}}{p}\right)^2 \cos i$$

The last item is the central angle-time relationship. Since the integrand is (essentially) now a function of u, one would prefer to integrate with respect to u. To the first approximation

$$du = \tilde{n} \, dt + O(e\epsilon^2) \quad \text{(A-6)}$$

so that the contribution of the $J_{n,m}$ term is given by

$$E_{n,m} \int \sin^{n-m-2\zeta+s} u \cos^{m-s} u \, [\gamma_1 \sin m(h^* - \lambda^*) - \gamma_2 \cos m(h^* - \lambda^*)] \times$$

$$\cos(p - 2q)(u - \omega) \, du$$

with‡

$$E_{n,m} = \frac{m\omega_\oplus}{\tilde{n}} \mu J_{n,m} \frac{R_\oplus^n}{a^{n+1}} \frac{1}{2^n n!} \sum_{\zeta=0}^{\nu} \sum_{p=0}^{n+1} \sum_{q=0}^{p} \sum_{s=0}^{m} \frac{(2n-2\zeta)!}{(n-m-2\zeta)!} \binom{n}{s} \binom{n+1}{p} \times$$

$$\binom{p}{q} \binom{m}{s} \left(-1\right)^{\zeta+\xi} \frac{e^p}{2^p(1 - e^2)^{p+1}} \times \cos^s i \sin^{n-m-2\zeta} i \tag{A-7}$$

or with $\tilde{h} = h^* - \lambda^*$,

$$E^*_{n,m} \int \cos(p - 2q)(u - \omega)[\gamma_1 \sin m\tilde{h} - \gamma_2 \cos m\tilde{h}][\delta_1 \cos(j + k - 2c - 2d)u$$

$$+ \delta_2 \sin(j + k - 2c - 2d)u]du$$

with

$$E^*_{n,m} = E_{n,m} \sum_{c=0}^{j} \sum_{d=0}^{k} \binom{j}{c} \binom{k}{d} \frac{(-1)^\ell}{2^{j+k}} \ .$$

Then let

$$B_o = -\omega_o(p - 2q)$$

$$B_1 = (1 - \omega')(p - 2q)$$

$$B_2 = j + k - 2c - 2d$$

$$B_3 = m(\Omega_o - \lambda^*)$$

$$B_4 = -m\left(-\Omega' + \frac{\omega_\oplus}{\tilde{n}}\right)$$

so that the integrand becomes

‡ If one prefers, the F and G functions of Kaula, Ref. 22, may be used instead.

$$\int \cos(B_o + B_1 u)[\delta_1 \cos B_2 u + \delta_2 \sin B_2 u][\gamma_1 \sin(B_3 + B_4 u) - \gamma_2 \cos(B_3 + B_4 u)]du \tag{A-8}$$

The following non-zero combinations arise in the above integral:

$$\begin{aligned}
I_1 &= \int \cos(B_o + B_1 u) \cos B_2 u \sin(B_3 + B_4 u)du \\
I_2 &= -\int \cos(B_o + B_1 u) \cos B_2 u \cos(B_3 + B_4 u)du \\
I_3 &= \int \cos(B_o + B_1 u) \sin B_2 u \sin(B_3 + B_4 u)du \\
I_4 &= -\int \cos(B_o + B_1 u) \sin B_2 u \cos(B_2 + B_4 u)du
\end{aligned} \tag{A-9}$$

which can all be evaluated explicitly to give

$$I_1 = \frac{1}{4}\left\{ \frac{1}{B_1-B_4-B_2} \cos[B_3-B_o+(B_2+B_4-B_1)u] + \frac{1}{B_1-B_4+B_2} \cos[B_o-B_3+(B_2+B_1-B_4)u] \right.$$

$$\left. + \frac{1}{B_2-B_4-B_1} \cos[B_3+B_0+(B_1+B_4-B_2)u] - \frac{1}{B_4+B_1+B_2} \cos[B_3+B_o+(B_2+B_4+B_1)u] \right\}$$

$$I_2 = -\frac{1}{4}\left\{ \frac{1}{B_2-B_4+B_1} \sin[B_o-B_3+(B_2-B_4+B_1)u] - \frac{1}{B_1-B_4-B_2} \sin[B_3-B_o+(B_2+B_4-B_1)u] \right.$$

$$\left. - \frac{1}{B_2-B_4-B_1} \sin[B_3+B_o+(B_1+B_4-B_2)u] - \frac{1}{B_4+B_1+B_2} \sin[B_3+B_o+(B_1+B_2+B_4)u] \right\}$$

$$I_3 = \frac{1}{4}\left\{ \frac{1}{B_1-B_4-B_2} \sin[B_3-B_o+(B_2+B_4-B_1)u] + \frac{1}{B_2-B_4+B_1} \sin[B_o-B_3+(B_2+B_1-B_4)u] \right.$$

$$\left. - \frac{1}{B_2-B_4-B_1} \sin[B_o+B_3+(B_1+B_4-B_2)u] - \frac{1}{B_4+B_1+B_2} \sin[B_3+B_o+(B_2+B_4+B_1)u] \right\}$$

$$I_4 = -\frac{1}{4}\left\{ \frac{1}{B_4-B_1-B_2} \cos[B_o-B_3+(B_2+B_1-B_4)u] + \frac{1}{B_4-B_2+B_1} \cos[B_o+B_3+(B_1+B_4-B_2)u] + \right.$$

$$+ \frac{1}{B_1-B_4-B_2} \cos[B_3-B_o+(B_2+B_4-B_1)u] - \frac{1}{B_4+B_1+B_2} \cos[B_3+B_o+(B_2+B_4+B_1)u]\Bigg\}$$

In the expressions of Ref. 16, it was assumed that i, ω, Ω and r were constants and in the test case the orbit was circular. The extensions of the above development cause no difficulty other than increasing the number of terms. However, any improvement of accuracy for non-zero eccentricity orbits is difficult to assess due to the approximation of Eq. (A-6). In order to define the error remaining as of order $e^2\epsilon^2$ one must include the e terms in (A-6).

PRECEDING PAGE BLANK NOT FILMED.

An Investigation of High Eccentricity Orbits About Mars*

by J. V. Breakwell and R. D. Hensley
Stanford University
Stanford, California

N67-29377

SUMMARY

Possible long-period fluctuations in the radius of pericenter, r_p, of an orbiter of Mars due to the solar gravitational field are investigated. The study is restricted to a "region of interest" defined by 6000 km $\geq r_p \geq$ 4000 km. Eleven different "critical" orbital inclinations are found for which the long-period fluctuations in eccentricity and inclination contain terms of vanishing frequency suggesting very large amplitudes. A closer analysis of these resonant situations near a critical inclination is accomplished by transforming the Hamiltonian into that of a simple pendulum problem. Maximum variations in the eccentricity, and hence radius of pericenter, are then obtained and curves of maximum change in radius of pericenter versus eccentricity plotted.

I. INTRODUCTION

For many of the problems in space-flight mechanics it is necessary to find an orbit which satisfies certain boundary or mission conditions. In this investigation, which considers the problem of an artificial satellite about the planet Mars, the size and shape of the orbit are dictated by two practical considerations: (1) the savings in fuel obtained by transferring from an interplanetary trajectory to an elliptical orbit and (2) the desire to pass as close as possible to the surface of the planet. This last consideration introduces the notion of the radius of pericenter r_p which, with the eccentricity

*NASA Grant NsG-133-61

e, will determine the size and shape of any orbit. The mission requirements can be expressed in terms of these two parameters, e.g., the orbits must lie in a range of interest for which $e > 0.5$ and $4000 \text{ km} \leq r_p \leq 6000 \text{ km}$ (see Fig. 1). Of major concern to the mission planner is the variation of r_p, over a long period of time, caused by a fluctuation in e.

This study investigates the possible long-period fluctuations in the radius of pericenter of an orbiter of Mars due to the solar gravitational field. For those orbits which satisfy the mission requirements the "secular" rotations of the orbital plane and the major-axis orientation (argument of pericenter ω) due to Mars' oblateness dominate those due to the solar field. On the other hand, these oblateness secular rotations are substantially slower than Mars' motion around the Sun. This latter consideration permits averaging of the perturbations not only over orbital revolutions but also over the Martian year to obtain equations describing "long-period" rates of change of inclination I, eccentricity, longitude of the node Ω, and argument of pericenter ω, the last two elements being driven mainly by oblateness. The long-period rates of inclination and eccentricity are sinusoidal in certain linear combinations of ω and Ω leading to fluctuations in inclination and eccentricity as sums of easily computable sinusoidal terms of known frequency. There are 11 different "critical" orbital inclinations where the frequency of a sinusoidal term vanishes leading to a very large amplitude.

An analysis of these resonant situations near a "critical" inclination I_c is accomplished by transforming the Hamiltonian into that of a simple pendulum problem.[1] And, as in the pendulum analysis, it is possible to obtain a phase-space contour describing the total libration of the system which, for this study, determines the maximum excursions of I and e near I_c. Once the maximum variation in e is known the maximum change in r_p $(\delta r_p)_{max}$ can be calculated, for those orbits meeting the mission conditions, and curves showing $(\delta r_p)_{max}$ versus e plotted.

As noted above, in this study the radius of pericenter lies in the range

$$4000 \text{ km} \leq r_p < 6000 \text{ km}$$

Introducing the semimajor axis a and the eccentricity e, a "region of interest" is defined by

$$\frac{4000}{(1-e)} \leq a \leq \frac{6000}{(1-e)}$$

This region is shown on Figure 1.

II. THE DISTURBING FUNCTION

In this investigation the motion of a body in the gravitational field of the planet Mars will be studied under the assumption that the only perturbing forces acting are the oblateness of Mars and the solar gravitational field. Ignoring, of course, the effects of the (small) mass of the orbiting body, then the equations of motion in a Mars-centered coordinate frame may be written in the form

$$\ddot{\bar{r}} = \nabla(R_M + R_S) \tag{1}$$

where $\bar{r}$ is the radius vector of the satellite, ∇ the gradient operator and R_M, R_S the potentials of Mars and the Sun respectively. The solar perturbing potential R_S may be written as

$$R_S = \mu_s \left[\frac{1}{|\bar{r}_s - r|} - \frac{\bar{r} \cdot \bar{r}_s}{r_s^3} \right] \tag{2}$$

where μ_s is the gravitational constant of the Sun and $\bar{r}_s$ the radius vector from Mars to the Sun. The potential of Mars can be expressed as

$$R_M = \frac{\mu_m}{r} + R_m = \frac{\mu_m}{r} \left[1 + \frac{J_2}{2} \left(\frac{R_e}{r} \right)^2 (1 - 3 \sin^2 \delta) + \text{higher order terms} \right] \tag{3}$$

with J_2 the second harmonic coefficient

R_e equatorial radius of Mars

δ the planetocentric latitude

R_m the perturbing potential due to the figure of Mars

The higher order terms indicated in (3) will henceforth be omitted and R_m will be taken as the perturbing potential due to J_2. Then the disturbing function as referred to henceforth will be given by $R_m + R_s$. Equation (3) then is easily rewritten in terms of the satellite's inclination I, (measured from the Martian equator), true anomaly v and the argument of pericenter ω as

$$R_M = \frac{\mu_m}{r}\left[1 + \frac{J_2}{2}\left(\frac{R_e}{r}\right)^2\left(1 - \frac{3}{2}\sin^2 I + \frac{3}{2}\sin^2 I \cos 2(v+\omega)\right)\right] \tag{4}$$

Since the effect of the terms periodic in the satellite's mean anomaly M are not significant, R_M can be averaged over 2π on M. Transform, therefore, the true anomaly v to the mean anomaly M by the differential equation

$$\frac{dv}{dM} = \frac{a^2}{r^2}\sqrt{1-e^2} \tag{5}$$

and express the radius as a function of the semimajor axis a, eccentricity and true anomaly,

$$r = \frac{a(1-e^2)}{1 + e\cos v} \tag{6}$$

so that by obtaining from (5) and (6) the relations of Tisserand[2]

$$\overline{\left(\frac{a}{r}\right)^3} = \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{a}{r}\right)^3 dM = \frac{1}{(1-e^2)^{3/2}}$$

$$\overline{\left(\frac{a}{r}\right)^3 \sin 2v} = \overline{\left(\frac{a}{r}\right)^3 \cos 2v} = 0 \tag{7}$$

the averaged disturbing potential due to J_2 becomes

$$\bar{R}_m = \frac{1}{2\pi} \int_0^{2\pi} R_m \, dM = \frac{\mu_m J_2 R_e^2}{2a^3(1-e^2)^{3/2}} \left(1 - \frac{3}{2} \sin^2 I\right) \tag{8}$$

Note that $\bar{R}_m$ is also independent of the argument of pericenter ω.

Next, consider the expression for the solar perturbation potential R_s. If B denotes the angle between $\bar{r}$ and $\bar{r}_s$, then the first term of Eq. (2) can be developed in a series of Legendre polynomials as follows:

$$\frac{1}{|\bar{r}_s - \bar{r}|} = \frac{1}{r_s} \left\{1 + \frac{r}{r_s} \cos B + \left(\frac{r}{r_s}\right)^2 \left(\frac{3}{2} \cos^2 B - \frac{1}{2}\right) + \dots\right\} \tag{9}$$

so that, together with the fact that

$$\cos B = \frac{\bar{r} \cdot \bar{r}_s}{r r_s} \tag{10}$$

the potential R_s becomes

$$R_s = \frac{\mu_s}{2r_s^3} r^2 \left[3\left(\frac{\bar{r} \cdot \bar{r}_s}{r r_s}\right)^2 - 1\right] \tag{11}$$

From Fig. 2 the unit vectors $\bar{e}_r$ and $\bar{e}_s$, in the radial direction of the satellite and the Sun, may be written in component form as

$$\bar{e}_r = \frac{\bar{r}}{r} = \begin{bmatrix} \cos \Omega \cos \theta - \sin \Omega \cos I \sin \theta \\ \sin \Omega \cos \theta + \cos \Omega \cos I \sin \theta \\ \sin I \sin \theta \end{bmatrix} \tag{12}$$

and, with $C \triangleq \cos$ and $S \triangleq \sin$, a convention used henceforth:

$$\bar{e}_s = \frac{\bar{r}_s}{r_s} = \begin{bmatrix} C_{\Lambda_s} C_{L_s} \\ C_{L_s} S_{\Lambda_s} \\ S_{L_s} \end{bmatrix} \tag{13}$$

where Ω is the right ascension of the satellite's ascending node

θ is the satellite central angle

Λ_s is the right ascension of the Sun

L_s is the planetocentric latitude of the Sun

Thus the Sun's disturbing function becomes

$$R_s = \frac{\mu_s r^2}{2r_s^3}\left\{3\left[C_{L_s} C_{\Omega-\Lambda_s} C_\theta - S_{\Omega-\Lambda_s} C_{L_s} C_I S_\theta + S_{L_s} S_I S_\theta\right]^2 - 1\right\} \tag{14}$$

R_s is next written in terms of the satellite's true anomaly via

$$\begin{aligned}
C_\theta &= C_{v+\omega} \\
S_\theta &= S_{v+\omega} \\
C_\theta^2 &= C_v^2 C_\omega^2 - C_v S_v S_{2\omega} + S_v^2 S_\omega^2 \\
S_\theta C_\theta &= S_v C_v C_{2\omega} + \frac{1}{2} C_v^2 S_{2\omega} - \frac{1}{2} C_v^2 S_{2\omega}
\end{aligned} \tag{15}$$

so that

$$R_s = \frac{\mu_s r^2}{2r_s^3}\left\{\frac{3}{2}\,\mathcal{A}\, C_{2(v+\omega)} + 3\,\mathcal{B}\, S_{2(v+\omega)} + \frac{3}{2}\,\mathcal{C} - 1\right\} \tag{16}$$

where

$$\mathcal{A} = \left[\left(C_{L_s} C_{\Omega-\Lambda_s}\right)^2 - \left(C_I C_{L_s} S_{\Omega-\Lambda_s} - S_{L_s} S_I\right)^2\right]$$

$$\mathcal{B} = \left[C_{L_s} S_{L_s} S_I C_{\Omega-\Lambda_s} - C_{L_s}^2 C_I C_{\Omega-\Lambda_s} S_{\Omega-\Lambda_s}\right]$$

$$\mathcal{C} = \left[C_{L_s}^2 C_{\Omega-\Lambda_s}^2 + \left(C_I C_{L_s} S_{\Omega-\Lambda_s} - S_{L_s} S_I\right)^2\right]$$

Holding the Sun's variables L_s and Λ_s constant and integrating R_s with respect to time, the short-period terms depending on θ can be removed. The average then is

$$\bar{R}_s = \frac{1}{2\pi} \int_0^{2\pi} R_s \, dM \tag{17}$$

which may be readily integrated by introducing the eccentric anomaly E through

$$dM = (1 - e\,C_E)\, dE \tag{18}$$

and the relationships

$$C_v = \frac{(C_E - e)}{1 - e\,C_E}$$

$$C_{2v} = \frac{(2 - e^2)\,C_E^2 - 2e\,C_E + 2e^2 - 1}{(1 - e\,C_E)^2} \tag{19}$$

Thus

$$\begin{aligned}
\bar{R}_s &= \frac{1}{2\pi} \int_0^{2\pi} R_s (1 - e\,C_E)\, dE \\
&= \frac{\mu_s\, a^2}{4\pi\, r_s^3} \int_0^{2\pi} \frac{(1 - e^2)^2 (1 - e\,C_E)}{(1 + e\,C_v)^2} [1,\ C_{2v}]\, dE \\
&= \frac{\mu_s\, a^2}{4\pi\, r_s^3} \int_0^{2\pi} (1 - e\,C_E)^3 \left[1,\ \frac{(2 - e^2)\,C_E^2 - 2e\,C_E + 2e^2 - 1}{(1 - e\,C_E)^2}\right] dE
\end{aligned} \tag{20}$$

Integration yields

$$\frac{\mu_s a^2}{4\pi r_s^3} \int_0^{2\pi} (1 - e C_E)^3 \, dE = \frac{\mu_s a^2}{4 r_s^3} (2 + 3 e^2)$$

$$\frac{\mu_s a^2}{4\pi r_s^3} \int_0^{2\pi} \left[(2 - e^2) C_E^2 - 2e C_E + 2e^2 - 1\right](1 - e C_E) \, dE = \frac{5}{4} \mu_s \frac{a^2 e^2}{r_s^3} \tag{21}$$

so that the Sun's disturbing function with the short-period terms removed is

$$\overline{R}_s = \frac{\mu_s a^2}{4 r_s^3} \left[\frac{15}{2} e^2 \mathcal{A} C_{2\omega} + 15 e^2 \mathcal{B} S_{2\omega} + (2 + 3 e^2)\left(\frac{3}{2} \mathcal{C} - 1\right)\right] \tag{22}$$

or, substituting for $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$:

$$\begin{aligned} \overline{R}_s = \frac{\mu_s a^2}{4 r_s^3} \Bigg\{ & \frac{15}{2} e^2 \left[C_{L_s}^2 C_{\Omega-\Lambda_s}^2 - \left(C_I C_{L_s} S_{\Omega-\Lambda_s} - S_{L_s} S_I\right)^2\right] C_{2\omega} \\ & + 15 e^2 \left[C_{L_s} S_{L_s} S_I C_{\Omega-\Lambda_s} - C_{L_s}^2 C_I C_{\Omega-\Lambda_s} S_{\Omega-\Lambda_s}\right] S_{2\omega} \\ & + (2 + 3 e^2)\left[\frac{3}{2}\left[C_{L_s}^2 C_{\Omega-\Lambda_s}^2 + \left(C_I C_{L_s} S_{\Omega-\Lambda_s} - S_{L_s} S_I\right)^2\right] - 1\right]\Bigg\} \end{aligned} \tag{23}$$

The next step is to remove the medium period effects caused by the motion of the Sun. One assumes that the central angle of the Sun θ_s (see Fig. 2) varies considerably faster than either ω or Ω, and hence R_s may be averaged with respect to the motion of the Sun. The variations of ω and Ω during one revolution of the satellite are given by

$$\frac{d\omega}{d\theta} = \frac{3}{2} J_2 \left(\frac{R_e}{p}\right)^2 \left(2 - \frac{5}{2} S_1^2\right) \equiv J_2 \left(\frac{R_e}{p}\right)^2$$

$$\dot{\omega} \approx J_2 \left(\frac{R_e}{p}\right)^2 n \tag{24}$$

$$\frac{d\Omega}{d\theta} = -\frac{3}{2} J_2 \left(\frac{R_e}{p}\right)^2 C_I \approx -J_2 \left(\frac{R_e}{p}\right)^2$$

$$\dot{\Omega} \approx -J_2 \left(\frac{R_e}{p}\right)^2 n \tag{25}$$

where the mean orbital rate of the satellite is defined as $n = \sqrt{\mu_m/a^3}$ and the semi-latus rectum of the orbit p is written

$$p = a(1 - e^2) = r_p(1 + e) \tag{26}$$

Defining the mean motion of the Sun n_s to be

$$n_s = \sqrt{\frac{\mu_s}{a_s^3}} \tag{27}$$

with a_s the semi-major axis of the Martian orbit (equivalently, the apparent Sun's orbit around Mars), the condition that $\dot{\omega}$ and $\dot{\Omega}$ are (assumed) much smaller than n_s leads to

$$\frac{J_2 R_e^2 \mu_m^{1/2}}{a^{7/2} (1 - e^2)^2} \ll n_s \tag{28}$$

This condition then is satisfied throughout most of the "region of interest."

Introducing the solar coordinate θ_s and the inclination I_s (see Fig. 2) through

$$
\begin{aligned}
C_{\theta_s} &= C_{\Lambda_s} C_{L_s} \\
S_{L_s} &= S_{I_s} S_{\theta_s} \\
C_{L_s} S_{\Lambda_s} &= C_{I_s} S_{\theta_s}
\end{aligned}
\tag{29}
$$

the function $\overline{R}_s$ takes on the somewhat lengthy appearance

$$
\begin{aligned}
\overline{R}_s = & -\frac{\mu_s a^2}{4 a_s^3}\Bigg(\frac{a_s^3}{r_s^3}\Bigg)\Bigg\{(2 + 3e^2)\Bigg\{1 - \frac{3}{4}\Bigg(C_\Omega^2 C_{2\theta_s} + C_\Omega^2 + 2C_\Omega S_\Omega C_{I_s} S_{2\theta_s} \\
& - S_\Omega^2 C_{I_s}^2 C_{2\theta_s} + S_\Omega^2 C_{I_s}\Bigg) - \frac{3}{4}\Bigg[\Bigg(C_I^2 S_\Omega^2 C_{2\theta_s} + C_I^2 S_\Omega^2 - 2\Bigg(C_I^2 S_\Omega C_\Omega C_{I_s} \\
& + C_I S_I S_\Omega S_{I_s}\Bigg) S_{2\theta_s} - \Bigg(C_I^2 C_\Omega^2 C_{I_s}^2 + 2C_I S_I C_\Omega S_{I_s} + S_{I_s}^2 S_I^2\Bigg) C_{2\theta_s} \\
& + \Bigg(C_I^2 C_\Omega^2 C_{I_s}^2 + 2C_I S_I C_\Omega S_{I_s} C_{I_s} + S_I^2 S_{I_s}^2\Bigg)\Bigg)\Bigg]\Bigg\} + \frac{15}{4} e^2 \Bigg[C_I^2 S_\Omega^2 C_{2\theta_s} + C_I^2 S_\Omega^2 \\
& - 2\Bigg(C_I^2 S_\Omega C_\Omega C_{I_s} + C_I S_I S_\Omega S_{I_s}\Bigg) S_{2\theta_s} - \Bigg(C_I^2 C_\Omega^2 C_{I_s} + S_I S_{I_s}\Bigg)^2 C_{2\theta_s} \\
& + \Bigg(C_I^2 C_\Omega^2 C_{I_s} + S_I S_{I_s}\Bigg)^2 - C_\Omega^2 C_{2\theta_s} - C_\Omega^2 + 2C_\Omega S_\Omega C_{I_s} S_{2\theta_s} - S_\Omega^2 C_{I_s}^2 C_{2\theta_s} - S_\Omega^2 C_{I_s}^2\Bigg] C_{2\omega} \\
& + \frac{15}{2} e^2 \Bigg[C_I^2 S_\Omega^2 C_{2\theta_s} + C_I^2 S_\Omega^2 - 2\Bigg(C_I^2 S_\Omega C_\Omega C_{I_s} + C_I S_I S_\Omega S_{I_s}\Bigg) S_{2\theta_s} - \Bigg(C_I^2 C_\Omega^2 C_{I_s} + S_I S_{I_s}\Bigg)^2 C_{2\theta_s} \\
& + \Bigg(C_I^2 C_\Omega^2 C_{I_s} + S_I S_{I_s}\Bigg)^2 - C_\Omega^2 C_{2\theta_s} - 2C_\Omega^2 + C_\Omega S_\Omega C_{I_s} S_{2\theta_s} - S_\Omega^2 C_{I_s}^2 C_{2\theta_s} - S_\Omega^2 C_{I_s}^2\Bigg] C_{2\omega}\Bigg\}
\end{aligned}
\tag{30}
$$

As before, the following trigonometric identities are introduced:

$$C_{2\theta_s} = C_{2v_s} C_{2\omega_s} - S_{2v_s} S_{2\omega_s}$$

$$S_{2\theta_s} = S_{2v_s} C_{2\omega_s} - S_{2\omega_s} C_{2v_s} \tag{31}$$

and recalling relations 7, applied to the Sun:

$$\overline{\left(\frac{a_s}{r_s}\right)^3} = \frac{1}{(1 - e_s^2)^{3/2}}$$

$$\overline{\left(\frac{a_s}{r_s}\right)^3 C_{2v_s}} = \overline{\left(\frac{a_s}{r_s}\right)^3 S_{2v_s}} = 0 \tag{32}$$

A further simplification is introduced here by assuming that the Sun's orbit is circular so that

$$\overline{\left(\frac{a_s}{r_s}\right)^3} = 1 \tag{33}$$

The slowly varying disturbing function due to the Sun then becomes:

$$\begin{aligned}
\bar{\bar{R}}_s = \frac{n_s^2 a^2}{4} \Bigg\{ & [2 + 3e^2]\Bigg[\frac{3}{8} C_{2\Omega} S_{I_s}^2 S_I^2 + \frac{3}{2} C_I S_I S_{I_s} C_{I_s} C_\Omega \\
& + \frac{1}{4}\left(1 - \frac{3}{2} S_{I_s}^2\right)\left(3C_I^2 - 1\right)\Bigg] + \frac{15}{4} e^2 \Bigg[\frac{1}{4} S_{I_s}^2 \left(1 + C_I^2\right)\left(C_{2\Omega+2\omega} + C_{2\Omega-2\omega}\right) \\
& - C_I S_I C_{I_s} S_{I_s} \left(C_{\Omega+2\omega} + C_{\Omega-2\omega}\right) + S_I^2\left(1 - \frac{3}{2} S_{I_s}^2\right) C_{2\omega}\Bigg] \\
& - \frac{15}{2} e^2 \Bigg[\frac{1}{4} C_I S_{I_s}^2 \left(C_{2\Omega-2\omega} - C_{2\Omega+2\omega}\right) \\
& - \frac{1}{2} S_{I_s} S_I C_{I_s} \left(C_{\Omega-2\omega} - C_{\Omega+2\omega}\right)\Bigg]\Bigg\}
\end{aligned} \tag{34}$$

The total slowly varying disturbing function due to J_2 and the Sun perturbation is given thus in Keplerian elements of the satellite and the Sun by Eqs. (8) and (34).

III. LONG PERIOD VARIATIONS IN ECCENTRICITY AND INCLINATION

In the previous section the slowly varying disturbing function, periodic in only Ω and ω, was determined. The long-period changes in the orbit elements corresponding to that disturbing function can be readily determined by invoking the techniques of Hamiltonian mechanics and canonical transformations.(3) If one introduces the slowly varying Hamiltonian through

$$\bar{\mathcal{H}} = -(\bar{R}_m + \bar{\bar{R}}_s) - \frac{\mu_m^2}{2L^2} \tag{35}$$

where $L = \sqrt{\mu_m a}$,

then the canonical equations of motion in terms of (slowly varying) Delaunay elements are

$$\dot{L} = -\frac{\partial\bar{\mathcal{H}}}{\partial\ell} \qquad \dot{G} = -\frac{\partial\bar{\mathcal{H}}}{\partial g} \qquad \dot{H} = -\frac{\partial\bar{\mathcal{H}}}{\partial h}$$

$$\dot{\ell} = \frac{\partial\bar{\mathcal{H}}}{\partial L} \qquad \dot{g} = \frac{\partial\bar{\mathcal{H}}}{\partial G} \qquad \dot{h} = \frac{\partial\bar{\mathcal{H}}}{\partial H} \tag{36}$$

where $\ell = M$, $g = \omega$, $h = \Omega$ are generalized coordinates

$L = \sqrt{\mu a}$, $G = \sqrt{\mu a(1 - e^2)}$, $H = G \cos I$ are generalized momenta

The set of Eqs. (36) may be integrated if one can find a suitable canonical transformation, determined by a function $\mathcal{S}$, such that the transformed Hamiltonian is a function of the new momenta only. Such a transformation, from the old variables (L, G, H, ℓ, g, h) to a new set $(L', G', H', \ell', g', h')$, will be given by the function $\mathcal{S}$ which is assumed to be expandable in powers of the parameter J_2 ($\mathcal{S}_1$ is order J_2):

$$\mathcal{S} = \ell L' + gG' + hH' + \mathcal{S}_1(L', G', H', \ell, g, h) + \ldots \tag{37}$$

Since the canonical form of the equations of motion is preserved and the new Hamiltonian is a function of L', G', H' only, it follows that L', G', H' are constants and ℓ', g', h' are linear functions of time.[3] The old variables are related to the new by the formulas

$$L = \frac{\partial \mathcal{S}}{\partial \ell}, \quad G = \frac{\partial \mathcal{S}}{\partial g}, \quad H = \frac{\partial \mathcal{S}}{\partial h}$$

$$\ell' = \frac{\partial \mathcal{S}}{\partial L'}, \quad g' = \frac{\partial \mathcal{S}}{\partial G'}, \quad h' = \frac{\partial \mathcal{S}}{\partial H'} \tag{38}$$

In order to apply the theory to the present problem, the disturbing function must be written in terms of the Delaunay variables. It turns out to be convenient to write it as a sum of three parts, which follow from Eqs. (8), (34) and the definitions (36), and to introduce the perturbing Hamiltonian $\overline{\mathcal{H}}_p$:

$$-\overline{\mathcal{H}}_p = \overline{R}_{m+s} = \overline{\overline{R}}_{ss} + \overline{R}_m + \overline{\overline{R}}_{sp} \tag{39}$$

where the terms independent of g and h are

$$\overline{\overline{R}}_{ss} = \frac{n_s^2 L^4}{4 \mu_m^2} \left(5 - 3 \frac{G^2}{L^2}\right)\left(3 \frac{H^2}{G^2} - 1\right)$$

$$\overline{R}_m = \frac{\mu_m^4 J_2 R_e^2}{4L^3 G^3} \left(3 \frac{H^2}{G^2} - 1\right) \tag{40}$$

and those periodic in g and h

$$\overline{\overline{R}}_{sp} = \frac{n_s^2 L^4}{4 \mu_m^2} \left[\alpha_1 C_{2h} + \alpha_2 C_h + \alpha_3 C_{2g} + \alpha_4 C_{2h+2g} + \alpha_5 C_{2h-2g} - \alpha_6 C_{h+2g} + \alpha_7 C_{h-2g}\right] \tag{41}$$

The coefficients α_i are given by

$$\alpha_1 = \frac{3}{8} S_{I_s}^2 \left(5 - 3\frac{G^2}{L^2}\right)\left(1 - \frac{H^2}{G^2}\right)$$

$$\alpha_2 = \frac{3}{2} S_{I_s} C_{I_s} \left(5 - 3\frac{G^2}{L^2}\right)\left(\frac{H}{G}\right)\left(1 - \frac{H^2}{G^2}\right)^{1/2}$$

$$\alpha_3 = \frac{15}{4}\left(1 - \frac{3}{2} S_{I_s}^2\right)\left(1 - \frac{G^2}{L^2}\right)\left(1 - \frac{H^2}{G^2}\right)$$

$$\alpha_4 = \frac{15}{16} S_{I_s}^2 \left(1 - \frac{G^2}{L^2}\right)\left(1 + \frac{H}{G}\right)^2$$

$$\alpha_5 = \frac{15}{16} S_{I_s}^2 \left(1 - \frac{G^2}{L^2}\right)\left(1 - \frac{H}{G}\right)^2$$

$$\alpha_6 = \frac{15}{4} C_{I_s} S_{I_s} \left(1 - \frac{G^2}{L^2}\right)\left(1 - \frac{H^2}{G^2}\right)^{1/2}\left(1 + \frac{H}{G}\right)$$

$$\alpha_7 = \frac{15}{4} C_{I_s} S_{I_s} \left(1 - \frac{G^2}{L^2}\right)\left(1 - \frac{H^2}{G^2}\right)^{1/2}\left(1 - \frac{H}{G}\right) \tag{42}$$

The theory of this study assumes that the function $\bar{R}_m$ dominates $\bar{\bar{R}}_{ss}$. If the following approximations are used for $\bar{R}_m$ and $\bar{\bar{R}}_{ss}$

$$\bar{R}_m \cong \frac{\mu_m J_2 R_e^2}{a^3(1 - e^2)^{3/2}}$$

$$\bar{\bar{R}}_{ss} \cong n_s^2 a^2 (2 + 3e^2) \tag{43}$$

then

$$a \ll \left\{ \frac{\mu_m J_2 R_e^2}{n_s^2 (1 - e^2)^{3/2} (2 + 3e^2)} \right\}^{1/5} \tag{44}$$

This inequality holds throughout most of the "region of interest" (see Fig. 1).

The Hamiltonians associated with the three principal parts of the function $\bar{R}_{m+s}$ are

$$\begin{aligned} \bar{\mathcal{H}}_{sp} &= - \bar{R}_{sp} \\ \bar{\mathcal{H}}_{ss} &= - \bar{\bar{R}}_{ss} \\ \bar{\mathcal{H}}_{m} &= - \bar{R}_{m} \\ \bar{\mathcal{H}}_{s} &= - \bar{\bar{R}}_{ss} - \bar{\bar{R}}_{sp} \end{aligned} \tag{45}$$

Note that since ℓ is already ignorable we may assume $\mathcal{S}$ to be independent of ℓ so that $L = L' \equiv$ constant. Utilizing Eqs. (38), the perturbing Hamiltonian in terms of mixed variables (L', G', H', g, h) becomes

$$\bar{\mathcal{H}}_p = \bar{\mathcal{H}}_m \left(L', G' + \frac{\partial \mathcal{S}_1}{\partial g}, H' + \frac{\partial \mathcal{S}_1}{\partial h} \right) + \bar{\mathcal{H}}_{ss} \left(L', G' + \frac{\partial \mathcal{S}_1}{\partial g}, H' + \frac{\partial \mathcal{S}_1}{\partial h} \right)$$

$$+ \bar{\mathcal{H}}_{sp} \left(L', G' + \frac{\partial \mathcal{S}_1}{\partial g}, H' + \frac{\partial \mathcal{S}_1}{\partial h}, g, h \right) \tag{46}$$

Expanding in Taylor's series about G' and H' and retaining only one term in $\bar{\mathcal{H}}_{ss}$ and $\bar{\mathcal{H}}_{sp}$ yields

$$\bar{\mathcal{H}}_p = \bar{\mathcal{H}}_m (L', G', H') + \bar{\mathcal{H}}_{ss} (L', G', H') + \frac{\partial \bar{\mathcal{H}}_m}{\partial G'} \frac{\partial \mathcal{S}_1}{\partial g} + \frac{\partial \bar{\mathcal{H}}_m}{\partial H'} \frac{\partial \mathcal{S}_1}{\partial h}$$

$$+ \bar{\mathcal{H}}_{sp} (L', G', H', g, h) \tag{47}$$

In order that $\overline{\mathcal{H}}$ be a function of only (L', G', H') one requires that

$$\overline{\mathcal{H}}_{sp}\,(L', G', H', g, h) + \frac{\partial \overline{\mathcal{H}}_m}{\partial G'}\frac{\partial \mathcal{S}_1}{\partial g} + \frac{\partial \overline{\mathcal{H}}_m}{\partial H'}\frac{\partial \mathcal{S}_1}{\partial h} = 0 \tag{48}$$

This leads to the condition that

$$\frac{3\,\mu_m^4\,R_e^2\,J_2}{4L'^3\,G'^4}\left(1 - 5\,\frac{H'^2}{G'^2}\right)\frac{\partial \mathcal{S}_1}{\partial g} + \frac{3\,\mu_m^4\,R_e^2\,J_2\,H'}{2L'^3\,G'^5}\frac{\partial \mathcal{S}_1}{\partial h}$$

$$- \frac{n_s^2\,L'^4}{4\,\mu_m^2}\,[\alpha_1 C_{2h} + \alpha_2 C_h + \alpha_3 C_{2g} + \alpha_4 C_{2h+2g} + \alpha_5 C_{2h-2g}$$

$$- \alpha_6\,C_{h+2g} + \alpha_7 C_{h-2g}\,] = 0 \tag{49}$$

Therefore, choose $\mathcal{S}_1$ so that

$$\mathcal{S}_1 = -\frac{4L'^3\,G'^4}{3\,\mu_m^4\,J_2\,R_e^2}\cdot\frac{n_s^2\,L'^4}{4\,\mu_m^2}\,[\alpha_{11} S_{2h} + \alpha_{22} S_h + \alpha_{33} S_{2g}$$

$$+ \alpha_{44} S_{2h+2g} + \alpha_{55} S_{2h-2g} + \alpha_{66} S_{h+2g} + \alpha_{77} S_{h-2g}\,] \tag{50}$$

where the quantities α_{ij} are to be determined. Substituting $\mathcal{S}_1$ into Eq. (49) and taking the partial derivatives, the following condition is obtained:

$$\alpha_1 C_{2h} + \alpha_2 C_h + \alpha_3 C_{2g} + \alpha_4 C_{2h+2g} + \alpha_5 C_{2h-2g} - \alpha_6 C_{h+2g}$$

$$+ \alpha_7 C_{h-2g} + \left(1 - 5\,\frac{H'^2}{G'^2}\right)(2\alpha_{33} C_{2g} + 2\alpha_{44} C_{2h+2g} - 2\alpha_{55} C_{2h-2g}$$

$$+ 2\alpha_{66}C_{h-2g} - 2\alpha_{77}C_{h-2g}) + 2\,\frac{H'}{G'}\,(2\alpha_{11}C_{2h} + \alpha_{22}C_{h}$$

$$+ 2\alpha_{44}C_{2h+2g} + 2\alpha_{55}C_{2h-2g} + \alpha_{66}C_{h+2g} + \alpha_{77}C_{h-2g}) = 0 \tag{51}$$

Then equating coefficients

$$\alpha_{11} = \frac{1}{4}\,\alpha_1\,\frac{G'}{H'}$$

$$\alpha_{22} = \frac{1}{2}\,\alpha_2\,\frac{G'}{H'}$$

$$\alpha_{33} = \frac{1}{2}\,\alpha_3\left(1 - 5\,\frac{H'^2}{G'^2}\right)$$

$$\alpha_{44} = -\frac{\alpha_4}{2\left(5\,\frac{H'^2}{G'^2} - 2\,\frac{H'}{G'} - 1\right)}$$

$$\alpha_{55} = \frac{\alpha_5}{2\left(5\,\frac{H'^2}{G'^2} + 2\,\frac{H'}{G'} - 1\right)}$$

$$\alpha_{66} = \frac{\alpha_6}{2\left(5\,\frac{H'^2}{G'^2} - \frac{H'}{G'} - 1\right)}$$

$$\alpha_{77} = \frac{\alpha_7}{2\left(5\,\frac{H'^2}{G'^2} + \frac{H'}{G'} - 1\right)} \tag{52}$$

With the α_{ij}'s substituted into (50) the function δ_1 takes on the appearance

$$\mathcal{S}_1 = -\frac{n_s^2 L'^7 G'^4}{3\mu_m^6 J_2 R_e^2}\left\{\frac{1}{4}\alpha_1 \frac{G'}{H'} S_{2h} + \frac{1}{2}\alpha_2 \frac{G'}{H'} S_h + \frac{\alpha_3}{2\left(1 - 5\frac{H'^2}{G'^2}\right)} S_{2g}\right.$$

$$- \frac{\alpha_4}{2\left(5\frac{H'^2}{G'^2} - 2\frac{H'}{G'} - 1\right)} S_{2h+2g} + \frac{\alpha_5}{2\left(5\frac{H'^2}{G'^2} + 2\frac{H'}{G'} - 1\right)} S_{2h-2g}$$

$$\left. + \frac{\alpha_6}{2\left(5\frac{H'^2}{G'^2} - \frac{H'}{G'} - 1\right)} S_{h+2g} + \frac{\alpha_7}{2\left(5\frac{H'^2}{G'^2} + \frac{H'}{G'} - 1\right)} S_{h-2g}\right\} \qquad (53)$$

The coefficients of the trigonometric terms in $\mathcal{S}_1$ contain six critical divisors. An examination of these divisors reveals eleven critical inclinations which are summarized in Table 1.

From Eq. (53) the long-period variations in the elements ℓ, g, h, L, G, H can be found by appropriate partial differentiations. In this investigation, the variations in e and I are of primary importance, and are obtained as follows.

Form the differential of G from its definition (Eq. (36))

$$\delta G = \delta[\mu_m a(1 - e^2)]^{1/2}$$

or

$$\delta e = -\frac{G}{\mu_m a e}\delta G = -\frac{G}{L^2 e}\frac{\partial \mathcal{S}_1}{\partial g} \qquad (54)$$

Substituting for $\mathcal{S}_1$ the variation of e is given as

$$\delta e = \frac{n_s^2 L'^5 G'^5}{3 \mu_m^6 J_2 R_e^2 e'} \left\{ \frac{\alpha_3}{\left(1 - 5 \frac{H'^2}{G'^2}\right)} C_{2g} - \frac{\alpha_4}{\left(5 \frac{H'^2}{G'^2} - 2 \frac{H'}{G'} - 1\right)} C_{2g+2h} \right.$$

$$- \frac{\alpha_5}{\left(5 \frac{H'^2}{G'^2} + 2 \frac{H'}{G'} - 1\right)} C_{2h-2g} + \frac{\alpha_6}{\left(5 \frac{H'^2}{G'^2} - \frac{H'}{G'} - 1\right)} C_{h+2g}$$

$$\left. - \frac{\alpha_7}{\left(5 \frac{H'^2}{G'^2} + \frac{H'}{G'} - 1\right)} C_{h-2g} \right\} \tag{55}$$

which becomes

$$\delta e = \frac{5}{4}\left(\frac{n_s}{n}\right)^2 \left(\frac{p}{R_e}\right)^2 \frac{e'\sqrt{1-e'^2}}{J_2} \left\{ \frac{\left(1 - \frac{3}{2} S_{I_s}^2\right) S_{I'}^2}{\left(1 - 5 C_{I'}^2\right)} C_{2\omega'} \right.$$

$$- \frac{S_{I_s}^2 (1 + C_{I'})^2}{4\left(5 C_{I'}^2 - 2 C_{I'} - 1\right)} C_{2\Omega'+2\omega'} - \frac{S_{I_s}^2 (1 - C_{I'})^2}{4\left(5 C_{I'}^2 + 2 C_{I'} - 1\right)} C_{2\Omega'-2\omega'}$$

$$\left. + \frac{S_{I_s} C_{I_s} S_{I'} (1 + C_{I'})^2}{\left(5 C_{I'}^2 - C_{I'} - 1\right)} C_{\Omega'+2\omega'} - \frac{S_{I_s} C_{I_s} S_{I'} (1 - C_{I'})}{5 C_I^{2\prime} + C_{I'} - 1} C_{\Omega'-2\omega} \right\} \tag{56}$$

In order to simplify the equation and the checking of dimensions the parameters n (mean motion of satellite) and p (defined previously) have been introduced. It appears from an inspection of Eq. (56) that at a critical inclination the amplitude of a term in δe becomes infinite while the frequency approaches zero. Actually, as will be shown later, near any critical inclination, other than $I_c = 90°$, δe experiences a finite maximum variation.

To obtain a corresponding expression for δI it is convenient to express $\cos\ I'$ in terms of H' and G' by using

$$\frac{H'}{G'} = C_{I'} \tag{57}$$

Taking the differential of both sides

$$\delta\left\{\frac{H}{G}\right\} = \frac{\delta H}{G} - \frac{H}{G^2}\,\delta G$$

$$\delta I = \frac{1}{\sqrt{\mu_m a(1 - e^2)}\ S_I}\left\{\cos I\,\frac{\partial \mathcal{S}_1}{\partial g} - \frac{\partial \mathcal{S}_1}{\partial h}\right\} \tag{58}$$

To assist the reader in evaluating the equations of variation the partial derivatives of $\mathcal{S}_1$ with respect to the variables g and h are shown below

$$\frac{\partial \mathcal{S}_1}{\partial g} = -\frac{n_s^2\,L'^7\,G'^4}{3\,\mu_m^6\,J_2\,R_e^2}\left\{\frac{\alpha_3}{\left(1 - 5\,\frac{H'^2}{G'^2}\right)}C_{2g} - \frac{\alpha_4}{\left(5\,\frac{H'^2}{G'^2} - 2\,\frac{H'}{G'} - 1\right)}C_{2g+2h}\right.$$

$$\left. - \frac{\alpha_5}{\left(5\,\frac{H'^2}{G'^2} + 2\,\frac{H'}{G'} - 1\right)}C_{2h+2g} + \frac{\alpha_6}{\left(5\,\frac{H'^2}{G'^2} - \frac{H'}{G'} - 1\right)}C_{h+2g} - \frac{\alpha_7}{\left(5\,\frac{H'^2}{G'^2} + \frac{H'}{G'} - 1\right)}C_{h-2g}\right\}$$

$$\frac{\partial \mathcal{S}_1}{\partial h} = -\frac{n_s^2\,L'^7\,G'^4}{3\,\mu_m^6\,J_2\,R_e^2}\left\{\frac{1}{2}\,\alpha_1\,\frac{G'}{H'}\,C_{2h} + \frac{1}{2}\,\alpha_2\,\frac{G'}{H'}\,C_h\right.$$

$$- \frac{\alpha_4}{\left(5\,\frac{H'^2}{G'^2} - 2\,\frac{H'}{G'} - 1\right)}C_{2h+2g} + \frac{\alpha_5}{\left(5\,\frac{H'^2}{G'^2} + 2\,\frac{H'}{G'} - 1\right)}C_{2h-2g}$$

$$\left. + \frac{\alpha_6}{2\left(5\,\frac{H'^2}{G'^2} - \frac{H'}{G'} - 1\right)}C_{h+2g} + \frac{\alpha_7}{2\left(5\,\frac{H'^2}{G'^2} + \frac{H'}{G'} - 1\right)}C_{h-2g}\right\} \tag{59}$$

The variation in I written in canonical variables is

$$\delta I = - \frac{n_s^2 L'^7 G'^4}{3 a'^{1/2} \mu_m^{3/2} J_2 R_e^2 (1 - e'^2)^{1/2} S_{I'}} \left\{ \frac{\alpha_3 C_{I'}}{\left(1 - 5 \frac{H'^2}{G'^2}\right)} C_{2g} \right.$$

$$- \frac{\alpha_4 C_{I'}}{\left(5 \frac{H'^2}{G'^2} - 2 \frac{H'}{G'} - 1\right)} C_{2h+2g} - \frac{\alpha_5 C_{I'}}{\left(5 \frac{H'^2}{G'^2} + 2 \frac{H'}{G'} - 1\right)} C_{2h-2g}$$

$$+ \frac{\alpha_6 C_{I'}}{\left(5 \frac{H'^2}{G'^2} - \frac{H'}{G'} - 1\right)} C_{h+2g} - \frac{\alpha_7 C_{I'}}{\left(5 \frac{H'^2}{G'^2} + \frac{H'}{G'} - 1\right)} C_{h-2g}$$

$$+ \frac{1}{2} \frac{\alpha_1}{C_{I'}} C_{2h} + \frac{1}{2} \frac{\alpha_2}{C_{I'}} C_h - \frac{\alpha_4}{\left(5 \frac{H'^2}{G'^2} - 2 \frac{H'}{G'} - 1\right)} C_{2h+2g}$$

$$+ \frac{\alpha_5}{\left(5 \frac{H'^2}{G'^2} + 2 \frac{H'}{G'} - 1\right)} C_{2h-2g} + \frac{\alpha_6}{2\left(5 \frac{H'^2}{G'^2} - \frac{H'}{G'} - 1\right)} C_{h+2g}$$

$$\left. + \frac{\alpha_7}{2\left(5 \frac{H'^2}{G'^2} + \frac{H'}{G'} - 1\right)} C_{h-2g} \right\} \tag{60}$$

and in Keplerian elements

$$\delta I = \frac{5}{4} \left(\frac{n_s}{n}\right)^2 \left(\frac{a'}{R_e}\right)^2 \frac{e'^2 (1 - e'^2)^{3/2}}{J_2 S_{I'}} \left\{ \frac{\left(\frac{3}{2} S_{I_s}^2 - 1\right) S_{I'}^2 C_{I'}}{1 - 5 C_{I'}^2} C_{2\omega'} \right.$$

$$
\begin{aligned}
&+ \frac{s^2_{I_s}(1 + c_{I'})^2 c_{I'}}{4\left(5\,c^2_{I'} - c_{I'} - 1\right)}\, c_{2\Omega'+2\omega'} + \frac{s^2_{I_s}(1 - c_{I'})^2 c_{I'}}{4\left(5\,c^2_{I'} + 2\,c_{I'} - 1\right)}\, c_{2\Omega'-2\omega'} \\
&- \frac{s_{I_s} c_{I_s} s_{I'} c_{I'} (1 + c_{I'})}{\left(5\,c^2_{I'} - c_{I'} - 1\right)}\, c_{\Omega'+2\omega'} + \frac{s_{I_s} c_{I_s} s_{I'} c_{I'} (1 - c_{I'})}{\left(5\,c^2_{I'} + c_{I'} - 1\right)}\, c_{\Omega'-2\omega'} \\
&+ \frac{s^2_{I_s} s^2_{I'}}{4\,c_{I'}}\, c_{2\Omega'} + s_{I_s} c_{I_s} s_{I'} c_{\Omega'} \\
&- \frac{s^2_{I_s}(1 + c_{I'})^2}{4\left(5\,c^2_{I'} - 2\,c_{I'} - 1\right)}\, c_{2\Omega'+2\omega'} + \frac{s_{I_s}(1 - c_{I'})^2}{4\left(5\,c^2_{I'} + 2\,c_{I'} - 1\right)}\, c_{2\Omega'-2\omega'} \\
&\left.+ \frac{s_{I_s} c_{I_s} (1 + c_{I'}) s_{I'}}{2\left(5\,c^2_{I'} - c_{I'} - 1\right)}\, c_{\Omega'+2\omega'} + \frac{s_{I_s} c_{I_s} (1 - c_{I'}) s_{I'}}{2\left(5\,c^2_{I'} + c_{I'} - 1\right)}\, c_{\Omega'-2\omega'} \right\} \qquad (61)
\end{aligned}
$$

Note here that the inclination $I_c = 90°$ is critical for δI but does not appear in the expression for δe.

IV. LONG-PERIOD BEHAVIOR NEAR CRITICAL INCLINATION

The technique outlined in the previous section for obtaining long-period behavior of the elements proves unsatisfactory when the inclination is near one of the critical values (see Table 1). In this section a suitable analysis will be conducted to determine the behavior of the satellite's motion while in the neighborhood of a general critical inclination. From Eq. (41) one can write the periodic part of the Hamiltonian in the form

$$\overline{\overline{R}}_{sp} = \frac{n_s^2 L'^4}{4 \mu_m^2} \left[\sum_{i=-2}^{2} \gamma_i \cos (2g' + ih') + \sum_{j=1,2} \beta_j \cos jh' \right] \tag{62}$$

The critical inclinations are just those at which one of the trigonometric arguments in (62) has a zero rate due to J_2. Thus the terms of $\cos jh'$ yield a critical inclination of 90°; however, since they do not contain the variable g' they will contribute nothing to the variation of e'. Of course, these terms will contribute to the long-period behavior of the inclination, but since the primary emphasis in this investigation is centered on the radius of pericenter (equivalently e', since $\delta r_p = -a'\delta e'$), these terms are not of direct interest here. Thus in the following analysis these terms will also be considered to have been removed by a suitable generating function, so that the total Hamiltonian of interest will be

$$-\overline{\mathcal{H}}'_p = \frac{\mu_m^4 J_2 R_e^2}{4L'^3 G'^3} \left(3 \frac{H'^2}{G'^2} - 1\right) + \frac{n_s^2 L'^4}{4 \mu_m^2} \left(5 - 3 \frac{G'^2}{L'^2}\right)\left(3 \frac{H'^2}{G'^2} - 1\right)$$

$$+ \frac{n_s^2 L'^4}{4 \mu_m^2} \sum_{i=-2}^{2} \gamma_i \cos (2g' + ih') \tag{63}$$

Specifically, the coefficients γ_i are given in terms of the α_i of Eq. (42) as

$$\begin{aligned} \gamma_o &= \alpha_3 & \gamma_{-1} &= \alpha_7 \\ \gamma_1 &= -\alpha_6 & \gamma_{-2} &= \alpha_5 \\ \gamma_2 &= \alpha_4 & & \end{aligned} \tag{64}$$

In order to analyze the motion near the critical inclination defined by $2\dot{g}' + i\dot{h}' = 0$, a determining function $\mathcal{S}^*$ is introduced which is taken to be $\mathcal{S}$ with the i term of interest absent. Thus $\mathcal{S}^*$ will "remove" all the periodic terms from the Hamiltonian except the i term, so that $\bar{\mathcal{H}}_p (= -\bar{R}^*_{m+s})$ now reads

$$\bar{R}^*_{m+s} = -\bar{\mathcal{H}}_p = \bar{\bar{R}}_{ss} + \bar{R}_m + \frac{\hat{n}_s^2 L'^4}{4 \mu_m^2} \gamma_i \cos (2g' + ih') \tag{65}$$

Introduce now a contragredient transformation to new variables

$$\begin{bmatrix} g'' \\ h'' \end{bmatrix} = \begin{bmatrix} 1 & i/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} g' \\ h' \end{bmatrix} \tag{66}$$

$$\begin{bmatrix} G'' \\ H'' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -i/2 & 1 \end{bmatrix} \begin{bmatrix} G' \\ H' \end{bmatrix} \tag{67}$$

Then the Hamiltonian is of the form

$$\bar{R}^*_{m+s} = \bar{\bar{R}}^{\ddagger}_{ss} (G'', H'', L') + \bar{R}^{\ddagger}_m (G'', H'', L') + \frac{n_s^2 L'^4}{4 \mu_m^2} \gamma_i^{\ddagger} \cos 2g'' \tag{68}$$

Note that L' and H'' are constants and, since the Hamiltonian itself is a constant in the absence of explicit time terms it follows that $\bar{R}^*_{m+s}$ is also a constant. Define now G''_c as that function of L' and H'' for which

$$\left[\frac{\partial}{\partial G''} \bar{R}^*_m\right]_{G''=G''_c} = 0$$

and expand $\bar{R}^*_{m+s}$ about $(G'' - G''_c)$:

$$\begin{aligned} \bar{R}^*_{m+s} \cong\ & \bar{\bar{R}}^{\ddagger}_{ss} (G''_c, H'', L') + \bar{R}^{\ddagger}_m (G''_c, H'', L') + \left[\frac{\partial}{\partial G''} \bar{\bar{R}}^{\ddagger}_{ss}\right]_{G''=G''_c} (G'' - G''_c) \\ & + \frac{1}{2} \left[\frac{\partial^2}{\partial G''^2} \left(\bar{\bar{R}}^{\ddagger}_{ss} + \bar{R}^{\ddagger}_m\right)\right]_{G''=G''_c} (G''-G''_c)^2 + \left[\frac{n_s^2 L'^4}{4 \mu_m^2} \gamma_i^{\ddagger}\right]_{G''=G''_c} \cos 2g'' \end{aligned} \tag{69}$$

Ignoring the fluctuations in $\gamma_i^{\ddagger}$ by comparison with its value at G''_c is not valid for small eccentricity (see Appendix). Using the definitions

$$\mathcal{B}_1 = \frac{1}{2}\left[\frac{\partial^2}{\partial G''^2}\left(\bar{\bar{R}}^{\ddagger}_{ss} + \bar{R}^{\ddagger}_{m}\right)\right]_{G''=G''_c}$$

$$\mathcal{B}_2 = \left[\frac{\partial}{\partial G''}\bar{\bar{R}}^{\ddagger}_{ss}\right]_{G''=G''_c}$$

$$\mathcal{B}_3 = \left[\bar{\bar{R}}^{\ddagger}_{ss} + \bar{R}^{\ddagger}_{m}\right]_{G''=G''_c} ,$$

$$f = \left[\frac{n_s^2 L'^4}{4\mu_m^2}\gamma_i^{\ddagger}\right]_{G''=G''_c} , \tag{70}$$

equation (69) may be written as

$$\mathcal{B}_1(G'' - G''_c)^2 + \mathcal{B}_2(G'' - G''_c) + \mathcal{B}_3 + f \cos 2g'' = \text{constant} \tag{71}$$

Completing the square and suitable rearrangement yields the equation

$$\left[\left(\frac{G''}{H''} - \frac{G''_c}{H''}\right) + \frac{\mathcal{B}_2}{2\mathcal{B}_1 H''}\right]^2 + \frac{f}{\mathcal{B}_1 H''^2}\cos 2g'' = K = \text{constant} \tag{72}$$

which is now recognized as being in the form of the simple pendulum equation. The solution of this problem is well known and results in a plot of phase space contours which describe the motion of the pendulum. In the present problem, the variation of G''/H'' with the angle g'' follows from

$$\frac{G''}{H''} - \frac{G''_c}{H''} = \pm\left[K - \frac{f}{\mathcal{B}_1 H''^2}\cos 2g''\right]^{1/2} - \frac{\mathcal{B}_2}{2\mathcal{B}_1 H''} \tag{73}$$

Note that g'' can librate about $\pi/2$ or 0 according to whether $f/\mathcal{B}_1 > 0$ or < 0. As in the pendulum problem, the separatrix is determined by setting $K = f/\mathcal{B}_1 H''^2$ and varying g''. The motion bounded by this contour represents a libration of G''/H'' about some average $(G''/H'')_{ave}$, and as can be seen in Fig. 3, the maximum fluctuation of this variable follows a contour just inside the separatrix and is equal to $2\sqrt{2K}$.

Recall now that it was assumed earlier that $\bar{\bar{R}}^{\ddagger}_{ss} \ll \bar{R}^{\ddagger}_m$, so that $\mathcal{B}_1$ can be taken as

$$\mathcal{B}_1 = \frac{1}{2}\left[\frac{\partial^2}{\partial G''^2}\bar{R}^{\ddagger}_m\right]_{G''=G''_c} \tag{74}$$

which, written explicitly in terms of the old primed variables, is

$$\mathcal{B}_1 = \frac{3}{2}\frac{\mu_m^4 J_2 R_e^2}{L'^3 G'^5}\left(30\cos^2 I' - 10i\cos I' + \frac{i^2}{2} - 4\right) \tag{75}$$

The maximum variation in G''/H'' written in terms of the original orbital elements is

$$\delta\left(\frac{G''}{H''}\right) = 2\sqrt{2K} = 2\left[\left|\frac{2f}{\mathcal{B}_1 H''^2}\right|\right]^{1/2} \tag{76}$$

from which the maximum variation in the inclination is found to be

$$|\delta I|_{max} = \frac{|2\cos I'_c - i|e'(1-e'^2)^{3/4} n_s a'^{5/2}}{R_e \sin I'_c}\left[\frac{5\gamma_i^*}{\mu_m J_2\left|\left(30\cos^2 I'_c - 10i\cos I'_c + \frac{i^2}{2} + 4\right)\right|}\right]^{1/2} \tag{77}$$

since

$$\delta\left(\frac{G''}{H''}\right) = \frac{\sin I'_c}{\left(\cos I'_c - \frac{i}{2}\right)^2}\delta I = 2\left[\left|\frac{2f}{\mathcal{B}_1 H''^2}\right|\right]^{1/2}$$

with γ_i^* defined by $\gamma_i^* = \frac{4}{15e'^2} \gamma_i \left(I_s, e', I_c'\right)$

The theory developed near I_c' is valid if δI is small enough so that I' does not get too near any other neighboring I_c'. To prevent this occurrence, an upper limit is assigned to the maximum variation $|\delta I|_{max}$ such that

$$|\delta I|_{max} \leq \frac{1}{3} \left|I'_{c_i} - I'_{c_j}\right| = \Delta \tag{78}$$

where I'_{c_i} is the I_c' under investigation and I'_{c_j} is the nearest neighboring one.

The corresponding maximum variation in the eccentricity is found from the constancy of H'' i.e., $\delta H'' = 0$ which gives

$$|\delta e'|_{max} = \frac{2(1 - e'^2) \sin I_c'}{e'|2 \cos I_c' - i|} |\delta I| \tag{79}$$

or

$$|\delta e'|_{max} = \frac{2n_s\ a^{5/2}(1-e'^2)^{3/2}}{R_e} \left[\frac{5\ \gamma_i^*}{\mu_m J_2\left|\left(30 \cos^2 I_c' - 10i \cos I_c' + \frac{i^2}{2} - 4\right)\right|}\right]^{1/2} \tag{80}$$

Now define the constant K_i by

$$K_i = 2n_s\ R_e^{3/2} \left[\frac{5\ \gamma_i^*}{\mu_m J_2\left|\left(\cos^2 I_c' - 10i \cos I_c' + \frac{i^2}{2} - 4\right)\right|}\right]^{1/2} \tag{81}$$

so that

$$|\delta e'|_{max} = K_i(1 - e'^2)^{7/4} \left(\frac{a}{R_e}\right)^{5/2} \tag{82}$$

Then $|\delta I|_{max}$ can be written in terms of K_i as

$$|\delta I|_{max} = \frac{|2 \cos I_c' - i|}{2 \sin I_c'} K_i \, e'(1-e'^2)^{3/4} \left(\frac{a}{R_e}\right)^{5/2} \tag{83}$$

Invoking then the condition on a maximum allowable $|\Delta I|_{max}$ (Eq. (78)) one can write an upper limit a' as a function of e' consistent with the analysis as

$$\frac{|a|}{R_e} \leq \frac{|a|_{max}}{R_e} = \left\{\frac{2 \sin I_c' \, \Delta}{|2 \cos I_c' - i| K_i} \left[\frac{1}{e'(1-e'^2)^{3/4}}\right]\right\}^{2/5} \tag{84}$$

or, defining constants B_i by

$$B_i = \left[\frac{2 \sin I_c' \, \Delta}{|2 \cos I_c' - i| K_i}\right]^{2/5} \tag{85}$$

then

$$\frac{(r_p)_{max}}{R_e} = B_i \, e'^{-2/5} \, (1+e')^{-3/10} \, (1-e')^{+7/10} \tag{86}$$

The constants B_i and K_i for the different values of $i(-2 \leq i \leq 2)$ corresponding to the 10 different critical inclinations are listed in Table 2. The bounds (86) on the applicability of this analysis are shown in Fig. 1.

Finally, knowing the maximum fluctuations in e', the maximum variation in the radius of pericenter is determined as

$$|\delta r_p|_{max} = a' \, |\delta e'|_{max} = K_i \frac{r_p^{7/2}}{R_e^{5/2}} \left(\frac{1+e'}{1-e'}\right)^{7/4} \tag{87}$$

This maximum variation of the radius of pericenter with e' near the different critical inclinations is shown in Fig. 4 for r_p = 4000 and 6000 km. For r_p above the theoretical limits of Fig. 1, the corresponding portion of the curve is dotted in Fig. 4 for a specific index i, the theory being open to question since the fluctuations in I' may overlap a neighboring critical inclination. In case of overlap a simple analysis of $|\delta e|_{max}$ is not possible.

This investigation was purposely limited to large values of e' due to the mission requirements imposed on it; in fact, the analysis of Section IV was invalid for e' close to zero. For e' close to zero, however, the maximum variation in e' agrees with that found in this analysis; this agreement is demonstrated in the Appendix.

APPENDIX

In this section, the eccentricity is assumed to be of order $J_2^{1/2}$ and it is assumed that the inclination is near one of the critical inclinations I'_{c_1} other than 90°.

Recall that in Section IV it was found that the Hamiltonian could be written in the form

$$\mathcal{B}_1(G'' - G''_c)^2 + \mathcal{B}_2(G'' - G''_c) + \mathcal{B}_3 + f \cos 2g'' = \text{constant} \tag{A.1}$$

Then from the constancy of H''

$$H'' = H' - \frac{1}{2} G' = \text{const}$$

it follows that

$$\sqrt{1-e'^2}\left(\cos I' - \frac{1}{2}\right) = \text{const} \tag{A.2}$$

or

$$\ell n\ (1-e'^2) + 2\ \ell n \left(\cos I' - \frac{1}{2}\right) = \text{const}$$

$$e'^2 = 2\ \ell n \left(\cos I' - \frac{1}{2}\right) + \text{const} + 0(J_2^2) \tag{A.3}$$

Next, expand about $I' = I'_c$:

$$e'^2 = 2\left[\ln\left(\cos I'_c - \frac{1}{2}\right) + \frac{\sin I'_c}{\cos I'_c - \frac{1}{2}}(I'-I'_c) + \ldots\right]$$

$$e'^2 = \frac{2 \sin I'_c}{\cos I'_c - \frac{1}{2}}(I'-I'_c) + H_1 + O(J_2^2) \tag{A.4}$$

where H_1 is a small constant of order J_2 if the initial values of e'^2 and $I'-I'_c$ are $O(J_2)$. Therefore

$$I'-I'_c = \frac{\left(\cos I'_c - \frac{1}{2}\right)}{2 \sin I'_c}\left[e'^2 - H_1\right] \tag{A.5}$$

From the constancy of H''

$$\delta H'' = \delta G''\left(\cos I'_c - \frac{1}{2}\right) - G''_c \sin I'_c \,\delta I' = 0$$

so that

$$\delta G'' = \frac{G''_c \sin I'_c}{\cos I'_c - \frac{1}{2}}\,\delta I' \tag{A.6}$$

Substituting $\delta I' = I'-I'_c$ from (A.5)

$$G''-G''_c = \frac{G''_c}{2}\left[e'^2 - H_1\right] \tag{A.7}$$

the Hamiltonian now takes the form

$$\mathcal{B}_1\left[\frac{G''_c}{2}\left(e'^2 - H_1\right)\right]^2 + \mathcal{B}_2\,\frac{G''_c}{2}\left[e'^2 - H_1\right] + e'^2 f^* \cos 2g'' = \text{constant} \tag{A.8}$$

where† $f^* = f/e'^2$.

†This form of f^* does not, in fact, contain a $1/e'^2$ factor since inspection of definitions 70, 64, and 42 reveals that f has an e'^2 factor.

Combining powers of e', Eq. (A.8) is rewritten as

$$\bar{\mathcal{B}}_1 e'^4 + (\bar{\mathcal{B}}_2 - 2\,\bar{\mathcal{B}}_1 H_1)\, e'^2 + f^* e'^2 \cos 2g'' = \text{constant} \tag{A.9}$$

where $\bar{\mathcal{B}}_1 = \mathcal{B}_1 \dfrac{G''^2_c}{4}$, $\quad \bar{\mathcal{B}}_2 = \mathcal{B}_2 \dfrac{G''_c}{2}$.

The interplay between g'' and e' (and hence I') may be more easily visualized if one introduces the coordinates

$$\begin{aligned} \xi &= e' \cos g'' \\ \eta &= e' \sin g'' \end{aligned} \tag{A.10}$$

so that Eq. (A.9) becomes

$$\bar{\mathcal{B}}_1(\xi^2 + \eta^2)^2 + (\bar{\mathcal{B}}_2 - 2\,\bar{\mathcal{B}}_1 H_1)(\xi^2 + \eta^2) + f^*(\xi^2 - \eta^2) = K \tag{A.11}$$

The equilibrium points of the contours in the ξ-η plane are found from

$$\frac{\partial K}{\partial \xi} = \xi\left[2\,\bar{\mathcal{B}}_1(\xi^2 + \eta^2) + (\bar{\mathcal{B}}_2 - 2\,\bar{\mathcal{B}}_1 H_1) + f^*\right] = 0 \tag{A.12}$$

$$\frac{\partial K}{\partial \eta} = \eta\left[2\,\bar{\mathcal{B}}_1(\xi^2 + \eta^2) + (\bar{\mathcal{B}}_2 - 2\,\bar{\mathcal{B}}_1 H_1) + f^*\right] = 0 \tag{A.13}$$

Solution of these equations yields equilibrium points at the origin of the ξ, η plane and on the ξ and η axes at

$$(0, \eta_2) = \left(0, \left[H_1 - \frac{(\bar{\mathcal{B}}_2 + f^*)}{2\bar{\mathcal{B}}_1}\right]^{1/2}\right)$$

and

$$(\xi_2, 0) = \left(0, \left[H_1 + \frac{f^* - \bar{\mathcal{B}}_2}{2\,\bar{\mathcal{B}}_1}\right]^{1/2}, 0\right)$$

respectively.

From (A.11) one can see that the contours are symmetric about both axes; depending on the value of H_1 a variety of contours are possible about the equilibrium points, c.f., Fig. 5.

The "separatrices" between different types of contours are shown by dotted lines. As H_1 increases, (a) → (c) → (e) or (b) → (d) → (f) and the separatrices, after appearing in (c) or (d), grow and then change to circular form in (e) or (f). In (g) and (h) are shown their transition forms corresponding say, to $H_1 = H_{1c}$, and in Fig. 6 is shown the situation for H_1 slightly larger.

On contours outside the circle-pair, and for $f^*/\mathfrak{B}_1 < 0$ for example, from (A.11) one finds that e_{max} (corresponding to $y = 0$) is given by

$$e^2_{max} = \sqrt{P} + \sqrt{P+k} \tag{A.14}$$

where $\sqrt{P} = H_1 - \dfrac{\bar{\mathfrak{B}}_2 + f^*}{2\bar{\mathfrak{B}}_1}$

$$k = \frac{K}{\bar{\mathfrak{B}}_1}$$

while e_{min} (corresponding to $\xi = 0$) is

$$e^2_{min} = \sqrt{Q} + \sqrt{Q+k} \tag{A.15}$$

where $\sqrt{Q} = H_1 - \dfrac{\bar{\mathfrak{B}}_2 - f^*}{2\bar{\mathfrak{B}}_1}$

Then

$$(\delta e)_{max}\,(e_{max} + e_{min}) = e^2_{max} - e^2_{min} = \sqrt{P} - \sqrt{Q} + \frac{P - Q}{\sqrt{P+k} + \sqrt{Q+k}} \tag{A.16}$$

which is a decreasing function of k, while $e_{max} + e_{min}$ is an increasing function of k. It follows that $(\delta e)_{max}$ decreases as k increases, and hence is smaller outside the circle-pair than on it.

Thus it is apparent in Fig. 6 that the largest fluctuation, δe, in eccentricity occurs on a contour just inside a separatrix circle. It will be shown, moreover, that this $(\delta e)_{max}$ remains constant as H_1 increases further in spite of the growth of the radius of the separatrix circles. The overall maximum δe is thus just the diameter of the circles in (g) or (h). To evaluate this, note that for sufficiently large H_1 the family of contours (A.11) includes the circle-pair

$$[(\xi - \gamma)^2 + \eta^2 - \rho^2][(\xi + \gamma)^2 + \eta^2 - \rho^2] = 0, \qquad \frac{f^*}{\bar{\mathscr{B}}_1} < 0 \tag{A.17}$$

or

$$[\xi^2 + (\eta - \gamma)^2 - \rho^2][\xi^2 + (\eta + \gamma)^2 - \rho^2] = 0, \qquad \frac{f^*}{\bar{\mathscr{B}}_1} > 0 \tag{A.18}$$

where $\rho = \left(H_1 - \dfrac{\bar{\mathscr{B}}_2}{2\bar{\mathscr{B}}_1}\right)^{1/2}$

$$\gamma = \left[\left|\frac{f^*}{2\bar{\mathscr{B}}_1}\right|\right]^{1/2} < \rho$$

But $(\delta e)_{max} = (\rho + \gamma) - (\rho - \gamma) = 2\gamma$, independent of ρ. Thus

$$(\delta e)_{max} = 2\left[\left|\frac{f^*}{2\bar{\mathscr{B}}_1}\right|\right]^{1/2} = \frac{4}{e'G''_c}\left[\left|\frac{f}{2\mathscr{B}_1}\right|\right]^{1/2} \tag{A.19}$$

in agreement (for small e') with Eqs. (77) and (79) of the large e' theory, which yield

$$(\delta e)_{max} = \frac{4(1-e'^2)}{e'G''_c}\left[\left|\frac{f}{2\mathscr{B}_1}\right|\right]^{1/2} \tag{A.20}$$

TABLE I

CRITICAL INCLINATIONS

Critical Divisors	Critical Inclincations I_c	Index i
$\cos I_c$	90°	
$1 - 5\cos^2 I_c$	63.4°, 116.6°	0
$5\cos^2 I_c - 2\cos I_c - 1$	46.4°, 106.8°	2
$5\cos^2 I_c + 2\cos I_c - 1$	73.2°, 133.6°	-2
$5\cos^2 I_c - \cos I_c - 1$	56.1°, 111.0°	1
$5\cos^2 I_c + \cos I_c - 1$	69.0°, 123.9°	-1

TABLE 2

COEFFICIENTS B_i AND K_i

i	B_i	K_i	I_c
0	2.6	5.6×10^{-3}	63.4°, 116.6°
1	4.5	1.4×10^{-2}	56.1°, 111.0°
2	4.6	2.8×10^{-3}	46.4°, 106.8°
-1	2.6	2.4×10^{-3}	69.0°, 123.9°
-2	3.9	5.9×10^{-4}	73.2°, 133.6°

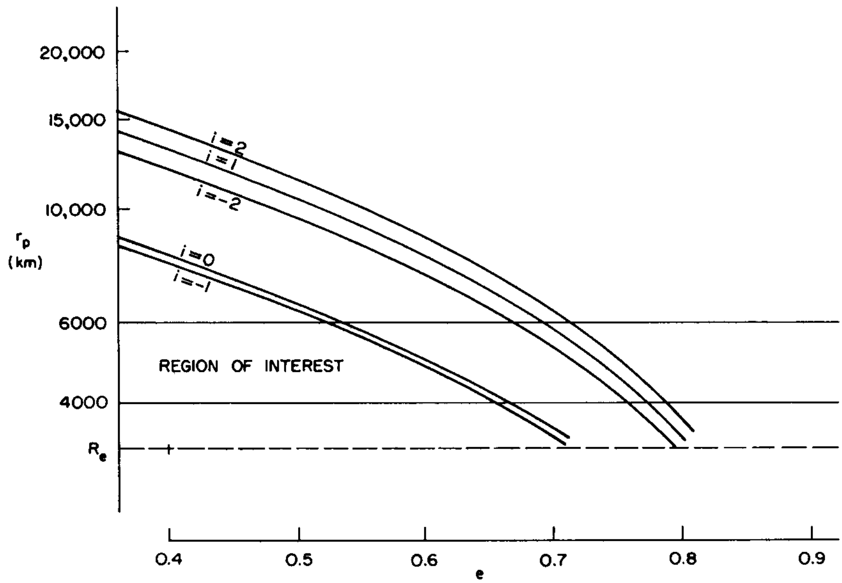
REFERENCES

1. Garfinkel, B., Astronomical Journal, 1960, Vol. 65.

2. Tisserand, F., Traite de Mecanique Celeste, 1889, Vol. 1

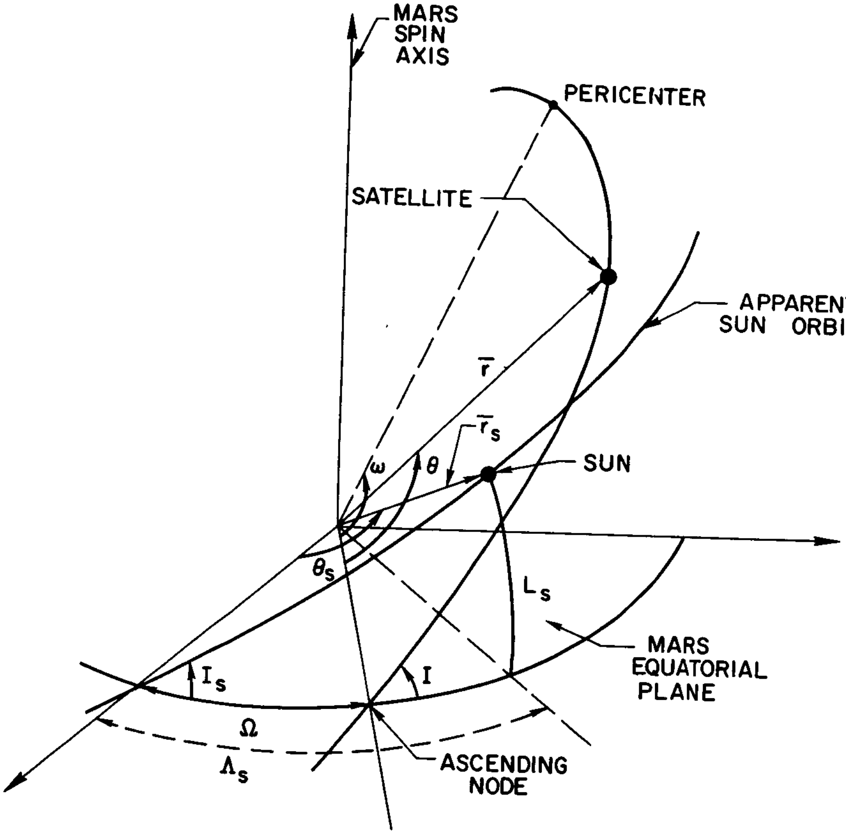
3. Brouwer, D., Astronomical Journal, 1959, Vol. 64.

National Aeronautics and Space Administration
Electronics Research Center
Cambridge, Massachusetts, December 1966
129-04-04-07

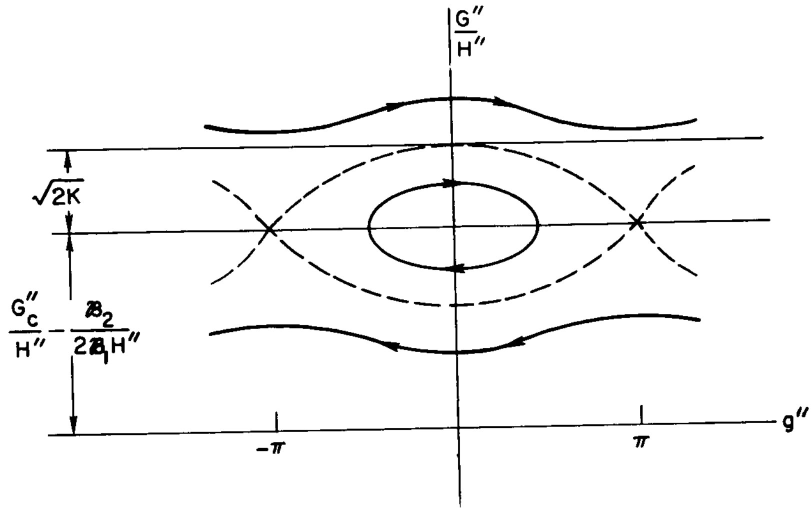


BOUNDARY CURVES FOR VALIDITY OF CRITICAL INCLINATION ANALYSIS

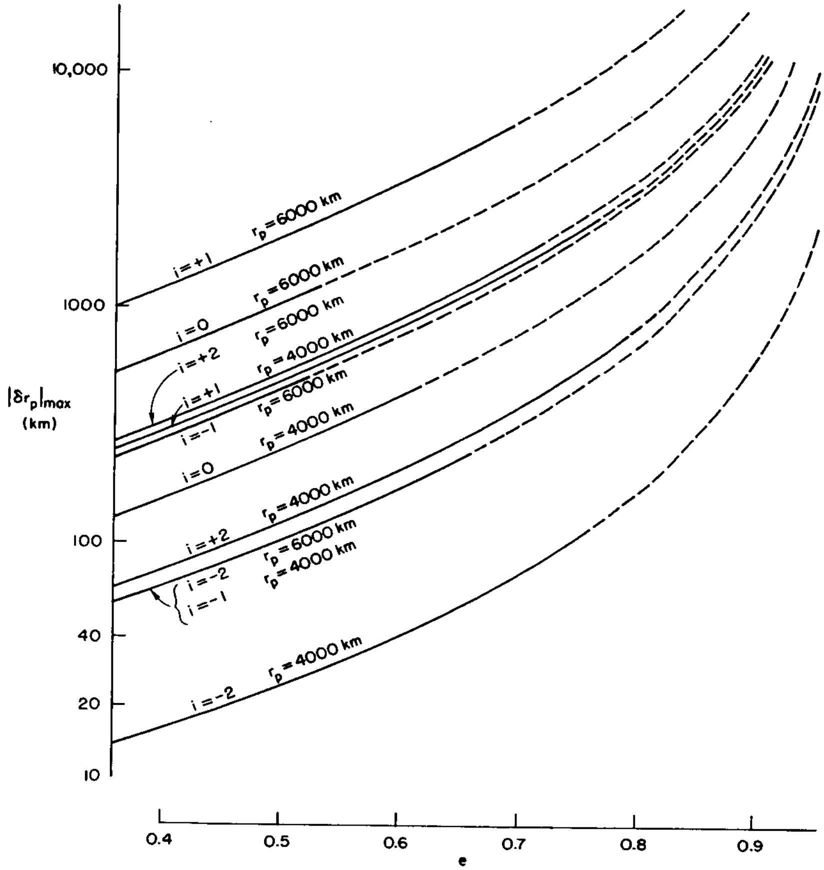
Figure 1



MARS-CENTERED ORBIT GEOMETRY
Figure 2

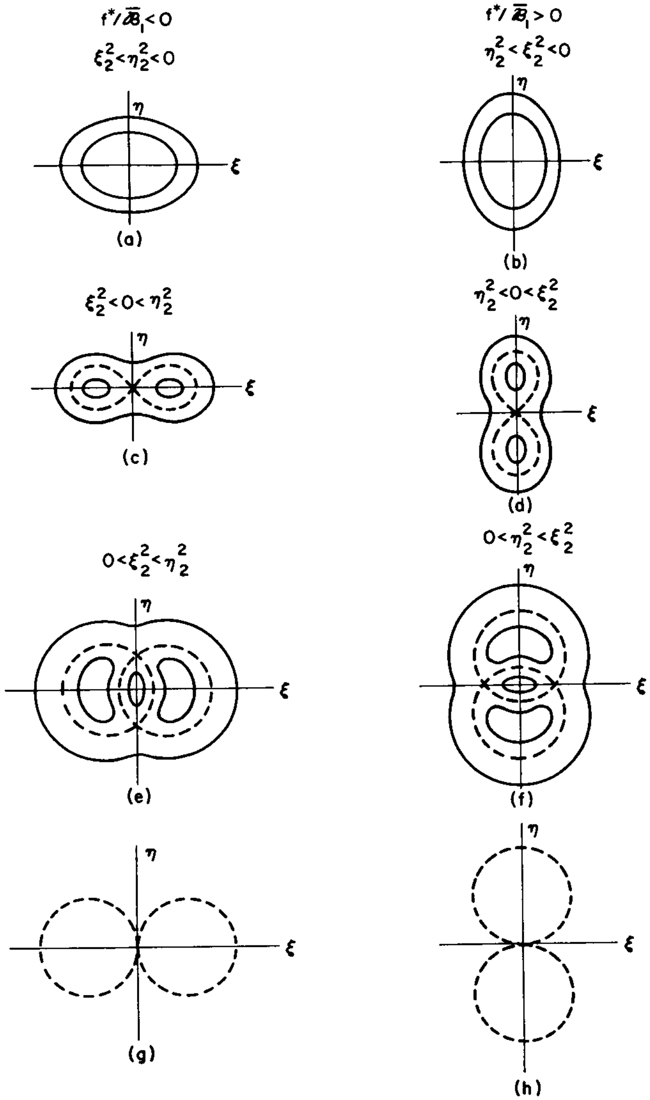


PHASE SPACE CONTOURS
Figure 3



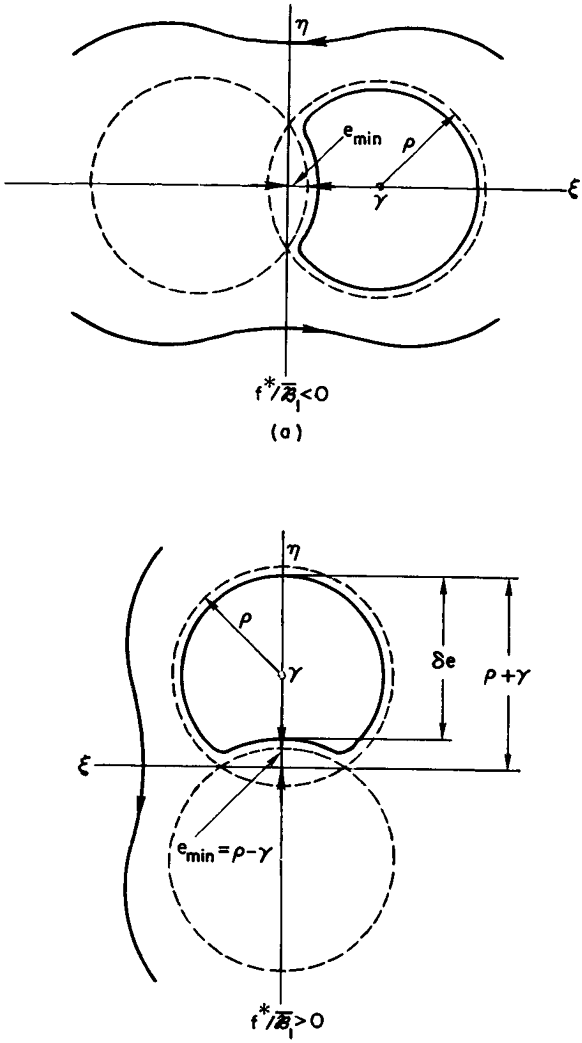
MAXIMUM VARIATION IN r_p

Figure 4



ξ–η CONTOURS

Figure 5



ξ–η CONTOURS FOR H_1 SLIGHTLY GREATER THAN H_{1c}
Figure 6

NASA-Langley, 1967 —— 30