AVAILABLE TO THE PUBLIC

ATMOSPHERIC AND NEAR PLANET TRAJECTORY OPTIMIZATION

BY THE VARIATIONAL STEEPEST-DESCENT METHOD

By Donald S. Hague

Distribution of this report is provided in the interest of information exchange. Responsibility for the contents resides in the author or organization that prepared it.

N70-32759

FACILITY FORM 602

(ACCESSION NUMBER)

124

(PAGES)

CR-73365

(NASA CR OR TMX OR AD NUMBER)

(THRU)

(CODE)

30

(CATEGORY)

ATMOSPHERIC AND NEAR PLANET TRAJECTORY OPTIMIZATION

BY THE VARIATIONAL STEEPEST-DESCENT METHOD

By Donald S. Hague

PREFACE

This document reports a study carried out by Aerophysics Research Corporation during the period from April to August, 1969, under Contract NAS 2-5383. The National Aeronautics and Space Administration Technical Monitor was Mr. Richard H. Petersen, Mission Analysis Division, Moffett Field, California. Mr. D. S. Hague functioned as Aerophysics Research Corporation's Project Leader for the study.

The trajectory program and basic steepest-descent optimization program which applies to multi-staged vehicles whose stage points occur at fixed times were originally developed under U.S. Air Force Funding Contracts AF 33(616)-6848 and AF 33(657)-8829. Mr. B. R. Benson of the Air Force Flight Dynamics Laboratory sponsored these developments in the period from 1959 to 1964. The optimization program was further extended under NASA Contract NAS 2-3691. Mr. Hubert Drake of the Mission Analysis Division, Moffett Field, California monitored the study.

Major extensions to program capability were undertaken in the period from 1964 to 1965 by the author while at McDonnell-Douglas Corporation, St. Louis. These extensions included solution of problems involving simultaneous determination of both optimal stage points and time varying control histories, the multiple-arc problem, and the extension to two-vehicle problems when the maneuvers of one vehicle are pre-determined, the "maneuvering target" problem. The program delivered to Ames Research Center under the present study does not include the optimal staging and maneuvering target capability. However, in the interest of information dissemination, the present report collects in one source the analytic approach employed in the point mass trajectory equations, the basic optimization formulation, the multiple-arc extension, and the maneuvering target extension.

Past contributors to the program development include Mr. Robert L. Mobley, now with the Fand Corporation, who programmed the optimization formulation and

    Mr. Robert C. Browne, McDonnell-Douglas Corp., St. Louis
    Mr. R. V. Brulle, McDonnell-Douglas Corp., St. Louis
    Mr. A. E. Combs, McDonnell-Douglas Corp., St. Louis
    Mr. K. N. Easley, now with Aerospace Corporation
    Mr. Ken Geib, McDonnell-Douglas Corp., St. Louis
    Mr. G. D. Griffin, now with Aerospace Corporation
    Mr. G. G. Grose, McDonnell-Douglas Corp., St. Louis
    Dr. H. L. Rozendaal, now with Lockheed Electronics, Houston
    Mr. F. W. Seubert, McDonnell-Douglas Corp., St. Louis
    Mr. N. E. Usher, McDonnell-Douglas Corp., St. Louis

Mrs. Jane Yonke of Aerophysics Research Corporation prepared
the present report.

The steepest-descent atmospheric trajectory optimization
programs described in this report are currently being extended
to the general two-vehicle trajectory optimization problem. In
the general case both vehicles exert time varying control with
either cooperative or conflicting objectives. This work will
be reported separately at a later date. Inquiries concerning
the development should be directed to Mr. B. R. Benson of the
Air Force Flight Dynamics Laboratory.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

ATMOSPHERIC AND NEAR PLANET TRAJECTORY OPTIMIZATION

BY THE VARIATIONAL STEEPEST-DESCENT METHOD

by Donald S. Hague
Aerophysics Research Corporation

SUMMARY

The variational steepest-descent method is described in detail. The analysis is first developed for a system consisting of a single-stage. Performance, constraint, and trajectory termination criteria may be any functions of terminal state and time.

The analysis is then extended to multi-stage systems. Here, a stage is bounded by two points selected from initial conditions, points of state derivative discontinuity, and terminal conditions. In the multi-stage analysis the stage-times replace the independent variable time as the independent variable time itself thus becomes an additional state variable.

The three-degree point-mass equations of motion for a vehicle maneuvering in the vicinity of a central planet are developed in rectangular rotating coordinates. Planetary characteristics include up to four harmonics in the gravitational field, flattening of the polar axis, layered atmosphere, and wind structure. Vehicle characteristics include generalized aerodynamic and thrust descriptions in terms of up to six control variables.

Weighting matrix or control variable metric tensor definitions are presented in some detail. The problem of false convergence induced by ill-chosen weighting matrices is discussed in terms of an order of magnitude analysis.

INTRODUCTION

Trajectory optimization by the Steepest-Descent Method is now a routine performance estimation at several government research establishments and major aerospace concerns. The computer program utilized for trajectory optimization studies in this report is capable of determining optimal three-dimensional flight paths for a wide variety of vehicles in the vicinity of a single planet. Atmospheric effects may be included, if desired. Past program applications include flight path optimization of

a. High performance supersonic aircraft
b. Spacecraft orbital transfer rendezvous and re-entry
c. Multi-stage booster ascent trajectories
d. Boost-glide re-entry vehicles
e. Advanced hypersonic cruise aircraft
f. Air-to-ground missiles.

Optimal control can be determined for any combination of the time varying variables

a. Angle-of-attack (or pitch angle)
b. Bank-angle
c. Side-slip
d. Throttle
e. Two thrust orientation angles

All the commonly employed terminal performance and constraint criteria may be specified. Inequality constraints may be imposed along the vehicle flight path.

Several options are available for specification of vehicle aerodynamic and propulsive options. Data and vehicle characteristics option can be modified at preselected stage points. An arbitrary number of stage points may be specified.

Planetary characteristics are nominally set to those of the earth. Up to four gravitational harmonics may be specified. Nominal planetary atmosphere employed is the 1959 ARDC. A variety of wind specification options are available. An ellipsoidal planetary shape may be specified.

The original trajectory optimization program is described in References 1 and 2. Equations of motion employed are described in References 3 and 4. Some past applications are described in References 5 and 6. An extension of program capability is described in Reference 7. An extension to simultaneously determine both optimal time varying control and discrete

stage points together with some applications are described in References 8 and 9. A guidance and control application, the so-called lambda-guidance scheme, is reported in Reference 10.

The optimization program of References 1 and 2 employs a second-order prediction scheme and several control variable "weighting matrix" options to assist convergence of the steepest-descent algorithm. These two features have also been included in a recently developed trajectory optimization, Reference 11. They are also retained as convergence options in an extended version of the program of Reference 1 and 2, reported in Reference 12.

THE STEEPEST DESCENT METHOD

Problem Statement

Point mass motion is governed by three second order differential equations of position together with a first order differential equation governing the mass. By suitably defining additional state variables, it is possible to reduce these equations to a set of first order differential equations. Point mass motion is, therefore, governed by a set of first order differential equations. The form of these equations is

$$\left\{\dot{x}_n(t)\right\} = \left\{f\left(x_n(t),\ \alpha_m(t),\ t\right)\right\}$$

$$n = 1,2 \ldots \ldots \ldots N$$
$$m = 1,2 \ldots \ldots \ldots M \qquad (1)$$

That is, there are N state variables whose derivatives $\dot{x}_n(t)$ are defined by N first order differential equations involving the state variables, together with M control variables, $\alpha_m(t)$, and t, the independent variable itself.

Constraints may be imposed on a set of functions of the state variables and time at the end of the trajectory. In this case, a set of constraint functions of the form

$$\left\{\psi_p\right\} = \left\{\psi_p\left(x_n(T),\ T\right)\right\} = 0$$

$$p = 1,2 \ldots \ldots \ldots P \qquad (2)$$

can be constructed which the final trajectory must satisfy. Any one of the constraints may be used as a cut-off function which, when satisfied, will terminate a particular trajectory. The cut-off function can, therefore, be written in the form

$$\Omega = \Omega\left(x_n(T),\ T\right) = 0 \qquad (3)$$

and determines the trajectory termination time T. In all, then, when the cut-off function is included, there are (P + 1) end constraints.

Finally, it may be that some other function of the state variables and time at the end of the trajectory is to be optimized. Hence, a pa_ off function

$$\phi = \phi\left(x_n(T)\right. \qquad (4)$$

which is to be maximized ed, can be constructed.

4

Now, suppose that a nominal trajectory is available. The requirements of this trajectory are modest; it must satisfy the cut-off condition, Equation (3), but it need not optimize the pay-off function or satisfy the constraint equations. To generate this nominal trajectory by integrating Equations (1), the vehicle characteristics, the initial state variable values, and a nominal control variable history must be known. Once this nominal trajectory is available, the steepest descent process can be applied. To do this, the trajectory showing the greatest improvement in the pay-off function, while at the same time eliminating a given amount of the end point errors as measured by Equations (2) for a given size of control variable perturbation, is obtained by application of the Variational Calculus.

Equations (2) provide an end point error measure, for they will only be satisfied if the end points have been achieved. Therefore, any non-zero ψ_p represents an end point error which must be corrected. A convenient measure of the control variable perturbation can be defined by the scalar quantity,

$$DP^2 = \int_{t_0}^{T} \lfloor \delta\alpha(t) \rfloor \left[W(t) \right] \left\{ \delta\alpha(t) \right\} dt \tag{5}$$

where W is any arbitrary symmetric matrix. In the case where all control variables have a similar ability to affect the trajectory, W is taken equal to the unit matrix, and DP^2 becomes the integrated square of the control variable perturbations $\delta\alpha(t)$. It might be noted that if Equation 5 is to have meaning, it is essential that all control variables have the same dimensions. To meet this condition, the control variables can be expressed in non-dimensional form.

The constraint on control variable perturbation size represented by Equation (5) is an essential element of the steepest descent process; for the optimum perturbation will be found by local linearization of the non-linear trajectory equations about the nominal path. To insure validity of the linearized approximation, the analysis must be limited to small control variable perturbations by means of Equation (5) which provides an integral measure of the local perturbation magnitudes.

Single Stage Analysis

The steepest descent process has been outlined above. To implement this method, an analysis of all perturbations about the nominal trajectory must be undertaken. In the present report, all perturbations will be linearized; only first order perturbations in the control and state variables will be considered. The objective of the linearized analysis is

-5-

determination of the optimum control variable perturbation in the sense discussed in the previous section.

Denoting variables on the nominal trajectory by a bar

$$\left\{ \alpha_m(t) \right\} \text{ nominal } = \left\{ \bar{\alpha}_m(t) \right\} \tag{6}$$

and

$$\left\{ x_n(t) \right\} \text{ nominal } = \left\{ \bar{x}_n(t) \right\} \tag{7}$$

where there are M control variables and N state variables.

Now consider a small perturbation to the control variable history, $\delta\alpha(t)$; this in turn will cause a small perturbation in the state variable history, $\delta x(t)$. The new values of the variables become

$$\left\{ \alpha(t) \right\} = \left\{ \bar{\alpha}(t) \right\} + \left\{ \delta\alpha(t) \right\} \tag{8}$$

and

$$\left\{ x(t) \right\} = \left\{ \bar{x}(t) \right\} + \left\{ \delta x(t) \right\} \tag{9}$$

The nominal state variable and perturbed state variable histories can also be written as

$$\left\{ \bar{x}(t) \right\} = \left\{ x(t_o) \right\} + \int_{t_o}^{t} \left\{ f\left(\bar{x}(t),\ \bar{\alpha}(t),\ t \right) \right\} dt \tag{10}$$

$$\left\{ x(t) \right\} = \left\{ x(t_o) \right\} + \int_{t_o}^{t} \left\{ f\left(\bar{x} + \delta x,\ \bar{\alpha} + \delta\alpha,\ t \right) \right\} dt \tag{11}$$

Subtracting Equation (10) from Equation (11) and using Taylor's expansion to first order,

$$\left\{ x(t) \right\} - \left\{ \bar{x}(t) \right\} = \int_{t_o}^{t} \left\{ \frac{\partial \bar{f}}{\partial x_n} \cdot \delta x^n + \frac{\partial \bar{f}}{\partial \alpha_m} \cdot \delta\alpha^m \right\} dt = \left\{ \delta x(t) \right\} \tag{12}$$

where

$$\bar{f} = f\left(\bar{x}(t),\ \bar{\alpha}(t),\ t \right) \tag{13}$$

and where the repeated index indicates a summation over all possible values. Differentiation leads to

$$\frac{d}{dt}\left\{ \delta x(t) \right\} = \left\{ \frac{\partial \bar{f}}{\partial x_n}\ \delta x^n + \frac{\partial \bar{f}}{\partial \alpha_m}\ \delta\alpha^m \right\} \tag{14a}$$

6

or in matrix form

$$\frac{d}{dt}\left\{\delta x(t)\right\} = \left[F\right]\left\{\delta x\right\} + \left[G\right]\left\{\delta \alpha\right\}$$ (14b)

where

$$F_{ij} = \frac{\partial \bar{f}_i}{\partial x_j} \quad \text{and} \quad G_{ij} = \frac{\partial \bar{f}_i}{\partial \alpha_j}$$ (15)

Here the $(i,j)^{th}$ element lies in the ith row and j^{th} column of the matrices; F is an N x N matrix and G is an N x M matrix.

The effect of these perturbations on pay-off, cut-off, and constraint functions must now be determined. A general method for obtaining these effects, known as the 'adjoint method,' Reference 13, is to define a new set of variables by the equations

$$\left[\dot{\lambda}(t)\right] = -\left[F(t)\right]'\left[\lambda(t)\right]$$ (16)

By specifying various boundary conditions on the λ, the changes in all functions of interest can be found in turn. To show this pre-multiply Equation (14) by λ' and Equation (16) by $\delta x'$, transpose the second of these equations and sum with the first giving

$$\left[\lambda\right]' \cdot \left\{\frac{d}{dt}(\delta x)\right\} + \left[\dot{\lambda}\right]'\left\{\delta x\right\} = \left[\lambda\right]'\left[F\right]\left\{\delta x\right\} + \left[\lambda\right]'\left[G\right]\left\{\delta \alpha\right\}$$
$$- \left[\lambda\right]'\left[F\right]\left\{\delta x\right\}$$ (17)

which may be written as

$$\left\{\frac{d}{dt}(\lambda'\delta x)\right\} = \left[\lambda\right]'\left[G\right]\left\{\delta \alpha\right\}$$ (18)

Integrating Equation (18) over the trajectory

$$\left\{\lambda'\delta x\right\}_T - \left\{\lambda'\delta x\right\}_0 = \int_{t_0}^{T}\left[\lambda\right]'\left[G\right]\left\{\delta \alpha\right\}dt$$ (19)

Now define three distinct sets of λ functions by applying the following boundary conditions at t = T:

$$\left\{\lambda(T)\right\} = \left\{\frac{\partial \phi}{\partial x_i}\right\}_T = \left\{\lambda_\phi(T)\right\}$$ (20a)

$$\left\{\lambda(T)\right\} = \left\{\frac{\partial \Omega}{\partial x_i}\right\}_T = \left\{\lambda_\Omega(T)\right\}$$ (20b)

$$\left[\lambda(T)\right] = \left[\frac{\partial \psi_j}{\partial x_i}\right]_T = \left[\lambda_\psi(T)\right]$$ (20c)

Equation (16) may now be integrated in the reverse direction (i.e., from T to t_o) to obtain the functions, $\{\lambda_\phi(t)\}$, $\{\lambda_\Omega(t)\}$, and $\{\lambda_\psi(t)\}$.

Substituting each of these functions into Equation (19) in turn and noting that

$$\lfloor \lambda_\phi(T) \rfloor \{\delta x\} = \lfloor \frac{\partial \phi}{\partial x} \rfloor \{\delta x\} = \delta\phi_{t=T} \tag{21a}$$

$$\lfloor \lambda_\Omega(T) \rfloor \{\delta x\} = \lfloor \frac{\partial \Omega}{\partial x} \rfloor \{\delta x\} = \delta\Omega_{t=T} \tag{21b}$$

$$\left[\lambda_\psi(T) \right]' \{\delta x\} = \left[\frac{\partial \psi_i}{\partial x_j} \right] \{\delta x\} = \{\delta\psi_{t=T}\} \tag{21c}$$

It follows that

$$\delta\phi_{t=T} = \int_{t_o}^{T} \lfloor \lambda_\phi \rfloor \left[G \right] \{\delta\alpha\} dt + \lfloor \lambda_\phi(t_o) \rfloor \{\delta x(t_o)\} \tag{22a}$$

$$\delta\Omega_{t=T} = \int_{t_o}^{T} \lfloor \lambda_\Omega \rfloor \left[G \right] \{\delta\alpha\} dt + \lfloor \lambda_\Omega(t_o) \rfloor \{\delta x(t_o)\} \tag{22b}$$

$$\{\delta\psi\}_{t=T} = \int_{t_o}^{T} \left[\lambda_\psi \right]' \left[G \right] \{\delta\alpha\} dt + \left[\lambda_\psi(t_o) \right]' \{\delta x(t_o)\} \tag{22c}$$

Now, Equations (22) give the changes in pay-off function, cut-off function and constraint functions at the terminal time of the nominal trajectory; however, on the perturbed trajectory, the cut-off will usually occur at some perturbed time, $T + \Delta T$. In this case, the total change in the above quantities becomes

$$d\phi = \int_{t_o}^{T} \lfloor \lambda_\phi \rfloor \left[G \right] \{\delta\alpha\} dt + \lfloor \lambda_\phi(t_o) \rfloor \{\delta x(t_o)\} + \dot{\phi}(T) \Delta T \tag{23a}$$

$$d\Omega = \int_{t_o}^{T} \lfloor \lambda_\Omega \rfloor \left[G \right] \{\delta\alpha\} dt + \lfloor \lambda_\Omega(t_o) \rfloor \{\delta x(t_o)\} + \dot{\Omega}(T) \Delta T \tag{23b}$$

$$\{d\psi\} = \int_{t_o}^{T} \left[\lambda_\psi \right]' \left[G \right] \{\delta\alpha\} dt + \left[\lambda_\psi(t_o) \right]' \{\delta x(t_o)\} + \{\dot{\psi}(T)\} \Delta T \tag{23c}$$

8

Equations (23) supply the change in pay-off, cut-off, and constraint functions on the perturbed trajectory.

The time perturbation in Equations (23a) and (23c) may be eliminated by noting that, by definition of the cut-off function, Equation (23b) must be zero.

$$\therefore \Delta T = -\frac{1}{\dot{\Omega}(T)}\left(\int_{t_0}^{T}\lfloor\lambda_\Omega\rfloor\left[G\right]\{\delta\alpha\}\,dt + \lfloor\lambda_\Omega(t_0)\rfloor\{\delta x(t_0)\}\right) \tag{24}$$

Substituting Equation (24) into Equations (23a) and (23c)

$$d\phi = \int_{t_0}^{T}\lfloor\lambda_{\phi\Omega}\rfloor\left[G\right]\{\delta\alpha\}\,dt + \lfloor\lambda_{\phi\Omega}(t_0)\rfloor\{\delta x(t_0)\} \tag{25a}$$

$$\{d\psi\} = \int_{t_0}^{T}\left[\lambda_{\psi\Omega}\right]'\left[G\right]\{\delta\alpha\}\,dt + \left[\lambda_{\psi\Omega}(t_0)\right]'\{\delta x(t_0)\} \tag{25b}$$

where

$$\{\lambda_{\phi\Omega}\} = \{\lambda_\phi\} - \frac{\dot{\phi}(T)}{\dot{\Omega}(T)}\{\lambda_\Omega\} \tag{26a}$$

$$\left[\lambda_{\psi\Omega}\right]' = \left[\lambda_\psi\right]' - \frac{\{\dot{\psi}(T)\}\lfloor\lambda_\Omega\rfloor}{\dot{\Omega}(T)} \tag{26b}$$

Equations (25) reveal the significance of the λ functions, originally defined by Equations (16) and (20). At time t_0, $\lambda_{\phi\Omega}$ gives the sensitivity of $\phi(T)$ to small perturbations in the state variables at t_0. Similarly, $\lambda_{\phi\Omega}(t)$ measures the sensitivity of $\phi(T)$ to small perturbations in the state variables at any time t. The sensitivity of the constraints $d\psi$ to small state variable perturbations at any time is likewise defined by each row of the function $\lambda_{\psi\Omega}(t)$.

A measure of the sensitivity of a trajectory to control variable perturbations can be obtained from the quantities $\lambda_\phi'_\Omega G$ and $\lambda_\psi'_\Omega G$. Consider a pulse control variable perturbation at time t', that is, $\delta(t-t')$, where δ is the Dirac delta function. With this type of control variable perturbation, it can be seen from Equations (25) that the changes in pay-off and constraint functions will be $\lambda_{\phi\Omega}(t')'G(t')$ and $\lambda_{\psi\Omega}(t')'G(t')$, respectively, for fixed initial conditions.

In order to apply the steepest-descent process, the performance function change, Equation (20a), must be maximized; subject to specified changes in the constraints, Equation (25b); and a

given size perturbation to the control variables, Equation (5).
This can be achieved by constructing an augmented function in
the manner of Lagrange which is to be maximized instead of $d\phi$.
For the present problem, the augmented function is

$$
U = \int_{t_0}^{T} \lfloor \lambda_{\phi\Omega} \rfloor [G] \{\delta\alpha\} \, dt + \lfloor \lambda_{\phi\Omega}(t_0) \rfloor \{\delta x(t_0)\}
$$

$$
+ \lfloor \upsilon \rfloor \left\{ \int_{t_0}^{T} \left[\lambda_{\psi\Omega} \right]' [G] \{\delta\alpha\} \, dt + \left[\lambda_{\psi\Omega}(t_0) \right]' \{\delta x(t_0)\} \right\}
$$

$$
+ \mu \int_{t_0}^{T} \lfloor \delta\alpha \rfloor [W] \{\delta\alpha\} \, dt
$$

(27)

where the υ are P undetermined Lagrangian multipliers, and μ
is a single undetermined Lagrangian multiplier. The objective
now is to find that variation of the control variable history
which will maximize U.

Consider a variation of $\delta\alpha$, that is a $\delta(\delta\alpha)$. Then,
it is always possible to write any $\delta\alpha$ distribution in the form

$$
\{\delta\alpha\} = \{A(t)\} \, k, \quad \text{or} \quad \lfloor \delta\alpha \rfloor = \lfloor A(t) \rfloor k
$$

(28)

where A(t) prescribes the perturbation shape;and k, its mag-
nitude. Now that part of Equation (27) which depends on $\delta\alpha$,
the perturbation in the control variable, can be written in
the form

$$
\bar{U} = k \int_{t_0}^{T} \lfloor \lambda_{\phi\Omega} \rfloor [G] \{A(t)\} \, dt + k \lfloor \upsilon \rfloor \int_{t_0}^{T} \left[\lambda_{\psi\Omega} \right]' [G] \{A(t)\} \, dt
$$

$$
+ k^2 \mu \int_{t_0}^{T} \lfloor A(t) \rfloor [W] \{A(t)\} \, dt
$$

(29)

So that

$$
\frac{\partial \bar{U}}{\partial k} = \int_{t_0}^{T} \lfloor \lambda_{\phi\Omega} \rfloor [G] \{A(t)\} \, dt + \lfloor \upsilon \rfloor \int_{t_0}^{T} \left[\lambda_{\psi\Omega} \right]' [G] \{A(t)\} \, dt
$$

(30)

$$
+ 2k\mu \int_{t_0}^{T} \lfloor A(t) \rfloor [W] \{A(t)\} \, dt
$$

10

or

$$\delta\bar{U} = \int_{t_o}^{T} \left(\left\lfloor \lambda_{\phi\Omega} \right\rfloor \left[G \right] \left\{ \delta k \cdot A(t) \right\} + \left\lfloor \nu \right\rfloor \left[\lambda_{\psi\Omega} \right]' \left[G \right] \left\{ \delta k \cdot A(t) \right\} \right.$$
$$\left. + 2\mu \left\lfloor k \cdot A(t) \right\rfloor \left[W \right] \left\{ \delta k \cdot A(t) \right\} \right) \, dt$$

$$= \int_{t_o}^{T} \left\lfloor \left\lfloor \lambda_{\phi\Omega} \right\rfloor \left[G \right] + \left\lfloor \nu \right\rfloor \left[\lambda_{\psi\Omega} \right]' \left[G \right] + 2\mu \left\lfloor \delta\alpha \right\rfloor \left[W \right] \right\rfloor \left\{ \delta(\delta\alpha) \right\} \, dt$$

(31)

where it has been noted from Equation (28) that

$$\delta(\delta\alpha) = A(t) \, \delta k \tag{32}$$

Now, since Equation (31) holds for any $A(t)$, it follows that it is a general relationship. Further, for $\bar{U}$ to be an extremal, $\delta\bar{U}$ must be zero.

If $\bar{U}$ has been maximized by means of a control variable perturbation $\delta\alpha$, $\delta\bar{U}$ must be stationary for all small perturbations to the $\delta\alpha$, that is, for all $\delta(\delta\alpha)$. The only way in which Equation (31) can be zero for all $\delta(\delta\alpha)$ is for the coefficient of $\delta(\delta\alpha)$ to be identically zero. That this last statement is true follows from considering the case where, over some finite time interval between t_o and T, the coefficient of $\delta(\delta\alpha)$ is, say, positive. If this were the case, we could choose a $\delta(\delta\alpha)$ distribution that was also positive in this same interval and zero elsewhere between t_o and T. It would follow that $\bar{U}$ was also positive, and, hence, $\bar{U}$ could not be maximum. A similar argument holds when $\delta(\delta\alpha)$ is negative over any interval in t_o to T. Hence, the coefficient of $\delta(\delta\alpha)$ must be identically zero in the whole interval $t_o \leq t \leq T$. This argument is essentially based on that presented by Goldstein, Reference 14. It follows that

$$\left\lfloor \left\lfloor \lambda_{\phi\Omega} \right\rfloor + \left\lfloor \nu \right\rfloor \left[\lambda_{\psi\Omega} \right]' \right\rfloor \left[G \right] = -2\mu \left\lfloor \delta\alpha \right\rfloor \left[W \right] \tag{33}$$

Transposing, noting that W is symmetric, and solving for $\delta\alpha$,

$$\left\{\delta\alpha\right\} = -\frac{1}{2\mu}\left[W\right]^{-1}\left[G\right]'\left\{\left\{\lambda_{\phi\Omega}\right\} + \left[\lambda_{\psi\Omega}\right]\left\{\nu\right\}\right\}$$ (34)

Substituting Equation (34) into Equation (25b)

$$\left\{d\beta\right\} = -\frac{1}{2\mu}\left\{\left\{I_{\psi\phi}\right\} + \left[I_{\psi\psi}\right]\left\{\nu\right\}\right\}$$ (35a)

where

$$\left\{d\beta\right\} = \left\{d\psi\right\} - \left[\lambda_{\psi\Omega}(t_o)\right]'\left\{\delta x(t_o)\right\}$$ (35b)

and

$$\left[I_{\psi\psi}\right] = \int_{t_o}^{T}\left[\lambda_{\psi\Omega}\right]'\left[G\right]\left[W\right]^{-1}\left[G\right]'\left[\lambda_{\psi\Omega}\right]dt$$ (36a)

$$\left\{I_{\psi\phi}\right\} = \int_{t_o}^{T}\left[\lambda_{\psi\Omega}\right]'\left[G\right]\left[W\right]^{-1}\left[G\right]'\left\{\lambda_{\phi\Omega}\right\}dt$$ (36b)

For subsequent use define the integral

$$I_{\phi\phi} = \int_{t_o}^{T}\left\lfloor\lambda_{\phi\Omega}\right\rfloor\left[G\right]\left[W\right]^{-1}\left[G\right]'\left\{\lambda_{\phi\Omega}\right\}dt$$ (36c)

The multipliers ν can be expressed in terms of the multipliers μ by Equation (35a)

$$\left\{\nu\right\} = -\left[I_{\psi\psi}\right]^{-1}\left\{2\mu\left\{d\beta\right\} + \left\{I_{\psi\phi}\right\}\right\}$$ (37)

Substituting Equation (34) into Equation (5)

$$DP^2 = \frac{1}{4\mu^2}\left(I_{\phi\phi} + \left\lfloor I_{\psi\phi}\right\rfloor\left\{\nu\right\} + \left\lfloor\nu\right\rfloor\left\{I_{\psi\phi}\right\} + \left\lfloor\nu\right\rfloor\left[I_{\psi\psi}\right]\left\{\nu\right\}\right)$$ (38)

Transposing the second term in the right hand side bracket

$$DP^2 = \frac{1}{4\mu^2}\left(I_{\phi\phi} + 2\left\lfloor\nu\right\rfloor\left\{I_{\psi\phi}\right\} + \left\lfloor\nu\right\rfloor\left[I_{\psi\psi}\right]\left\{\nu\right\}\right)$$ (39)

Substituting Equation (37) in Equation (39)

and noting that $\left[I_{\psi\psi}\right]^{-1}$ is symmetrical gives

$$4\mu^2\, DP^2 = I_{\phi\phi} - \lfloor I_{\psi\phi} \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\} + 4\mu^2 \lfloor d\beta \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ d\beta \right\} \qquad (40)$$

So that

$$2\mu = \pm \sqrt{\frac{I_{\phi\phi} - \lfloor I_{\psi\phi} \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\}}{DP^2 - \lfloor d\beta \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ d\beta \right\}}} \qquad (41)$$

Substituting Equation (41) into Equation (37), the remaining Lagrangian multipliers are obtained in the form

$$\left\{ \nu \right\} = -\left[I_{\psi\psi} \right]^{-1} \left\{ \left\{ I_{\psi\phi} \right\} \pm \sqrt{\frac{I_{\phi\phi} - \lfloor I_{\psi\phi} \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\}}{DP^2 - \lfloor d\beta \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ d\beta \right\}}} \left\{ d\beta \right\} \right\} \qquad (42)$$

The optimum control perturbation is found by substituting Equations (41) and (42) back into Equation (34) and is

$$\left\{ \delta\alpha \right\} = \pm \left[W \right]^{-1} \left[G \right]' \left\{ \left\{ \lambda_{\phi\Omega} \right\} - \left[\lambda_{\psi\Omega} \right] \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\} \right\}$$

$$\times \sqrt{\frac{DP^2 - \lfloor d\beta \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ d\beta \right\}}{I_{\phi\phi} - \lfloor I_{\psi\phi} \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\}}}$$

$$+ \left[W \right]^{-1} \left[G \right]' \left[\lambda_{\psi\Omega} \right] \left[I_{\psi\psi} \right]^{-1} \left\{ d\beta \right\} \qquad (43)$$

With this equation the steepest-descent control perturbation has been determined. Perturbing the control variables according to Equation (43) gives the optimum change in the trajectory as discussed in the section entitled, "Problem Statement," with the added effect of changes in the initial value of the state variables included through the term in $d\beta$. The appropriate sign to use on the first term of equation (43) can be determined by evaluating $d\phi$. Substituting the optimum control perturbation into Equation (25a) results in the equation shown on the following page.

$$d\phi = \pm\sqrt{\left(I_{\phi\phi} - \left\lfloor I_{\psi\phi}\right\rfloor\left\lfloor I_{\psi\psi}\right\rfloor^{-1}\left\{I_{\psi\phi}\right\}\right)\left(DP^2 - \left\lfloor d\beta\right\rfloor\left\lfloor I_{\psi\psi}\right\rfloor^{-1}\left\{d\beta\right\}\right)}$$
$$+ \left\lfloor I_{\psi\phi}\right\rfloor\left\lfloor I_{\psi\psi}\right\rfloor^{-1}\left\{d\beta\right\} + \left\lfloor \lambda_{\phi\Omega}(t_o)\right\rfloor\left\{\delta x(t_o)\right\} \qquad (44)$$

As the quantity in the radical must be positive to assure the change in ϕ is real, it follows that the negative sign must be taken when minimizing the payoff function and the positive sign when maximizing the payoff function.

An Alternative Analysis Using the Independent Variable for Cut-Off

In the analysis of the previous section, it is implied that any function of the form

$$\Omega(x_n(T), T) = 0 \qquad (45)$$

will suffice to terminate the trajectory. While this is true in an analytic sense, in practice any function passing through zero more than once in the cut-off region may be difficult to employ for cut-off purposes.
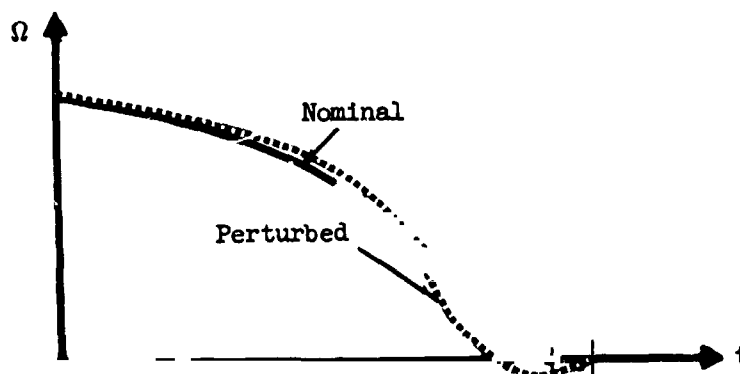
Figure 1.-- Double Valued Cut-Off Function

Figure 1 presents a nominal cut-off function history which decreases monotonically. The perturbed cut-off function history, shown dotted, behaves in a different manner in that it passes through zero twice in the cut-off region. As the trajectory must be integrated numerically, there is a danger that cut-off will occur the first time Ω passes through zero instead of the second, thereby introducing both errors in the linearized perturbations and preventing the build-up of the anticipated cut-off function history.

One method for overcoming this difficulty is to terminate the trajectory at the linearly predicted cut-off value of the independent variable, $\lambda + \Delta T$. This revises the analysis of the previous section; for by terminating the trajectory in this manner, a small error will exist in the value of the cut-off function, say $\Delta \Omega$. In this case, Equation (24) must be modified to account for the cut-off function error. Allowing for this effect, we can write in place of Equation (24)

$$\Delta T = \frac{\left(\Delta \Omega - \int_{t_0}^{T} \lfloor \lambda_\Omega \rfloor [G] \{\delta\alpha\} dt - \lfloor \lambda_\Omega(t_0) \rfloor \{\delta x(t_0)\} \right)}{\dot{\Omega}(T)}$$

(46)

Substituting into Equations (23a) and (23b)

$$d\phi = \int_{t_0}^{T} \lfloor \lambda_{\phi\Omega} \rfloor [G] \{\delta\alpha\} dt + \lfloor \lambda_{\phi\Omega}(t_0) \rfloor \{\delta x(t_0)\} + \frac{\dot{\phi}}{\dot{\Omega}} \Delta\Omega$$

(47a)

and

$$\{d\psi\} = \int_{t_0}^{T} \lfloor \lambda_{\psi\Omega} \rfloor [\cdot] \{\delta\alpha\} dt + \lfloor \lambda_{\psi\Omega}(t_0) \rfloor' \{\delta x(t_0)\} + \frac{\{\dot{\psi}(T)\}}{\dot{\Omega}(T)} \Delta\Omega$$

(47b)

The additional term plays no part in the equations which result from taking a variation in $\delta\alpha$; hence, the remaining equations in this section up to Equation (34) are still correct. Equation (34) must now be substituted into Equation (47b) instead of Equation (25b) with the result

$$\{d\beta\} - \frac{\{\dot{\psi}(T)\}}{\dot{\Omega}(T)} \Delta\Omega = \frac{-1}{2\mu} \left\{ \{I_{\psi\phi}\} + [I_{\psi\nu}] \{\nu\} \right\}$$

(48)

Defining $d\beta^*$ by

$$\{d\beta^*\} = \{d\beta\} - \frac{\{\dot{\psi}(T)\}}{\dot{\Omega}(T)} \Delta\Omega$$

(49)

and substituting $d\beta^*$ for $d\beta$ in the remainder of the analysis of the previous section, the entire analysis remains correct up to and including Equation (43). Equation (43) still gives the optimum control variable perturbation. The change in ϕ, however, becomes

$$d\phi = \mp\sqrt{\left(I_{\phi\phi} - \lfloor I_{\psi\phi}\rfloor\left[I_{\psi\psi}\right]^{-1}\{I_{\psi\phi}\}\right)\left(DF^2 - \lfloor d\beta^*\rfloor\left[I_{\psi\psi}\right]^{-1}\{d\beta^*\}\right)}$$

$$+ \lfloor I_{\psi\phi}\rfloor\left[I_{\psi\psi}\right]^{-1}\{d\beta^*\} + \lfloor\lambda_{\phi\Omega}(t_0)\rfloor\{\delta x(t_0)\} + \frac{\dot\phi(T)}{\dot\Omega(T)}\Delta\Omega$$

(50)

16

MULTI-STAGE ANALYSIS

Outline of Multi-Stage Analysis

Trajectories in which some of the state variables or state variable derivatives have a discontinuity for some value, t', of the independent variable, t, are frequently encountered. Such a point will be called a stage point. That portion of a trajectory preceding a stage point is in a different "stage" to that following the stage point. The first stage will be that portion of a trajectory lying between t_0 and the first stage point; the sth stage will be that portion of the trajectory lying between the $(s-1)^{th}$ and s^{th} stage points. The final stage will be that portion of the trajectory lying between the last stage point and the final cut-off time.

It is convenient in the analysis of multi-stage trajectories to define a new independent variable to replace t. This new independent variable, the stage time τ, is defined separately for each stage in the following manner. Let the sth stage commence at time t'_{s-1} and terminate at time t'_s. Then,

$$\tau_s = t - t'_{s-1} \tag{51}$$

so that when

$$t = t'_{s-1} \ , \quad \tau_s - 0 \tag{52}$$

The sth stage is terminated by a cut-off function Ω_s, assumed to be of the form

$$\Omega_s = \Omega_s\left(x(T_s), \ T_s \right) = 0 \tag{53}$$

where T_s is the stage time at cut-off.

The analysis of the preceding section, pages 4 to 16, no longer holds for a staged trajectory, unless the stage points are determined by cut-off functions of the form

$$\Omega_s \ (T_s) = 0 \tag{54}$$

That is, the stages are of fixed length in the independent variable.

For suppose the nominal trajectory has an sth stage lying in the region

$$t'_{s-1} \leq t \leq t'_s \tag{55}$$

Then on the perturbed trajectory unless the sth stage and the $(s-1)^{th}$ stage are terminated in the manner of Equation (54), the sth stage will occupy the region

$$t'_{s-1} + \Delta t'_{s-1} \leq t \leq t'_s + \Delta t'_s \tag{56}$$

by virtue of the perturbations in the state variables. This, in turn, will mean that estimates of the optimum α-perturbation based on the analysis of the preceding section, pages 4 to 16 will be in error due to the fact that the F and G matrices are incorrect in the regions between

$$(t'_{s-1}, \ t'_{s-1} + \Delta t'_{s-1})$$

and

$$(t'_s, \ t'_s + \Delta t'_s)$$

In such a situation, a new factor enters the optimization problem; for control may be exercised over the position of all or some of the stage points. In this case not just the optimum control variable perturbation is sought by the analyst but rather the optimum combination of control variable and stage point perturbations when considered simultaneously. The objective in the following section is to obtain this optimum combined perturbation.

Changes in Payoff and Constraint Functions in Combined Perturbation

Given a nominal multi-stage trajectory, suppose the control variable histories and the stage point positions are simultaneously perturbed throughout the whole trajectory. Considering the first stage, the effect of this combined perturbation, when the stage terminates, will appear as a modification in the state variables values. This perturbation in the first stage final state variable values may be looked upon as a perturbation in the initial state variable values to the second stage. The combined effect of stage point, control variables, and initial state variable perturbations in the second stage will be to produce a perturbation in the state variable values at the second stage termination. These effects can be in like manner until the last stage is reached. Perturbations in a typical stage are illustrated in Figure 2.
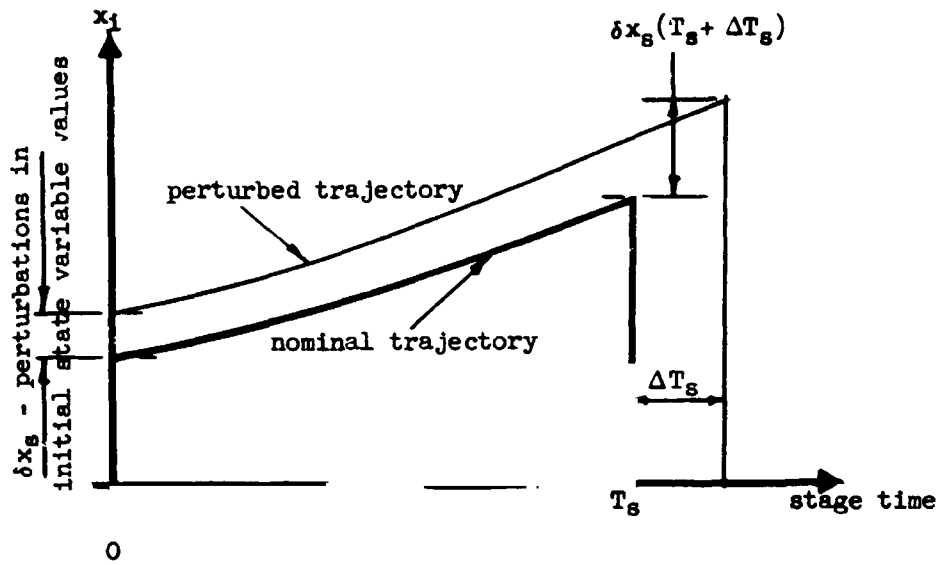
Figure 2.-- Perturbations in the sth Stage

Consider the last stage. Since the result of trajectory perturbations ahead of this stage appear solely as initial state variable perturbations, the optimum combined perturbation along the whole trajectory can be found by optimizing the last stage in the presence of initial state variable perturbations which are a function of the previous stage perturbations.

Consider the sth stage of the perturbed trajectory. As noted above, the effect of perturbations in all stages preceding the sth appear as some perturbation in the initial state variable values of the sth stage, δx_s. Suppose the sth stage is terminated on the nominal trajectory by some function of the form

$$\Omega_s = \Omega_s\left(x(T_s), T_s\right) = 0 \tag{57}$$

Then, provided the stage time, τ_s, is used as the independent variable instead of t, the change in any function of the state variables and τ_s at cut-off can be found by

$$\left\{\delta\psi(T_s)\right\} = \int_0^{T_s}\left[\lambda_{\psi\Omega_s}\right]'\left[G\right]\left\{\delta\alpha\right\}d\tau_s + \left[\lambda_{\psi\Omega_s}(0)\right]'\left\{\delta x_s(0)\right\}$$

where ψ_s is any function of the form

$$\left\{\psi_s\right\} = \left\{\psi_s\left(x(T_s), T_s\right)\right\} \tag{58}$$

19

and

$$\left[\lambda \Psi \Omega_s\right] = \left[\lambda \Psi_s\right] - \left\{\lambda \Omega_s\right\} \frac{\left|\dot{\Psi}_s(T_s)\right|}{\dot{\Omega}_s(T_s)} \tag{59}$$

Here $\lambda \psi_s$ and $\lambda \Omega_s$ are obtained by integrating the adjoint equations in stage time through the sth stage subject to the boundary conditions

$$\left[\lambda \Psi_s(T_s)\right] = \left[\frac{\partial \Psi_s}{\partial x_j}\right]' \tag{60}$$

$$\left\{\lambda \Omega_s(T_s)\right\} = \left\{\frac{\partial \Omega_s}{\partial x_j}\right\} \tag{61}$$

So far, the cut-off function has not been perturbed; this can be achieved by terminating the trajectory when

$$\Omega_s + \Delta \Omega_s = 0 \tag{62}$$

instead of by Equation (53). Perturbing the cut-off function will cause a change in the trajectory stage time at cut-off given by,

$$\Delta T_s = \frac{\Delta \Omega_s}{\dot{\Omega}_s(T_s)} \tag{63}$$

The total change in ψ_s at the termination of the sth stage will then be given by

$$\left\{\delta \Psi_s(T_s + \Delta T_s)\right\} = \int_0^{T_s} \left[\lambda \Psi \Omega_s\right]' \left[G\right]\left\{\delta \alpha\right\} d\tau_s + \left[\lambda \Psi \Omega_s(0)\right]' \left\{\delta x_s(0)\right\}$$

$$+ \left\{\dot{\Psi}_s(T_s)\right\} \Delta T_s \tag{64}$$

Suppose the state variables are selected as ψ_s, so that

$$\left\{\Psi_s\right\} = \left\{x(T_s)\right\} \tag{65}$$

With this choice of ψ_s, Equation (60) becomes

$$\left[\lambda \Psi_s(T_s)\right] = \left[\frac{\partial x_i}{\partial x_j}\right]' = \left[I\right] \tag{66}$$

where I is the unit matrix. Denoting the $\lambda \psi_s$ resulting from this particular choice of boundary conditions by λx_s, it follows from Equation (64) that the change in state variables

at termination of the sth stage is

$$\left\{ \delta x_s(T_s + \Delta T_s) \right\} = \int_0^{T_s} \left[\lambda_{x\Omega_s} \right]' \left[G \right] \left\{ \delta \alpha \right\} d\tau_s + \left[\lambda_{x\Omega_s}(0) \right]' \left\{ \delta x_s(0) \right\} + \left\{ \dot{x}(T_s) \right\} \Delta T_s$$

(67)

These perturbations are the state variable changes to the left of s + 1 stage point.

The δ x to the right of a stage point, $\delta x^{(+)}$ are not necessarily equal to those on the left, $\delta x^{(-)}$, Figure 3, but a matrix P_s can be defined which will transform the left-hand perturbations into the right hand ones.

$$\left\{ \delta x_{s+1}^{(+)} \right\} = \left[P_s \right] \left\{ \delta x_{s+1}^{(-)} \right\} = \left[P_s \right] \left\{ \delta x_s(T_s + \Delta T_s) \right\}$$

(68)

Typically, for example, consider the case of a multi-stage boost vehicle with a fixed amount of fuel in each stage. The mass of the remaining portions of the vehicle at the commencement of the sth stage is the sum of the empty weights of the remaining stages, together with the sum of the fuel contained in those stages. Perturbations in mass at the termination of the $(s-1)^{th}$ stage reflect changes in the burning time of that stage. It will usually be physically impossible to transfer any fuel remaining in the $(s-1)^{th}$ stage across the interface with the sth stage, and, hence, changes in the state variable of mass to the left of the stage point may fail to cause a corresponding change to the right of the stage point. In such a case, the P matrix will have a null row for that particular state variable. On the other hand, changes in the state variables of position to the left of a stage point will always appear directly as changes to the right of a stage point. The corresponding row in the P matrix will have unity on the diagonal element and zero elsewhere. This is also true of the state variables of velocity, provided impulsive forces are absent at the stage point.

Substituting Equation (67) into Equation (68)

$$\left\{ \delta x_{s+1} \right\} = \left[P_s \right] \left\{ \int_0^{T_s} \left[\lambda_{x\Omega_s} \right]' \left[G \right] \left\{ \delta \alpha \right\} d\tau_s + \left[\lambda_{x\Omega_s}(0) \right]' \left\{ \delta x_s(0) \right\} + \left\{ \dot{x}(T_s) \right\} \Delta T_s \right\}$$

$$= \left[P_s \right] \left\{ \left\{ K_s \right\} + \left[\bar{\lambda}_{x\Omega_s} \right]' \left\{ \delta x_s(0) \right\} + \left\{ \dot{\bar{x}}_s \right\} \Delta T_s \right\}$$

(69)

where

$$\left\{ K_s \right\} = \int_0^{T_s} \left[\lambda_{x\Omega_s} \right]' \left[G \right] \left\{ \delta\alpha \right\} d\tau_s \tag{70}$$

$$\left[\overline{\lambda}_{x\Omega_s} \right] = \left[\lambda_{x\Omega_s}(0) \right] \tag{71}$$

and

$$\left\{ \overline{\dot{x}}_s \right\} = \left\{ \dot{x}_s(T_s) \right\} \tag{72}$$
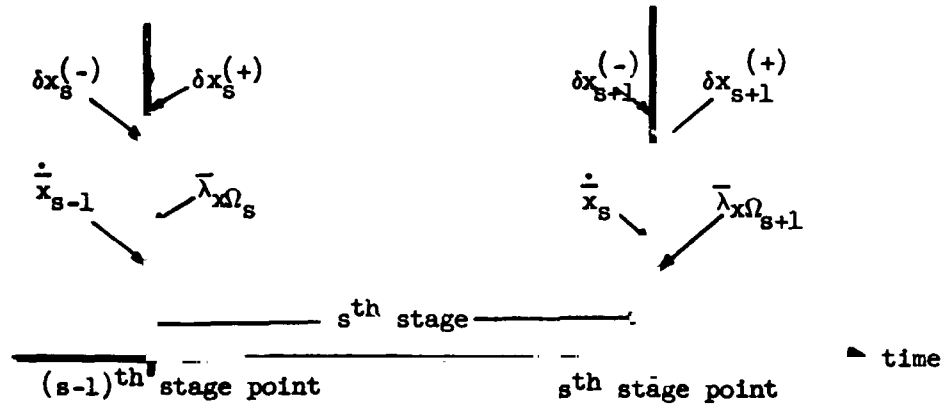
Figure 3.-- Position of Functions Defined in the sth Stage

In some cases additional perturbations in the state variables to the right of a stage point may be specified. For example, returning to the case of a multi-stage booster, in a more sophisticated analysis of booster capability, one may wish to consider variations in the initial mass of fuel contained within each stage. Typically, in a given iteration the amount of fuel consumed in the $s+1th$ stage may be either less than or greater than the total amount of fuel available in that stage. If this or a similar situation arises, the initial amount of fuel contained in the stage must be adjusted in the next iteration. It is essential to have a mechanism within the optimization analysis which will permit these required changes to be specified.

The P matrix, as described above, is unable to provide this mechanism, for the additional changes may clearly be functions of the state variable perturbations at the termination of the $(s+1)^{th}$ stage rather than at the termination of the sth stage. The P matrix is primarily introduced to convert changes in the state variables at the termination of the

s^{th} stage into state variable changes at the beginning of the $(s+1)^{th}$ stage. Accordingly, a set of additional state variable perturbations $\{\Delta x_{s+1}\}$ which are specified directly may be introduced. The complete expression for the state variable perturbations, therefore, becomes

$$\left\{\delta x_{s+1}\right\} = \left[P_s\right]\left\{\left\{K_s\right\} + \left[\overline{\lambda}_{x\Omega_s}\right]'\left\{\delta x_s(0)\right\} + \left\{\dot{\overline{x}}_s\right\}\Delta T_s\right\} + \left\{\Delta x_{s+1}\right\}$$
(73)

With Equation (73) the first objective of this analysis is achieved: a recursion formula which determines initial state variable perturbations in the $(s+1)^{th}$ stage when the perturbations in the sth stage are known and additional changes are directly specified in the initial values of the $(s+1)^{th}$ stage state variables.

The recursion formula Equation (73) can be applied to each stage in turn, commencing with the first. In the case of the first stage, there are no perturbations in the initial state due to prior stages, but there may be perturbations to the state variables if a search for the optimum trajectory initial conditions is being made.

These initial value perturbations will be some combination of state variable vectors dictated by the particular problem under consideration. For example, suppose the optimal launch point for a mobile mission is sought. The nominal launch point may be specified in terms of latitude and longitude. On successive iterations the launch point is perturbed towards the optimal position. The perturbation in latitude is a state variable vector having components corresponding to the position state variables and zeros elsewhere

$$\left\{\Delta x_1^{\phi}\right\} = \begin{Bmatrix} \dfrac{\partial \phi}{\partial X_e} \\ \dfrac{\partial \phi}{\partial Y_e} \\ \dfrac{\partial \phi}{\partial Z_e} \\ ----- \\ 0 \end{Bmatrix}$$
(74a)

Here X_e, Y_e, and Z_e are rectangular coordinates rotating with the planet, as discussed in a later section, pages 45 to 47.

Similarly, the components for longitude changes are

$$\left\{\Delta x_1^{\theta}\right\} = \begin{Bmatrix} \dfrac{\partial \theta}{\partial X_e} \\ \dfrac{\partial \theta}{\partial Y_e} \\ \dfrac{\partial \theta}{\partial Z_e} \\ --- \\ 0 \end{Bmatrix}$$
(74b)

In addition to this type of initial point perturbation, there may be changes resulting from previous iterations, Δx_1^P. Combining both types of perturbation, the total change in the first stage state variables is

$$\left\{\delta x_1\right\} = \left\{\Delta x_1\right\} = \sum_1 \left\{\Delta x_1^i\right\} + \left\{\Delta x_1^P\right\} \tag{75}$$

Knowing the δx_1, the total initial value perturbations in the second stage may be computed from Equation (73).

$$\left\{\delta x_2\right\} = \left[P_1\right]\left\{\left\{K_1\right\} + \left[\bar{\lambda}_{x\Omega_1}\right]' \left\{\delta x_1\right\} + \left\{\dot{\bar{x}}_1\right\}\Delta T_1\right\} + \left\{\Delta x_2\right\} \tag{76}$$

Proceeding to the next stage,

$$\left\{\delta x_3\right\} = \left[P_2\right]\left\{\left\{K_2\right\} + \left[\bar{\lambda}_{x\Omega_2}\right]' \left\{\delta x_2\right\} + \left\{\dot{\bar{x}}_2\right\}\Delta T_2\right\} + \left\{\Delta x_3\right\}$$

$$= \left[P_2\right]\left\{K_2\right\} + \left[P_2\right]\left[\bar{\lambda}_{x\Omega_2}\right]'\left[P_1\right]\left\{K_1\right\}$$

$$+ \left[P_2\right]\left\{\dot{\bar{x}}_2\right\}\Delta T_2 + \left[P_2\right]\left[\bar{\lambda}_{x\Omega_2}\right]'\left[P_1\right]\left\{\dot{\bar{x}}_1\right\}\Delta T_1$$

$$+ \left[P_2\right]\left[\bar{\lambda}_{x\Omega_2}\right]'\left[P_1\right]\left[\bar{\lambda}_{x\Omega_1}\right]'\left\{\Delta x_1\right\} + \left[P_2\right]\left[\bar{\lambda}_{x\Omega_2}\right]'\left\{\Delta x_2\right\} + \left\{\Delta x_3\right\} \tag{77}$$

and the next

$$\left\{\delta x_4\right\} = \left[P_3\right]\left\{\left\{K_3\right\} + \left[\bar{\lambda}_{x\Omega_3}\right]' \left\{\delta x_3\right\} + \left\{\dot{\bar{x}}_3\right\}\Delta T_3\right\} + \left\{\Delta x_4\right\}$$

$$= Q_3 Q_2 P_1 K_1 + Q_3 P_2 K_2 + P_3 K_3$$

$$+ Q_3 Q_2 P_1 \dot{\bar{x}}_1 \Delta T_1 + Q_3 P_2 \dot{\bar{x}}_2 \Delta T_2 + P_3 \dot{\bar{x}}_3 \Delta T_3$$

$$+ Q_3 Q_2 Q_1 \Delta x_1 + Q_3 Q_2 \Delta x_2 + Q_3 \Delta x_3 + \Delta x_4 \tag{78}$$

where

$$\left[Q_s\right] = \left[P_s\right]\left[\bar{\lambda}_{x\Omega_s}\right]' \tag{79}$$

In general, the total state variable perturbation at the commencement of the $(s+1)^{th}$ stage is

$$\left\{\delta x_{s+1}\right\} = \left(Q_s Q_{s-1}\cdots\cdots Q_2 P_1 K_1 + Q_s Q_{s-1}\cdots\cdots Q_3 P_2 K_2 + \text{---}\right.$$

$$\text{----} + Q_s P_{s-1} K_{s-1} + P_s K_s\Big)$$

$$+ \left(Q_s Q_{s-1}\cdots\cdots Q_2 P_1 \dot{\bar{x}}_1 \Delta T_1 + Q_s Q_{s-1}\cdots\cdots Q_3 P_2 \dot{\bar{x}}_2 \Delta T_2 + \text{---}\right.$$

$$\text{---} + Q_s P_{s-1} \dot{\bar{x}}_{s-1} \Delta T_{s-1} + P_s \dot{\bar{x}}_s \Delta T_c\Big)$$

$$+ \left(Q_s Q_{s-1}\cdots\cdots Q_1 \Delta x_1 + Q_s Q_{s-1}\cdots\cdots Q_2 \Delta x_2 + \text{----}\right.$$

$$\text{---} + Q_s Q_{s-1} \Delta x_{s-1} + Q_s \Delta x_s + \Delta x_{s+1}\Big) \tag{80}$$

24

At the commencement of the last (Nth) stage

$$\left\{\delta x_N\right\} = \sum_{s=1}^{N-1} \left[A_s^N\right]\left\{K_s\right\} + \sum_{s=1}^{N-1} \left\{B_s^N\right\}\Delta T_s + \sum_{s=1}^{N} \left[C_s^N\right]\left\{\Delta x_s\right\} \tag{81}$$

where

$$\left[A_s^N\right] = Q_{N-1}Q_{N-2} \cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot Q_{s+1}P_s , \quad S < N-1$$

$$= P_s \qquad\qquad\qquad , \quad S = N-1 \tag{82}$$

$$\left\{B_s^N\right\} = Q_{N-1}Q_{N-2} \cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot Q_{s+1}P_s \dot{\overline{x}}_s , \quad S < N-1$$

$$= P_s \dot{\overline{x}}_s \qquad\qquad\qquad , \quad S = N-1 \tag{83}$$

$$\left[C_s^N\right] = Q_{N-1}Q_{N-2} \cdot\cdot\cdot\cdot\cdot\cdot\cdot Q_s \qquad , \quad S < N$$

$$= I \qquad\qquad\qquad , \quad S = N \tag{84}$$

Knowing the initial perturbations to the last stage, Equation (25) can be applied in stage time to find changes in the pay-off function ϕ, and the constraints.

$$\delta\phi = \int_0^{T_N} \left\lfloor\lambda_\phi\Omega_N\right\rfloor \left[G\right]\left\{\delta\alpha\right\} d\tau_N + \left\lfloor\lambda_\phi\Omega_N(0)\right\rfloor\left\{\delta x_N\right\} \tag{85}$$

$$\left\{\delta\psi\right\} = \int_0^{T_N} \left[\lambda_\psi\Omega_N\right]\dot{}\left[G\right]\left\{\delta\alpha\right\} d\tau_N + \left[\lambda_\psi\Omega_N(0)\right]\dot{}\left\{\delta x_N\right\} \tag{86}$$

and on substituting for δx_N

$$\delta\phi = K_{N_\phi} + \left\lfloor\overline{\lambda}_\phi\Omega_N\right\rfloor\left\{\sum_{s=1}^{N-1} \left[A_s^N\right]\left\{K_s\right\} + \sum_{s=1}^{N-1} \left\{B_s^N\right\}\Delta T_s + \sum_{s=1}^{N} \left[C_s^N\right]\left\{\Delta x_s\right\}\right\} \tag{87}$$

$$\left\{\delta\psi\right\} = \left\{K_{N_\psi}\right\} + \left[\overline{\lambda}_\psi\Omega_N\right]\dot{}\left\{\sum_{s=1}^{N-1} \left[A_s^N\right]\left\{K_s\right\} + \sum_{s=1}^{N-1} \left\{B_s^N\right\}\Delta T_s + \sum_{s=1}^{N} \left[C_s^N\right]\left\{\Delta x_s\right\}\right\} \tag{88}$$

where

$$K_{N_\phi} = \int_0^{T_N} \lfloor \lambda \phi \Omega_N \rfloor [G] \{\delta\alpha\} \, d\tau_N \tag{89}$$

$$\{K_{N_\psi}\} = \int_0^{T_N} [\lambda \psi \Omega_N]' [G] \{\delta\alpha\} \, d\tau_N \tag{90}$$

It is convenient to combine the integrals, K_s, through each stage into a single set of integrals throughout the complete trajectory. To accomplish this, define

$$\lfloor \Lambda_{\phi\Omega} \rfloor = \lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor [A_s^N] [\lambda_{x\Omega_s}]' ; \quad s < N$$

$$= \lfloor \lambda_{\phi\Omega_N} \rfloor ; \qquad s = N \tag{91}$$

$$[\Lambda_{\psi\Omega}]' = [\bar{\lambda}_{\psi\Omega_N}]' [A_s^N] [\lambda_{x\Omega_s}]' ; \quad s < N \tag{92}$$

$$= [\lambda_{\psi\Omega_N}]' \qquad s = N$$

Expanding Equation (87)

$$\delta\phi = K_{N_\phi} + \left(\lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor \sum_{s=1}^{N-1} [A_s^N] \{K_s\} \right) + \left(\lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor \left\{ \sum_{s=1}^{N-1} \{B_s^N\} \Delta T_s + \sum_{s=1}^{N} [C_s^N] \{\Delta x_s\} \right\} \right)$$

$$= \int_0^{T_N} \lfloor \lambda_{\phi\Omega_N} \rfloor [G] \{\delta\alpha\} \, d\tau_N + \sum_{s=1}^{N-1} \lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor [A_s^N] \left\{ \int_0^{T_s} [\lambda_{x\Omega_s}]' [G] \{\delta\alpha\} \, d\tau_s \right\}$$

$$+ \lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor \left\{ \sum_{s=1}^{N-1} \{B_s^N\} \Delta T_s + \sum_{s=1}^{N} [C_s^N] \{\Delta x_s\} \right\} \tag{93}$$

$$\therefore \delta\phi = \sum_{s=1}^{N} \int_0^{T_s} \lfloor \Lambda_{\phi\Omega} \rfloor [G] \{\delta\alpha\} \, d\tau_s + \lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor \left\{ \sum_{s=1}^{N-1} \{B_s^N\} \Delta T_s + \sum_{s=1}^{N} [C_s^N] \{\Delta x_s\} \right\}$$

$$\tag{94}$$

26

The first term, which is the summation of a set of integrals throughout each stage on the unperturbed trajectory, can be combined into one integral by reverting to the original independent variable t. That is

$$\delta\phi = \int_{t_o}^{T} \lfloor \Lambda_{\phi\Omega} \rfloor \lfloor G \rfloor \{\delta\alpha\} \, dt + \lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor \left\{ \sum_{s=1}^{N-1} \{B_s^N\} \Delta T_s + \sum_{s=1}^{N} \lfloor C_s^N \rfloor \{\Delta x_s\} \right\} \quad (95)$$

Similarly,

$$\{\delta\psi\} = \int_{t_o}^{T} \lfloor \Lambda_{\psi\Omega} \rfloor' \lfloor G \rfloor \{\delta\alpha\} \, dt + \lfloor \bar{\lambda}_{\psi\Omega_N} \rfloor' \left\{ \sum_{s=1}^{N-1} \{B_s^N\} \,_s + \sum_{s=1}^{N} \lfloor C_s^N \rfloor \{\Delta x_s\} \right\} \quad (96)$$

Equations (95) and (96) give the total change in payoff and constraint functions when the trajectory simultaneously undergoes perturbations in the control variable histories, stage point positions and initial state variable values in each stage. The sensitivity of payoff and constraint functions to these variations is immediately apparent. For pulse variations in the control variables at time t = t', the individual sensitivities of the payoff function are the elements of the row matrix

$$\lfloor \phi_\alpha(t') \rfloor = \lfloor \Lambda_{\phi\Omega}(t') \rfloor \lfloor G(\,\cdot\,) \rfloor \quad (97)$$

The individual sensitivities of the constraint functions to control variable pulse perturbations are similarly the elements of the rectangular matrix

$$\lfloor \psi_\alpha(t') \rfloor = \lfloor \Lambda_{\psi\Omega} \rfloor' \lfloor G \rfloor \quad (98)$$

The sensitivity of the payoff function with respect to stage point variations follows directly from the second term of Equation (95). If the sth stage alone is perturbed by ΔT_s,

$$\phi_{\Delta T_s} = \lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor \{B_s^N\} \quad (99)$$

Similarly, the constraint sensitivities with respect to stage point perturbations are obtained from Equation (96)

$$\{\psi_{\Delta T_s}\} = \lfloor \bar{\lambda}_{\psi\Omega_N} \rfloor' \{B_s^N\} \quad (100)$$

27

Finally, the sensitivities to initial state variable value perturbations in each stage can be obtained from the last terms in Equations (95) and (96). The payoff function sensitivities for the sth stage are the elements of

$$\lfloor \phi_{\Delta x_s} \rfloor = \lfloor \lambda_{\phi \Omega_N} \rfloor \left[c_s^N \right] \qquad (101)$$

and the constraint function sensitivities for the sth stage are the elements of the rectangular matrix

$$\left[\psi_{\Delta x_s} \right] = \left[\lambda_{\psi \Omega_N} \right]' \left[c_s^N \right] \qquad (102)$$

In general two types of stage point must be considered: those whose perturbation is prescribed and those which are free to optimize. Hence,

$$\sum_{s=1}^{N-1} \left\{ B_s^N \right\} \Delta T_s = \sum_{s'=1}^{S'} \left\{ B_{s'}^N \right\} \Delta T_{s'} + \sum_{\bar{s}=1}^{\bar{S}} \left\{ B_{\bar{s}}^N \right\} \Delta T_{\bar{s}} \qquad (103)$$

where the $\Delta T_{s'}$ are prescribed and the $\Delta T_{\bar{s}}$ are to be optimized. Substituting Equation (103) into Equation (95)

$$\delta \phi = \int_{t_0}^{T} \lfloor \Lambda_{\phi \Omega} \rfloor \left[G \right] \left\{ \delta \alpha \right\} dt + \lfloor \bar{\lambda}_{\phi \Omega_N} \rfloor \left\{ \sum_{\bar{s}=}^{\bar{S}} \left\{ B_{\bar{s}}^N \right\} \Delta T_{\bar{s}} \right\}$$

$$+ \lfloor \bar{\lambda}_{\phi \Omega_N} \rfloor \left\{ \sum_{s'=1}^{S'} \left\{ B_{s'}^N \right\} \Delta T_{s'} + \sum_{s=1}^{N} \left[c_s^N \right] \left\{ \Delta x_s \right\} \right\} \qquad (104)$$

Substituting Equation (103) into Equation (96)

$$\left\{ \delta \Gamma \right\} = \int_{t_0}^{T} \left[\Lambda_{\psi \Omega} \right]' \left[G \right] \left\{ \delta \alpha \right\} dt + \left[\bar{\lambda}_{\psi \Omega_N} \right]' \left\{ \sum_{s=1}^{\bar{S}} \left\{ B_s^N \right\} \Delta T_{\bar{s}} \right\} \qquad (105)$$

where all the quantities specified directly are grouped together in the term,

$$\left\{ \delta \Gamma \right\} = \left\{ \delta \psi \right\} - \left[\bar{\lambda}_{\psi \Omega_N} \right]' \left\{ \sum_{s'=1}^{S'} \left\{ B_{s'}^N \right\} \Delta T_{s'} + \sum_{s=1}^{N} \left[c_s^N \right] \left\{ \Delta x_s \right\} \right\} \qquad (106)$$

With Equations (104) and (105), expressions for the change in the payoff and constraint functions resulting from a general perturbation of control variable histories and stage point perturbations are obtained. The optimal perturbation has yet to be found. This task is considered in the next section.

Derivation of Variational Equations

The preceding section derived expressions for payoff and constraint function changes when a combined perturbation was introduced in the control variable histories, stage point positions, and initial state variable values in each stage. The perturbation which optimizes payoff function change while at the same time producing specified changes in the constraints $\{\delta\psi\}$ and the initial state variable values in each stage $\{\Delta x_s\}$ will now be obtained. In order to obtain a meaningful solution, constraints are placed on perturbation magnitudes. Control variable perturbations are limited in the manner of the preceding section by introducing the constraint

$$DP^2 = \int_{t_o}^{T} \left(\lfloor \delta\alpha \rfloor \left[W \right] \{ \delta\alpha \} \right) dt$$

(5)

A similar constraint (DT^2) is introduced to limit total stage point perturbation.

$$DT^2 = \left(\sum_{\overline{s}=1}^{\overline{S}} V_{\overline{s}} \ \Delta T_{\overline{s}}^2 \right)$$

(107)

Here the $V_{\overline{s}}$ are a set of weighting functions used to modulate optimal stage point perturbations.

Proceeding as in the previous section, pages 10 to 13 an augmented function is constructed in the manner of Lagrange and minimized (maximized). In the present case, the augmented function is

$$U = \int_{t_o}^{T} \lfloor \Lambda_{\phi\Omega} \rfloor [G]\{\delta\alpha\} \, dt + \lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor \left\{ \sum_{s=1}^{\bar{S}} \{B_{\bar{s}}^N\} \, \Delta T_{\bar{s}} \right\}$$

$$+ \lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor \left\{ \sum_{s'=1}^{S'} \{B_{s'}^N\} \Delta T_{s'} + \sum_{s=1}^{N} [C_{s'}^N] \{\Delta x_s\} \right\}$$

$$+ \lfloor \vartheta \rfloor \left\{ \int_{t_o}^{T} [\Lambda_{\psi\Omega}]'[G]\{\delta\alpha\} dt + [\bar{\lambda}_{\psi\Omega_N}]' \left\{ \sum_{s=1}^{\bar{S}} \{B_{\bar{s}}^N\} \Delta T_{\bar{s}} \right\} \right.$$

$$+ [\bar{\lambda}_{\psi\Omega_N}]' \left\{ \sum_{s'=1}^{S'} \{B_{s'}^N\} \Delta T_{s'} + \sum_{s=1}^{N} [C_s^N]\{\Delta x_s\} \right\}$$

$$+ \mu \int_{t_o}^{T} \lfloor \delta\alpha \rfloor [W]\{\delta\alpha\} \, dt + \omega \left(\sum_{\bar{s}=1}^{\bar{S}} V_{\bar{s}} \, \Delta T_{\bar{s}}^2 \right)$$

$$(108)$$

where ω is a Lagrangean Multiplier introduced for the stage point perturbation constraint.

First, differentiate with respect to each stage point perturbation being optimized. This results in $\bar{S}$ equations,

$$\frac{\partial \cdots}{\partial (\Delta T_{\bar{s}})} = \lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor \{B_{\bar{s}}^N\} + 2\omega V_{\bar{s}} \Delta T_{\bar{s}} + \lfloor \vartheta \rfloor [\bar{\lambda}_{\psi\Omega_N}]'\{B_{\bar{s}}^N\}$$

$$(109)$$

These expressions must disappear for U to be an extremal. Solving for the $\Delta T_{\bar{s}}$

$$\Delta T_{\bar{s}} = -\frac{1}{2\omega V_{\bar{s}}} \left[\lfloor \bar{\lambda}_{\phi\Omega_N} \rfloor + \lfloor \vartheta \rfloor [\bar{\lambda}_{\psi\Omega_N}]' \right] \{B_{\bar{s}}^N\}$$

$$(110)$$

Squaring both sides of Equation (110), multiplying throughout by $V_{\bar{s}}$, summing the $\bar{S}$ equations, and using Equation (107)

30

$$\sum_{\overline{s}=1}^{\overline{S}} V_{\overline{s}} \, \Delta T_{\overline{s}}^2 = DT^2 = \frac{1}{4\omega^2} \sum_{\overline{s}=1}^{\overline{S}} V_{\overline{s}}^{-1} \left(\left\lfloor \left\lfloor \overline{\lambda}_{\phi\Omega_N} \right\rfloor + \lfloor \vartheta \rfloor \left[\overline{\lambda}_{\psi\Omega_N} \right] \right]' \left\{ B_{\overline{s}}^N \right\} \right)^2$$

$$= \frac{1}{4\omega^2} \sum_{\overline{s}=1}^{\overline{S}} \left\lfloor \left\lfloor \overline{\lambda}_{\phi\Omega_N} \right\rfloor + \lfloor \vartheta \rfloor \left[\overline{\lambda}_{\psi\Omega_N} \right] \right]' \frac{\left\{ B_{\overline{s}}^N \right\} \left\lfloor B_{\overline{s}}^N \right\rfloor}{V_{\overline{s}}} \left\{ \left\{ \overline{\lambda}_{\phi\Omega_N} \right\} + \left[\overline{\lambda}_{\psi\Omega_N} \right] \{ \vartheta \} \right\}$$

(111)

$\overline{}$ .ranging Equation (111) and taking the summation into the matrix product

$$\omega^2 = \frac{1}{4\,DT^2} \left(L_{\phi\phi} + 2 \lfloor L_{\psi\phi} \rfloor \{ \vartheta \} + \lfloor \vartheta \rfloor \left[L_{\psi\psi} \right] \{ \vartheta \} \right)$$

or

$$4\omega^2 DT^2 - L_{\phi\phi} - 2 \lfloor L_{\psi\phi} \rfloor \{ \vartheta \} - \lfloor \vartheta \rfloor \left[L_{\psi\psi} \right] \{ \vartheta \} = 0$$

(112)

where

$$L_{\phi\phi} = \lfloor \overline{\lambda}_{\phi\Omega_N} \rfloor [D] \{ \overline{\lambda}_{\phi\Omega_N} \}$$

(113)

$$\lfloor L_{\psi\phi} \rfloor = \lfloor \overline{\lambda}_{\phi\Omega_N} \rfloor [D] \left[\overline{\lambda}_{\psi\Omega_N} \right]$$

(114)

$$\left[L_{\psi\psi} \right] = \left[\overline{\lambda}_{\psi\Omega_N} \right]' [D] \left[\overline{\lambda}_{\psi\Omega_N} \right]$$

(115)

and

(116)

$$[D] = \sum_{\overline{s}=1}^{\overline{S}} V_{\overline{s}}^{-1} \left\{ B_{\overline{s}} \right\} \lfloor B_{\overline{s}} \rfloor$$

Second, take a variation of $\delta\alpha$ to Equation (108) in a similar manner to the preceding section

$$\delta U_{(\alpha)} = \int_{t_o}^{T} \lfloor \Lambda_{\phi\Omega} \rfloor [G] \{ \delta^2\alpha \} dt + \lfloor \vartheta \rfloor \int_{t_o}^{T} \left[\Lambda_{\psi\Omega} \right]' [G] \{ \delta^2\alpha \} dt + 2\mu \int_{t_o}^{T} \lfloor \delta\alpha \rfloor [W] \{ \delta^2\alpha \} dt = 0$$

(117)

$$\therefore \lfloor \Lambda_{\phi\Omega} \rfloor [G] + \lfloor \vartheta \rfloor \left[\Lambda_{\psi\Omega} \right]' [G] + 2\mu \lfloor \delta\alpha \rfloor [W] = 0$$

(118)

$$\therefore \{ \delta\alpha \} = -\frac{1}{2\mu} [W]^{-1} [G]' \left\{ \{ \Lambda_{\phi\Omega} \} + \left[\Lambda_{\psi\Omega} \right] \{ \vartheta \} \right\}$$

(119)

31

Substituting into Equation (105)

$$\left\{ \delta \Gamma \right\} = -\frac{1}{2\mu} \left\{ \left\{ J_{\psi\phi} \right\} + \left[J_{\psi\psi} \right] \left\{ \vartheta \right\} \right\} + \left[\bar{\lambda}_{\psi\Omega_N} \right]' \left\{ \sum_{s=1}^{\bar{s}} \left\{ B_{\bar{s}}^N \right\} \Delta T_{\bar{s}} \right\} \qquad (120)$$

where we define

$$J_{\phi\phi} = \int_{t_o}^{T} \lfloor \Lambda_{\phi\Omega} \rfloor \left[G \right] \left[W \right]^{-1} \left[G \right]' \left\{ \Lambda_{\phi\Omega} \right\} dt \qquad (121)$$

$$\left\{ J_{\psi\phi} \right\} = \int_{t_o}^{T} \left[\Lambda_{\psi\Omega} \right]' \left[G \right] \left[W \right]^{-1} \left[G \right]' \left[\Lambda_{\phi\Omega} \right] dt \qquad (122)$$

$$\left[J_{\psi\psi} \right] = \int_{t_o}^{T} \left[\Lambda_{\psi\Omega} \right] \left[G \right] \left[W \right]^{-1} \left[G \right]' \left[\Lambda_{\psi\Omega} \right] dt \qquad (123)$$

Transposing Equation (110), substituting into Equation (120), and using Equation (116)

$$\left\{ \delta \Gamma \right\} = -\frac{1}{2\mu} \left\{ \left\{ J_{\psi\phi} \right\} + \left[J_{\psi\psi} \right] \left\{ \vartheta \right\} \right\} - \frac{1}{2\omega} \left[\bar{\lambda}_{\psi\Omega_N} \right]' \left[D \right] \left\{ \left[\bar{\lambda}_{\psi\Omega_N} \right] \left\{ \vartheta \right\} + \left\{ \bar{\lambda}_{\phi\Omega_N} \right\} \right\} \qquad (124)$$

Substituting Equations (114) and (115) into Equation (124), multiplying throughout by $2\omega\mu$ and collecting terms on the left

$$2\mu\omega \left\{ \delta \Gamma \right\} + \omega \left\{ \left\{ J_{\psi\phi} \right\} + \left[J_{\psi\psi} \right] \left\{ \vartheta \right\} \right\} + \mu \left\{ \left[L_{\psi\psi} \right] \left\{ \vartheta \right\} + \left\{ L_{\psi\phi} \right\} \right\} = 0 \qquad (125)$$

Third, substitute Equation (119) into Equation (5)

$$DP^2 = \frac{1}{4\mu^2} \int_{t_o}^{T} \left[\lfloor \Lambda_{\phi\Omega} \rfloor + \lfloor \vartheta \rfloor \left[\Lambda_{\psi\Omega} \right] \right] \left[G \right] \left[W \right]^{-1} \left[G \right]' \left\{ \left\{ \Lambda_{\phi\Omega} \right\} + \left[\Lambda_{\psi\Omega} \right] \left\{ \vartheta \right\} \right\} dt \qquad (126)$$

Using Equations (121), (122), and (123), multiplying by $4\mu^2$ and collecting on the left

$$4DP^2\mu^2 - J_{\phi\phi} - 2 \lfloor J_{\psi\phi} \rfloor \left\{ \vartheta \right\} - \lfloor \vartheta \rfloor \left[J_{\psi\psi} \right] \left\{ \vartheta \right\} = 0 \qquad (127)$$

32

Equations (112), (125), and (127) are the variational equations which must be solved to obtain the Lagrangean multipliers and, hence, the optimal perturbations. The solution to the optimal staging problem is immediately available in terms of the Lagrangian Multipliers. Equation (119) provides the optimal control variable perturbation

$$\{\delta\alpha\} = -\frac{1}{2\mu} \left[W\right]^{-1} \left[G\right]' \left\{\{\Lambda_{\phi\Omega}\} + \left[\Lambda_{\psi\Omega}\right]\{\vartheta\}\right\} \tag{119}$$

The optimum stage point perturbations are given by Equations (110). Rearranging and combining Equations (110) into one matrix equation

$$\{\Delta T_s\} = -\frac{1}{2\omega} \left|V_{\overline{s}}\right|^{-1} \left[B_{\overline{s}}^N\right]' \left\{\{\overline{\lambda}_{\phi\Omega_N}\} + \left[\overline{\lambda}_{\psi\Omega_N}\right]\{\vartheta\}\right\} \tag{128}$$

where the columns of $\left[B_s^N\right]$ are the $\{B_s^N\}$. The similarity of form between the expressions for $\{\delta\alpha\}$ and $\{\Delta T_s\}$ is immediately apparent.

The optimal payoff function change is found by substituting Equation (119) and (128) into Equation (104)

$$\delta\phi = -\frac{1}{2\mu} \left\{J_{\phi\phi} + \left[J_{\psi\phi}\right]\{\vartheta\}\right\} - \frac{1}{2\omega}\left\{L_{\phi\phi} + \left[L_{\psi\phi}\right]\{\vartheta\}\right\}$$

$$+ \left[\overline{\lambda}_{\phi\Omega_N}\right]\left\{\sum_{s'=1}^{S'}\{B_{s'}^N\} \Delta T_{s'} + \sum_{s=1}^{N}\left[C_s^N\right]\{\Delta x_s\}\right\} \tag{129}$$

where the J terms are defined by Equations (121) to (123) and the L terms by Equations (113) to (116).

The equations presented here in this section constitute the general solution to the optimal staging problem. A closed form solution to these equations cannot be obtained. For straightforward elimination by Sylvester's Method leads to a high order polynomial equation indicating that a closed form solution is unattainable. Accordingly, in the next section a further assumption permitting attainment of a closed form solution is made.

Solution Using Combined Step-Size Parameter

The fundamental reason a closed form solution to the general optimal staging equations could not be obtained in the preceding section, pages 29 to 33 , was the introduction of separate Lagrangean Multipliers for the control variable

and stage point perturbations. If these perturbations are combined and a single step-size parameter defined by the expression

$$DC^2 = \int_{t_0}^{T} \lfloor \delta\alpha(t) \rfloor \left[W(t) \right] \left\{ \delta\alpha(t) \right\} dt + \sum_{\bar{s}=1}^{\bar{s}} V_{\bar{s}}^- \Delta T_{\bar{s}}^2 \tag{130}$$

this difficul.y can be avoided. It should be noted that the weighting functions, $V_{\bar{s}}^-$, must be dimensional quantities to insure compatibility between both portions of Equation (130).

Returning to the augmented function of Equation (108) it follows that in the last line ω can now be replaced by μ. This substitution can also be made in Equations (109) to (112). Equations (112) and (127) can then be combined to obtain

$$4\mu^2 \cdot DC^2 - (J_{\phi\phi} + I_{\phi\phi}) - \epsilon \left[\lfloor J_{\psi\phi} \rfloor + \lfloor L_{\psi\phi} \rfloor \right] \left\{ \vartheta \right\}$$

$$- \lfloor \vartheta \rfloor \left[\left[J_{\psi\psi} \right] + \left[L_{\psi\psi} \right] \right] \left\{ \vartheta \right\} \tag{131}$$

Substituting μ for ω into the equations leading to Equation (125) results in the expression

$$2\mu \left\{ \delta\Gamma \right\} + \left\{ \left\{ J_{\psi\phi} \right\} + \left\{ L_{\psi\phi} \right\} \right\} + \left[\left[J_{\psi\psi} \right] + \left[L_{\psi\psi} \right] \right] \left\{ \vartheta \right\} = 0 \tag{132}$$

It follows by comparison with Equations (39) and (35a) that μ and $\{\nu\}$ are given by Equations (40) and (42) provided the $\{d\beta\}$ are replaced by $\{\delta\Gamma\}$ and the I by the appropriate $(J + L)$. For example,

$$I_{\phi\phi} \longrightarrow J_{\phi\phi} + L_{\phi\phi}$$

The optimal control variable perturbation is then given by Equation (43) with the I replaced by the $(J + L)$ and the λ by the appropriate Λ. For example,

$$\left\{ \lambda_{\phi\Omega} \right\} \longrightarrow \left\{ \Lambda_{\phi\Omega} \right\}$$

Substituting μ and $\{\nu\}$ into Equation (128), the optimal stage point perturbations are obtained in the form

$$\left\{ \Delta T_{\bar{s}} \right\} = \mp \left\lceil V_s \right\rfloor^{-1} \left[B_s^N \right]' \left\{ \left\{ \bar{\lambda}_{\phi\Omega N} \right\} - \left[\bar{\lambda}_{\psi\Omega N} \right] \left[J_{\psi\psi} + L_{\psi\psi} \right]^{-1} \left\{ J_{\psi\phi} + L_{\psi\phi} \right\} \right\}$$

$$\times \sqrt{ \frac{ DC^2 - \lfloor \delta\Gamma \rfloor \left[J_{\psi\psi} + L_{\psi\psi} \right]^{-1} \left\{ \delta\Gamma \right\} }{ (J_{\phi\phi} + L_{\phi\phi}) - \lfloor J_{\psi\phi} + L_{\psi\phi} \rfloor \left[J_{\psi\psi} + L_{\psi\psi} \right]^{-1} \left\{ J_{\psi\phi} + L_{\psi\phi} \right\} } }$$

$$+ \left\lceil V_s \right\rfloor^{-1} \left[B_s^N \right]' \left[\bar{\lambda}_{\psi\Omega N} \right] \left[J_{\psi\psi} + L_{\psi\psi} \right]^{-1} \left\{ \delta\Gamma \right\} \right\} \tag{133}$$

With these equations a general solution to the optimal staging problem has been obtained. The main difficulty in applying the results will undoubtedly lie in determining suitable weighting matrices.

POINT MASS TRAJECTORY EQUATIONS

Basic State Variables

Preceding portions of this report derived a successive approximation scheme for computation of optimum trajectories generated by a set of first order differential equations. The analysis is quite general and holds for trajectories generated by any set of first order differential equations. The object of the present section is to specialize the analysis to point mass vehicle trajectory problems. This will be accomplished when a suitable set of state variables, together with their derivatives, the control variables, and the forces associated with the control variables has been specified. First, a suitable coordinate system is selected, and Newton's Laws in this system are used to define the vehicle's motion.

Several suitable coordinate systems are available for point mass trajectory computations. The basic set of coordinates used in the present analysis is a rectangular set rotating with the earth, (X_e, Y_e, Z_e). This coordinate system is illustrated in Figure 4.
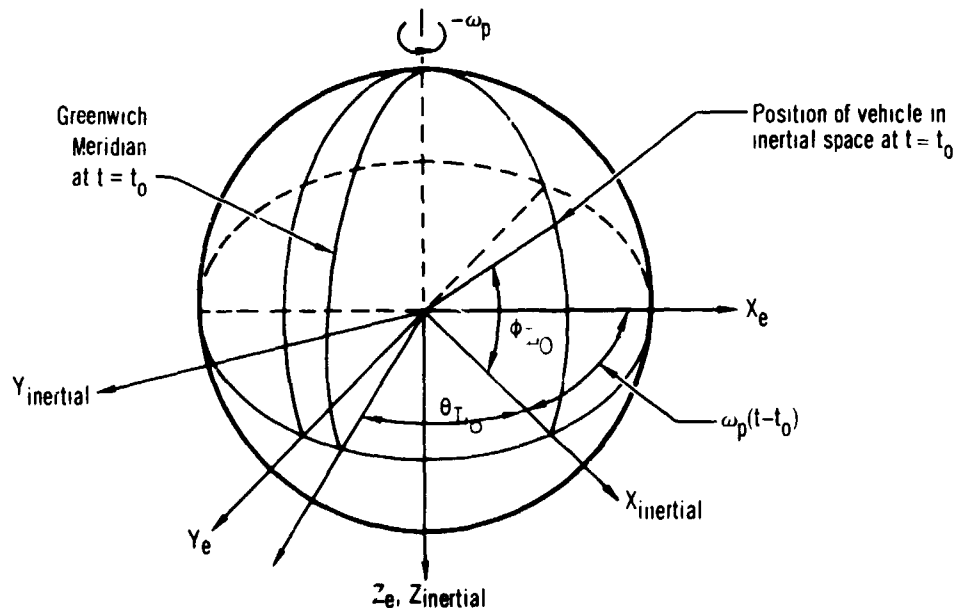
Figure 4.— Basic Coordinate System

The X_e and Y_e axes lie in the equatorial plane, the positive X_e axis being initially chosen as the intersection of this plane with the vehicle longitudinal plane at $t = t_o$. Y_e is 90° to the west of X_e, and Z_e is positive through the South Pole. Denoting the radius vector from the center of the earth to the point mass vehicle by $\mathbf{R}$, its magnitude is given by,

$$|\mathbf{R}| = \sqrt{X_e^2 + Y_e^2 + Z_e^2}$$

(134)

The angle between $\mathbf{R}$ and the North Pole is given by

$$\phi' = 90 - \phi_L$$

(135)

where ϕ_L is the latitude of the vehicle. As a result of the earth's rotation, an observer in the (X_e, Y_e, Z_e) system would detect an apparent motion of the point mass if it were at rest in inertial space. In time Δt the apparent displacement of such a vehicle would be

$$\delta R_{apparent} = R \sin \phi' \cdot \omega_p \Delta t$$

(136)

to the west. In vector notation

$$\delta \mathbf{R}_{apparent} = \mathbf{R} \times \omega_p \Delta t = -\omega_p \times \mathbf{R} \Delta t$$

(137)

This apparent displacement is independent of the vehicle's motion and exists whether or not the vehicle is at rest in inertial space. In general, then, to an observer in the rotating coordinate system,

$$(\delta \mathbf{R})_e = (\delta \mathbf{R})_{inertial} + (\delta \mathbf{R})_{apparent}$$

(138)

$$\therefore \quad (\delta \mathbf{R})_{inertial} = (\delta \mathbf{R})_e + \omega_p \times \mathbf{R} \Delta t$$

(139)

Dividing Equation (139) by ΔT and taking the limit, it follows that

$$\left(\frac{d\mathbf{R}}{dt}\right)_{inertial} = \left(\frac{d\mathbf{R}}{dt}\right)_e + \omega_p \times \mathbf{R}$$

(140

or

$$\mathbf{V}_{inertial} = \mathbf{V}_e + \omega_p \times \mathbf{R}$$

(141)

37

The vector $\mathbf{R}$ in Equation (140) could equally well be taken as any vector; the arguments of Equations (134) to (141) still hold. Therefore, in general, for any vector quantity the operational equality

$$\left(\frac{d}{dt}\right)_{inertial} = \left(\frac{d}{dt}\right)_e + \omega_p \times \qquad (142)$$

can be defined. Applying Equation (142) to Equation (141), the inertial acceleration is given by

$$\left(\frac{dV}{dt}\right)_{inertial} = \left(\left(\frac{d}{dt}\right)_e + \omega_p \times\right)\left(\left(\frac{dR}{dt}\right)_e + \omega_p \times R\right)$$

$$= \left(\frac{d^2R}{dt^2}\right)_e + \omega_p \times \left(\frac{dR}{dt}\right)_e + \omega_p \times \omega_p \cdot \times R \qquad (143)$$

Now Newton's Law applies in inertial space so that in the rotating system

$$\frac{F}{m} = \left(\frac{d^2R}{dt^2}\right)_e + 2\omega_p \times \left(\frac{dR}{dt}\right)_e + \omega_p \times \omega_p \times R \qquad (144)$$

Here $\mathbf{F}$ is the total force acting on the vehicle. This vector equation can be expressed in component form using the relationships

$$\mathbf{R} = X_e \cdot \mathbf{i} + Y_e \cdot \mathbf{j} + Z_e \cdot \mathbf{k} \qquad (145)$$

$$\omega_p = -\omega_p \cdot \mathbf{k} \qquad (146)$$

$$\mathbf{F} = F_{x_e} \cdot \mathbf{i} + F_{y_e} \cdot \mathbf{j} + F_{z_e} \cdot \mathbf{k} \qquad (147)$$

Here $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are unit vectors aligned along the X_e, Y_e, and Z_e axes, respectively. Equating components on either side of Equation (144)

38

$$\frac{F_{x_e}}{m} = \ddot{X}_e + 2\omega_p \dot{Y}_e - \omega_p^2 \, X_e \qquad (148)$$

$$\frac{F_{y_e}}{m} = \ddot{Y}_e - 2\omega_p \dot{X}_e - \omega_p^2 Y_e \qquad (149)$$

$$\frac{F_{z_e}}{m} = \ddot{Z}_e \qquad (150)$$

These equations are not yet in a suitable form for the steepest descent analysis to be applied, for they are not in first order form. The transformation of Equations (148) to (150) into first order form is immediately accomplished if the following quantities are defined as state variables:

$$\{x\} = \begin{Bmatrix} X_e \\ Y_e \\ Z_e \\ u_e \\ v_e \\ w_e \end{Bmatrix} \qquad (151)$$

where

$$V_e = \frac{dR_e}{dt} = u_e \cdot \mathbf{i} + v_e \cdot \mathbf{j} + w_e \cdot \mathbf{k} \qquad (152)$$

With this set of state variables the following expressions for the state variable derivatives are obtained from Equations (148) to (152)

$$\dot{X}_e = u_e \qquad (153$$

$$\dot{Y}_e = v_e \qquad (154$$

$$\dot{Z}_e = w_e \qquad (155)$$

$$\dot{u}_e = \frac{F_{x_e}}{m} - 2\omega_p v_e + \omega_p^2 \, X_e \qquad (156)$$

$$\dot{v}_e = \frac{F_{y_e}}{m} + 2\omega_p u_e + \omega_p^2 \, Y_e \qquad (157)$$

$$\dot{w}_e = \frac{F_{z_e}}{m} \qquad (158)$$

39

se equations are in the same form as Equation (1)
pr̥   the total forc˷ is a function of the state variables,
a s˯̮ ̮ control variab̤ ˳̣, and time. When the mass is vari-
able, it too must be introdu˰ed as a state variable. Any
expression for ᵗhe rate of change of mass of the form

$$\dot{m} = \dot{m}\left(x_n(t), \alpha_i(t), t\right) \tag{159}$$

may be used in the analysis. The above state variables, X_e,
Y_e, Z_e, u_e, v_e, w_e, and m will be referred to as the basic
state variables. In certain problems it becomes necessary to
specify additional state variables; these will be discussed
later in this section of the report.

Control Variables

The total force acting on the vehicle has three distinct
sourc˄s: first, aerodynamic force as a result of interaction
between the vehicle surfaces and the planetary atmosphere;
second, gravitational force as a result of vehicle and planetary
mass interaction; and finally, thrust forces introduced by the
vehicle propulsion system.

Befoꞁe aꞓ˯odynamic forces can be computed, the atmospheric
properties, ꞏvehicle velocity relative to the atmosphere, and
vehicle att   ꞏe must be specified. Atmospheric properties can
usually be specified as a function of altitude which, in turn,
is a function of the state variables X_e, Y_e, Z_e. ꞏvehicle velo-
city relative to ˡhe atmosphere is also a function of the state
variables, for u_e, v_e, and w_e are the vehicle veloc˙ty compo-
nents in a rotating system. The first and second factors
determining aerodynamic forces are, therefore, f ꞏctions of
the basic state variables.

The remaining factor entering into aerodynamic force
determination, the vehicle attitude, is clearly not a function
of the basic state variables. Fꞏ given the vehicle's position
and velocity, we are still quite ꞏ˳ee to specify its angular
orientation in space. The angles which determine vehicle
orientation may, therefore, be utilized as control variables by
which aerodynamic forces may ꞁe modulated. Any set of three
independent angles could be utilized for this purpose. Conven-
tion suggests use of the vehicle angle-of-attack and angle-of-
sideslip to orient the vehicle reference axis with respect tꞅ
the velocity vector. Angle-of-attack, (α) , is the angle
ꞃetween the velocity vector and the vehicle reference axis when
viewed in thꞁ vehicle side elevation. That is in a rectangular
coordinate system, x, y, z with x along the vehicle reference
axis, poꞅꞏtꞏve forwaꞃd, y perpendicular to the vehicle plane of

40

symmetry, positive to starboard, and z completing a right hand system, a view normal to the x-z plane is considered. If u, v, w are the components of the vehicle velocity with respect to the atmosphere in this body axis system

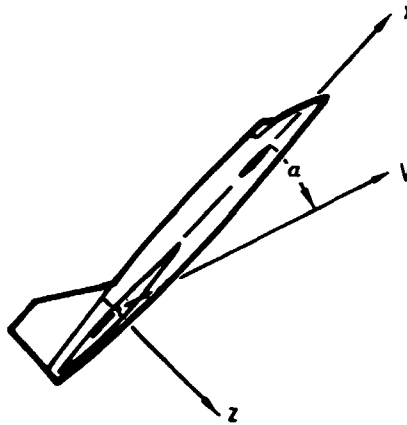$$\alpha = \tan^{-1} (\frac{w}{u})$$ (160)

Figure 5.——Angle of Attack

Sideslip angle (β) is the angle between the velocity vector and the reference axis when looking down on the vehicle planform, that is along the z axis. In this case,

$$\beta = \tan^{-1} (\frac{v}{u})$$ (161)
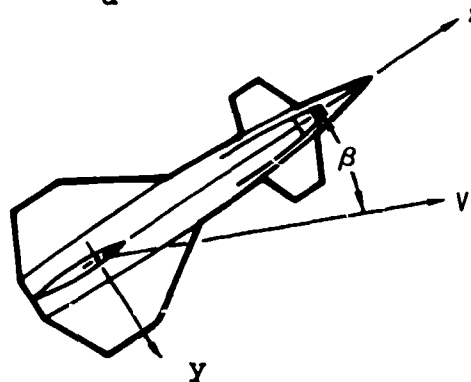
Figure 6.—— Sideslip Angle

Angle of attack and sideslip completely define the attitude of the vehicle with respect to the velocity vector. The third angle required to establish vehicle orientation in space is a rotation about the velocity vector. This last angle, bank angle

41

(B_A), will be taken as zero when the vehicle plane of symmetry is vertical and the vehicle upright. Positive bank angle will be taken as a positive rotation about the velocity vector, as in Figure 7.
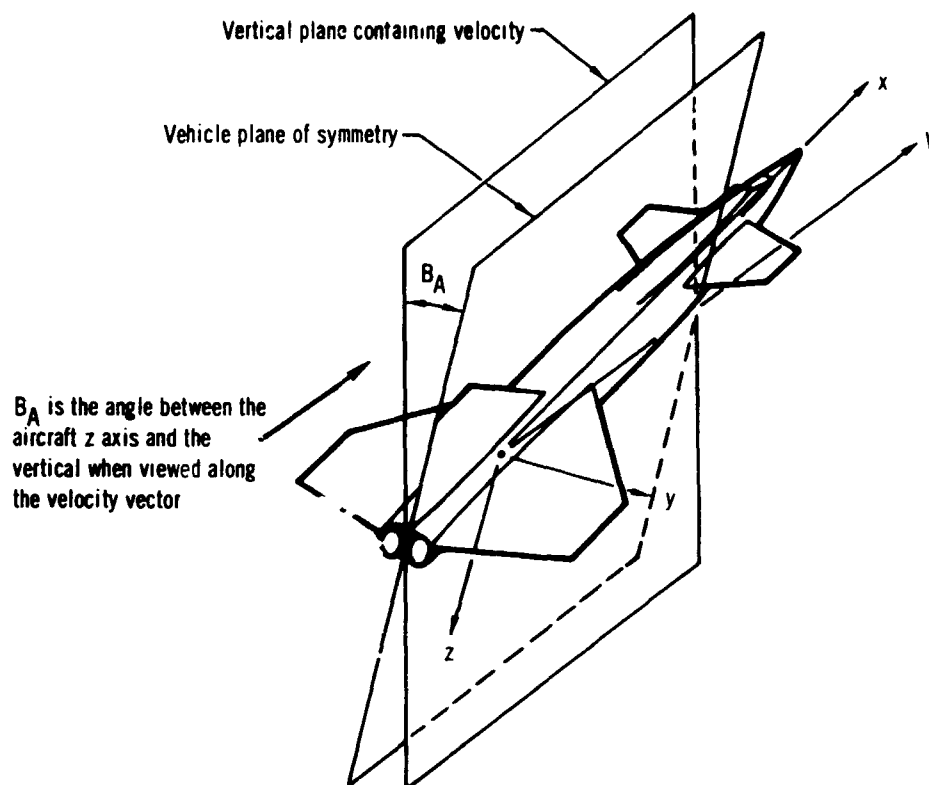

Figure 7.— Bank Angle

With the above set of angles to describe vehicle attitude, the velocity vector known and a given atmosphere, the aerodynamic forces can be completely specified.

Returning to the second source of vehicle force, gravitation, from Newton's Laws, this is merely a function of position and mass. It is, therefore, completely defined in terms of the state variables and, hence, introduces no new control variable.

The final source of vehicle force, thrust from the propulsion system, involves the atmospheric properties, either due to the atmospheric back pressure degrading the vacuum thrust or by virtue of the atmospheric fluid used in the combustion process which creates thrust. The propulsion unit efficiency may be

affected by Mach number and, hence, velocity so that thrust forces depend on the basic state variables of position and velocity in a similar manner to aerodynamic forces. If the propulsion system has a fixed orientation within the vehicle, the control variables introduced to describe aerodynamic forces suffice to describe thrust forces also. It may be, however, that the propulsion unit has a variable orientation within the vehicle. In this case, additional control variables to describe the relative position of the propulsion unit with respect to the vehicle are required. With vehicle attitude already specified by α, β and B_A, two additional angles are sufficient to orient the thrust. These may conveniently be taken as the cone angle from the reference axis, λ_T, and the inclination about the reference axis, ϕ_T. This latter angle will be measured positively about the reference axis and be zero when the thrust force is perpendicular to the port side of the vehicle plane of symmetry, as illustrated in Figure 8.
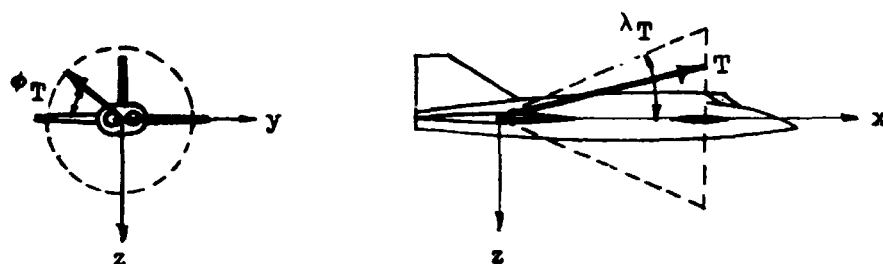

Figure 8.— Thrust Angles

One other control variable for thrust remains to be specified; this is the throttle setting, N, which serves to determine the propulsion unit power setting on variable thrust engines.

In all then, to specify the forces acting on a point mass vehicle with a single propulsion unit, six control variables, α, β, B_A, λ_T, ϕ_T, and N are required. If there is more than one independently controllable propulsion unit, additional λ_T, ϕ_T, and N must be defined.

Coordinates and Coordinate Transformations

Local geocentric-horizon coordinates.— Components of
the planet-referenced acceleration are integrated to obtain
the planet-referenced velocity components ($\dot{X}_e$, $\dot{Y}_e$, $\dot{Z}_e$). Vehicle
position in this coordinate system is determined by integration
of these velocities. Vehicle position in the planet-referenced
spherical coordinate system will now be determined. The spher-
ical coordinates are longitude, geocentric latitude, and dis-
tance from the center of the planet. Angle "C" represents the
change in vehicle longitude and may be written

$$C = \theta_{L_C} - \theta_L \qquad (162)$$

Angle C is related to the vehicle position by the expression

$$C = \mathrm{Tan}^{-1} \left(\frac{Y_e}{X_e}\right) \qquad (163)$$

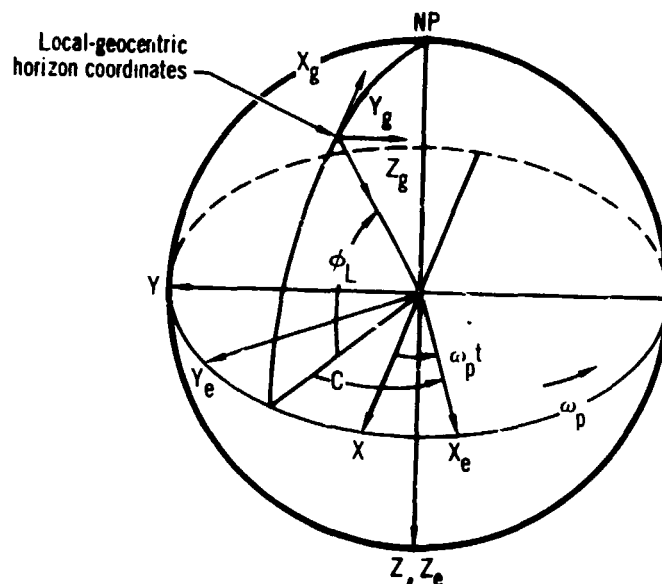The relationships are illustrated in Figure 9.

Figure 9.—Relation between Local-Geocentric, Inertial and
Earth-Referenced Coordinates for Point-Mass Problems

To describe body motion relative to the planet, a local-
geocentric-horizon coordinate system is employed. The Z_g-axis
of this system is along a radial line passing through the body
center of gravity and is positive toward the center planet.
The X_g-axis of this system is normal to the Z_g-axis and is posi-
tive northward; Y_g forms a right-handed system. Figure 9 shows
the relation of this coordinate system to the other systems
employed.

To locate the X_g-Y_g-Z_g axes with respect to the X_e-Y_e-Z_e axes, first rotate about Z_e by an angle $(180° + C)$ and then rotate about Y_g through the angle $(90° - \phi_L)$. The first rotation defines the intermediate coordinate system shown in Figure 10. The transformation is given by

$$\begin{vmatrix} \bar{1}_{X'} \\ \bar{1}_{Y_g} \\ \bar{1}_{Z_e} \end{vmatrix} = \begin{vmatrix} \text{Cos }(180° + C) & \text{Sin }(180° + C) & 0 \\ -\text{Sin }(180° + C) & \text{Cos }(180° + C) & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \bar{1}_{X_e} \\ \bar{1}_{Y_e} \\ \bar{1}_{Z_e} \end{vmatrix}$$

or

$$\begin{vmatrix} \bar{1}_{X'} \\ \bar{1}_{Y_g} \\ \bar{1}_{Z_e} \end{vmatrix} = \begin{vmatrix} -\text{Cos } C & -\text{Sin } C & 0 \\ \text{Sin } C & -\text{Cos } C & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \bar{1}_{X_e} \\ \bar{1}_{Y_e} \\ \bar{1}_{Z_e} \end{vmatrix} \tag{164a}$$

The second rotation is shown in Figure 11. The transformation matrix for the second rotation is given by

$$\begin{vmatrix} \bar{1}_{X_g} \\ \bar{1}_{Y_g} \\ \bar{1}_{Z_g} \end{vmatrix} = \begin{vmatrix} \text{Cos }(90° - \phi_L) & 0 & -\text{Sin }(90° - \phi_L) \\ 0 & 1 & 0 \\ \text{Sin }(90° - \phi_L) & 0 & \text{Cos }(90° - \phi_L) \end{vmatrix} \begin{vmatrix} \bar{1}_{X'} \\ \bar{1}_{Y_g} \\ \bar{1}_{Z_e} \end{vmatrix}$$

or

$$\begin{vmatrix} \bar{1}_{X_g} \\ \bar{1}_{Y_g} \\ \bar{1}_{Z_g} \end{vmatrix} = \begin{vmatrix} \text{Sin } \phi_L & 0 & -\text{Cos } \phi_L \\ 0 & 1 & 0 \\ \text{Cos } \phi_L & 0 & \text{Sin } \phi_L \end{vmatrix} \begin{vmatrix} \bar{1}_{X'} \\ {}_{\lrcorner}Y_g \\ \bar{1}_{Z_e} \end{vmatrix} \tag{164b}$$

Figure 10- Coordinate System Transformation ... red to Local-Geocentric Coordinates
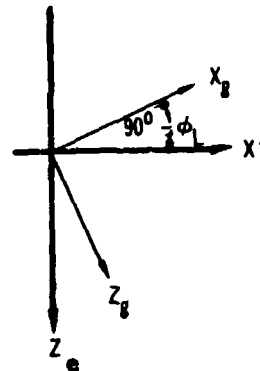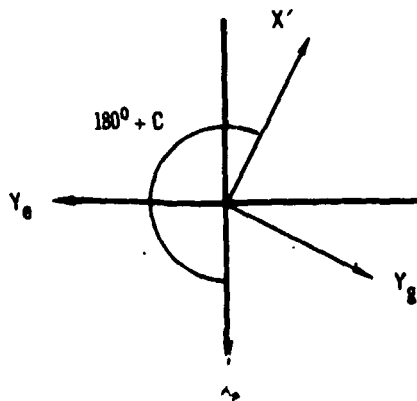
Figure 11- Final Rotation in Transformation from Earth-Referenced to Local-Geocentric Coordinates

45

In this analysis, a positive rotation is defined in the sense adopted for vector cross products in a right-handed system. That is, a positive rotation about the z-axis occurs when the x-axis rotates into the y-axis; positive rotation about the x-axis when y-axis rotates into the z-axis; and positive rotation about the y-axis when the z-axis rotates into the x-axis. The intermediate coordinate system (X', Y_g, Z_e) is eliminated by the met od of successive rotation. The complete transformation is given by

$$
\begin{vmatrix} \bar{1}_{X_g} \\ \bar{1}_{Y_g} \\ \bar{1}_{Z_g} \end{vmatrix} = \begin{vmatrix} \sin \phi_L & 0 & -\cos \phi_L \\ 0 & 1 & 0 \\ \cos \phi_L & 0 & \sin \phi_L \end{vmatrix} \begin{vmatrix} -\cos C & -\sin C & 0 \\ \sin C & -\cos C & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \bar{1}_{X_e} \\ \bar{1}_{Y_e} \\ \bar{1}_{Z_e} \end{vmatrix} \tag{165}
$$

This can be reduced to the single transformation matrix

$$
\begin{vmatrix} \bar{1}_{X_g} \\ \bar{1}_{Y_g} \\ \bar{1}_{Z_g} \end{vmatrix} = \begin{vmatrix} -\sin \phi_L \cos C & -\sin \phi_L \sin C & -\cos \phi_L \\ \sin C & -\cos C & 0 \\ -\cos \phi_L \cos C & -\cos \phi_L \sin C & \sin \phi_L \end{vmatrix} \begin{vmatrix} \bar{1}_{X_e} \\ \bar{1}_{Y_e} \\ \bar{1}_{Z_e} \end{vmatrix} \tag{166}
$$

which defines a direction cosine set (i, j, k) by the equation

$$
\begin{vmatrix} \bar{1}_{X_g} \\ \bar{1}_{Y_g} \\ \bar{1}_{Z_g} \end{vmatrix} = \begin{vmatrix} i_1 & j_1 & k_1 \\ i_2 & j_2 & k_2 \\ i_3 & j_3 & k_3 \end{vmatrix} \begin{vmatrix} \bar{1}_{X_e} \\ \bar{1}_{Y_e} \\ \bar{1}_{Z_e} \end{vmatrix} \tag{167}
$$

Planet referenced velocity in the local-geocentric coordinate system is given by

$$
\begin{vmatrix} \dot{X}_g \\ \dot{Y}_g \\ \dot{Z}_g \end{vmatrix} = \begin{vmatrix} i_1 & j_1 & k_1 \\ i_2 & j_2 & k_2 \\ i_3 & j_3 & k_3 \end{vmatrix} \begin{vmatrix} \dot{X}_e \\ \dot{Y}_e \\ \dot{Z}_e \end{vmatrix} \tag{168}
$$

and

$$
V_g = \sqrt{\dot{X}^2_g + \dot{Y}^2_g + \dot{Z}^2_g}. \tag{169}
$$

Flight path angle₅ are computed by

$$
\sigma = \tan^{-1} \left(\frac{\dot{Y}_g}{\dot{X}_g} \right) \tag{170}
$$

46

and
$$\gamma = \sin^{-1}\left(\frac{-\dot{Z}_g}{V_g}\right) \tag{171}$$

Here σ is the heading angle and λ is the flight path angle.

Wind Axis Coordinates.— Aerodynamic and thrust forces for point-mass problems are conveniently summed in a wind-axis coordinate system, (X_A, Y_A, Z_A). Since the equations of motion are solved in (X_e, Y_e, Z_e) coordinates, the wind-axis components of force must then be resolved into this basic system.

When winds exist, defined by atmospheric velocity components along the local geocentric axes, vehicle velocity relative to the atmosphere is the vector difference of vehicle geocentric velocity and wind velocity. The wind axis system is then determined by the vehicle airspeed, V_A, and the flight path angles relative to the atmosphere λ_A and σ_A. If wind velocity is zero, $V_A = V_g$, $\lambda_A = \lambda$ and $\sigma_A = \sigma$. If there is a wind, with velocity components (X_{gw}, Y_{gw}, Z_{gw}), then

$$V_A = \sqrt{(\dot{X}_g - \dot{X}_{gw})^2 + (\dot{Y}_g - \dot{Y}_{gw})^2 + (\dot{Z}_g - \dot{Z}_{gw})^2} \tag{172}$$

$$\gamma_A = \sin^{-1}\left[-(\dot{X}_g - \dot{X}_{gw})/V_A\right] \tag{173}$$

$$\sigma_A = \tan^{-1}\left[(\dot{Y}_g - \dot{Y}_{gw})/(\dot{X}_g - \dot{X}_{gw})\right] \tag{174}$$

Forces are first resolved from wind axes to the local geocentric coordinates. The wind axes are defined relative to the local geocentric axes by three angles: heading, σ_A; flight path attitude, λ_A, defined above; together with angle, B_A.
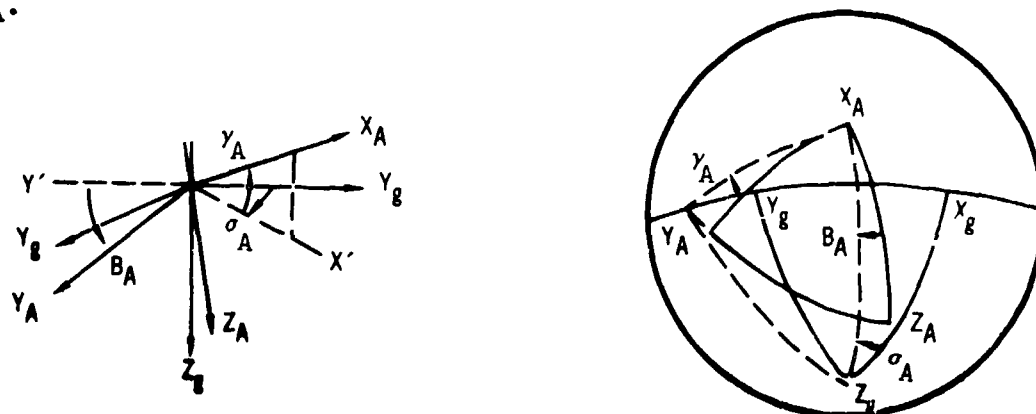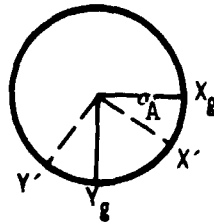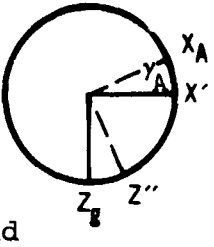

Figure 12.—
Relationship between Local-Geocentric Axes and Wind Axes

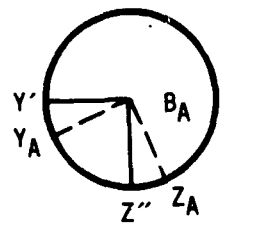Appropriate transformations are

$$
\begin{vmatrix} X' \\ Y' \\ Z_g \end{vmatrix} = \begin{vmatrix} \cos \sigma_A & \sin \sigma_A & 0 \\ -\sin \sigma_A & \cos \sigma_A & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} X_g \\ Y_g \\ Z_g \end{vmatrix} \tag{175}
$$

$$
\begin{vmatrix} X_A \\ Y' \\ Z'' \end{vmatrix} = \begin{vmatrix} \cos \gamma_A & 0 & -\sin \gamma_A \\ 0 & 1 & 0 \\ \sin \gamma_A & 0 & \cos \gamma_A \end{vmatrix} \begin{vmatrix} X' \\ Y' \\ Z_g \end{vmatrix} \tag{176}
$$

and

$$
\begin{vmatrix} X_A \\ Y_A \\ Z_A \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos B_A & \sin B_A \\ 0 & -\sin B_A & \cos B_A \end{vmatrix} \begin{vmatrix} X_A \\ Y' \\ Z'' \end{vmatrix} \tag{177}
$$

The complete transformation from local geocentric horizon coordinates to wind axes then is

$$
\begin{vmatrix} X_A \\ Y_A \\ Z_A \end{vmatrix} = \begin{vmatrix} \cos \gamma_A \cos \sigma_A & \cos \gamma_A \sin \sigma_A & -\sin \gamma_A \\ -\sin \sigma_A \cos B_A + \sin \gamma_A \cos \sigma_A \sin B_A & \cos \sigma_A \cos B_A + \sin \gamma_A \sin \sigma_A \sin B_A & \cos \gamma_A \sin B_A \\ \sin \sigma_A \sin B_A + \sin \gamma_A \cos \sigma_A \cos B_A & -\cos \sigma_A \sin B_A + \sin \gamma_A \sin \sigma_A \cos B_A & \cos \gamma_A \cos B_A \end{vmatrix} \begin{vmatrix} X_g \\ Y_g \\ Z_g \end{vmatrix}
$$

which defines a direction cosine set

$$
\begin{vmatrix} X_A \\ Y_A \\ Z_A \end{vmatrix} = \begin{vmatrix} r_1 & s_1 & t_1 \\ r_2 & s_2 & t_2 \\ r_3 & s_3 & t_3 \end{vmatrix} \begin{vmatrix} X_g \\ Y_g \\ Z_g \end{vmatrix} \tag{179}
$$

The resolution of forces from wind axes to local geocentric then becomes

$$
\begin{vmatrix} F_{X_g} \\ F_{Y_g} \\ F_{Z_g} \end{vmatrix} = \begin{vmatrix} r_1 & r_2 & r_3 \\ s_1 & s_2 & s_3 \\ t_1 & t_2 & t_3 \end{vmatrix} \begin{vmatrix} F_{X_A} \\ F_{Y_A} \\ F_{Z_A} \end{vmatrix}
\tag{180}
$$

For the rotating-planet, the local geocentric components must be resolved into the X_e-Y_e-Z_e system. The required direction cosines are given by Equation (168)

$$
\begin{vmatrix} F_{X_e} \\ F_{Y_e} \\ F_{Z_e} \end{vmatrix} = \begin{vmatrix} i_1 & i_2 & i_3 \\ j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \end{vmatrix} \begin{vmatrix} F_{X_g} \\ F_{Y_g} \\ F_{Z_g} \end{vmatrix}
\tag{181}
$$

The combined transformation from wind axes to local geocentric can be defined as a single matrix transformation

$$
\begin{vmatrix} F_{X_e} \\ F_{Y_e} \\ F_{Z_e} \end{vmatrix} = \begin{vmatrix} o_1 & o_2 & o_3 \\ p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \end{vmatrix} \begin{vmatrix} F_{X_A} \\ F_{Y_A} \\ F_{Z_A} \end{vmatrix} + \begin{vmatrix} mg_{X_e} \\ mg_{Y_e} \\ mg_{Z_e} \end{vmatrix}
\tag{182}
$$

Body-axis coordinates.-- Origin of this system is the vehicle center of gravity with x-axis along the geometric longitudinal axis of the body. Positive direction of the x-axis is from center of gravity to the front of the body. The y-axis is positive to starboard extending from the center of gravity in a water-line plane. The z-axis forms a right-handed orthogonal system. To permit the use of body (x, y, z) axes aerodynamic data, and to convert the body axes components of thrust to the wind axes system, a coordinate transformation must be made. The coordinate transformation shown in Figure 13 involves rotation first through angle of attack, α, then through an auxiliary angle, β'. Noting that

$$
\tan \beta' = \frac{v}{u} \cos \alpha = \tan \beta \cos \alpha
\tag{183}
$$

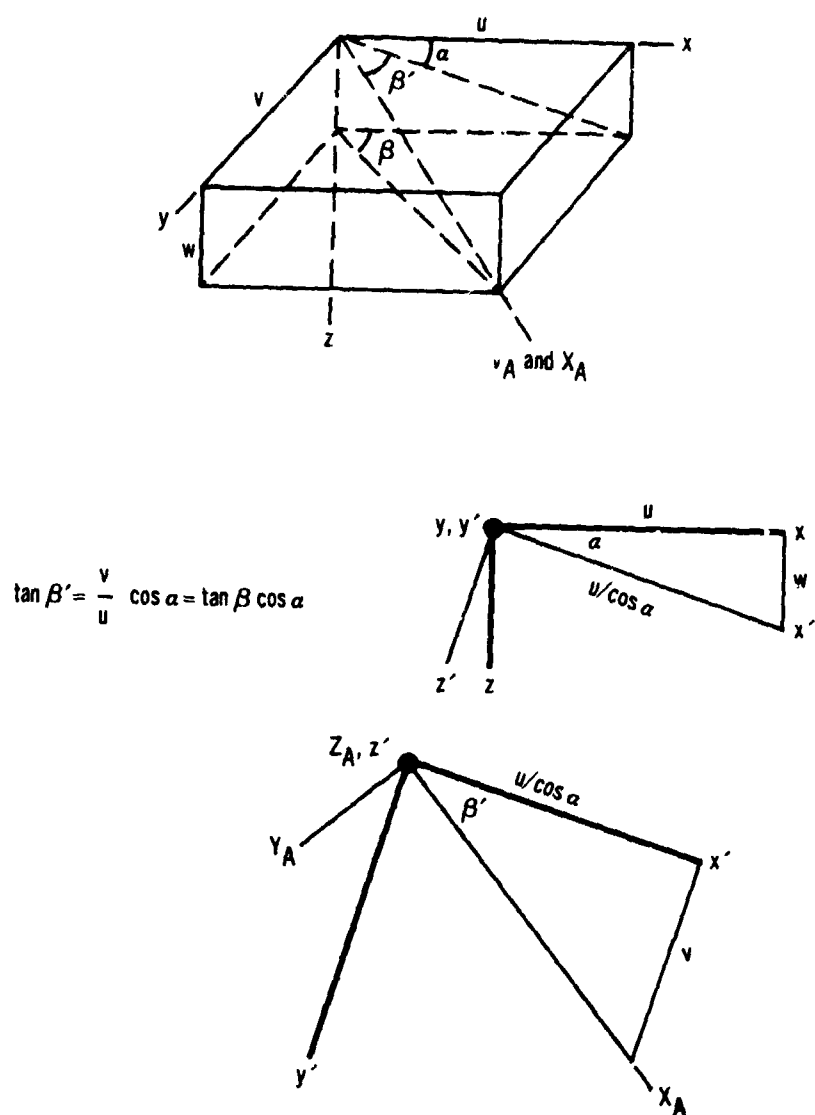$$\tan \beta' = \frac{v}{u} \quad \cos \alpha = \tan \beta \cos \alpha$$

Figure 13.— Relationship Between Body Axes and Wind Axes

the transformation is

$$
\begin{vmatrix} x' \\ y' \\ z' \end{vmatrix} = \begin{vmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{vmatrix} \begin{vmatrix} x \\ y \\ z \end{vmatrix}
$$

$$
\begin{vmatrix} X_A \\ Y_A \\ Z_A \end{vmatrix} = \begin{vmatrix} \cos \beta' & \sin \beta' & 0 \\ -\sin \beta' & \cos \beta' & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} x' \\ y' \\ z' \end{vmatrix}
$$

$$
= \begin{vmatrix} \cos \beta' \cos \alpha & \sin \beta' & \cos \beta' \sin \alpha \\ -\sin \beta' \cos \alpha & \cos \beta' & -\sin \beta' \sin \alpha \\ -\sin \alpha & 0 & \cos \alpha \end{vmatrix} \begin{vmatrix} x \\ y \\ z \end{vmatrix} \qquad (184)
$$

which defines the (u, v, w) direction cosines

$$
\begin{vmatrix} X_A \\ Y_A \\ Z_A \end{vmatrix} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \begin{vmatrix} x \\ y \\ z \end{vmatrix} \qquad (185)
$$

which define the force coefficient transformation

$$
\begin{vmatrix} -C_D \\ C_{Y_A} \\ -C_L \end{vmatrix} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & u_2 & u_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \begin{vmatrix} -C_A \\ C_y \\ -C_N \end{vmatrix} \qquad (186)
$$

The relationship between body and wind-axes aerodynamic coefficients is now established. Note the negative directions of the coefficients relative to the axes.

Inertial coordinates.-- The selected inertial coordinates coincide with the earth references (X_e, Y_e, Z_e) system at time zero. At a later time they differ by the rotation of the earth,

51

$\omega_p t$. The transformation between inertial velocities and planet referenced velocities is derived as follows.

Let $\bar{R}$ be the displacement of the point-mass, (See Figure 9).

In inertial coordinates

$$R = X\bar{1}_X + Y\bar{1}_Y + Z\bar{1}_Z \tag{187}$$

and

$$\bar{V} = \dot{\bar{R}} = \dot{X}\bar{1}_X + \dot{Y}\bar{1}_Y + \dot{Z}\bar{1}_Z \tag{188}$$

In planet-referenced coordinates

$$\bar{R} = X_e\bar{1}_{X_e} + Y_e\bar{1}_{Y_e} + Z_e\bar{1}_{Z_e}$$

However, due to the rotation of the X_e, Y_e, Z_e coordinate system, the velocity is

$$V = \dot{\bar{R}} = \frac{\delta\bar{R}}{\delta t} + \bar{\omega}_p \times \bar{R} \tag{189}$$

where

$$\frac{\delta\bar{R}}{\delta t} = \dot{X}_e\bar{1}_{X_e} + \dot{Y}_e\bar{1}_{Y_e} + \dot{Z}_e\bar{1}_{Z_e} \tag{190}$$

The planet's rotation is about the Z-axis which is also the Z_e-axis. Therefore,

$$\bar{\omega}_p = -\omega_p\bar{1}_Z = -\omega_p\bar{1}_{Z_e}$$
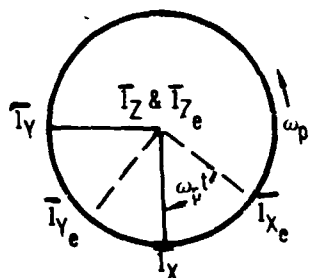
and the required cross product is

$$\bar{\omega}_p \times \bar{R} = \begin{vmatrix} \bar{1}_{X_e} & \bar{1}_{Y_e} & \bar{1}_{Z_e} \\ 0 & 0 & -\omega_p \\ X_e & Y_e & Z_e \end{vmatrix} = (Y_e\omega_p)\bar{1}_{X_e} - (X_e\omega_p)\bar{1}_{Y_e} \tag{191}$$

If Equations (188), (190), and (191) are substituted into Equation (189), it follows that

$$\dot{X}\bar{1}_X + \dot{Y}\bar{1}_Y + \dot{Z}\bar{1}_Z = (\dot{X}_e + \omega_p Y_e)\bar{1}_{X_e} + (\dot{Y}_e - \omega_p X_e)\bar{1}_{Y_e} + (\dot{Z}_e)\bar{1}_{Z_e} \tag{192}$$

The relation between the unit vectors in the inertial system and unit vectors in the planet referenced system are obtained by a single rotation about the Z-axis.

52

The transformation matrix is

$$\begin{vmatrix} \bar{1}_{X_e} \\ \bar{1}_{Y_e} \\ \bar{1}_{Z_e} \end{vmatrix} = \begin{vmatrix} \cos \omega_p t & -\sin \omega_p t & 0 \\ \sin \omega_p t & \cos \omega_p t & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \bar{1}_X \\ \bar{1}_Y \\ \bar{1}_Z \end{vmatrix}$$

(193)

The transformation from planet-referenced velocities to inertial velocities is made with the inverse of the matrix of Equation (193) and the component relations derived in Equation (192)

$$\begin{vmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{vmatrix} = \begin{vmatrix} \cos \omega_p t & \sin \omega_p t & 0 \\ -\sin \omega_p t & \cos \omega_p t & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \dot{X}_e + \omega_p Y_e \\ \dot{Y}_e - \omega_p X_e \\ \dot{Z}_e \end{vmatrix}$$

(194)

The components of inertial velocities are used to calculate the inertial speed of the body as

$$V_I = \sqrt{\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2}$$

(195)

Equation (195) is valid regardless of the inertial coordinate system involved.

Local-geocentric to geodetic coordinates.-- Positions on the planet are specified in terms of geodetic latitude and altitude (for a given longitude) while the motion of the body is computed in a planetocentric system which is independent of the surface. In the computer program, flight-path angle λ and heading angle σ are calculated with respect to the local geocentric coordinates. By definition λ_D and σ_D are angles measured with respect to the local geodetic. Although the maximum difference that can exist between the two coordinate systems is 11 minutes of arc, it may be desirable to know λ_D and σ_D more accurately than is obtained when measured from the local geocentric.

It will be necessary to resolve the geocentric latitude to geodetic latitude for an accurate determination of position. Figure 14 presents the geometry required for describing the position of a point in a meridian plane of a planet shaped in the form of an oblate spheroid.
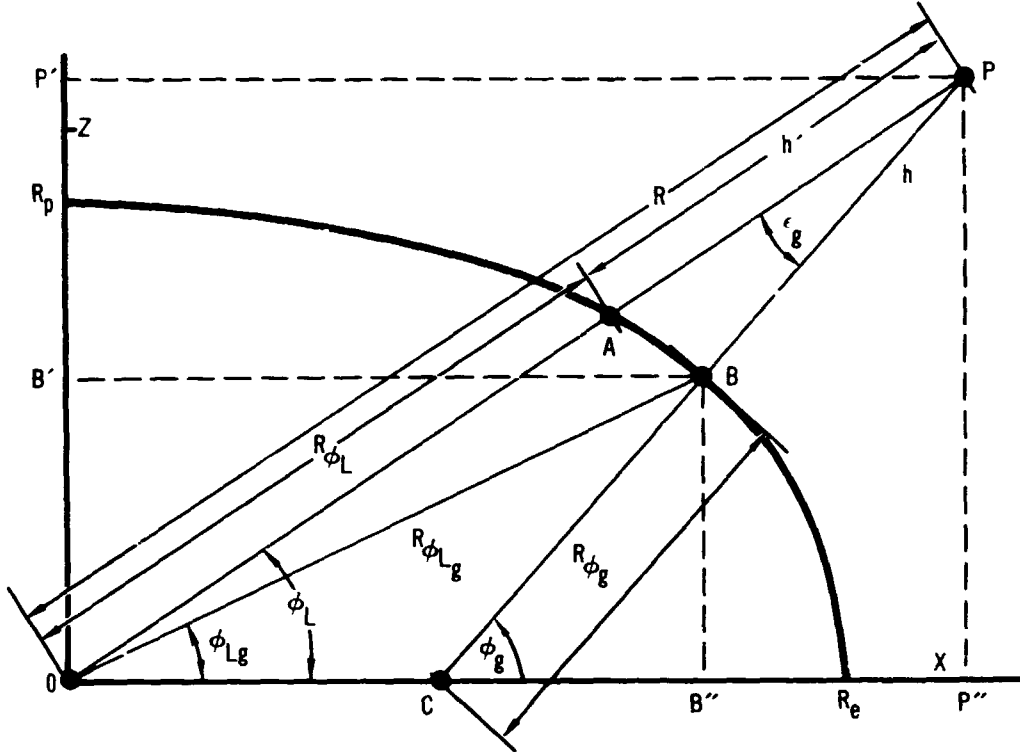
Figure 14.— Planet-Oblateness Effect on Latitude and Altitude

It is apparent from this figure that the most significant difference between the geocentric referenced position and the geodetic position is the distance $\overline{AB}$ on the surface of the reference spheroid. The distance can be defined by a knowledge of the angle ϕ_L; the geocentric latitude; ϕ_g, the geodetic latitude; the corresponding radii; and the distance $\overline{OC}$.

The relationship between the geocentric and geodetic latitude of a point on the surface of a planet which is an oblate spheroid is obtained as follows. The equation for the surface in a meridian plane is

$$\frac{X^2}{R_e^2} + \frac{Z^2}{R_p^2} = 1 \qquad (196)$$

The tangent of the geodetic latitude can be found by determining the negative reciprocal of the slope of a tangent to this ellipse. The expression for this tangent is

$$\text{Tan } \phi_g = - \left. \frac{1}{\frac{d(-Z)}{dX}} \right|_B = - \frac{R_e^2 \, Z_B}{R_p^2 \, X_B} \qquad (197)$$

Note that Z_B is a negative number in the northern hemisphere.

54

The tangent of the geocentric latitude of point B is

$$\text{Tan } \phi_{L_g} = - \frac{Z_B}{X_B} \tag{198}$$

Substituting Equation (198) into Equation (197) gives the required relation

$$\text{Tan } \phi_g = \frac{R_e^2}{R_p^2} \text{ Tan } \phi_{L_g} \tag{199}$$

The expression for the radius of the planet at point B in terms of the geocentric latitude of the point and the equatorial and polar radii is obtained by the rectangular to polar coordinate transformation

$$-Z_B = R_{\phi_{L_g}} \text{ Sin } \phi_{L_g} \tag{200}$$

$$X_B = R_{\phi_{L_g}} \text{ Cos } \phi_{L_g} \tag{201}$$

and, solving for R_{ϕ_L} by substituting Equations (200) and (201) into Equation (196) ^{L_g} gives

$$R_{\phi_{L_g}} = \frac{R_e R_p}{\sqrt{R_p^2 \text{ Sin}^2 \phi_{L_g} + R_e^2 \text{ Cos}^2 \phi_{L_g}}} \tag{202}$$

$$= \frac{\cos \phi_L}{\cos \phi_{L_g}} \; R_e \Big/ \sqrt{[(R_e/R_p)(\tan \phi_{L_g}/\tan \phi_L)]^2 \sin^2 \phi_L + \cos^2 \phi_L}$$

It may be seen from Figure 14 that

$$\overline{F''P'} = \overline{OP'} - \overline{OB'} \tag{203}$$

or

$$h \sin \phi_g = \overline{OP} \sin \phi_L - R_{\phi_{L_g}} \sin \phi_{L_g} \tag{204}$$

Likewise

$$\overline{B''P''} = \overline{OP''} - \overline{OG''} \tag{205}$$

or

$$h \cos \phi_g = \overline{OP} \cos \phi_L - R_{\phi_{L_g}} \cos \phi_{L_g} \tag{206}$$

If Equation (204) is divided by Equation (206) and then the quotient is divided by tan ϕ_{L_g}, there results

$$\left(\frac{\tan \phi_g}{\tan \phi_{L_g}}\right) = \left[\overline{OP} \left(\frac{\sin \phi_L}{\sin \phi_{L_g}}\right) - R\phi_{L_g}\right] / \left[\overline{OP} \left(\frac{\cos \phi_L}{\cos \phi_{L_g}}\right) - R\phi_{L_g}\right] \quad (207)$$

or

$$(R_e/R_p)^2 \left(\frac{\cos \phi_L}{\cos \phi_{L_g}}\right) = \left(\frac{\cos \phi_L}{\sin \phi_{L_g}}\right) + \left[(R_e^2 - R_p^2)/R_p^2\right] \left[R\phi_{L_g}/\overline{OP}\right] \quad (208)$$

Finally, if Equation (208) is multiplied by $(R_p \sin \phi_{L_g})/(R_e \sin \phi_L)$, it follows that

$$\left(\frac{R_e}{R_p}\right) (\tan \phi_{L_g}/\tan \phi_L) = \left(\frac{R_p}{R_e}\right) + \left[1 - (R_p/R_e)^2\right] \left(\frac{R_e \sin \phi_{L_g}}{R_p \sin \phi_L}\right) (R\phi_{L_g}/\overline{OP}) \quad (209)$$

Let

$$U = (R_e \tan \phi_{L_g}/R_p \tan \phi_L)$$

$$= (R_p \tan \phi_g/R_e \tan \phi_L) \quad (210)$$

Then it follows from Equations (202) and (209) that

$$U = \left(\frac{R_p}{R_e}\right) + [R_e/\overline{OP}] \left\{U/\sqrt{U^2 \sin^2 \phi_L + \cos^2 \phi_L}\right\} [1 - (R_p/R_e)^2] \quad (211)$$

Equation (211) is solved by an iterative scheme.

Then

$$\phi_g = \tan^{-1} \left[\left(\frac{R_e U}{R_p}\right) \tan \phi_L\right] \quad (212)$$

The flight-path and heading angles corrected to the local geodetic latitude are computed by

$$\gamma_D = \sin^{-1}\left(\frac{-\dot{Z}_{g_1}}{V_{g_1}}\right) = \sin^{-1}\left(\frac{-\dot{Z}_g - \{\dot{X}_g(\phi_g - \phi_L)\}}{V_g}\right) \quad (213)$$

Since the magnitude of vector V_g is equal to the magnitude of vector V_{g_1}

and

$$\sigma_D = \sin^{-1}\left(\frac{\dot{Y}_g}{\sqrt{\dot{X}_{g_1}^2 + \dot{Y}_{g_1}^2}}\right) = \sin^{-1}\left(\frac{\dot{Y}_g}{\sqrt{\{\dot{X}_g + \dot{Z}_g(\phi_g - \phi_L)\}^2 + \dot{Y}_g^2}}\right)$$

$$(214)$$

56

Auxiliary Computations

In addition to the computations which can be made from the problem formulation as presented in preceding sections, several other quantities are available as optional calculations.

a. Planet-surface referenced range, R_D

b. Great-circle range, R_g

c. Down- and cross-range, X_D and Y_r

d. Theoretical burnout velocity, V_{theo}

e. Velocity losses, V_P, V_{grav}, V_D, and V_{ML}

f. Orbital variables and satellite target

Planet-surfaced referenced range.-- The total distance traveled over the surface of the planet is computed as the integrated surface range. If the distance traveled by the vehicle over a given portion of the trajectory is

$$R_D' = \int_{t_1}^{t_2} V_g \, dt \tag{215}$$

then the curvilinear planet surface referenced range is

$$R_D = \int_{t_1}^{t_2} \frac{R_{\phi_L}}{R} V_g \cos \gamma \, dt \tag{216}$$

The flight-path angle, λ, is referenced to local geocentric coordinates for this computation.

Great-circle range.-- Great-circle distance from the launch point to the instantaneous vehicle position, R_g, may also be required. Expressions for this distance are derived as follows.

By spherical trigonometry, (see Figure 15)

$$\cos \frac{R_g}{R'} = \cos(90-\phi_L)\cos(90-\phi_{L_0}) + \sin(90-\phi_L)\sin(90-\phi_{L_0}) \cos(\theta_L-\theta_{L_0}) \tag{217}$$

or simplifying

$$\cos \frac{R_g}{R'} = \sin \phi_L \sin \phi_{L_0} + \cos \phi_L \cos \phi_{L_0} \cos(\theta_L-\theta_{L_0}) \tag{218}$$

Therefore,

$$R_g = R' \cos^{-1} \left[\sin \phi_L \sin \phi_{L_0} + \cos \phi_L \cos \phi_{L_0} \cos(\theta_L-\theta_{L_0}) \right] \tag{219}$$
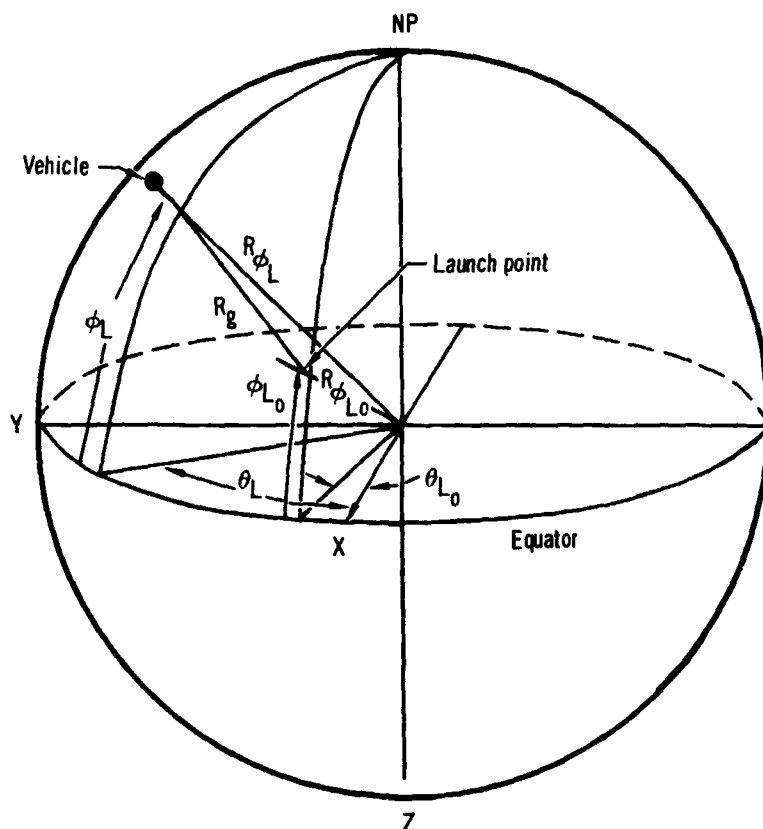
57

Figure 15.——Great-Circle Range

However, since the planets are generally oblate spheroids, R'
is not a constant radius. An approximation may be obtained by
averaging the planet's radius at the launch point and at the
vehicle's position. Therefore, define the average radius, R',
as

$$R' = \frac{R_{\phi_L} + R_{\phi_{L_O}}}{2} \tag{220}$$

and the surface-referenced great-circle range from the launch
point to the vehicle is

$$R_g = \left[\frac{R_{\phi_L} + R_{\phi_{L_O}}}{2}\right] \cos^{-1}\left[\sin\phi_L \sin \phi_{L_O} + \cos \phi_L \cos \phi_{L_O} \cos(\theta_L - \theta_{L_O})\right] \tag{221}$$

Down- and cross-range.-- Down- and cross-range from the
initial great circle can be determined. The initial great
circle is determined from the input quantities σ_O, ϕ_{L_O}, and

58

θ_{L_O} (see Figure 16)   Then the cross range of a particular trajectory point is defined as the perpendicular distance from the point to the initial great circle. The downrange is then the distance along the initial great circle from the initial point to the point P at which the cross range is measured. From the spherical triangle, Figure 16, the great circle range LF to the point F is computed by Equation (221).

The right spherical triangle LPF is solved for the downrange, X_D, and the cross range, Y_D.

$$X_D = R' \cos^{-1} \left(\frac{\text{Cos LF}}{\text{Cos} (\sin^{-1}(\sin \text{LF} \sin \xi))} \right) \qquad (222)$$

$$Y_D = R' \sin^{-1} (\sin \text{LF} \sin \xi) \qquad (223)$$

where

$$\xi = \zeta - \sigma_o \qquad (224)$$
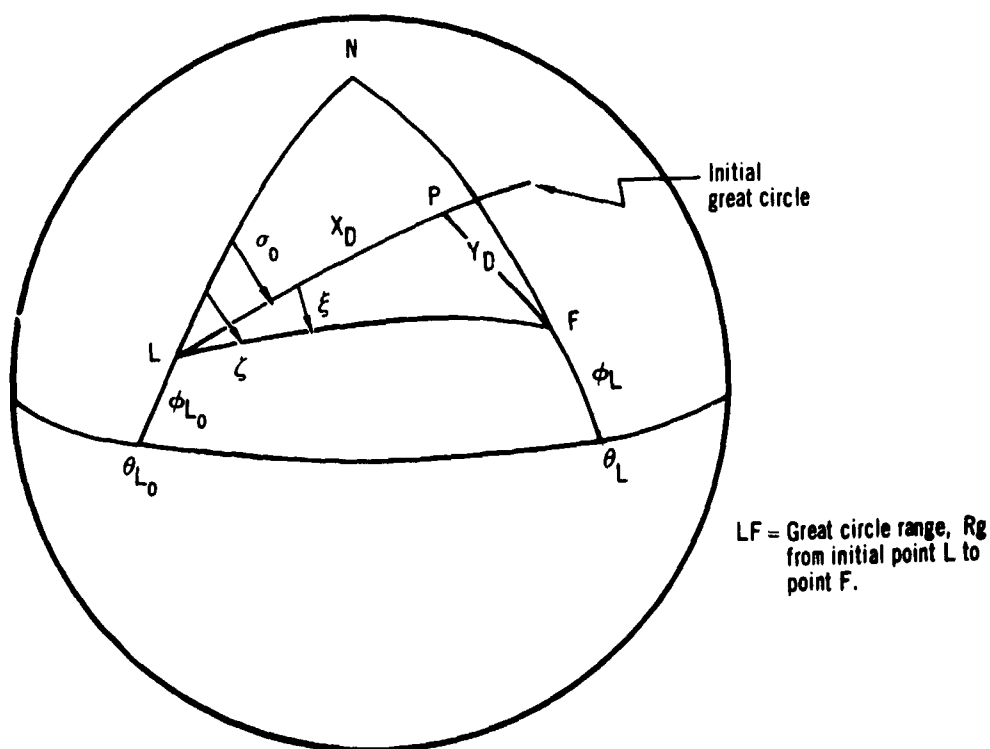
R' is defined by Equation (220)

Figure 16.—Downrange and Crossrange Geometry

Theoretical burnout velocity and losses.-- For trajectory
and performance optimization studies, it is convenient to know
the theoretical burnout velocity possible and the velocity
losses due to gravity, aerodynamic drag, and atmospheric back
pressure upon the engine nozzle. These quantities may be com-
puted as follows:

Theoretical Velocity

$$V_{theo} = \int_{t_1}^{t_2} \frac{T_{VAC}}{m} dt$$

(225)

Speed Loss Due to Gravity

$$V_{grav} = \int_{t_1}^{t_2} -g_{Zg} \sin \gamma \, dt$$

(226)

Speed Loss Due to Aerodynamic
Drag

(227)

$$V_D = \int_{t_1}^{t_2} \frac{D}{m} dt$$

Speed Loss Due to Atmosphere
Back Pressure Upon the Engine Nozzle

(228)

$$V_P = \int_{t_1}^{t_2} - \frac{PA_e}{m} dt$$

(229)

Maneuvering Losses

$$V_{ML} = \int_{t_1}^{t_2} \left(\frac{T_{VAC} - PA_e}{m} \right) \left(\cos \alpha - 1 \right) dt.$$

The resultant velocity $V_g'(t_2)$ is obtained by adding the components
computed to the initial value $V_g'(t_1)$

$$V_g'(t_2) = V_g'(t_1) + V_{theo} + V_{grav} + V_D + V_P \quad (230)$$

The maneuvering losses are valid only if λ_T is zero for the
engine.

Orbital variables and satellite target.-- Certain functions
of use in orbital trajectory calculations have been added to the
point mass equations of motion used in the Steepest Descent Opti-
mization Program. These functions permit the specification of
terminal conditions in inertial space when this is convenient. A
further set of functions will permit rendezvous calculations with
a satellite in a circular orbit about a central planet.

Orbital variable calculations commence immediately after
the calculation of vehicle inertial velocity. Flight path
angles in inertial space are computed from the expressions

$$\sigma_I = \tan^{-1}\left(\frac{\dot{Y}_g + \omega_p\,|R|\,\cos\phi_L}{\dot{X}_g}\right) \tag{231}$$

$$\gamma_I = \sin^{-1}\left(\frac{\dot{Z}_g}{|v_I|}\right) \tag{232}$$

The inclination angle, i, is the angle between the plane con-
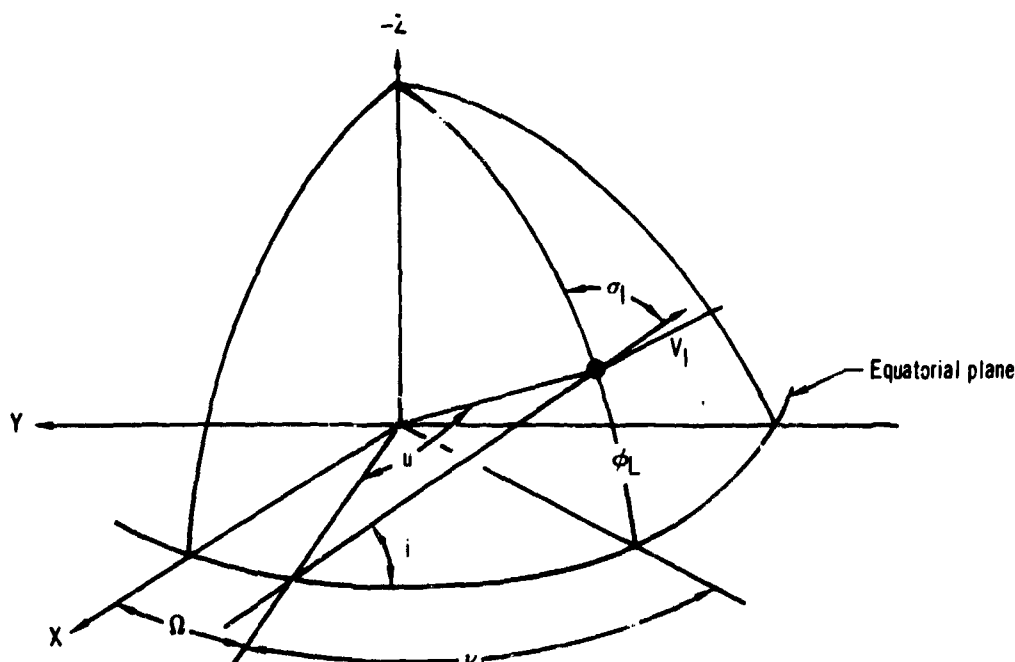taining the velocity vector and the center of the earth, and
the equatorial plane.

Figure 17.— Orbital Plane Geometry

61

Applying spherical trigonometry to Figure 17, we obtain the relationship

$$\cos i = \cos \phi_L \sin \sigma_I \qquad (233)$$

The difference in longitude between the vehicle and the ascending node, ν, is given by

$$\tan \nu = \sin \phi_L \tan \sigma_I \qquad (234)$$

The inertial longitude is given by

$$\theta_I = \theta_L - \omega_p t \qquad (235)$$

and the inertial longitude of the ascending node by

$$\Omega = \theta_I - \nu \qquad (236)$$

It is convenient to know the central angle, u, in the orbital plane. Measuring from the ascending node, we obtain

$$\tan u = \frac{\tan \phi_L}{\cos \sigma_I} \qquad (237)$$

The orbital variable calculation introduces positional and velocity information from a second body. This body is a satellite considered in a circular orbit about the earth. Its orbital height, h_s, is specified and remains constant. Position in the orbit is computed from an initial central angle, ϕ_{s_O}, by the expression

$$\phi_s = \phi_{s_O} + \omega_s t \qquad (238)$$

The satellite angular velocity is obtained from the satellite inertial velocity, V_{c_s}, where

$$V_{c_s} = \sqrt{\frac{\mu_g}{(R_e + h_s)}} \qquad (239)$$

where μ_g is the gravitational potential constant and R_e the earth radius. It should be noted that Equation (239) assumes a spherical earth; for the earth radius is taken as constant, and none of the higher order gravitational harmonics are included. Knowing V_{c_s}, it follows that

$$\omega_s = \frac{V_{c_s}}{R_e + h_s} \qquad (240)$$

The variables of this section provide sufficient information to either rendezvous with or terminate the trajectory in a specified position relative to the satellite.

VEHICLE CHARACTERISTICS

Methods by which the aerodynamic, propulsive, and physical characteristics of a vehicle are introduced into the computer program are presented in this section. Form and preparation of the input data are discussed, together with methods by which stages and staging may be used to increase the effective data storage area allotted to a description of the vehicle's properties.

Aerodynamic Coefficients

The primary objective of the aerodynamic data input subprogram is to provide for a complete accounting of the various contributions to the aerodynamic forces and moments regardless of the flight conditions of the vehicle being considered. Two techniques are available for use in the digital computer program: (a) an n-dimensional table look-up and interpolation and (b) an m-order polynomial function of n variables prepared by "curve fit" techniques. In the first method, the proper value for each term is obtained by an interpolation in "n" dimensions where the number of dimensions is taken to be the number of parameters to be varied independently plus the dependent variable. This method has the advantage of accurately describing most non-linear variations with reasonable preparation effort. The amount of storage space which must be allocated to such a method, however, can achieve unreasonable proportions and may require substantial computing time for the interpolation as the number of dimensions are increased. The second method has essentially the opposite characteristics; that is, a large amount of data may be represented with a small amount of storage space, and computation time is held to reasonable limits, but the data variations which may be represented must be regular. A substantial amount of effort can be required for the preparation of data by a curve-fit technique. Both these methods are very convenient when the amount of data to be handled is moderate, but tend to become unmanageable when large amounts of data are required. This usually occurs when the program, having several degrees of freedom, is committed to one or the other of these two techniques. Therefore, the computer program incorporates both of the techniques discussed as a compromise to take advantage of the more desirable features of both. To do this, a general set of data equations have been programmed which define each of the aerodynamic forces. In general, the coefficients for these equations will be obtained from a curve-read interpolation. Several simplifications may be made to the equations depending on the flight condition and vehicle to be considered.

63

Often the particular application will not require some of
the terms listed in order to describe the flight path and vehicle
under consideration. The subprogram is arranged so that the com-
puter will assign a constant value to any curve for which the
data has not been supplied. For most curves, the constant value
will be zero. This technique may be used to reduce the time
required for the preparation of data. Values intermediate to
those introduced in a tabular listing will be obtained by linear
interpolation.

Aerodynamic Forces

Aerodynamic forces are customarily defined by three mutu-
ally perpendicular forces. These are lift (L), drag (D), and
side force (Y). Lift force is perpendicular to the velocity
vector in a vertical plane; drag force is measured along the
velocity vector but in opposite direction; side force is meas-
ured in the horizontal plane, positive toward the right, pro-
vided the bank angle is zero. If the bank angle is not zero,
L and Y will be rotated by $-B_A$ about the velocity vector.
Coordinates are shown in Figure 18.
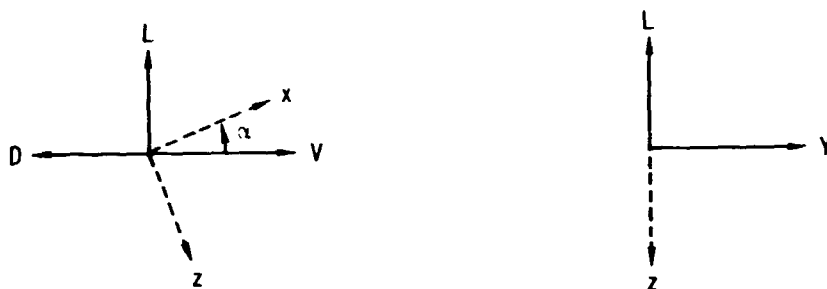
Figure 18.— Aerodynamic Forces - Wind Axes

These forces may be expressed in the form

$$L = q(V,h) \; SC_L(V,h,\alpha,\beta) \tag{241}$$

$$D = q(V,h) \; SC_D(V,h,\alpha,\beta) \tag{242}$$

$$Y = q(V,h) \; SC_Y(V,h,\alpha,\beta) \tag{243}$$

where q is the dynamic pressure and S a convenient reference
area. The aerodynamic coefficients C_L, C_D, and C_Y may be ex-
pressed in terms of the aerodynamic derivatives.

64

$$C_L = C_{L_o} + C_{L_\alpha} \alpha + C_{L_\alpha}2 \, \alpha|\alpha| + C_{L_\beta} |\beta|$$

$$+ C_{L_\beta}2\beta^2 + C_{L_{\alpha\beta}} \alpha|\beta| \tag{244}$$

$$C_D = C_{D_o} + C_{D_\alpha} |\alpha| + C_{D_\alpha}2 \, \alpha^2 + C_{D_\beta} |\beta|$$

$$+ C_{D_\beta}2 \, \beta^2 + C_{D_{\alpha\beta}} |\alpha||\beta| \tag{245}$$

$$C_Y = C_{Y_o} + C_{Y_\alpha}|\alpha| + C_{Y_\alpha}2 \, \alpha^2 + C_{Y_\beta}\beta$$

$$+ C_{Y_\beta}2 \, \beta|\beta| + C_{Y_{\alpha\beta}}|\alpha|\beta \tag{246}$$

Alternatively, the aerodynamic derivatives may be expressed as tabular functions of Mach number (M_N), α, and β, that is, functions of the state variables and the control variables.

On occasion, it may be convenient to measure the aerodynamic forces in the body axis coordinate system introduced in a preceding section, pages 49 to 51. In this case, normal force, (n_f), is measured along the -z axis, side force (y) along the y axis, and axial force (a) along the -x axis, as in Figure 19.
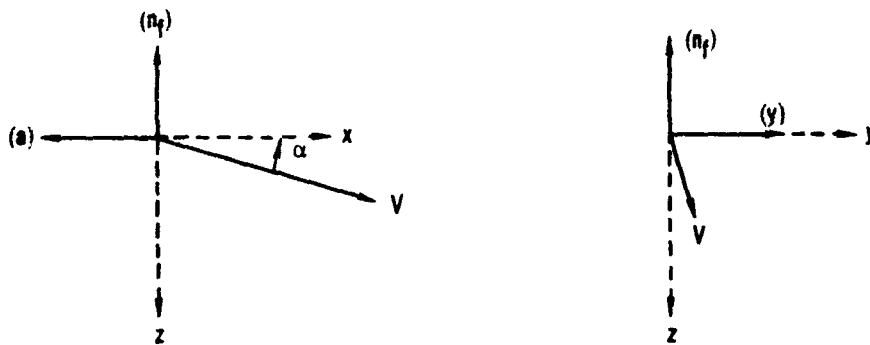

Figure 19.— Aerodynamic Force in Body Axes

The specification of forces in the body axis system is similar to that in the wind axis system

$$n_f = qSC_N \tag{247}$$

$$a = qSC_A \tag{248}$$

$$y = qSC_y \tag{249}$$

65

where the body axis aerodynamic coefficients are

$$C_N = C_{N_0} + C_{N_1} + C_{N_\alpha}2 \; \alpha \; |\alpha|$$
$$+ \; C_{N_\beta} |\beta| + C_{N_\beta}2 \; \beta^2 + C_{N_{\alpha\beta}} \; \alpha \; |\beta|$$

$$C_A = C_{A_0} + C_{A_\alpha} |\alpha| + C_{A_\alpha}2 \; \alpha^2 \tag{250}$$
$$+ \; C_{A_\beta} |\beta| + C_{A_\beta}2 \; \beta^2 + C_{A_{\alpha\beta}} \; |\alpha\beta|$$

$$C_y = C_{y_0} + C_{y_\alpha} |\alpha| + C_{y_\alpha}2\alpha^2 \tag{251}$$
$$+ \; C_{y_\beta}\beta + C_{y_\beta}2 \; \beta |\beta| + C_{y_{\alpha\beta}} \; |\alpha|\beta$$

$$\tag{252}$$

Thrust and Fuel Flow Data

The techniques employed to introduce thrust and fuel-flow data into the equations of motion are developed in an approach similar to that employed for aerodynamic data. An n-dimensional tabular listing and interpolation technique is used with the independent variables being defined by the type of propulsion unit being considered. For the present formulation, the propulsion units are grouped into the following options: (1) rocket, (2) air breathing engine.

Propulsion option (1) rocket.-- The thrust of a rocket motor is assumed variable with stage time, altitude, and, if the rocket is controllable, it will also vary with throttle setting. The altitude effect is determined by the exit area of the nozzle, A_e, and the ambient pressure, P. If the thrust is specified for some constant ambient air pressure, the altitude correction can be calculated within the subprogram. If the rocket motor is uncontrolled, the vacuum thrust, in pounds, will be introduced by a tabular listing as a function of time, in seconds, and corrected as follows:

$$T = Max \; [T_{vac} - PA_e, \; 0] \tag{253}$$

The propellant consumption rate is specified by a tabular listing in slugs per second as a function of time, in seconds, for the single engine options, or computed from the thrust and the engine specific impulse, I_{SP}, for the multiple engine options.

If the rocket is controlled, the propellant mass flow rate $\dot{m}_f$ is introduced by a tabular listing as a function of throttle setting. The thrust is then specified by a tabular listing as a function of mass flow rate.

66

Propulsion option (2) air breathing engines.—An air-breathing engine is strongly affected by the environmental conditions under which it is operating. Engines which would be grouped in this classification are turbojets, ramjets, pulsejets, turboprops, and reciprocating machines. The parameters considered significant in the program are

(a) Altitude (h-ft)

(b) Mach number (M_N)

(c) Angle of attack (α-degrees), and

(d) Throttle setting (N-units defined by problem)

Both the thrust and fuel flow are functions of these variables. In order to accommodate these variables, a five-dimensional tabular listing and interpolation are used to obtain both thrust and fuel flow. The thrust has no further correction as the effects of all parameters are assumed included in the interpolated value.

Engine perturbation factors.—The engine options include provision for two data scaling factors for use in parametric studies. These are in the form

$$T = \varepsilon_{13} \, T_{VAC} + \varepsilon_{14} \qquad (254)$$

Components of the thrust vector.—The equations used to reduce the thrust vector to its components along the body axes are

$$T_x = T \cos\lambda_T \qquad (255)$$

$$T_y = -T \sin\lambda_T \cos\phi_T \qquad (256)$$

and

$$T_z = -T \sin\lambda_T \sin\phi_T \qquad (257)$$

ϕ_T and λ_T are defined and explained in the control variable section.

Reference weight and propellant consumed.—Rate of change of vehicle mass, m, is set equal to the negative of the total mass flow rate, $-\dot{m}_t$. $\dot{m}$ is integrated to give variation of vehicle mass, m. The instantaneous mass is used in the computation of the body motion. The reference weight is obtained by an auxiliary calculation

$$W_T = m(32.174) \qquad (258)$$

67

The propellant consumed is computed as

$$m_f = m_o - m \qquad\qquad (259)$$

where m_o is a reference mass input equal to the initial vehicle mass

Stages and Staging

A problem common in missile performance analyses and encountered frequently in airplane performance work is that of staging or the release of discrete masses from the continuing airframe. The effect of dropping a booster rocket or fuel tanks is often great enough to require that the complete set of aerodynamic data be changed. Configuration changes at constant weight, such as extending drag brakes or turning on afterburners, may also require revising the aerodynamic or physical characteristics of the vehicle. Another use of the staging technique is possible with the present computer program which does not involve physical changes to the configuration; this technique may be used to revise the aerodynamic descriptors as a function of aerodynamic attitude or Mach number. With this use of the stage concept, accurate descriptions of the forces acting upon the vehicle may be maintained over wide attitude ranges, if required.

VEHICLE ENVIRONMENT

The models for simulating the environment in which a
vehicle will operate are presented in this section. This
environment includes the atmosphere properties, wind velocity,
and the field associated with the planet over which the ve-
hicle is moving. The shape of the planet and the conversion
from geodetic to geocentric latitudes are also considered.
In the discussions which follow, the descriptions of vehicle
environment pertain to the planet Earth. The environmental
simulation may be extended to any planet by replacing appro-
priate constants in the describing equations.

Atmosphere

The concept of a model atmosphere was introduced many
years ago, and over the years several models have been de-
veloped. Reference 20 outlines the historical background of
the gradual evolution of the ARDC model. The original (1956)
ARDC model (Reference 20) was revised to reflect the density
variation with altitude that was obtained from an analysis
of artificial satellite orbit data. This revision is the
widely used 1959 ARDC Model Atmosphere and is the basic option
in the present program.

The advantage of a model atmosphere is that it provides
a common reference upon which performance calculations can
be based. The model is not intended to be the "final word"
on the properties of the atmosphere for a particular time
and location. The atmosphere properties are quite variable
and are affected by many parameters other than altitude.
At the present time, the "state-of-the-art" is not advanced
to the point where these parameters can be accounted for;
it may be several years before the effects of some parameters
can be evaluated.

1959 ARDC Model Atmosphere.-- The 1959 ARDC Model Atmos-
phere is specified in layers assuming either isothermal or
linear temperature lapse-rate sections. This construction
makes it very convenient to incorporate other atmospheres,
either from specifications for design purposes or for other
planets. The relations which mathematically specify the
1959 ARDC Model Atmosphere are as follows (Reference 21)
the 1959 ARDC Model Atmosphere is divided into 11 layers as
noted in the table below.

Layer	H_b-Lower Altitude (Geopotential) Meters	Upper Altitude (Geopotential) Meters
1	0	11,000
2	11,000	25,000
3	25,000	47,000
4	47,000	53,000
5	53,000	79,000
6	79,000	90,000
7	90,000	105,000
8	105,000	160,000
9	160,000	170,000
10	170,000	200,000
11	200,000	700,000

For layers 1, 3, 5, 7, 8, 9, 10, and 11, a linear molecular-scale temperature lapse-rate is assumed and the following equations are used:

$$H_{gp} = \frac{.3048h}{1 + .3048h/6356766} \qquad \text{Meters} \qquad (260)$$

$$T_M = (T_M)_b \left[1 + K_1(H_{gp} - H_b) \right] \qquad {}^{\circ}R \qquad (261)$$

$$T = T_M \left[A - B \tan^{-1} \left(\frac{H_{gp} - C}{D} \right) \right] \qquad {}^{\circ}R \qquad (262)$$

$$P = P_b \left[1 + K_1(H_{gp} - H_b) \right]^{-K_2} \qquad \text{Lb./Ft.}^2 \qquad (263)$$

$$\rho = \rho_b \left[1 + K_1(H_{gp} - H_b) \right]^{-(1 \cdot K_2)} \qquad \text{Slugs/Ft.}^3 \qquad (264)$$

$$V_s = 49.021175(T_M)^{1/2} \qquad \text{Ft./Sec.} \qquad (265)$$

$$\nu = 2.269681 \times 10^{-8} \left[\frac{T^{3/2}}{(T+198.72)\rho} \right] \qquad \text{Ft.}^2/\text{Sec.} \qquad (266)$$

For the isothermal layers 2, 4, and 6, the following changes are made

$$P = P_b e^{-K_3(H_{gp}-H_b)} \qquad (267)$$

$$\rho = \rho_b e^{-K_3(H_{gp}-H_b)} \qquad (268)$$

70

Values of the temperature, pressure, density, and altitude at the base of each altitude layer are listed below along with the appropriate values K_1, K_2, and K_3.

Layer

Quantity	1	2	3	4	5	6
K_1	$-.22556913 \times 10^{-4}$	0	$.13846580 \times 10^{-4}$	0	$-.15920187 \times 10^{-4}$	0
K_2	-5.2561222	0	11.388265	0	-7.5921765	0
K_3	0	$.15768852 \times 10^{-3}$	0	$.12086887 \times 10^{-3}$	0	$.20623442 \times 10^{-3}$
T_b	518.688	389.988	389.988	508.788	508.788	298.188
P_b	2116.2170	472.67599	51.975418	2.5154578	1.2180383	2.1082485×10^{-2}
ρ_b	2.37692×10^{-3}	7.0611078×10^{-4}	7.7643892×10^{-5}	2.8803201×10^{-6}	1.3947125×10^{-7}	4.1190042×10^{-9}
H_b	0	11000.	25000.	47000.	53000.	79000.

Layer

Quantity	7	8	9	10	11
K_1	$.24245841 \times 10^{-4}$	$.88628910 \times 10^{-4}$	$.75434123 \times 10^{-5}$	$.35071476 \times 10^{-5}$	$.22212914 \times 10^{-5}$
K_2	8.5411986	1.7082397	3.4164794	6.8329589	9.7613698
K_3	0	0	0	0	0
T_b	298.188	406.188	2386.188	2566.188	2836.188
P_b	2.1811754×10^{-3}	1.5564912×10^{-4}	7.5604667×10^{-6}	5.8971644×10^{-6}	2.9769746×10^{-6}
ρ_b	4.2614856×10^{-9}	$2.2324424 \times 10^{-10}$	$1.8458849 \times 10^{-12}$	$1.3397990 \times 10^{-12}$	$6.1150607 \times 10^{-13}$
H_b	90000.	105000.	160000.	170000.	200000.

Values of the appropriate constants to be applied in the temperature equation, Equation (262), are listed below.

H_{gp}(Km)	A	B	C	D
0-90	1.	0.	-	-
90-180	.75951115	.17416404	220,000.	25,000.
180-1200	.93578678	.27396592	180,000.	140,000.

U. S. Standard Atmosphere, 1962.--The part of the U.S. Standard Atmosphere, 1962, below 90 kilometers geometric altitude (295,276 ft. altitude) is defined in the same way as the 1959 model--by the hydrostatic equation and a piecewise linear variation of temperature with geopotential altitude. Equations (260) to (268) are, therefore, applicable with a different set of constants. These constants, based on the published tabulation of atmosphere properties (Reference 22) at the base altitudes, are presented below. The 1962 model uses a different set of relationships above 90 kilometers. These have not been included. The tables define 1962 model properties between sea level and 295,800 ft. geometric altitude.

Values of the temperature, pressure, density, and altitude at the base of each altitude layer are listed below along with the appropriate values of K_1, K_2, and K_3.

	Layer			
Quantity	1	2	3	4
K_1	$-.2255877 \times 10^{-4}$	0	$.48012406 \times 10^{-5}$	$.12199559 \times 10^{-4}$
K_2	$-.5255871 \times 10^{1}$	0	$.32844801 \times 10^{2}$	$.12202470 \times 10^{2}$
K_3	0	$.1576958 \times 10^{-3}$	0	0
T_b	518.67	389.97	389.97	413.104
P_b	2116.217	472.6812	114.3431	17.22518
ρ_b	$.2377002 \times 10^{-2}$	$.7061512 \times 10^{-3}$	$.1708202 \times 10^{-3}$	$.2429209 \times 10^{-4}$
H_b	0	10999.474	19999.191	32354.854

	Layer			
Quantity	5	6	7	8
K_1	0	$-.7383899 \times 10^{-5}$	$-.1572230 \times 10^{-4}$	0
K_2	0	$-.1709562 \times 10^{+2}$	$-.8602817 \times 10$	0
K_3	$.1262323 \times 10^{-3}$	0	0	$.1891214 \times 10^{-3}$
T_b	487.17	487.17	454.668	325.170
P_b	2.302550	1.226346	.3766873	.2106440
ρ_b	$.2753526 \times 10^{-5}$	$.146637 \times 10^{-5}$	$.4826665 \times 10^{-6}$	$.3773977 \times 10^{-7}$
H_b	47051.501	52042.023	61077.348	79192.936

Within the altitude range considered, T and T_M (Equation (262)) are equal.

Atmosphere limitations.-- The validity of the 1959 ARDC model is limited to altitudes below 700 km; although the program is arranged to extrapolate the relationships to greater altitudes, if desired. Extrapolation to greater altitudes is accomplished by altering the cutoff altitude.

At an altitude greater than 2.6×10^6 feet, no calculations are made, and the program sets kinematic viscosity, speed of sound, pressure, temperature, and density to zero. At and below

sea level the parameters, pressure, temperature, and density are set to the values below. Other terms are computed as normal.

$$\text{Pressure} = 2116.2170 \ \text{Lb/Ft}^2 \tag{269}$$

$$\text{Temperature} = 518.688 \ {}^\circ\text{R} \tag{270}$$

$$\text{Density} = 2.37692 \times 10^{-3} \ \text{Slugs/Ft}^3 \tag{271}$$

At altitudes between 90 kilometers and 2.6×10^6 feet, the speed of sound is set to 846.50255, and kinematic viscosity is set to 2.3519252×10^{-7} over density. Other terms are computed as normal.

The 1962 model is limited to altitudes below 295,8000 feet (90 kilometers). It is suggested that zero values be returned above that altitude. At and below sea level, the sea level values should be employed. When the atmosphere constants are determined from the published tabulations at the base altitude, the calculated values at intermediate altitudes may not agree with the tabulated values to the number of significant figures in the tables. This has been allowed for in the 1959 model by developing coefficients with the necessary extra precision to give agreement between the calculated values and published tables at all altitudes. The values calculated by the 1962 model are good to about four significant figures, which should be adequate for most purposes.

Kinematic viscosity and speed of sound lose their physical significance at very high altitudes, and are not normally defined by model atmospheres above 90 kilometers. The constant values by the 1959 model option were added to provide data required by the aerodynamic heating routine. The aerodynamic heating calculation should not be used with the 1962 model option above 90 kilometers. The constant values of ν and V_s in the 1959 model will give reasonable values of Mach number and Reynolds number for use in the aerodynamics calculations to altitudes somewhat above 90 kilometers, say 350,000 feet, above which constant aerodynamic coefficients should be used.

Winds Aloft

The winds-aloft subprogram provides for three separate methods of introducing the wind vector: as a function of altitude, a function of range, and a function of time. This facilitates the investigation of wind effects for the conventional performance studies. The wind vector is approximated by a series of straight line segments for each of the methods mentioned above.

Four options are used to define the wind vector in the computer program. The three components of the wind vector in a geodetic horizon coordinate system can be specified as tabular listings with linear interpolations (curve reads) in the following options.

Wind options (0).-- In this option the wind vector is zero throughout the problem. This allows the analyst the option of evaluating performance without the effects of wind. This option causes the winds-aloft computations to be bypassed.

Wind option (1).-- In this option the components of the wind vector are specified as a function of time. Wind speeds are specified in feet per second and time in seconds.

Wind option (2).-- The three components of the wind vector are introduced as a function of altitude in this option. Wind speed is specified in feet per second and altitude in feet.

Wind option (3).-- In this option the components of the wind vector are introduced as a function of range. Wind speed is specified in feet per second and range in nautical miles. The range utilized in this computation is the great-circle range.

By staging of the wind option, it is possible to switch from one method of reading wind data to another during the computer run. Care must be exercised in this operation, however, as the switching will introduce sharp-edged gusts if there are sizeable differences in the wind vector from one option to another at the time of switching. This effect should be avoided except in cases where gust effects are being studied.

Gravity

This section presents the equations necessary for the introduction of the gravity components into the equations of motion. These components were determined by taking partial derivatives of the gravity potential equation. The potential equation adopted has been recommended for use in the Six-Degree-of-Freedom Flight-Path Study computer program by AFCRC. Constants for the potential equation were determined from References 23,24,and 25.

Spherical harmonics are normally used to define the gravity potential field of the Earth, References 23 through 26. Each harmonic term in the potential is due to a deviation of the potential from that of a uniform sphere. In the present analysis the second-, third-, and fourth-order terms are considered. The first-order term, which would account for the error introduced by assuming that the mass center of the Earth

74

is at the origin of the geocentric coordinate system is assumed to be zero. With this assumption

$$U = \frac{\mu_g}{R} \left[1 + \frac{J}{3} \left(\frac{R_e}{R} \right)^2 P_2 + \frac{H}{5} \left(\frac{R_e}{R} \right)^3 P_3 + \frac{K}{30} \left(\frac{R_e}{R} \right)^4 P_4 + \cdots \right] \quad (272)$$

where P_2, P_3, and P_4 are Legendre functions of geocentric latitude ϕ_L expressed as

$$P_2 = 1 - 3 \sin^2 \phi_L$$

$$P_3 = 3 \sin \phi_L - 5 \sin^3 \phi_L$$

$$P_4 = 3 - 30 \sin^2 \phi_L + 35 \sin^4 \phi_L \quad (273)$$

The gravitational acceleration along any line is the partial derivative of U along that line. At this point, it should be noted that the three mutually perpendicular directions in the spherical coordinate system are identical (other than sign) to those in the local-geocentric-horizon coordinate system which is defined previously. Therefore, the acceleration in the ϕ_L direction is identical to gx_g, and the acceleration in the R direction is identical to $-gz_g$. Or in the equation form:

$$gz_g = - \frac{\partial U}{\partial R} = - \frac{\mu_g}{R} \left[- \frac{2J}{3} \left(\frac{R_e^2}{R^3} \right) P_2 - \frac{3H}{5} \left(\frac{R_e^3}{R^4} \right) P_3 - \frac{4K}{30} \left(\frac{R_e^4}{R^5} \right) P_4 \right]$$

$$+ \frac{\mu_g}{R^2} \left[1 + \frac{J}{3} \left(\frac{R_e}{R} \right)^2 P_2 + \frac{H}{5} \left(\frac{R_e}{R} \right)^3 P_3 + \frac{K}{30} \left(\frac{R_e}{R} \right)^4 P_4 \right] \quad (274)$$

$$gx_g = \frac{1}{R} \frac{\partial U}{\partial \phi_L} = \frac{\mu_g}{R^2} \left[\frac{J}{3} \left(\frac{R_e}{R} \right)^2 (-6 \sin \phi_L \cos \phi_L) \right.$$

$$+ \frac{H}{5} \left(\frac{R_e}{R} \right)^3 (3 \cos \phi_L - 15 \sin^2 \phi_L \cos \phi_L)$$

$$\left. + \frac{K}{30} \left(\frac{R_e}{R} \right)^4 (-60 \sin \phi_L \cos \phi_L + 140 \sin^3 \phi_L \cos \phi_L) \right] \quad (275)$$

Collecting terms:

$$gz_g = \frac{\mu_g}{R^2} \left[1 + J \left(\frac{R_e}{R} \right)^2 P_2 + \frac{4H}{5} \left(\frac{R_e}{R} \right)^3 P_3 + \frac{K}{6} \left(\frac{R_e}{R} \right)^4 P_4 \right] \quad (276)$$

$$gx_g = \frac{\mu_g}{R^2} \left[-2J \left(\frac{R_e}{R} \right)^2 P_5 + \frac{3H}{5} \left(\frac{R_e}{R} \right)^3 P_6 + \frac{2K}{3} \left(\frac{R_e}{R} \right)^4 P_7 \right] \quad (277)$$

75

where

$$P_5 = \sin \phi_L \cos \phi_L$$

$$P_6 = \cos \phi_L \ (1 - 5 \sin^2 \phi_L)$$

$$P_7 = \sin \phi_L \cos \phi_L \ (-3 + 7 \sin^2 \phi_L) \qquad (278)$$

Equations (276) and (277) are used in the gravity subroutine with the following values recommended for the constants:

$$\mu_g = 1.407698 \times 10^{16} \ \text{ft.}^3/\text{sec.}^2$$

$$R_e = 20,925,631. \ \text{ft.}$$

$$J = 1623.41 \times 10^{-6}$$

$$K = 6.37 \times 10^{-6} \qquad (279)$$

It should be noted that these constants and equations pertain to the planet Earth; however, it is possible to use these same equations for any other planet. For this reason, the values of these constants is an input to the program so that the applicable constants may be inserted for the planet under consideration. Due to limited knowledge cf the gravitational fields of other planets, it is probable that zero values would be assigned to some of the harmonic coefficients when the program is used for entry studies on other planets.

The above equations are applicable to a non-rotating planet as the centrifugal relieving effects caused by the planet's rotation are included in the equations of motion. In addition, the effects of local anomalies must be added if it is desired to make a weight-to-mass conversion based on a measured weight. The program has the options of retaining the first, third, and fourth order terms.

WEIGHTING MATRICES

The perturbation constraint, Equation (5),limits perturbed control histories to the neighborhood of a time varying "volume" lying in the neighborhood of the nominal trajectory. At any time point in the control perturbation predicted by steepest-descent analysis is directly proportional to the product of instantaneous weighting matrix (metric tensor) inverse and partial derivative magnitudes. This is readily apparent in the unconstrained solution; for in this case, Equation (43) reduces to

$$\{\delta\alpha\} = \mp \; [W]^{-1} [G]' \{\lambda_{\phi\Omega}\} \sqrt{\frac{DP^2}{I_{\phi\phi}}} \tag{280}$$

Now consider any point t' along the trajectories, Figures 20 and 21 in the two variable case;at this point the integrand of Equation (5) defines an inclined ellipse in the $(s_{\alpha_1}, s_{\alpha_2})$ plane. In the usual case, when off-diagonal weighting matrix elements are zero, the principle axes of this ellipse parallel the $(s_{\alpha_1}, s_{\alpha_2})$ axes, and the length of the major and minor axes are inversely proportional to the weighting matrix diagonal elements (directly proportional to the inverse diagonal elements).

Now the steepest-descent analysis presented earlier in this report can be applied to the case of time pulses in the controls at any time t'. This type of problem reduces the analysis to an ordinary parameter optimization problem, References 15 and 16 in the variables s_{α_1} and s_{α_2}. In this case, the steepest-descent direction is clear, for it lies along the line joining the ellipse center and the ellipse tangency point to the local performance contours. Further details may be obtained from Reference 15. It is clear, however, in the case considered that by suitably choosing the weighting matrix elements at t = t', any descending direction can be made that of steepest-descent.

In general, this behavior persists in the case of a time-distributed control perturbation. A badly chosen weighting matrix can either greatly inhibit convergence rate or,in an extreme case, cause convergence failure due to the limited accuracy of digital computation. Examples illustrating this behavior in the parameter optimization case can be found in Reference 17.

Some weighting matrix options available in the program of References 1 and 2 are described below.
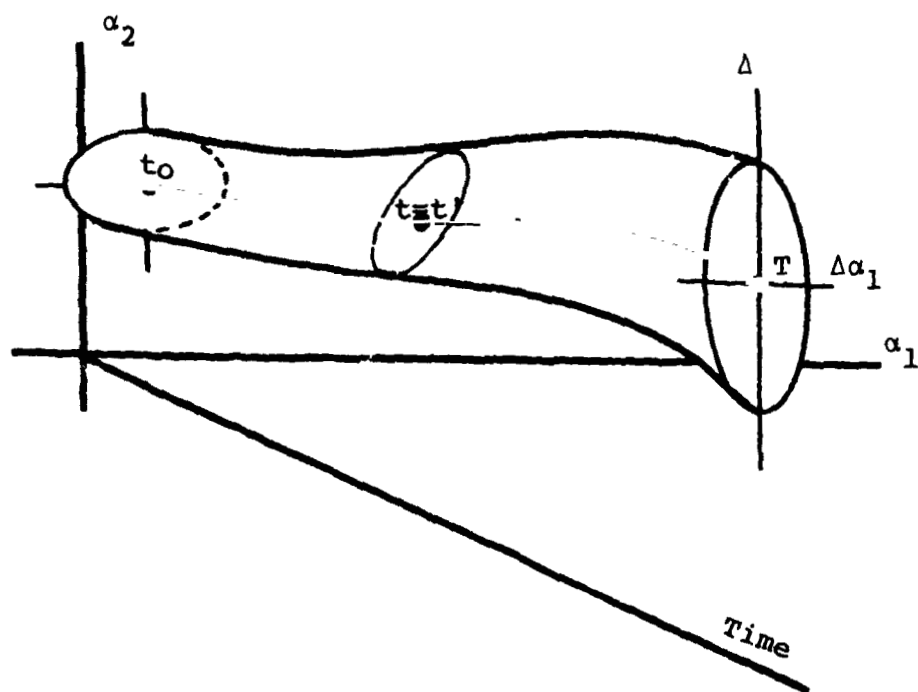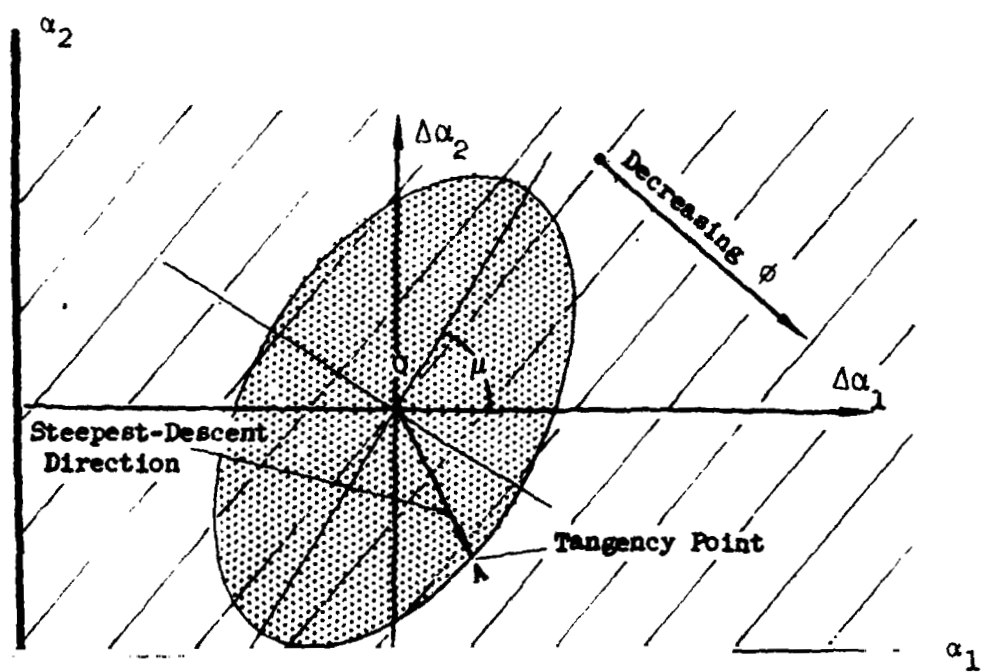
77

Figure 20.— Time Varying Control Values

Figure 21.— Control Measure at the Instant t = t'

Multiple Control Variable Optimization

The most insidious types of convergence failures are
those in which the payoff function fails to reach the optimal
value, while at the same time the terminal constraints are
achieved. This problem is prevalent among optimization prob-
lems involving multiple control variables when a unit weighting
matrix is employed. The reason for this behavior becomes
apparent from consideration of an optimization problem involving
two control variables, α_1 and α_2, where the weighting function
$W(t)$ is taken as the unit matrix and α_1 is consistently more
powerful than α_2. More powerful implies that a small change
in α_1 will produce a greater change in the payoff function
than an equal change in α_2 will produce for the type of per-
turbation of interest. In this situation, the greater control
variable perturbation will tend to appear in α_1 rather than α_2.
The total perturbation in the first control variable over a
series of steepest-descent steps will, therefore, always tend
to be greater than the total perturbation in the second control
variable, provided α_1 remains the more powerful of the two
control variables, no matter how many steps in the descent have
been taken. Now, the total required perturbation in control
variables during convergence from the nominal trajectory to
the optimum trajectory is purely a function of the particular
problem under consideration and the nominal path chosen. There
is no reason for supposing the total perturbation required in
the powerful control variable to be either greater than or less
than that of the less powerful one. It follows that when the
steepest descent process is presented with a situation in which
the converse is true, i.e., the weaker control variable re-
quires the greatest total perturbation, there will of neces-
sity be a high risk of false convergence.

The argument can be made more specific. First create a
measure of the total perturbation required during convergence
from the chosen nominal to the optimal solution for each con-
trol variable. For example, this can be achieved by separately
integrating the absolute value of the perturbation required
along the trajectory, i.e.,

$$\left\{ \overline{\Delta P_i^*} \right\} = \left\{ \int_{t_0}^{T} \left\{ \left| \Delta \alpha_i \right| \right\} dt \right\} \tag{2$\cdots$}$$

where

$$\left\{ \Delta \alpha_i \right\} = \left\{ \alpha_{i_{optimum}} \right\} - \left\{ \alpha_{i_{nominal}} \right\} \tag{282}$$

Total perturbation achieved by the steepest descent
method after C descents can be expressed in the form

$$\left\{\overline{\Delta P_i}(C)\right\} = \left\{\left|\int_{t_0}^{T} \left|\sum_{j=1}^{C} \delta\alpha_{ij}(t)\right| dt\right|\right\} \tag{283}$$

Here $\delta\alpha_{ij}(t)$ is the perturbation of the ith control variable in the jth descent at time t. Now suppose that the rth control variable perturbations are consistently greater than the sth, by some order of magnitude P, so that

$$\delta\alpha_{rj}(t) = 0(P) \quad \delta\alpha_{sj}(t) \tag{284}$$

Inverting this

$$\delta\alpha_{sj}(t) = 0(-P) \quad \delta\alpha_{rj}(t) \tag{285}$$

On substituting Equation (285) into Equation (283)

$$\overline{\Delta P_s}(C) = \int_{t_0}^{T} \left|\sum_{j=1}^{C} \delta\alpha_{sj}(t)\right| dt \leq \int_{t_0}^{T} \sum_{j=1}^{C} \left|0(-P) \delta\alpha_{rj}(t)\right| dt$$

or

$$\overline{\Delta P_s}(C) \leq 0(-P) \int_{t_0}^{T} \sum_{j=1}^{C} \left|\delta\alpha_{rj}(t)\right| dt \tag{286}$$

Monotonic Descents

Consider the case in which the successive control variable perturbations at any instant are monotonic as the number of descents increases. From Equation (286) the total change in the sth control variable will always be P orders of magnitude less than that in the rth control variable, no matter how many descents are made. In this case, dispensing with the inequality in Equation (286)

$$\overline{\Delta P_s}(C) = 0(-P) \int_{t_0}^{T} \sum_{j=1}^{C} \left|\delta\alpha_{rj}(t)\right| dt \tag{287}$$

The same remarks are true of Equation (283). On substituting Equation (283) with i = r into Equation (287)

$$\overline{\Delta P_s}(C) = 0(-P) \quad \overline{\Delta P_r}(C) \tag{288}$$

That is, the total change in the sth control variable after C descents depends only on the change in the more powerful control variable and the ratio of their powers. In such a case,

once the constraints are satisfied, the rth control variable
will approach its final history with regularly diminishing
steps. The final history of the rth control variable may will
be near optimum. The sth control variable history will, of
necessity, be perturbed by smaller amounts on each successive
descent during this period until it finally approaches its
limiting value of $\Delta P_s(\infty)$. It follows from Equation (288) that

$$\overline{\Delta P_s}(\infty) = 0(-P)\ \overline{\Delta P_r}(\infty) \tag{289}$$

In general there is assurance that either

$$\overline{\Delta P_r}(\infty) = \overline{\Delta P_r}^* \tag{290}$$

or

$$\overline{\Delta P_s}(\infty) = \overline{\Delta P_s}^* \tag{291}$$

If the original total perturbation required in the sth
control variable, $\overline{\Delta P_s}^*$, is P orders of magnitude less than
that required in the rth control variable, as it might be if
previous knowledge of the optimum history of the control
variable were absent. Convergence would tend to be at least
one order of magnitude worse in the weaker (s^{th}) control
variable than the rth control variable.

If, on the other hand, the total perturbation required
in the sth control variable had been Q orders of magnitude
greater than that of the rth control variable, we would have

$$\overline{\Delta P_s}^* = 0(Q)\ \overline{\Delta P_r}^* \tag{292}$$

Combining with Equations (289) and (290), in this case

$$\overline{\Delta P_s}(\infty) = 0(-P)\ \overline{\Delta P_r}^* = 0(-(P+Q))\ \overline{\Delta P_s}^* \tag{293}$$

If the mean perturbation obtained in the sth control
variable history after the descent is $\overline{\Delta \alpha_s}$, and that required
for convergence is $\overline{\Delta \alpha_s}^*$, then

$$\overline{\Delta \alpha_s} = \frac{\overline{\Delta \alpha_s}^*}{0(P+Q)} \tag{294}$$

Now problems in which the control variable powers are
in a ratio of 10^3:1 are not uncommon in trajectory optimi-
zation. It is also fairly common to create a nominal trajec-
tory in which the weaker control variable has, say, ten times
greater total required perturbation than the more powerful one.
In such a case, from Equation (294), when convergence is com-
pleted, the weaker control variable may be practically unper-
turbed from the nominal history.

In practice, the successive descents need not be monotonic. It is, therefore possible for the weaker control variable to increase its total perturbation while the more powerful control variable oscillates. However, it seems reasonable to assume that the descent is "almost monotonic." In this sense, the above analysis is "almost correct," and, hence, provides at least a qualitative insight into the general behavior of the steepest descent process with multiple control variables. It should also be noted that the arguments of this section hinge on the persistence of unequal control variable powers.

The possibility of failing to converge to the desired end constraints is somewhat more remote than that of failing to converge the payoff function. The dominant control variables for the payoff function are very often the dominant control variables for the constraints and, hence, will continue to be perturbed until the constraints are achieved. In addition, the control variables usually need not be optimized to achieve the end constraints. In any case, failure to achieve the end constraints is immediately obvious, whereas the only reliable method of checking the payoff function convergence is to obtain the same result from as different and widely removed as nominal as possible or to apply a time dependent equivalent of the topologically invariant warping of Reference 15.

Control Variable Power

Previously in this section the concept of control variable power has been used; specifically this is a measure of the ability of a control variable to influence the final value of the payoff function.

It may be recalled from a previous section that the change in payoff function is given by

$$d\phi = \int_{t_0}^{T} \lfloor \lambda_{\phi\Omega} \rfloor [G] \{\delta\alpha\} dt + \lfloor \lambda_{\phi\Omega}(t_0) \rfloor \{\delta x(t_0)\} \qquad (25a)$$

When considering changes in the control variables alone, the second term can be ignored.

Suppose at time t' we create a pulse, i.e., a Dirac Delta function of unit magnitude in each of the control variables. The change in ϕ produced by these pulses will be

$$\delta(d\phi) = \lfloor \lambda_{\phi\Omega}(t') \rfloor [G(t')] \{1\} \qquad (29S)$$

where {1} is a unit column matrix. The elements of the row matrix $\lambda_{\phi\Omega}G$ indicate the effect of separate pulses in each of the control variables. These elements will be referred to as the instantaneous payoff function sensitivities, $_s\phi_c$, or

82

control impulse response functions with respect to performance

$$\left\lfloor s_{\alpha}^{\phi}(t)\right\rfloor = \left\lfloor \lambda_{\phi\Omega}(t)\right\rfloor \left[G(t)\right]$$

(296)

These quantities measure the ability of a control variable to affect the terminal payoff function when constraint changes are not considered. The instantaneous payoff function sensitivities, $\lambda_{\phi\Omega}(t)$, are intimately connected with the optimum control variable perturbations. In the case of no terminal constraints from Equation (43) the optimum control variable perturbation is given by

$$\left\{\delta\alpha\right\} = \dot{\div}\left[W\right]^{-1}\left[G\right]'\left\{\lambda_{\phi\Omega}\right\}\sqrt{\frac{DP^2}{I_{\phi\phi}}}$$

(297)

Substituting Equation (296) into Equation (297)

$$\left\{\delta\alpha\right\} = \dot{\div}\left[W\right]^{-1}\left\{s_{\alpha}^{\phi}\right\}\sqrt{\frac{DP^2}{I_{\phi\phi}}}$$

(298)

That is, the optimum perturbation varies directly with the instantaneous sensitivities and the inverse weighting matrix. If the problem being investigated involves terminal constraints, it follows from Equation (43) and (296) that

$$\left\{\delta\alpha\right\} = \dot{\div}\left[W\right]^{-1}\left\{\left\{s_{\alpha}^{\phi}\right\} - \left[G\right]'\left[\lambda_{\psi\Omega}\right]\left[I_{\psi\psi}\right]^{-1}\left\{I_{\psi\phi}\right\}\right\}$$
$$\sqrt{\frac{DP^2 - \left\lfloor d\psi\right\rfloor\left[I_{\psi\psi}\right]^{-1}\left\{d\psi\right\}}{I_{\phi\phi} - \left\lfloor I_{\psi\phi}\right\rfloor\left[I_{\psi\psi}\right]^{-1}\left\{I_{\psi\phi}\right\}}}$$
$$+ \left[W\right]^{-1}\left[G\right]'\left[\lambda_{\psi\Omega}\right]\left[I_{\psi\psi}\right]^{-1}\left\{d\psi\right\}$$

(299)

These results suggest an approach to the problem of false convergence. Since the problem is due to small perturbations in weak control variables, an inverse weighting matrix based on the control variable sensitivities may be employed to accentuate the weaker control variables. Effectively changing the basis of optimization from that perturbation having the greatest change in ϕ for a given total perturbation magnitude to that perturbation having the greatest change in ϕ assuming all control variables are equally important and must, therefore, be perturbed by a reasonable amount.

When concerned with terminal constraint variations, the above definition of control variable power may be modified. In this case interest lies in control variable perturbations

83

which improve the payoff function while providing a prescribed change in the constraints. These control variable perturbations may be either one of two components or "modes." The first mode considered is one in which constraints undergo prescribed changes with the minimum control variable perturbation possible. From Equation (299) this is seen to be when

$$(DP)_1^2 = \lfloor d\psi \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ d\psi \right\} \tag{300}$$

with a corresponding mode shape of

$$\left\{ \delta\alpha_1 \right\} = \left[W \right]^{-1} \left[G \right]' \left[\lambda_{\psi\Omega} \right] \left[I_{\psi\psi} \right]^{-1} \left\{ d\psi \right\} \tag{301}$$

The second mode considered is one in which the payoff function is improved while holding the terminal constraints constant. From Equation (299) this mode is given by

$$\left\{ \delta\alpha_2 \right\} = \pm \left[W \right]^{-1} \left\{ \left\{ s_\alpha^\phi \right\} - \left[G \right]' \left[\lambda_{\psi\Omega} \right] \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\} \right\} \sqrt{ \frac{DP_2^2}{ I_{\phi\phi} - \lfloor I_{\psi\phi} \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\} } } \tag{302}$$

which may be written in the form

$$\left\{ \delta\alpha_2 \right\} = \pm \left[W \right]^{-1} \left\{ \left\{ s_\alpha^\phi \right\} - \left\{ s_\psi^\phi \right\} \right\} \sqrt{ \frac{DP_2^2}{ I_{\phi\phi} - \lfloor I_{\psi\phi} \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\} } } \tag{303}$$

where

$$\left\{ s_\psi^\phi \right\} = \left[G \right]' \left[\lambda_{\psi\Omega} \right] \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\} \tag{304}$$

Substituting pulse variations of this second type into Equation (25a) and using Equation (296)

$$\delta(d\phi) = \pm \lfloor s_\alpha^\phi \rfloor \left[W \right]^{-1} \left\{ \left\{ s_\alpha^\phi \right\} - \left\{ s_\psi^\phi \right\} \right\} \sqrt{ \frac{DP_2^2}{ I_{\phi\phi} - \lfloor I_{\psi\phi} \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\} } } \tag{305}$$

$$= \pm \lfloor s_\alpha^\phi \rfloor \left[W \right]^{-1} \left\{ s_{\alpha\psi}^\phi \right\} \sqrt{ \frac{DP_2^2}{ I_{\phi\phi} - \lfloor I_{\psi\phi} \rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ I_{\psi\phi} \right\} } } \tag{306}$$

where $s_{\alpha\psi}^\phi$, the mixed control variable instantaneous sensitivities, are defined by

$$\left\{ s_{\alpha\psi}^\phi \right\} = \left\{ s_\alpha^\phi \right\} - \left\{ s_\psi^\phi \right\} \tag{307}$$

84

Substituting Equation (307) into Equation (303) we obtai..

$$\left\{\delta\alpha_2\right\} = \mp \left[W\right]^{-1}\left\{s^\phi_{\alpha\psi}\right\} \sqrt{\frac{DP^2_2}{I_{\phi\phi} - \lfloor I_{\psi\phi}\rfloor\lfloor I_{\psi\psi}\rfloor^{-1}\left\{I_{\psi\phi}\right\}}} \qquad (308)$$

From Equation (308) it follows that when the payoff function is improved while leaving the terminal constraints unaltered, the control variable perturbations at any point will vary directly as the product of the inverse weighting matrix and the mixed sensitivity matrix.

Equations (298) and (308) enable establishment of rational methods for choosing weighting functions to insure payoff function convergence. When limited to diagonal weighting matrices to insure reasonable perturbations in all the control variables, one need only increase those diagonal elements of W^{-1} corresponding to the weaker control variable elements or decrease those corresponding to the powerful elements. Further, the elements of $\left\{s^\phi_\alpha\right\}$ or $\left\{s^\phi_{\alpha\psi}\right\}$ can be used to decide in which class a particular control variable belongs at any instant. End point convergence could be improved by basing W^{-1} on $G'\lambda_{\psi\Omega}$, Equation (301). To date, this has not been attempted.

By integrating the absolute value of the instantaneous sensitivities over the whole trajectory, a measure of the total control variable power is obtained. If the terminal constraint variations are ignored

$$\left\{s^\phi_\alpha\right\} = \left\{\int^T_{t_0} \left| s^\phi_\alpha \right| dt\right\} \qquad (309)$$

If the terminal constraint variations are held to zero

$$\left\{s^\phi_{\alpha\psi}\right\} = \left\{\int^T_{t_0} \left| s^\phi_{\alpha\psi} \right| dt\right\} \qquad (310)$$

The elements of these column matrices will be referred to as the integrated payoff function sensitivities.

The integrated payoff function sensitivities based on Equation (309) should approach zero at the optimum; those based on Equation (310) do not, of necessity, approach zero. Either form, in its own way, serves to measure the overall ability of a control variable to affect the payoff function and is, therefore, a measure of the control variable power previously defined.

If there were perfect numerical accuracy in the steepest descent process, the control variable histories would continue

85

to be perturbed until such time as all the control variable powers as measured by integrated payoff sensitivities based on Equation (310) were zero. In practice, this condition is practically impossible to achieve; in fact, it is often difficult to reduce these control variable powers by more than an order of magnitude when the weighting function is absent. This, then, is the basic reason for the weighting function matrix, for without one a high risk of failure to converge the weaker control variables is present unless foreknowledge of the required total perturbations $\Delta \bar{P}_i^*$ is available.

It will generally be impossible to obtain the desired total perturbations $\Delta \bar{P}_i^*$ directly, for to do this would require a knowledge of the optimum control variable history. In lieu of this knowledge, make the assumption that the ΔP_i^* all have the same order of magnitude. Reasonable convergence can then be assured by choosing weighting matrices based on this assumption. Several such weighting matrices based on the payoff function sensitivities will be described in the remainder of this section; to date, only diagonal matrices have been utilized in this manner.

Weighting Functions Based on Integrated Sensitivity

Choose a diagonal weighting matrix in the form,

$$\left\lceil W_{ii} \right\rfloor^{-1} = \frac{1}{M+1} \left[A_{ii} + B_{ii} \frac{\sum_{j=1}^{M} S_{\alpha\psi_j}^{\phi}}{S_{\alpha\psi}^{\phi}} \right] \tag{311}$$

where M is the number of control variables.

With equally powerful control variables, the unit matrix is obtained with $A_{ii} = B_{ii} = 1$, for then

$$\left\lceil W_{ii} \right\rfloor^{-1} = \left[\frac{1 + \frac{M S^{\phi}}{S^{\phi}}}{M+1} \right] = \left\lfloor 1 \right\rfloor \tag{312}$$

where S^{ϕ} is a typical sensitivity.

In the case of unequal sensitivities this form of the weighting function will insure that we have perturbations of similar orders of magnitude in each of the control variables. For example, suppose there are

r   control variables with $S^{\phi} = O(R)$,

s   control variables with $S^\phi = O(S)$, and

t   control variables with $S^\phi = O(T)$

then

$$\left\{\delta\alpha_2(t)\right\} \sim \left[\frac{1 + \dfrac{\sum S^\phi}{S^\phi_1}}{M + 1}\right] \left\{s^\phi_{\alpha\psi}(t)\right\} \tag{313}$$

Integrating

$$\int_{t_o}^{T} \left\{\left|\delta\alpha_2(t)\right|\right\} dt \sim \left[\frac{1 + \dfrac{\sum S^\phi}{S^\phi_1}}{M + 1}\right] \left\{S^\phi_1\right\} = \left\{\frac{S^\phi_1 + \sum S^\phi_i}{M + 1}\right\} \tag{314}$$

Partitioning the matrix according to the power of the control variables and considering orders of magnitude

$$\frac{1}{M+1} \left\{ \begin{array}{c} S^\phi_r + \sum S^\phi \\ \overline{S^\phi_s} + \overline{\sum S^\phi} \\ \overline{S^\phi_t} + \overline{\sum S^\phi} \end{array} \right\} = \frac{1}{(M+1)} \left\{ \begin{array}{c} (r+1)O(R) + sO(S) + tO(T) \\ \overline{rO(R)} + (s+1)O(S) + \overline{tO(T)} \\ \overline{rO(R)} + sO(S) + (t+1)O(T) \end{array} \right\} \tag{315}$$

where S^ϕ_r, S^ϕ_s, and S^ϕ_t are typical sensitivities of order R,S, and T, respectively.

Suppose that R>>S and T, then

$$\int_{t_o}^{T} \left\{\left|\delta\alpha_2(t)\right|\right\} dt \sim \frac{O(R)}{(M+1)} \left\{ \begin{array}{c} (r+1) \\ \overline{r} \\ r \end{array} \right\} \tag{316}$$

In the extreme case when there is one control variable of O(R) and several of O(S) and O(T), after the descent is complete

$$\left\{ \begin{array}{c} \overline{\Delta P}_r(\infty) \\ \overline{\Delta P}_s(\infty) \\ \overline{\Delta F}_t(\infty) \end{array} \right\} \sim \left\{ \begin{array}{c} 2 \\ \overline{1} \\ 1 \end{array} \right\} \tag{317}$$

87

if the descent is monotonic.

In the absence of a weighting function in the same example, from Equation (308)

$$\left\{\delta\alpha_2(t)\right\} \sim \left\{s^{\phi}_{\alpha\psi}(t)\right\}$$

(318)

On integrating

$$\int_{t_o}^{T} \left\{ |\delta\alpha_2(t)| \right\} dt \sim \left\{s^{\phi}\right\} = \left\{ \begin{array}{c} O(R) \\ \hline O(S) \\ \hline O(T) \end{array} \right\}$$

(319)

so that the weaker control variables would be practically unperturbed in each descent provided the assumption of R>>S, T is retained. On summing over the entire descent, it follows that the total perturbation in the weak control variables will be negligible compared to those in the powerful control variables.

Weighting matrices based on the integrated instantaneous payoff function sensitivities S^{ϕ}_{α} act in a similar manner by emphasizing the first term of $\delta\alpha_2$, Equation (303), instead of the complete expression. It is difficult to arrive at a quantitative result similar to that of Equations (317) and (303) for this type of weighting matrix. For the present it must suffice to mention that several cases of false convergence in the weaker control variables have been eliminated by the use of this type of weighting function.

Weighting Function Based on Instantaneous Sensitivity

Suppose there is a single control variable α_1, and that the power of this control variable varies drastically along the trajectory. In region A of the trajectory, let the power of α_1 be several orders of magnitude greater than in region B. The greater perturbation will tend to appear in region A and, should the discrepancy in control power persist throughout the steepest descent convergence, the greater total perturbation will always occur in region A. However, the total perturbation required in regions A and B are functions of the nominal control variable histories created of the region of interest and the problem at hand. Therefore, once more a false convergence can occur.

To be more specific, let region A be that region in which $t_o \leq t \leq t'$ and region B be that region in which $t' \leq t \leq T$.

Let the power of the control variable in region A be $O(P)$ greater than that in region B. In the absence of a weighting function the perturbation mode which improves the payoff function directly is proportional to the mixed sensitivities payoff function sensitivity, $s_{\alpha\psi}^{\phi}$. Therefore,

$$\delta\alpha_j(t)_{t<t'} \sim O(P)\ \delta\alpha_j(t)_{t>t'} \tag{320}$$

where $\delta\alpha_j(t)$ is the perturbation at any point in the jth cycle of the descent.

Following the approach used previously, a W matrix can be used which will tend to equalize these perturbations. For example,

$$\left[W_{11}\right]^{-1} = \left[1 + \frac{s_{\alpha\psi\ max}^{\phi}}{s_{\alpha\psi}^{\phi}(t)}\right] \tag{321}$$

where $s_{\alpha\psi\ max}^{\phi}$ is the largest value of $s_{\alpha\psi}^{\phi}$ along the trajectory (in practice use the maximum value of $s_{\alpha\psi}^{\phi}$ from the preceding descent). In this case

$$\delta\alpha(t) \sim \left(1 + \frac{s_{\alpha\psi\ max}^{\phi}}{s_{\alpha\psi}^{\phi}(t)}\right) s_{\alpha\psi}^{\phi}(t) \tag{322}$$

Let $s_{\alpha\psi\ max}^{\phi}$ be $O(P)$ and let $s_{\alpha\psi\ min}^{\phi}$ be $O(Q)$.

At the point of greatest power

$$\delta\alpha(t)_{max} \sim (2)(O(P)) \tag{323}$$

and at the point of weakest power

$$\delta\alpha(t)_{min} \sim \left(1 + \frac{O(P)}{O(Q)}\right) O(Q) = O(Q) + O(P) \tag{324}$$

If $P \gg Q$ we, therefore, obtain

$$\frac{\delta\alpha_{max}}{\delta\alpha_{min}} \sim 2 \tag{325}$$

Without the weighting function

$$\frac{\delta\alpha\ max}{\delta\alpha\ min} \sim O(P-Q) \tag{326}$$

Therefore, the solution will be limited to extremely small control variable perturbations in region A, unless a weighting function is used to alleviate the discrepancy in control variable power in the two regions.

Combined Weighting Functions

In general, there may be several control variables whose individual sensitivities vary drastically both with respect to the independent variable and with regard to each other at any instant. The variation with the independent variable may be modulated by using the inverse W matrix which will equalize the total sensitivity, i.e., the sum of the ual control variable sensitivities, at each point along the trajectory. A time varying term of the form

$$\left[A_{ii} + \frac{B_{ii} \left(\sum_{j=1}^{M} s_{\alpha\psi_j}^{\phi} \right)_{max}}{\sum_{j=1}^{M} s_{\alpha\psi_j}^{\phi}} \right] \tag{327}$$

will achieve this effect.

The difference in sensitivity between the control variables at any instant may be equalized by utilizing a term similar to Equation (311) with instantaneous sensitivities replacing the integrated sensitivities

$$\left[C_{ii} + D_{ii} \frac{\sum_{j=1}^{M} s_{\alpha\psi_j}^{\phi}}{s_{\alpha\psi_1}^{\phi}} \right] \tag{328}$$

Combining Equation (327) and (328) and adjusting the matrix so that with equally powerful control variables throughout the trajectory, the unit matrix is obtained with $A_{ii} = B_{ii} = C_{ii} = D_{ii} = 1$, the inverse weighting matrix becomes

$$\left[W_{ii} \right]^{-1} = \frac{\left[A_{ii} + \frac{B_{ii} \left(\sum_{j=1}^{M} s_{\alpha\psi_j}^{\phi} \right)_{max}}{\sum_{j=1}^{M} s_{\alpha\psi_j}^{\phi}} \right] \left[C_{ii} + D_{ii} \frac{\sum_{j=1}^{M} s_{\alpha\psi_j}^{\phi}}{s_{\alpha\psi_1}^{\phi}} \right]}{2 \ (M + 1)} \tag{329}$$

Equation (329) is the most general weighting function obtained to date; it has been utilized successfully with both the mixed sensitivities, $s_{\alpha\psi}^{\phi}$, and with the sensitivities, s_{α}^{ϕ}. The first term in Equation (329) tends to equalize differences in total sensitivity at different instants. The second term tends to distribute the total perturbation at an instant more equally between all control variables.

90

STEEPEST-DESCENT STEP-SIZE CRITERIA

Application of the Steepest-Descent Method is reduced to a routine computation once an automatic scheme or control system for determining step-size has been devised. The present section describes a step-size control system which has consistently produced convergence in the calculation of several hundred diverse atmospheric trajectory optimization problems in the period since 1963. The control system is an integral part of the basic computer program described previously in this report. This program is limited to trajectories having all stages, except possibly the last, terminated by a fixed value of independent variable, time. The major principles of the control system apply to the more general problem of optimal staging described in the multi-stage analysis section of the present report.

A control system is basically a set of tests, mainly of a logical nature, which determine a suitable step-size and, hence, a control variable perturbation for each successive iteration. It can be seen from Equations (35b) and (43) that a step-size is determined when the control variable perturbation magnitude DP^2, the amount of each end point error to be eliminated, $d\psi_i$, and the initial state variable value perturbations, $\delta x(t_o)$ are specified. No further mention of the initial state variable value perturbations will be made, as this type of perturbation has not been included in the basic optimization program.

It might be thought with the solution of Equation 43 available that the choice of step-size is a trivial problem. This judgment would be false, for the solution of Equation 43 merely provides the optimal perturbation for a very small step: that is, it is a linearized solution only. In converging to an optimal trajectory, the analyst seeking to keep computer time expenditure within reasonable bounds will require the use of as large a step as possible. These large steps inevitably incur significant differences between linearized predictions of system behavior and the actual non-linear system behavior in the presence of large perturbations. This can be simply illustrated by considering the behavior of any one of the optimization functions, f, the payoff function or any of the terminal constraints. For an infinitesimal perturbation of the control variable, the linear prediction of the change in the function based on the adjoint method and the non-linear behavior, as determined by an actual trajectory integration will be identical and

$$df(\epsilon) = \Delta f(\epsilon) \qquad (330)$$

where df is the linear prediction and Δf is the actual non-linear change. Both are functions of step-size, which is assumed small and denoted by ϵ.

As the step-size is increased, the difference between df and Δf will tend to grow progressively. For a large enough step-size the linear prediction may cease to provide a reasonable guide to the actual effect of the perturbation on the optimization function and may, in fact, even be of the wrong sense. This behavior is illustrated in Figure 22 for a payoff function being maximized.
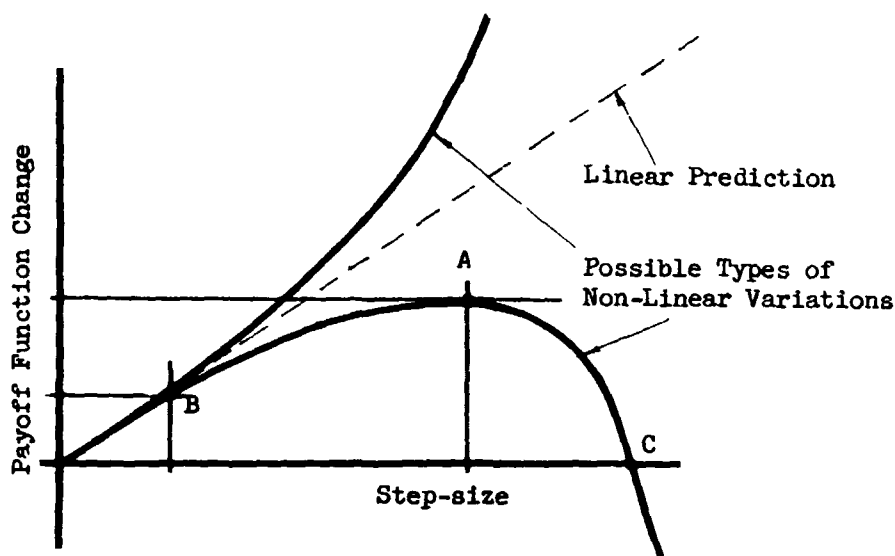

Figure 22.— Payoff Function Behavior Against Step-Size

Considering non-linear variations of the type denoted by the lower of the two solid lines in Figure 22, the best step-size to use is that resulting in the maximum payoff function change (Point A),provided this is the only function to be considered.

Since the linear prediction is available from the adjoint equations but all non-linear changes require integration of the trajectory equations, it is desirable to determine the maximum change in payoff function with the least computational effort. At Point A there is a considerable discrepancy between linear prediction of and actual change in the optimization function. Limitation of step size to a point such as B, where the linear and non-linear changes are in close agreement, results in significantly smaller step-size. If the payoff function was the only optimization function to be considered, point A would represent a reasonable upper limit on step-size, for any greater step-size would reduce the payoff function change. If step sizes greater than that denoted by C were to be taken, the payoff function would actually be degraded rather than improved.

Figure 22 illustrates one of the major problems confronting the analyst seeking to define a reasonable step size. Too small a step, as at B, will produce an extremely well behaved convergence, but an excessive expenditure of computer time results. Too large a step, one significantly greater than that of A, results in
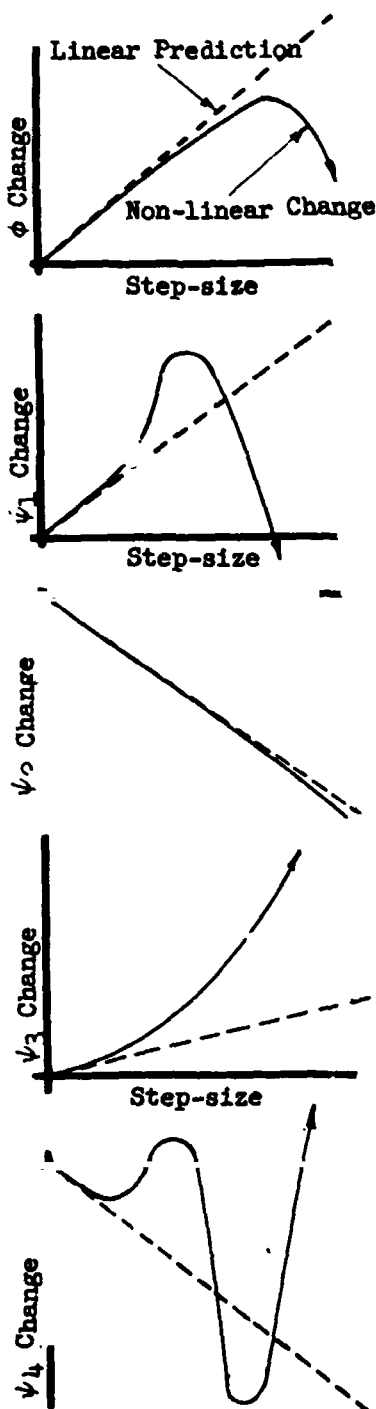
Figure 23.- Possible
Optimization Function
Behavior

the linear prediction becoming
meaningless,and convergence, if it
does occur, tends to be erratic.
Again, this tends to cause excessive
computer time expenditures. The
problem is further complicated by
the fact that usually, instead of the
single optimization function of
Figure 22, the behavior of several
functions must be considered. For
example, with a payoff function and
four constraints, a typical problem
may encounter the changes illustrated
in Figure 23. In such a case it is
not clear what step-size is reasonable.
In view of the wide variety of be-
havior exhibited by the optimization
functions, the need for a control
system which will make consistent
and logical choices of step-size.
becomes apparent.

In this section, it is assumed
that a single parameter can be chosen
to define the step-size and that for
small perturbations the predicted
changes in optimization functions
will vary linearly with this parameter.
For example, choose the parameter as
follows: let nominal values be available
for DP^2 and $\{d\psi\}$ and let these values be
denoted by DP_o^2 and $\{d\psi_o\}$. Now take a
parameter k to determine step-size in
the manner

$$DP^2 = k^2 \, DP_o{}^2 \qquad (331)$$

$$\{d\psi\} = k \, \{d\psi_o\} \qquad (332)$$

If $\{\delta\alpha_o\}$ is the control variable his-
tory generated by the nominal choice
of step-size parameter k=1, then from
Equations (43) and (44)

$$\{\delta\alpha_k\} = k \, \{\delta\alpha_o\} \qquad (333)$$

and

$$d\phi_k = k \, d\phi_o \qquad (334)$$

It follows that the perturbations are
linear with the parameter k, as was
desired.

93

Control System Philosophy

There are two philosophies which may be followed in most complex decision-making situations. A person may attempt to reach a conclusion directly by asking what is the correct course to follcw, or indirectly by asking which courses are not to be followed. The steepest-descent step-size control system follows the latter course. The direct approach may at first sight appear the more attractive method; however, it should be borne in mind that it is usually easier tc determine which courses of action should not be followed than it is to determine the particular course of action which should be followed.

Major problems involved in the design of a step-size control system are failure to converge and false convergence. The first type of failure is immediately apparent, but the latter may be difficult to detect. For example, suppose that a case involves a single constraint which, after the first M iterations, has effectively met the desired terminal value. If the constraint value is not permitted to drift away from the desired value subsequent perturbations will be small by virtue of problem non-linearity. In a severe case, this will result in behavior easily mistaken for convergence. On the other hand, by permitting the constraints to drift off the desired value by means of an indirect test, this difficulty may be avoided; this type of behavior is illustrated by Figure 24.

In view of the above and similar types of phenomena, the step-size control system has been constructed as a group of very loose tests, in the sense that a set of almost obvious decisions as to step-size magnitude lead indirectly to a choice of ⁻ ꞓp.

Basic Control System Principles

Second Order System Behavior.— Each iteration commences with a trial trajectory; this trajectory is distinguished from the final trajectory of an iteration in that the computation of the partial derivative matrices, F and G, is omitted. The computation of these matrices usually requires somewhat more computer time than the trajectory integration itself.

This trial trajectory is defined by a step-size parameter, k, where k=0 denotes the previous final trajectory and k=1 denotes the step-size magnitude used to obtain the trial trajectory.

On completing the trial trajectory, the non-linearities of the payoff function and constraints are computed. These are non-dimensional measures of the difference between the actual and linear predictions of the change in these functions. The payoff function non-linearity is defined below.
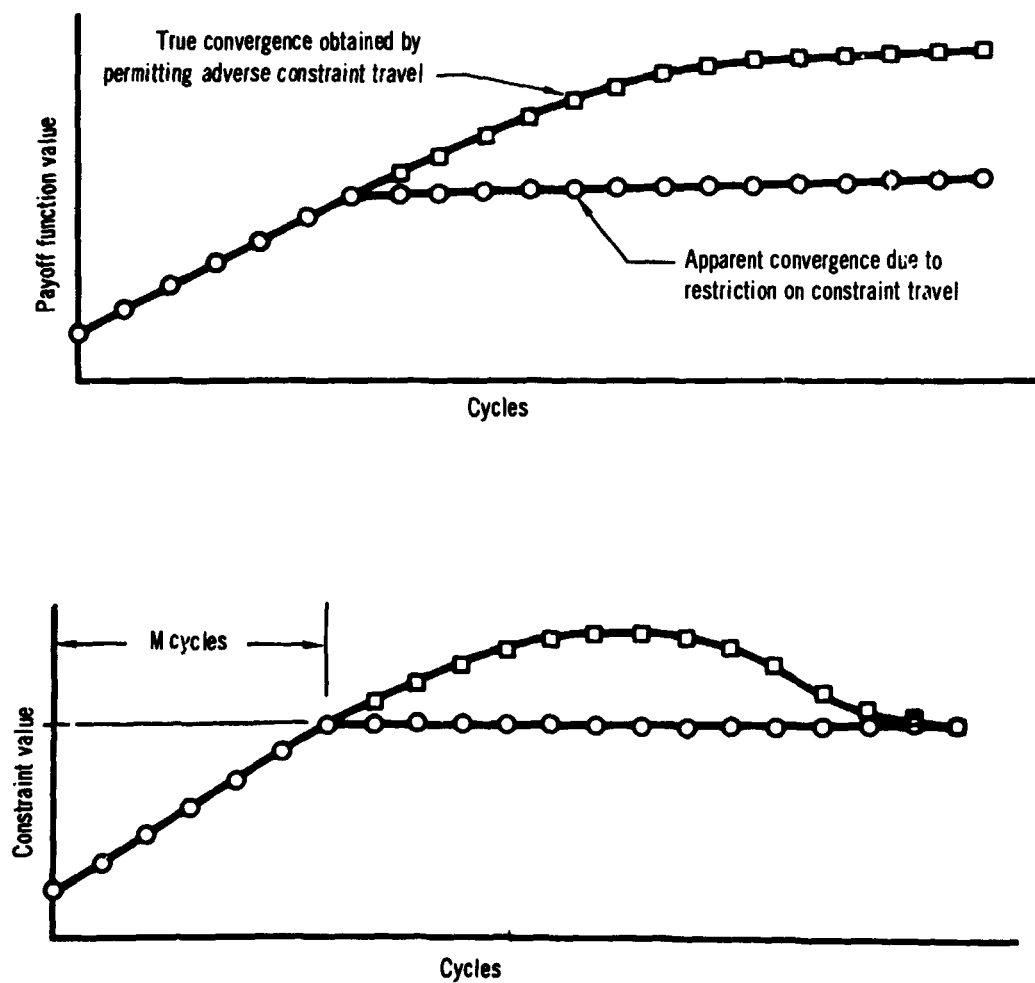
94

Figure 24.—An Example of False Convergence

$$\phi_{NL} = \left| \frac{\Delta\phi - d\phi}{d\phi} \right| \tag{335}$$

and the constraint non-linearities are defined as

$$\left\{ \psi_{NL} \right\} = \left\{ \left| \frac{\Delta\psi - d\psi}{d\psi} \right| \right\} \tag{336}$$

Here $d\phi$ and $d\psi$ denote the linear predicted change in ϕ and ψ, and $\Delta\phi$ and $\Delta\psi$ denote the actual change between the previous final trajectory and the present trial trajectory.

From the previous discussion for a reasonable step in any of these variables, the corresponding non-linearity must be neither too small nor too great.

On completing the trial trajectory, an approximation to the actual non-linear variation of the optimization functions with step-size parameter k, can be obtained from Taylor's Theorem by making the assumption that the behavior of each function is parabolic. The three conditions defining each of the parabolic variations are as follows:

$$k = 0; \quad \Delta\phi \ , \left\{ \Delta\psi \right\} = 0 \tag{337}$$

$$k = 1; \quad \Delta\phi = \Delta\phi_0, \left\{ \Delta\psi \right\} = \left\{ \Delta\psi \right\}_0 \tag{338}$$

$$k = 0; \quad \frac{d(\Delta\phi)}{dk} = \frac{d(d\phi)}{dk} \ , \left\{ \frac{d(\Delta\psi)}{dk} \right\} = \left\{ \frac{d(d\psi)}{dk} \right\} \tag{339}$$

The last of these equations follows from Equation (330). Equations (337) and (338) define two points on a parabola; Equation (339) equates the predicted linear slope at the first point to the parabolic slope at that point. Applying these conditions, the approximate non-linear variations are obtained as functions of k.

$$\Delta\phi(k) = \left(\Delta\phi_0 - d\phi_0 \right) k^2 + d\phi_0 \cdot k \tag{340}$$

and

$$\left\{ \Delta\psi(k) \right\} = \left\{ \left(\Delta\psi_0 - d\psi_0 \right) k^2 + d\psi_0 \, k \right\} \tag{341}$$

Now, the value of k which will provide a specified non-linearity in the payoff or constraint functions can be found. Substituting Equations (340) and (341) into Equations (335) and (336)

$$\left\{ k_\psi \right\} = \left\{ \frac{\psi_{NL}}{\psi_{NL_0}} \right\} \tag{342}$$

and

$$k\phi = \frac{\phi_{NL}}{\phi_{NL_o}}$$

(343)

That is, the desired value of k for each quantity is the desired value of its non-linearity divided by its non-linearity on the trial trajectory. A reasonable value for the non-linearity desired can be obtained from the geometry of a parabolic variation. Consider any of the parabolic approximations to the optimization function f, as shown in Figure 25.
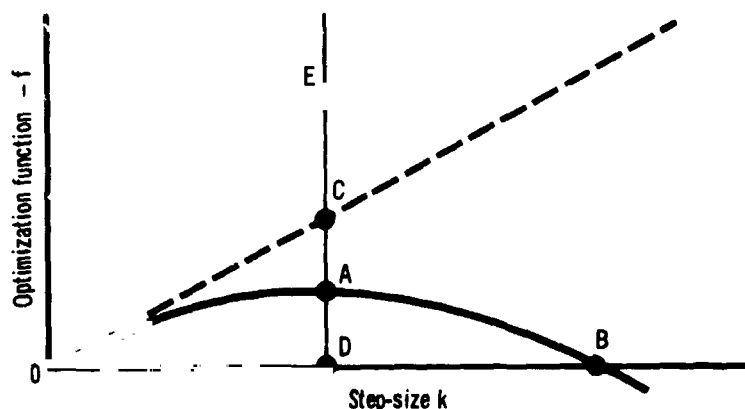

Figure .5.—Parabolic Variations

For a curve such as OAB, the maximum gain in the function occurs at A,and if OAB is parabolic, the non-linearity is

$$f_{NL} = \frac{DA - DC}{DC} = .5$$

(344)

Accordingly, a reasonable non-linearity for the payoff function allowing for the approximation involved is about .45. For the constraints a more conservative value of .3 can be employed. If the curve is of the nature of OE, these values still provide a reasonable step-size guide. With these assumptions, the step-size parameter which gives the desired non-linearity for each function by use of Equations (342) and (343) can be computed. A basic principle of the step-size control system is to base the step-size on the optimization function having the largest k. If all the desired non-linearities were equal, this would be equivalent to controlling step-size with the function exhibiting the most linear behavior. Additional trial trajectories are made when the resulting k<.5 or >2, due to the increased possibility of the parabolic assumption being in error if it is either extrapolated or interpolated too far. For example, consider Figure 26 where an interpolation from a trial value causes a reduction in f(k) rather than an increase. Similarly, in Figure 27, an extrapolation has the same effect. It may be noted from Equations (332) and (334) that the trial trajectory corresponds to a k=1.

97

If the largest computed k is less than .5 then take another trial trajectory with k=.5 and repeat the above logic. Similarly, if k is greater than 2, a trial with k=2 is indicated; however, before making such a trial, the control system proceeds to various other tests which may reduce the value of k; these tests will be described later.
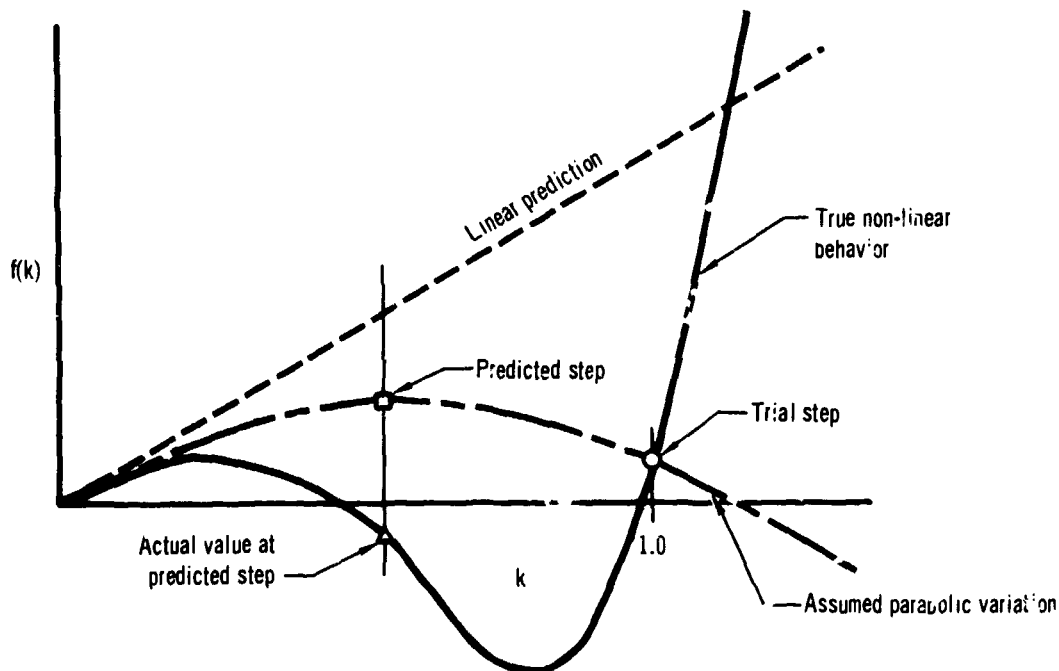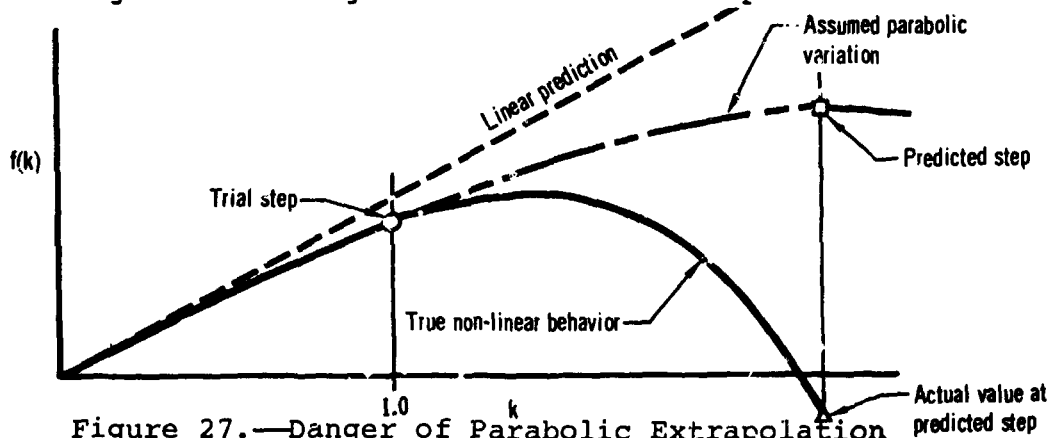

Figure 26.—Danger of Parabolic Interpolation

Figure 27.—Danger of Parabolic Extrapolation

To summarize these tests: their purpose is merely to assure that at least one of the optimization functions is reasonably linear, a modest requirement for a reasonable perturbation. The use of non-linearity in the above manner is the first basic principle of the control system.

The second basic principle is that of correcting constraint errors gradually. There are several reasons for eliminating constraint errors by a small amount on each iteration, rather than by attempting to eliminate the entire error in the first iteration.

First, working with non linear equations, the large steps which are required to eliminate the entire constraint error will frequently lie outside the linear range. Hence, after a set of time consuming trials of decreasing step-size, the analyst will ultimately be reduced to the gradual elimination of the errors.

Second, it should be noticed from Equation (43) that of the control variable perturbation magnitude DP^2 an amount equal to $\lfloor d\psi \rfloor \lfloor I_{\psi\psi} \rfloor^{-1} \{d_\psi\}$ is required to provide the desired constraint changes. If this portion of DP^2 is too large, the payoff function Equation (44) will be primarily the result of constraint changes rather than an inherent improvement in the trajectory characteristics. In this case, there is a danger that the optimization will degenerate into a more terminal constraint search.

Third, it must be noted that it is possible to introduce local extremals into a problem by the means of terminal constraints. This becomes clear from an elementary example in the ordinary calculus. Consider the problem of maximizing a function $z(x,y)$ which has a single optimal value as in Figure 28.

Now seek extremal values of $z(x,y)$ subject to a constraint

$$g(x,y,z) = 0 \qquad (345)$$

It is clear from diagram that in the particular case considered, there are two solutions: one at A and one at B, the global optimum being that at B. Now consider the solution of this problem by the method of steepest descent commencing from Point C. If achieving the constraint is the dominating influence in choosing a step, the solution will tend to traverse a path of the nature CDA, and, hence, the lower extremal solution will be located. On the other hand, if in initially choosing the step-size, one pays little attention to the constraint, the likelihood of traversing a path such as CEB and locating the global extremal is increased.

From this discussion, it is apparent that there are sound reasons for not attempting to eliminate the complete end point error at each step; accordingly the control system initially attempts to remove constraint errors of magnitude

$$\{d\psi\} = \Delta C_\psi \cdot \{\psi\} \qquad (346)$$

where ΔC_ψ is a small non-dimensional quantity.

99

After N iterations, if certain requirements are met, the control system will be attempting to eliminate an error of

$$\{d\psi\} = N \cdot \Delta C_\psi \cdot \{\psi\} = C_\psi \{\psi\} \qquad (347)$$

provided $N \cdot \Delta C_\psi \leqslant 1$. When $N \cdot \Delta C_\psi \geqslant 1$, the amount of constraint error to be removed is given by
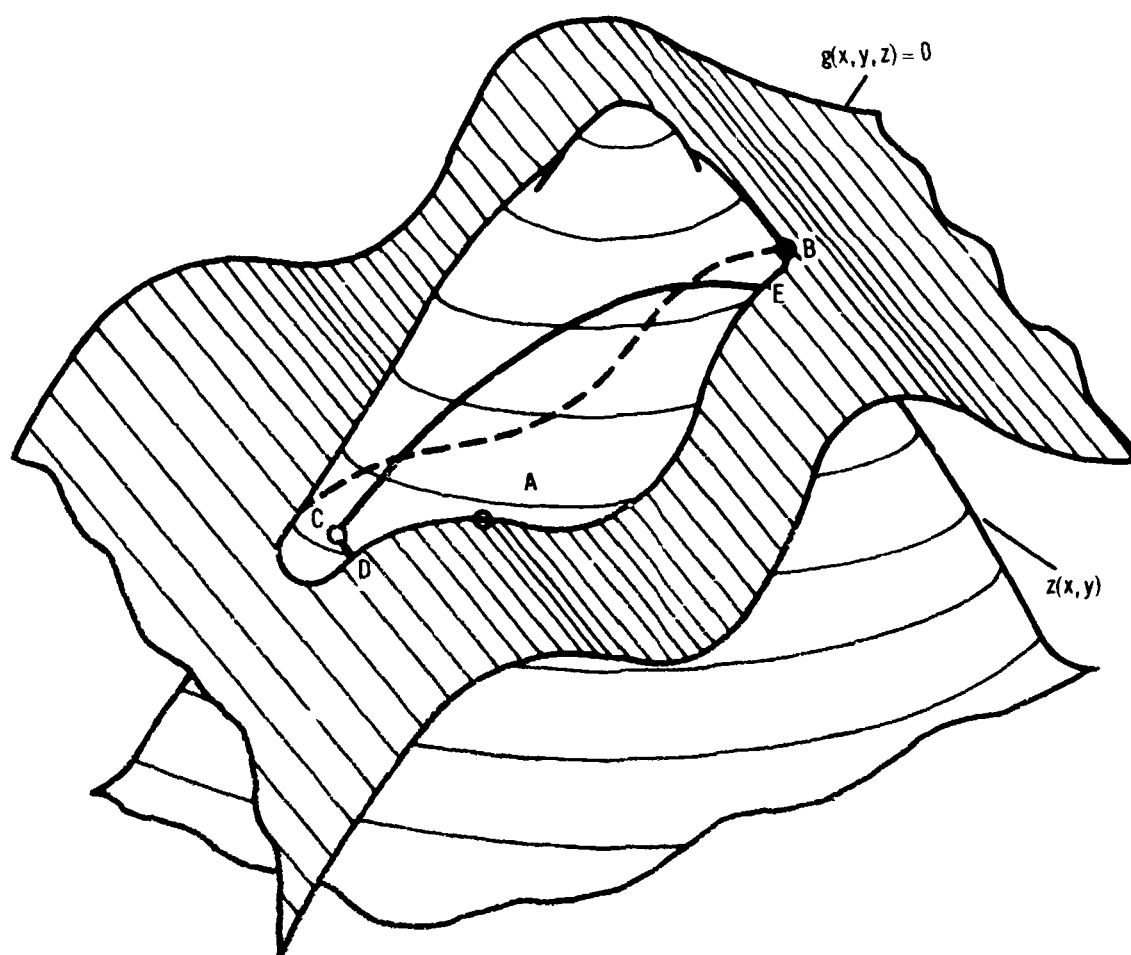
$$\{d\psi\} = \{\psi\} \qquad (348)$$

Figure 28.—Local Extremal Introduced by Constraint Function

This is the second basic principle of the step-size control system, the gradual removal of constraint errors in order to emphasize the payoff function role, in the initial iterations of a descent.

The two principles of this section are not adequate to insure convergence. It has been necessary to add many other logical tests to the control system; some of these are described below.

<center>Secondary Tests</center>

The two principles outlined above are far from sufficient to insure convergence to the correct solution. They must be supplemented by many secondary decisions, mainly of an indirect nature. The more important ones will now be listed, not necessarily in the order in which they occur in the actual control system. For details of the control system as programmed, reference must be made to the appropriate flow charts in Reference 2.

Determination of Step-Size Magnitude for First Trial of Each Iteration.— The step-size magnitude $DP_O{}^2$ used in the first trial of each iteration except the first is automatically based on the values used in preceding iterations. For the first iteration, an arbitrary value must be specified by the analyst; this value should be chosen on the large size; the control system will very quickly determine the correct value by making trial trajectories.

After each iteration in an attempt to inhibit any tendency to a gradual decrease in DP^2, the control system determines the trial value from the value finally used on the previous iteration, DP^2_{N-1} using the expression

$$DP_O{}^2 = 2DP^2_{N-1} \qquad (349)$$

provided certain other conditions have been met. These other conditions are as follows:

(a) That the value given by Equation (349) is at least great enough to provide the constraint change being sought. If it is not, then,

$$DP^2 = \text{Min}\left(\left\lfloor d\psi \right\rfloor \left[I_{\psi\psi} \right]^{-1} \left\{ d\psi \right\}, \ \text{Max}\left(10^{-4}T, \ DP_O{}^2 \right) \right) \qquad (350)$$

That is, DP^2 is taken as the minimum value which will provide the constraint change desired, provided that this is not greater than 10^{-4} times the expected trajectory cut-off, unless this in turn is a smaller quantity than that given by Equation (349). The reasoning behind this test is as follows:

<center>101</center>

(i)   Assume first that the rules for con-
      straint changes, to be described in detail
      later, are functioning correctly, and,
      hence, DP^2 should be no less than the
      quantity $\lfloor d\psi \rfloor \lfloor I_{\psi\psi} \rfloor^{-1} \{d\psi\}$.

(ii) In some cases, the constraint changes
      may be excessive; in this case as $10^{-4}T$
      should insure a reasonable size of per-
      turbation, the control uses this value
      as an upper limit.

(iii) If, however, the previous iteration used
      a perturbation magnitude DP^2_{N-1} greater than
      $10^{-4}T$, then the control system uses DP^2_{N-1}
      as the upper limit on step-size magni-
      tude.

(b) The quantity,

$$\text{grad}\,\phi = I_{\phi\phi} - \lfloor I_{\psi\phi} \rfloor \left[I_{\psi\psi} \right]^{-1} \{ I_{\psi\phi} \} \tag{351}$$

is the gradient of ϕ with respect to DP^2 if the
constraint changes are zero; that is, it is the
measure of how close any trajectory is to the
optimal trajectory having the same end-points.
Now grad ϕ is usually the difference of two very
large numbers, and these numbers are the result
of lengthy numerical computations. In this
situation, small numerical errors can lead to
the difference between positive and negative re-
sults for the value of gradϕ when a trajectory
approaches the optimal trajectory for a particular
set of end-points. As these may not be the
desired set of end-points and as gradϕ is essen-
tially a positive quantity (see Equation 44), it
must be recognized that negative values of gradϕ
merely mean that a trajectory is practically the
optimal one to the current end-points. All that
remains in such a situation is to perturb the end
points towards their desired values. This is
accomplished by setting

$$DP^2 = \lfloor d\psi \rfloor \left[I_{\psi\psi} \right]^{-1} \{ d\psi \}, \; DP_0^2 \leq \lfloor d\psi \rfloor \left[I_{\psi\psi} \right]^{-1} \{ d\psi \} \tag{352}$$

or by following the logic of (a) above if the
inequality is not satisfied.

(c) On occasion, an idiosyncrasy in a particular
     trajectory may cause the step-size to become
     severely reduced; this will usually be accompanied
     by an excessive number of trials. After six or

102

eight trials, depending on the circumstance,
the control system will force a final trajectory
to be computed. In the next iteration, the
magnitude of the control variable perturbation
DP^2 for the first trial trajectory will then be
computed by the expression

$$DP_o^2 = \sqrt{(DP_{N-1}^2)(DP_{N-2}^2)} \qquad (353)$$

instead of by Equation (349). This value is
used in an attempt to maintain a reasonable
perturbation magnitude should an excessive
number of trials occur.

(d) If sense switch 4 on the IBM 7094 computer
console is depressed, the step-size magnitude
for the first trial of an iteration is auto-
matically given by

$$DP_o^2 = 10^{-4}T \qquad (354)$$

This feature can be used towards the termination
of a calculation as convergence appears complete,
to artificially attempt a large step. It is a
safety device designed to provide assurance that
a false convergence has not occurred; for by
this means, a complete family of step-sizes will
be attempted as the step-size is reduced to an
acceptable level.

(e) Two limits are placed on a trajectory; a time
limit at

$$t = T_{max} \qquad (355)$$

and an altitude limit at

$$h = h_{min} \qquad (356)$$

If time is integrated beyond T_{max}, or the alti-
tude below h_{min}, then the trajectory integration
is terminated. Should either of these situations
occur on a nominal trajectory, the control system
assumes that the nominal trajectory is unsatis-
factory and terminates the convergence immediately.
If these limits are violated on any trajectory
subsequent to the nominal trajectory computation,
it is assumed that this is the result of a per-
turbation being too great, and another trajectory
is computed with

$$k = .5 \qquad (357)$$

103

at this point, the value of DP^2 which caused the violation of a trajectory limit is saved (DP^2_{save}) and is then used as the nominal value on the first trial of the next iteration; for the violation of the trajectory limit may be of a transient nature, and a major object of the control system is to keep the step-size as large as possible.

Determination of Step-Size Magnitude After First Trial.— After the first trial, the step-size magnitude is basically controlled by the step-size parameter k, according to the expression,

$$DP^2 = k^2 . DP^2_o \qquad (358)$$

There is an exception to this rule when the step-size is "bouncing". "Bouncing" means that either a value of DP^2 equal to or smaller than one already demonstrated to be too small or a value of DP^2 equal to or greater than one already demonstrated to be too great is again predicted. Figure 29 demonstrates one way this phenomena can arise. Here a value of DP^2 has been computed from a value of the step-size parameter k_{low}; a trial is made and the extrapolation is made and a value of k corresponding to point C is computed. If this value of k is beyond the point
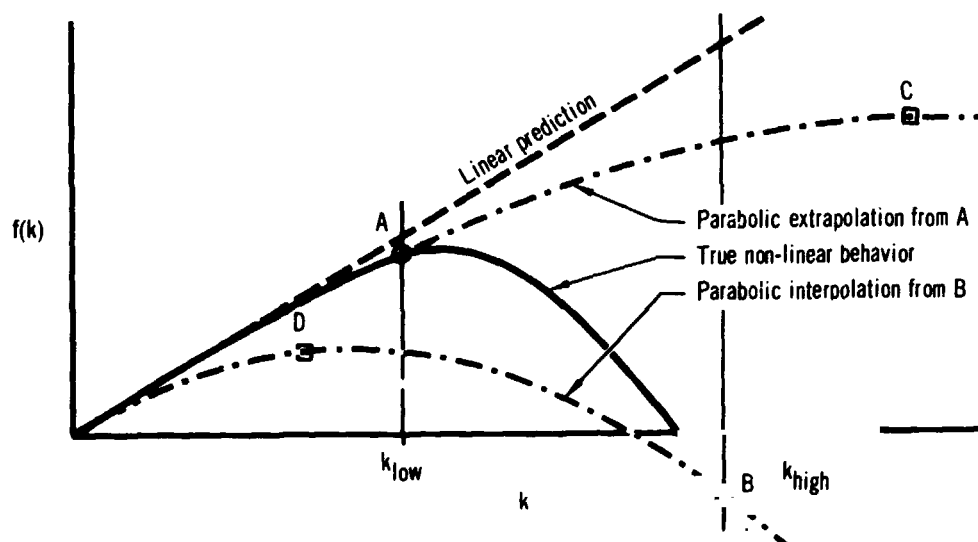
$$k_{high} = 2k_{low} \qquad (359)$$

Figure 29.—Step-Size Bounce Induced by Parabolic Approximation

104

The control system will compute a new trial trajectory corresponding to a step-size of k_{high}, i.e.,

$$k = k_{high} \tag{360}$$

On completing the trial, the controlling function takes on the value at B. The parabolic interpolation from this point predicts a value at D less than k_{low} and without a "bounce test," a trial would be taken with

$$k = k_{low} \tag{361}$$

and a closed loop established. Accordingly, if a situation of this nature arises, Equation (360) is overruled, and the step-size magnitude is determined by a midpoint search

$$DP^2 = \frac{DP^2_{high} + DP^2_{low}}{2} \tag{362}$$

Limits on Dimensional Travel of Payoff Function.—The step-size parameter k is determined by the first principle described previously, that is control with the most linear of the optimization functions. This decision is over-ruled if such a step causes the dimensional travel of any of the optimization functions to become excessive. Constraints are placed on the dimensional travel of the payoff function in the following manner:

(a) If the problem at hand is one involving maximization of the payoff function and

$$\Delta\phi \geq -\phi_{ADV} = \Phi_{ADV} \tag{363}$$

(b) If the problem at hand is one involving minimization of the payoff function and

$$\Delta\phi \leq \phi_{ADV} = \Phi_{ADV} \tag{364}$$

The permissable adverse ϕ travel magnitude, ϕ_{ADV}, is determined by the expression

$$\phi_{ADV} = Max\left(\frac{\phi_{N-1}}{10}, \frac{\phi_{max}}{50}\right) \tag{365}$$

where ϕ_{N-1} is the value of the payoff function at the termination of the last iteration, and ϕ_{max} is the greatest value of the payoff functio absolute value obtained at the termination of any of the previous iterations.

The adverse ϕ travel test described above has its basis in the principle of emphasizing the payoff function behavior.

105

Problems are often encountered in which, due to the initial terminal constraint errors, the unconstrained performance, as measured by the payoff function, is better on the nominal trajectory than it is on the final optimal trajectory. A problem of this nature inevitably involves the loss of performance during the major portion of the descent. Now the greatest obstacles facing the analyst in applying the Steepest-Descent method are false convergence and failure to converge in a reasonable number of cycles. Both these phenomena are inhibited by the adverse ϕ travel test when performance has to be given up in order to achieve the end points; Fig 30 demonstrates how the test inhibits false convergence in a problem of this type. Without the adverse ϕ travel tests, the first M iterations are spent in reducing the constraint error at the expense of ϕ. At that point (A) in the convergence, if all went well, emphasis would return to the payoff function and the optimal trajectory obtained at point B. This type of behavior is illustrated by the lines OAB. At point B, however, there is a risk of false convergence and the descent may continue in the manner of OAC. The adverse ϕ travel test, on the other hand, will not permit the initial rapid loss of ϕ and convergence with this test included is far more likely to be of the nature of the broken line OD.

Again, in a problem where the performance must tend to deteriorate as the constraints improve, a very irregular convergence may result. This is demonstrated in Figure 31; initially, a decline in performance occurs as the constraints are improved until the point A_1 is reached. At this point emphasis returns to the payoff function; a set of steps which improve performance at the expense of the constraint are undertaken until the point B_1 is reached. Here emphasis returns to the constraint and the process repeats. The resulting convergence tends to have the appearance of the lines $OA_1B_1A_2B_2$. . .; the adverse ϕ travel test inhibits this irregular behavior and tends to lead to a convergence of the nature of OC.

It should be noted that without the second part of the decision of Equation (365), ϕ would be unable to change sign; if the end points were attainable with $\phi = 0$, a false convergence such as OD would result. This provides a simple example of how an over-restrictive rule in the control system can lead to false convergence.

Whenever the ϕ travel fails to satisfy the appropriate inequality of Equations (363) or (364), the parabolic assumption is applied to compute a value of k that will result in a satisfactory step by solving the equation

$$(\Delta\phi - d\phi)k^2 + d\phi \cdot k = \Phi_{ADV} \qquad (366)$$

resulting in

$$k = \frac{-d\phi \pm \sqrt{(d\phi)^2 + 4(\Delta\phi - d\phi) \cdot \Phi_{ADV}}}{2(\Delta\phi - d\phi)} \qquad (367)$$
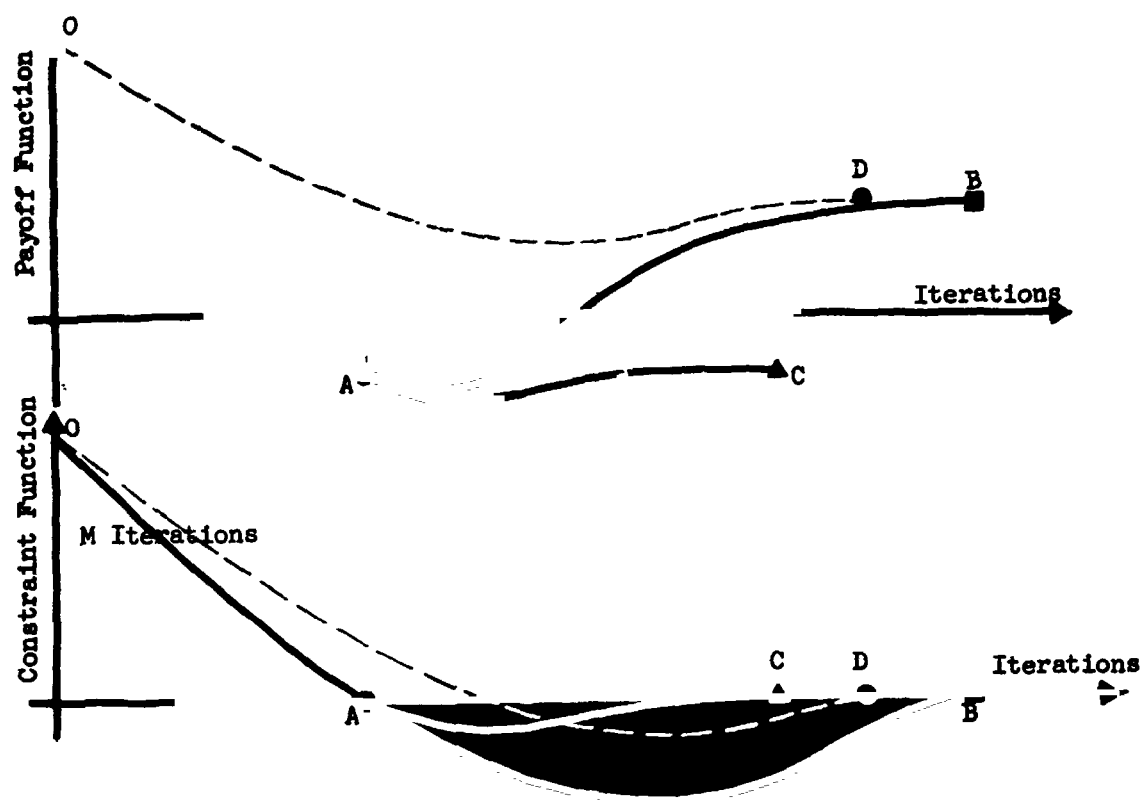
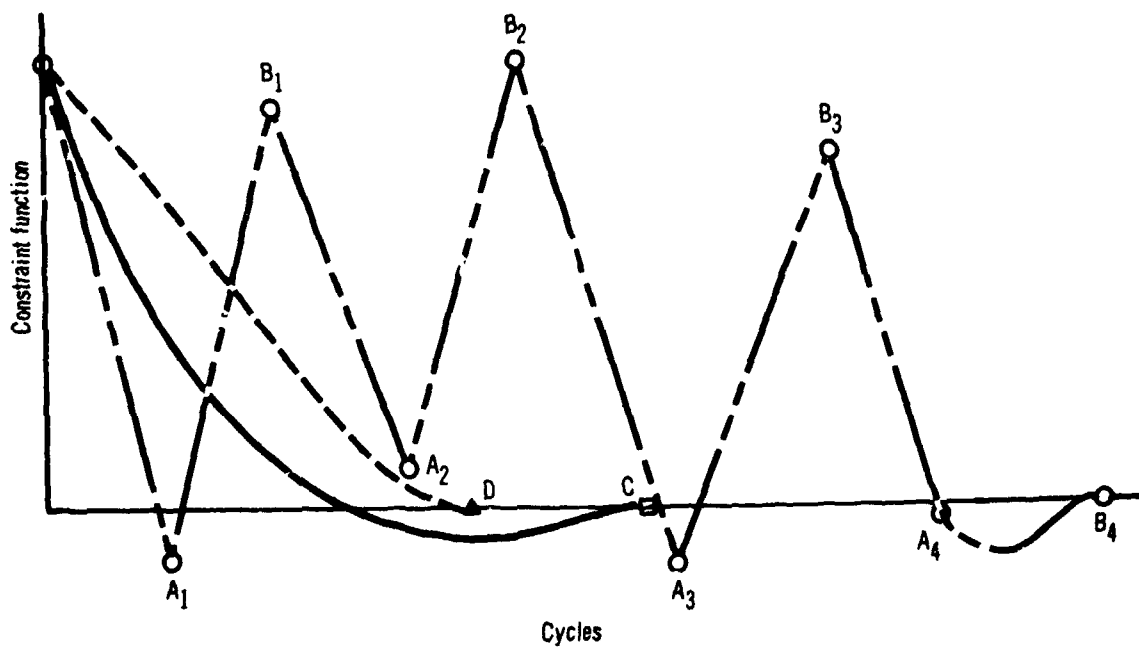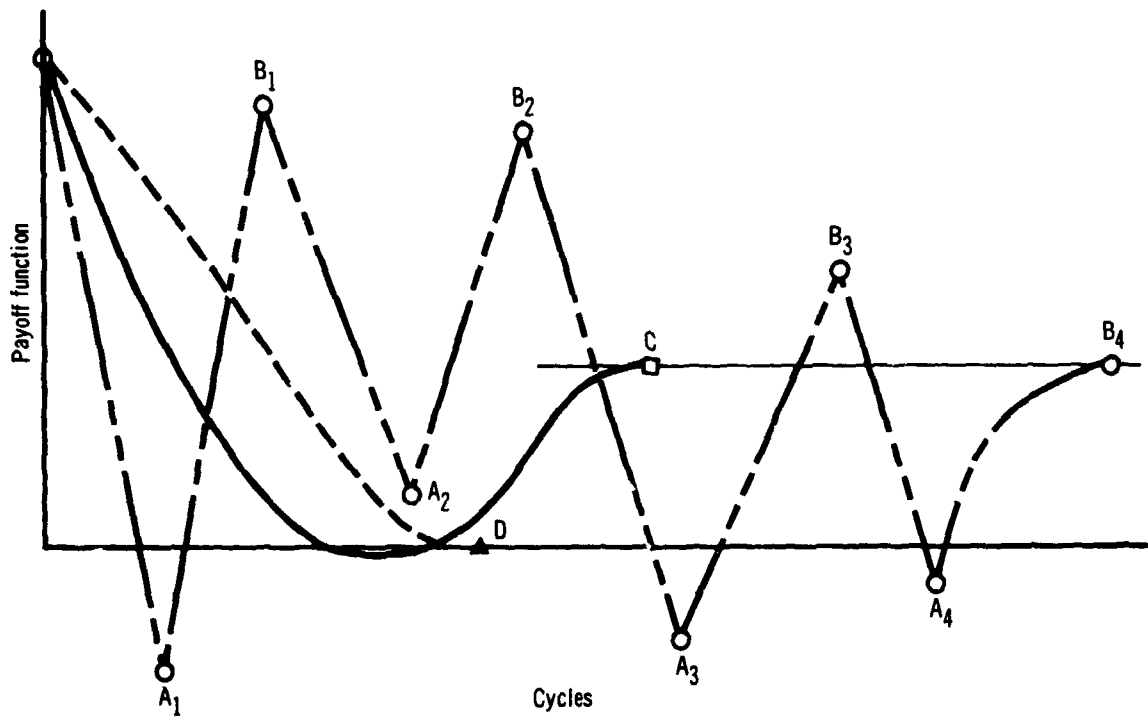Figure 30.—Adverse φ Travel Tests Inhibit False Convergence

Figure 31.—Adverse φ Travel Test Inhibits Irregular Convergence

The solution must have one positive and one negative root, provided the calculation is performed only when the adverse travel is too great, as can be seen from Figure 32. When k has been computed from Equation (367) it is multiplied by a factor of .9 in view of the approximations involved so that finally the acceptable value of k, based on adverse ϕ travel, is given by

$$k\phi_{TVL} = .45 \left(\frac{-d\phi \pm \sqrt{(d\phi)^2 + 4(\Delta\phi - d\phi)\Phi_{ADV}}}{\Delta\phi - d\phi} \right) \qquad (368)$$

Limits on Dimensional Travel of Constraints.— Rules which specify the amount of end-point error to be eliminated on each trajectory have been given previously. Due to the non-linear nature of the trajectory equations and the necessity of attempting to take large steps at each iteration, the actual constraint changes may differ considerably from those asked for. Accordingly another set of rules which specify acceptable limits on the end-point travel must be used; it has proven convenient to state these rules in the form

$$\psi_{BWD_i} d\psi_i \leq \Delta\psi_i \leq \psi_{FWD_i} d\psi_i , \quad \psi_i \leq 0 \qquad (369)$$

$$\psi_{BWD_i} d\psi_i \geq \Delta\psi_i \geq \psi_{FWD_i} d\psi_i , \quad \psi_i \geq 0 \qquad (370)$$

The permissable non-dimensional limits on adverse constraint travel, ψ_{BWD}, and favorable constraint travel, ψ_{FWD}, are functions of the amount of non-dimensional constraint error being eliminated, the number of iterations completed, and the number of iterations since the particular constraint error changed sign. If less than ten iterations have been completed since the constraint error changed sign then:

$$\psi_{BWD} = 1$$

$$\psi_{FWD} = 3 , \qquad\qquad c_\psi \leq .5 \qquad\qquad (371)$$

$$\psi_{BWD} = .5$$

$$\psi_{FWD} = 2.5 , \qquad\qquad .5 < c_\psi < 1 \qquad\qquad (372)$$

$$\psi_{BWD} = .025$$

$$\psi_{FWD} = 1.5 , \qquad\qquad c_\psi \geq 1 \qquad\qquad (374)$$

If more than ten iterations have elapsed since the constraint error changed sign, it is assumed that some difficulty in meeting the constraint exists. In this case, ψ_{FWD} and ψ_{BWD} are based on the number of completed iterations.
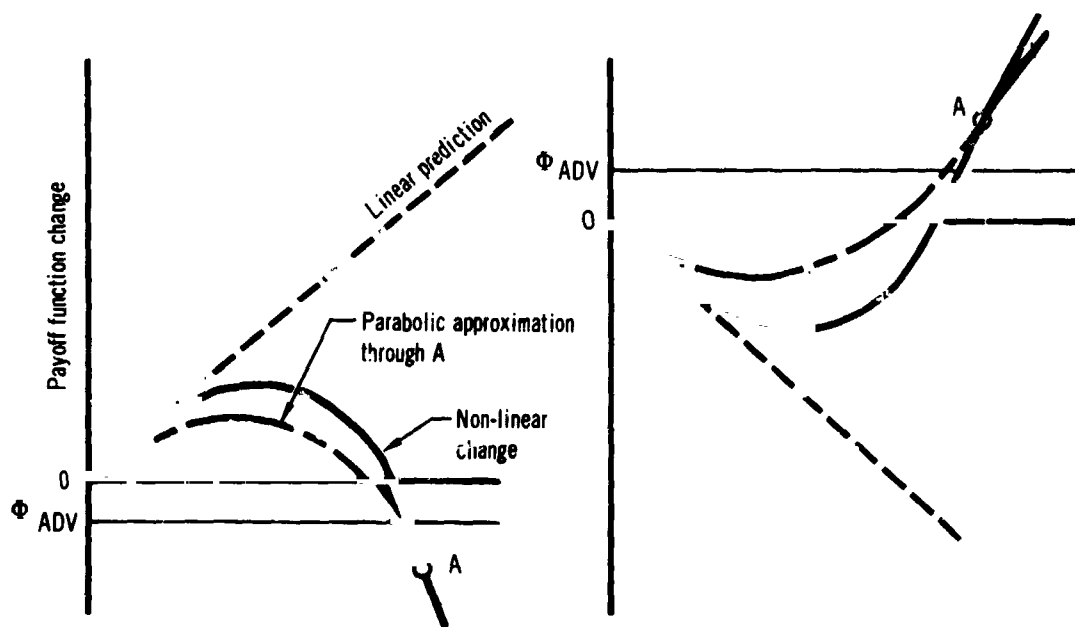
Figure 32.—Application of Parabolic Approximation
to Adverse ϕ Travel

$$\psi_{FWD_i} = 2.5$$

$$\psi_{BWD_i} = .5, \qquad N < 20 \tag{374}$$

$$\psi_{FWD_i} = 1.5$$

$$\psi_{BWD_i} = .025, \qquad N \geq 20 \tag{375}$$

36

It has previously been indicated that normally

$$C_\psi = N \Delta C_\psi \tag{376}$$

There is an exception to this rule. The exception occurs when the step-size magnitude on the first trial of an iteration is less than the amount required to provide the desired constraint change and when $\text{grad}\phi$ is positive. When this condition occurs, C_ψ is successively reduced by ΔC_ψ until the constraint change is less than the amount the DP^2 is capable of providing.

With ψ_{BWD} and ψ_{FWD} specified, the control system merely checks in which direction each constraint is travelling and computes by the now familiar parabolic approximation what values of k, if any, will cause each constraint to reach the boundary towards which it is travelling. The method is demonstrated in Figure 33 for a constraint which must be increased and has moved in the correct direction.
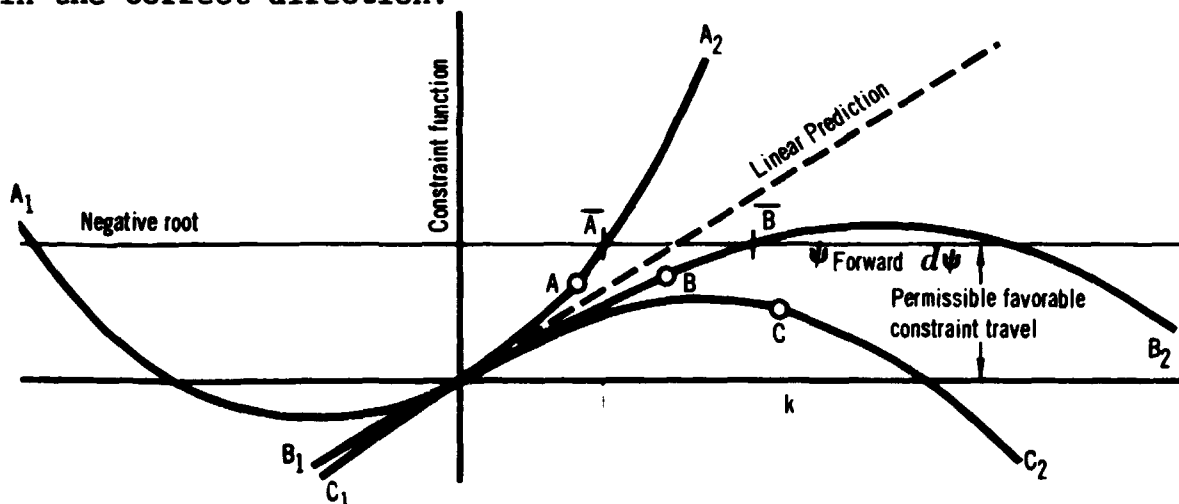

Figure 33.—Application of Parabolic Approximation, Constraint Moving in Desired Direction

111

If the trial point is at A, the parabolic approximation must behave in the manner of A_1A_2. The solution sought is at $\bar{A}$, and the negative root may be ignored. If the trial is at B, the approximate solution behaves in the manner of B_1B_2, and the point $\bar{B}$ is sought. If the trial is at C, the curve behaves in the manner of C_1C_2, and there is no real value of k which will produce a point on the forward boundary; in this case, the limit on k is ignored by setting $k_{\psi FWD} = \infty$.

In Figure 34 a constraint which must be increased is considered; here, however, the motion is adverse. A trial such as that at point D indicates that the point sought is at $\bar{D}$; the negative solution may be ignored. Similar sketches to those of Figure 33 and 34 may be drawn for a constraint which must be decreased.
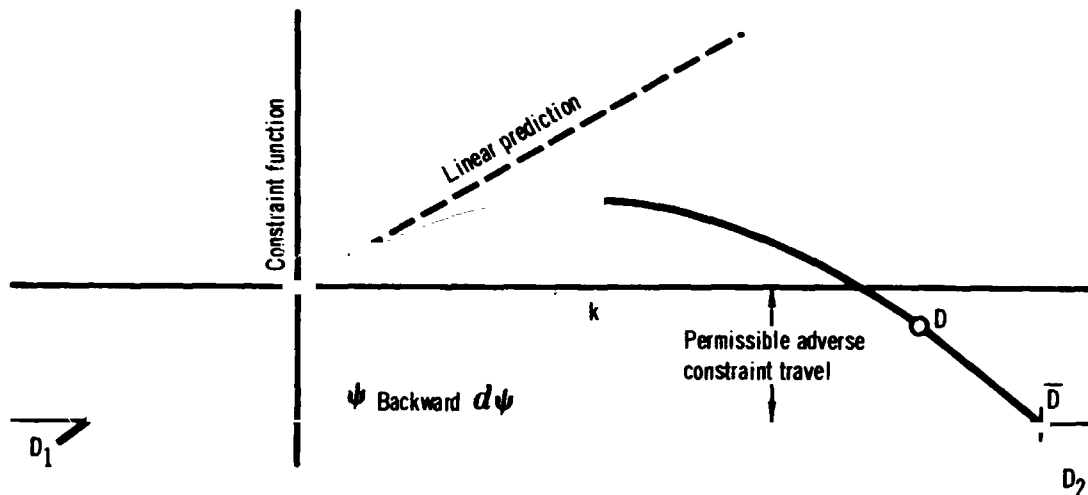

Figure 34.—Application of Parabolic Approximation, Constraint Moving in Wrong Direction

Let the value of k which places a constraint at the appropriate boundary on its travel be denoted by $k_{\psi TVL}$. If the solution is complex, adopt the convention that $k_{\psi TVL} = \infty$, so that

$$k_{\psi TVL} = \frac{-d\psi \pm \sqrt{(d\psi)^2 + 4(\Delta\psi - d\psi)\psi_{TVL}}}{2(\Delta\psi - d\psi)}$$

when

$$k_{\psi TVL} = \infty \ , \quad d\psi^2 + 4(\Delta\psi - d\psi)\psi_{TVL} < 0$$

(377)

and

$$d\psi^2 + 4(\Delta\psi - d\psi)\psi_{TVL} \geq 0$$

(378)

where $\psi_{TVL} = \psi_{BWD}d\psi$, if the constraint has moved in the wrong direction

$= \psi_{FWD}d\psi$, if the constraint has moved in the right direction

112

In Equation (377) the smallest positive root must be taken.

Conditions for Ignoring Dimensional Constraint Change Test.—
In some circumstances, the limits on end-point travel are ignored.
For example, if a constraint has been obtained within the accep-
table limits, ψ_{TOL}, which are specified by the analyst for the
particular problem, then its end-point travel ceases to be
monitored unless the constraint once again drifts outside the
acceptable limits. This decision is made in order to avoid
the possibility of limiting step-size magnitude on the basis of
a constraint which is essentially met, while considerable errors
still remain in other constraints or a significant amount of
performance gain remains.

The limits on constraint travel are also ignored for a
constraint error which is being reduced more rapidly than another
constraint error. This is achieved by creating a measure of
the end point errors at the termination of the nominal trajec-
tory. These errors are denoted by ψ_{ERR}. When, after a number
of iterations all the constraint errors have been halved, ψ_{ERR}
is also halved. This process is repeated until the computed
ψ_{ERR} are less than ψ_{TOL}; at this point ψ_{ERR} is set equal to ψ_{TOL}.
Any time a particular end-point error is less than ψ_{ERR}, its
dimensional end point travel will not be tested during the
following iteration.

It should be noted that if the ψ_{TOL} are zero, a danger of
false convergence exists; for if the constraints are essentially
satisfied before the greatest performance is obtained, a situa-
tion of the nature of that depicted in Figure 24 exists, and
the limits on constraint travel may inhibit the development of
performance.

It should be noted that whenever the controlling function
(the one with the greatest k based on linearity) is a constraint,
its end-point travel is always checked, for there is no point
in controlling with a constraint beyond the permissable limits
on its travel. If the limits on the travel of a controlling
constraint cause the step-size to be less than it would be if
based on linearity and that constraint is within ψ_{ERR}, then an
attempt to control with the next most linear function is made.
As the limits on the first controlling function travel can then
be ignored, it is possible that a larger step will result from
the use of the second controlling function. The larger of the
two step-sizes obtained in this manner is then used; if necessary
this process is repeated with the next most linear function, etc.

Majority Vote Test.— Only those trajectories on which at
least half the optimization functions of interest improve will
be considered satisfactory. The optimization functions of in-
terest are defined as the constraint functions having errors

greater than their respective ψ_{ERR}, and the payoff function provided the number of optimization functions of interest is odd or zero.

A trial trajectory which fails to satisfy the majority vote test is not permitted to lead to the final trajectory of an iteration (valid step). A valid step which fails to satisfy the test is over-ruled by another valid step. In either case, the new trajectory is computed with a step-size based on k = .5

Summary.— The control system essentials have been presented; it may be noted that wherever possible, the payoff function change is emphasized at the expense of the constraint changes.

Choice of step-size after the first trial is based on both linearity and dimensional changes. A careful examination of the various tests will reveal that the step-size parameter is basically given by the expression

$$k = \text{Min}\left(\text{Max}\left(\underset{\psi \leq \psi_{ERR}}{\text{Min}}\left(k_\psi, k_{\psi TVL}\right), k_\phi, k_{\psi \,>\, \psi_{ERR}}\right), k_{\psi TVL}, k_{\psi TVL \,\psi \,>\, \psi_{ERR}}\right) \quad (379)$$

This value of k must then be checked against the bounce test and the majority vote test. If

$$.5 \leq k \leq 2.0 \quad (380)$$

a final trajectory is computed; otherwise, a further trial trajectory at the appropriate limit is computed.

After a final trajectory, the majority vote test and the adverse travel test must both be satisfied; if they are not, then the final trajectory is recomputed with a step-size determined by k = .5 or on a computed $k_{\psi TVL}$.

The control system has shown extreme reliability to date provided only that the integration technique employed is adequate for the problem under consideration. Should a convergence failure be encountered by an analyst using the program of References 1 and 2, it is strongly recommended that the first course of action to follow is a critical examination of the integration technique and integration step being employed.

REFERENCES

1. Hague, D. S.: Three-Degree-of-Freedom Problem Optimization Formulation - Analytical Development. Part I, vol. 3, FDL-TDR-64-1, McDonnell-Douglas Corporation, October 1964.

2. Mobley, R. L. and Vorwald, R. R.: Three-Degree-of-Freedom Optimization Formulation - User's Manual. Part II, vol.3, FDL-TDR-64-1, McDonnell-Douglas Corporation, October 1964.

3. Brown, Robert C.; Brulle, R. V.; Combs, A. E.; and Griffin, G.D.: Six-Degree-of-Freedom Flight Path Study Generalized Computer Program. Part I, vl.1, FDL-TDR-64-1, McDonnell-Douglas Corporation, October 1964.

4. Seubert, F. W. and Usher, Newell E.: Six-Degree-of-Freedom Flight Path Study Generalized Computer Program - User's Manual. Part II, WADD Technical Report 60-781, McDonnell-Douglas Corporation, May 1961.

5. Landgraf, S. K.: Some Practical Applications of Performance Optimization Techniques to High-Performance Aircraft. AIAA Paper 64-288, July 1964.

6. Hague, D. S.: An Outline and Operating Instructions for the Steepest-Descent Trajectory Optimisation Programme - STOP. Aerodynamics Methods Note 1, McDonnell-Douglas Corporation, March 11, 1963.

7. Hague, D. S.; Geib, Ken; Ballew, L.; and Witherspoon, J.: Two Vehicle Optimization - Theoretical Outline and Program User's Manual. Report B983, McDonnell-Douglas Corporation, 1965.

8. Hague, D. S.: The Optimization of Multiple-Arc Trajectories by the Steepest-Descent Method. Recent Advances in Optimization Techniques. Lavi and Vogl, ed., John Wiley, 1966, pp. 489-517.

9. Hague, D. S. and Glatt, C. R.: Study of Navigation and Guidance of Launch Vehicles Having Cruise Capability. vol. 2, Boeing Document D2-113016-5, The Boeing Company, April 1967.

10. Retka, J.; Harder, D.; Hague, D. S.; Glatt, C. R.; Seavoy, T.; and Minden, D.: Study of Navigation and Guidance of Launch Vehicles Having Cruise Capability. vol. 4, Boeing Document D2-113016-7, April 1967.

11. Stein, H.; Mathews, M. L.; and Frenck, J. W.: STOP - A
    Computer Program for Supersonic Transport Trajectory
    Optimization. NASA CR-793, May 1967.

12. Petersen, L. D.: Trajectory Optimization by the Method
    of Steepest Descent. vol. 1, AFFDL-TR-67-108, April
    1968.

13. Moulton, Forest R.: Methods of Exterior Ballistics.
    Dover, New York.

14. Goldstein, H.: Classical Mechanics, Edison Wesley, 1950.

15. Hague, D. S. and Glatt, C. R.: An Introduction to
    Multivariable Search Techniques for Parameter Optimi-
    zation (and Program AESOP), NASA CR-73200, April 1968.

16. Hague, D. S. and Glatt, C. R.: A Guide to the Automated
    Engineering and Scientific Optimization Program -
    AESOP. NASA CR-73201, April 1968.

17. Hague, D. S. and Glatt, C. R.: Application of Multivar-
    iable Search Techniques to the Optimal Design of
    Hypersonic Cruise Vehicles. NASA CR-73202, April 1968.

18. Gregory, T. J.; Petersen, R. H.; and Wyss, J. A.: Perfor-
    mance Trade-Offs and Research Problems for Hypersonic
    Transports. Journal of Aircraft. vol. 2, no. 4, July
    August 1965.

19. Petersen, R. H.; Gregory, T. J., and Smith, C. L.: Some
    Comparisons of Turboramjet-Powered Hypersonic Aircraft
    for Cruise and Boost Missions. Journal of Aircraft.
    vol. 3, no. 5, September-October 1966.

20. Minzner, R. A. and Ripley, W. S.: The ARDC Model Atmos-
    phere. AFCRC-TN-56-204, United States Air Force,
    Cambridge Research Center, December 1956.

21. Champion, K. S. W.; Minzner, R. A.; and Pond, H. T.: The
    ARDC Model Atmosphere, 1959. AFCRC-TR-59-267, United
    States Air Force, Cambridge Research Center, August
    1959.

22. Anon.: U. S. Standard Atmosphere, 1962. U. S. Government
    Printing Office, 1962.

23. Heiskanen, W. and Meinesz, Vening: The Earth and Its
    Gravity Field. McGraw-Hill Book Company., 1958.

24. Jeffreys, H: The Earth. Fourth ed., Cambridge University
    Press, 1959.

116

25. O'Keefe, J. A.; Eckels, A.; and Squires, R. K.: Vanguard
    Measurements Give Pear-Shaped Component of Earth's
    Figure. Science vol. 129, p. 565.