—6—

Method of Gradients

HENRY J. KELLEY
Grumman Aircraft Engineering Corporation,
Bethpage, New York

Page

6.0 Introduction

The method of gradients or "method of steepest descent," as it is sometimes called, is an elementary concept for the solution of minimum problems. It dates back to Cauchy[1–3] and, in variational version, to Hadamard.[4,5] A particularly clear and attractive exposition of the method was given by Courant in a 1941 address to the American Mathematical Society.[5] In recent years the computational appeal of the method has led to its adoption in a variety of applications—multivariable minimum problems of ordinary calculus,[3,6–9] solution of systems of algebraic equations,[10,11] integral equations,[12] and variational problems.[9,13]

The gradient method has been applied to variational problems of flight path optimization by the present writer in the investigation of reference 14, a main source of material for this chapter. A similar scheme has been developed independently by Bryson and his colleagues.[15]

We will first discuss some of the main features of the gradient method in the context of ordinary minimum problems subject to constraints. Although this class of problems is chosen primarily for simplicity of explanation, it is one which is increasingly of interest per se in aeronautical and astronautical applications. We will then turn to variational problems of flight performance, introducing Green's functions in the role played by partial derivatives in ordinary minimum problems, and attempting to preserve an analogy between the two classes of problems in the subsequent development. Some numerical results illustrating the computational successive approximation procedure in examples will then be presented.

6.1 Gradient Technique in Ordinary Minimum Problems

6.11 The Continuous Descent Process

To present the basic idea of the gradient method we consider a function f of several variables $x_1, \cdots, x_n$, defined on an open domain, which possesses continuous partial derivatives with respect to these variables. Starting at some point $x_i = \bar{x}_i$, $i = 1, \cdots, n$, we move a small distance ds defined in the Euclidean sense

$$ds^2 = \sum_{i=1}^{n} dx_i^2 \tag{6.1}$$

Seeking to move toward a minimum of f, we consider directions in which the rate of change of f with respect to s

$$\frac{df}{ds} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{dx_i}{ds} \tag{6.2}$$

is negative. In fact we may find the direction of "steepest descent" (most negative df/ds) among the directions which make (6.2) stationary subject to (6.1).

Proceeding formally, we write the constraint (6.1)

$$1 - \sum_{i=1}^{n} \left(\frac{dx_i}{ds}\right)^2 = 0 \tag{6.3}$$

in terms of direction cosines dx_i/ds and adjoin it to (6.2) by means of a Lagrange multiplier λ_0, forming

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{dx_i}{ds} + \lambda_0 \left[1 - \sum_{i=1}^{n} \left(\frac{dx_i}{ds}\right)^2\right] \tag{6.4}$$

Equating partial derivatives taken with respect to the dx_i/ds to zero,

$$\frac{\partial f}{\partial x_i} + \lambda_0 \left(-2 \frac{dx_i}{ds}\right) = 0, \qquad i = 1, \cdots, n \tag{6.5}$$

we obtain

$$\frac{dx_i}{ds} = \frac{1}{2\lambda_0} \frac{\partial f}{\partial x_i}, \qquad i = 1, \cdots, n \tag{6.6}$$

From the constraint equation (6.3) the multiplier λ_0 is determined as

$$\lambda_0 = \pm \tfrac{1}{2} \left[\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2\right]^{1/2} \tag{6.7}$$

Provided the partial derivatives $\partial f/\partial x_i$ are not all zero, there are two distinct sets of direction numbers which make df/ds stationary, namely,

$$\frac{dx_i}{ds} = \pm \frac{\partial f}{\partial x_i} \left[\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2\right]^{-1/2}, \qquad i = 1, \cdots, n \tag{6.8}$$

Inspection of the expression for the directional derivative df/ds in the two cases

$$\frac{df}{ds} = \pm \left[\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2\right]^{1/2} \tag{6.9}$$

enables identification of the two stationary directions as those of steepest ascent and steepest descent.

A simple geometric interpretation of these formulas is possible if the various quantities are regarded in vector terms: the x_i as components of a vector $\mathbf{X}$, the direction cosines dx_i/ds as components of a unit vector $d\mathbf{X}/ds$, and the partial derivatives $\partial f/\partial x_i$ as components of a gradient vector. The derivative df/ds is then the dot product

$$\frac{df}{ds} = \operatorname{grad} f \cdot \frac{d\mathbf{X}}{ds} \tag{6.10}$$

and, with the direction of motion oriented along the gradient as per Eq. (6.8), the magnitude is equal to that of the gradient vector as given by Eq. (6.9). Thus, steepest ascent corresponds to motion in the gradient direction and steepest descent to that along the negative gradient.

We now introduce a time parameter σ and consider motion along the negative gradient direction as a continuous process. For motion in n-space at a velocity of magnitude V

$$\frac{ds}{d\sigma} = \left[\sum_{i=1}^{n} \left(\frac{dx_i}{d\sigma}\right)^2\right]^{1/2} = V \tag{6.11}$$

The expressions for the velocity components $dx_i/d\sigma$ become

$$\frac{dx_i}{d\sigma} = -V\left[\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2\right]^{-1/2} \frac{\partial f}{\partial x_i}, \qquad i = 1, \cdots, n \tag{6.12}$$

as a consequence of (6.8). These expressions become particularly simple if the velocity magnitude is taken proportional to that of the gradient

$$V = k\left[\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2\right]^{1/2}, \qquad k > 0 \tag{6.13}$$

$$\frac{dx_i}{d\sigma} = -k\,\frac{\partial f}{\partial x_i}, \qquad i = 1, \cdots, n \tag{6.14}$$

so that motion in the negative gradient direction is assured by merely setting the time derivatives of the coordinates proportional to the partial derivatives of f.

It is clear that in this continuous process, wherein the point $\mathbf{X}$ moves according to the system of ordinary differential equations (6.14), the process will for $\sigma \to \infty$ approach a position for which $\operatorname{grad} f = 0$, if f is bounded below. The stationary value so approached will correspond to a minimum of f if the x_i remain finite in the limit; otherwise a lower bound is approached.

6.12 Stepwise Version

As an alternative to the continuous procedure described by Eqs. (6.14) we may proceed stepwise, correcting a set of approximations to the solution $\partial f/\partial x_i = 0$ by increments proportional to the negative of the gradient:

$$x_i^{(p+1)} = x_i^{(p)} - k \frac{\partial f}{\partial x_i} \Delta\sigma, \qquad i = 1, \cdots, n \tag{6.15}$$

It is clear that in this stepwise process, the proportionality constant k may be absorbed in the step size $\Delta\sigma$; hence we take $k = 1$.

Since the determination of the partial derivatives $\partial f/\partial x_i$ may be expensive in terms of volume of numerical computations in case the number of variables n is large, it is desirable to exploit each calculation of local gradient direction to the utmost, taking $\Delta\sigma$ as large as possible. One procedure is to follow the local gradient direction until f reaches a minimum, i.e., evaluate Eqs. (6.15) and the function f for a number of step sizes, determining $\Delta\sigma$ for minimum f by some suitable one-dimensional search technique. A new gradient direction is then calculated and the procedure repeated.

In this way an n-dimensional minimum problem is reduced to a sequence of one-dimensional problems. The gradient method shares this feature with another computational technique to be described in later chapters.

The continuous and stepwise processes are contrasted in the sketch of Fig. 1 which depicts the two types of motion as they may occur in the

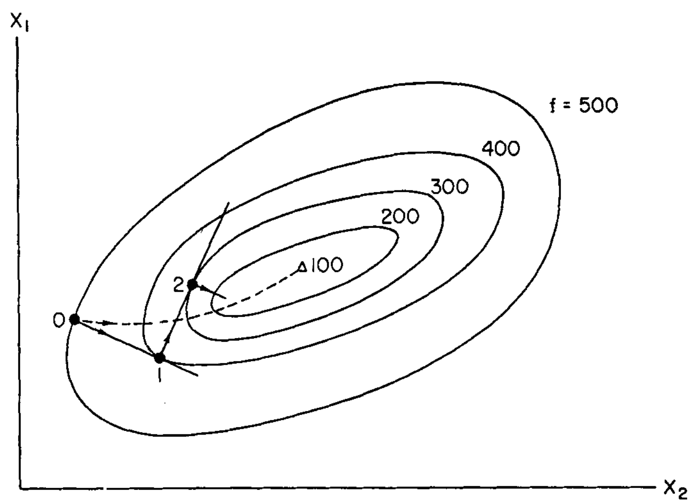


FIG. 1. Continuous and stepwise descent processes.

vicinity of a minimum of a function of two variables $f(x_1, x_2)$. The gradient direction, as shown, is normal to a contour while the local minimum in the gradient direction is attained at a point of tangency to a contour. These characteristics are also common to higher dimensional cases.

It is apparent that the stepwise path is not independent of the coordinate system selected. If a transformation of coordinates could be found which maps the oval contours of Fig. 1 into circles, for example, the number of steps required to attain the minimum would be reduced to one. In the usual situation, of course, insufficient information is available *a priori* to permit a sophisticated choice of coordinates; otherwise the character of the surface would be known and gradient determination of minima unnecessary.

With typical engineering problems, however, one will often have some idea of the "practical range" of the variables x_i, and this will facilitate the introduction of normalized variables for gradient computations. Such a procedure is virtually a necessity where the variables have different dimensions and are numerically of different orders of magnitude. As a result of normalization some semblance of meaning may be attached to the concept of distance in the n-space in which the operations are performed.

6.13 Ordinary Minimum Problems with Constraints: Gradient Projection Technique

In many problems of practical interest, we may wish to determine a minimum of the function $f(x_1, \cdots, x_n)$ subject to subsidiary conditions relating the variables x_i. An important class of subsidiary conditions are *equations*

$$g_j(x_1, \cdots, x_n) = 0, \qquad j = 1, \cdots, m \tag{6.16}$$

numbering $m < n$, sometimes referred to in the literature as equality constraints.

If the functions g_j are given analytically, it may be possible to solve the set of Eqs. (6.16) for m of the variables x_i in terms of the others and, by their elimination from the function f, to reduce the problem to one without constraints. Such an approach will very often not be practicable in applications.

We may, however, consider a version of this procedure "in the small," i.e., in a neighborhood of the starting point $x_i = \bar{x}_i$. Retaining only first-order terms in a Taylor expansion

$$g_j(x_1, \cdots, x_n) \cong g_j(\bar{x}_1, \cdots, \bar{x}_n) + \sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i} (\bar{x}_1, \cdots, \bar{x}_n)(x_i - \bar{x}_i), \qquad j = 1, \cdots, m \tag{6.17}$$

and assuming that Eqs. (6.16) are satisfied at the point $x_i = \bar{x}_i$, we obtain a system of linear equations in the increments $\Delta x_i = x_i - \bar{x}_i$ for the vanishing of the g_j to first order:

$$\sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i} \Delta x_i = 0, \qquad j = 1, \cdots, m \tag{6.18}$$

We may then attempt to solve this system for m of the increments Δx_i in terms of the remaining ones. This will be possible unless *all* of the $m \times m$ Jacobian determinants of the system vanish at the point $x_i = \bar{x}_i$. The procedure breaks down at such *singular points*.[16]

If solution is possible, the $n - m$ partial derivatives of f with respect to the $n - m$ remaining variables may be calculated by the chain rule of differentiation, and the gradient of f subject to the constraints thus determined. Geometrically, this is the projection of the free gradient vector upon the $n - m$ subspace determined by the intersection of the m hyperplanes (6.18). The terminology *gradient projection* for this scheme, as employed in the nonlinear programming literature[7,8] in connection with related problems involving inequality constraints, seems appropriate.

We may perform an analysis similar to that of Section 6.11 to determine the projected gradient direction. Again introducing direction cosines dx_i/ds, we seek stationary directions of df/ds as given by (6.2) subject to (6.3) and to

$$\frac{dg_j}{ds} = \sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i} \frac{dx_i}{ds} = 0, \qquad j = 1, \cdots, m \tag{6.19}$$

Introducing Lagrange multipliers λ_0 and λ_j, $j = 1, \cdots, m$, we form

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{dx_i}{ds} + \lambda_0 \left[1 - \sum_{i=1}^{n} \left(\frac{dx_i}{ds}\right)^2\right] + \sum_{j=1}^{m} \lambda_j \sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i} \frac{dx_i}{ds} \tag{6.20}$$

Equating partial derivatives taken with respect to the dx_i/ds to zero, we obtain

$$\frac{\partial f}{\partial x_i} - 2\lambda_0 \frac{dx_i}{ds} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \qquad i = 1, \cdots, n \tag{6.21}$$

$$2\lambda_0 = \left[\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}\right)^2\right]^{1/2} \tag{6.22}$$

$$\frac{dx_i}{ds} = \frac{\left[\dfrac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \dfrac{\partial g_j}{\partial x_i}\right]}{\left[\sum_{i=1}^{n} \left(\dfrac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \dfrac{\partial g_j}{\partial x_i}\right)^2\right]^{1/2}} \tag{6.23}$$

Substituting this result into (6.19), we obtain relations determining the λ_j:

$$\sum_{j=1}^{m} \lambda_j \sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i} \frac{\partial g_k}{\partial x_i} = - \sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i} \frac{\partial f}{\partial x_i}, \qquad k = 1, \cdots, m \tag{6.24}$$

This system of linear equations in the multipliers will have a solution provided that the $m \times m$ matrix whose elements are

$$a_{kj} = \sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i} \frac{\partial g_k}{\partial x_i} \tag{6.25}$$

is nonsingular. This statement is precisely the Gram determinant criterion for linear independence of the expressions (6.19) and it is equivalent to the earlier assertion concerning Jacobian determinants (see Courant and Hilbert,[17] p. 34). Fulfillment of this condition requires that the magnitudes of the vectors grad g_j be nonzero and that their directions be distinct from one another.

Mechanization of the gradient projection procedure thus requires numerical solution of the system of linear equations (6.24). The relations appropriate to a stepwise process analogous to that given by (6.15) are then

$$x_i^{(p+1)} = x_i^{(p)} - \left(\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}\right) \Delta\sigma, \qquad i = 1, \cdots, n \tag{6.26}$$

The gradient projection technique is an excellent one where the constraints (6.16) are linear in the x_i, or nearly so. There is a possibility of difficulty associated with ill-conditioning of the linear system (6.24) if the intersections of the hypersurfaces $g_j = 0$ are poorly defined, e.g., if two or more of the tangent hyperplanes are nearly parallel.

A considerable complication arises from significant nonlinearities in the constraints. For large step sizes $\Delta\sigma$, Eqs. (6.16), having been satisfied only in linearized version, may long since have been violated before minimum f is reached. A correction cycle designed to restore the constraints must then be introduced. Typically this will take the form of an iterative adjustment of the variables x_i. This feature of the gradient projection technique may entail a considerable increase in computation time.

6.14 Ordinary Minimum Problems with Constraints: An Approximation Technique

Another approach to the handling of constraints is provided by the idea of approximating a minimum problem subject to constraints by another problem without constraints. Thus, in lieu of the problem of the preceding section, we consider the problem of minimizing

$$f + \tfrac{1}{2} \sum_{j=1}^{m} K_j g_j^2 \tag{6.27}$$

where the K_j are positive constants.

It is intuitively reasonable that the "penalty" terms of the second member of (6.27) will have the effect of making the constraint "violations" small in this problem, owing to the fact that these terms are nonnegative. For increasingly large positive K_j, it may be anticipated that the solution of this minimum problem will tend toward the desired solution of the minimum problem for f subject to the constraints (6.16). This idea is due to Courant.[5] It has been placed upon a rigorous basis in terms of an approximation theorem by Moser (see Courant[18]) for the case of a single constraint, and exploited in a particular class of variational problems involving multiple constraints by Rubin and Ungar.[19]

In the employment of this idea for computational purposes, numerically large values of the constants K_j are to be assigned. The choice of values must be decided on the basis of permissible approximations to the constraints. Thus "tolerances" may be set from physical considerations.

The penalty function technique is quite compatible with the successive approximation process provided by the gradient method. A plausible technique for computer operations is comparison of the constraint "violations" with preassigned "tolerances" at the end of each descent step, followed by appropriate adjustment of K_j values if necessary for the succeeding step.

Seeking a guide to the estimation and adjustment of the values of the K_j, we now examine the magnitudes of the "violations" g_j at a minimum of (6.27). We equate to zero the derivatives of (6.27) in the m-directions determined by the gradients of the g_j. The direction cosines of the gradient of g_k are

$$\frac{dx_i}{ds} = \frac{\partial g_k}{\partial x_i} \left[\sum_{l=1}^{n} \left(\frac{\partial g_k}{\partial x_l} \right)^2 \right]^{-1/2}, \qquad i = 1, \cdots, n \tag{6.28}$$

and the derivative of (6.27) in this direction is assumed to vanish:

$$\frac{\left[\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} K_j g_j \frac{\partial g_j}{\partial x_i} \right) \frac{\partial g_k}{\partial x_i} \right]}{\left[\sum_{i=1}^{n} \left(\frac{\partial g_k}{\partial x_i} \right)^2 \right]^{1/2}} = 0, \qquad k = 1, \cdots, m \tag{6.29}$$

We may regard this as a system of linear equations in the products $K_j g_j$:

$$\sum_{j=1}^{m} K_j g_j \sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i} \frac{\partial g_k}{\partial x_i} = -\sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i} \frac{\partial f}{\partial x_i}, \qquad k = 1, \cdots, m \tag{6.30}$$

This system bears a strong resemblance to that governing the multipliers of the (converged) gradient projection process (6.24), the distinction being that the partial derivatives in the present case are evaluated at a minimum of (6.27) for which the constraint equations are satisfied only approximately.

The system (6.30) may be solved numerically under conditions mentioned previously in connection with determination of the multipliers. In the limit as $K_j \to \infty$, evidently

$$g_j \to \frac{\lambda_j}{K_j}, \qquad j = 1, \cdots, m \tag{6.31}$$

and the "violations" are seen to vary inversely with the K_j values in the neighborhood of the solution.

This analysis suggests two plausible schemes for adjustment of the K_j values. The simplest is employment of the absolute value of "violation"/"tolerance" ratio as a factor to increase or decrease the "current" K_j values at the end of each descent step. As an alternative, one may estimate the λ_j from (6.24) on the basis of "current" partial derivatives and compute

$$K_j = \left| \frac{\lambda_j \text{ (est.)}}{\epsilon_j} \right|$$

where ϵ_j is the "tolerance" set on the constraint $g_j = 0$. Preliminary indications from numerical experiments with analogous adjustment schemes in variational problems indicate that the second is preferable to the first in regard to speed of convergence.

The attractive feature of the "penalty function" scheme is that it avoids the need for iterative corrections to assure satisfaction of the constraint equations as the descent process continues. By the same token, the need for separate determination of a starting point $\bar{x}_i$ which satisfies the constraints is obviated. At the time of the present writing, there is insufficient computational experience with this method to permit comparison with other techniques in ordinary minimum problems; however, some experience with a similar technique in solution of variational problems will be reported in a later section.

6.15 Ordinary Minimum Problems with Constraints: Inequality Constraints

Constraints which take the form of inequalities

$$g_j(x_1, \cdots, x_n) \leq 0, \qquad j = 1, \cdots, m \tag{6.32}$$

are increasingly of interest in applications. Ordinary minimum problems featuring such constraints form the basis of nonlinear programming theory, for which the reader is referred to Rosen[7] and Wolfe.[8]

In employing the gradient projection idea for solution of such problems, one must first find a suitable starting point for which relations (6.32) are satisfied. If the starting point is an interior point [a point for which strict inequality signs apply in (6.32)], one then proceeds in the negative free gradient direction until a minimum of f is reached or until one of the g_j changes sign. When the threshold of a constraint is reached, one then employs gradient projection, regarding the constraint as an equality. The subsequent possibilities for motion subject to various degrees of constraint are numerous, and it is clear that extensive testing and provision for projecting upon various combinations of constraints will require quite a sophisticated computer program if several inequality constraints are to be dealt with simultaneously.

The "penalty function" notion may also be applied to minimum problems featuring inequality constraints. To the function f are added terms comprising a "penalty function" of the form

$$f + \tfrac{1}{2} \sum_{j=1}^{m} K_j g_j^2 H(g_j) \tag{6.33}$$

Here $H(g_j)$ is the Heaviside unit step function of argument g_j. With the K_j chosen as positive constants, the second member is nonnegative. Note that the partial derivatives of the second member are continuous if the partial derivatives of the functions g_j are continuous. This feature favors the retention of the square law form of penalty even though it is not required to make the second member nonnegative in the case of inequality constraints.

This scheme possesses an attractive simplicity when employed in conjunction with the gradient method in that computer logic is minimal. In gradient calculations, the influence of a particular inequality constraint is automatically nil if the constraint is satisfied and increasingly large as it is violated. As in the case of equality constraint "penalties," adjustment of the constants K_j should be performed systematically, after each local minimum is attained, on the basis of comparison between "violations" and preassigned "tolerances."

An analysis to determine the products $K_j g_j$ for adjustment purposes may be carried out in a manner similar to that of the preceding section. One will automatically lose a row and a column of the matrix of coefficients of the system analogous to (6.30) for each strict inequality satisfied at the point under examination, in that the partial derivatives

$$\frac{\partial}{\partial x_i} g_k^2 H(g_k), \qquad i = 1, \cdots, n$$

vanish at points for which $g_k < 0$.

6.2 Gradient Technique in Flight Path Optimization Problems

An earlier chapter has developed the classical "indirect" method of the calculus of variations which is based upon the reduction of variational problems to differential equations. Although many interesting results have been forthcoming from analytical solutions of the Euler-Lagrange differential equations governing optimal flight, the idealizing assumptions usually invoked limit their applicability in practical situations. Under more realistic assumptions, a numerical attack on these equations is required and in this approach a serious difficulty may arise in the satisfaction of two-point boundary conditions (see, for example, Mengel,[20] Irving and Blum,[21] and Faulders[22]). This difficulty becomes a limiting factor where the order of the differential equations governing the basic system is four or higher.

This situation has provided the motivation for attack on variational problems of flight performance by means of the gradient technique. An application of gradient method to fixed end-point variational problems was given by Stein.[13] The class of problems featuring differential equations as constraints is more complex, and our development will be heuristic in character. We will, in the following, assume whatever continuity and differentiability properties may be necessary to avoid difficulty.

6.21 Problem Formulation

For present purposes it will be assumed that the system of differential equations to be satisfied along the flight path is given in first-order form:

$$\dot{x}_m = g_m(x_1, \cdots, x_n, y, t), \qquad m = 1, \cdots, n \tag{6.34}$$

These equations relate velocities and positions, forces and accelerations, mass and flow of propellants and coolants, and the like. The x_m are termed problem or "state" variables, and y the control variable. Differentiation with respect to the independent variable, time t, is denoted by a superscribed dot.

An important class of problems is that in which the performance quantity to be minimized is expressed as a function of the final values of the variables x_m and t:

$$P = P(x_{1_f}, \cdots, x_{n_f}, t_f) \tag{6.35}$$

At a specified initial time t_0 as many as n boundary conditions on the x_m may be stipulated. Since an entire function $y(t)$ is at our disposal, we may reasonably consider problems in which numerous conditions are imposed upon the x_m at various subsequent t values. In the following we will restrict attention to conditions imposed at the terminal point of the flight path. Among the $n + 1$ quantities consisting of the n final values of the x_m plus the final time t_f, no more than n relations may be specified in order that the value of P not be predetermined.

This problem statement is essentially that employed in an earlier chapter in connection with the Mayer formulation of variational problems.

6.22 Neighboring Solutions and Green's Functions

We now assume that a solution of Eqs. (6.34) is available which does not minimize P. This solution is required to satisfy the specified initial conditions; it may or may not be required to also satisfy the specified conditions at the terminal point depending on the version of the gradient method to be adopted, as discussed in later sections. Denoting the solution by $x_m = \bar{x}_m(t)$, $y = \bar{y}(t)$, we examine behavior in the neighborhood of this solution by setting $x_m = \bar{x}_m + \delta x_m$, $y = \bar{y} + \delta y$ and linearizing:

$$\delta \dot{x}_m = \sum_{j=1}^{n} \frac{\partial g_m}{\partial x_j} \delta x_j + \frac{\partial g_m}{\partial y} \delta y, \qquad m = 1, \cdots, n \tag{6.36}$$

The partial derivatives of the g_m are evaluated along $x_m = \bar{x}_m$, $y = \bar{y}$ and are therefore known functions of the independent variable t. The functions δx_m and δy are the *variations* of x_m and y in the neighborhood of $\bar{x}_m$, $\bar{y}$. The motivation for study of the linearized system (6.36) is the alteration of the control function $y(t)$ by means of a gradient process such as to obtain a reduction in the function P whose minimum is sought.

A formal solution of Eqs. (6.36) may be written in the form:

$$\delta x_m = \sum_{p=1}^{n} \delta x_p(t_0) \xi_{mp}(t_0, t) + \int_{t_0}^{t} \mu_m(\tau, t)\, \delta y(\tau)\, d\tau, \qquad m = 1, \cdots, n \tag{6.37}$$

where the first member represents solution of the homogeneous system of equations and the second a superposition of control variable effects. The functions μ_m are Green's functions or influence functions; $\mu_m(\tau, t)$ may be regarded as the solution for δx_m corresponding to a unit impulse (Dirac delta function) in control introduced at time τ.[23] The μ_m are related to the functions ξ_{mp} of the homogeneous system by

$$\mu_m(\tau, t) = \sum_{p=1}^{n} \xi_{mp}(\tau, t) \frac{\partial g_p}{\partial y}, \qquad m = 1, \cdots, n \tag{6.38}$$

Since interest centers on final values of the x_m, we evaluate expressions (6.37) at $t = t_f$; however, to provide for determination of the effects of small variations in terminal time δt_f from the terminal time $t_f = \bar{t}_f$ of solution $\bar{x}_m$, $\bar{y}$, we include a first-order correction term:

$$\delta x_{m_f} = \sum_{p=1}^{n} \delta x_p(t_0)\xi_{mp}(t_0, \bar{t}_f) + \int_{t_0}^{\bar{t}_f} \mu_m(\tau, \bar{t}_f)\, \delta y(\tau)\, d\tau + \bar{g}_{m_f}\, \delta t_f$$

$$m = 1, \cdots, n \tag{6.39}$$

Here the symbol $\bar{g}_{m_f}$ denotes the derivative $\dot{\bar{x}}_m(\bar{t}_f) = g_m[\bar{x}_1(\bar{t}_f), \cdots, \bar{x}_n(\bar{t}_f), \bar{y}(\bar{t}_f)]$ evaluated at the terminal point of the nonminimal solution.

6.23 The Adjoint System

Since computation of the functions $\mu_m(\tau, t)$ over a complete range of both arguments will be found unnecessary, only their evaluation at $t = \bar{t}_f$ being required for subsequent calculations, it is reasonable to seek a means for performing the special computation which avoids the labor of the more general one. The following development relates the functions $\mu_m(\tau, t)$ to solutions of an adjoint system of equations through an application of Green's theorem. The scheme is based on the work of Bliss[24] and Goodman and Lance.[25]

We rewrite Eqs. (6.36) employing a subscript notation suitable to our immediate purpose:

$$\delta \dot{x}_i = \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j} \delta x_j + \frac{\partial g_i}{\partial y} \delta y, \qquad i = 1, \cdots, n \tag{6.40}$$

and write the system of equations adjoint to this system

$$\dot{\lambda}_i = -\sum_{j=1}^{n} \frac{\partial g_j}{\partial x_i} \lambda_j, \qquad i = 1, \cdots, n \tag{6.41}$$

which by definition is the system obtained from the homogeneous system by transposing the matrix of coefficients and changing the sign.

The solutions of the two systems are related by

$$\frac{d}{dt}\sum_{i=1}^{n}\lambda_i\,\delta x_i = \sum_{i=1}^{n}\lambda_i\frac{\partial g_i}{\partial y}\,\delta y \tag{6.42}$$

as may be verified directly by evaluating the derivative on the left:

$$\frac{d}{dt}\sum_{i=1}^{n}\lambda_i\,\delta x_i = \sum_{i=1}^{n}\dot{\lambda}_i\,\delta x_i + \sum_{i=1}^{n}\lambda_i\,\delta\dot{x}_i$$

$$= -\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial g_j}{\partial x_i}\lambda_j\,\delta x_i + \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\frac{\partial g_i}{\partial x_j}\,\delta x_j + \sum_{i=1}^{n}\lambda_i\frac{\partial g_i}{\partial y}\,\delta y \tag{6.43}$$

and noting the cancellation of terms arising by interchange of subscripts i and j in the double summations. After integration of both left and right members between definite limits t_0 and $\bar{t}_f$, we find

$$\sum_{i=1}^{n}\lambda_i(\bar{t}_f)\;\delta x_i(\bar{t}_f) - \sum_{i=1}^{n}\lambda_i(t_0)\;\delta x_i(t_0) = \int_{t_0}^{\bar{t}_f}\sum_{i=1}^{n}\lambda_i\frac{\partial g_i}{\partial y}\,\delta y\,dt \tag{6.44}$$

This is the one-dimensional form of Green's theorem.

We now consider numerical solution of the adjoint system with all boundary values specified at $t = \bar{t}_f$. To the special solutions corresponding to

$$\begin{aligned} \lambda_i(\bar{t}_f) &= 0, \qquad i \neq m \\ \lambda_i(\bar{t}_f) &= 1, \qquad i = m \end{aligned} \tag{6.45}$$

we assign the symbols $\lambda_i^{(m)}(t)$. In this fashion n expressions for the values of the $\delta x_m(\bar{t}_f)$ are obtained from Eq. (6.44):

$$\delta x_m(\bar{t}_f) = \sum_{i=1}^{n}\lambda_i^{(m)}(t_0)\;\delta x_i(t_0) + \int_{t_0}^{\bar{t}_f}\sum_{i=1}^{n}\lambda_i^{(m)}\frac{\partial g_i}{\partial y}\,\delta y\,dt, \qquad m = 1, \cdots, n \tag{6.46}$$

By comparison of Eqs. (6.39) and (6.46) we may now relate the functions ξ_{mp} to the unit adjoint solutions defined by Eq. (6.45) as

$$\xi_{mp}(t_0, \bar{t}_f) = \lambda_p^{(m)}(t_0), \qquad \begin{aligned} m &= 1, \cdots, n \\ p &= 1, \cdots, n \end{aligned} \tag{6.47}$$

[Note that the dependence of the right-hand members upon $\bar{t}_f$ is implicit in the definition of (6.45).] The Green's functions μ_m may also be expressed in terms of unit adjoint solutions as

$$\mu_m(\tau, \bar{t}_f) = \sum_{p=1}^{n} \lambda_p^{(m)}(\tau) \frac{\partial g_p}{\partial y}(\tau), \qquad m = 1, \cdots, n \tag{6.48}$$

In the preceding development the choice of symbols λ for the variables of the adjoint system is deliberate, for Eqs. (6.41) are precisely those governing the Lagrange multiplier functions of the "indirect" theory. We note the important distinction, however, that the coefficients of (6.41) employed in the "indirect" theory are evaluated along a minimal solution of Eqs. (6.34), whereas in gradient computations they correspond to nonminimal paths.

The close relationship between Green's functions or influence functions and the "error coefficients" of guidance theory has drawn attention to the usefulness of the adjoint system technique in guidance analysis.[26–28]

6.24 The Gradient in Function Space

Introducing, as before, a second independent variable σ, we seek analogous means of performing gradient computations. We first evaluate the slope of descent of the performance quantity P at a "point" in function space determined by the nonminimal solution $x_m = \bar{x}_m(t)$, $y = \bar{y}(t)$,

$$\frac{dP}{d\sigma} = \sum_{m=1}^{n} \frac{\partial P}{\partial x_{m_f}} \frac{dx_{m_f}}{d\sigma} + \frac{\partial P}{\partial t_f} \frac{dt_f}{d\sigma} \tag{6.49}$$

If the σ derivatives of the initial and final values of the x_m, the final time t_f, and the function $y(t, \sigma)$ are taken and evaluated at $\sigma = \bar{\sigma}$, corresponding to the nonminimal solution, Eqs. (6.39) become

$$\frac{dx_{m_f}}{d\sigma} = \sum_{p=1}^{n} \frac{dx_{p_0}}{d\sigma} \xi_{mp}(\bar{t}_f) + \int_{t_0}^{\bar{t}_f} \mu_m \frac{\partial y}{\partial \sigma} d\tau + \bar{g}_{m_f} \frac{dt_f}{d\sigma}, \qquad m = 1, \cdots, n \tag{6.50}$$

and expression (6.49) may then be written in the following form:

$$\frac{dP}{d\sigma} = \sum_{p=1}^{n} \frac{dx_{p_0}}{d\sigma} \sum_{m=1}^{n} \frac{\partial P}{\partial x_{m_f}} \xi_{mp}(\bar{t}_f) + \frac{dt_f}{d\sigma}\left(\frac{\partial P}{\partial t_f} + \sum_{m=1}^{n} \frac{\partial P}{\partial x_{m_f}} \bar{g}_{m_f}\right) + \int_{t_0}^{\bar{t}_f} \left(\sum_{m=1}^{n} \frac{\partial P}{\partial x_{m_f}} \mu_m\right) \frac{\partial y}{\partial \sigma} d\tau \tag{6.51}$$

In the type of problem presently under consideration, the performance quantity P depends implicitly upon a finite number of parameters, the initial values of the x_m and the final time t_f; it also depends upon the function $y(t)$. Hence the problem is of "mixed" type, having partly the character of an ordinary minimum problem and partly a variational character. We will momentarily assume that the initial x_m values and that of the final time t_f are fixed:

$$\frac{dx_{p_0}}{d\sigma} = 0, \qquad p = 1, \cdots, n$$
$$\frac{dt_f}{d\sigma} = 0 \tag{6.52}$$

in order to examine independently the variational aspect.

Under these circumstances the expression $dP/d\sigma$ takes the form of the integral

$$\frac{dP}{d\sigma} = \int_{t_0}^{t_f} \left(\sum_{m=1}^{n} \frac{\partial P}{\partial x_{mf}} \mu_m \right) \frac{\partial y}{\partial \sigma} \, d\tau \tag{6.53}$$

We wish to determine the "direction" $\partial y(\tau)/\partial\sigma$ of steepest descent in the function space $y(\tau)$. In a procedure analogous to the earlier treatment of ordinary minimum problems, we consider a differential distance ds in the function space $y(\tau)$, defined by

$$1 - \int_{t_0}^{t_f} \left(\frac{\partial y}{\partial s} \right)^2 d\tau = 0 \tag{6.54}$$

and seek stationary values of

$$\frac{dP}{ds} = \int_{t_0}^{t_f} \left(\sum_{m=1}^{n} \frac{\partial P}{\partial x_{mf}} \mu_m \right) \frac{\partial y}{\partial s} \, d\tau \tag{6.55}$$

subject to (6.54) as a constraint.

We form

$$\frac{dP}{ds} + \Lambda_0 \left[1 - \int_{t_0}^{t_f} \left(\frac{\partial y}{\partial s} \right)^2 d\tau \right] \tag{6.56}$$

and set the derivative of this expression with respect to $\partial y(\tau)/\partial s$ to zero for all τ, obtaining

$$2\Lambda_0 \frac{\partial y}{\partial s} = \sum_{m=1}^{n} \frac{\partial P}{\partial x_{mf}} \mu_m \tag{6.57}$$

From (6.54) the multiplier Λ_0 is evaluated as

$$2\Lambda_0 = \pm\left[\int_{t_0}^{t_f}\left(\sum_{m=1}^{n}\frac{\partial P}{\partial x_{m\,f}}\,\mu_m\right)^2 d\tau\right]^{1/2} \tag{6.58}$$

and

$$\frac{\partial y}{\partial s} = \frac{\pm\sum_{m=1}^{n}\dfrac{\partial P}{\partial x_{m\,f}}\,\mu_m}{\left[\int_{t_0}^{t_f}\left(\sum_{m=1}^{n}\dfrac{\partial P}{\partial x_{m\,f}}\,\mu_m\right)^2 d\tau\right]^{1/2}} \tag{6.59}$$

Taking the "velocity" of motion in function space as

$$\frac{ds}{d\sigma} = k\left[\int_{t_0}^{t_f}\left(\sum_{m=1}^{n}\frac{\partial P}{\partial x_{m\,f}}\,\mu_m\right)^2 d\tau\right]^{1/2} \tag{6.60}$$

we obtain

$$\frac{\partial y}{\partial \sigma} = \pm k\sum_{m=1}^{n}\frac{\partial P}{\partial x_{m\,f}}\,\mu_m \tag{6.61}$$

We thus identify

$$[P]_y = \sum_{m=1}^{n}\frac{\partial P}{\partial x_{m\,f}}\,\mu_m \tag{6.62}$$

as the "free" gradient direction and assure motion in the negative gradient direction by choosing the negative sign in (6.61). With $k = 1$

$$\frac{\partial y}{\partial \sigma} = -[P]_y \tag{6.63}$$

This development has been carried out by analogy with a characteristic property of a vector gradient, for a discussion of which the reader is referred to Courant[5] and to Courant and Hilbert[17] (pp. 222–224).

In this subsection the most convenient combination of boundary conditions has been assumed for simplicity of explanation, namely: t_0 and t_f fixed, x_{m_0} fixed for all m, $x_{m\,f}$ not appearing in the function P unspecified. The handling of other types of boundary conditions will be the subject of the next two subsections.

6.25 Boundary Conditions as Constraints

We consider boundary conditions of separated type, i.e., equations relating either initial values or final values. Terminal values, for example, may be variable on a surface typified by

$$\mathfrak{J}_j(x_{1\mathrm{f}}, \cdots, x_{n\mathrm{f}}, t_\mathrm{f}) = 0 \tag{6.64}$$

and there may be a number of such relations specified $j = 1, \cdots, l < n + 1$ if the final time is variable and $l < n$ if it is fixed.

Linearized versions of such constraints are given by

$$\bar{\mathfrak{J}}_j + \delta\mathfrak{J}_j = \bar{\mathfrak{J}}_j + \sum_{m=1}^{n} \frac{\partial \mathfrak{J}_j}{\partial x_{m\mathrm{f}}} \delta x_{m\mathrm{f}} + \frac{\partial \mathfrak{J}_j}{\partial t_\mathrm{f}} \delta t_\mathrm{f} = 0, \qquad j = 1, \cdots, l \tag{6.65}$$

Similar constraints relating initial values x_{m_0} and t_0 may also be specified. For the present we confine attention to the case of fixed initial values and fixed final time t_f. Hence we assume

$$\frac{\partial \mathfrak{J}_j}{\partial t_\mathrm{f}} = 0 \quad \text{and} \quad l < n \tag{6.66}$$

In an analysis similar to that given in the preceding section, we write the constraints (6.65) in the form

$$\frac{d\mathfrak{J}_j}{ds} = \int_{t_0}^{t_\mathrm{f}} \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{J}_j}{\partial x_{m\mathrm{f}}} \mu_m \right) \frac{\partial y}{\partial s} \, d\tau = 0 \tag{6.67}$$

and adjoin them to (6.56) by means of Lagrange multipliers $\Lambda_j, j = 1, \cdots, l$. The omission of zero-order terms in expressions (6.65) corresponds to an assumption that the solution $\bar{x}_m$, $\bar{y}$ satisfies the boundary conditions (6.64).

Leaving details of the derivation to the interested reader, we state the principal result for the "projected" gradient direction

$$[P]_y = \sum_{m=1}^{n} \frac{\partial P}{\partial x_{m\mathrm{f}}} \mu_m + \sum_{j=1}^{l} \Lambda_j \sum_{m=1}^{n} \frac{\partial \mathfrak{J}_j}{\partial x_{m\mathrm{f}}} \mu_m \tag{6.68}$$

$$\frac{\partial y}{\partial \sigma} = -[P]_y \tag{6.69}$$

The multipliers Λ_j, $j = 1, \cdots, l$ are determined by the system of linear equations

$$\sum_{j=1}^{l} \Lambda_j \int_{t_0}^{t_\mathrm{f}} \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{J}_j}{\partial x_{m\mathrm{f}}} \mu_m \right) \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{J}_k}{\partial x_{m\mathrm{f}}} \mu_m \right) d\tau$$

$$= -\int_{t_0}^{t_\mathrm{f}} \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{J}_k}{\partial x_{m\mathrm{f}}} \mu_m \right) \left(\sum_{m=1}^{n} \frac{\partial P}{\partial x_{m\mathrm{f}}} \mu_m \right) d\tau, \qquad k = 1, \cdots, l \tag{6.70}$$

This system of equations will have a solution if the matrix of coefficients $A = (a_{kj})$

$$a_{kj} = \int_{t_0}^{t_f} \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \mu_m \right) \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_k}{\partial x_{mf}} \mu_m \right) d\tau, \qquad j, k = 1, \cdots, l \tag{6.71}$$

is nonsingular, and this will be the case if the matrix $B = (b_{kj})$

$$b_{kj} = \sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \frac{\partial \mathfrak{Z}_k}{\partial x_{mf}}, \qquad j, k = 1, \cdots, l \tag{6.72}$$

is nonsingular, i.e., if the boundary conditions are not locally redundant, and, further, if the matrix $C = (c_{mp})$

$$c_{mp} = \int_{t_0}^{t_f} \mu_m \mu_p \, d\tau, \qquad m, p = 1, \cdots, n \tag{6.73}$$

is of rank $r \geq l$.* A proof of this due to Norman Greenspan of Grumman's Research Department is given in Appendix A.

The technique for handling terminal constraints just discussed is essentially that employed by Bryson.[15] It may be termed "gradient projection" by analogy with the technique discussed earlier in connection with ordinary minimum problems. A somewhat related scheme employed by the present writer in the investigation of reference 14 is presented as follows.

If in lieu of introduction of the multipliers Λ_j, the control variable y is broken down as

$$y(t, \sigma) = \phi(t, \sigma) + \sum_{q=1}^{l} a_q(\sigma) f_q(t) \tag{6.74}$$

then the l constants a_q may be employed for the purposes of satisfying the

---

* The latter test requires that at least l of the n functions μ_m be linearly independent (Courant and Hilbert,[17] pp. 61–62). Two circumstances where this requirement may be violated should be mentioned. The first concerns the case in which the system of differential equation subsidiary conditions (6.34) contains a holonomic condition, i. e., a differential equation obtainable by differentiation of a finite condition relating the variables x_m and t. In such a case the matrix of (6.73) may degenerate in rank for all values of the upper limit of the integrals (see Courant and Hilbert,[17] p. 221). More subtle cases of degeneracy may arise for special values of the upper limit corresponding to other *abnormality* phenomena and to the occurrence of *conjugate points,* as defined in connection with Jacobi's necessary condition.

constraints (6.67) which we rewrite in the form

$$\frac{d\mathfrak{Z}_j}{d\sigma} = \int_{t_0}^{\bar{t}_f} \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \mu_m\right)\left(\frac{\partial \phi}{\partial \sigma} + \sum_{q=1}^{l} \frac{da_q}{d\sigma} f_q\right) d\tau = 0, \qquad j = 1, \cdots, l \quad (6.75)$$

These equations may be rearranged as

$$\sum_{q=1}^{l} \frac{da_q}{d\sigma} \sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \int_{t_0}^{\bar{t}_f} \mu_m f_q \, d\tau = -\sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \int_{t_0}^{\bar{t}_f} \mu_m \frac{\partial \phi}{\partial \sigma} \, d\tau, \qquad j = 1, \cdots, l \quad (6.76)$$

and regarded as a system of simultaneous equations in the $da_q/d\sigma$. This system will have a solution if the matrix $D = (d_{jq})$

$$d_{jq} = \sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \int_{t_0}^{\bar{t}_f} \mu_m f_q \, d\tau, \qquad j, q = 1, \cdots, l \quad (6.77)$$

is nonsingular.

The $f_q(t)$ are arbitrary functions to be chosen so that requirements for the system to have a solution are met. The $da_q/d\sigma$ determined by simultaneous solution of (6.76) may then be substituted into

$$\frac{dP}{d\sigma} = \int_{t_0}^{\bar{t}_f} \left(\sum_{m=1}^{n} \frac{\partial P}{\partial x_{mf}} \mu_m\right)\left(\frac{\partial \phi}{\partial \sigma} + \sum_{q=1}^{l} \frac{da_q}{d\sigma} f_q\right) d\tau \quad (6.78)$$

and the gradient $[P]_\phi$ determined as the collected coefficient of $\partial\phi/\partial\sigma$.

Both of the schemes so far described in this section suffer from "drift" of terminal values, for large steps in $\Delta\sigma$, as a result of the boundary linearizations. The terminal values must be restored in the course of the descent process, and, in the gradient projection case, this is accomplished by reinstatement of the zero-order terms in the linearized constraint expressions (6.65). In the case of the process just described, the coefficients a_q provide a natural choice of parameters by which the necessary adjustments may be performed. However, a difficulty in performing corrections may arise which is associated with the choice of the arbitrary functions f_q, as will be illustrated in an example.

The employment of a "penalty function" scheme has the advantage of "built-in" corrections. The constraints on terminal values will be satisfied only approximately, although within any desired tolerances. One seeks a minimum of

$$P' = P + \tfrac{1}{2} \sum_{j=1}^{l} K_j \mathfrak{Z}_j^2 \quad (6.79)$$

Just as in the ordinary minimum case, the products $K_j\mathfrak{Z}_j$ will satisfy a system resembling that determining the multipliers Λ_j [Eq. (6.70)], namely,

$$\sum_{j=1}^{l} K_j \mathfrak{Z}_j \int_{t_0}^{t_f} \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{m\,f}} \mu_m \right) \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_k}{\partial x_{m\,f}} \mu_m \right) d\tau$$

$$= - \int_{t_0}^{t_f} \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_k}{\partial x_{m\,f}} \mu_m \right) \left(\sum_{m=1}^{n} \frac{\partial P}{\partial x_{m\,f}} \mu_m \right) d\tau, \qquad k = 1, \cdots, l \quad (6.80)$$

where the μ functions and the various partial derivatives are evaluated at a minimum of (6.79), for which the boundary constraints are satisfied only approximately. If the matrix whose coefficients are given by (6.71) is nonsingular, the errors in the terminal constraint equations (6.64) may be reduced to within desired tolerances by appropriate adjustment of the K_j as described in Section 6.14 in connection with ordinary minimum problems.

The expression for the gradient $[P']_y$ takes a form similar to (6.68), namely,

$$[P']_y = \sum_{m=1}^{n} \frac{\partial P}{\partial x_{m\,f}} \mu_m + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{m\,f}} \mu_m \qquad (6.81)$$

From the vanishing of $[P']_y$ along the path approached in the limit of the descent process, we may estimate the effects of small changes in specified terminal values of the $\mathfrak{Z}_j$. In fact, for an *arbitrary* small variation in the control variable $\delta y(\tau)$, we may obtain the relationship

$$\delta P = - \sum_{j=1}^{l} K_j \mathfrak{Z}_j \, \delta \mathfrak{Z}_j \qquad (6.82)$$

by multiplication of (6.81) by δy and integration from t_0 to t_f. Thus the quantities $-K_j\mathfrak{Z}_j$ play the role of trade-off slopes in penalty function computations as do the quantities $-\Lambda_j$ in the gradient projection case [see Eq. (6.68)]. This remarkable property of optimal paths which holds for small variations in their neighborhood has been noted by Cicala[29] in connection with the indirect theory.

6.26 Variable Terminal Time: Boundary Values as Parameters

The way in which the terminal value of the independent variable time, t_f, enters into problems of the type presently under discussion is such as to require special treatment if it is not fixed. If, in a particular application, one of the variables x_m behaves monotonically and has fixed initial and terminal values, then its adoption as independent variable provides a simple means of avoiding complication. Such a choice may very often not be available, and in these circumstances the matter is one to be left to the ingenuity of the individual investigator according to the application and the version of gradient scheme adopted. A particular scheme suited to the technique involving arbitrary functions discussed in the previous section will be illustrated later in an example. In this section we will present a treatment of t_f as a free parameter in connection with the use of a penalty function for handling terminal constraints. It is convenient to consider simultaneously a similar means of handling unspecified initial values.

In order to avoid undue complication in the expressions to follow, we will assume that the initial value of time, t_0, is fixed and that initial conditions on certain of the variables x_m are given as fixed values, the remainder being parameters free for optimization purposes. We proceed to determine the derivative with respect to σ of the performance function given by (6.79) as follows:

$$\frac{dP'}{d\sigma} = \sum_{m=1}^{n} \frac{\partial P}{\partial x_{m\mathrm{f}}} \frac{dx_{m\mathrm{f}}}{d\sigma} + \frac{\partial P}{\partial t_\mathrm{f}} \frac{dt_\mathrm{f}}{d\sigma} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \left(\sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{m\mathrm{f}}} \frac{dx_{m\mathrm{f}}}{d\sigma} + \frac{\partial \mathfrak{Z}_j}{\partial t_\mathrm{f}} \frac{dt_\mathrm{f}}{d\sigma} \right) \quad (6.83)$$

Making use of the expressions (6.50), we expand this to the following:

$$\begin{aligned} \frac{dP'}{d\sigma} = & \int_{t_0}^{\bar{t}_\mathrm{f}} \left[\sum_{m=1}^{n} \left(\frac{\partial P}{\partial x_{m\mathrm{f}}} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial x_{m\mathrm{f}}} \right) \mu_m \right] \frac{\partial y}{\partial \sigma} \, d\tau \\ & + \left[\sum_{m=1}^{n} \left(\frac{\partial P}{\partial x_{m\mathrm{f}}} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial x_{m\mathrm{f}}} \right) \bar{g}_{m\mathrm{f}} + \frac{\partial P}{\partial t_\mathrm{f}} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial t_\mathrm{f}} \right] \frac{dt_\mathrm{f}}{d\sigma} \\ & + \sum_{p=1}^{n} \left[\sum_{m=1}^{n} \left(\frac{\partial P}{\partial x_{m\mathrm{f}}} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial x_{m\mathrm{f}}} \right) \xi_{mp}(\bar{t}_\mathrm{f}) \right] \frac{dx_{p_0}}{d\sigma} \quad (6.84) \end{aligned}$$

(The summation over p is to be understood to range over only those initial values x_{p_0} which are unspecified.)

We are now in a position to identify a "mixed" gradient direction on the basis of the coefficients of the various σ derivatives appearing in (6.84). We set

$$\frac{\partial y}{\partial \sigma} = -[P']_y = -\left[\sum_{m=1}^{n} \left(\frac{\partial P}{\partial x_{m\mathrm{f}}} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial x_{m\mathrm{f}}} \right) \mu_m \right] \quad (6.85)$$

$$\begin{aligned} \frac{dt_\mathrm{f}}{d\sigma} = -[P']_{t_\mathrm{f}} = & -\left[\sum_{m=1}^{n} \left(\frac{\partial P}{\partial x_{m\mathrm{f}}} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial x_{m\mathrm{f}}} \right) \bar{g}_{m\mathrm{f}} \right. \\ & \left. + \frac{\partial P}{\partial t_\mathrm{f}} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial t_\mathrm{f}} \right] \quad (6.86) \end{aligned}$$

$$\frac{dx_{p_0}}{d\sigma} = -[P']_{x_{p_0}} = -\left[\sum_{m=1}^{n} \left(\frac{\partial P}{\partial x_{m\mathrm{f}}} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial x_{m\mathrm{f}}} \right) \xi_{mp}(\bar{t}_\mathrm{f}) \right] \quad (6.87)$$

and, in the course of the descent process, we change the final time t_f and the free initial values x_{p_0}, along with the function y, linearly with increments in the descent parameter σ according to the slopes given by these expressions. In such a fashion we combine the features of variational problems and ordinary minimum problems in a single gradient optimization process.

A related scheme for determination of optimal t_f is employment of the vanishing of the expression (6.86) as run termination criterion. This may be viewed as a one-dimensional search for a minimum of P' versus t_f which may be performed simultaneously with numerical integration of trajectories.

We note again in passing that the quantities $\xi_{mp}(\bar{t}_f)$, the partial derivatives of the terminal x_m values with respect to initial x_m values are conveniently evaluated from unit solutions of the adjoint system of equations as previously discussed in Section 6.23.

6.27 Optimization with Respect to Configuration and System Parameters

In many practical engineering applications, optimal performance is sought not only in terms of flight path selection but also in terms of vehicle and system parameters. We have, for clarity, avoided complicating the preceding analytical work by such considerations. It is of interest to note, however, that such parameters may conveniently be handled as initial conditions by means of the following artifice which is due to Cicala.[29]

We characterize the parameters e_i as initial values of additional system variables x_i which are governed by the differential equations

$$\begin{aligned} \dot{x}_i &= 0 \\ x_i(t_0) &= e_i \end{aligned} \qquad i = n + 1, \cdots \tag{6.88}$$

Considering these equations as additional members of the basic system (6.34), we may obtain the partial derivatives of terminal x_m values with respect to the e_i in terms of unit solutions of the (now expanded) adjoint system as previously described. This offers a convenient means compatible with the computational scheme previously suggested for simultaneous treatment of variational and ordinary minimum problems.

6.28 Inequality Constraints

If to the problem statement of Section 6.21 one or more requirements in the form of inequality constraints on the variables x_m, y, and t are added,

$$Q_i(x_1, \cdots, x_n, y, t) \leq 0 \tag{6.89}$$

the problem becomes of nonclassical type. If the control variable y appears in the expressions Q_i, certain techniques developed along the lines of the "indirect" variational method are applicable, namely, the techniques of Valentine[30] and Pontryagin,[31] discussed in other chapters. Relatively little theory is available, however, for treatment of cases in which the functions Q_i do not depend on the control variable y. Such cases are of great practical interest, e.g., in connection with air vehicle flight paths subject to a minimum altitude limit and to air speed/altitude envelope boundaries arising from structural and power plant limitations.

Examining first the situation where the control variable y appears in the Q_i, we assume that the inequalities are of the form

$$y_1 \leq y \leq y_2 \tag{6.90}$$

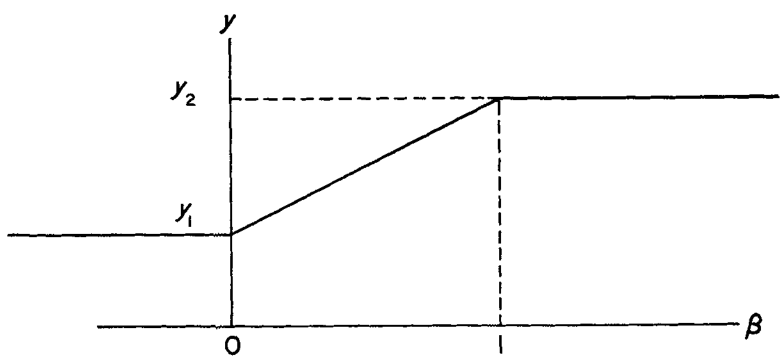
which is usually the case in applications. We now introduce a parameter $\beta(t, \sigma)$ as a new control variable by defining a function $y(\beta)$ as

$$y = y_1, \qquad \beta \leq 0 \tag{6.91}$$

$$y = y_1 + (y_2 - y_1)\beta, \qquad 0 \leq \beta \leq 1 \tag{6.92}$$

$$y = y_2, \qquad 1 \leq \beta \tag{6.93}$$

This is shown in the following sketch.



We may now apply the theory developed earlier in the absence of any constraint on the variable β. As a result of the transformation from $y(t, \sigma)$ to $\beta(t, \sigma)$ as control variable, the integrand typically arising in the integral expression for $dP/d\sigma$ becomes

$$[P]_y \frac{\partial y}{\partial \sigma} = [P]_y \frac{dy}{d\beta} \frac{\partial \beta}{\partial \sigma} \tag{6.94}$$

and motion in the negative gradient direction is determined by

$$\frac{\partial\beta}{\partial\sigma} = -[P]_y \frac{dy}{d\beta} \tag{6.95}$$

The derivative $dy/d\beta$ is given by the expression

$$\frac{dy}{d\beta} = (y_2 - y_1)[H(\beta) - H(\beta - 1)] \tag{6.96}$$

where H is the Heaviside unit step function [$H(\alpha) = 0$ for $\alpha < 0$, $H(\alpha) = 1$ for $\alpha > 0$]. The derivative is undefined at the points $\beta = 0$ and $\beta = 1$. These ambiguities may be removed conveniently by setting

$$\frac{dy}{d\beta} = (y_2 - y_1)H(-[P]_y), \qquad \beta = 0 \tag{6.97}$$

$$\frac{dy}{d\beta} = (y_2 - y_1)H([P]_y), \qquad \beta = 1 \tag{6.98}$$

which decides the question on whether or not the negative gradient direction leads into or out of the interval $0 \leq \beta \leq 1$.

In a continuous descent process, such a formulation will succeed in holding the control variable y in the desired region $y_1 \leq y \leq y_2$ and the parameter β will automatically remain in the region $0 \leq \beta \leq 1$. There will be difficulty, however, with a stepwise version, since the right-hand member of (6.95) is evaluated along the solution $x = \bar{x}(t)$, $y = \bar{y}(t)$ and with finite step size $\Delta\sigma$, the control parameter β will not, in general, remain within limits. Consequently, it will be necessary before each calculation of $\partial\beta(t)/\partial\sigma$ to alter the function $\beta(t)$ obtained in the course of the preceding descent computations to conform to the inequality $0 \leq \beta \leq 1$; otherwise the right-hand member of (6.95) will vanish in all subsequent computations at points for which the inequality is violated.

Constraints $Q_i \leq 0$ which are independent of y may be handled approximately by construction of suitable penalty functions. If an additional variable x_i is introduced in connection with each constraint according to the equations

$$\begin{aligned} \dot{x}_i &= \tfrac{1}{2}Q_i^2 H(Q_i) \\ x_i(t_0) &= 0 \end{aligned} \qquad i = n+1, n+2, \cdots, s \tag{6.99}$$

where H is the Heaviside unit step function, then the terminal values of these x_i will be

$$x_i(t_f) = \tfrac{1}{2}\int_{t_0}^{t_f} Q_i^2 H(Q_i)\, dt, \qquad i = n+1, n+2, \cdots, s \tag{6.100}$$

representing integral squares of the "violations" taken over those segments of the flight path which violate the constraints. Appropriate penalty terms proportional to these terminal values may then be added to Eq. (6.79), giving it the form

$$P' = P + \tfrac{1}{2} \sum_{j=1}^{l} K_j \mathfrak{Z}_j^2 + \sum_{i=n+1}^{s} K_i x_i(t_f), \qquad l < n \tag{6.101}$$

Evidently "tolerances" for comparison with the x_i terminal values and adjustment of constants K_i must be specified in similar terms. If ϵ_i is a specified error tolerance in the sense of an rms average over instantaneous values, then the appropriate "tolerance" for comparison with x_{if} is

$$E_i = \tfrac{1}{2} \int_{t_0}^{t_f} \epsilon_i^2 H(Q_i)\, dt \tag{6.102}$$

The ratio $x_i(t_f)/E_i$ will then serve as a "violation"/"tolerance" index for adjustment of K_i in the course of the descent process.

6.29 Penalty Functions: Error Estimates

In connection with computer mechanization of the penalty function scheme, it is useful to have available a means for estimation of the error in the functional P in terms of the constraint "violations." Thus those "violations" which have a particularly strong effect on P may be noted and a more rational basis provided for the establishment of the "tolerances" discussed previously.

The increment in P due to small changes in terminal values of the x_m and t is given to first order as

$$\begin{aligned} \Delta P &= \sum_{m=1}^{n} \frac{\partial P}{\partial x_{mf}} \delta x_{mf} + \frac{\partial P}{\partial t_f} \delta t_f \\ &= \int_{t_0}^{\bar{t}_f} \left(\sum_{m=1}^{n} \frac{\partial P}{\partial x_{mf}} \mu_m \right) \delta y \, d\tau + \left(\frac{\partial P}{\partial t_f} + \sum_{m=1}^{n} \frac{\partial P}{\partial x_{mf}} \bar{g}_{mf} \right) \delta t_f \end{aligned} \tag{6.103}$$

Assuming that a minimum of P' as given by (6.101) has been attained for certain fixed values of the penalty constants, we have the following equation for the vanishing of the gradient $[P']_y$:

$$[P']_y = \sum_{m=1}^{n} \frac{\partial P}{\partial x_{mf}} \mu_m + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \mu_m + \sum_{i=n+1}^{s} K_i \mu_i = 0 \tag{6.104}$$

From the condition t_f open,

$$[P']_{t_f} = \sum_{m=1}^{n} \left(\frac{\partial P}{\partial x_{mf}} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \right) \bar{g}_{mf} + \sum_{i=n+1}^{s} K_i \bar{g}_{if} + \frac{\partial P}{\partial t_f} + \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial t_f} = 0 \quad (6.105)$$

Making use of these expressions, we obtain for ΔP

$$\Delta P = -\int_{t_0}^{\bar{t}_f} \left(\sum_{j=1}^{l} K_j \mathfrak{Z}_j \sum_{m=1}^{n} \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \mu_m + \sum_{i=n+1}^{s} K_i \mu_i \right) \delta y \, d\tau - \left[\sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial t_f} + \sum_{m=1}^{n} \sum_{j=1}^{l} K_j \mathfrak{Z}_j \frac{\partial \mathfrak{Z}_j}{\partial x_{mf}} \bar{g}_{mf} + \sum_{i=n+1}^{s} K_i \bar{g}_{if} \right] \delta t_f$$

$$= -\sum_{j=1}^{l} K_j \mathfrak{Z}_j \, \Delta \mathfrak{Z}_j - \sum_{i=n+1}^{s} K_i \, \Delta x_{if} \quad (6.106)$$

If we now consider corrections in control $\delta y(\tau)$ and in terminal time δt_f such as to produce increments

$$\Delta \mathfrak{Z}_j = -\mathfrak{Z}_j \quad (6.107)$$

$$\Delta x_{if} = -x_{if} \quad (6.108)$$

which will null the "violations," then the resulting increase in P may be estimated as

$$\Delta P = \sum_{j=1}^{l} K_j \mathfrak{Z}_j^2 + 2 \sum_{i=n+1}^{s} K_i x_{if} \quad (6.109)$$

where the factor of 2 in the second member is introduced to account approximately for the square-law behavior of the integrand in (6.100) near $Q_i = 0$.

The increment in P may be conveniently evaluated in terms of known terminal values. The contributions of the various components of ΔP provide a basis for the setting of "tolerances" and hence for the adjustment of the K_j and K_i. In terms of P and P' evaluated for finite K_j and K_i, a "best" estimate of the limiting value of P is

$$P + \Delta P = P + 2(P' - P) = 2P' - P \quad (6.110)$$

as obtained from (6.101) and (6.109).

6.3 Solar Sailing Example

For the purpose of exploring the computational aspect of the gradient optimization technique, we have chosen a planar case of transfer between planetary orbits by means of the interesting solar sailing scheme. The potential capabilities of solar sail propulsion have been investigated in the papers of Garwin,[32] Cotter,[33] Tsu,[34] and London.[35] This problem has the simplicity appropriate to an exploration of method, yet sufficient complexity to render analytical solution quite difficult unless drastic simplifications are introduced.

6.31 System Equations

The equations of motion and kinematic relations are given in a notation nearly the same as that of Tsu.[34] With reference to the schematic of Fig. 2, these are as follows:

Radial acceleration

$$\dot{u} = g_1 = \frac{v^2}{R} - A_0\left(\frac{R_0}{R}\right)^2 + \alpha\left(\frac{R_0}{R}\right)^2 |\cos^3\theta| \tag{6.111}$$

Circumferential acceleration

$$\dot{v} = g_2 = -\frac{uv}{R} - \alpha\left(\frac{R_0}{R}\right)^2 \sin\theta\cos^2\theta \tag{6.112}$$

Radial velocity

$$\dot{R} = g_3 = u \tag{6.113}$$

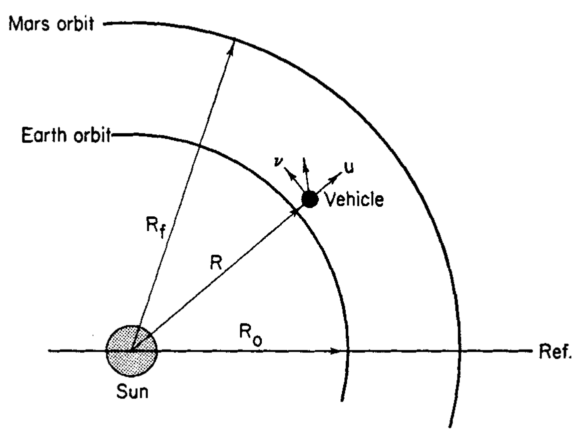


FIG. 2. Orbital transfer schematic.

Circumferential angular velocity

$$\dot{\psi} = \frac{v}{R} \tag{6.114}$$

Since the heliocentric angle ψ does not appear in the first three equations, nor will it appear in the statements of boundary conditions to be considered, Eq. (6.114) may be ignored for purposes of gradient optimization. This amounts to an assumption that terminal matching of the heliocentric angles of vehicle and "target" planet is accomplished by selection of launch time.

6.32 Boundary Values

Seeking minimum-time transfer, we identify the functional P as

$$P = t_f \tag{6.115}$$

The functions u, v, R are the variables x_m of the theoretical development, and the sail angle θ appears in the role of the control variable y.

As initial conditions we specify velocity components u, v, and radius R corresponding to motion in the earth's orbit approximated as a circle:

$$t_0 = 0 \tag{6.116}$$

$$u(0) = u_0 = u_E = 0 \tag{6.117}$$

$$v(0) = v_0 = v_E \tag{6.118}$$

$$R(0) = R_0 = R_E \tag{6.119}$$

We consider terminal conditions corresponding to arrival at the orbit of the planet Mars (also taken as a circle) with prescribed velocity components:

$$u(t_f) = u_f \tag{6.120}$$

$$v(t_f) = v_f \tag{6.121}$$

$$R(t_f) = R_f = R_M \tag{6.122}$$

6.33 Correction Functions

For fixed boundary values of u, v, and R, the equations corresponding to Eqs. (6.39) of the preceding theoretical development are

$$\delta u_f = \int_0^{\bar{t}_f} \mu_1 \, \delta\theta \, d\tau + \bar{g}_{1f} \, \delta t_f = 0 \tag{6.123}$$

$$\delta v_f = \int_0^{\bar{t}_f} \mu_2 \, \delta\theta \, d\tau + \bar{g}_{2f} \, \delta t_f = 0 \tag{6.124}$$

$$\delta R_f = \int_0^{\bar{t}_f} \mu_3 \, \delta\theta \, d\tau + \bar{g}_{3f} \, \delta t_f = 0 \tag{6.125}$$

In this case the number of functions f_q and coefficients a_q required is

$$r = n - 1 - s = 3 - 1 - 0 = 2 \tag{6.126}$$

We select the functions f_q as

$$f_1(t) = t \tag{6.127}$$

$$f_2(t) = t^2 \tag{6.128}$$

and the control function $\theta(t)$ is broken down as

$$\theta = \phi + a_1 t + a_2 t^2 \tag{6.129}$$

The system of equations (6.123)–(6.125) becomes

$$\left[\int_0^{\bar{t}_f} f_1(\tau)\mu_1 \, d\tau\right] \frac{da_1}{d\sigma} + \left[\int_0^{\bar{t}_f} f_2(\tau)\mu_1 \, d\tau\right] \frac{da_2}{d\sigma} + \bar{g}_{1f} \frac{dt_f}{d\sigma} = -\int_0^{\bar{t}_f} \mu_1 \frac{\partial\phi}{\partial\sigma} \, d\tau \tag{6.130}$$

$$\left[\int_0^{\bar{t}_f} f_1(\tau)\mu_2 \, d\tau\right] \frac{da_1}{d\sigma} + \left[\int_0^{\bar{t}_f} f_2(\tau)\mu_2 \, d\tau\right] \frac{da_2}{d\sigma} + \bar{g}_{2f} \frac{dt_f}{d\sigma} = -\int_0^{\bar{t}_f} \mu_2 \frac{\partial\phi}{\partial\sigma} \, d\tau \tag{6.131}$$

$$\left[\int_0^{\bar{t}_f} f_1(\tau)\mu_3 \, d\tau\right] \frac{da_1}{d\sigma} + \left[\int_0^{\bar{t}_f} f_2(\tau)\mu_3 \, d\tau\right] \frac{da_2}{d\sigma} + \bar{g}_{3f} \frac{dt_f}{d\sigma} = -\int_0^{\bar{t}_f} \mu_3 \frac{\partial\phi}{\partial\sigma} \, d\tau \tag{6.132}$$

This 3 × 3 case may conveniently be inverted analytically. There seems little point in listing the inverse elements here, however.

$$\begin{bmatrix} \dfrac{da_1}{d\sigma} \\ \dfrac{da_2}{d\sigma} \\ \dfrac{dt_f}{d\sigma} \end{bmatrix} = - \begin{bmatrix} \\ C \\ \\ \end{bmatrix} \begin{bmatrix} \displaystyle\int_0^{\bar{t}_f} \mu_1 \frac{\partial\phi}{\partial\sigma}\, d\tau \\ \displaystyle\int_0^{\bar{t}_f} \mu_2 \frac{\partial\phi}{\partial\sigma}\, d\tau \\ \displaystyle\int_0^{\bar{t}_f} \mu_3 \frac{\partial\phi}{\partial\sigma}\, d\tau \end{bmatrix} \tag{6.133}$$

The slope of descent $dt_f/d\sigma$ is given by

$$\frac{dt_f}{d\sigma} = - \int_0^{\bar{t}_f} [C_{31}\mu_1 + C_{32}\mu_2 + C_{33}\mu_3] \frac{\partial\phi}{\partial\sigma}\, d\tau \tag{6.134}$$

and the gradient of P by

$$[P]_\phi = -[C_{31}\mu_1 + C_{32}\mu_2 + C_{33}\mu_3] \tag{6.135}$$

Accordingly, we set

$$\frac{\partial\phi}{\partial\sigma} = -[P]_\phi = C_{31}\mu_1 + C_{32}\mu_2 + C_{33}\mu_3 \tag{6.136}$$

and proceed with stepwise descent.

A particularly suitable case for a first illustration of computational technique is the one in which terminal velocity components are unspecified—"free" boundary conditions. Here Eqs. (6.123) and (6.124) may be deleted and

$$r = n - 1 - s = 3 - 1 - 2 = 0 \tag{6.137}$$

so that no functions f_q are needed. Hence

$$\theta = \phi \tag{6.138}$$

and

$$\frac{\partial\phi}{\partial\sigma} = \frac{\mu_3}{\bar{g}_{3f}} \tag{6.139}$$

$$\frac{dt_f}{d\sigma} = -\frac{1}{\bar{g}_{3f}^2} \int_0^{\bar{t}_f} \mu_3^2\, d\tau \tag{6.140}$$

in this case.

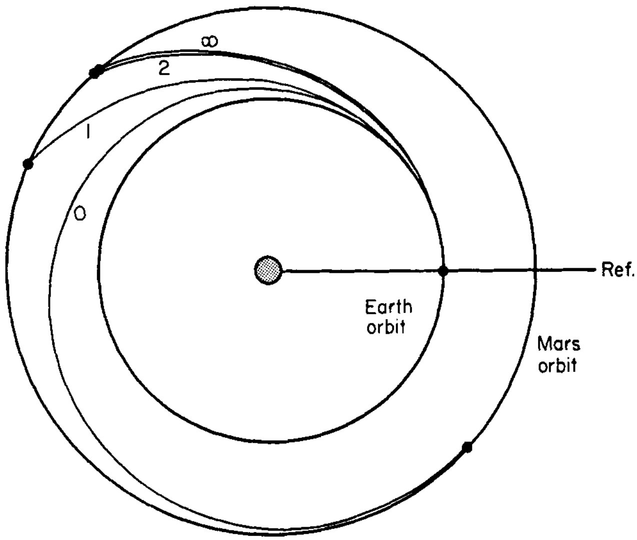


FIG. 3. Successive approximations to optimal transfer path, terminal velocity components open.

Computations have employed numerical values of the various constants from Tsu's paper, with

$$\alpha = 0.1 \text{ cm/sec}^2 = 3.28 \times 10^{-3} \text{ ft/sec}^2 \tag{6.141}$$

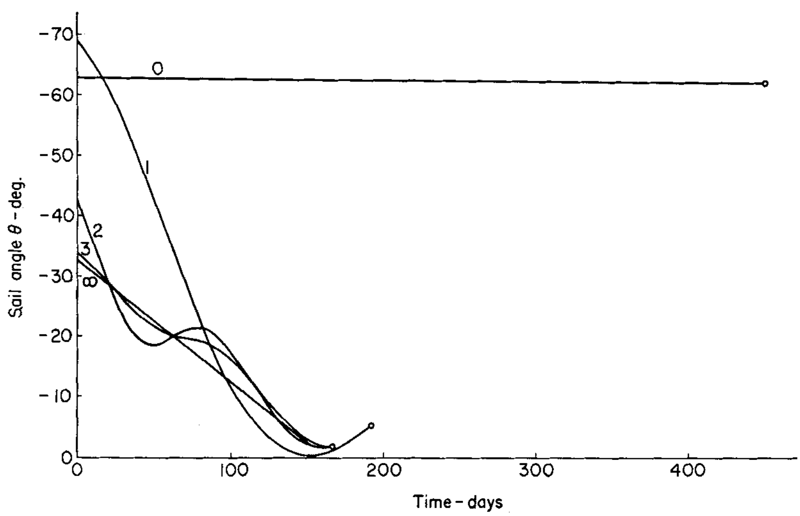


FIG. 4. Successive approximations to optimal sail angle program, terminal velocity components open.

This value corresponds to about $10^{-4}g$ thrust acceleration developed by the sail when oriented broadside to the sun ($\theta = 0$) at earth's orbit radius, or about 17% of the sun's gravitational attraction.

6.34 Orbit Transfer Computations

Results of descent computations for the case of "open" terminal velocity components are shown in Figs. 3–5. The control program of the original

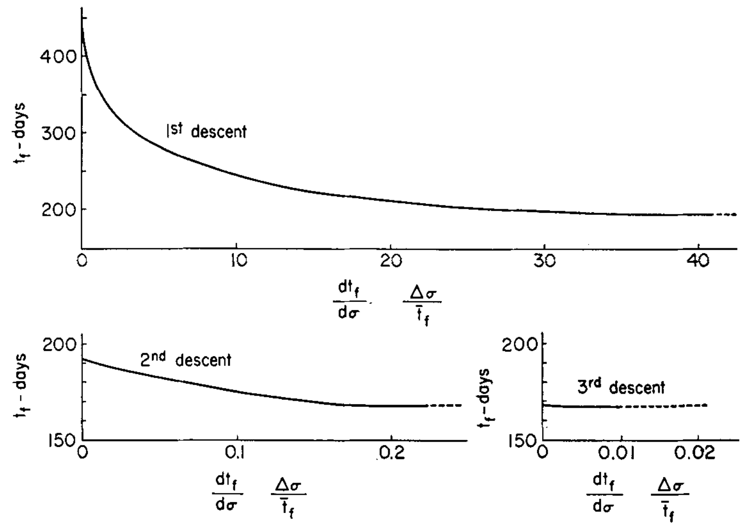


FIG. 5. Descent curves—solar sail transfer, terminal velocity components open.

flight path (Figs. 3 and 4), chosen arbitrarily, was far from optimal in that the radial velocity component at crossing of Mars orbit was small. The greatest reduction in flight time—more than half of the original—is seen to be obtained in the course of the first descent (Fig. 5). In three descents minimum flight time has been attained for practical purposes, although small changes in the detailed structure of the control program are still in evidence.

Results for the case of terminal velocity components matched to the target planet

$$u_f = u_M, \; v_f = v_M \tag{6.142}$$

are presented in Figs. 6–9. The tendency of the terminal values to depart from the prescribed values is shown in Fig. 6. These were restored via an iterative correction process employing increments in the coefficients a_1

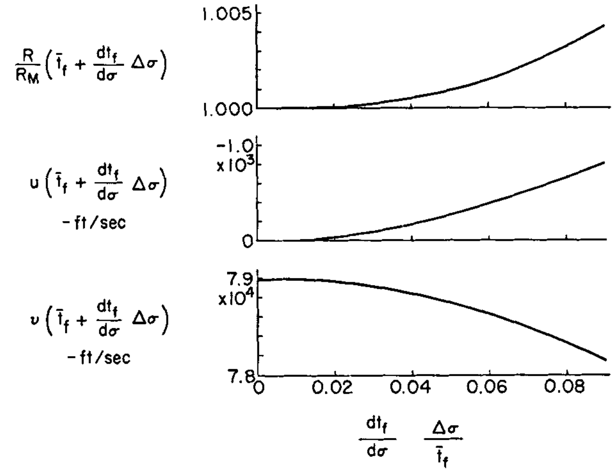


FIG. 6. Departure of terminal values, "matched" terminal velocity components.

and a_2 of Eq. (6.129). Typically, two or three iteration cycles were required to correct each point. Descent curves are shown in Fig. 7. The approach to the minimum-time solution is depicted in Figs. 8 and 9.

6.35 Convergence Considerations

The first attempt at computations for this "matched velocity" case employed functions f_q constant and linear with time. This met with near-

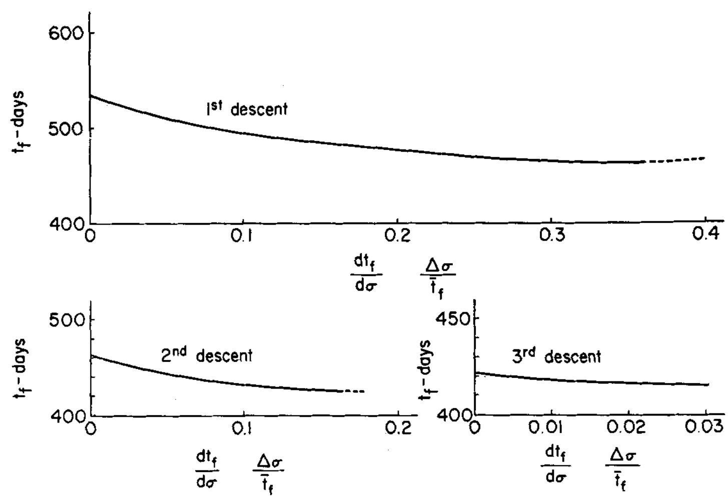


FIG. 7. Descent curves—solar sail transfer, "matched" terminal velocity components.

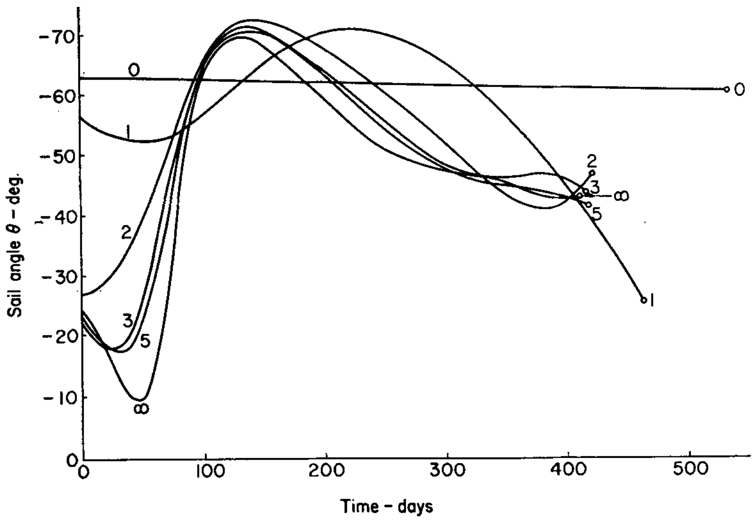


FIG. 8. Successive approximations to optimal sail angle program, "matched" terminal velocity components.

zero determinant difficulty, whereas the combination of linear and square-law corrections indicated above was successful in avoiding this difficulty.

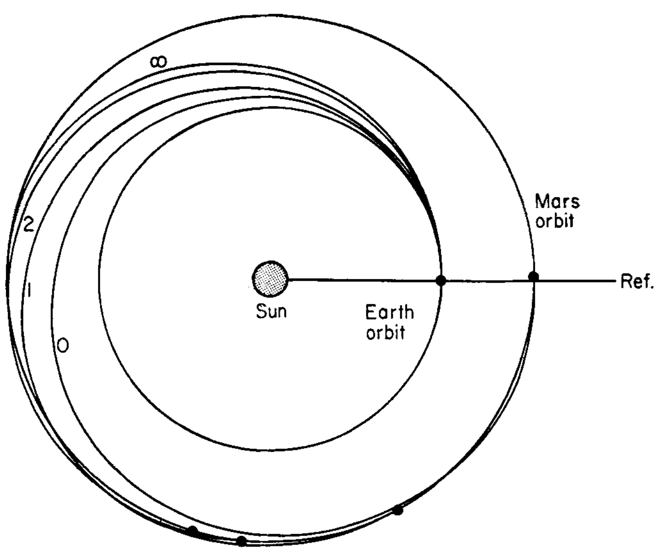


FIG. 9. Successive approximations to optimal transfer path, "matched" terminal velocity components.

A subsequent examination of the behavior of the determinant in question during the course of the descent process was instructive on the matter of choice of correction functions. Values of the determinant of the matrix on the left of Eqs. (6.130)–(6.132), which is required in calculation of the inverse matrix C of Eq. (6.133), were tabulated for the first few descent steps and for the converged trajectory as shown in Table I.

TABLE I

DETERMINANT BEHAVIOR

Descent no. i	Case: $f_1 = 1$, $f_2 = t$ Δ_i/Δ_∞	Case: $f_1 = t$, $f_2 = t^2$ Δ_i/Δ_∞
1	−2.383	3.983
2	−2.695	3.268
3	−0.770	1.527
4	−0.340	1.104
∞	1.000	1.000

A check showed that the numerical values of both determinants evaluated for the converged trajectory were large enough to permit terminal value adjustments without difficulty of ill-conditioning. Thus the convergence of a descent process which employs correction functions depends upon a choice of functions for which the determinant may not change sign. With the benefit of hindsight gained from this experiment, we observe that a choice of influence functions μ_m in the role of correction functions f_q appears attractive in that the determinant in question will be well behaved except in the special circumstances noted in the footnote of Section 6.25.

6.4 Low-Thrust Example

6.41 System Equations and Boundary Values

A second, closely related, example for which numerical computations have been performed concerns orbit transfer by means of low-thrust propulsion, e.g., ion rocket or plasma jet. In this case the penalty function technique for treatment of terminal constraints is adopted. The results of this section are due to Lindorfer and Moyer.[36]

The equations of motion and kinematic relations resemble those for the solar sail example, with differences appearing only in the expressions for thrust components:

Radial acceleration

$$\dot{u} = g_1 = \frac{v^2}{R} - A_0\left(\frac{R_0}{R}\right)^2 + \frac{T}{m}\sin\theta \tag{6.143}$$

Circumferential acceleration

$$\dot{v} = g_2 = -\frac{uv}{R} + \frac{T}{m}\cos\theta \tag{6.144}$$

Radial velocity

$$\dot{R} = g_3 = u \tag{6.145}$$

Circumferential angular velocity

$$\dot{\psi} = g_4 = \frac{v}{R} \tag{6.146}$$

The thrust magnitude T has been taken as constant at a value corresponding to an initial thrust acceleration

$$\frac{T}{m_0 g} = 0.846 \times 10^{-4} \tag{6.147}$$

and the mass has been assumed to decrease linearly with time

$$m = m_0 - Qt \tag{6.148}$$

at a rate corresponding to

$$\frac{Q}{m_0} = 1.29 \times 10^{-3} \text{ per day} \tag{6.149}$$

These estimates are taken from the paper of Edwards and Brown.[37] The thrust direction θ is measured from the circumferential direction in the above.

Boundary conditions considered in the computations are those for the "matched velocity" case of transfer from earth's orbit to the Martian orbit. These orbits are idealized as circular as in the example of the preceding section.

6.42 Orbit Transfer and Rendezvous Computations

The orbit transfer case, ψ_f open, was examined first. The convergence of the successive approximation process was fairly rapid. Although the number of descents required was larger than in the solar sail example of the preceding section, the absence of any need for correction of the "drift" exhibited in Fig. 6 resulted in an over-all reduction in computation time. Experimentation with penalty constant adjustment schemes indi-

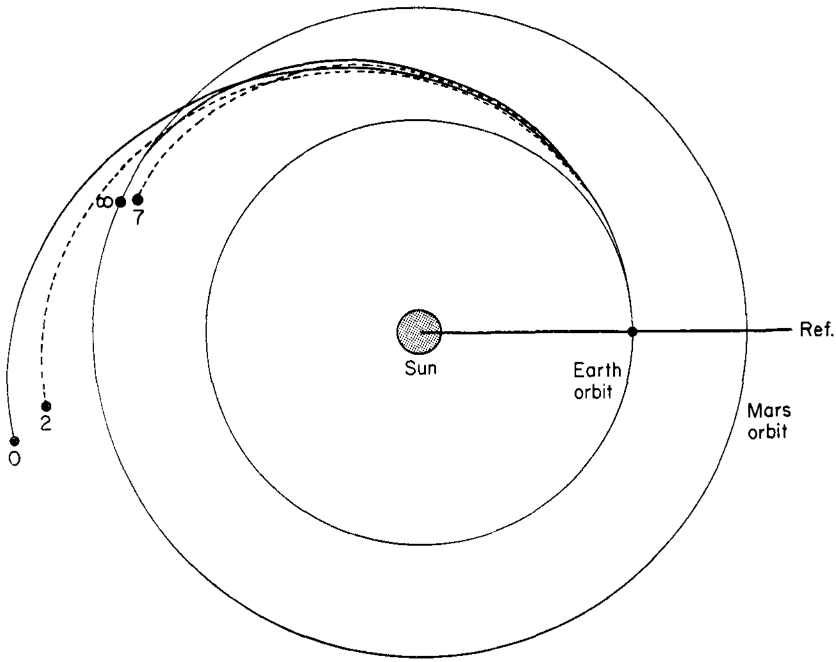


FIG. 10. Successive approximations to optimal low-thrust transfer, terminal heliocentric angle open.

cated that convergence is improved if the values of the penalty constants are initially taken small—"loose tolerances"—and the descent process allowed to proceed through a number of steps before the tolerances are "tightened." Thus in Figs. 10 and 11 the first 25 descent steps were performed with somewhat low values of the K_j and a minimum of P' approached. The K_j were then increased in proportion to the absolute value of error/tolerance ratio for each variable (radial velocity excepted, since this was employed as run termination variable) and the process continued as shown.

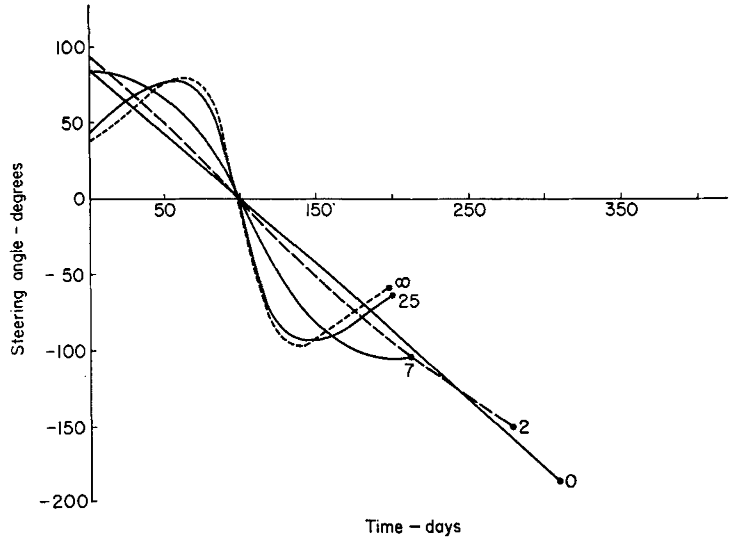


FIG. 11. Successive approximations to optimal low-thrust direction program, terminal heliocentric angle open.

For the planetary rendezvous problem it is required that the heliocentric angle of the vehicle, ψ, match that of Mars at the terminal point. Figures 12 and 13 present results obtained for various assumed initial configurations of the vehicle(earth)-Mars system. Minimum transfer time is plotted

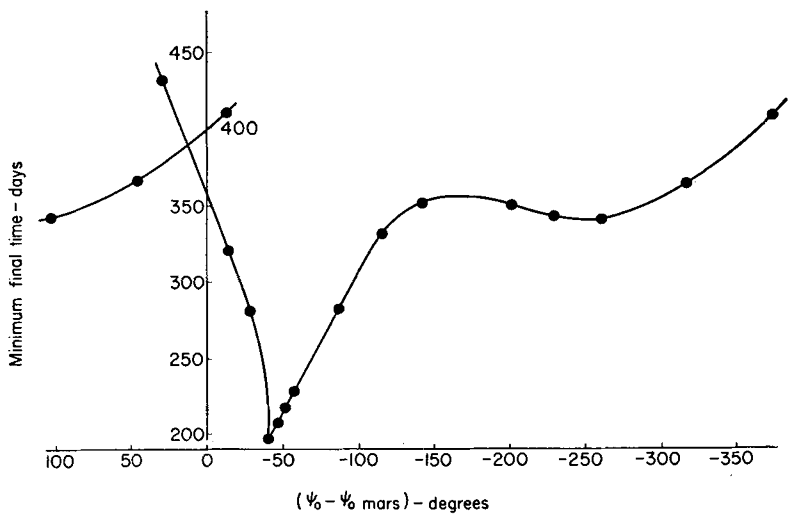


FIG. 12. Minimum time for orbital rendezvous.

against the initial configuration angle in Fig. 12. The lowest point on this curve corresponds to the minimum time for the orbit transfer case (terminal angle open). On the basis of the few data points shown, the minimum appears to have a cusplike character.

This particular point also divides the solutions to the rendezvous problem into two operationally distinct classes. One class ($\psi_0 - \psi_{0\,\text{Mars}} > -46°$) consists of solutions in which the vehicle "waits" for Mars to overtake it. To accomplish this, the vehicle flies out past the Martian orbit, decreasing

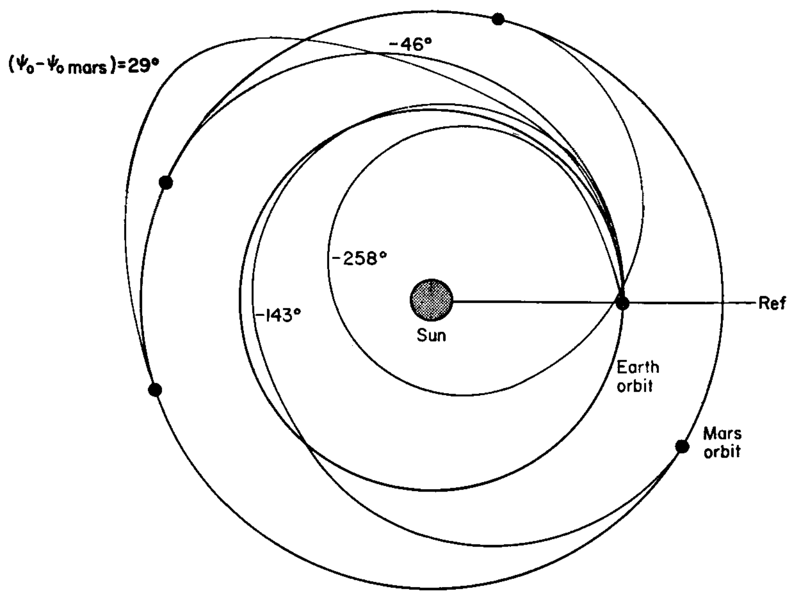


FIG. 13. Optimal transfer paths for rendezvous, various initial configurations.

circumferential velocity until it becomes lower than Martian circular velocity. As the relative heliocentric angle decreases, the vehicle's circumferential velocity is increased to match that of Mars.

The second class of solutions ($\psi_0 - \psi_{0\,\text{Mars}} < -46°$) is characterized by an initial inward motion toward the sun resulting in velocity build-up and final tangential approach to the Mars orbit from within.

The results of Fig. 12 indicate that transit time as a function of launch configuration of the planets exhibits a fairly sharp minimum, and that on a majority of "unfavorable" launch dates the technique of pursuit from behind permits faster transit than the "waiting" technique.

6.5 Remarks on the Relative Merits of Various Computational Techniques

The reader may have correctly inferred that the present chapter amounts to a status report on research still underway rather than a definitive exposition of a standard technique. Indeed, the diversity of viewpoints in the various chapters of this volume serves to indicate the current general uncertainty regarding the applicability and relative merits of the numerous schemes in actual numerical computations. Accordingly, the remarks of the present section will be qualitative, provisional, and confined to the gradient and "indirect" methods with which the author has firsthand experience.

The writer has occasionally been asked to compare gradient methods with the classical "indirect" method (numerical solution of the Euler-Lagrange equations) in regard to speed and ease of computation. It is not easy to draw such a comparison since gradient methods have a "hammer-and-tongs" character while the numerical solution of the two-point boundary-value problem for the Euler equations is very much an art. In problems simple enough to yield to a survey of a one- or two-parameter family of Euler solutions, the classical approach is effective; yet there are few problems of practical interest which can be manipulated into so simple a form. Treatments of the more complicated problems reported in the literature have centered on determination of the mapping between initial and final values, either by mechanized iterative procedures or by extensive cross-plotting of boundary values. This sort of process may be complicated, depending on the particular application, by extremely high sensitivity to small changes in initial values and/or by encounter with near-singular matrices governing successive adjustments. Even the most successful procedures depend upon first obtaining a trial Euler solution whose terminal values lie somewhere in the vicinity of those specified; and the preliminary search for such a solution may consume much time and effort.

It seems possible and even likely that improved means for solution of the mapping problem will materialize. The state-of-the-art in nonlinear differential equations is perhaps not very encouraging, however, in regard to the early development of a powerful general scheme.

Another approach based upon the Euler equations and Newton's method should be mentioned. This is a proposed iterative technique due to Hestenes[38] designed to lead to an Euler solution which satisfies the specified boundary conditions. This idea has been pursued by Kalaba[39] whose theoretical studies have indicated favorable convergence properties within the limits of a convexity assumption. Computational experience with the method and examination of the practical implications of the convexity assumption are presently lacking.

Insofar as comparison between the various versions of the gradient method presented herein are concerned, the active competition appears to be between gradient projection and the penalty function scheme. The former requires an expensive error correction cycle which is unnecessary with the latter. On the other hand, the concentration of effort with the penalty function technique is overly heavy on reduction of errors in terminal values. If "tight" tolerances are employed, the process is oscillatory and convergence is slowed. The employment of "loose" tolerances in initial computations, with later tightening, appears to represent an effective compromise. The writer has considered, but not yet tried, the two versions in combination, i.e., in alternate cycles. This would relieve the need for corrections during the projection cycle and perhaps combine the good features of the two versions. In general, the computational experience with gradient methods has been sufficiently encouraging throughout the exploratory work reported herein to warrant future attack on comparatively large scale problems.

The writer and his colleagues have compared the gradient and Euler solutions for a limited number of cases in the problems described in the preceding sections. The motivation for this work happened to be a desire for an "exact" solution as a test extremal in some work on the second variation, rather than a comparison of technique. The first lesson from these experiments was that the control variable time history for a gradient solution converged within "engineering" accuracy (say 1% of minimum P) did not agree well over certain time intervals with an Euler solution. The disagreement was confined to those portions of the path over which the terminal values affecting P are relatively insensitive to control variations (see, for example, Fig. 14). From an engineering viewpoint this is unimportant if only the value of P is of main interest, as in flight performance work. It is inconvenient if a family of neighboring extremals are required, as, for example, in connection with a guidance study, for this requires that additional computations be performed to converge the control variable history to within the desired accuracy.

It has been speculated that an appropriate procedure for treatment of this situation is transition to a scheme for systematic numerical solution of the Euler equations, and this appears plausible on first consideration. One finds, however, that the appropriate linear combinations of adjoint solutions do *not* yield a good approximation to the multiplier functions of the indirect theory, and, in particular, the initial values of the multipliers may be sufficiently in error to cause difficulty in an iterative adjustment process.

In this connection, we will present in the following section a scheme intended to refine the control program of a near-minimal gradient solution

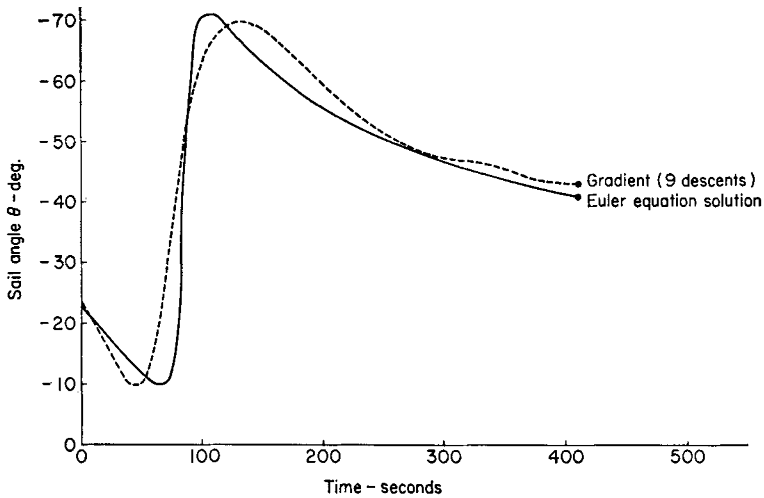


FIG. 14. Comparison of gradient and Euler equation solutions for solar sail orbital transfer example.

into a close approximation to an Euler solution. Although this development may appear to be an afterthought, as is actually the case, an account seems worthwhile in the spirit of the present volume. While computational experience with this technique is extremely limited at the present writing, its potential as a primary computational scheme as well as a refinement scheme is perhaps of future interest.

6.6 A Successive Approximation Scheme Employing the Min Operation

While the method outlined in this section is not, properly speaking, a gradient method, it is a close relative both in concept and operation. We consider the problem of minimizing the function P' of terminal values given by (6.79), referring the reader to earlier discussions on the matter of penalty constant determination. Initial values of the x_i are presumed fixed. A nonminimal solution corresponding to $y = \bar{y}(t)$ is generated by numerical integration of the basic system (6.34). For run termination we may employ the criterion $dP'/dt_f = 0$, this feature being equally applicable to computations with the penalty function version of the gradient technique. At the time $t = \bar{t}_f$ so determined, the terminal conditions

$$\lambda_i(\bar{t}_f) = \frac{\partial P'}{\partial x_i} \tag{6.150}$$

are imposed upon the λ_i and a solution of the adjoint system (6.41) generated by numerical integration proceeding from $\bar{t}_f$ to t_0.

We designate this solution by $\lambda_i^*(t)$, and from (6.44) we then have that

$$\delta P' = \sum_{i=1}^{n} \frac{\partial P'}{\partial x_{if}} \delta x_i(\bar{t}_f) = \int_{t_0}^{\bar{t}_f} \sum_{i=1}^{n} \lambda_i^* \frac{\partial g_i}{\partial y} \delta y \, dt \qquad (6.151)$$

(Note that the term $(dP'/dt_f)\delta t_f$ missing from the $\delta P'$ expression vanishes by virtue of the run termination criterion chosen above.)

We now define a function $H^*(y, t)$ as

$$H^* = \sum_{i=1}^{n} \lambda_i^*(t) g_i(\bar{x}_1, \cdots, \bar{x}_n, y, t) \qquad (6.152)$$

where the functions $\bar{x}_i(t)$ correspond to the nonminimal solution $y = \bar{y}(t)$. We get then that

$$\delta P' = \int_{t_0}^{\bar{t}_f} \frac{\partial H^*}{\partial y} \delta y \, dt \qquad (6.153)$$

where the partial derivative $\partial H^*(y, t)/\partial y$ is evaluated at $y = \bar{y}(t)$. The argument to this point is essentially identical with that of Section 6.5, and if we were to specify steepest descent subject to a Euclidean metric,

$$ds = \int_{t_0}^{\bar{t}_f} \delta y^2 \, dt \qquad (6.154)$$

[cf. Eq. (6.54)] we would then have a gradient method, leading to

$$\delta y = k \frac{\partial H^*}{\partial y} \qquad (k < 0) \qquad (6.155)$$

A shortcoming of such a process is that over intervals in which $\partial H^*/\partial y$ is small in magnitude, the corresponding changes in y will be small. After several steps y may still be far from its optimal value over such "insensitive" intervals owing to this feature of the gradient process. This, of course, stems from the rather arbitrary imposition of the Euclidean distance measure.

We may choose an equally arbitrary alternative, dropping the distance constraint altogether and adding a term

$$\delta P' = \int_{t_0}^{\bar{t}_f} \left[\frac{\partial H^*}{\partial y} \delta y + \frac{1}{2} \frac{\partial^2 H^*}{\partial y^2} \delta y^2\right] dt \qquad (6.156)$$

The only justification which can be offered for this alteration is that the term provides assurance that $\delta P'$ possesses a minimum for some finite δy, if indeed this is the case. By a formal process we obtain

$$\delta y = -\frac{\partial H^*/\partial y}{\partial^2 H^*/\partial y^2} \tag{6.157}$$

as the value of δy making $\delta P'$ stationary. The stationary value will yield a negative integrand only if $\partial^2 H^*/\partial y^2 > 0$. If $\partial^2 H^*/\partial y^2$ is zero or negative, the integrand is not bounded below and the minimum problem for $\delta P'$ ill-posed. Evidently this first attempt at modification, then, is a failure.

If, however, we add enough higher order terms to the integrand we will approach the operation

$$\min_y H^* \tag{6.158}$$

in the limit as the solution of the minimum problem for $\delta P'$, and this appears more promising since H^* can be expected to possess a minimum except in the unusual case where H^* does not depend upon y (a so-called abnormal case).

In adopting the control $y = y^*(t)$ generated by min H^* as our next approximation, we must risk the violation of our linearizing assumptions, for this may represent a large step process. For the purpose intended, the refinement of a near-minimal solution, this represents a calculated risk. However, being conservative, we may elect to replace the large step by an exploratory series of small ones, setting

$$y = \bar{y}(t) + \zeta[y^*(t) - \bar{y}(t)] \tag{6.159}$$

and evaluating P' versus ζ, a one-dimensional search analogous to that versus σ in gradient computations.

The reasoning leading up to the min H^* scheme for successive approximations is presented here as a matter of interest. The result might equally have been arrived at by analogy with the Pontryagin principle of the next chapter. (The difference between min H and Pontryagin's max H is one of sign convention between East and West.) Note that as the process converges the function H^* tends toward the function H, the generalized Hamiltonian.

In limited experience obtained with a single low-thrust example as of this writing, it appears that the method of this section converges to a solution which is a much better approximation to an Euler solution than normally obtainable with a gradient method as far as details of the control variable time history $y(t)$ are concerned, the improvement being in the "insensitive" regions discussed earlier. The speed of convergence of the method appears to make it competitive with the gradient/penalty function scheme, although insufficient evidence exists as yet to support any firm conclusion.

Some recent publications relating to the material of this chapter are listed as references 40–49.

ACKNOWLEDGMENTS

The writer is pleased to acknowledge the contributions of William P. O'Dwyer and H. Gardner Moyer of Grumman's Computation Facility in handling the computational phase of these studies on the IBM 704, and of Mrs. Agnes Zevens in checking and preparing the numerical results for publication.

Portions of the material of this chapter were generated in connection with USAF Contracts AF 29(600)–2671 and AF 29(600)–2733 monitored by AFOSR Directorate of Research Analysis, Holloman Air Force Base, New Mexico, and NASA Contract NAS 8–1549 with the Aeroballistics Division of the Marshall Space Flight Center, Huntsville, Alabama.

Appendix A

The rank of the matrix A whose elements are given by Eq. (6.71) may be examined by recourse to factorization as follows:

$$A = \beta C \beta'$$

where the prime indicates the transpose matrix. The elements of β are

$$\beta_{jm} = \frac{\partial \mathfrak{J}_j}{\partial x_{m_f}}, \qquad \begin{matrix} m = 1, \cdots, n \\ j = 1, \cdots, l \end{matrix}$$

and the elements of C are given by Eq. (6.73).

We denote the ranks of β and C as s and r, respectively, and observe that the three-way product above must have rank l for A to be nonsingular. Since the rank of this product may not exceed the rank of the premultiplicative matrix β, we have that

$$l \leq s = R(\beta)$$

(see reference 50, Chapter 3). Further, since the rank of β may not exceed the number of rows,

$$R(\beta) = s \leq l$$

and it follows that

$$R(\beta) = s = l$$

Since the rank of the product may not exceed the rank of the post-multiplicative matrix, we have that

$$R(\beta C\beta') = l \leq R(C\beta')$$

and since the rank of a matrix may not exceed the number of columns,

$$R(C\beta') \leq l$$

from which we conclude

$$R(C\beta') = l$$

Similar considerations establish that

$$R(C\beta') = l \leq r = R(C)$$

The rank of the product

$$B = \beta\beta'$$

may be deduced from the special case in which C is the identity matrix

$$R(\beta\beta') = R(A)$$

The results of interest in the text may thus be summarized as

$$l = R(A) \leq R(C) = r$$

and

$$\bar{l} = R(A) = R(B)$$

References

1. A. L. Cauchy, Méthode générale pour la résolution systèmes d'équations simultanées, *Compt. rend. acad. sci.* **25,** 536–538 (1847).
2. K. Levenberg, A method for the solution of certain non-linear problems in least squares, *Quart. Appl. Math.* **2,** 164–168 (1944).
3. H. B. Curry, The method of steepest descent for non-linear minimization problems, *Quart. Appl. Math.* **2,** 258–261 (1944).
4. J. Hadamard, Mémoire sur le problèm d'analyse relatif à l'équilibre des plaques élastiques encastrées, *Mém. prés. acad. sci. France* [2] **33,** No. 4 (1908).
5. R. Courant, Variational methods for the solution of problems of equilibrium and vibrations, *Bull. Am. Math. Soc.* **49,** 1–23 (1943).
6. G. E. P. Box and K. B. Wilson, On the experimental attainment of optimum conditions, *J. Roy. Statist. Soc. Ser. B* **13,** 1–45 (1951).
7. J. B. Rosen, The gradient projection method for nonlinear programming. Part I. Linear constraints, *J. Soc. Ind. Appl. Math.* **8,** 181–217 (1960).
8. P. Wolfe, The present status of nonlinear programming, RAND-University of California Symposium on Mathematical Optimization Techniques, Santa Monica, California, October 18–20, 1960.

9. C. B. Tompkins, Methods of steep descent, *in* "Modern Mathematics for the Engineer" (E. F. Beckenbach, ed.), Chapter 18. McGraw-Hill, New York, 1956.
10. M. R. Hestenes and E. Stiefel, Method of conjugate gradients for solving linear systems, *J. Research Natl. Bur. Standards* **49**, 409–436 (1952).
11. A. I. Forsythe and G. E. Forsythe, Punched-card experiments with accelerated gradient methods for linear equations, *Natl. Bur. Standards Appl. Math. Ser.* **39**, 55–69, (1954).
12. C. Müller, A new method for solving Fredholm integral equations, Rept. No. BR-15, New York University Institute of Mathematical Sciences, Division of Electromagnetic Research (1955).
13. M. L. Stein, On methods for obtaining solutions of fixed end-point problems in the calculus of variations, *J. Research Natl. Bur. Standards,* **50**, 277–297 (1953).
14. H. J. Kelley, Gradient theory of optimal flight paths, presented at Am. Rocket Soc. Semi-Annual Meeting, Los Angeles, California, May 9–12, 1960; *ARS Journal* **30**, 947–954 (1960).
15. A. E. Bryson, F. J. Carroll, K. Mikami, and W. F. Denham, Determination of the lift or drag program that minimizes re-entry heating with acceleration or range constraints using a steepest descent computation procedure, presented at IAS 29th Annual Meeting, New York, New York, January 23–25, 1961.
16. R. Courant, "Differential and Integral Calculus," Vol. II, Chapter 3. Interscience, New York, 1936.
17. R. Courant and D. Hilbert, "Methods of Mathematical Physics," first English ed. Interscience, New York, 1953.
18. R. Courant, Calculus of variations and supplementary notes and exercises," 1945–1946. Revised and amended by J. Moser, New York University Institute of Mathematical Sciences, New York (mimeographed lecture notes), 1956–1957.
19. H. Rubin and P. Ungar, Motion under a strong constraining force, *Communs. Pure Appl. Math.* **10**, 65–87 (1957).
20. A. S. Mengel, Optimum trajectories, *Proc. Project Cyclone Symposium 1 on REAC Techniques*, Reeves Instrument Corp. and USN Special Devices Center, New York, March 15–16, 1951.
21. J. H. Irving and E. K. Blum, Comparative performance of ballistic and low-thrust vehicles for flight to Mars, *in* "Vistas in Astronautics," Vol. II, pp. 191–218. Pergamon, New York, 1959.
22. C. R. Faulders, Low-thrust steering program for minimum time transfer between planetary orbits, SAE National Aeronautic Meeting, Los Angeles, California, September 29–October 4, 1958.
23. B. Friedman, "Principles and Techniques of Applied Mathematics." Wiley, New York, 1956.
24. G. A. Bliss, "Mathematics for Exterior Ballistics." Wiley, New York, 1944.
25. T. R. Goodman and G. N. Lance, The numerical integration of two-point boundary value problems, *Math. Tables and Other Aids to Computation* **10**, 82–86 (1956).
26. W. Kizner, Perturbation theory and Green's functions, Section Rept. No. 12-152, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California (October 21, 1957).
27. C. G. Pfeiffer, Guidance for space missions, External Publ. No. 656, Jet Propulsion Laboratory, Pasadena, California (June 23, 1959).
28. J. C. Dunn, Green's functions for space trajectory perturbation analysis, *J. Astronaut Sci.* **8**, 95–103 (1961).
29. P. Cicala, "An Engineering Approach to the Calculus of Variations," Levrotto-Bella, Torino, 1957.

30. F. A. Valentine, The problem of Lagrange with differential inequalities as added side conditions, Dissertation, Department of Mathematics, University of Chicago, Chicago, Illinois, 1937.
31. L. I. Rozonoer, L. S. Pontryagin principle in the theory of optimum systems, *Avtomat. i Telemekh.* **20** (1959). English Translation: *Automation Remote Control U.S.S.R.* **21,** 1288–1302, 1405–1421, 1517–1532 (1960).
32. R. L. Garwin, Solar sailing—a practical method of propulsion within the solar system, *Jet Propulsion* **28,** 188–190 (1958).
33. T. P. Cotter, Solar sailing, Sandis Research Colloquium, April, 1959, Albuquerque, New Mexico.
34. T. C. Tsu, Interplanetary travel by solar sail, *ARS Journal* **29,** 422–427 (1959).
35. H. S. London, Some exact solutions of the equations of motion of a solar sail with constant sail setting, *ARS Journal* **30,** 198–200 (1960).
36. W. Lindorfer and H. G. Moyer, An application of a low-thrust trajectory optimization scheme to planar Earth-Mars transfer, *ARS Journal* **32,** 260–262 (1962).
37. R. N. Edwards and H. Brown, Ion rockets for small satellites, Am. Rocket Soc. Controllable Satellites Conference, Massachusetts Institute of Technology, Cambridge, Massachusetts, April 30–May 1, 1959.
38. M. R. Hestenes, Numerical methods of obtaining solutions of fixed end-point problems in the calculus of variations, RAND Rept. No. RM 102, The RAND Corporation, Santa Monica, California (August, 1949).
39. R. Kalaba, On nonlinear differential equations, the maximum operation, and monotone convergence, *J. Math. and Mech.* **8,** 519–584 (1959).
40. H. J. Kelley, R. E. Kopp, and H. G. Moyer, Successive approximation techniques for trajectory optimization, presented at the IAS Vehicle Systems Optimization Symposium, Garden City, New York, November 28–29, 1961.
41. Y.-C. Ho, A successive approximation technique for optimal control systems subject to input constraints, presented at the 1961 Joint Automatic Control Conference, Boulder, Colorado, June 28–30, 1961.
42. Y.-C. Ho, A computational procedure for optimal control problems with state variable inequality constraints, Rept. No. P–2402, the RAND Corp., Santa Monica, California, September 1961 (to be published in *J. Math. Anal. and Appl.*).
43. A. E. Bryson and W. F. Denham, A steepest ascent method for solving optimum programming problems, Am. Soc. Mech. Eng. Conf. on Applied Mechanics, Chicago, Illinois, June 1961 (to be published in *J. Appl. Mech.*).
44. C. Saltzer and C. W. Fetheroff, A direct variational method for the calculation of optimum thrust programs for power-limited interplanetary flight, *Astronaut. Acta* **7,** 8–20 (1961).
45. T. Butler and A. V. Martin, On a method of Courant for minimizing functionals (to be published).
46. S. Dreyfus, Variational problems with inequality constraints, Rept. No. P-2357, the RAND Corp., Santa Monica, California (to be published in *J. Math. Anal. and Appl.*).
47. S. Dreyfus, Numerical Solution of variational problems, Rept. No. P-2374, the RAND Corp., Santa Monica, California (to be published in *J. Math. Anal. and Appl.*).
48. I. M. Gel'fand and M. L. Tsetlin, The organized search principle in systems with automatic stabilization, *Doklady Akad. Nauk S.S.S.R.* **137**(2), 295–298 (1961).
49. N. S. Bromberg, Maximization and minimization of complicated multivariable functions, *Communication and Electronics* **58,** 725–730 (1962).
50. S. Perlis, "Theory of Matrices." Addison-Wesley, Reading, Massachusetts, 1952.