NASA CR-1003

GUIDANCE, FLIGHT MECHANICS AND TRAJECTORY OPTIMIZATION

Volume IV - The Calculus of Variations and Modern Applications

By M. Mangad and M. D. Schwartz

Issued by Originator as Report No. SID 65-1200-4

Prepared under Contract No. NAS 8-11495 by
NORTH AMERICAN AVIATION, INC.
Downey, Calif.

for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

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

PRECEDING PAGE BLANK NOT FILMED.

FOREWORD

This report was prepared under contract NAS 8-11495 and is one of a series intended to illustrate analytical methods used in the fields of Guidance, Flight Mechanics, and Trajectory Optimization. Derivations, mechanizations and recommended procedures are given. Below is a complete list of the reports in the series.

Volume I	Coordinate Systems and Time Measure
Volume II	Observation Theory and Sensors
Volume III	The Two Body Problem
Volume IV	The Calculus of Variations and Modern Applications
Volume V	State Determination and/or Estimation
Volume VI	The N-Body Problem and Special Perturbation Techniques
Volume VII	The Pontryagin Maximum Principle
Volume VIII	Boost Guidance Equations
Volume IX	General Perturbations Theory
Volume X	Dynamic Programming
Volume XI	Guidance Equations for Orbital Operations
Volume XII	Relative Motion, Guidance Equations for Terminal Rendezvous
Volume XIII	Numerical Optimization Methods
Volume XIV	Entry Guidance Equations
Volume XV	Application of Optimization Techniques
Volume XVI	Mission Constraints and Trajectory Interfaces
Volume XVII	Guidance System Performance Analysis

The work was conducted under the direction of C. D. Baker, J. W. Winch, and D. P. Chandler, Aero-Astro Dynamics Laboratory, George C. Marshall Space Flight Center. The North American program was conducted under the direction of H. A. McCarty and G. E. Townsend.

1.0 STATEMENT OF THE PROBLEM

The analysis of deterministic systems common in studies of trajectories and control frequently requires that a particular measure of performances (time of operation, propellant expended, etc.) be extremized on some path satisfying all of the imposed boundary conditions. The fundamental problem to which this monograph is thus addressed can be stated as: "determine an arc from the family of all admissible arcs (piecewise smooth) which join the points $(x_1 \vec{y}_1)$ and $(x_2 \vec{y}_2)$ (where $\vec{y} = y(x)$ and where all values of x and $\vec{y}$ lie in a given region, G) and which extremizes the integral (or functional)

$$I(\vec{y}) = \int_{x_1}^{x_2} f(x, \vec{y}, \vec{y}')\, dx ,$$

where f is a real continuous function of $x, \vec{y}$ and $\frac{d}{dx}\vec{y}$." This monograph is intended to provide the theory sufficient to accomplish the formulation and solution of such problems.

The discussions presented in the text progress from the fundamental lemma to the Euler Lagrange equations, to the transversality condition, to the incorporation of constraint equations, to other necessary conditions for extrema, to the problems of Bolza, Mayer and Lagrange, and finally to the bounded and unbounded control problems.

In addition to providing the basis for formulating this class of problems, the monograph is intended to serve as a necessary introduction to subsequent discussions of the optimization problem. These other monographs will present the formulation of problems in stochastic control, of problems involving inequality constraints on the state variables, and of singular arc problems. Still another monograph will discuss numerical optimization techniques.

2.0 STATE OF THE ART IN OPTIMIZATION

2.1 MAXIMUM AND MINIMUM OF FUNCTIONALS EMPLOYING ORDINARY CALCULUS

This section presents a brief review of the results dealing with the extremization (maximization and minimization) of functions and certain types of functionals by means of ordinary calculus. This review was prepared as an introductory exposition for the purpose of providing a background to the Calculus of Variations. The information given in this chapter is not intended to be complete and assumes a degree of familiarity with ordinary calculus.

2.1.1 Maximization and Minimization of Functions

This exposition is concerned with the problem of finding, for a given continuous function $f(x) \equiv f(x_1, \ldots, x_n)$ in a given closed region, **G** , a point $\bar{x} \equiv (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ in G at which $f(x)$ attains its extremum (maximum or minimum) with respect to all points of G in a neighborhood of $\bar{x}$. This problem is known to have a solution due to Weierstrass theorem stating that "every continuous function in a closed region G possess a largest and a smallest value in the interior or on the boundary of G ." The solution to this problem may be found if the function $f(x)$ is differentiable in G and if an extreme (maximum or minimum) value is attained at an interior point $\bar{x}$. Indeed, if $\bar{x}$ yields an extremum for $f(x)$ (i.e., $f(\bar{x}) \leq f(x)$ for all $x \neq \bar{x}$ if $\bar{x}$ minimuzes f ; $f(\bar{x}) \geq f(x)$ for all $x \neq x$ if $\bar{x}$ maximizes f), then by expanding $f(x) = f(\bar{x} + \epsilon) \equiv f(x_1 + \epsilon_1, x_2 + \epsilon_2, \ldots, x_n + \epsilon_n)$ in a Taylor series around the point $\bar{x}$, it is seen that for $\bar{x}$ to extremize f it is necessary that the first order partial derivatives of $f(x)$ with respect to each variable x_i , $i = 1, 2, \ldots, n$, vanish, i.e., $\bar{x}$ is determined from the relation

$$\frac{\partial f(x)}{\partial x_i} = 0, \; i = 1, 2, \ldots, n. \qquad (2.1.1)$$

This necessary condition is not sufficient, as can be seen from the existence of points of inflection or saddle points in the following examples:

$$f(x) = x^3 \text{ at } \bar{x} = 0$$

$$f(x, y) = xy \text{ at } \bar{x} = 0, \bar{y} = 0.$$

PRECEDING PAGE BLANK NOT FILMED.

TABLE OF CONTENTS

LIST OF FIGURES

Sufficient conditions for more general type problems will be presented in Section 2.5.

Another maximum or minimum situation which is handled by ordinary calculus is the problem of extremizing

$$f(x) \equiv f(x_1, \ldots, x_n) \tag{2.1.2}$$

subject to the constraints

$$g_k(x) = 0 \quad , k = 1, 2, \ldots, m < n \tag{2.1.3}$$

In this situation the necessary conditions for an extremum are obtained by means of the so-called Lagrange multipliers. This method consists first of introducing m new parameters, λ_1, $\lambda_2, \ldots, \lambda_m$ and constructing a new function

$$F(x, \lambda) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x) + \cdots + \lambda_m g_m(x). \tag{2.1.4}$$

Next, by using a heuristic approach, the following equivalency relation may be deduced from constraint (2.1.3)

$$\text{Minimum } F(x) \equiv \text{ Minimum } [f(x) + 0] = \text{ Minimum } f(x)$$

Thus, the problem of extremizing $f(x)$ under the restrictions $g_1(x) = 0, \ldots, g_m(x) = 0$ is equivalent to the problem of extremizing $F = f + \sum_{k=1}^{n} \lambda_k g_k$. Finally, since the independent variables of F are $x_1, \ldots, x_2$; $\lambda_1, \lambda_2, \ldots, \lambda_m$, it follows from Equation (2.1.1) that the solution to the constrained problem expressed in Equations (2.1.2) and (2.1.3) is given by

$$\begin{aligned} &\frac{\partial F}{\partial x_i} = 0, \ \frac{\partial F}{\partial \lambda_k} = 0, \quad i = 1, \ldots, n \ ; k = 1, \ldots m; \\ &F(\lambda, x) = f(x) + \sum_{k=1}^{n} \lambda_k g_k(x). \end{aligned} \tag{2.1.5}$$

These $(n + m)$ equations will yield $\bar{x}_i$ and the undetermined constant multipliers λ_k .

2.1.2 Elementary Extremum Problems

The problem to be considered here is that of extremizing an integral of the form

$$I(y)=\int_a^b f(x, y_1, \ldots, y_n)\, dx \equiv \int_a^b f(x,y)\,dx \tag{2.1.6}$$

subject to some restrictions of the form

$$J_k(y)=\int_a^b g_k(x,y)\,dx = C_k, \quad k=1,2,\ldots,m. \tag{2.1.7}$$

Such problems are solvable by means of ordinary calculus as the following theorems show.

<u>Theorem 2.1.1.</u> A necessary and sufficient condition for $y=\bar{y}(x)$ to yield a minimum or a maximum of $I(y)$ (as defined in Equation (2.1.6)) is that $\bar{y}(x)$ yields a minimum or a maximum respectively, to $f(x,y)$. That is, if $y=\bar{y}(x)$ is the extremizing arc for $I(y)$, then

$$I(x,\bar{y})=\begin{cases}\text{min}\\ \text{or}\\ \text{max}\end{cases} I(x,y) \text{ if and only if } f(x,\bar{y})=\begin{cases}\text{min}\\ \text{or}\\ \text{max}\end{cases} f(x,y) \tag{2.1.8}$$

Proof. Consider the case of minimization and suppose $\bar{y}(x)$ minimizes $I(y)$ (the maximum version may be treated analogously). Next, consider

$$\frac{I[y(x)]-I[\bar{y}(x)]}{\delta} \qquad a<s \le x \le s+\delta < b,\ \delta>0.$$

Then,

$$0 \le \frac{I(y)-I(\bar{y})}{\delta} = \frac{1}{\delta}\int_s^{s+\delta} [f(x,\ y(x)) - f(x,\ \bar{y}(x))]\,dx$$

Now taking the limit as $\delta \to 0$, it follows that

$$0 \le f(s,\ y^*) - f(s,\ \bar{y})$$

for all points s (where $y^*(x)$ is some arbitrary admissible function). Hence, the arc $y=\bar{y}(x)$ minimizes $f(x,y)$.

The converse follows immediately.

Corollary. A necessary condition for minimizing or maximizing

$$I(y)=\int_a^b f(x,y)\,dx$$

is that

$$\frac{\partial f(x, y)}{\partial y_i} \equiv \frac{\partial f}{\partial y_i}(x, y_1, \ldots, y_n) = 0, \; i = 1, 2, \ldots, n \tag{2.1.9}$$

In regard to the more general problem given in Equations (2.1.6) and (2.1.7) the following results hold :

Theorem 2.1.2. Suppose $y = \bar{y}(x)$ minimizes $I(y)$ subject to the conditions $J_k(y) = C_k$ given in Equations (2.1.6) and (2.1.7). Then, there exist multipliers $\lambda_1, \ldots, \lambda_m$ such that for a $a \le x \le b$, $\bar{y}(x)$ minimizes

$$F(x, y, \lambda) = f(x, y) + \sum_{k=1}^{m} \lambda_k g_k(x, y) \tag{2.1.10}$$

The following corollary is a result of theorem 2.1.2:

Corollary. A necessary condition for minimizing

$$I(y) = \int_a^b f(x, y)\, dx \qquad J_k(y) = \int_a^b g_k(x, y)\, dx = C_k$$

$k = 1, 2, \ldots, m$ is that

$$\frac{\partial F}{\partial y_i} = 0, \; i = 1, 2, \ldots, n, \tag{2.1.11}$$

for each y_i and all values $a \le x \le b$.

2.2 THE FIRST NECESSARY CONDITION WITH FIXED END POINTS

2.2.1 Statement and Formulation of the Fundamental problem

The fundamental problem of the classical calculus of variations may be formulated as follows: Let there be given in region G of real variables $(x, y_1, y_2, \ldots, y_n) \equiv (x, Y)$ and a real continuous function $f(x, y_1, y_2, \ldots, y_n; z_1, z_2, \cdots, z_n) \equiv f(x, Y, Z)$. Now consider the collection of all piecewise-smooth (a piecewise-smooth function is a function whose first derivative is piecewise continuous) admissible curves $Y = Y(x)$ lying in the given region G which join the points $(x_1, Y_1) \equiv (x_1, y_{11}, y_{12}, \ldots, y_{1n})$ and $(x_2, Y_2) \equiv (x_2, y_{21}, y_{22}, \cdots, y_{2n})$, and for which the integral (or functional)

$$I(Y) = \int_{x_1}^{x_2} f(x, Y, Y')\, dx \tag{2.2.1}$$

has an extremum [where $Y' \equiv (y_1', y_2', \ldots, y_n')$, denotes differentiations with respect to x]. It is thus seen in the statement of the fundamental problem that, in contrast to the ordinary calculus, the calculus of variations is concerned with maximizing or minimizing functionals rather than ordinary functions.

This problem is sometimes restated in a less rigorous manner as that of finding an arc $Y = \bar{Y}(x)$ which extremizes the functional $I(Y)$ in Equation (2.2.1) subject to the boundary conditions.

$$\bar{Y}(x_1) = \bar{Y}_1 \,, \quad \bar{Y}(x_2) = \bar{Y}_2 \tag{2.2.2}$$

The functional (I) in Equation (2.2.1) is defined by a function $Y(x)$ and is denoted by $I[Y(x)]$, {i.e., from a certain class of admissible functions, there corresponds a value or a number $I[Y(x)]$}. Further, the variation (δY) of the argument $Y(x)$ of a functional $I(Y)$ is the difference of two functions: $\delta Y \equiv Y(x) - Y^*(x)$, where it is assumed that the argument $Y(x)$ runs through a certain class of functions [i.e., the variation δY is analogous to the differential $dx = x - x^*$ in ordinary functions $f(x)$]. Thus, the first variation of the functional (δI) is defined by

$$\delta I = I[Y(x) + \delta Y] - I[Y(x)] = \frac{\partial}{\partial \varepsilon} I[Y(x) + \varepsilon \delta Y]\Big|_{\varepsilon = 0} \tag{2.2.3}$$

Before proceeding with the discussion, relative, strong, and weak extrema will be introduced and defined to facilitate their use in the various sections of this monograph.

An extremum which is attained in the whole region G under consideration is called an absolute extremum. Otherwise, it is a relative extremum.

The extremal continuous vector function $y=\bar{y}\ (x)$ constitutes a strong extremum if the continuous curve $y=\bar{y}(x)$ in a neighborhood of $y=\bar{y}(x)$ satisfies the norm condition that

$$\|y(x)-\bar{y}(x)\| \equiv \sqrt{[y_1(x)-\bar{y}_1(x)]^2+\ldots+[y_n(x)-\bar{y}_n(x)]^2} \leqslant \epsilon \quad (2.2.4)$$

holds for $a \leq x \leq b$ [where $\varepsilon > 0$ characterizes the neighborhood of the extremal $y=\bar{y}(x)$]. If the extremal curve $y=\bar{y}(x)$ is piecewise-smooth and if, in addition to satisfying Equation (2.2.4), it satisfies the additional norm condition

$$\|y'(x)-\bar{y}'(x)\| \leq \varepsilon \quad (2.2.5)$$

for $a \leq x \leq b$, then the extremal is said to be a weak extremal. Extensions to a weaker extremal may be obtained by considering higher order derivatives of $\bar{y}(x)$ and its neighboring extremals $y(x)$.

<u>The Fundamental Lemma of Variational Calculus</u>

If x_1 and x_2 $(>x_1)$ are fixed constants and $G(x)$ is a particular continuous function for $x_1 \leq x \leq x_2$ and if

$$\int_{x_1}^{x_2} \eta(x)\, G(x)\, dx = 0$$

for every choice of the continuously differentiable function $\eta(x)$ for which

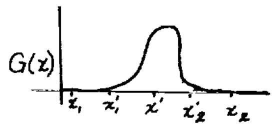
$$\eta(x_1) = \eta(x_2) = 0$$

then

$$G(x) = 0 \quad \text{identically in} \quad x_1 \leq x \leq x_2$$

Proof. The proof of this basic lemma rests upon demonstration of the existence of at least one suitable function $\eta(x)$ for which $\int_{x_1}^{x_2} \eta(x)\, G(x)\, dx \neq 0$ when $G(x) \neq 0$ throughout the interval $x_1 < x < x_2$. So, assume $G(x) \neq 0$; then, there is a particular x' ($x_1 < x' < x_2$) for which $G(x') \neq 0$ and for simplicity assume $G(x') > 0$. Since $G(x)$ is continuous, there must be an interval surrounding x' (say, $x_1' \leq x \leq x_2'$) in which $G(x) > 0$ everywhere. But the hypothesis $\int_{x_1}^{x_2} \eta(x)\, G(x)\, dx = 0$ cannot then hold for every choice of $\eta(x)$. For example, consider the choice

$$\eta(x) = \begin{cases} 0 & x_1 \le x \le x_1' \\ (x-x_1')^2 (x-x_2')^2 & x_1' \le x \le x_2' \\ 0 & x_2' \le x \le x_2 \end{cases}$$



which satisfies the assumed constraints. But, for this choice of $\eta(x)$

$$\int_{x_1}^{x_1} \eta(x)\, G(x)\, dx = \int_{x_1'}^{x_2'} (x-x_1')^2 (x-x_2)^2\, G(x) > 0$$

which contradicts the hypothesis. A similar contradiction is reached if $G(x')$ is assumed to be <0 . The lemma is hereby proved.

NOTES.(a) In some applications of the basic lemma, a more restrictive form will be required. It is required, for example, that an integral of the form $\int_{x_1}^{x_2} \eta(x) G(x)\, dx$ vanishes for every continuously twice-differentiable $\eta(x)$ for which $\eta(x_1) = \eta(x_2) = 0$.

To prove the necessity of $G(x) = 0$, again suppose $G(x') > 0$, in $x_1' \le x' \le x_2'$ but choose $\eta(x)$ to equal $(x - x_1')^3 \; (x_2' - x)^3$ in $x_1' \le x \le x_2'$ and zero elsewhere.

(b) Thus, the basic lemma can be extended to all cases where $\eta(x)$ is required to possess continuous derivatives up to and including any given order.

(c) If D is a domain of the xy-plane, the vanishing of the double integral

$$\iint_D \eta(x,y)\, G(x,y)\, dx\, dy$$

for every continuously differentiable $\eta(x,y)$ that vanishes on the boundary C of D necessititates the identical vanishing of $G(x,y)$ (G is assumed continuous in D). The proof of this extension is, in essence, the same as the proof given above. Further, the lemma is still valid if $\eta(x,y)$ possess continuous partial derivatives of any given order.

2.2.2 The First Necessary Condition

The classical method for solving problems of the calculus of variations resembles the approach used in ordinary calculus for solving elementary extremum problems. As a first step toward obtaining solutions to the problem, it is noted that any necessary condition for a weak relative extremum is also a necessary condition for a strong relative extremum and *a'* fortiori, an absolute extremum. Hence, in order to derive the first necessary conditions, the weak relative extremum will be examined.

2.2.2.1 Euler-Lagrange Equations

Let there be given a piecewise-smooth vector arc $y = \bar{y}(x) = (\bar{y}_1(x), \bar{y}_2(x), \ldots, \bar{y}_n(x))$ which yields a weak relative extremum for $I(y)$ in Equation (2.2.1) joining the fixed boundary points (x_1, y_1) and (x_2, y_2). Next, choose the arbitrary vector function $\eta(x) \equiv (\eta_1(x), \ldots, \eta_n(x))$ to satisfy the boundary conditions

$$\eta_i(x_1) = \eta_i(x_2) = 0 \quad , \quad i = 1, 2, \ldots, n \tag{2.2.6}$$

and consider neighboring arcs defined by the equation

$$y = \bar{y}(x) + \varepsilon \cdot \eta(x) \tag{2.2.7}$$

where $|\varepsilon|$ is sufficiently small so that $y = \bar{y}(x) + \varepsilon \cdot \eta(x)$ lies in a weak neighborhood (a neighborhood of weak extremals) of $y = \bar{y}(x)$. (Conditions (2.2.6) guarantee that the end point conditions (2.2.2) are satisfied.) Therefore, the functional $I(y)$ defined in Equation (2.2.1) has an extremum for y when $\varepsilon = 0$. Now, since $y = \bar{y}(x)$, $I(\bar{y} + \varepsilon\eta) = I(\varepsilon)$, (i.e., I is a function of ε alone), it follows that $dI/d\varepsilon\big|_{\varepsilon=0} = 0$ on the extremizing arc. Or

$$\begin{aligned} 0 = I'(\varepsilon)\big|_{\varepsilon=0} &= \int_{x_1}^{x_2} \sum_{i=1}^{n} \left\{ \frac{\partial f(x, y, y')}{\partial y_i} \eta_i(x) + \frac{\partial f(x, y, y')}{\partial y_i'} \eta_i'(x) \right\} dx \Big|_{\varepsilon=0} \\ &= \int_{x_1}^{x_2} \sum_{i=1}^{n} \left\{ \frac{\partial \bar{f}}{\partial y_i} \eta_i + \frac{\partial \bar{f}}{\partial y_i'} \eta_i' \right\} dx \end{aligned} \tag{2.2.8}*$$

where $\bar{f} = f(x, \bar{y}, \bar{y}')$. Integrating now by parts yields

$$\int_{x_1}^{x_2} \frac{\partial \bar{f}}{\partial y_i} \eta_i(x)\, dx = \left[\eta_i(x) \int_{x_1}^{x} \frac{\partial \bar{f}}{\partial y_i} dx \right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta_i'(x) \left[\int_{x_1}^{x} \frac{\partial f}{\partial y_i} dx \right] dx .$$

But, $\eta(x) = 0$ at x_1 and x_2, thus Equation (2.2.8) reduces to

$$\int_{x_1}^{x_2} \sum_{i=1}^{n} \left\{ \frac{\partial f}{\partial y_i'} - \int_{x_1}^{x} \frac{\partial \bar{f}}{\partial y_i} dx \right\} \eta_i'(x)\, dx = 0 \tag{2.2.8a}$$

Now, since the vector function $\eta(x)$ was arbitrary (satisfying conditions (2.2.6) only), it can be specified as

$$\eta_i(x) = \int_{x_1}^{x} [M_i(t) - C_i]\, dt \quad , i = 1, 2, \ldots, n \tag{2.2.9}$$

where

$$M_i(x) = \frac{\partial \bar{f}}{\partial y_i'} - \int_{x_1}^{x} \frac{\partial \bar{f}}{\partial y_i}\, dx \,, \qquad C_i = \frac{1}{x_2 - x_1}\int_{x_1}^{x_2} M_i(x)\, dx \tag{2.2.10}$$

Under this substitution, Equation (2.2.8a) becomes

$$\int_{x_1}^{x_2} \sum_{i=1}^{n} M_i(x)\,[M_i(x) - C_i]\, dx = 0$$

or

$$\int_{x_1}^{x_2} \sum_{i=1}^{n} [M_i(x) - C_i]^2\, dx = 0 \tag{2.2.11}$$

It follows, therefore, from Equation (2.2.11) that

$M_i(x) = C_i$

which, in turn, yields by means of Equation (2.2.10)

$$\frac{\partial f(x, y, y')}{\partial y_i'} - \int_{x_1}^{x} \frac{\partial f(x, y, y')}{\partial y_i}\, dx = C_i \quad , i = 1, 2, \ldots, n \tag{2.2.12}$$

Equations (2.2.12) are the Euler-Lagrange equations (or the first necessary condition) for the extrema of functionals of the form (2.2.1). This significant result is contained in the following theorem.

Theorem 2.2.1. If a piecewise-smooth curve $y = \bar{y}(x)$ lies in a given region and provides a weak relative extremum for the functional $I(y)$ in (2.2.1), then there exists constants C_i for which $y = \bar{y}(x)$ satisfies the Euler-Lagrange Equations (2.2.12).

As an immediate consequence of theorem 1, the following (most commonly used) result governing the extremal arcs can be stated.

Theorem 2.2.2. Wherever $\bar{y}'(x)$ is continuous in the inverval $[x_1, x_2]$ these extremizing functions $\bar{y}(x)$ also satisfy the Euler-Lagrange equations in the form (commonly used)

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right) = 0 \quad , i = 1, 2, \ldots, n, \tag{2.2.13}$$

$$\frac{d}{dx}\left(f - \sum_{i=1}^{n} y_i' \frac{\partial f}{\partial y_i'}\right) - \frac{\partial f}{\partial x} = 0 \tag{2.2.14}$$

Moreover, the so-called **Weierstrass-Erdmann corner conditions**

$$\frac{\partial \bar{f}}{\partial y_i'}\Big|_{x_0^-} = \frac{\partial \bar{f}}{\partial y_i'}\Big|_{x_0^+} \quad , i = 1, 2, \ldots, n \tag{2.2.15a}$$

$$\left\{\bar{f} - \sum_{i=1}^{n} \bar{y}_i \frac{\partial \bar{f}}{\partial y_i'}\right\}\Big|_{x_0^-} = \left\{\bar{f} - \sum_{i=1}^{n} \bar{y}_i \frac{\partial f}{\partial y_i}\right\}\Big|_{x_0^+} \tag{2.2.15b}$$

hold at any corner point (a point where $\bar{y}'$ is discontinuous) x_0 of $\bar{y}(x)$.

Proof. First note that the continuity of $\bar{y}'(x)$ in some interval implies the continuity of $\partial \bar{f}/\partial y_i$ and $(\partial \bar{f})/(\partial x)$ in the same interval.

Therefore, the functions $\int_{x_1}^{x} \frac{\partial \bar{f}}{\partial y_i} dx$ and $\int_{x_1}^{x} \frac{\partial f}{\partial x} dx$ have first order continuous derivatives in the interval (x_1 , x). This, in turn, implies that the Euler-Lagrange Equations (2.2.12) may be differentiated to obtain the desired Equations (2.2.13).

Next, note that the differentiation of the left-hand side of Equation (2.2.14) explicitly leads to

$$\frac{d}{dx}\left[f - \sum_{i=1}^{n} y_i' \frac{\partial f}{\partial y_i'}\right] - \frac{\partial f}{\partial x} = \sum_{i=1}^{n}\left\{-y_i' \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right) + y_i' \frac{\partial f}{\partial y_i}\right\}$$
$$= \sum_{i=1}^{n} -y_i\left\{\frac{d}{dx}\left(\frac{\partial f}{\partial y_i'} - \frac{\partial f}{\partial y_i}\right)\right\} = 0$$

Hence, Equation (2.2.14) is valid.

The Weierstrass-Erdmann corner conditions (2.2.15) follow from the fact that $\int_{x_1}^{x} \frac{\partial \bar{f}}{\partial y_i} dx$ and $\int_{x_1}^{x} \frac{\partial \bar{f}}{\partial x} dx$ are continuous at every point so that the quantities in Equation (2.2.15) are also continuous. This completes the proof of the theorem.

The Euler-Lagrange Equations (2.2.13) constitute the most fundamental and useful result of variational calculus. The more restricted E-L equations [Eq (2.2.12)] are seldom used since they apply to extremal functions $y(x)$ whose second derivatives need not exist. Every integral curve (extremal) of equation (2.2.13) is either smooth or composed of broken extremal parts: extremals satisfying the Weierstrass-Erdmann corner conditions (2.2.15).

2.2.2.2 First Integrals and Degenerate Solutions of E-L Equations

Two corollaries follow as immediate consequences of Equations

(2.2.13) and (2.2.14). These are:

Corollary 1. If the integrand function f does not contain any of the components y_i, but does contain y_i', i.e., if $f=f(x, y')$ then the E-L equations have a first integral:

$$\frac{\partial f}{\partial y_i'} = C_i, \quad i = 1, 2, \ldots, n$$

Further, if $f=f(y, y')$, then in view of Equation (2.2.14) a weak extremum arc $\bar{y}(x)$ satisfies the equation

$$f(y, y') - \sum_{i=1}^{n} y_i' \frac{\partial f}{\partial y_i'} = C$$

These deductions have special significance for the case $n=1$ in which the extremals are defined by a single equation.

Corollary 2. If $y(x)$ is twice differentiable, the E-L Equations (2.2.13) and (2.2.14) are second order differential equations; Equation (2.2.13) implied Equation (2.2.14)

Proof. The first assertion is proved by explicitly differentiating the E-L equations. Next, by differentiating Equation (2.2.14) there results

$$\sum_{i=1}^{n} y_i' \left[\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'} \right) \right] = 0 \tag{2.2.16}$$

so that the second assertion of the corollary follows as well.

NOTE: Equation (2.2.16) allows the following assertion: If a twice differentiable function $y(x)$ satisfies Equation (2.2.14) in some interval and if it also satisfies all but the K-th of the E-L equations (2.2.13), then if $y_k'(x) \neq 0$ in the interval, $y(x)$ must satisfy the K-th E-L equations as well.

The system of n second order E-L equations, requires $2n$ arbitrary constants for their complete solution. They are provided by the $2n$ boundary conditions(2.2.2)

2.2.2.3 Parametric Representation: $f=f(t, x, \dot{x})$

The results of 2.2.2.1 and 2.2.2.2 are directly applicable to problems involving parametric representation (i.e., problems where the time t is the independent variable). Indeed, such a representation of both the integrand f and the extremal arc is essential in treating dynamical problems. Therefore, the previous analyses will be extended to the problem of extremizing the functional

$$I(\bar{x}) = \int_{t_1}^{t_2} f(t, x, \dot{x})\, dt \equiv \int_{t_1}^{t_2} f(t, x, \ldots, x_n; \dot{x}, \ldots, \dot{x}_n)\, dt$$

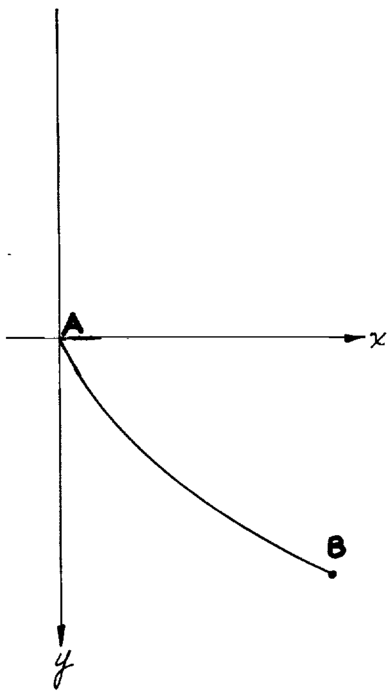
where $\dot{x} = \frac{dx}{dt}$ and with the boundary conditions

This problem is identical to that presented in Equations (2.2.1) and (2.2.2) except for a change in variables: Here t replaces x and χ replaces y. Thus, the corresponding E-L equations are:

$$\frac{\partial f}{\partial \chi_i} = \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{\chi}_i}\right) , \quad i = 1, 2, \ldots, n. \tag{2.2.13*}$$

Example 1.

The Brachistochrone Problem. Find a curve joining two given points A and B so that a particle moving under gravity along this curve starting at A reaches B in the shortest time. (Friction and resistance of the medium are neglected.)



Solution. Place point A at the origin of the coordinate system as shown in Figure 1: the x-axis being horizontal and y-axis, vertical directed downwards. The velocity of the particle $v \equiv \frac{ds}{dt} = \sqrt{2gy}$ so that

$$dt = \frac{ds}{\sqrt{2gy}} = \frac{\sqrt{1+y'^2}}{\sqrt{2gy}}\,dx .$$

Therefore, the time needed to reach point $B(x_2, y_2)$ is

$$t[y(x)] = 1/\sqrt{2g}\int_0^{x_2} \frac{\sqrt{1+y'^2}}{\sqrt{y}}\,dx , \quad y(0) = 0 , \quad y(x_2) = y_2$$

but (2.2.13*) yields

$$\frac{\sqrt{1+y'^2}}{\sqrt{y}} - \frac{y'^2}{\sqrt{y(1+y'^2)}} = C \quad \text{or} \quad y(1+y'^2) = C,$$

At this point, if the substitution $y' = \cot t$ is made

$$y = \frac{C_1}{1+\cot^2 t} = C_1 \sin^2 t = \frac{C_1}{2}(1 - \cos 2t)$$

$$dx = \frac{dy}{y'} = \frac{2C_1 \sin t \cos t \, dt}{\cot t} = 2C_1 \sin^2 t \, dt = 2C_1(1-\cos^2 t)\,dt$$

$$x = C_1\left(t - \frac{\sin 2t}{2}\right) + C_2 = \frac{C_1}{2}(2t - \sin 2t) + C_2$$

Hence, the parametric equation of the curve is

$$x - C_2 = \frac{C_1}{2}(2t - \sin 2t), \quad y = \frac{C_1}{2}(1 - \cos 2t)$$

Now, by observing that $C_2 = 0$ for $y = 0$ when $x = 0$, and by modifying the parameter by setting $2t = T$, a family of cycloids (in the usual form) is obtained.

$$x = \frac{C_1}{2}(T - \sin T), \quad y = \frac{C_1}{2}(1 - \cos T)$$

where $C_1/2$ is a radius of the rolling circle, which may be determined from the cycloid passing through the point $B(x_2, y_2)$. The Brachistochrone is, therefore, a cycloid.

2.2.3 Functionals Involving Higher Order Derivatives

Attention will now turn to the analysis of extremals of functionals where integrands contain second and higher derivatives. For simplicity, the one-dimensional case will be considered first. Extension to higher dimensional cases will then follow.

2.2.3.1 The One-Dependent Variable Case

The objective is to investigate the extremal of the functional

$$I[y(x)] = \int_{x_1}^{x_2} f(x, y, y', y'', \ldots, y^{(n)})\, dx \qquad (2.2.17)$$

It is assumed that f possesses $(n+2)$ partial derivatives with respect to all its arguments and the extremal $y = y(x)$ possesses $2n$ derivatives satisfying the following prescribed boundary conditions.

$$y(x_i) = y_i, \; y'(x_i) = y_i', \ldots, y^{(n-1)}(x_i) = y_i^{(n-1)} \qquad (2.2.18)$$
$$i = 1, 2$$

Let $y = y^*(x)$ be a neighboring comparison curve to the extremal $y = \bar{y}(x)$ and consider the expression

$$y(x,\varepsilon) = \bar{y}(x) + \varepsilon[y^*(x) - \bar{y}(x)] \equiv \bar{y}(x) + \varepsilon\delta\bar{y}$$

where

$$y(x,0) = \bar{y}(x) \text{ and } y(x,1) = y^*(x).$$

Since only the values of the family of curves $y = y(x,\varepsilon)$ are considered in $I(y)$, $I(y)$ then becomes a function of ε alone: $I = I(\varepsilon)$.Hence, the extremum of I is achieved from

$$\delta I \equiv I'(\varepsilon)\Big|_{\varepsilon=0} = \frac{d}{d\varepsilon}\left[I\{y(x,\varepsilon)\}\right]_{\varepsilon=0} = 0 ,$$

i.e.,

$$\delta I = \frac{d}{d\varepsilon}\int_{x_1}^{x_2} f\left[x, y(x,\varepsilon), y'(x,\varepsilon), \ldots, y^{(n)}(x,\varepsilon)\right] dx\Big|_{\varepsilon=0}$$

$$= \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta y' + \ldots + \frac{\partial f}{\partial y^{(n)}}\delta y^{(n)}\right] dx = 0.$$

Integrating by parts, the second term on the right hand side (once), the third term (twice), etc., and the last term ($n-1$ times), the following expressions are obtained:

$$\int_{x_1}^{x_2} f_{y'}\,\delta y'\,dx = \left[f_{y'}\,\delta y\right]_{x_1}^{x_2} - \int_{x_1}^{x_2}\frac{d}{dx}\{f_{y'}\}\,\delta y\,dx,$$

$$\int_{x_1}^{x_2} f_{y''}\,\delta y''\,dx = \left[f_{y''}\,\delta y'\right]_{x_1}^{x_2} - \left[\frac{d}{dx}\{f_{y''}\}\,\delta y\right]_{x_1}^{x_2} + \int_{x_1}^{x_2}\frac{d^2}{dx^2}\{f_{y''}\}\,\delta y\,dx,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\int_{x_1}^{x_2} f_{y^{(n)}}\,\delta y^{(n)}\,dx = \left[f_{y^{(n)}}\,\delta y^{(n-1)}\right]_{x_1}^{x_2} - \left[\frac{d}{dx}\{f_{y^{(n)}}\}\,\delta y^{(n-2)}\right] + \ldots$$

$$+ (-1)^n\int_{x_1}^{x_2}\frac{d^n}{dx^n}\left\{f_{y^{(n)}}\right\}\delta y\,dx$$

Now remembering that the end points (x_1, y_1) and (x_2, y_2) are fixed so that the variations vanish at these points: $\delta y = \delta y' = \ldots = \delta y^{(n-1)} = 0$, $\delta I = 0$ reduces to

$$\delta I = \int_{x_1}^{x_2}\left[f_y - \frac{d}{dx}f_{y'} + \frac{d^2}{dx^2}f_{y''} + \ldots (-1)^n\frac{d^n}{dx^n}f_{y^{(n)}}\right]\delta y\,dx = 0$$

with arbitrary δy . Finally using the fundamental lemma of the calculus of variation (Section 2.2.1) the E-L equation follows

$$f_y - \frac{d}{dx}f_{y'} + \frac{d}{dx}f_{y''} + \ldots + (-1)^n\frac{d^n}{dx^n}f_{y^{(n)}} = 0 \qquad (2.2.19)$$

This is the differential equation (of order $2n$) for integral curves (solutions) which are the desired extremals. The $2n$ arbitrary constants are determined from the $2n$ boundary conditions given in Equation (2.2.18).

Example: Find the extremal of

$$I(y)=\int_0^1 (1+y''^2)\,dx \quad \text{with} \quad y(0)=0\ ,\ y'(0)=1\ ,\ y(1)=1\ ,\ y'(1)=1.$$

The Euler equation is $\dfrac{d^2}{dx^2}(2y'')=0$ or $\dfrac{d^4y}{dx^4}=0$

The general solution is $y=C_3x^3+C_2x^2+C_1x+C_0$

From the boundary conditions $\qquad C_1=1\ ,\ C=C_2=C_3=0$.

Hence, the extremal curve is $y=x$.

2.2.3.2 The Two and Higher Dependent Variable Cases

In the same manner, if the functional $[I]$ has the form

$$I[y(x),z(x)]=\int_{x_1}^{x_2} f(x,y,y',\dots y^{(n)},z,z',\dots,z^{(m)})\,dx$$

then varying only $y(x)$ while keeping $z(x)$ fixed, reveals that any pair of functions $y(x)$ and $z(x)$ that yield an extremum for $I[y(x),\ z(x)]$ must satisfy the Euler equation

$$f_y-\frac{d}{dx}f_{y'}+\dots+(-1)^n\frac{d^n}{dx^{(n)}}f_{y^{(n)}}=0 \tag{2.2.19}$$

Similarly, by varying $z(x)$ while keeping $y(x)$ fixed,

$$f_z-\frac{d}{dx}f_{z'}+\dots+(-1)^n\frac{d^m}{dx^m}f_{z^{(m)}}=0 \tag{2.2.20}$$

Consequently, the functions $z(x)$ and $y(x)$ should satisfy the system of two equations (2.2.19) - (2.2.20).

An identical line of argument applies in the discussion of extrema of similar functionals depending on an arbitrary number of functions

$$I[y_1,y_2,\dots,y_m]=\int_{x_1}^{x_2} f(x,y_1,\dots,y_1^{(n_1)};y_2,y_2',\dots y_n^{(n_2)};\dots; y_m,y_m',\dots,y_m^{(n_m)})\,dx$$

Varying any function $y_i(x)$, and keeping the remaining ones fixed, the Euler equations are

$$f_{y_i} - \frac{d}{dx} f_{y_i'} + \cdots + (-1)^{n_i} \frac{d^{n_i}}{dx^{n_i}} f_{y_i^{(n_i)}} = 0 \quad , i = 1, 2, \ldots, m. \qquad (2.2.21)$$

2.2.4 Relations to Dynamics of Particles - Hamilton Principle, Canonical Forms

The material of the ensuing section is based upon an assumed knowledge of basic concepts of particle dynamics. The aim here is to provide a glimpse of the role played by the calculus of variations in a small segment of dynamics as well as demonstrating their relationships.

2.2.4.1 Hamilton's Principle

One of the most important principles in mechanics, commonly known as Hamilton's principle, is variational in nature. Consider the Kinetic energy $[T]$ of a mass m

$$T = \tfrac{1}{2} m \left(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2 \right)$$

and the potential energy $[V]$ of the particle in a conservative field of force.

$$V = V(x_1, \ldots, x_3)$$

so that the equations of motion of a particle are given by

$$m \ddot{x}_i = - \frac{\partial V}{\partial x_i} \, , \quad i = 1, 2, 3. \qquad (2.2.22)$$

Thus, if Lagrange's function $[L]$ is formed

$$L(x, \dot{x}) = T - V$$

it is seen that

$$\frac{\partial L}{\partial \dot{x}_i} = m \dot{x}_i \, , \quad \frac{\partial L}{\partial x_i} = - \frac{\partial V}{\partial x_i}$$

Equation (2.2.22), therefore, becomes

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}_i} \right) = \frac{\partial L}{\partial x_i} \, , \quad i = 1, 2, 3. \qquad (2.2.23)$$

which are the Euler equations for the integral

$$I = \int_{t_1}^{t_2} L(x, \dot{x}) \, dt \qquad (2.2.24)$$

and which is independent of the choice of coordinates. This result suggests the following:

Hamilton's Principle - In a conservative force field, a particle moves so as to extremize the integral

$$I = \int_{t_1}^{t_2} L \; dt$$

An equivalent expression of Hamilton's Principle is the "Principle of "Least Action." Since it has been assumed that the potential energy does not depend explicitly on time (i.e., $V = V(x_1, x_2, x_n)$) the total energy is a constant. Hence

$$L = T - V = 2T - E, \qquad E = T + V = \text{constant}$$

and extremizing the integral

$$I = \int_{t_1}^{t_2} L \; dt$$

is equivalent to extremizing the integral

$$I_1 = \int_{t_1}^{t_2} 2T \, dt \tag{2.2.25}$$

subject to the constraint that the energy is constant along the path. The integral I_1 , is usually referred to on the "action" integral; hence, the name "Principle of Least Action."

In most cases, the motion of the particle is such as to minimize the action. For this reason, Hamilton's Principle and the "Least Action Principle" are usually referred to as minimum principles. However, it can be shown that the action is a definite minimum only if the interval of integration $[t_1, t_2]$ is sufficiently small. For large intervals, the action may be only stationary. This point is best illustrated by an example. Consider the gravitational field in the xy-plane with the x-axis as ground and the y-axis directed upward. (The potential per unit mass is $V = gy$.) Assume a bullet is shot vertically from the origin with an initial speed V_0 at an angle A with respect to the horizontal. Then $E = \frac{1}{2} V_0^2$ and the integral (2.2.25) is

$$I_1 = \int_{t_1}^{t_2} \sqrt{(E - gy)(\dot{x}^2 + \dot{y}^2)} \;\; dt \tag{2.2.26}$$

Now employing Euler Equation (2.2.23) reveals that the trajectories are of the form

$$\frac{E}{g} - y = b\left[1 + \tfrac{1}{4}\left(\frac{x-a}{b}\right)^2\right].$$

The particular values

$$b = \frac{E}{g}\cos^2 A \ , \ a = \frac{E}{g}\sin 2A$$

yield the following family of trajectories

$$y = x \tan A - \frac{g}{4E} x^2 \sec^2 A \quad + \quad E/g$$

The family has as its envelope the

$$y = \frac{E}{g} - \frac{g}{4E} x^2 .$$

A particle shot upward from the origin cannot cross this envelope, but a point (x_2, y_2) under the envelope may be reached by two paths with inclinations A_1 and A_2, respectively, say $A_1 > A_2$. The path determined by A_2 yields a relative minimum to the integral (2.2.26), while the other does not. In fact, if the point 2 is sufficiently close to the envelope, there are paths below the line $y = E/g$ which give the integral a smaller value than either of these two trajectories. These results indicate that the principle of least action is applicable, as a minimum principle, only to arcs that are sufficiently small. The least action principle is, none the less, a very important principle in mechanics.

As an illustration of Hamilton's principle, consider a particle of unit mass acted upon by a central force which is inversely proportional to the square of the distance from the center. Assuming planar motion and using polar coordinates with the pole at the center of force, the following is valid:

$$T = \frac{1}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right), \ V = -\frac{C}{r}$$

Now employing the Euler equation

$$\frac{\partial L}{\partial \theta} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right), \quad \frac{\partial L}{\partial r} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right)$$

it follows first that the energy $E = T + V$ is again constant. However, for this problem, a second constant also exists. This fact can be shown by examining the first of the two equations of motion.

$$\frac{\partial L}{\partial \dot{\theta}} = r^2\dot{\theta} = \text{constant or } r^2\dot{\theta} = h$$

Next, since $\dot{r} = \dot{\theta}\,\frac{dr}{d\theta}$, T becomes

$$T = \frac{1}{2}\left[\left(\frac{dr}{d\theta}\right)^2 + r^2\right]\dot{\theta}^2 = \frac{1}{2}\left[\left(\frac{dr}{d\theta}\right)^2 + r^2\right]\frac{h^2}{r^4}$$

Assuming that $h \neq 0$, then from the relations $T + V = E$

$$a^2\left[\left(\frac{dr}{d\theta}\right)^2 + r^2\right] = Er^4 + Cr^3, \quad a^2 = \frac{h^2}{2}$$

which when integrated, has a solution of the form

$$r = \frac{K}{1 + e\cos(\theta - \gamma)}$$

where K, e, and γ are constants. The path is accordingly a conic.

2.2.4.2 Hamiltonians in Mechanics. Canonical Forms

For "Natural Systems" (i.e., those system for which the Lagrangian does not contain linear functions of the velocity) the Hamiltonian is equal to the total energy of the system and the constraints on the system are independent of time. Consider such a system:

$$H = \sum_{i=1}^{n} \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} - L = T + V \tag{2.2.27}$$

in place of the Lagrangian L. In this case, the momenta

$$P_i = \frac{\partial L}{\partial \dot{x}_i}(X, \dot{X}), \quad i = 1, 2, \ldots, n \tag{2.2.28}$$

are used instead of the variables $\dot{x}_i$. Denoting the solutions of Equations (2.2.28) by

$$\dot{x}_i = Q_i\,(P, X), \tag{2.2.29}$$

and substituting this result into Equation (2.2.27), then

$$H(X, P) = \sum_{i=1}^{n} Q_i P_i - L(X, Q)$$

as a formula for the Hamiltonian H in terms of the "canonical variables" (x_i, P_i).

It is not difficult to see that under the transformation (2.2.29)

$$\frac{\partial H}{\partial X_i} = -\frac{\partial L}{\partial X_i}, \quad \frac{\partial H}{\partial P_i} = \dot{x}_i, \quad i = 1, 2, \ldots, n.$$

Using these relations, the E-L Equations (2.2.23) reduce to the following "canonical form"

$$\dot{x}_i = \frac{\partial H}{\partial P_i}, \quad \dot{P}_i = \frac{-\partial H}{\partial x_i}, \quad i = 1, 2, \ldots, n. \tag{2.2.30}$$

These equations are fundamental in mechanics.

It is of interest to note that Equations (2.2.30) are equivalent to the E-L equations corresponding to extremizing the integral

$$\int_{t_1}^{t_2} \left[\sum_{i=1}^{n} \dot{x}_i p_i - H(X,P) \right] dt \tag{2.2.31}$$

Accordingly, an alternative form of Hamilton's principle can be stated as follows:

"A trajectory

$$x_i(t),\ p_i(t),\ t_1 \le t \le t_2,\ i = 1, 2, \ldots, n, \tag{2.2.32}$$

in a conservative field of force is a stationary (extremal) curve of the integral (2.2.31) where $H(X, P)$ is the associated Hamiltonian. Along this stationary curve, H is a constant".

The term "stationary" is used instead of "extremal" since it is clear that the integral (2.2.31) has no maximizing or minimizing arcs when x_i and p_i are free variables.

2.2.5 Variational Problems Involving Multiple Integrals

The mathematical theory for variational problems associated with extremizing multiple integrals is not as simple as that for the problems described previously. It is,however, a relatively simple matter to obtain the E-L equations associated with the solution which the extremum must satisfy. The procedure will be illustrated in a formal way for a double integral problem in xyz -space. The integral to be extremized for this class of problem is

$$I = \iint_A f(x, y, z, p, q)\, dx\, dy,\ p = \frac{\partial z}{\partial x},\ q = \frac{\partial z}{\partial y} \tag{2.2.33}$$

for a class of surfaces

$$z = z(x,y),\ (x,y) \text{ in } A \tag{2.2.34}$$

having the same boundary values. Proceeding as before, by considering the first variation δI of the integral (I) along an extremizing surface, then

$$\delta I = \iint_A \left[\frac{\partial f}{\partial z} \delta z + \frac{\partial f}{\partial p} \delta z_x + \frac{\partial f}{\partial q} \delta z_y \right] dx\, dy = 0 \tag{2.2.35}$$

for all variations z that vanish on the boundary of A. By Green's formula, Equation (2.2.35) reduces to

$$\delta I = \iint_A \left[\frac{\partial f}{\partial z} - \frac{\partial}{\partial x} f_p - \frac{\partial}{\partial y} f_q \right] \delta z\, dx\, dy = 0 \tag{2.2.36}$$

In view of Note (c) of Section 2.2.1, this is possible only if

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial p}\right) - \frac{\partial}{\partial q}\left(\frac{\partial f}{\partial p}\right) = 0 \tag{2.2.37}$$

Hence, "an extremizing surface having continuous second partial derivatives must satisfy the Euler Equation (2.2.37)."

One of the best known multiple integrals of this type is the Dirichlet integral

$$\iint_A \left[\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2\right] dx\, dy$$

In this minimization case, Eq (2.2.37) yields

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0$$

whose solution is a harmonic function. This fact suggests the following variational principle, commonly known as Dirichlet's principle: "A harmonic function minimizes the Dirichlet integral in the class of functions having the same boundary values."

No further studies involving the extremization of multiple integrals (such as the second variation for example) will be reported at this time since the efforts are quite involved and since the generalization of the Euler-Lagrange problem has not as yet been discussed.

2.3 THE FIRST NECESSARY CONDITION WITH FREE END POINT. THE TRANSVERSALITY CONDITION

2.3.1 Introduction and Formulation of the Problem

In the analyses presented in Sections 2.1 and 2.2, it was assumed that the end points $(x_1, y_1) \equiv (x_1, y_{11}, \ldots, y_{1n})$ and $(x_2, Y_2) \equiv (x_2, y_{21}, y_{22}, \cdots, y_{2n})$ of the functional

$$I = \int_{x_1}^{x_2} f(x, Y, Y')\, dx$$

were held fixed. This discussion is intended to relax this assumption and allow the end points to move freely along lines and curves. The inclusion of such extremal curves (assuming their existence) will in turn enlarge the class of admissible and comparison curves (curves in a small neighborhood of the extremizing one) having the same end points as the extremizing curve. In view of this extension, if $Y = \bar{Y}(x)$ yields an extremum for a problem with variable end points, then the same curves yield an extremum with respect to a more restricted class of curves having the same end points as the curve $Y = \bar{Y}(x)$. Hence, the Euler-Lagrange Equations are still valid for the free end points problem:

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right) = 0, \quad i = 1, 2, \ldots, n$$

The general solution of E-L equations involve 2n arbitrary constants. In the fixed end point problem, these 2n constants were provided by the conditions:

$$y_i(x_1) = y_{i1}, \; y_i(x_2) = y_{i2}, \quad i = 1, 2, \ldots, n$$

In the movable end-points case, some or all of these conditions are not satisfied. Therefore, additional conditions to determine the arbitrary constants in E-L equations are needed. These conditions are derived from the fundamental necessary condition that the first variation vanishes:

$$\delta I \equiv \frac{d}{d\varepsilon}\left[I(\bar{Y} + \varepsilon \cdot \eta)\right]\Big|_{\varepsilon=0} = 0$$

The conditions will be derived first for the one dimensional case (n = 1) for the purposes of simplicity. The generalization to any arbitrary dimensions will follow.

2.3.2 The One Dependent Variable Case

In the one dependent variable case, the complete solution to E-L equations involves two arbitrary constants: a, b say; $y = y(x, a, b)$. Consider next the integral $I = I[y(x, a, b)]$ evaluated along such an integral curve. When thus evaluated, the functional (I) becomes a function of five variables: $y(x), a, b, x_1$ and x_2 so that the variation of that functional coincides with the differential of such a function; i.e.,

$$\delta I = \frac{d}{d\varepsilon} I(y + \varepsilon \eta) = dI(y, a, b, x_1, x_2) \equiv \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial a} da + \dots + \frac{\partial I}{\partial x_2} dx_2$$

For simplicity, assume first that only the end point $x = x_2$ is free while $x = x_1$ is kept fixed; i.e., the point (x_2, y_2) moves to $(x_2 + \Delta x_2, y_2 + \Delta y_2)$ or $(x_2 + \delta x_2, y_2, \delta y_2)$ which is the usual designation in the calculus of variations. In this case, $I = I(y(x), a, y_2)$ and hence,

$$\delta I = \int_{x_1}^{x_2+\delta x_2} f(x, y+\delta y, y'+\delta y')\, dx - \int_{x_1}^{x_2} f(x, y, y')\, dx \qquad (2.3.1)$$

$$= \int_{x_2}^{x_2+\delta x_2} f(x, y+\delta y, y'+\delta y')\, dx$$

$$+ \int_{x_1}^{x_2} [f(x, y+\delta y, y'+\delta y') - f(x, y, y')]\, dx$$

Now using the mean value theorem,

$$\int_{x_2}^{x_2+\delta x_2} f(x, y+\delta y, y'+\delta y')\, dx = f\Big|_{x = x_2 + \theta \delta x_2} \delta x_2$$

But f is assumed piece wise continuous. Thus,

$$f\Big|_{x_2+\theta\delta x_2} = f(x\ y\ \ y')\Big|_{x=x_2} + \varepsilon_2$$

where $\varepsilon_2 \to 0$ as $\delta x_2 \to 0$ and $\delta y_2 \to 0$. Consequently,

$$\int_{x_2}^{x_2+\delta x_2} f(x, y+\delta y, y'+\delta y')dx = f(x, y, y')\Big|_{X=X_2} \delta x_2 + \varepsilon_2 \delta x_2$$

Now turning attention to the second expression in Eq. (2.3.1), the utilization of Taylor series **yields**

$$\int_{x_1}^{x_2} [f(x, y+\delta y, y'+\delta y') - f(x, y, y')]\, dx$$
$$= \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}(x, y, y')\, \delta y + \frac{\partial f}{\partial y'}(x, y, y')\, \delta y' \right] dx + R$$

where R is an infinitesimal of higher order than δy and $\delta y'$.
But $\int_{x_1}^{x_2} \frac{\partial f}{\partial y'} \delta y'\, dx$ can be integrated by parts to yield

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial y'} \delta y'\, dx = \frac{\partial f}{\partial y'} \delta y \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) \delta y\, dx$$

Hence,

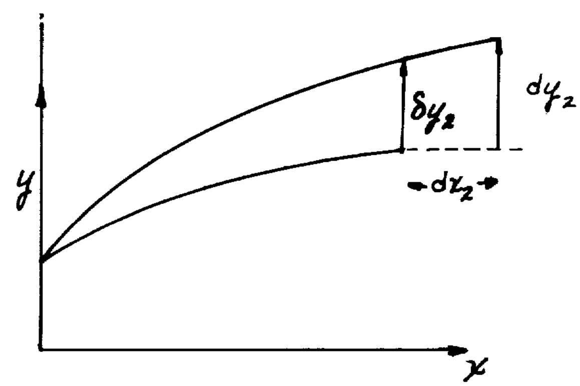
$$\int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} \delta y + \frac{\partial F}{\partial y'} \delta y' \right] \delta y\, dx = \frac{\partial f}{\partial y'} \delta y \Big|_{x_1}^{x_2} + \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] \delta y\, dx.$$

Now noting that the values of the functional are taken along extremals, i.e. $\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0$ and that the end point (x_1, y_1) is fixed, (i.e., $\delta y\big|_{x=x_1} = 0$) reduces this second term of Eq. (2.3.1) to

$$\int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\,\delta y + \frac{\partial f}{\partial y'}\,\delta y'\right] dx = \frac{\partial f}{\partial y'}\,\delta y\Big|_{x=x_2}$$

Before proceeding any further, it is considered necessary to note that $\delta y_2 \neq dy_2$ since δy_2 is an increment of y at $x = x_1$ whereas dy_2 is the change of the y coordinate of an extremal produced at the point $x = x_2$, see sketch). The relations between δy_2 and dy_2 are therefore

$$\delta y_2 \approx (dy - y'dx)_{x=x_2}$$



Consequently, $\int_{x_2}^{x_2+\delta x_2} f\,dx \approx f\Big|_{x=x_2} dx_2$ so that

$$\int_{x_1}^{x_2}[f(x, y+\delta y, y'+\delta y') - f(x, y, y')\,dx \approx \frac{\partial f}{\partial y'}\Big|_{x=x_2}\left(dy_2 - y'(x_2)\,dx_2\right),$$

where all approximate equations (symbolized by $\approx$) hold true to within infinitesimals of second or higher order in dx_2 or dy_2 .

Finally, it follows that Eq. (2.3.1) reduces to

$$\delta I = f(x, y, y')\Big|_{x=x_2} dx_2 + \frac{\partial f}{\partial y'}(x, y, y')\Big|_{x=x_2} (dy_2 - y'(x_2)\, dx_2)$$

$$= \left(f - y' \frac{\partial f}{\partial y'}\right)\Big|_{x=x_2} dx_2 + \frac{\partial f}{\partial y'}\Big|_{x=x_2} dy_2 .$$

But the fundamental necessary condition $\delta I = 0$ implies that the following end-point condition must exist.

$$\left(f - y' \frac{\partial f}{\partial y'}\right)\Big|_{x=x_2} dx_2 + \frac{\partial f}{\partial y'}\Big|_{x=x_2} dy_2 = 0 \tag{2.3.2}$$

or for the case where dx_2 and dy_2 are independent

$$\left(f - y' \frac{\partial f}{\partial y'}\right)\Big|_{x=x_2} = 0 \tag{2.3.3 a}$$

$$\frac{\partial f}{\partial y'}\Big|_{x=x_2} = 0 \tag{2.3.3 b}$$

These equations are generally referred to as the "transversality condition" associated with the free point (x_2, y_2).

For example, assume that the end point (x_2, y_2) is moving along the curve $y = \varphi(x)$. Then $dy_2 \approx \varphi'(x_2) dx_2$ and the transversality condition (2.3.2) becomes

$$\left[f + (\varphi' - y') \frac{\partial f}{\partial y'}\right]_{x=x_2} dx_2 = 0$$

or, since dx_2 varies arbitrarily

$$\left[f + (\varphi' - y') \frac{\partial f}{\partial y'}\right]_{x=x_2} = 0 \tag{2.3.4}$$

If the left boundary point (x_1, y_1) is now held free while (x_2, y_2) is kept fixed, identical arguments lead to the following transversality condition:

$$\left(f - y'\frac{\partial f}{\partial y'}\right)\Big|_{x=x_1} dx_2 + \frac{\partial f}{\partial y'}\Big|_{x=x_1} dy_2 = 0 \tag{2.3.5}$$

And, if the end point (x_1, y_1) moves along a curve $y = \Psi(x)$, then condition (2.3.5) reduces to

$$\left[f + (\Psi' - y')\frac{\partial f}{\partial y'}\right]_{x=x_1} = 0 \tag{2.3.6}$$

Revising the analyses which led to Eq. (2.3.2) to include the more general case in which both end points (x_1, y_1) and (x_2, y_2) are free, produces the following general transversality condition:

$$\left[\left(f - y'\frac{\partial f}{\partial y'}\right)dx + \frac{\partial f}{\partial y'}\, dy\right]_{x_1}^{x_2} = 0 \tag{2.3.7}$$

Thus, if the two end point (x_1, y_1) and (x_2, y_2) are independent of one another, then both Eqs. (2.3.2) and (2.3.5) hold. And if the end points (x_1, y_1) and (x_2, y_2) both lie on the curves $y = \Psi(x)$ and $y = \varphi(x)$, respectively, then condition (2.3.7) reduces to conditions (2.3.6) and (2.3.4).

The transversality conditions obtained above provide the boundary conditions that the extremal must satisfy and along with E-L equations, they constitute the complete solution of the variational problem considered here.

Example 1: Consider the functional

$$I = \int_{x_1}^{x_2} F(x,y)\sqrt{1+y'^2}\, dx$$

with $y(x_1) = y_1$ while (x_2, y_2) moves along $y_2 = \varphi(x_2)$

The transversality condition (2.3.4) is given by

$$\left[F(x,y)\sqrt{1+y'^2} + (\varphi' - y')\frac{F(x,y)\,y'}{\sqrt{1+y'^2}}\right]_{x=x_2} = 0$$

or

$$\frac{F(x,y)(1+\varphi' y')}{\sqrt{1+y'^2}} = 0$$ implying therefore that $y'\varphi' = -1$

since $F(x, y) \neq 0$. This means that the transversality condition in this particular case reduces to the orthogonal condition.

Example 2: Consider the problem of extremizing $I = \int_0^{x_2} \frac{\sqrt{1+y'^2}}{y}\, dx$ with $y(0) = 0$ and $y_2 = x_2 - 5$. The solutions to the E-L equations are the circles $(x-a)^2 + y^2 = b^2$ and the boundary condition $y(0) = 0$ implies $a = b$. Thus, since Example 1 proved that the transversality condition reduces to orthogonality condition, it follows that $a = 5 = b$. Hence, $(x-5)^2 + y^2 = 25$ so that an extremum may occur only along the paths $y = \pm\sqrt{10x - x^2}$.

2.3.3 The Two and More Dependent Variable Cases

The concern here will be with the extremizations of functionals of the form

$$I(y, z) = \int_{x_1}^{x_2} f(x, y, z, y', z')\, dx,$$

where one or both of the boundary points $A(x_1, y_1, z_1)$ and $B(x_2, y_2, z_2)$ are free. To simplify matters, it will be assumed that the right end point B is free while the left boundary point A is held fixed. Then, as in the previous section, it is evident that an extremum may be attained only along the integrals of E-L equations:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \quad , \quad \frac{\partial f}{\partial z} - \frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right) = 0$$

The general solution to E-L equations in this case, involve four arbitrary constants, two of which may be found from the boundary condition imposed on the fixed point $A(x_1, y_1, z_1)$. The remaining two constants will be determined (as shown previously) from the fundamental condition: $\delta I = 0$ where

$$I = I(x_2, y_2, z_2)$$

For this case, the variation in the function (ΔI) is:

$$\Delta I = \int_{x_1}^{x_2+\delta x_2} f(x, y+\delta y, z+\delta z, y'+\delta y', z'+\delta z')\,dx - \int_{x_1}^{x_2} f(x, y, z, y', z')\,dx$$

$$= \int_{x_1}^{x_2} [f(x, y+\delta y, z+\delta z, y'+\delta y', z+\delta z') - f(x, y, z, y', z')]\,dx$$

$$+ \int_{x_2}^{x_2+\delta x_2} f(x, y+\delta y, z+\delta z, y'+\delta y', z'+\delta z')\,dx$$

And, as before, the application of the mean value theorem and Taylor's expansion theorem to the first of these integrals, yields:

$$\delta I = f\Big|_{x=x_2} d x_2 + \int_{x_1}^{x_2} \Big[\frac{\partial f}{\partial y}\,\delta y + \frac{\partial f}{\partial z}\,\delta z + \frac{\partial f}{\partial y'} + \frac{\partial f}{\partial z'}\,\delta z'\Big]\,dx$$

Now, integrating the last two terms (those involving the derivatives) by parts, produces

$$\delta I = f\Big|_{x=x_2} d x_2 + \Big[\frac{\partial f}{\partial y'}\,\delta y\Big]_{x=x_2} + \Big[\frac{\partial f}{\partial z'}\,\delta z\Big]_{x=x_2} + \int_{x_1}^{x_2} \Big\{\Big[\frac{\partial f}{\partial y} - \frac{d}{dx}\Big(\frac{\partial f}{\partial y'}\Big)\Big]\delta y$$

$$+ \Big[\frac{\partial f}{\partial z} - \frac{d}{dx}\Big(\frac{\partial f}{\partial z'}\Big)\Big]\,\delta z\Big\}\,dx$$

but since the integral I was considered along the extremals where the E-L equations are valid

$$\delta I = f\Big|_{x=x_2} \delta x_2 + \Big[\frac{\partial f}{\partial y'}\,\delta y\Big]_{x=x_2} + \Big[\frac{\partial f}{\partial z'}\,\delta z\Big]_{x=x_2}$$

Using now the fact that

$$\delta y\Big|_{x=x_2} \approx d y_2 - y'(x_2)\,d x_2\,,\quad \delta z\Big]_{x=x_2} \approx d z_2 - z'(x_2)\,d x_2\,,$$

and equating δI to zero reduces this equation to the condition

$$\left[f - y'\frac{\partial f}{\partial y'} - z'\frac{\partial f}{\partial z'}\right]_{x=x_2} dx_2 + \left[\frac{\partial f}{\partial y'}\right]_{x=x_2} dy_2 + \left[\frac{\partial f}{\partial z'}\right] dz_2 = 0 \quad (2.3.8)$$

This condition is called the "transversality" condition associated with the boundary point $B(x_2, y_2, z_2)$.

If the variations dx_2 , dy_2 , and dz_2 are independent, then condition (2.3.8) becomes

$$\left[f - y'\frac{\partial f}{\partial y'} - z'\frac{\partial f}{\partial z'}\right]_{x=x_2} = 0, \quad \left[\frac{\partial f}{\partial y'}\right]_{x=x_2}, \left[\frac{\partial f}{\partial z'}\right] = 0 \qquad (2.3.9)$$

If the boundary point $B(x_2, y_2, z_2)$ moves along a curve $y_2 = \varphi(x)$, $z_2 = \psi(x_2)$ so that $dy_2 = \varphi'(x_2)dx_2$ and $dz_2 = \psi'(x_2)dx_2$, then the transversality condition (2.3.8) reduces to

$$\left[f + (\varphi' - y')\frac{\partial f}{\partial y'} + (\psi' - z')\frac{\partial f}{\partial z}\right]_{x=x_2} = 0 \qquad (2.3.10)$$

However, this transversality condition for the problem of extremizing $I = \int_{x_1}^{x_2} f(x, y, z, y', z')\,dx$ along with the equations $y_2 = \varphi(x_2)$, $z_2 = \psi(x_2)$ is not sufficient to evaluate all arbitrary constants in the general solution of a system of Euler equations. The remaining required conditions are produced by considering that the boundary point $B(x_2, y_2, z_2)$ can move on a surface $z_2 = \varphi(x_2, y_2)$; thus, $dz_2 = \frac{\partial \varphi}{\partial x_2}dx_2 + \frac{\partial \varphi}{\partial y_2}dy_2$, where the variations dx_2 and dy_2 are arbitrary and independent. Now employing the condition $\delta I = 0$ transforms Eq. (2.3.8) to

$$\left[f - y'\frac{\partial f}{\partial y'} - z\frac{\partial f}{\partial z'} + \frac{\partial \varphi}{\partial x}\frac{\partial f}{\partial z'}\right]_{x_2} dx_2 + \left[\frac{\partial f}{\partial y'} + \frac{\partial f}{\partial z'}\frac{\partial \varphi}{\partial y}\right]_{z_2} dy = 0 \quad (2.3.11)$$

And, since δx_2 and δy_2 are independent, relation (2.3.11) reduces to

$$\left[f - y'\frac{\partial f}{\partial y'} + \left(\frac{\partial \varphi}{\partial x} - z'\right)\frac{\partial f}{\partial z'}\right]_{x_2} = 0\ ; \quad \left[\frac{\partial f}{\partial y'} + \frac{\partial f}{\partial z'}\frac{\partial \varphi}{\partial y}\right]_{x_2} = 0 \quad (2.3.12)$$

These two relations (Eq. (2.3.12)) together with $z_2 = \varphi(x_2, y_2)$, are in general, adequate to evaluate the two remaining arbitrary constants of the E-L equations.

Identical relations to those presented in Eqs. (2.3.8) and (2.3.12) for replacing (x_2, y_2, z_2) by (x_1, y_1, z_1) will hold when the left boundary point $A(x_1, y_1, z_1)$ is held free except that the point (x_2, y_2, z_2) will be replaced by the point (x_1, y_1, z_1) .

The extension to the general case follows from the consideration of

$$I = \int_{x_1}^{x_2} f(x, Y, Y')\, dx \equiv \int_{x_1}^{x_2} f(x; y_1, \cdots, y_n; y_1', \cdots, y_n')\, dx$$

with the end point $B(x_2, Y_2) \equiv B(x_2; y_{12}, y_{22}, \cdots, y_{n2})$ being free, using similar arguments to those given above. The following general transversality condition is stated without proof

$$\left[f - \sum_{i=1}^{n} y_i' \frac{\partial f}{\partial y_i'}\right] dx_2 + \left[\sum_{i=1}^{n} \frac{\partial f}{\partial y_i'}\right] dy_{i2} = 0 \tag{2.3.13}$$

As before, a similar relation holds at the left end point $x = x_1$ when it is free.

<u>Example 1:</u> Examine the extrema of the functional

$$I = \int_0^{x_2} [y'^2 + z'^2 + 2yz]\, dx$$

where $y(0) = z(0) = 0$ and the point (x_2, y_2, z_2) can move on the plane $x = x_2$.

The E-L equations are $z'' - y = 0$, $y'' - z = 0$. Hence,

$$y(x) = C_1 \cosh x + C_2 \sinh x + C_3 \cos x + C_4 \sin x$$

and

$$z(x) = C_1 \cosh x + C_2 \sinh x + C_3 \cos x - C_4 \sin x$$

But, since $y(0) = z(0) = 0$ it follows that $C_1 + C_3 = 0$, $C_1 - C_3 = 0$ and thus $C_1 = C_3 = 0$.

The transversality condition Eq. (2.3.8) reduces in this case (due to the independency of dy_2 and dz_2 and since $dx_2 = 0$) to

$$\left.\frac{\partial f}{\partial y'}\right|_{x=x_2} = 0 \qquad \left.\frac{\partial f}{\partial z'}\right|_{x=x_2} = 0,$$

i.e., $y_1(x_2) = 0$ and $z'(x_2) = 0$. This fact yields

$$C_2 \cosh x_2 + C_4 \cos x_2 , \quad C_2 \cosh x_2 - C_4 \cos x_2 = 0$$

Thus, if $\cos x_2 \neq 0$, then $C_2 = C_4 = 0$ so that extremum may occur only along the line $y = 0$, $z = 0$. On the other hand, if $\cos x_2 = 0$, $x_2 = \frac{\pi}{2} + 2n\pi$, $n = \pm 1, \pm 2, \ldots$, and it follows that $C_2 = 0$ and that C_4 is an arbitrary constant. Therefore, the solution is

$$\bar{y} = C_4 \sin x \;, \quad \bar{z} = C_4 \sin x$$

It is readily verified that this latter case satisfies $\delta I = 0$ for arbitrary C_4.

2.3.4 Higher Derivative Considerations

The problem under consideration here is that of extremizing the functional

$$I = \int_{x_1}^{x_2} f(x, y, y', y'')\, dx ,$$

where it is assumed that the values of the function and its derivative at one boundary point is given: say $y(x_1) = y_1, y'(x_1) = y_1'$,

while the quantities $y(x)$ and $y'(x)$ vary at the other end point (x_2, y_2, y_2'). Once again, the curve $y = y(x)$ which extremizes I satisfies the Euler equation:

$$f_y - \frac{d}{dx}(f_{y'}) + \frac{d^2}{dx^2}(f_{y''}) = 0$$

The general solution of this fourth-order differential equation involves four arbitrary constants: $y(x) = y(x, c_1, c_2, c_3, c_4)$. Two of these constants may be obtained from the given boundary values of y and y' at the point $x = x_1$ which are fixed in advance: $y = (x_1) = y_1$, $y'(x_1) = y_1'$. The two remaining constants may be obtained from the condition $\delta I = 0$ as follows: First note that

$$\Delta I = \int_{x_1}^{x_2+\delta x_2} f(x, y+\delta y, y'+\delta y', y''+\delta y'')dx - \int_{x_1}^{x_2} f(x, y, y', y'')dx$$

$$= \int_{x_2}^{x_2+\delta x_2} f(x, y+\delta y, y'+\delta y', y''+\delta y'')\,dx$$

$$+ \int_{x_1}^{x_2} [f(x, y+\delta y, y'+\delta y', y''+\delta y'') \quad f(x, y, y', y'')]\,dx$$

Thus, by applying the mean value theorem, it follows that

$$\Delta I = f(x, y, y'')\Big]_{x_2} dx_2 + \int_{x_1}^{x_2} [f_y \delta y + f_y \delta y' + f_y .. \delta y'']dx + R,$$

where R is an infinitesimal of order higher than the maximum of the absolute values $|\delta x_2|$, $|\delta y_2|$, $|\delta y_2'|$, and $|\delta y_2''|$. Hence,

$$\delta I = f\Big|_{x_2} \delta x_2 + \int_{x_1}^{x_1} [f_y \delta y + f_{y'} \delta y' + f_{y''} \delta y'']dx$$

Integrating (by parts) the second term of the integrand once and the third term twice and then recalling that

$$\delta y\Big|_{x_1} = 0, \quad \delta y'\Big|_{x_1} = 0, \text{ and } \quad f_y - \frac{d}{dx}(f_{y'}) + \frac{d^2}{dx^2}(f_{y''}) \equiv 0,$$

yields

$$\delta I = [f \delta x + f_{y'} \delta y + f_{y''} \delta y' - \frac{d}{dx}(f_{y''}) \delta y]_{x=x_2}$$

Now using the relations

$$dy_2 = y'(x_2)\,dx_2 + [\delta y]_{x=x_2}$$

and

$$dy_2' = y''(x_2)\,dx_2 + [\delta y']_{x=x_2},$$

produces

$$\delta I = \left[f - y' f_{y'} - y'' f_{y''} + y' \frac{d}{dx}\left(f_{y''}\right)\right]_{x=x_2} dx_2 + \left[f_{y'} - \frac{d}{dx}\left(f_{y''}\right)\right]_{x=x_2} \delta y_2 + f_{y''}\Big]_{x=x_2} dy_2'$$

Consequently, the fundamental condition (the transversality condition) of an extremum $\delta I = 0$ takes the form

$$\left[f - y' f_{y'} - y'' f_{y''} + y' \frac{d}{dx}(f_{y''})\right] dx_2 + \left[f_{y'} - \frac{d}{dx}(f_{y''})\right]_{x_2} \delta y_2 + f_{y''}\Big] dy_2' = 0 \tag{2.3.14}$$

for the free end point (x_2, y_2, y_2'). Similar results exist for the other end point $x = x_1$.

As before, if dx_2, dy_2 and dy_2' are independent, then their coefficients should vanish at the point $x = x_2$. However, if there is some relation between them: say $y_2 = \varphi(x_2)$ and $y_2' = \psi(x_2)$ then $dy_2 = \varphi'(x_2)dx_2$ and $dy_2' = \psi'(x_2)\, dx_2$. Substituting these values into Eq. (2.3.14) and noting that dx_2 is arbitrary, reduces Eq. 2.3.16

$$\left\{f - y' f_{y'} - y'' f_{y''} + y' \frac{d}{dx}(f_{y''}) + \left[f_{y'} - \frac{d}{dx}(f_{y''})\right]\varphi' + f_{y''}\psi'\right\}_{x=x_2} = 0 \tag{2.3.15}$$

This condition along with the conditions $y_2 = \varphi(x_2)$, $y_2' = \psi(x_2)$ are in general sufficient to determine the values of x_2, y_2 and y_2'.

If x_2, y_2, y_2' are related through one equation $\varphi(x_2, y_2, y_2') = 0$ then two of the variations dx_2, dy_2, dy_2' are arbitrary and the remaining one is given by

$$\varphi'_{x_2}\, dx_2 + \varphi'_{y_2}\, dy_2 + \varphi'_{y_2'}\, dy_2' = 0$$

or

$$dy_2' = \frac{-\varphi'_{x_2}\, dx_2 - \varphi'_{y_2}\, dy_2}{\varphi_{y_2'}} \qquad \varphi'_{y_2'} \neq 0$$

Substituting this equation into Eq. (2.3.14) and recalling that the coefficients of the independent variation dx_2 and dy_2 must vanish, yields two relations at the point $x = x_2$ which in general, together with $\varphi(x_2, y_2, y_2')$ allows the point (x_2, y_2, y_2') to be evaluated.

If the end point $A(x_1, y_1)$ is also variable, then the analogous conditions may be obtained for this point.

Example: Examine the extrema of the functional

$$I = \int_0^1 (1 + y''^2)\, dx , \qquad \text{where } y(0) = 0 ,\ y'(0) = 1 ,\ y(1) = 1$$

and $y'(1)$ is a variable.

Solution: The E-L equation here is

$$\frac{d^4 y}{dx^4} = 0$$

so that the extremum is

$$y = C_0 + C_1 x + C_2 x^2 + C_3 x^3$$

Now, employing the boundary conditions $y(0) = 0$, $y'(0) = 1$, and $y(1) = 1$ yield respectively $C_0 = 0$, $C_1 = 1$, and $C_2 + C_3 = 0$

But, since $dx_2 = dy_2 = 0$ and dy_2' arbitrary, it follows that the transversality condition (2.3.11) reduces to

$$f_{y''}\Big]_{x_2 = 1} = 0 \quad \text{or} \quad y''(1) = 0$$

yielding the relation

$$2C_2 + 6C_3 = 0$$

Since $C_2 + C_3 = 0$ then $C_2 = C_3 = 0$

Therefore, the extremum occurs along $y = x$.

2.3.5 Functionals Depending on Boundary Conditions

This section considers the more general problem of extremizing the functional

$$I = \int_{x_1}^{x_2} f(x, y, y')\, dx + \Phi(x_1, y_1, x_2, y_2) \tag{2.3.16}$$

with the end points (x_1, y_1), (x_2, y_2) subjected to the conditions: $y_1 = \varphi(x_1)$ and $y_2 = \psi(x_2)$

The method of solution is the same as for the solution of problems presented earlier, but with more complicated boundary conditions. Indeed, the extremum may occur only along the integral solutions of E-L equation:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0$$

Since the curve $y = \bar{y}(x)$ yields an extremum for I in Eq. (2.3.16) (the variable end-point problem), such a curve will also do the same with respect to a more restricted class of admissible curves, the end point of which coincides with the end points of the curve $y = \bar{y}(x)$. Further, along such curves the functional (2.3.16) differs from $\int_{x_1}^{x_2} f(x, y, y')\,dx$ by, at most, a constant term having no influence on the extremal properties of the functional.

By considering δI by methods analogous to those presented in previous sections, the variation is determined to be

$$\delta I = \left[\left(f - y'\frac{\partial f}{\partial y'}\right)dx_2 + \frac{\partial f}{\partial y'}\,\delta y\right]_{x=x_2} - \left[\left(f - y'\frac{\partial f}{\partial y'}\right)dx + \frac{\partial f}{\partial y'}\,\delta y\right]_{x=x_1} \qquad (2.3.17)$$

$$+ \frac{\partial \Phi}{\partial x}\,dx_1 + \frac{\partial \Phi}{\partial y'}\,dy_1 + \frac{\partial \Phi}{\partial x_2}\,dx_2 + \frac{\partial \Phi}{\partial y_2}\,dy_2 = 0$$

Now, if the boundary points (x_1, y_1) and (x_2, y_2) move along the curves $y_1 = \varphi(x_1)$, $y_2 = \psi(x_2)$ then $dy_1 = \varphi'(x_1)\,dx_1$ and $dy_2 = \psi'(x_2)\;x_2$. Thus $\delta I = 0$ reduces to

$$\left[f + \frac{\partial f}{\partial y'}(\psi' - y') + \frac{\partial \Phi}{\partial v_2} + \frac{\partial \Phi}{\partial y_2}\,\psi'\right]_{x=x_2} dx_2$$

$$-\left[f + \frac{\partial f}{\partial y'}(\varphi' - y') + \frac{\partial \Phi}{\partial x_1} - \frac{\partial \Phi}{\partial y'}\,\varphi'\right]_{x=x_1} dx_1 = 0$$

And since δx_1 and δx_2 are independent, two equations result

$$\left[f + \frac{\partial f}{\partial y'}(\psi' - y') + \frac{\partial \Phi}{\partial x_2} + \frac{\partial \Phi}{\partial y_2}\,\psi'\right]_{x=x_2} = 0 \qquad (2.3.18)$$

and

$$\left[f + \frac{\partial f}{\partial y'}(\varphi' - y') + \frac{\partial \Phi}{\partial x_i} - \frac{\partial \Phi}{\partial y_1}\varphi'\right]_{x=x_1} = 0 \tag{2.3.19}$$

The generalization to a vector function y is obvious. The transversality condition (2.3.17) is

$$\left[d\Phi + \left(f - \sum_{i=1}^{n} \frac{\partial f}{\partial y_i'} y_i'\right) dx + \sum_{i=1}^{n} \frac{\partial f}{\partial y_i'} dy_i\right]_{x_1}^{x_2} = 0 \tag{2.3.20}$$

2.4 VARIATIONAL PROBLEMS WITH CONSTRAINING CONDITIONS. FIXED AND FREE END POINTS CONSIDERATIONS

In the simpler variational problem considered earlier, the class of admissible extremal curves was specified (apart from certain smoothness requirements) by conditions imposed on the end points (either fixed or free) of the extremal curves. However, many applications of the calculus of variations lead to problems in which not only boundary conditions, but also conditions of quite a different type known as subsidiary conditions (synonymously, side conditions or constraints) are imposed on the extremal curves.

This section is concerned with the problem of extremizing some functional $I(Y)$ under three types of subsidiary conditions; (1) Function Constraints, (2) Differential Equations Constraints, (3) Integral Constraints. Each of the three types will be treated in detail.

2.4.1 Function Constraints

This section considers the problem of extremizing

$$I = \int_{x_1}^{x_2} f(x, Y, Y')\, dx \tag{2.4.1}$$

such that the conditions

$$F_j(x, Y) = 0 \quad , j = 1, 2, \cdots, m < n \tag{2.4.2}$$

are satisfied. As in finding maxima or minima of functions with side constraints (see section 2.1) the method of solution here calls for some function of the form of Lagrange multipliers.

Consideration of the fundamental necessary condition for an extrema, $\delta I = 0$ once again yields

$$\int_{x_1}^{x_2} \sum_{i=1}^{n} \left[\frac{\partial f}{\partial y_i} \delta y_i + \frac{\partial f}{\partial y_i'} \delta y_i' \right] dx = 0$$

But, the second term in each bracket can be integrated by parts. Further, if the two end points are fixed so that $[\delta Y_i]_{x = x_1, x_2} = 0$, then,

$$\int_{x_1}^{x_2} \sum_{i=1}^{n} \left[\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right)\right] \delta y_i \, dx = 0 \tag{2.4.3}$$

Now, since $y_1, \cdots, y_n$ are subjected to condition Eq. (2.4.3), the variations δy_i are not arbitrary (i.e., the Fundamental Lemma of the calculus of variations cannot be applied) and must satisfy the following relation (obtained from varying the constraining relations $F_j = 0$)

$$\sum_{i=1}^{n} \frac{\partial F_j}{\partial y_i} \delta y_i = 0 \quad , \quad j = 1, 2, \ldots, m \tag{2.4.4}$$

Multiplying each of these equations successively by $\lambda_j(x)$ and integrating from x_1 to x_2 yields

$$\int_{x_1}^{x_2} \lambda_j(x) \sum_{i=1}^{n} \frac{\partial F_j}{\partial y_i} \delta y_i \, dx = 0 \quad , j = 1, 2, \ldots, m.$$

Now adding all m equations along with Eq. (2.4.3) produces

$$\int_{x_1}^{x_2} \sum_{i=1}^{n} \left(\frac{\partial f}{\partial y_i} + \sum_{j=1}^{m} \lambda_j(x) \frac{\partial F_j}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right)\right) \delta y_i \, dx = 0$$

Finally, adopting the notation

$$f^* = f + \sum_{j=1}^{m} \lambda_j(x) F_j ,$$

the integral assumes the form

$$\int_{x_1}^{x_2} \sum_{i=1}^{n} \left[\frac{\partial f^*}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y_i'}\right)\right] \delta y_i \, dx = 0 \tag{2.4.5}$$

It is noted again that the fundamental lemma of the calculus of variations cannot be applied for this problem since the δy_i are not independent. To overcome this dependence, the m multipliers $\lambda_1(x), \ldots, \lambda_m(x)$ are chosen satisfying the m equations

$$\frac{\partial f^*}{\partial y_j} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y_j'}\right) = 0 \quad , \; j = 1, 2, \ldots, m \tag{2.4.6}$$

or

$$\frac{\partial f}{\partial y_j} + \sum_{i=1}^{m} \lambda_i(x) \frac{\partial F_i}{\partial y_j} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_j'}\right) = 0 \; , j = 1, 2, \ldots, m$$

These m equations (based on Eq. (2.4.4)) state that n-m from among the variations δy_i may be considered independent; for instance

$$\delta y_{m+1}, \ldots, \delta y_n .$$

Equations (2.4.6) constitute a system of linear equations with respect to $\lambda_i(x)$ with non-vanishing Jacobian

$$\frac{J(F_1, \ldots, F_m)}{J(y_1, \ldots, y_n)} \neq 0$$

Hence, the system Eq. (2.4.6) has a solution for $\lambda_1(x), \ldots, \lambda_m(x)$. With this choice of $\lambda_j(x)$, the fundamental necessary condition $\delta I = 0$ in Eq. (2.4.3) takes the form

$$\int_{x_1}^{x_2} \sum_{j=m+1}^{n} \left[\frac{\partial f^*}{\partial y_j} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y'_j} \right) \right] \delta y_j \, dx = 0$$

Next, since for functions $y_1(x), \ldots, y_n(x)$ that make the functional $[I]$ have an extremum, this functional equation becomes an identity even if δy_j, $j = m+1, m+2, \ldots, n$ are arbitrary. It follows, therefore, that the fundamental lemma of the calculus of variations can be applied. Accordingly, setting all but one of the $\delta y_j = 0$ and applying the fundamental lemma, and repeating this procedure (n-m) times yields:

$$\frac{\partial f^*}{\partial y_j} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y'_j} \right) = 0, \qquad j = m+1, m+2, \ldots, n. \tag{2.4.7}$$

Summarizing, the functions which make the functional [I] have a constrainted extrema, and the multipliers, $\lambda_1(x), \ldots, \lambda_m(x)$ satisfy Eq. (2.4.2), (2.4.6) and (2.4.7). This important result is noted in the following theorem.

Theorem 2.4.1

A set of functions $\bar{y}_1(x) \cdots \bar{y}_n(x)$ which yields an extremum to the functional Eq. (2.4.1) with constraining conditions Eq. (2.4.2) satisfies for suitably chosen multipliers $\lambda_1(x) \cdots \lambda_m(x)$, the E-L equations for the functional

$$I^* = \int_{x_1}^{x_2} \left[f + \sum_{j=1}^{m} \lambda_j(x) F_j \right] dx \equiv \int_{x_1}^{x_2} f^* dx \tag{2.4.8}$$

The functions $\lambda_j(x)$ and $\bar{y}_i(x)$ are determined from the E-L equations:

$$\frac{\partial f^*}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y'_i} \right) = 0, \qquad i = 1, 2, \ldots, n \tag{2.4.9}$$

$$F_j(x, y) = 0, \qquad j = 1, 2, \ldots, m.$$

The equation $F_j(x, y) = 0$ are assumed to be independent with non-vanishing Jacobian.

Example: Find the shortest distance between two points $A(x_1, y_1, z_1)$ and $B(x_2, y_2, z_2)$ lying on a surface $F(x, y, z) = 0$.

Solution: The distance between two points on a surface is given by $I = \int_{x_1}^{x_2} \sqrt{1 + y'^2 + z'^2} \, dx$ and the constraining equation is $F(x, y, z) = 0$.

So, in view of theorem 2.4.1, form $I^{*} = \int_{x_1}^{x_2} \left[\sqrt{1+y'^2+z'^2} + \lambda(x) F(x,y,z)\right] dx$

where E-L equations are:

$$\lambda(x)\ \frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2+z'^2}}\right) = 0$$

$$\lambda(x)\ \frac{\partial F}{\partial z} - \frac{d}{dx}\left(\frac{z'}{\sqrt{1+y'^2+z'^2}}\right) = 0$$

$$F(x,y,z) = 0$$

This system of equations determines $\bar{y}(x)$, $\bar{z}(x)$ that may give a constrained minimum of I, as well as the multiplier $\lambda(x)$.

Specifically, let the surface $F(x, y, z) = 0$ be that of a sphere, i.e., $F \equiv x^2+y^2+z^2-a^2=0$. Then, substitution yields $x - c_2 y + c_1 z = 0$ which is the equation of a plane through the center of the sphere.

Similar results to those given in section 2.3 may be claimed here if the end points are free rather than fixed. The only change to be made involves replacing f with the auxiliary or (augmented) function f^*.

2.4.2 Differential Equation Constraints

In this section, the problem is to extremize the functional

$$I = \int_{x_1}^{x_2} f(x, Y, Y')\,dx \tag{2.4.1}$$

with the constraining relations being the differential equations

$$F_j(x, Y, Y') = 0 \quad , \qquad j = 1, 2, \dots, m \tag{2.4.11}$$

In this case,as in the case given in Section (2.4.1), a similar principle of undetermined multipliers may be proved stating that constrained extrema of the functional I can occur only along those curves that extremizes I*:

$$I^* = \int_{x_1}^{x_2} \left[f(x, Y, Y') + \sum_{j=1}^{m} \lambda_j(x) F_j(x, Y, Y')\right] dx = \tag{2.4.12}$$

$$= \int_{x_1}^{x_2} f^*(x, Y, Y')\,dx$$

That is, the complete solution to the problem presented here is given by the (m+n) system of equations

$$\frac{\partial f^*}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y_i'}\right) = 0 \quad , \qquad i = 1, 2, \ldots, n \qquad (2.4.13)$$

and

$$F_j(x\, ; y_1, \ldots, y_n\, ; y_1', \ldots, y_n') = 0 \quad , \qquad j = 1, 2, \ldots, m \qquad (2.4.11)$$

where

$$f^* = f + \sum_{j=1}^{m} \lambda_j(x) F_j \qquad (2.4.14)$$

The method of proof required is similar (except for minor changes) to the method given in proving theorem 2.4.1 when a weaker assertion is made: The curves along which the functional [I] attains constrained extrema are the extremals of the functional[I*], for suitably chosen $\lambda_j(x)$ given in Eq. (2.4.6) with f* given in Eq. (2.4.14).

2.4.3 Integral Constraints: The Isoperimetric Problem

The Isoperimetric Problem is a variational problem having integral constraints; i.e., it is concerned with finding extrema of a functional

$$I = \int_{x_1}^{x_2} f(x, y, y')\, dx \qquad (2.4.1)$$

with the conditions

$$\int_{x_1}^{x_2} F_j(x, y, y')\, dx = L_j \quad ; \qquad j = 1, 2, \ldots, m \qquad (2.4.15)$$

where L_j are constants.

By proper substitutions, isoperimetric problems may be reduced to problems of constrained extrema, considered in the preceeding section. Indeed, define

$$Z_j(x) = \int_{x_1}^{x_2} F_j\, dx \quad , \qquad j = 1, 2, \ldots, m \qquad (2.4.16)$$

so that $Z_j(x_1) = 0$, and $Z_j(x_2) = L_j$. Next, observe that

$$\frac{dZ_j}{dx} \equiv Z_j'(x) = F_j(x, y, y') \quad , \qquad j = 1, 2, \ldots, m$$

Thus, the integral isoperimetric constraints Eq. (2.4.15) can be replaced by the differential constraints

$$Z_j'(x) - F_j(x, y, y') = 0\, , \; j = 1, 2, \ldots, m \qquad (2.4.17)$$

and the problem is, therefore, completely solved: The solution of the Isoperimetric Problem is obtained by extremizing

$$I^* = \int_{x_1}^{x_2} \Big[f + \sum_{j=1}^{m} \lambda_j(x)(Z_j' - F_j)\Big] dx = \int_{x_1}^{x_2} f^* dx, \tag{2.4.18}$$

where

$$f^* = f + \sum_{j=1}^{n} \lambda_j(x)(Z_j' - F_j).$$

The E-L equations are thus,

$$\frac{\partial f^*}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y_i'}\right) = 0, \qquad i = 1, 2, \ldots, n \tag{2.4.19}$$

$$\frac{\partial f^*}{\partial z_j} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial z_j'}\right) = 0, \qquad j = 1, 2, \ldots, m$$

or

$$\frac{\partial f}{\partial y_i} + \sum_{j=1}^{m} \lambda_j(x) \frac{\partial f_j}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'} + \sum_{j=1}^{m} \lambda_j \frac{\partial f_j}{\partial y_i'}\right) = 0 \qquad i = 1, 2 \cdots n \tag{2.4.20}$$

$$\frac{d}{dt}\left(\lambda_j(x)\right) = 0, \qquad j = 1, 2, \ldots, m$$

The last m equations imply that all λ_j are **constants.** The first n equations of Eq. (2.4.20) are the E-L equations for the functional

$$I^{**} = \int_{x_1}^{x_2} \Big(f + \sum_{j=1}^{m} \lambda_j F_j\Big) dx \equiv \int_{x_1}^{x_2} f^{**} dx \tag{2.4.21}$$

The following result has, therefore, been obtained: The extremum of the functional $I = \int_{x_1}^{x_2} F(x, y, y')dx$ with constraining relations $\int_{x_1}^{x_2} F_j(x, y, y')dx = L_j$, $j = 1, \ldots, m$ is obtained from extremizing I^{**} of Eq. (2.4.21) where λ_j are constants.

Example: Find a curve of given length, L, which **encloses** a maximum area with one fixed end point. Using parametric representation, the area enclosed by the curve $x = x(t)$, $y = y(t)$ is given by

$$I = \frac{1}{2}\int_{t_1}^{t_2} (x\dot{y} - y\dot{x})\, dt$$

and the length of the curve is given by

$$\int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2}\, dt = L$$

To find the curve, form **the augmented function,** $f^* = \frac{1}{2}(x\dot{y} - y\dot{x}) + \lambda\sqrt{\dot{x}^2 + \dot{y}^2}$

Substituting this equation into the E-L equations

$$\frac{\partial f^*}{\partial x} - \frac{d}{dt}\left(\frac{\partial f^*}{\partial \dot{x}}\right) = 0, \qquad \frac{\partial f^*}{\partial y} - \frac{d}{dt}\left(\frac{\partial f^*}{\partial \dot{y}}\right) = 0$$

$$\dot{y} - \frac{d}{dt}\left(-\frac{y}{2} + \frac{\lambda\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = 0, \qquad -\frac{\dot{x}}{2} - \frac{d}{dt}\left(\frac{x}{2} + \frac{\lambda\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = 0$$

Direct integration of these equations yields

$$y - \lambda \dot{x}/\sqrt{\dot{x}^2+\dot{y}^2} = C_1 \; , \qquad x + \lambda \dot{y}/\sqrt{\dot{x}^2+\dot{y}^2} = C_2$$

where C_1 and C_2 are arbitrary constants.

Finally, note that $\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}}$ and solve for $(y - C_1)$ and $(x - C_2)$; Now, squaring and adding the equations obtained yields

$$(x - C_2)^2 + (y - C_1)^2 = \lambda^2$$

Thus, the curve in question is a circle and the Lagrange multiplier is the radius $(L = 2\pi\lambda)$. Since the location of the circle is immaterial, C_1 and C_2 remain arbitrary.

2.4.4 Constrained Extrema with Free End Points

A combination of arguments carried out in this section (2.4) together with those presented in Section 2.3 lead to the following free end conditions by substituting $f^* = f + \sum_{k=1}^{m} \lambda_k(x) F_k$ for f

in any of the three cases of this section. Thus, (1) if neither end point value is prescribed for extremizing I with any one of the constraints discussed, then

$$\frac{\partial f^*}{\partial y'} = 0 \text{ for } x = x_1 \text{ and } x = x_2$$

(2) if one end point lies on a curve $\varphi(x, y) = 0$, then

$$\left[\frac{\partial f^*}{\partial y'} - \frac{\left(\frac{\partial \varphi}{\partial y}\right) f^*}{\frac{\partial \varphi}{\partial x} + y' \frac{\partial \varphi}{\partial y}}\right]_{\substack{x = x_1 \\ x = x_2}} = 0$$

2.5 SUFFICIENT AND OTHER NECESSARY CONDITIONS FOR AN EXTREMUM

This section is concerned with deriving sufficient conditions for extrema of functionals and further necessary conditions. By means of these additional necessary conditions, it will be possible to obtain additional information as to whether the extremum under question is weak or strong.

2.5.1 Existence of Fields of Extremals. Jacobi Condition

2.5.1.1 Field of Extremals

The first step in deriving sufficient conditions for extrema of functionals is the introduction of the notion of a field of extremals.

Definition: Let there be a region D in the xy-plane. If to each point of D there is exactly one curve that belongs to the family of curves $y = y(x, c)$, then this family of curves is called a field. The slope $p(x, y)$ of the tangent to a curve $y = y(x, c)$ at a point (x, y), is called the slope of the field at the point (x, y).

Example: Let D be the circle $x^2 + y^2 = 1$ and let first the family of curves be the straight lines $y = x + c$. These straight lines do not intersect one another inside the circle and hence they constitute a field. The slope of this field is $p(x, y) = 1$. If, on the other hand, the family of curves had been the parabolas $y = (x - c)^2 - 1$, no field would exist inside the same circle since some of the parabolas intersect others.

Definition. If all the curves belonging to a family of curves $y = y(x, c)$ pass through a certain point (x_0, y_0) and if none of these curves intersect anywhere in a region D, except at the point (x_0, y_0), then, it is said that the family of curves $y = y(x, c)$ is a central field (central because it is centered at the point (x_0, y_0)).

Definition. If a field or a central field is generated by a family of extremals for some variational problem, then it is called a field of extremals.

Definition. A family $y_i = y_i(x, c_1, \ldots, c_n)$, $i = 1, 2, \ldots, n$ is a field in a region D of the space $x, y_1, \ldots, y_n$, if through each point' of D there passes exactly one curve of the family $y_i = y_i(x, c_1, \ldots, c_n)$. The slope functions denoted by $p_i(x, y_i, \ldots, y_n)$ are given by $p_i = \frac{\partial y_i}{\partial x}$, $i = 1, 2, \ldots, n$. The central field may be defined likewise.

Definition. Let the curve $y = y(x)$ yield an extremum for $I = \int_{x_1}^{x_2} f(x, y, y')dx$ where the end points (x_1, y_1) and (x_2, y_2) are fixed. The curve $y = y(x)$ is said to be admissible in a field of extremals, if there is a family of extremals $y = y(x, c)$ which is a field and which for a particular value of $C = C_0$ it becomes the extremal $y = y(x)$ not lying on the boundary of the region D in which the family of $y = y(x, c)$ is a field. A similar definition is valid for admissible curves in central fields. The slope of the tangent at the point (x_1, y_1) may be utilized as a parameter of the family.

Example. Let $I = \int_0^a (y'^2 - y^2)\, dx$. Find a central field of extremals admitting the extremal arc $y = 0$, joining the points $(0,0)$ and $(a, 0)$, $0 < a < \pi$.

E-L. equation is $y'' + y = 0$ so that $y = C_1 \cos x + C_2 \sin x$. The condition $y(0) = 0$ yields $C_1 = 0$ and hence $y = C_2 \sin x$. Therefore, in an interval $0 \le x \le a < \pi$, the family of curves $y = C_2 \sin x$ form a central field, including the case $C_2 = 0$. C_2 for this problem is $\frac{\partial y}{\partial x}(0, 0)$. If $a \geq \pi$, then the family $y = C_2 \sin x$ does not form a field since the family of curves intersect at the point $x = \pi$.

2.5.1.2 Conjugate Points and the Jacobi Conditions

Let $y = y(x, c)$ be the equation of a family of extremals eminating from a central point $A(x_1, y_1)$ (where the parameter C may be thought of as the slope y' of an extremal of the family at the point $A(x_1, y_1)$). But, the C-discriminant locus (on the envelope of the family of curves) is determined by the equations:

$$y = y(x, c), \qquad \frac{\partial y(x,c)}{\partial c} = 0$$

where $\frac{\partial y}{\partial c}(x, c)$ evaluated along an arbitrary curve (i.e., C = constant) of the family becomes a function of x only, i.e., $\frac{\partial y}{\partial c}(x, c) = u(x)$. Therefore, $u' = \frac{\partial^2 y}{\partial c \partial x}$. But, the function $y = y(x, c)$ is a solution of the E-L. equation implying that

$$\frac{\partial f}{\partial y}\left[x, y(x, c), y'_x(x, c)\right] - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}(x, y(x,c), y'_x(x, c\right] \equiv 0$$

Differentiating this identity with respect to C and setting $u = \frac{\partial y}{\partial c}(x, c)$ yields

$$\left[\frac{\partial^2 f}{\partial y^2} - \frac{d}{dx}\left(\frac{\partial^2 f}{\partial y \partial y'}\right)\right] u - \frac{d}{dx}\left(\frac{\partial^2 f}{\partial y \partial y'} u'\right) = 0 , \qquad (2.5.1)$$

The functions $\frac{\partial^2 f}{\partial y^2}(x, y, y')$, $\frac{\partial^2 f}{\partial y \partial y'}$, $\frac{\partial^2 f}{\partial y'^2}$ are linear functions of x, for $y = y(x, c_0)$ is a solution of E-L. equation with $C = C_0$. The linear homogeneous second order differential Eq. (2.5.1) is called Jacobi's equation.

If a solution of this equation $u = \frac{\partial y(x, c)}{\partial c}$ vanishes at the center of the family of curves, when $x = x_1$, (the center of the family always belongs to the C-discriminant locus), and vanishes at some other point x^* of the interval $x_1 \leq x \leq x_2$, then the point x^* is called a conjugate point to x_1, and is determined by:

$$y = y(x, c_0) \text{ and } \frac{\partial y}{\partial c}(x, c) = 0 \text{ or } u = 0$$

If there is a solution of Eq. (2.5.1) which vanishes at $x = x_1$, and does not vanish for any other point of the interval $x_1 \leq x \leq x_2$, then there is no point conjugate to $A(x_1, y_1)$, on the extremal arc AB; i.e., the Jacobi condition is satisfied and the extremal arc AB can be admitted into a field of extremals centered at $A(x_1, y_1)$.

Note: It can be shown that the Jacobi condition is a necessary condition (i.e., for a curve AB giving an extremum to [I], no conjugate point to the point A can be in the interval $x_1 < x < x_2$.

Example: Investigate the Jacobi condition for $I = \int_0^a (y'^2 - y^2)\, dx$ passing through $A(0,0)$ and $B(a,0)$.

The Jacobi equation is $u'' + u = 0$ so that $u = C_1 \sin(x - C_2)$. However, since $u(0) = 0$, it follows that $C_2 = 0$ and $u = C_1 \sin x$ (u vanishes at the points $k\pi$, $k = 0, 1, 2, \ldots$).

Thus, if $0 < a < \pi$, there is only one point $x = 0$ in the interval $0 \le x \le a$ at which $u = 0$ (Jacobi's condition holds). Jacobi's condition does not hold for $a \ge \pi$.

2.5.2 Weierstrass Conditions

Consider the simplest problem of extremizing

$$I = \int_{x_1}^{x_2} f(x, y, y')\,dx, \qquad y(x_1) = y_1, \; y(x_2) = y_2$$

where the Jacobi condition holds (i.e., the extremal arc C can be admitted into a central field with a slope function $p(x, y)$). The variation in the functional for this problem can be written as

$$\Delta I = \int_{x_1\,C^*}^{x_2} f(x, y, y')\,dx - \int_{x_1\,C}^{x_2} f(x, y, y')\,dx$$

where C^* is a neighboring arc to C.

Now consider an auxiliary functional:

$$I^* = \int_{C^*} \left[f(x, y, p) + \left(\frac{dy}{dx} - p \right) \frac{\partial f}{\partial p}(x, y, p) \right] dx$$

which becomes $\int f(x, y, y')\,dx$ along $C^* = C$, for $\frac{dy}{dx} = p$ along extremals of the field and, note that I^* is the integral of an exact differential:

$$I^* = \int_{C^*} \left[f(x, y, p) - p \frac{\partial f}{\partial p}(x, y, p) \right] dx + \frac{\partial f}{\partial p}(x, y, p)\,dy \quad (2.5.2)$$

In fact (see the transversality conditions in Section 2.3 :

$$dI^* = \left[f(x, y, y') - y' \frac{\partial f}{\partial y'}(x, y, y') \right] dx + \frac{\partial f}{\partial y'}(x, y, y')\,dy$$

Therefore, $\int_{C^*}[f+(y'-p)\frac{\partial f}{\partial p}]dx = \int_C f(x,y,y')dx$

Now since $\int_{C^*}[f+(y'-p)\frac{\partial f}{\partial p}]dx$ is an integral of an exact differential and is therefore independent of the path of integration, it follows that:

$$\int_C f(x,y,y')\,dx = \int_{C^*}\left[f(x,y,p)+(y'-p)\frac{\partial f}{\partial p}\right]dx$$

not only when $C^* = C$, but for arbitrary C^*. Thus,

$$\Delta I \equiv \int_{C^*} f(x,y,y')dx - \int_C f(x,y,y')dx$$

$$= \int_{C^*} f(x,y,y')\,dx - \int_{C^*}\left[f(x,y,p)+(y'-p)\frac{\partial f}{\partial p}(x,y,p)\right]dx$$

$$= \int_{C^*}\left[f(x,y,y')-f(x,y,p)-(y'-p)\frac{\partial f}{\partial p}(x,y,p)\right]dx$$

$$\equiv \int_{C^*} E(x,y,y',p)\,dx.$$

The integrand, denoted by $E(x,y,p,y')$, is called the Weierstrass function:

$$E(x,y,p,y') = f(x,y,y') - f(x,y,p) - (y'-p)\frac{\partial f}{\partial p}(x,y,p) \qquad (2.5.3)$$

It is obvious that the requirement that the function E be non-negative (together with the validity of the Jacobi condition) is a sufficient condition for I to have a minimum along the curve C (if $E \geq 0$, then $\Delta I \geq 0$). A sufficient condition of maximum is that $E \leq 0$, for then $\Delta I \leq 0$. For a weak minimum (see Section 2.2.1), it is sufficient to have $E(x,y,P,y') \geq 0$ (or $E \leq 0$ for maximum) for all values x, y that are close to the values of x and y along the extremal C, and for all values of y' which are close to $P(x,y)$ along this same extremal. For a strong minimum this same inequality holds for the same values of x and y for arbitrary y'.

It can be shown that the Weierstrass condition is also necessary.
Example: Consider the extremization of

$$I = \int_0^a y'^3 dx \; , \quad y(0) = 0 \; , \quad y(a) = b \; , \quad a > 0 \; , \; b > 0 .$$

The extremals are $y = C_1 x + C_2$ which upon substitution of the boundary conditions yields $y = b/a \, x$. The family of lines $y = cx$ with center at (O,O) forms a central field including the extremal $y = b/a \, x$.

The Weierstrass function E is (Eq. 2.5.3)

$$E = (x, y, p, y') = y'^3 - p^3 - 3p^2(y'-p)^2 (y' + 2p) .$$

Along the extremal $y = b/a \, x$, the slope function is $p = b/a > 0$ thus, if y' takes on values that are close to $p = b/a$, then $E \geq 0$, and a sufficient condition for a weak minimum resulted. On the other hand, if y' can assume arbitrary values, then $y' + 2p$ has arbitrary sign, and the function E may change sign. The sufficient condition for a strong minimum does not, therefore, hold.

2.5.3 Legendre Condition

The Legendre conditions which will be derived below, will constitute a simpler check of sufficiency (also necessity) than the one of Weierstrass. As a first step, assume that $f(x, y, y')$ has third derivative with respect to y' and expand it in a Taylor series

$$f(x, y, y') = f(x, y, p) + (y'-p) \frac{\partial f}{\partial p}(x, y, p) + \frac{(y'-p)^2}{2!} \frac{\partial^2 f}{\partial y'^2}(x, y, q)$$

where q is a value between p and y'. Substituting this expansion for $f(x, y, y')$ in Eq. (2.5.3) yields

$$E(x, y, p, y') = \frac{(y'-p)^2}{2!} \frac{\partial^2 f}{\partial y'^2}(x, y, q) \tag{2.5.4}$$

Thus, E and $\frac{\partial^2 f}{\partial y'^2}(x, y, q)$ have the same sign.
Further, for weak extrema $\frac{\partial^2 f}{\partial y'^2}(x, y, q)$ should not change sign when x and y are close to the extremal in question if q is close to $p(x, y)$.

Therefore, if $\frac{\partial^2 f(x, y, y')}{\partial y'^2} \neq 0$ along the extremal arc C , then, since C is continuous, this function retains a given sign for all points which are close to the extremal C , and for all such values of y' that are close to the values of y' taken at the corresponding points of C . Hence, the condition $E \geq 0$ (or $E \leq 0$) may be replaced by the condition $\frac{\partial^2 f}{\partial y'^2} > 0$ (or $\frac{\partial^2 f}{\partial y'^2} < 0$) taken along the extremal C . This condition (the so-called Legendre conditions) is a necessary condition for an extremum since the Weierstrass condition is a necessary condition. By the same token, the Legendre condition is sufficient provided the Jacobi condition is valid.

Example: Consider $I = \int_0^a (y'^2 - y^2)\, dx$ with $y(0) = y(a) = 0$, $a > 0$. The E-L equation is $y'' + y = 0$ so that $y = C_1 \cos x + C_2 \sin x$.

Now substituting the boundary conditions yield $C_1 = C_2 = 0$, provided $a \neq k\pi$, where k is an arbitrary integer. Thus, if $a \neq k\pi$ an extremum occurs only along the line $y = 0$. If $a < \pi$, then as shown earlier, the family of extremals $y = C$ sin x centered at (O, O) forms a central field (i.e., when $a < \pi$, the Jacobi condition is satisfied).

As for the Legendre condition in the case $a < \pi$, since $f = y'^2 - y^2$ possesses a third derivative with respect to y' for arbitrary y' and since $\frac{\partial^2 f}{\partial y'^2} = 2 > 0$ regardless of y' , it follows that in this case $y = 0$ furnishes I with a strong minimum.

2.5.4. Sufficiency Criteria

This section summarizes the results of Sections 2.5.1 through 2.5.3 by listing sufficient conditions for a minimum of the functional.

(These conditions are also necessary.)

$$I = \int_{x_1}^{x_2} f(x, y, y')\,dx \qquad y(x_1) = y_1\,,\; y(x_2) = y_2$$

2.5.4.1 Weak Minimum

The following conditions are sufficient (jointly) for I to have a weak minimum

(i) Jacobi condition or
Existence of a field of extremals

(ii) Legendre condition: $\frac{\partial^2 f}{\partial y'^2} > 0$ along the extremal arc or

Weierstrass condition: $E(x, y, p, y') \geq 0$ for all the points (x, y) sufficiently close to the extremal arc, and all y' sufficiently close to p (x, y).

2.5.4.2 Strong Minimum

The following conditions are sufficient (jointly) for I to have a strong minimum

(i) Jacobi condition or the existence of fields of extremals including the extremal arc under question.

(ii) Legendre condition: $\frac{\partial^2 f}{\partial y'^2}(x, y, y') \geq 0$ for these points (x, y) which are close to the extremal under examination and for arbitrary values of y' , or

Weierstrass condition: $E(x, y, p, y') \geq 0$ for all points (x, y) sufficiently close to the extremal under examination, and arbitrary y'.

Sufficient conditions for a weak (strong) maximum, are obtained by reversing the sense of the inequalities.

2.5.5 Weierstrass and Legendre Conditions in r-Dimensional Space

The theory given earlier may be extended, without any essential changes, to functions of the form:

$$I[y] = \int_{x_1}^{x_n} f(x, y\ y')\,dx \equiv \int_{x_1}^{x_n} f(x; y_1, \ldots, y_n; y_1', \ldots, y_n')\,dx \quad (2.5.5)$$

with

$$y_i(x_1) = y_{i1}, \quad y_i(x_2) = y_{i2}, \qquad i = 1, 2, \ldots, n.$$

The Weierstrass function E takes the form:

$$E = f(x; y_1, \ldots, y_n; y_1', \ldots, y_n') - f(x; y_1, \ldots, y_n; p_1, \ldots, p_n) - \sum_{i=1}^{n} (y_i' - p_i) \frac{\partial f}{\partial p_i}(x; y_1, \ldots, y_n; p_1, \ldots, p_n), \quad (2.5.6)$$

where the p_i are slope functions subject to some restrictions.

The Legendre conditions $\frac{\partial^2 f}{\partial y'^2} \geq 0$ must be replaced by the set of inequalities:

$$\frac{\partial^2 f}{\partial y_1'^2} \geq 0, \quad \det \begin{vmatrix} \frac{\partial^2 f}{\partial y_1'^2} & \frac{\partial^2 f}{\partial y_1' \partial y_2'} \\ \frac{\partial^2 f}{\partial y_2' \partial y_1'} & \frac{\partial^2 f}{\partial y_2'^2} \end{vmatrix} \geq 0; \ldots; \quad (2.5.7)$$

$$
\det \begin{vmatrix} \frac{\partial^2 f}{\partial y_1' \partial y_1'} & \frac{\partial^2 f}{\partial y_1' \partial y_2'} & \cdots & \frac{\partial^2 f}{\partial y_1' \partial y_n'} \\ \frac{\partial^2 f}{\partial y_2' \partial y_1'} & \frac{\partial^2 f}{\partial y_2' \partial y_2'} & \cdots & \frac{\partial^2 f}{\partial y_2 \partial y_n'} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial y_n' \partial y_1'} & \cdots & \cdots & \frac{\partial^2 f}{\partial y_n' \partial y_n'} \end{vmatrix} \geq 0
$$

2.5.6. Second Variation and Legendre Condition

Legendre conditions for weak extrema may also be obtained from the sign of the second variations of the functional under investigation. This is proven as follows:

By Taylor theorem

$$
\begin{aligned}
\Delta I &= \int_{x_1}^{x_2} [f(x, y+\delta y, y'+\delta y') - f(x, y, y')] dx \\
&= \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial y'} \delta y'\right] dx + \frac{1}{2} \int_{x_1}^{x_2} \left[\frac{\partial^2 f}{\partial y^2} \delta y^2 + \frac{\partial^2 f}{\partial y \partial f'} \delta y \delta y' \right. \\
&\quad \left. + \frac{\partial^2 f}{\partial y'^2} \delta y'^2 \right] dx + R
\end{aligned} \tag{2.5.8}
$$

where R is an infinitesimal of order greater than two with respect to δy and $\delta y'$ (δy and $\delta y'$ are sufficiently small that the sign of the increment ΔI is determined by the sign of the term on the right side of the equation involving the lowest powers of δy and $\delta y'$). But, the first variation vanishes along an extremal, i.e.,

$$
\int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial y'} \delta y' \right) dx = 0
$$

Therefore, in general, ΔI and the second variation

$$
\delta^2 I \equiv \int_{x_1}^{x_2} \left[\frac{\partial^2 f}{\partial y^2} \delta y^2 + \frac{\partial^2 f}{\partial y \partial y'} \delta y \delta y' + \frac{\partial^2 f}{\partial y'^2} \delta y'^2 \right] dx \tag{2.5.9}
$$

have the same sign.

Now consider the integral

$$\int_{x_1}^{x_2} [\omega'(x)\,\delta y^2 + 2\,\omega(x)\,\delta y\,\delta y']\,dx = 0 \tag{2.5.10}$$

where ω (x) is an arbitrary differentiable function. The vanishing of the integral (2.5.10) is due to

$$\int_{x_1}^{x_2} [\omega'(x)\,\delta y^2 + 2\,\omega(x)\,\delta y\,\delta y']\,dx = \int_{x_1}^{x_2} \frac{d}{dx}(\omega\,\delta y^2)\,dx$$
$$= [\omega(x)\,\delta y^2]_{x_1}^{x_2} = 0$$

for $\delta y|_{x_1} = 0 = \delta y|_{x_2}$. Thus, adding the integral (2.5.10) to $\delta^2 I$ yields

$$\delta^2 I = \int_{x_1}^{x_2} \Big[\Big(\frac{\partial^2 f}{\partial y^2} + \omega'\Big)\delta y^2 + 2\Big(\frac{\partial^2 f}{\partial y\,\partial y'} + \omega\Big)\delta y\,\delta y' + \frac{\partial^2 f}{\partial y'^2}\,\delta y'^2\Big]dx \tag{2.5.11}$$

But, since ω (x) was chosen arbitrarily, let it be selected such that the integral in Eq. (2.5.11) becomes a perfect square (up to a certain factor) i.e., let ω (x) satisfy the equation

$$\frac{\partial^2 f}{\partial y'^2}\Big(\frac{\partial^2 f}{\partial y^2} + \omega'\Big) - \Big(\frac{\partial^2 f}{\partial y\,\partial y'} + \omega\Big)^2 = 0$$

With this choice of ω (x)

$$\delta^2 I = \int_{x_1}^{x_2} \frac{\partial^2 f}{\partial y'^2}\Big[\delta y' + \Big\{\Big(\frac{\partial^2 f}{\partial y\,\partial y'} + \frac{\omega}{\frac{\partial^2 f}{\partial y'^2}}\Big)\delta y\Big\}^2\Big]dx$$

and consequently, the signs of $\delta^2 I$ and $\frac{\partial^2 f}{\partial y'^2}$ are the same.

2.6 THE GENERAL PROBLEMS OF BOLZA, MAYER, AND LAGRANGE

This section is concerned with the so-called problems of Bolza, Mayer, and Lagrange, which were partially treated earlier (Sections 2.1 through 2.5) though no special name was given to them. These problems are the most general problems of the calculus of variations and their treatment here may be viewed as a summary of all of the previous sections. Even though these problems will be shown to be theoretically equivalent, the procedure will be to first introduce the Bolza problem and then to derive from this material the problems of Mayer and Lagrange. A complete discussion for the Bolza problem follows:

2.6.1 The Problem of Bolza

The general problem of Bolza may be formulated as follows: Consider the sequence of functions:

$$y_i(x) \qquad i = 1, 2, \cdots, n \qquad x_1 \leq x \leq x_2 \tag{2.6.1}$$

satisfying the constraining equations

$$F_j(x, Y, Y') = 0 \qquad j = 1, 2 \cdots, m < n \tag{2.6.2}$$

and which are consistent with the end-point conditions

$$\varphi_r(x_1, Y_1, x_2, Y_2) = 0 \qquad r = 1, 2 \cdots, p \leq 2n+2 \tag{2.6.3}$$

The problem of Bolza is then to find the special sequence of functions $y_i(x)$ which minimizes the functional

$$I = \int_{x_1}^{x_2} f(x, Y, Y')\, dx + \Phi(x_1, Y_1, x_2, Y_2) \tag{2.6.4}$$

By carefully reviewing the material presented in previous sections, it is noticed that the minimizing criteria [Eq (2.6.4)] and the constraining equations [Eq (2.6.2)] were treated earlier in Section 2.3.5 and Section 2.4.2, respectively.

Hence, the solution of Bolza's problem goes as follows:

(i) The Euler-Lagrange Equations

The extremal arcs $y_K(x)$ satisfy not only the boundary conditions [Eq (2.6.2)] but also the following Euler-Lagrange equations.

$$\frac{\partial f^*}{\partial y_K} - \frac{d}{dx}\left(\frac{\partial f^*}{\partial y'_K}\right) = 0 \qquad K = 1, 2 \cdots, n \tag{2.6.5}$$

where f^* is the augmented (or subsidiary) function given by

$$f^*(x, Y, Y') = f(x, Y, Y') + \sum_{j=1}^{m} \lambda_j(x) \; F_j(x, Y, Y') \tag{2.6.6}$$

and $\lambda_1(x), \cdots, \lambda_m(x)$ are the Lagrange multipliers.

The validity of assertion (i) above follows by combining results of both Section 2.3.5 and Section 2.4.2.

(ii) The Transversality (Natural) Condition

The system of differential equations involved consist of m constraining equations [Eq (2.6.2)] and n-Euler-Lagrange equations [Eq (2.6.5)] which are subjected to 2n+2 boundary conditions (the E-L. equations are a system of second order equations). Of these boundary conditions, p are supplied by Eq (2.6.3) and in view of Eq (2.3.20) and Section 2.4.2 the remaining 2n+2-p conditions are supplied by the following transversality (natural) condition:

$$\left[d\Phi + \left(f^* - \sum_{K=1}^{n} \frac{\partial f^*}{\partial y'_K} y'_K \right) dx + \sum_{K=1}^{n} \frac{\partial f^*}{\partial y'_K} d y_K \right]_{x=x_1}^{x=x_2} = 0 \tag{2.6.7}$$

This equation must be satisfied identically for all systems of infinitesimal displacements consistent with the boundary conditions [Eq (2.6.3)] .

For the particular case in which f* of Eq (2.6.6) is explicitly independent of x, the following first integral of E-L. equations is valid:

$$- f^* + \sum_{K=1}^{n} \frac{\partial f^*}{\partial y'_K} y'_K = C \tag{2.6.8}$$

where C is a constant of integration. Equation (2.6.8) follows immediately from Eq (2.2.14) when f is replaced by f*. By the same token, the transversality condition in this particular case reduces to

$$\left[d\Phi - C\,dx + \sum_{K=1}^{n} \frac{\partial f^*}{\partial y'_K} d y_K \right]_{x=x_1}^{x=x_2} = 0 \tag{2.6.9}$$

(iii) Erdmann-Weierstrass Corner Conditions

As it was indicated in Section 2.2 of this monograph, many variational problems are characterized by discontinuous solutions, i.e., solutions in which one or more of the derivatives (y_K') experience a finite jump at a finite number of points (called corner points). In view of Eq (2.2.15), therefore, the Erdmann-Weierstrass corner conditions for the problem of Bolza become:

$$\left.\frac{\partial f^*}{\partial y_K'}\right|_{x_0^-} = \left.\frac{\partial f^*}{\partial y_K'}\right|_{x_0^+} \qquad K = 1, 2, \cdots, n \tag{2.6.10a}$$

$$\left.\left(-f^* + \sum_{K=1}^{n} \frac{\partial f^*}{\partial y_K'} y_K'\right)\right|_{x_0^-} = \left.\left(-f^* + \sum_{K=1}^{n} \frac{\partial f^*}{\partial y_K'} y_K'\right)\right|_{x_0^+} \tag{2.6.10b}$$

These conditions must be satisfied at every corner point x_0 of the extremal arc.

(iv) The Legendre and Weierstrass Necessary Conditions

The final step in obtaining a complete solution for the Bolza problem is the consideration of the necessary condition due to Legendre (presented in Section 2.5) for the purpose of determining whether the functional I attains a minimum or a maximum value. Analytically, the Legendre condition states that a necessary condition for I in Eq (2.6.1) to attain a weak minimum is that the following inequality (called the Legendre-Clebsch condition) is satisfied at all points of the extremal arc consistent with Eq (2.6.2):

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f^*}{\partial y_i' \partial y_j'} \delta y_i' \, \delta y_j' \geq 0 \tag{2.6.11}$$

The consistency condition requires the $\delta y'$ to satisfy the perturbed form of the constraint equations:

$$\sum_{i=1}^{n} \frac{\partial F_j}{\partial y_i'} \delta y_i' = 0 \qquad j = 1, 2, \cdots, m$$

The Weierstrass necessary condition for the functional I to attain a strong minimum is that

$$\Delta f^* - \sum_{i=1}^{n} \frac{\partial f^*}{\partial y_i'} \Delta y_i' \geq 0 \tag{2.6.12}$$

for all systems of strong variations $\Delta y_i'$ consistent with Eq (2.6.2).

It should be noted that both the Legendre and Weierstrass conditions take a slightly modified form of all the derivatives y_i' $(i = 1, n)$ to not appear explicitly in the function f^* ; that is if there are one or more derivatives, y_k' , say, for which $\frac{\partial f^*}{\partial y_k'} \equiv 0$. In this case, Eq (2.611) becomes

$$\sum_{i \neq k} \sum_{j \neq k} \frac{\partial^2 f}{\partial y_i' \partial y_j'} \delta y_i' \delta y_j' + 2 \sum_{i \neq k} \frac{\partial^2 f^*}{\partial y_i' \partial y_k} \delta y_i' \delta y_k + \frac{\partial^2 f^*}{\partial y_h^2} \delta y_h^2 \geq 0 \qquad (2.611a)$$

with $\delta y_i'$ and δy_h required to satisfy the consistence equations

$$\sum_{i \neq h} \frac{\partial F_j}{\partial y_i'} \delta y_i' + \frac{\partial F_j}{\partial y} \delta y_k = 0 \qquad j = 1, m$$

while the Weierstrass condition becomes

$$f^*(x, y_i Y_i', Y_k) - f^*(x, y_i, y_i', y_k)$$

$$- \sum_{i \neq k} \frac{\partial f^*}{\partial y_i'} (Y_i' - y_i') - \frac{\partial f^*}{\partial y_k} (Y_k - y_k) \geq 0 \qquad (2.612a)$$

where Y_i' and Y_k must be consistent with the constraint equations

$$f^*(x, y_i Y_i', Y_k) = 0, \qquad j = 1, m$$

A similar situation holds if two or more derivatives do not appear explicitly in f^*.

2.6.2 The Problem of Lagrange

The problem of Lagrange has already been treated in Sections 2.4.2 and 2.6.1. In fact, it is a particular case of the Bolza problem occurring when

$$\phi = 0 \tag{2.6.13}$$

For this problem, Eq (2.6.5) through (2.6.12) still hold. In fact, the transversality condition [Eq (2.6.7)] is simplified since $d\phi = 0$.

2.6.3 The Problem of Mayer

The variational problem due to Mayer is that problem of Bolza for which the integral f is identically zero:

$$f \equiv 0 \tag{2.6.14}$$

For this problem, Eq (2.6.5) through (2.6.12) are still valid, with the further simplification that

$$f^* = \sum_{j=1}^{m} \lambda_j F_j (x, y, y') \tag{2.6.15}$$

for all admissible paths, and hence, for the extremal path.

The justification to this can be made either by means of the Bolza problem (as indicated above) or by means of Sections 2.3.5 and 2.4.2.

2.6.4 The Equivalency of the Problems of Bolza, Lagrange, and Mayer

It was noted in the preceding sections that the problems of Mayer and Lagrange are special cases of the Bolza problem. In the material which follows, it will be shown that the problems of Lagrange and Mayer reduce to one another as well as to the problem of Bolza, by simple substitutions, thus proving the equivalency of these three problems. Indeed, it is easy to see, with the help of very simple considerations, that the problem of Bolza is transformable into a problem of either Lagrange or Mayer with variable end points. First, it is equivalent to the problem of Mayer having a sequence of arcs

$$y_i(x) \quad ; \quad y_{n+1}(x) \qquad i = 1, 2 \cdots n \qquad x_1 \leq x \leq x_2$$

subjected to conditions of the form

$$F_j(x, Y, Y') = 0 \qquad j = 1, 2 \cdots m < n$$
$$y'_{n+1} - f(x, Y, Y') = 0$$

with boundary conditions

$$\varphi_r(x_1, Y_1, x_2, Y_2) = 0 \qquad r = 1, 2, \cdots p < 2n+2$$

$$y_{n+1}(x_1) = 0$$

and having the functional

$$I = \Phi + y_{n+1}(x_2)$$

to be minimized.

Secondly, it is equivalent to the problem of Lagrange with a sequence of arcs

$$y_i(x), \quad y_{n+1}(x) \qquad i = 1, 2, \cdots n \qquad x_1 \leq x \leq x_2$$

satisfying conditions of the form

$$F_j(x, Y, Y') = 0 \qquad y'_{n+1}(x) = 0$$

$$\varphi_r(x_1, Y_1, x_2, Y_2) = 0 \qquad y'_{n+1}(x_1) - \Phi/(x_2 - x_1) = 0$$

and with an integral to be minimized of the form

$$I = \int_{x_1}^{x_2} (f + y_{n+1})\, dx$$

Since the problem of Bolza contains the problems of Mayer and Lagrange as specializations, and since the set of transformations of the previous paragraphs produce the problems of Mayer and Lagrange from that of Bolza, the three problems are equivalent and possess the same degree of generality.

2.7 THE UNBOUNDED CONTROL PROBLEM

2.7.1 Introduction

The previous sections have dealt with the classical variational calculus in which the most general minimization problem considered was the problem of Bolza (2.6.1). This problem consisted of minimizing a functional of the form

$$I = \int_{x_1}^{x_2} f(x, y, y')\, dx + \phi(x_1, y_1, x_2, y_2)$$

subject to the differential constraints

$$F_j(x, y, y') = 0 \quad , \qquad j = 1, m$$

The independent variable is the scalar quantity x while the dependent variable is a vector with n components (i.e., $y = (y_1, y_2, \cdots y_n)$).

In the modern treatment of variational problems, the dependent vector y is assumed to consist of two distinct types of variables, namely "state" and "control" variables. This distinction is of little importance since all the theorems previously developed are valid regardless of how the dependent variables are chosen or renamed. However, the distinction is of some importance from both the physical and computational point of view.

The "state" of a physical system is usually defined as the least amount of information needed to know about the system at some instant in time in order to be able to predict the systems' future behavior. For example, the state of a point mass is given by its velocity and position. From a knowledge of these two quantities at some instant of time, the subsequent motion of the system is defined by Newton's second law. Thus, to control the system's motion it must be possible to control some of the forces that are acting. For example, control over a chemical rocket can be achieved by modulating the thrust vector of the vehicle either by steering the engine (direction control) or by throttling (magnitude control). The term "control variable" is used to describe the variables on which the controllable forces depend. In the chemical rocket problem, the control variables are the steering angle and the throttle setting.

Since the motion of a system must satisfy the differential equations for the motion (in the example, Newton's second law) the state variables are governed by a set of differential equations. The control variables appear in these equations as forcing functions and can take on any values subject to the constraints imposed on the system. Thus in the differential constraints (the $F_j\ (x, y, y') = 0$ in the Bolza problem) the derivatives of the state variables appear explicitly while the derivatives of the control variables do not. This fact leads to the requirement that the state of the system must be continuous in time (a sensible physical requirement) but allows for the control action to be discontinuous.

In the following sections, the term t is used to denote the independent variable, with the state and control vectors denoted by $x(t)$ and $u(t)$, respectively, (x is an n dimensional vector and u is an r dimensional vector). Furthermore, the differential equations of constraint are required to take the form

$$\dot{x}_i = g_i\,(t;\, x_1, x_2, \ldots, x_n;\ u_1, u_2, \ldots, u_n) \tag{2.7.1}$$

The theorems of the previous sections carry over directly with the change in nomenclature

$$x \longrightarrow t$$
$$y \longrightarrow \begin{matrix} x \\ u \end{matrix}$$
$$F_j\,(x, y, y') = 0 \longrightarrow \dot{x}_j - g_j\,(t, x, u) = 0$$

2.7.2 Problem Formulation

Consider the class of functions

$$x_i\,(t) \text{ and } u_j\,(t) \tag{2.7.2}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, r$ where each x_i is a state variable, and each u_j is a control variable. The x_i are constrained by the following set of differential equations which describe the dynamic process under consideration on the time interval of system operation from t_0 to t_1 :

$$\dot{x}_i = g_i\,(t;\, x_1, \ldots, x_n;\, u_1, \ldots, u_n) \equiv g_i\,(t, x, u) \tag{2.7.3}$$

for $i = 1, \ldots, n$.

The objective is to find that special set of extremal arcs (x° and u_j°) which minimize the following criterion of system performance:

$$I = G(x_1, \ldots, x_n)\Big|_{t=t_1} + \int_{t_0}^{t_1} H(t; x_1, \ldots, x_n; u_1, \cdots, u_r)\,dt \qquad (2.7.4)$$

This problem (referred to earlier as the Bolza problem with constraints) can be treated as indicated in the previous sections by introducing a set of variable Lagrange multipliers λ_k for $k = 1, \ldots, n$. The augmented function from Section 2.6.1, with suitable change in variables, then becomes

$$F = H + \sum_{i=1}^{n} \lambda_i(\dot{x}_i - g_i) \qquad (2.7.5)$$

The extremal arcs (the $x_1, \ldots, x_n$ and $u_1, \ldots, u_r$ which extremize I) must satisfy not only Eq. (2.7.3) but also the Euler-Lagrange Equations.

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}_i}\right) - \frac{\partial F}{\partial x_i} = 0 \qquad (2.7.6a)$$

$$\frac{\partial F}{\partial u_j} = 0 \qquad (2.7.6b)$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, r$.

The system of differential equations is composed of the differential equation constraints (Eq. (2.7.3)) and the Euler-Lagrange equations. It consists of $2n + r$ equations and unknowns. The solution yields the $n + r$ dependent variables $x_1, \ldots, x_n$; $u_1, \ldots, u_r$ and n Lagrange multipliers $\lambda_1, \ldots, \lambda_n$ simultaneously.

Most practical engineering problems require numerical methods to obtain a solution to the Eq. (2.7.3) and (2.7.6) since analytical solutions are possible only in special cases. An additional difficulty results from the fact that the variational problems of interest are always of the mixed boundary value type; that is, problems with conditions prescribed in part at the initial time and in part at the terminal time. Thus, in the case where closed form solutions cannot be obtained, trial and error techniques or directed search methods must be employed. These procedures consist of guessing the missing initial conditions or learning them through prediction techniques. The assumed initial conditions are then used in the numerical integration of the Euler-Lagrange equations and constraining equations, and the difference between the resulting final conditions and the specified final conditions are determined. Since these differences will not in general be zero, the process must be repeated several times until these differences are less than some specified amount.

The discussion in this monograph will be limited to a class of problems which are sufficiently simple as to be amenable to treatment by analytic techniques and yet realistic enough to be of practical interest.

2.7.3 Control Policy

A control policy is a mathematical function which specifies the control as a function of t, x_i and u_k for i = 1, ..., n and k = 1, ..., r. In this monograph the control policy satisfies only necessary conditions for a weak extremal of I. The methods of Section 2.5 can then be used to determine whether a weak minimum (or maximum) of I was actually obtained.

The optimal control policy plays a dominant role in determining the composition of the extremal arcs and is obtained by substituting Eq. (2.7.5) into (2.7.6b)

$$-\frac{\partial H}{\partial u_j} + \sum_{i=1}^{n} \lambda_i \frac{\partial g_i}{\partial u_j} = 0 \qquad j = 1, \ldots, r \qquad (2.7.7)$$

Now, the u_j (which satisfies Eq. (2.7.7) in terms of t, x_i and λ_i) satisfies the Euler-Lagrange equations and will be called the <u>optimal control</u>. The optimal control will be denoted by

$$u_j^o = f_j(t; x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_n) \qquad (2.7.8)$$

for j = 1, ..., r.

Eq. (2.7.8) is the optimal control policy which specifies u_j as a function of t, x_i and λ_i for i = 1, ..., n.

The optimal control policy (Eq. (2.7.8)) is used to reduce the system of differential equations to 2n equations. This is accomplished by substituting f_j for u_j. Using Equations (2.7.4), (2.7.5), (2.7.6), and (2.7.8), the system of differential equations then becomes

$$\dot{x}_i = g_i$$

$$\dot{\lambda}_i = \frac{\partial H}{\partial x_i} - \sum_{k=1}^{n} \lambda_k \frac{\partial g_k}{\partial x_i} \qquad (2.7.9a)$$

for i = 1, ..., n, where

$$g_i = g_i(t; x_1, \ldots, x_n; f_1, \ldots f_n)$$
$$H = H(t; x_1, \ldots, x_n; f_1, \ldots f_n) \tag{2.7.9b}$$

There are now 2n dependent variables $x_1, \ldots, x_n$ and $\lambda_1, \ldots, \lambda_n$ and one independent variable t. The augmented function F from Eq. (2.7.5) is given by

$$F = H(t; x_1, \ldots, x_n; f_1, \ldots, f_r)$$
$$+ \sum_{i=1}^{n} \lambda_i \left[\dot{x}_i - g_i(t; x_1, \ldots, x_n; f_1, \ldots, f_n)\right]$$

(This procedure of reducing the number of differential equations by using the optimal control policy is also applied in the bounded control problem of Section 2.8 and the optimal time problem of Section 2.9.)

Using the optimal control policy described in Eq. (2.7.8), the first integral, corner conditions and transversality condition can be written as follows (see Section 2.6).

First Integral

The first integral from Section 2.6.1 is given by

$$\frac{d}{dt}\left(-H + \sum_{i=1}^{n} \lambda_i g_i\right) + \frac{\partial F}{\partial t} = 0 \tag{2.7.10a}$$

For problems in which $\partial F/\partial t = 0$, the first integral reduces to

$$-H + \sum_{i=1}^{n} \lambda_i g_i = C \tag{2.7.10b}$$

where C is an integration constant.

Corner Conditions

When discontinuities occur in the derivatives of the solution of the variational problem (as in sections 2.8 and 2.9), a mathematical criterion is needed to join the different portions of the solution. This criterion is obtained by using Equations (2.7.5), (2.7.8), and (2.7.9) and the corner conditions from Section 2.6.1 as follows:

$$(\lambda_i)_- = (\lambda_i)_+ \tag{2.7.11a}$$

for i = 1, ..., n and

$$\left(-H + \sum_{i=1}^{n} \lambda_i g_i\right)_- = \left(-H + \sum_{i=1}^{n} \lambda_i g_i\right)_+ \tag{2.7.11b}$$

The negative and positive signs denote conditions immediately before and after a corner point, respectively.

Transversality Condition

Using Eq. (2.7.5), (2.7.8), and (2.7.9), the transversality condition from Section 2.6.1 is given by

$$\left[dG - \left(-H + \sum_{i=1}^{n} \lambda_i g_i\right) dt + \sum_{i=1}^{n} \lambda_i \, dx_i\right]\Bigg|_{t_0}^{t_1} = 0$$

and, for the special case in which $\partial F/\partial t = 0$, the transversality condition reduces to

$$\left[dG - C\,dt + \sum_{i=1}^{n} \lambda_i \, dx_i\right]\Bigg|_{t_0}^{t_1} = 0 \tag{2.7.12}$$

The dependent variables $x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_n$ can be found by solving the system of differential equations (2.7.9) and using Eq. (2.7.10), (2.7.11) and (2.7.12). When $x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_n$ are known, the optimal control variables $u_1, \ldots, u_r$ can be obtained from the optimal control policy described in Eq. (2.7.8).

After the extremal arcs have been evaluated, it is necessary to investigate whether the I attains a weak minimum or a maximum. In this connection, the second variation of I can be used to give a necessary condition for a weak minimum.

2.7.4 Second Variation

To apply the Legendre condition [Eq. (2.6.11)] the change in nomenclature

$$u_1 = \dot{z}_1$$
$$u_2 = \dot{z}_2$$
$$\vdots$$
$$u_r = \dot{z}_r$$

is introduced. This change is required since the development of Eq. (2.6.11) involved the tacit assumption that

$$\frac{\partial f^*}{\partial y_i'} \qquad i = 1, \ldots, n$$

are not identically zero. The second variation then requires that for a weak minimum

$$\sum_i^r \sum_K^r \frac{\partial^2 F}{\partial \dot{z}_i \partial \dot{z}_K} \delta \dot{z}_i \delta \dot{z}_K = \sum_i^r \sum_K^r \frac{\partial^2 F}{\partial u_i \partial u_K} \delta u_i \delta u_K \geq 0 \qquad (2.7.13)$$

at all points along the extremizing arc.

2.7.5 Example Problems

Utilizing the optimal control policy, a set of 2n differential equations can be formulated. The solution to these differential equations gives $x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_n$. Since the optimal control is a function of t; $x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_n$, it can now be evaluated analytically or synthesized by a controller. The analytic procedure is illustrated by the following three examples:

1. Linear first order, time invariant dynamic process.
2. Linear second order, time invariant dynamic process.
3. Nonlinear first order, time invariant dynamic process.

Example 1

Consider the control of the dynamic process shown in Figure 2.7.1 which is characterized by the differential equation

$$\dot{x}_1 = -2\sqrt{2}\, x_1 + u_1$$

Let the process be initially in the state $x_1(0) = 2$. The objective is to determine the control u_1 over the time interval from $t_0 = 0$ to $t_1 = 3$ so that the criterion of system performance

$$I = \tfrac{1}{2}\int_0^3 (x_1^2 + u_1^2)\, dt$$

has a minimum.

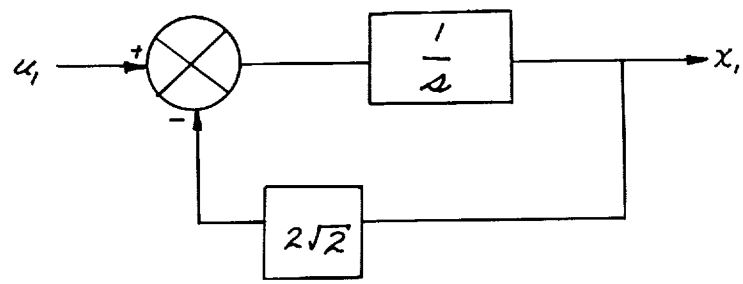


Figure 2.7. 1 Linear First Order System

The control attempts to optimize two aspects of dynamic performance. First, it attempts to minimize the area under the square of the response $x_1(t)$. This means that the control will tend to drive the $x_1(t)$ to zero. Secondly, it attempts to minimize the amount of control energy expended in performing the first task where it is assumed that the integral of the control squared is proportional to the control energy expended. Actually, u_1 can take on any value, but its amplitude is indirectly limited by the attempt to minimize the control energy expended. Unequal weighting between x_1^2, and u_1^2, can be specified depending upon which aspect of performance is considered more important; however for simplicity, equal weighting is assumed.

For this problem,

$$x = (x_1) \text{ with } n=1,\ u=(u_1) \text{ with } r=1$$

$$H = \tfrac{1}{2}(x_1^2 + u_1^2)\ ;\ G=0 \qquad \text{Eq.}[\ (2.7.4)]$$

$$g_1 = -2\sqrt{2}\, x_1 + u_1 \qquad \text{Eq.}[\ (2.7.3)]$$

Substituting these equations into Eq. (2. 7. 7), the optimal control policy becomes

$$-u_1 + \lambda_1 = 0$$

or from **Eq.** (2.7.8)

$$u_1 = f_1(\lambda_1) = \lambda_1$$

The system of differential equations from Eq. (2.7.9) then becomes

$$\begin{aligned} \dot{x}_1 &= -2\sqrt{2}\, x_1 + \lambda_1 \\ \dot{\lambda}_1 &= 2\sqrt{2}\, \lambda_1 + x_1 \end{aligned} \qquad (2.7.15)$$

where u_1 is replaced by λ_1 from the optimal control policy. There are now two dependent variables x_1 and λ_1 and one independent variable t.

The first integral from Eq. (2.7.10) is given by

$$-\tfrac{1}{2}(x^2 + \lambda_1^2) + \lambda_1(-2\sqrt{2}\, x_1 + \lambda_1) = C$$

The transversality condition from Eq. (2.7.12) is given by

$$\left[-C\,dt + \lambda_1\, dx_1 \right] \Big|_{t_0}^{t_1} = 0$$

This condition must be satisfied identically. That is,

$$-C\,dt_1 = 0 \qquad \lambda_1(t_1)\, dx_1(t_1) = 0$$

$$-C\,dt_0 = 0 \qquad \lambda_1(t_0)\, dx_1(t_0) = 0$$

But

$$dt_1 = 0\ , \quad dt_0 = 0 \ \text{ and } \ dx_1(t_0) = 0$$

because t_1, t_0 and $x_1(t_0)$ are specified and cannot be varied. The C and $\lambda_1(t_0)$ are arbitrary; however, $\lambda_1(t_1)$ must be zero since $x_1(t_1)$ is not specified. The boundary conditions now become

$$\begin{aligned} x_1(t_0) &= 2 & \qquad x_1(t_1) &= ? \\ \lambda_1(t_0) &= ? & \qquad \lambda_1(t_1) &= 0 \\ t_0 &= 0 & \qquad t_1 &= 3 \end{aligned}$$

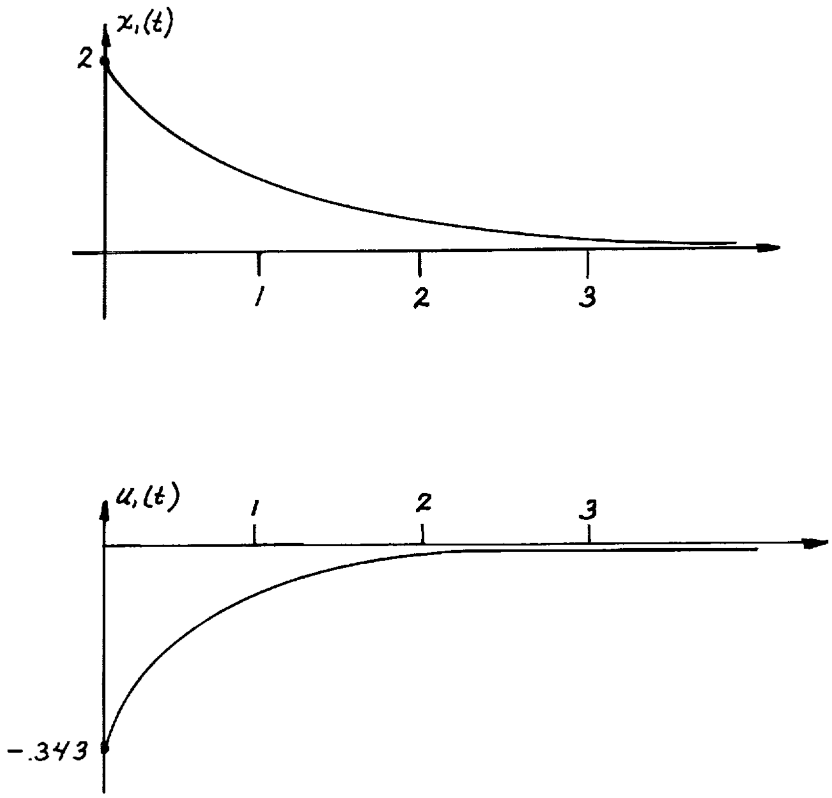


Figure 2.7.2 **Extremal Arcs**

where the question mark denotes an unspecified number.

Solving the system of differential equations from Eq. (2.7.15) subject to the boundary conditions, the following solutions are obtained:

$$x_1(t) = .0864(10^{-8})\, e^{3t} + 2e^{-3t}$$

$$\lambda_1(t) = u_1^\circ(t) = .508(10)^{-8} e^{3t} - .343 e^{-3t}$$

The second variation of I is obtained from Eq. (2.7.14) and is given by

$$(\delta u_1)^2 \geq 0$$

where the augmented function F from Eq. (2.7.5)

$$F = \tfrac{1}{2}(x_1^2 + u_1^2) + \lambda_1(\dot{x}_1 + \sqrt{2}\, x_1 - u_1)$$

was used.

Thus, this representation of $u_1^\circ(t)$ satisfies a necessary condition for a weak minimum of I since the second variation is ≥ 0. The resulting x_1 and u_1 are shown in Figure 2.7.2.

Example 2

Consider the control of the dynamic process shown in Figure 2.7.3 which is characterized by the differential equation

$$\ddot{x}_1 = \alpha^2 x_1 + u_1$$

or

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = \alpha^2 x_1 + u_1$$

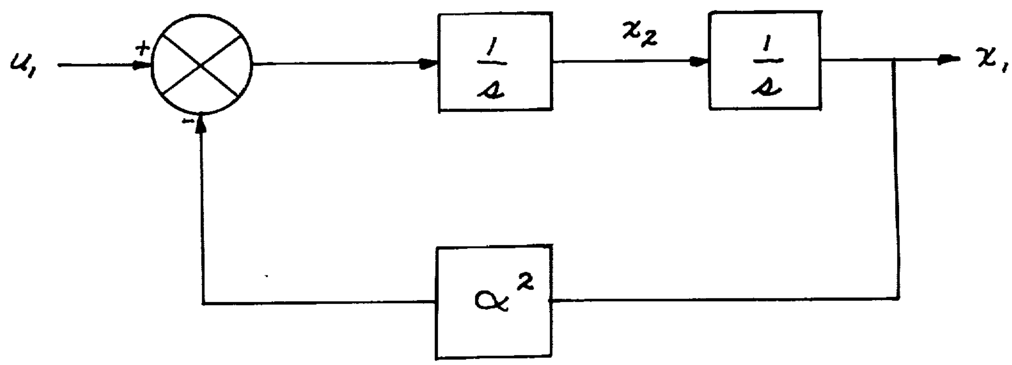


Figure 2.7.3 Linear Second Order System

Let the process be initially in the state $x_1(0) = A$ and $x_2(0) = B$. The objective is to determine the control u_1 over the time internal from $t_0 = 0$ to $t_1 = T$ so that $x_1(T) = 0$, $x_2(T) = 0$ and the criterion of system performance is minimized.

$$I = \tfrac{1}{2}\int_0^T u_1^2 \, dt$$

Note that $x_1(0)$, $x_2(0)$, $x_1(T)$, $x_2(T)$ and T are specified. In addition, let $\alpha T >> 1$. In this problem, the control attempts to minimize the amount of control energy expended in transforming the state of the system to the origin in state space. The control u_1 can take on any value, but its amplitude is indirectly limited by the criterion I.

For this problem

$x = (x_1, x_2)$ with $n = 2$, $u = (u_1)$ with $r = 1$

$H = \tfrac{1}{2} u_1^2$; $G = 0$ [Eq. (2.7.4)]

$g_1 = x_2$ [Eq. (2.7.3)]

$g_2 = \alpha^2 x_1 + u_1$

Substituting these equations into Eq. (2.7.7) the optimal control policy becomes

$$-u_1 + \lambda_2 = 0$$

or from Eq. (2.7.8)

$$u_1^\circ = f_1(\lambda_2) = \lambda_2$$

The system of differential equations from Eq. (2.7.9) then becomes

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \alpha^2 x_1 + \lambda_2 \\ \dot{\lambda}_1 &= -\alpha^2 \lambda_2 \\ \dot{\lambda}_2 &= -\lambda_1 \end{aligned} \tag{2.7.16}$$

where u_1 is replaced by λ_2 from the optimal control policy. There are now 4 dependent variables x_1, x_2, λ_1, and λ_2 and one independent variable t.

The first integral from Eq. (2.7.10) is given by

$$-\tfrac{1}{2}\lambda_2^2 + \lambda_1(\dot{x}_1 - x_2) + \lambda_2(\dot{x}_2 - \alpha^2 x_1 - \lambda_2) = C$$

The transversality condition (Eq. (2.7.12)) is given by

$$\left[-C\,dt + \lambda_1 dx_1 + \lambda_2\, dx_2\right]\Big|_{t_0}^{t_1} = 0$$

This condition specifies that

$$-C\,dt_1 = 0 \qquad\qquad \lambda_1(t_1)\,dx(t_1) = 0$$
$$-C\,dt_0 = 0 \qquad\qquad \lambda_1(t_0)\,dx(t_0) = 0$$
$$\lambda_2(t_1)\,dx_2(t_1) = 0$$
$$\lambda_2(t_0)\,dx_2(t_0) = 0$$

But

$$dt_1 = 0, \quad dt_0 = 0, \quad dx_1(t_1) = 0$$
$$dx_1(t_0) = 0, \quad dx_2(t_1) = 0, \quad dx_2(t_0) = 0$$

because

$$t_1 \,,\; t_0 \,,\; x_1(t_1) \,,\; x_1(t_0) \,,\; x_2(t_1) \,, \text{ and } x_2(t_0)$$

are specified and cannot be varied. Hence $C, \lambda_1(t_1), \lambda_1(t_0), \lambda_2(t_1)$ and $\lambda_2(t_0)$ are arbitrary, and the boundary conditions are given by

$$\begin{array}{ll} x_1(t_0) = A & \lambda_1(t_0) = ? \\ x_1(t_1) = 0 & \lambda_1(t_1) = ? \\ x_2(t_0) = B & \lambda_2(t_0) = ? \\ x_2(t_1) = 0 & \lambda_2(t_1) = ? \\ t_1 \quad = T & \\ t_0 \quad = 0 & \end{array}$$

Solving the system of differential equations from Eq. (2.7.16) subject to the boundary conditions, the following solutions are obtained when $\alpha T >> 1$:

$$x_1(t) = \left[A + (\alpha A + B)t\right] e^{-\alpha t}$$
$$x_2(t) = \left[B - \alpha(\alpha A + B)t\right] e^{-\alpha t}$$
$$\lambda_1(t) = -2\alpha^2(\alpha A + B)e^{-\alpha t}$$

$$\lambda_1(t) = u_1^\circ(t) = -2\alpha(\alpha A + B)e^{-\alpha t}$$

The second variation is obtained from Eq. (2.7.14) as

$$(\delta u_1)^2 \geq 0$$

where the augmented function F from Eq. (2.7.5) is

$$F = +\tfrac{1}{2}u_1^2 + \lambda_1(\dot{x}_1 - x_2) + \lambda_2(\dot{x}_2 - \alpha^2 x_1 - u_1)$$

Thus, this solution for $u_1(t)$ satisfies a necessary condition for a weak minimum of I (the second variation is ≥ 0).

Example 3

Consider the control of the dynamic process shown in Figure 2.7.4 which is characterized by the nonlinear differential equation

$$x_1 = -a x_1^2 + u_1$$

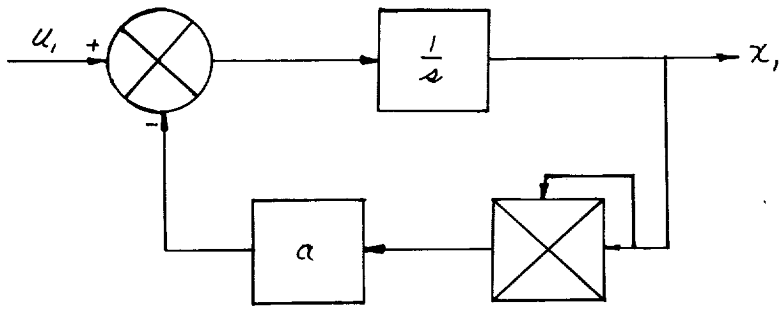


Figure 2.7.4 Nonlinear First Order System

Let the process be initially in the state $x_1(0) = A$. The objective is to determine the control u_1 over the time interval from $t_0 = 0$ to $t_1 = T$ so that the criterion of system performance

$$I = \frac{1}{2}\int_0^T (x_1^2 + u_1^2)\,dt$$

is minimized. The T and $x_1(o)$ are specified, but $x_1(t)$ is not. As in example 1, the control attempts to optimize two aspects of dynamic performance.

For this problem,

$$x = (x_1) \text{ with } n=1, \quad u = (u_1) \text{ with } r=1$$

$$H = \frac{1}{2}(x_1^2 + u_1^2)\ ;\ G = 0 \qquad \text{[Eq. (2.7.4)]}$$

$$g_1 = -a x_1^2 + u_1 \qquad \text{[Eq. (2.7.3)]}$$

Upon substituting these relations into Eq. (2.7.7), the optimal control policy becomes

$$-u_1 + \lambda_1 = 0$$

or the optimal control is [from Eq. (2.7.8)] given by

$$\mu_1^\circ = f_1(\lambda_1) = \lambda_1$$

The system of differential equations from Eq. (2.7.9) then becomes

$$\begin{aligned} \dot{x}_1 &= -a x_1^2 + \lambda_1 \\ \dot{\lambda}_1 &= 2a\lambda_1 x_1 + x_1 \end{aligned} \qquad (2.7.17)$$

where u_1 is replaced by λ_1 from the optimal control policy. There are now two dependent variables x_1 and λ_1, and one independent variable t.

The first integral from Eq. (2.7.10) is given by

$$-\tfrac{1}{2}(x_1^2 + \lambda_1^2) + \lambda_1(-a x_1^2 + \lambda_1) = C$$

The transversality condition from Eq. (2.7.12) is given by

$$\left[-c\,dt + \lambda_1 dx_1 \right] \Big|_{t_0}^{t_1} = 0$$

where the transversality condition must be satisfied identically. That is,

$$-c\,dt_1 = 0 \qquad \lambda_1(t_1)\,dx_1(t_1) = 0$$

$$-c\,dt_0 = 0 \qquad \lambda_1(t_0)\,dx_1(t_0) = 0$$

But

$$dt_1 = 0\,, \quad dt_0 = 0 \quad \text{and} \quad dx_1(t_0) = 0$$

because t_1, t_o and $x_1(t_o)$ are specified and cannot be varied. The C and $\lambda_1(t_o)$ are arbitrary; however, $\lambda_1(t_1)$ must be zero since $x_1(t_1)$ is not specified. The boundary conditions now become

$$\begin{aligned} x_1(t_0) &= A & \qquad x_1(t_1) &= ? \\ \lambda_1(t_0) &= ? & \qquad \lambda(t_1) &= 0 \\ t_0 &= 0 & \qquad t_1 &= T \end{aligned}$$

The solution to the system of differential equations from Eq. (2.7.17) is obtained by solving this two-point boundary value problem with the pair of simultaneous nonlinear ordinary differential equations. The solution gives $x_1(t)$ and $\lambda_1(t)$ for the optimal control $u_1^\circ(t) = \lambda_1(t)$. The use of a digital computer is required for this problem since there is no analytic solution.

2.8 CALCULUS OF VARIATIONS FOR THE BOUNDED CONTROL PROBLEM

2.8.1 Introduction

This section extends the optimal control problem of Section 2.7 to include the bounded control. The bounded control problem originates in engineering problems where the control variables are constrained by physical limitations.

For the bounded control problem, the class of admissible controls is enlarged by relaxing the requirement that the controls be smooth on open subintervals. The variational problem then becomes: Among all state variables which are piecewise smooth on the interval $[t_o, t_1]$ and all the control variables which are smooth on the open subintervals of (t_o, t_1) but not necessarily continuous on the closed interval $[t_0, t_1]$ find the state variables and controls for which the criterion has a weak extremum subject to amplitude constraints on the controls. This problem is treated by replacing the control variables with the optimal control policy (i.e., a function of the state variables and the Lagrange multipliers).

With the control variables replaced by functions of the state variables and Lagrange multipliers, the state variables which determine a weak extremum for the criterion are evaluated. The Euler-Lagrange equations must be satisfied on an open subintervals where 1) all the state and control variables are smooth, and 2) the criterion, the differential equation constraints, and the constraining equations are continuous with respect to the first and second partial derivatives of all its arguments. The corner conditions are then applied to tie together the subintervals.

Consider the interval from t_a to t_b. Let t_c $(t_a < t_c < t_b)$ represent a point which violates the condition of continuous differentiability. The criterion I on the interval from t_a to t_b can be written as the sum of two functionals

$$I(a,b) = I(a,c) + I(c,b)$$

where I (a, c) represents the criterion on the interval from t_a to t_c etc. The variation of I is calculated as two separate terms

$$\delta I(a,b) = \delta I(a,c) + \delta I(c,b)$$

The state variables and the Lagrange multipliers must join continuously at t_c, but otherwise the state at t_c can move freely. Note that the control variables neet not be continuous at t_c since the control variables are no longer in the variational problem. They have been replaced by functionals of the state variables and the Lagrange multipliers. However the corner conditions must be satisfied. This requirement places constraints on g and H from Eqs. (2.7.3) and (2.7.4). The resulting corner conditions require that Eqs. (2.7.11a) and (2.7.11b) be satisfied at t_c. Since the δ I (a, c) and δ I (c, b) involve the same increments $\delta \chi(t_c)$ and δt_c (χ denotes the state vector).

In a similar manner, the two subintervals can be extended to any finite number of subintervals. The corner conditions given in Eq (2.7.11a) and (2.7.11b) are then applied to connect the subintervals.

2.8.2 Problem Formulation

In this section and in Section 2.9 the control vector will be restricted as follows. First, the control vector $u \equiv (u_1, u_2, \cdots u_r)$ must lie within some closed set U at all times. This requirement is usually expressed in the form of a constraint on the u_k for $k = 1, \ldots, r$. For example, if r = 2, then a typical constraint is

$$\min u_1 \leq u_1 \leq \max u_1$$

$$\min u_2 \leq u_2 \leq \max u_2$$

for all time where minimum u_1, maximum u_1, minimum u_2 and maximum u_2 are given. Geometrically, this is interpreted as requiring the vector u to lie within or on the boundary of the rectangle as shown in Figure 2.8.1.

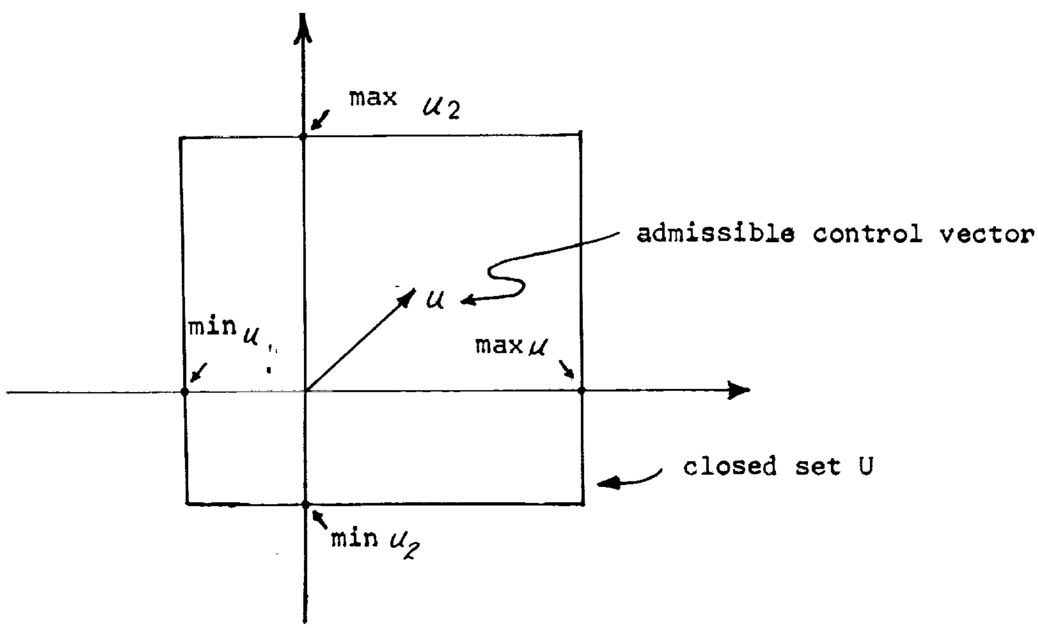


Figure 2.8.1 Admissible Control Region

Second, the u_k must be smooth on each open subinterval. Control variables which satisfy these two conditions will be called admissible controls. When some of the states of the process are specified, the admissible control must also give a solution to the system of differential equations that satisfies these specified states.

As in Section 2.7, consider the following dynamic process over the time interval from t_0 to t_1:

$$\dot{x}_i = g_i\,(t;\, x_1, \ldots x_n;\, u_1, \cdots, u_k)$$

for $i = 1, \cdots, n$. The objective is to find that special set of extremal arcs from X and the control arcs from the set of admissible controls which give a weak minimum for the following criterion:

$$I = G\,(x_1, \cdots, x_n)\Big|_{t=t_1} + \int_{t_0}^{t_1} H(t;\, x_1, \cdots, x_n;\, u_1, \cdots u_r)\, dt \qquad (2.8.1)$$

Control problems where Eq (2.8.1) is applicable are:

1. The initial and desired terminal states of the process are known. It is desired to find the control vector μ which minimizes $t_1 - t_0$, the time of system operation, subject to constraints on the control. This is the familiar time optimal problem and is discussed in Section 2.9.

2. The initial state of the process and the time interval $t_1 - t_0$ are known. It is desired to find the control vector μ which minimizes the value of I subject to constraints on the control.

2.8.3 Optimal Bounded Control Problem

If the set of differential constraint equations which describe the system, is given by

$$\dot{x}_i = g_i\,(t;\, x_1, \cdots, x_n;\, u_1, \cdots, u_k) \qquad i = 1, \ldots, n \quad (2.8.2)$$

and the constraints on the control variables are

$$\min u_k \le u_k \le \max u_k \qquad k = 1, \cdots, r$$

these constraints can be represented as equations

$$(u_k - \min u_k)(\max u_k - u_k) - \alpha_k^2 = 0 \qquad (2.8.3)$$

where α_k is an arbitrary real variable on the time interval from t_0 to t_1. Note that a u_k which satisfies Eq (2.8.3) is bounded above and below by maximum u_k and minimum u_k, respectively. The dependent variable α_k has no physical meaning in the problem and is used for convenience in generating a constraint equation.

Now, a set of variable Lagrange multipliers $\lambda_j(t)$ for $j=1,\ldots,n+r$ is introduced, and the augmented function F becomes

$$F = H + \sum_{i=1}^{n} \lambda_i(\dot{x}_i - g_i)$$

$$+ \sum_{k=1}^{r} \lambda_{k+n}\left[(u_k - \min u_k)(\max u_k - u_k) - \alpha_k^2\right] \qquad (2.8.4)$$

The dependent variables $x_1,\ldots,x_n; u_1,\ldots,u_r;$ and $\alpha_1,\ldots,\alpha_r$ must satisfy the constraint Equations (2.8.2) and (2.8.3) and the Euler-Lagrange equations. Using Eq (2.8.4), the Euler-Lagrange equations can be written as

$$\dot{\lambda}_i = \frac{\partial H}{\partial x_i} - \sum_{v=1}^{n} \lambda_v \frac{\partial g_v}{\partial x_i} \qquad (2.8.5a)$$

$$K_k(u_k) = \lambda_{k+n}(\max u_k + \min u_k - 2u_k) \qquad (2.8.5b)$$

$$0 = \lambda_{k+n}\alpha_k \qquad (2.8.5c)$$

for $k=1,\ldots,r$ and $i=1,\ldots,n$ where the term $K_k(u_k)$ is defined as

$$K_k(u_k) = -\frac{\partial H}{\partial u_k} + \sum_{i=1}^{n} \lambda_i \frac{\partial g_i}{\partial u_k}$$

The system of differential Equations (2.8.2), (2.8.3), and (2.8.5) consists of 2n+3r equations. Its solution yields n+2r dependent variables $x_1,\ldots,x_n; u_1,\ldots,u_r; \alpha_1,\ldots,\alpha_r$ and $n+r$ Lagrange multipliers $\lambda_1,\ldots,\lambda_{n+r}$.

These solutions are valid over each subinterval, if the corner conditions are satisfied.

In order to reduce the number of equations in the system of differential equations, the optimal control policy will be developed in Section 2.8.6. The optimal control policy specifies the optimal control u_k° for $k = 1, \ldots, r$ as a function of $t; x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_n$. This u_k° satisfies the Euler-Lagrange equations and the constraint equations on each piecewise smooth subinterval.

Now, if the u_k in the system of differential equations is replaced by the u_k°, the differential Eqs. (2.8.2) are no longer a function of u_k. Furthermore, the constraining Eqs. (2.8.3) are no longer required since the u_k° will satisfy the constraints. Consequently, only n Lagrange multipliers are required. The resulting system of differential equations will contain 2n equations for n dependent variables $x_1, \ldots, x_n$ and n Lagrange multipliers $\lambda_1, \ldots \lambda_n$.

2.8.4 Second Variation

Following a development similar to that used in Eq. (2.7.13) it can be shown that the Legendre condition takes the form

$$\sum_i^r \sum_j^r \frac{\partial^2 F}{\partial u_i \partial u_j} \delta u_i \delta u_j + \sum_i^r \sum_j^r \frac{\partial^2 F}{\partial \alpha_i \partial \alpha_j} \delta \alpha_i \delta \alpha_j \geq 0 \tag{2.8.6}$$

the quantities δu and $\delta \alpha$ must satisfy the variational equations

$$(\max u_k + \min u_k - 2u_k)\,\delta u_k - \alpha_k\, \delta\alpha_k = 0 \tag{2.8.7}$$

2.8.5 Selection of the Lagrange Multipliers

Consider the special variation in which $\delta u_i = 0 \quad (i = 1, \ldots r \quad)$. From Eq (2.8.7) it follows that either the $\delta\alpha_k$ or α_k are zero. Thus, from Eq (2.8.6) it follows that $\lambda_{k+n}\ (k = 1, \ldots, r)$ must be less than or equal to zero for a relative minimum. If α_k is zero, then from Eq 2.8.5c $\lambda_{k+n} = 0 (k = 1, \ldots, r \quad)$ Thus the Legendre condition yields the necessary condition

$$\lambda_{k+n} \leq 0 \qquad k = 1, \ldots, r$$

which must hold at each point along the extremal arc.

2.8.6 Control Policy

In Section 2.8.5 it was demonstrated that if $\lambda_{k+n} \leq 0$ for $k = 1, \dots, r$ a necessary condition for a weak minimum of I would be satisfied. The control policy will now be developed based on this necessary condition and will be called here the optimal control policy.

The λ_{k+n} for $k = 1, \dots, r$ can take on two values: $\lambda_{k+n} = 0$ or $\lambda_{k+n} < 0$. If $\lambda_{k+n} = 0$, the Eq (2.8.5b) becomes

$$K_k (u_k = w_k) = 0$$

where w_k is the value of u_k that makes $K_k = 0$ and

$$K_k(u_k) = \frac{-\partial H}{\partial u_k} + \sum_{i=1}^{n} \lambda_i \frac{\partial g_i}{\partial u_k} \tag{2.8.8}$$

$$\min u_k \leq w_k \leq \max u_k$$

If, on the other hand, $\lambda_{k+n} < 0$, the α_k from Eq (2.8.5c) must be zero. This means that u_k is either $\max u_k$ or $\min u_k$. That is, when $\alpha_k = 0$, the control constraint Eq (2.8.3) has $u_k = \max u_k$ or $\min u_k$ as a solution. Hence, with $\lambda_{k+n} < 0$, Eq (2.8.5b) becomes

$$\begin{aligned} &\text{for } u_k = \max u_k \\ &\qquad K_k(u_k = \max u_k) > 0 \\ &\text{for } u_k = \min u_k \\ &\qquad K_k(u_k = \min u_k) < 0 \end{aligned} \tag{2.8.9}$$

The optimal control policy is then given by

$$u_k^{\circ} = f_k\left(t; x_1, \dots, x_n; \lambda_1, \dots, \lambda_n\right) \tag{2.8.10a}$$

with

$$
f_k = \begin{cases} \max u_k & \text{whenever } K_k(u_k = \max u_k) > 0 \\ \min u_k & \text{whenever } K_k(u_k = \min u_k) < 0 \\ w_k & \text{whenever } K_k(u_k = w_k) = 0 \end{cases}
$$

where

$$
\min u_k \leq w_k \leq \max u_k
$$

Equation (2.8.10a) represents the optimal control policy and can be interpreted as follows: First, the optimal control is $\max u_k$ whenever

$$
\left[\frac{-\partial H}{\partial u_k} + \sum_{i=1}^{n} \lambda_i \frac{\partial g_i}{\partial u_k} \right] \Bigg|_{u_k = \max u_k} > 0
$$

Second, the optimal control is $\min u_k$ whenever

$$
\left[\frac{-\partial H}{\partial u_k} + \sum_{i=1}^{n} \lambda_i \frac{\partial g_i}{\partial u_k} \right] \Bigg|_{u_k = \min u_k} < 0
$$

Third, the optimal control is w_k whenever

$$
\left[\frac{-\partial H}{\partial u_k} + \sum_{i=1}^{n} \lambda_i \frac{\partial g_i}{\partial u_k} \right] \Bigg|_{u_k = w_k} = 0
$$

That is, w_k is the value of u_k such that the above expression is zero.

Many practical engineering problems have the following property: $\dot{K}_k(u_k = w_k) \neq 0$, where this statement can be written as

$$
\frac{d}{dt}\left(\frac{\partial H}{\partial u_k}\right) \neq \sum_{i=1}^{n} \left[\dot{\lambda}_i \frac{\partial g_i}{\partial u_k} + \lambda_i \frac{d}{dt}\left(\frac{\partial g_i}{\partial u_k}\right)\right]
$$

and where $u_k = w_k$. That is, when $K_k(u_k = w_k) = 0$, $\dot{K}_k(u_k = w_k) \neq 0$.

Under these conditions $K_k(u_k = w_k)$ cannot be zero for any length of time, and the optimal control policy is then given by

$$f_k = \begin{cases} \max u_k & \text{whenever} \quad K_k(u_k = \max u_k) > 0 \\ \min u_k & \text{whenever} \quad K_k(u_k = \min u_k) < 0 \end{cases} \tag{2.8.10b}$$

In other words, f_k is instantaneously changing between $\min u_k$ and $\max u_k$. For most practical systems, this change from $\min u_k$ to $\max u_k$ or from $\max u_k$ to $\min u_k$ will not produce a discontinuity in the state variables $x_1, \ldots, x_n$.

2.8.7 System of Differential Equations

The optimal control policy from Eq (2.8.10) is used to reduce the number of differential equations to 2n. This is accomplished by substituting f_k for u_k. The system of differential equations now becomes

$$\begin{aligned} \dot{x}_i &= g_i \\ \dot{\lambda}_i &= \frac{\partial H}{\partial x_i} - \sum_{k=1}^{n} \lambda_k \frac{\partial g_k}{\partial x_i} \end{aligned} \tag{2.8.11}$$

for $i = 1, \ldots, n$ where g_i and H are now given by

$$\begin{aligned} g_i &= g_i(t; x_1, \ldots, x_n; f_1, \ldots, f_n) \\ H &= H(t; x_1, \ldots, x_n; f_1, \ldots, f_n) \end{aligned}$$

The control constraint equations are no longer required since the optimal control u° satisfies the constraints. There are now n dependent variables $x_1, \ldots, x_n$, n Lagrange multipliers $\lambda_1, \ldots, \lambda_n$ and one independent variable t.

The first integral, corner conditions and transversality condition [Eq (2.7.10), (2.7.11), and (2.7.12)] can be simplified by replacing u_k with f_k (i.e., by using the optimal control policy). The solutions $x_1, \ldots, x_n$ for Eq (2.8.11) represent the extremal arcs if the corner conditions are satisfied.

2.8.8 Example Problems

Two example problems are presented. The first is a simplified linear second order model of an attitude control system with bounded control. The objective is to minimize the fuel expended in transforming the initial state of the system to the specified final state. The second example is a nonlinear thrust programming problem for a rocket in vertical flight. The objective is to find the propellant mass flow, which can be considered as the control, so that the final vertical rise of the rocket is maximized.

Example 1 Attitude Control System

Consider the control of the dynamic process shown in Figure 2.8.2 (a simplified model of an attitude control system) which is characterized by the differential equation

$$\ddot{x}_1 = u_1$$

or by

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = u$$

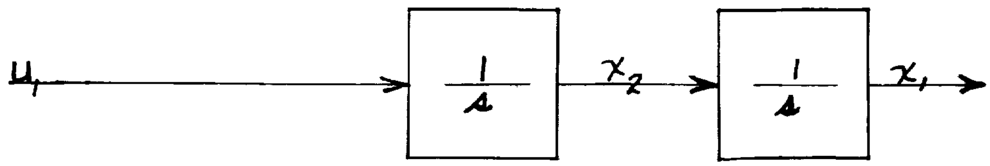


Figure 2.8.2 Linear Second Order System with Bounded Control

Let the process be initially in the state $x_1(0) = A$ and $x_2(0) = B$. The objective is to determine the control u_1 over the time interval from $t_0 = 0$ to $t_1 = T$ (where T is given) so that $x_1(T) = x_2(T) = 0$ and the criterion

$$I = \int_0^T |u_1| dt$$

is minimized

$$-1 \leq u_1 \leq 1$$

The control attempts to minimize the amount of control fuel expended in transforming the initial state of the system to the desired final state (it is assumed that the integral of the absolute value of the control is proportional to the control fuel expended).

For this problem.

$x = (x_1, x_2)$ with $n = 2$, $u = (u_1)$ with $r = 1$

$H = |u_1|$; $G = 0$ [Eq. (2.7.4)]

$g_1 = x_2$ [Eq. (2.7.3)]

$g_2 = u_1$

Substituting these equations into Eq (2.8.8) the $K_1(u_1)$ is given by

$$K_1(u_1) = -\frac{\partial |u_1|}{\partial u_1} + \lambda_2$$

The optimal control policy from Eq (2.8.10a) then becomes

$$u_1^\circ = f_1(\lambda_2) = \begin{cases} 1 = \max u_1 & \text{whenever} \quad \lambda_2 > 1 \\ -1 = \min u_1 & \text{whenever} \quad \lambda_2 < -1 \\ 0 = w_1 & \text{whenever} \quad -1 \leq \lambda_2 \leq 1 \end{cases}$$

where

$$K_1(u_1 = 1) = -1 + \lambda_2 > 0$$

$$K_1(u_1 = -1) = 1 + \lambda_2 < 0$$

$$K_1(u_1 = w_1) = \frac{-\partial |u_1 = w_1|}{\partial u_1} + \lambda_2 = 0$$

Figure 2.8.3 shows the optimal control policy as a function of λ_2.

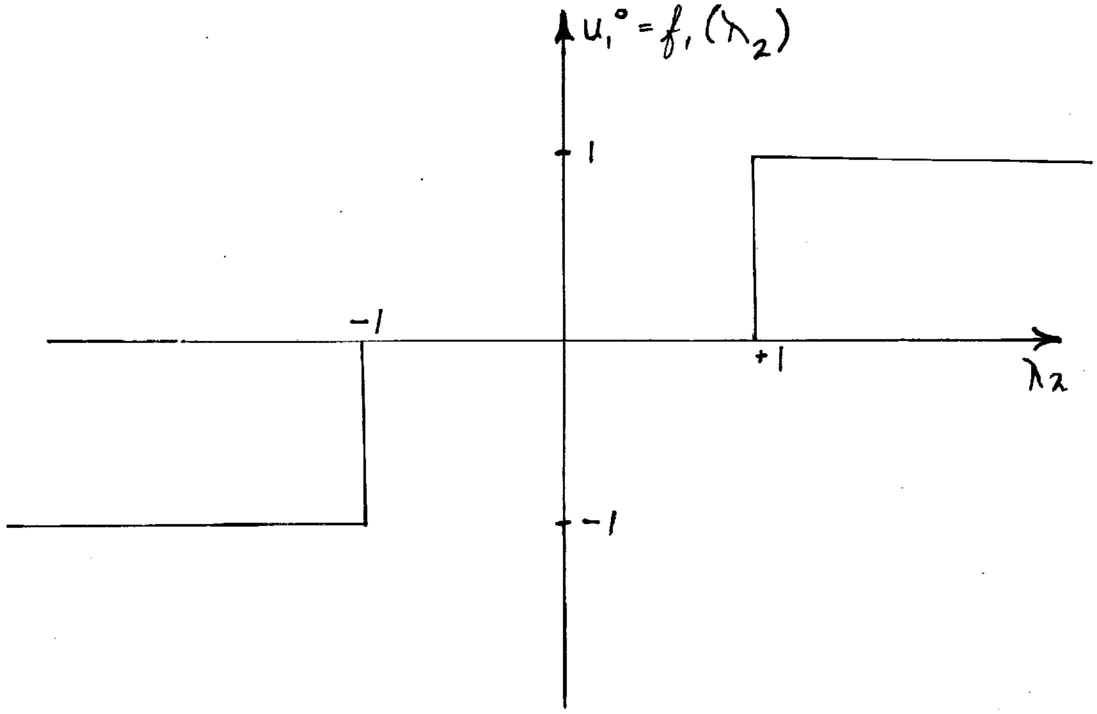


Figure 2.8.3 Relay with Deadband Optimal Control Policy as a Function of λ_2

Equation (2.8.10b) was not used here for the optimal control policy because K can be zero on a subinterval. In fact, $\dot{K} = 0$ for $t_2 < t < t_3$ where t_2 and t_3 are defined in Figure 2.8.4.

The system of differential Equations (2.8.11) then becomes

$$\dot{x}_1 = x_2 \qquad \dot{\lambda}_1 = 0$$

$$\dot{x}_2 = f_1(\lambda_2) \qquad \dot{\lambda}_2 = -\lambda_1 \tag{2.8.12}$$

where u_1 is replaced by $F(\lambda_2)$. There are now 4 dependent variables x_1, x_2, λ_1 and λ_2 and one independent variable t.

The first integral from Eq (2.7.10b) is given by

$$-\left| F(\lambda_2) \right| + \lambda_1 x_2 + \lambda_2 f_1(\lambda_2) = C$$

The corner conditions must be satisfied whenever the control is discontinuous and are obtained from Eq (2.7.11) as

$$(\lambda_1)_- = (\lambda_1)_+$$

$$(\lambda_2)_- = (\lambda_2)_+$$

$$(C)_- = (C)_+$$

Thus, λ_1, λ_2 and C must be continuous in order for the extremal arcs to be valid even though the control can be discontinuous. The transversality condition Eq (2.7.12) indicates that λ_1 and λ_2 at $t_0 = 0$ and $t_1 = T$ are arbitrary. The boundary conditions are

$$x_1(0) = A \qquad x_1(T) = 0$$

$$x_2(0) = B \qquad x_2(T) = 0$$

$$\lambda_1(0) = ? \qquad \lambda_1(T) = ?$$

$$\lambda_2(0) = ? \qquad \lambda_2(T) = ?$$

$$t_0 = 0 \qquad t_1 = T$$

Since $\dot{\lambda}_1 = 0$ from Eq (2.8.12) and λ_1 is continuous, λ_1 is constant even when the control is discontinuous. But, if λ_1 is constant, λ_2 from Eq (2.8.12) is linear with time as follows:

$$\lambda_2 = Dt + E$$

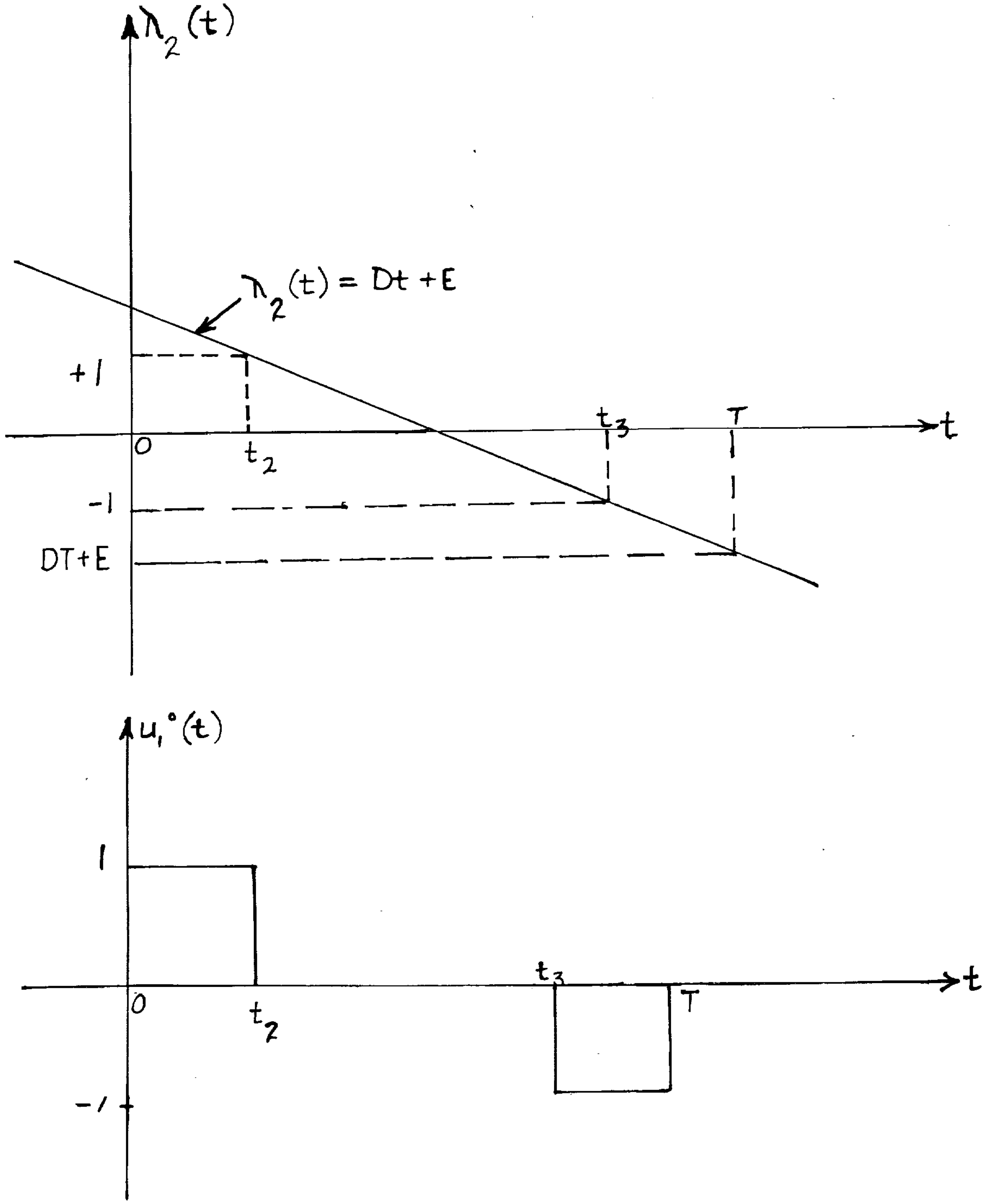


Figure 2.8.4 Possible Forms of $\lambda_2(t)$ and Optimal Control

where D and E are constants. One possible λ_2 as a function of time is shown in Figure 2.8.4. The time intervals from 0 to t_1, from t_1 to t_3, and from t_3 to T contain an applied control of 1, 0 and -1, respectively. Note that no more than 2 discontinuities are allowed in u_1^*.

The problem reduces to a two point boundary value problem which requires a digital computer for the solution since the system of differential equations is nonlinear. The computer procedure is based on selecting D and E by trial and error and solving Eq (2.8.12) repeatedly until the boundary conditions are satisfied. Once the boundary conditions are satisfied, the optimal control can be evaluated. Then the optimal control is substituted into the second variation of I to determine if the necessary condition for a weak minimum is satisfied.

Example 2 Rocket in Vertical Flight

The problem of determining a fuel burning program so that the maximum altitude is achieved for a rocket in vertical flight is of particular significance to this series of monographs. Several examples have been presented in the literature where the effect of aerodynamic forces have been considered. However, this example is confined to the limiting case in which these forces are negligible, so that the rocket moves under the combined effects of thrust and gravity. This limitation is required to assure that the solution will be analytic.

The following hypotheses are employed:

1. The earth is flat and the acceleration of gravity is constant.
2. The flight takes place in a vacuum.
3. The trajectory is vertical.
4. The thrust is tangent to the flight path.
5. The equivalent exit velocity of gases from the rocket engine is constant.
6. The engine is capable of delivering all mass flows between a lower limit and an upper limit.

Under these assumptions, the motion of the rocket vehicle is governed by

$$\begin{aligned} m\ddot{x}(t) &= T(t) - m(t)g \\ \dot{m}(t) &= -u(t) \end{aligned} \qquad (2.8.13)$$

where $T(t) =$ Thrust $= c\,u(t)$
$c =$ equivalent exit velocity
$u(t) =$ propellant mass flow $(0 \leq u \leq \max u)$
$m(t) =$ instantaneous rocket mass
$g =$ acceleration of gravity
$x(t) =$ altitude of rocket

This system of differential equations can be written as three first order differential equations as follows:

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = (c u_1)/x_3 - g$$

$$\dot{x}_3 = -u_1$$

where

$$x_1 = h \qquad x_2 = \dot{x}_1$$

$$x_3 = m \qquad u = u_1$$

Let the process be initially in the state $x_1(0) = 0$, $x_2(0) = 0$, $x_3(0) = m(0)$ (where m (o) is given), and $u_1(0) = A$ (where $0 < A \leq \max u$). The objective is to determine the control u_1 over the time interval from $t_0 = 0$ to t_1 where t_1 is not specified so that the criterion

$$I = -x_1(t_1)$$

is minimized. That is, find the propellant mass flow which maximizes the altitude obtained in vertical flight. (Note that maximizing $x_1(t_1)$ is equivalent to minimizing $-x_1(t_1)$.) This variational problem is subject to the constraint

$$0 \leq u_1 < \max u_1$$

The final state variables other than $x_2(t_1)$ are not specified $(x_2(t_1) = 0)$.

For this problem

$$x = (x_1, x_2, x_3) \qquad n = 3$$

$$u = (u_1) \qquad r = 1$$

$$H = 0 \quad ; \quad G = -x_1(t_1) \qquad [\text{Eq } (2.7.4)]$$

$$g_1 = x_2$$
$$g_2 = (cu_1)/x_3 - g \qquad \text{[Eq. (2.7.3)]}$$
$$g_3 = -u_1$$

Substituting these equations into Eq (2.8.8), the $K_1(u)$ is given by

$$K_1 = \lambda_2 c / x_3 - \lambda_3$$

Note that K_1 is not a function of u_1. The optimal control policy from Eq (2.8.10a) is given by

$$u_1^* = f(\lambda_2, x_3 \lambda_3) = \begin{cases} \max u_1 & , K_1 > 0 \\ 0 & , K_1 < 0 \\ w_1 & , K_1 = 0 \end{cases}$$

where

$$0 \leq w_1 \leq \max u_1$$

Using this optimal control policy, the system of differential Equations (2.8.11) then becomes

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = (c f_1)/x_3 - g$$
$$\dot{x}_3 = -f_1$$
$$\dot{\lambda}_1 = 0$$
$$\dot{\lambda}_2 = -\lambda_1$$
$$\dot{\lambda}_3 = (\lambda_2 c f_1)/x_3^2 - K_1 \frac{\partial f_1}{\partial x_3}$$

where u_1 is replaced by $f_1(\lambda_2, x_3 \lambda_3)$. There are now 6 dependent variables $x_1, x_2, x_3, \lambda_1, \lambda_2, \lambda_3$ and one independent variable t. Again, using the optimal control policy, the augmented function F is

$$F = \lambda_1(\dot{x}_1 - x_2) + \lambda_2\left(\dot{x}_2 - \frac{cf_1}{x_3} + g\right) + \lambda_3(\dot{x}_3 + f_1)$$

The first integral [Eq (2.7.10)] then becomes

$$\lambda_1 x_2 + \lambda_2 \left(\frac{c f_1}{x_3} - g \right) - \lambda_3 f_1 = C$$

where C is an integration constant. The corner conditions which must be satisfied whenever the control is discontinuous are obtained from Eq (2.7.11) as

$$(\lambda_i)_- = (\lambda_i)_+ \quad - \quad i = 1,2,3$$

and

$$(C)_- = (C)_+$$

and the transversality condition from Eq (2.7.12) is given by

$$\left[(\lambda_1 - 1) dx_1 - C\, dt + \lambda_2 dx_2 + \lambda_3 dx_3 \right] \Big|_{t_0}^{t_1} = 0$$

Since $x_1(t_1)$, $x_3(t_1)$ and t_1 were not specified,

$$C = \lambda_3(t_1) = 0$$

$$\lambda_1(t_1) = 1$$

The $\dot{\lambda}_1 = 0$, hence $\lambda_1 = 1$ for all time.

The optimal control policy states that $u_1^0 = w_1$, whenever $K_1 = 0$ where $0 \le w_1 \le \max u_1$. However, $K_1 \neq 0$ over any length of time. This fact can be shown by proving that $\dot{K}_1 \neq 0$ when $K_1 = 0$ as follows: With $K_1 = 0$, the K_1 can be written as

$$K_1 = \frac{\lambda_2 c}{x_3} - \dot{\lambda}_3 = 0$$

Thus, the derivative of K_1 with respect to time is given by

$$\dot{K}_1 = \frac{-\lambda_2 c \dot{x}_3}{x_3^2} + \frac{\dot{\lambda}_2 c}{x_3} - \ddot{\lambda}_3$$

But, from the system of differential equations with $K_1 = 0$ and $\lambda_1 = 1$,

$$\dot{\lambda}_3 = \frac{\lambda_2 c f_1}{x_3^2}$$

$$\dot{\lambda}_2 = -1$$

$$\dot{x}_3 = -f_1$$

Thus, $\dot{K}_1 = \dfrac{-c}{\chi_3}$

Now since $0 < \chi_3(t) \leq m(0)$ and $c > 0$, $\dot{K}_1 < 0$ when $K_1 = 0$,
K can be zero at any instant of time but cannot be zero over any subinterval. This means that $u_1^{\circ} \neq \omega_1$ over any subinterval and Eq (2.8.10b) must be used

$$u_1^{\circ} = f_1 = \begin{cases} \max u_1 & \text{whenever } K_1 > 0 \\ 0 & \text{whenever } K_1 < 0 \end{cases}$$

(At the instant of time at which K_1 passes through zero, f_1 is changing between $\max u_1$, or 0).

Now if f_1 is either $\max u_1$ or 0, $\partial f_1/\partial \chi_3$ is either zero or infinite. In the latter case, the set of extremal arcs terminates and a new set begins, and the corner conditions must be employed to connect the two set of arcs.

For a set of extremal arcs with arbitrary K_1, and $\partial f_1/\partial \chi_3 = 0$, it can be shown that $\dot{K}_1 = -c/\chi_3$. This means that $\dot{K}_1 = -c/\chi_3$ for any K_1. Further, since λ_2 and λ_3 are continuous from the corner conditions and χ_3 is continuous from physical reasoning, K_1 is continuous. Thus, a plot of K_1 versus t should resemble Figure 2.8.5 a (t_c denotes the instant of time where $K_1 = 0$ and when the u_1° changes from $\max u_1$ to 0). Figure 2.8.5 was drawn with the assumption that the initial $u_1^{\circ} = \max u_1$. This case doesn't represent a loss in generality since u_1° is either $\max u$ or 0. But, from physical reasoning the initial $u_1^{\circ} \neq 0$. The maximum number of changes in amplitude of u_1° is one since $\dot{K}_1 < 0$. Thus, the u_1° is

$$u_1^{\circ} = \begin{cases} \max u_1 & \text{for } 0 \leq t \leq t_c \\ 0 & \text{for } t_c < t \leq t_1 \end{cases}$$

The u_1° and χ_3 are shown in Figures 2.8.5b and 2.8.5c.

Solving for K_1, χ_2 and χ_3,

$$K_1 = \begin{cases} -\dfrac{c}{m(t_1)}(t - t_c), & \text{for } t_c < t \leq t_1 \\ \dfrac{c}{\max u} \log\left(\dfrac{m(0) - t \max u_1}{m(0) - t_c \max u_1}\right), & \text{for } 0 \leq t \leq t_c \end{cases}$$

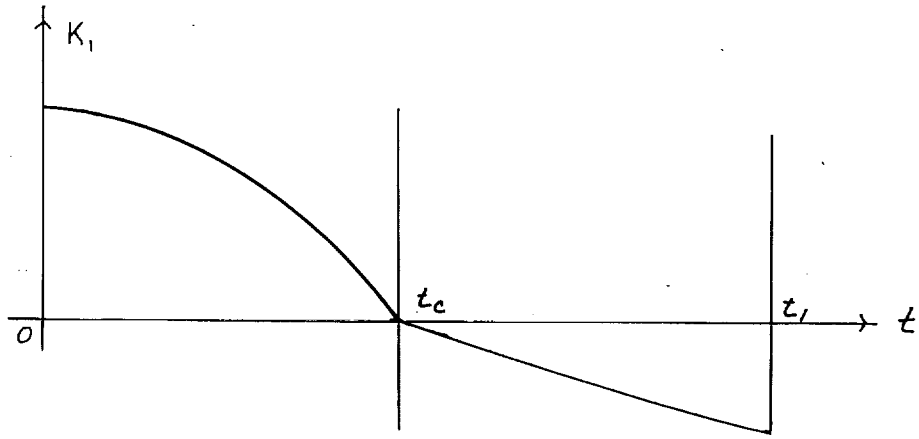


Figure 2.8.5a

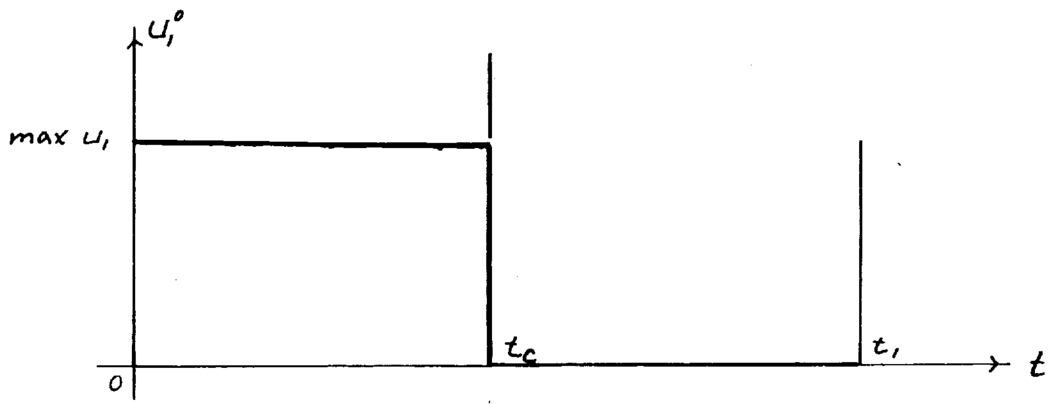


Figure 2.8.5b

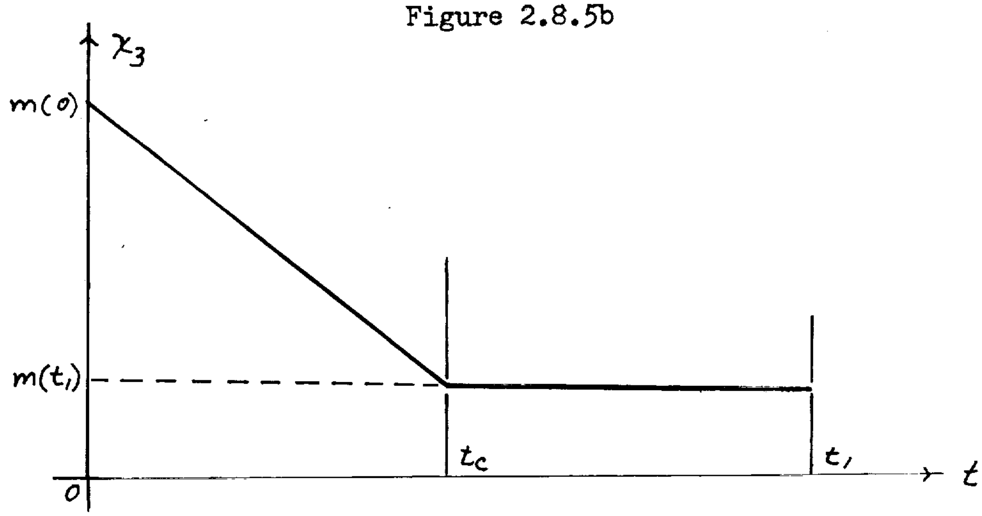


Figure 2.8.5c

Figure 2.8.5 Arcs for the Variational Problem

$$\chi_1 = \begin{cases} \chi_1(t_c) - \dfrac{g(t-t_c)^2}{2} + \chi_2(t_c)(t-t_c) & \text{for } t_c < t \le t_1 \\ c\left\{\left(\dfrac{m(o) - t \max u_1}{\max u_1}\right) \log\left(\dfrac{m(o) - t \max u_1}{m(o)}\right) + t\right\} - \dfrac{gt^2}{2} & \text{for } 0 \le t \le t_c \end{cases}$$

$$\chi_3 = \begin{cases} m(t_1), & \text{for } t_c < t \le t_1 \\ -t \max u_1 + m(o), & \text{for } 0 \le t \le t_c \end{cases}$$

$$\chi_2 = \begin{cases} -gt + \chi_2(t_c), & \text{for } t_c < t \le t_1 \\ c \log\left(\dfrac{m(o)}{m(o) - t \max u_1}\right) - gt, & \text{for } 0 \le t \le t_c \end{cases}$$

In a similar manner λ_1, λ_2 and λ_3 can be evaluated. Then, t_c and t_1 are found so that the following boundary values are satisfied:

$\chi_1(0) = 0$	$\chi_1(t_1) = ?$
$\chi_2(0) = 0$	$\chi_2(t_1) = 0$
$\chi_3(0) = m(0)$	$\chi_3(t_1) = ?$
$\lambda_1(0) = 1$	$\lambda_1(t_1) = 1$
$\lambda_2(0) = ?$	$\lambda_2(t_1) = ?$
$\lambda_3(0) = ?$	$\lambda_3(t_1) = 0$
$t_0 = 0$	$t_1 = ?$

2.9 CALCULUS OF VARIATIONS FOR THE TIME OPTIMAL CONTROL PROBLEM

2.9.1 Introduction

This chapter extends the bounded control problem of Section 2.8 to include the time optimal control problem. The time optimal situation originates in engineering problems where (1) the control variables are constrained by physical limitations and (2) the time required to reach the desired state is of paramount importance. The purpose of the control is to transform the given initial state of the system to the given final or desired state in such a way that the time required is minimized.

2.9.2 Problem Formulation

Consider the following dynamic process over the time interval from $t_0 = 0$ to t_1:

$$\dot{x}_i = g_i(t;\ x_1, \cdots x_n;\ u_1, \cdots u_r) \qquad i = 1, \ldots n$$

Let a new state variable x_{n+1} be defined where $x_{n+1}(0) = 0$. Furthermore let $x_{n+1} = t$ so that

$$\dot{x}_{n+1} = 1$$

The differential equation $\dot{x}_{n+1} = 1$ augments the above dynamic process so that the augmented state vector becomes $x = (x_1, \cdots x_{n+1})$.

There are now $n + r + 1$ dependent variables and one independent variable t. The objective is to find that special set of extremal arcs $x_1, \cdots x_{n+1}$ and the control arcs $u_1, \cdots, u_r$ from the set of admissible controls so that the criterion

$$I = x_{n+1}(t_1) = t_1$$

is minimized where

$$G(x_{n+1})\Big|_{t=t_1} = x_{n+1}(t_1)$$

and $H = 0$

and such that the boundary conditions on $x_i(0)$ and $x_i(t_1)$ for $i = 1, \cdots, n$ are satisfied. Note that minimizing I minimizes the time of system operation (t_1).

2.9.3 Time Optimal Problem

The differential equation constraints, the augmented differential equation, and the control constraining equations are given by

$$\dot{x}_i = g_i(t\ ;\ x_1, \cdots, x_n\ ;\ u_1, \cdots u_r)$$

$$\dot{x}_{n+1} = 1 \tag{2.9.1}$$

$$(u_K - \min u_K)(\max u_K - u_K) - \alpha_K^2 = 0$$

for $i = 1, \ldots, n$ and $k = 1, \cdots, r$.

Now, a set of variable Lagrange multipliers λ_j for $j = 1, \cdots n+r+1$ is introduced, and the augmented function F is then given by

$$F = \sum_{i=1}^{n} \lambda_i (\dot{x}_i - g_i) + \lambda_{n+1}(\dot{x}_{n+1} - 1)$$
$$+ \sum_{K=1}^{r} \lambda_{K+n+1}[(u_K - \min u_K)(\max u_K - u_K) - \alpha_K^2] \tag{2.9.2}$$

The dependent variables $x_1, \cdots, x_{n+1}; u_1, \cdots, u_r$ and $\alpha_1, \cdots \alpha_K$ must satisfy Eq. (2.9.1) and the Euler-Lagrange Equations over each piecewise smooth subinterval. Using Eq. (2.9.2) the Euler-Lagrange Equations can be written as

$$\dot{\lambda}_i = - \sum_{v=1}^{n} \lambda_v \frac{\partial g_v}{\partial x_i}$$

$$\dot{\lambda}_{n+1} = 0 \tag{2.9.3}$$

$$K_K(u_K) = \lambda_{K+n+1} [\max u_K + \min u_K - 2 u_K]$$

$$\lambda_{K+n+1} \alpha_K = 0$$

for $K = 1, \cdots, r$ and $i = 1, \cdots n$ where the term $K_K(u_K)$ is defined as

$$K_K(u_K) = \sum_{i=1}^{n} \lambda_i \frac{\partial g_i}{\partial u_K} \tag{2.9.4}$$

The system of differential equations which composes the constraining equations and the Euler-Lagrange equations represent 2n + 3r + 2 equations. Its solution yields n + 2r + 1 dependent variables $x_1, \cdots x_{n+1}$; $u_1, \cdots, u_r$; $\alpha_1, \cdots, \alpha_r$ and $n+r+1$ Lagrange multipliers $\lambda_1, \cdots, \lambda_{n+r+1}$.

2.9.4 Control Policy

In a manner similar to that of section 2.8, it can be shown that if $\lambda_{K+n+1} \leq 0$ for $K=1, \dots r$, a necessary condition for a weak minimum of I is satisfied. Thus, the resulting optimal control policy is identical to the u_K^* for $K=1,\dots r$ given in Eq. (2.8.10)

2.9.5 System of Differential Equations

The optimal control policy from Eq. (2.8.10) is used to reduce the number of differential equations to 2n + 2. This step is accomplished by substituting f_K for u_K . The system of differential equations now becomes

$$\dot{x}_i = g_i$$

$$\dot{x}_{n+1} = 1$$

$$\dot{\lambda}_i = -\sum_{K=1}^{n} \lambda_K \frac{\partial g_K}{\partial x_i} \tag{2.9.5}$$

$$\dot{\lambda}_{n+1} = 0$$

for $i = 1, \cdots, n$ where

$$g_i = g_i (t ; x_1 , \cdots x_n ; f_1 , \cdots f_n) \tag{2.9.6}$$

where the 2n + 2 dependent variables are $x_1, \cdots, x_{n+1}$, λ_1 , $\cdots$ λ_{n+1} and where the independent variable is t.

Using the optimal control policy and Eq. (2.9.2), the first integral, corner conditions and transversality condition can be reduced to the following:

First Integral

$$\frac{d}{dt} \left(\sum_{i=1}^{n} \lambda_i g_i + \lambda_{n+1} \right) + \frac{\partial F}{\partial t} = 0 \tag{2.9.7}$$

For problems in which $\frac{\partial F}{\partial t} = 0$, the first integral reduces to

$$\sum_{i=1}^{n} \lambda_i g_i + \lambda_{n+1} = C \tag{2.9.8}$$

where C is an integration constant.

Corner Conditions

$$(\lambda_i)_- = (\lambda_i)_+$$

$$(\lambda_{n+1})_- = (\lambda_{n+1})_+ \qquad (2.9.9)$$

and

$$\left(\sum_{i=1}^{n} \lambda_i g_i + \lambda_{n+1}\right)_- = \left(\sum_{i=1}^{n} \lambda_i g_i + \lambda_{n+1}\right)_+ \qquad i = 1, \cdots, n$$

Transversality Condition

$$(\lambda_{n+1}+1)\, dx_{n+1} + \left(-\sum_{i=1}^{n} \lambda_i g_i - \lambda_{n+1}\right) dt + \sum_{i=1}^{n} \lambda_i\, dx_i \Bigg|_{t_0}^{t_1} = 0 \qquad (2.9.10)$$

The extremal arcs $x_1, \cdots, x_{n+1}$ and the Lagrange multipliers $\lambda_1, \cdots, \lambda_{n+1}$ can be found by solving the system of differential Eq. (2.9.5) and using Eq. (2.9.7), (2.9.9) and (2.9.10). Knowing $x_1, \cdots x_{n+1}$; $\lambda_1, \cdots \lambda_{n+1}$, the optimal control variables $u_1^o, \cdots, u_r^o$ can be obtained from the optimal control policy described in Eq. (2.8.10).

2.9.6 Example

Consider the control of the dynamic process shown in Figure 2.7.1 that is characterized by the differential equation

$$\dot{x}_1 = -2\sqrt{2}\, x_1 + u_1$$

Let the process by initially in the state $x_1(0) = 2$. The objective is to determine the control u_1 on the time interval from $t_0 = 0$ to t_1 such that: (1) the final state $x_1(t_1) = 1$, and (2) the final or terminal time t_1 has a minimum subject to the constraint $-1 \le u_1 \le 1$. The criterion is $I = x_2(t_1)$ where the $x_2(t)$ represents the new state variable $x_2(t) = t$. The augmented state vector x is now given by

$$x = (x_1, x_2)$$

and the process is described by the following differential equation constraints and the control constraining equation:

$$\dot{x}_1 = -2\sqrt{2}\, x_1 + u_1$$

$$\dot{x}_2 = 1$$

$$(u_1+1)(1-u_1) - \alpha_1^2 = 0$$

Thus,

$$x = (x_1, x_2) \quad \text{with} \quad n = 2 \; ; \quad u = (u_1) \quad \text{with} \quad r = 1$$

$$H = 0 \quad ; \quad G = X_2(t) \qquad \text{from} \quad \text{Eq} \, (2.7.4)$$

$$g_1 = -2\sqrt{2}\, x_1 + u_1 \qquad \text{from} \quad \text{Eq} \, (2.7.3)$$

$$g_2 = 1$$

Substituting these equations into Eq. (2.9.4) yields

$$K_1(u_1) = K_1 = \lambda_1$$

where K_1 is not a function of u_1 .

The optimal control policy from Eq. (2.8.10a) is, then, given by

$$u_1^o = \left\{ \begin{array}{lll} 1 & \text{when} & K_1 = \lambda_1 > 0 \\ -1 & \text{when} & K_1 = \lambda_1 < 0 \\ w_1 & \text{when} & K_1 = \lambda_1 = 0 \end{array} \right\} = \operatorname{sgn} \lambda_1$$

where

$$-1 \leq w_1 \leq 1$$

Figure (2.9.1) shows the optimal control policy as a function of λ_1 .

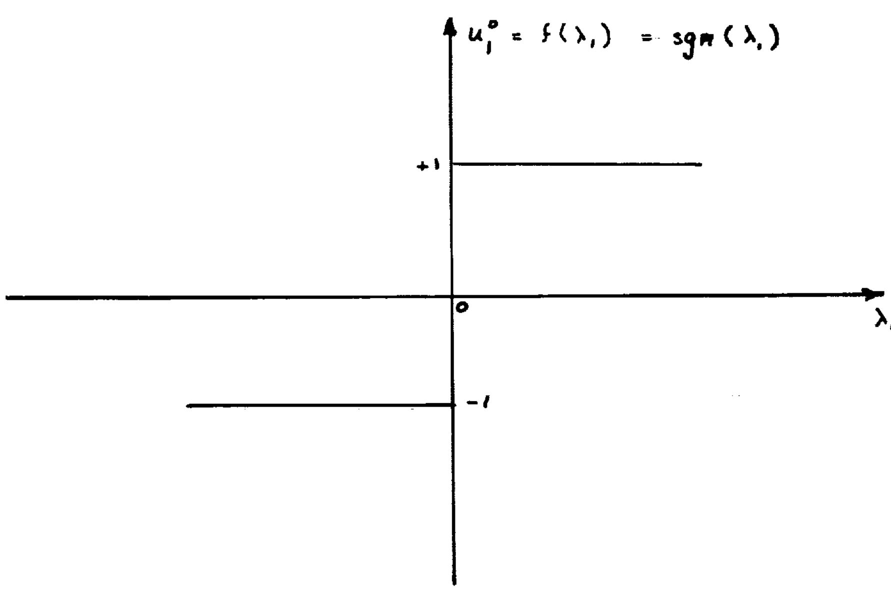


Figure 2.9.1 Relay: Time Optimal Control

The system of differential equations (2.9.5) then becomes

$$\dot{x}_1 = -2\sqrt{2}\ x_1 + \operatorname{sgn}\lambda_1$$

$$\dot{x}_2 = 1$$

$$\dot{\lambda}_1 = 2\sqrt{2}\ \lambda_1$$

$$\dot{\lambda}_2 = 0$$

where $u_1^\circ = \operatorname{sgn}(\lambda_1)$. The first integral from Eq. (2.9.8) is given by

$$C = \lambda_1(-2\sqrt{2}\ x_1 + \operatorname{sgn}\lambda_1\ ) + \lambda_2$$

where C is an integration constant. Now applying the transversality condition [Eq. (2.9.10)]

$$\left[(\lambda_2+1)\, dx_2 - C\, dt + \lambda_1\, dx_1 \right] \Big|_{t_0}^{t_1} = 0$$

indicates that $\lambda_2(t_1) = -1$ since $dx_2(t_1)$ is arbitrary. In addition, $C(t_1) = 0$ since dt_1 is arbitrary. But, the $\lambda_2(t)$ and $c(t)$ are constants; thus $\lambda_2 = -1$ and $C = 0$ for all t. The first integral thus becomes

$$0 = \lambda_1\left[-2\sqrt{2}\ x_1 + \operatorname{sgn}\lambda_1 \ \right] - 1$$

Now at $t = t_1$ with $x_1(t_1) = 1$,

$$0 = \lambda_1\left[-2\sqrt{2}\ x_1(t_1) + \operatorname{sgn}\lambda(t_1) \ \right] - 1$$

or

$$\lambda_1(t_1) = -.361$$

Since $\dot{\lambda}_1 = 2\sqrt{2}\lambda_1$ and $\lambda_1(t_1) = -.361$, then $\lambda_1(t) < 0$. Further,

$$u_1^\circ(t) = \operatorname{sgn}\lambda_1 = -1 \qquad (t_0 \le t \le t_1)$$

Finally, using the boundary conditions

$x_1(0) = 2$	$x_1(t_1) = 1$
$x_2(0) = 0$	$x_2(t_1) = ?$
$\lambda_1(0) = ?$	$\lambda_1(t_1) = -.361$
$\lambda_2(0) = -1$	$\lambda_2(t_1) = -1$
$t_0 = 0$	$t_1 = ?$

the system of differential equations yields

$$x_1(t) = -.354 + 2.354\, e^{-2.828\, t}$$

$$u_1(t) = -1$$

$$t_1 = .196$$

The $x_1(t_1)$, $\lambda_1(t)$ and $u_1^{\circ}(t)$ are shown in Figure 2.9.2 . An examination of the second variation of I indicates that the extremal arc x_1 is a weak extremum for I.

The differentiability constraints are satisfied on the interval from t_0 to t_1; that is, x_1 , x_2 and u_1 are smooth, etc.

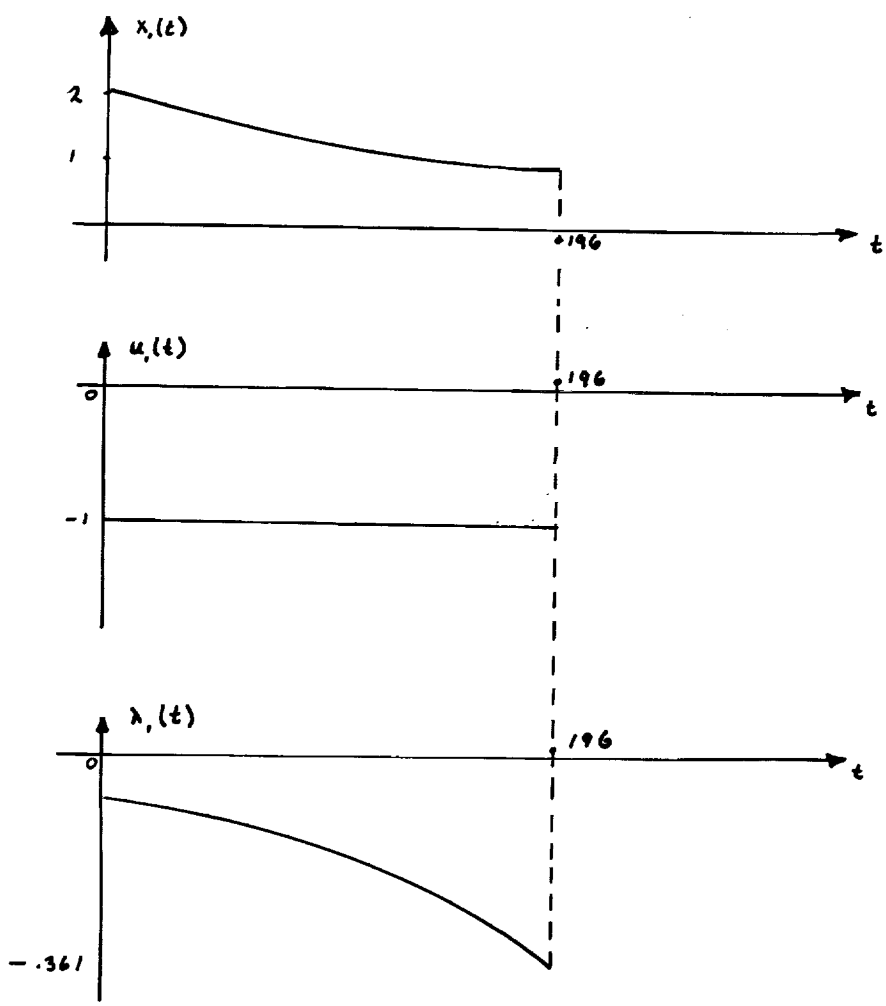


Figure 2.9.2: Solutions to System of Differential Equations

3.0 RECOMMENDED PROCEDURES

This monograph treats the optimization of deterministic systems and includes inequality constraints on the control variables. The formulations are, thus, primarily applicable to problems similar to those encountered in the trajectory and control analysis.

The procedures developed and illustrated in the previous sections of this monograph concern themselves with the problem of extremizing the functional

$$I(Y) = \int_{x_1}^{x_2} f(x, Y, Y')\, dx$$

where f is a continuous function, Y is a real piecewise smooth function of x defined over the region G which contains all Y of interest, (which satisfies the boundary conditions imposed on the problem) and where Y' is

$$Y' \equiv \frac{d}{dx}(Y(x)) .$$

Attention then turns to extending the analysis to include more general representations of f (e.g., f $(x, Y, Y', Y'' \ldots)$), the inclusion of variable boundary conditions, etc. In the process, several necessary conditions are derived (Euler-Lagrange, Legendre, Weierstrass, Jacobi) as are the transversality conditions, the corner conditions and the first integral. Sufficient conditions are also derived.

These formulations must be applied to problems as illustrated in the numerous samples presented in the previous sections of the monograph.

However, the difficulties in optimization theory lie primarily in the area of problem solution rather than problem formulation since the application of the variational calculus to an optimization problem in general leads to a set of equations of the two point boundary value type. These equations are readily solvable only if they are linear. If they are non-linear, then a solution must be effected iteratively through the use of numerical computation techniques. At present there are three iterative techniques based on linear theory available

(i) - Neighboring Extremal
(ii) - Quasilinearization
(iii) - Gradient or Steepest Ascent Method

Since these schemes are linear, the iterative process must begin "fairly close" to the actual solution of the problem if convergence is to be realized. Just how "close" depends, of course, on the particular problem and the particular set of boundary conditions. Since these techniques and other purely numerical techniques are the subject of a future monograph, the material of this monograph, while adequate to formulate most deterministic problems, is not complete unto itself for solving these problems. Thus, recommendations regarding this latter class of problems and means of obtaining initial "guesses" will be deferred.

4.0 REFERENCES

The developments presented in the text (Section 2.0) discuss the classical approach to optimization problems and extend this approach to modern problems of control. These discussions are believed to constitute a thorough treatment of the subject matter. However, since the text was intended primarily as a summary and since many versions of variational principles have been developed and applied in the literature, a short list of references will be presented. These references in conjunction with other monographs of this series provide an extremely thorough appreciation of the difficulties encountered in solving general optimization problems.

Akhiezer, N. I. The Calculus of Variations, translated by A. H. Frink, Blaisdell Publishing Co., New York (1962).

Blakrishman, A. V. and L. W. Neustadt. Computing Methods in Optimization Problems, Chapter 1: "Variational Theory and Optimal Control Theory," by M. R. Hestenes, Academic Press, New York (1964).

Becker, M. The Principles and Applications of Variational Methods, MIT Press, Cambridge, Massachusetts (1964).

Bliss, G. A. Lectures on the Calculus of Variations, University of Chicago Press, Chicago, Illinois, (1946) p. 296.

Bliss, G. A. "The Problem of Lagrange in the Calculus of Variations," American Journal of Mathematics, 52, p. 673-744 (1930).

Bliss, G. A. "Sufficient Conditions for a Minimum with Respect to One-Sided Variations," Trans. Am. Math. Soc. 5, p. 477-492 (1904).

Bolza, O. Lectures on the Calculus of Variations, Dover Publications, Inc., New York (1961).

Breakwell, J. V., et. al. "Optimization and Control of Nonlinear Systems Using The Second Variation," SIAM Journal Series A: Control 1, p. 193 (1963).

Breakwell, J. V. "The Optimization of Trajectories," Journal of Mathematical Analysis & Applications, 7, p. 215-247 (1959).

Bryson, A. E., et. al. "Optimal Programming Problems with Inequality Constraints I: Necessary Conditions for Extremal Solutions," AIAA Journal, number 1, p. 2544-2550 (1963).

Denham, W. F. and Bryson, A. E. "Optimal Programming Problems with Inequality Constraints II: Solution by Steepest-Ascent," AIAA Journal, number 2, p. 25-34 (1964).

Edelbaum, T. N. "Theory of Maxima and Minima," Optimization Techniques, Edited by G. Leitmann, Academic Press, Inc., New York, Chapter I, (1962).

Elsgole, L. E. Calculus of Variation, Addison-Wesley Publishing Company, Inc. Reading, Massachusetts (1962).

Garfinkel, B. "A Solution of the Goddard Problem," SIAM Journal, Series A: Control 2, p. 135 (1964).

Gelfand, I. M., and S. V. Fomin. Calculus of Variations (translated from Russian), Prentice-Hall, Inc., New Jersey (1963).

Hestenes, M. R. "Numerical Methods of Obtaining Solutions of Fixed Endpoint Problems in the Calculus of Variations," Rand Report RM-102 (August 1949).

Hestenes, M. R. "Variational Theory and Optimal Control Theory," Computing Methods in Optimization Problems, edited by A. V. Balakrishnan and L. W. Neustadt, Academic Press, Inc., New York (1964).

Hestenes, M. R. "A General Problem in the Calculus of Variations with Applications to Paths of Least Time," Rand Rm-100, Rand Corp., Santa Monica, California (March 1949).

Jurovics, S. and J. E. McIntyre. "The Adjoint Method and its Application to Trajectory Optimization," ARS Journal 32, p. 1354 (1962).

Kalman, R. E. "The Theory of Optimal Control and the Calculus of Variations," Research Institute for Advanced Studies, Baltimore, Maryland, TR 61-3 (1961).

Kalman, R. E. and R. W. Koepche. "Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indices," Trans. Am. Soc. Mech. Engrs., p. 1120 (November 1958).

Kelley, H. J. "An Investigation of Optimal Zoom-Climb Techniques," Journal of Aerospace Sciences, 26, pp. 794-802, 824 (1959).

Kelley, H. J. "Gradient Theory of Optimal Flight Paths," ARS Journal, 30, p. 947-954 (1960).

Kelley, H. "Optimization Techniques," Method of Gradients, edited by G. Leitman, Academic Press, Inc., New York (1962), Chapter 6.

Kelley, J. J. "Guidance Theory and Extremal Fields," Inst. Radio Engrs., Trans. Auto. Control AC-7 (October 1962).

Kipiniak, W. Dynamic Optimization and Control: A Variational Approach, Technical Press, Massachusetts Institute of Technology, Cambridge, Mass. and John Wiley and Sons, Inc., New York (1961).

Krasovskii, N. N. Sufficient Conditions for Optimization," Journal Applied Mathematics and Mechanics, 23, 839, p. 43 (1959).

McIntyre, J. E. and L. Crocco. "Optimal Transfer Between Close Circular Orbits," AIAA Paper 64-29 (January 1964).

McIntyre, J. E. and B. J. Paiewonsky. "On Optimal Control with Bounded State Variables," Aeronautical Research Associates of Princeton Inc. Report 60, (July 1964).

Merse, M. "The Calculus of Variations in the Large," American Mathematical Society Colloquim, Publication Vol. 18, American Mathematical Society, New York (1934).

Miele, A. Flight Mechanics: Theory of Flight Paths, Addison-Wesley Publishing Co., Inc., Reading, Massachusetts (1962).

Miele, A. "Generalized Variational Approach to the Optimum Thrust Programming for the Vertical Flight of a Rocket," Purdue University, School of Aerospace Engineering A-57-1 (March 1957).

Pars, L. A. Calculus of Variations, John Wiley and Sons, Inc., New York (1963).

Ross, S. "Minimality for Problems in Vertical and Horizontal Rocket Flight," Jet Propulsion, 28, p. 55-56 (1958).

Schwartz, L. "Optimation Techniques - A Comparison," U. S. Air Force Flight Dynamics Lab. TDR-64-21, Wright - Patterson Air Force Base, Ohio (March 1964).

Thompkins, C. B. "Methods of Steep Descent, Modern Mathematics for the Engineer, edited by E. F. Beckenbach, McGraw-Hill Bood Co., Inc., New York, (1956), Chapter 18.

Valentine, F. A. "The Problem of Lagrange with Differential Inequalities as Added Side Conditions," Contributions to the Calculus of Variations, 1933-1937, University of Chicago Press, Chicago, Illinois (1937).

Warga, J. "Relaxed Variational Problems," Journal of Mathematical Analysis, 4, p. 111-128 (1962).

Weinstock, R. Calculus of Variations with Applications to Physics and Engineering, McGraw-Hill Book Co., Inc., New York (1952).

NASA-Langley, 1968 —— 30 CR-1003