NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Technical Report 32-1043

The Conjugate Point and Dynamic Programming

W. G. Melbourne

GPO PRICE $ ____

CFSTI PRICE(S) $ ____

Hard copy (HC) 3.00

Microfiche (MF) .65

ff 653 July 65

N67 17825
(ACCESSION NUMBER) (THRU) 1
17 (PAGES) (CODE)
CR-81638 (NASA CR OR TMX OR AD NUMBER) 19 (CATEGORY)
FACILITY FORM 602

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

September 15, 1966

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Technical Report 32-1043

The Conjugate Point and Dynamic Programming

W. G. Melbourne

Approved by:

T. W. Hamilton

T. W. Hamilton, Manager
Systems Analysis Section

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA
September 15, 1966

JPL TECHNICAL REPORT 32-1043

Copyright © 1967
Jet Propulsion Laboratory
California Institute of Technology

Prepared Under Contract No. NAS 7-100
National Aeronautics & Space Administration

Contents

Abstract

The continuity properties of the second partials of the optimal performance function $S(x,y)$ play a crucial role in making the transition from the recursive formulation of the principle of optimality of dynamic programming to the Hamilton-Jacobi equation and to the Euler and Weierstrass necessary conditions of the calculus of variations. At a singularity point of an extremal, S_{yy} has a first-order pole. The asymptotic properties of S_{yy} near a pole are discussed and used to demonstrate the necessity of the boundedness of S_{yy} on the interior of a minimizing extremal. A one-to-one relationship is demonstrated between a pole of S_{yy} on the interior of an extremal and the conjugate point as classically defined in the calculus of variations. The geometric properties of envelopes and the envelope theorem are reviewed. Finally, it is shown that when the end point of an extremal is sufficiently close to an envelope, such an extremal is not a globally minimizing curve.

The Conjugate Point and Dynamic Programming

I. Introduction

It has been shown by Dreyfus (Ref. 1) that the principle of optimality of dynamic programming leads to the various minimizing necessary conditions of the calculus of variations, with the exception of the Jacobi condition concerning conjugate points on extremals. The conjugate-point phenomenon in the dynamic-programming approach is reflected in the second partial derivatives of the optimal performance function $S(x,y)$ becoming unbounded at a point. Dreyfus has observed that although the Jacobi condition that there be no conjugate points on an extremal between the end points is a demonstratable necessary condition when using the calculus of variations, it has not yet been shown that with the sole use of a dynamic-programming approach the existence of a point of singularity of the second partials of $S(x,y)$ on the interior of an extremal necessarily implies a nonminimizing character for such an extremal.

It is the purpose of this Report to indicate the necessity of this dynamic-programming condition and to describe related properties of the conjugate-point phenomenon. For the sake of completeness, this Report contains considerable material of a tutorial nature. Only the simplest problem of the calculus of variations will be considered. Hence, the problem is the minimization of a functional in the form

$$\min_{\{y(x)\}} \left[\int_{x_1}^{x_2} f(x,y,y')\, dx \right] \tag{1}$$

with fixed end points. It will be assumed that f has at least piece-wise continuous third partial derivatives and that the resulting extremals for this problem are regular (i.e., $f_{y'y'} \neq 0$ for all x).

II. Dynamic Programming and the Necessary Conditions of the Calculus of Variations[1]

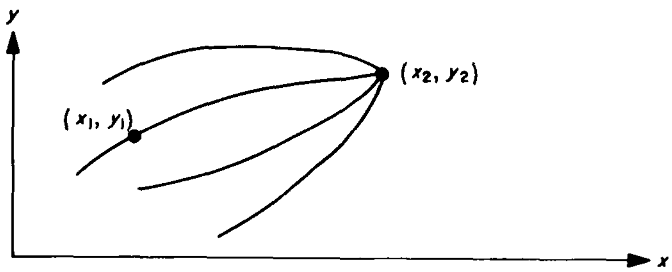
A brief and heuristic exposition of the dynamic-programming approach to the calculus of variations is provided in this Section. We define $S(x_1,y_1)$ to be the minimum value of expression (1), where we consider the final point (x_2,y_2) to be fixed and we let (x_1,y_1) be free to

[1] See Ref. 1, Chap. 3, for a more detailed treatment of the material in this Section.

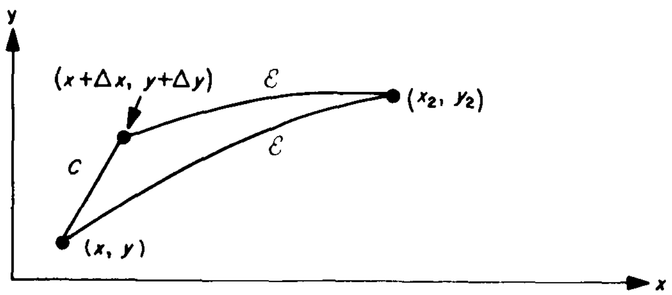
change. With this approach we have imbedded our original problem of finding the minimizing path from (x_1,y_1) to (x_2,y_2) in the more general family of problems of finding the minimizing curve from an arbitrary point (x,y) to the fixed final point (x_2,y_2) (Sketch 1). We hypothesize the existence of a minimizing path from (x,y) to (x_2,y_2) and we seek to characterize the minimizing properties of such a path. For each point (x,y) there is a unique value S associated with this point and it is well known that $S(x,y)$ satisfies the Hamilton Jacobi equation. The principle of optimality implies the following relation for all points from which a minimizing curve to (x_2,y_2) exists:

$$S(x, y) = \min_{\{y(x)\}} \left[\int_x^{x+\Delta x} f(x, y, y')\, dx + S(x + \Delta x, y + \int_x^{x+\Delta x} y'\, dx) \right] \tag{2}$$

where the minimization is with respect to the $y(x)$ over the interval $(x,\ x + \Delta x)$ (Sketch 2). By discretizing this expression, we are lead directly to the fundamental recursive relation which is the keystone of the dynamic-programming computational philosophy.



Sketch 1



Sketch 2

Since we here are interested in the relationship to the calculus of variations, we let Δx become small and expand the right-hand side of Eq. (2) in a Taylor series about (x,y). Although there are some fine points in this procedure, it has nevertheless been put on a rigorous basis by Berkovitz and Dreyfus (Ref. 2). We treat the curve C as a straight line connecting the two infinitesimally close points (x, y) and $(x + \Delta x, y + \Delta y)$. Hence, the minimization in Eq. (2) with respect to $y(x)$ now becomes a problem of finding the optimal slope at (x, y). That is, Eq. (2) may be approximated by

$$S(x, y) = \min_{\{y'\}} [hf(x, y, y') + S(x + h, y + y'h)] \tag{3}$$

where h has been written for Δx. Assuming that the second partial derivatives of $S(x, y)$ exist, we may expand through the first order to obtain

$$0 = \min_{\{y'\}} [hf(x, y, y') + hS_x + hS_y y' + 0(h^2)] \tag{4}$$

and dividing by h and letting $h \rightarrow 0$, we obtain

$$\min_{\{y'\}} [f(x, y, y') + S_x + S_y y'] = 0 \tag{5}$$

The minimizing value of y' is seen from Eq. (5) to be dependent on (x,y) and S_y. But S_y is implicitly a function of (x,y), and so we define the slope function $p(x,y)$ to be the minimizing value of y'. Because there are no constraints on y' in our problem, we may differentiate Eq. (5) to obtain a necessary condition for a minimum of Eq. (5) with respect to y'. Hence we obtain two results:

$$0 = f(x,y,p) + S_x + S_y\, p = 0 \tag{6a}$$

$$f_{y'}(x,y,p) + S_y = 0 \tag{6b}$$

Eliminating p from Eq. (6a) and (6b) yields the Hamilton-Jacobi equation for $S(x,y)$. If we assume that there is a unique minimizing extremal from (x,y) to (x_2,y_2), then except at corners of such extremals, $p(x,y)$ will be a single-valued function indicating the slope of the extremal passing through the point (x,y).

From Eq. (6), which is the fundamental result, it is a simple matter to derive the Euler equation, and the Legendre and Weierstrass necessary conditions of the calculus of variations. The Euler equation, which may be thought of as a first-order partial differential equation in

$p(x,y)$ describing its behavior *along* an extremal, is obtained as follows: One differentiates Eq. (6a) with respect to y to obtain

$$f_y+f_{y'}p_y+S_{xy}+S_{yy}p+S_y p_y = 0 \qquad (7)$$

and using Eq. (6b) we have

$$f_y+S_{xy}+S_{yy}p = 0 \qquad (8)$$

Since we have assumed that the second partials of S exist, we may interchange the differentiation order of S in Eq. (8). We denote the total derivative of a quantity taken along an extremal by $(d/dx)_p$ and thus we have

$$\left(\frac{dS_y}{dx}\right)_p = S_{yx} + S_{yy}\left(\frac{dy}{dx}\right)_p = S_{yx} + S_{yy}\,p \qquad (9)$$

Then from Eq. (6b), (8), and (9) we have the result

$$\left(\frac{d}{dx}f_{y'}(x,y,p)\right)_p - f_y(x,y,p) = 0 \qquad (10)$$

which is the Euler equation defining the change in $p(x,y)$ along a particular extremal.

The Legendre condition follows from noting that $p(x,y)$ minimizes Eq. (6a) and hence the second partial of Eq. (6a) with respect to p must be nonnegative. It follows that $f_{y'y'}(x,y,p)\geq 0$. The Weierstrass condition is deduced from the fact that zero is a minimum value for $f(x,y,y')+S_x+S_y y'$ which is attained for $y' = p$. Then, by replacing S_y with $-f_{y'}$ in Eq. (6a), we have the inequality

$$f(x,y,y') - f(x,y,p) - (y'-p)f_{y'}(x,y,p)\geq 0 \qquad (11)$$

which is the Weierstrass condition.

III. The Conjugate Point

The derivation of Eq. (6a) and (6b) required two assumptions; namely, that the second partials of $S(x,y)$ were bounded at (x,y) and that a minimizing extremal from (x,y) to (x_2,y_2) existed. Suppose now that as one travels backward along an extremal from (x_2,y_2) one reaches finally a point (x_0,y_0) where the second partials of S become infinite. It is not guaranteed that as one approaches (x_0,y_0) from (x_2,y_2) along an extremal $\mathcal{E}_0$, $\mathcal{E}_0$ will remain a globally minimizing extremal. Therefore, we must be somewhat careful in phrasing the minimization operation in Eq. (3) for (x,y) near a conjugate point. We will relax our definition of $S(x,y)$ to be simply the value of the integral along an extremal that is assumed to be globally minimizing for (x,y) near (x_2,y_2) and is at least relatively minimizing for (x,y) near a conjugate point.

Let us consider here the asymptotic behavior along $\mathcal{E}_0$ of the partial derivatives S_{xx}, S_{xy}, and S_{yy} near (x_0,y_0); these will be used in later considerations. On $\mathcal{E}_0$, for $x>x_0$, the expressions in Eq. (6a) and (6b) are valid and we can differentiate them because of the continuity properties assumed for $f(x,y,y')$. Hence we obtain from various differentiations

$$\left.\begin{aligned} f_y + S_{xy} + S_{yy}p &= 0 \\ f_x + S_{xx} + S_{yx}p &= 0 \\ f_{y'y} + f_{y'y'}p_y + S_{yy} &= 0 \\ f_{y'x} + f_{y'y'}p_x + S_{yx} &= 0 \end{aligned}\right\} \qquad (12)$$

Since f and its derivatives are bounded, we see that near (x_0,y_0) the ratios S_{xx}/S_{yy} and S_{xy}/S_{yy} are defined and, in fact,

$$\left.\begin{aligned} \lim_{(x,y)\to(x_0,y_0)}\left(\frac{S_{xx}}{S_{yy}}\right) &= p^2(x_0,y_0) \\ \lim_{(x,y)\to(x_0,y_0)}\left(\frac{S_{xy}}{S_{yy}}\right) &= -\,p(x_0,y_0) \end{aligned}\right\} \qquad (13)$$

From Eq. (13) it follows that we need only study the asymptotic behavior of S_{yy}. For since $p(x_0,y_0)$ is bounded in the simplest problem, if any of the second partials of $S(x,y)$ become unbounded, S_{yy} will be included. We need an expression describing the behavior of S_{yy} along $\mathcal{E}_0$. We note from Eq. (6b) and our continuity assumptions for $f(x,y,p)$ that S_y has at least piece-wise continuous second derivatives at all points on $\mathcal{E}_0$ for which the derivation of Eq. (6b) is valid. From this it follows that $[d(S_{yy})/dx]_p$ exists on $\mathcal{E}_0$ at all such points, which enables us to obtain a first-order differential equation describing the behavior of S_{yy} along $\mathcal{E}_0$. Differentiating the first expression in Eq. (12) with respect to y, we obtain

$$f_{yy} + f_{yy'}p_y + S_{xyy} + S_{yyy}p + S_{yy}p_y = 0 \qquad (14)$$

Along $\mathcal{E}_0$ we also have

$$\left(\frac{dS_{yy}}{dx}\right)_p = S_{yyx} + S_{yyy}p \qquad (15)$$

and it follows that

$$\left(\frac{dS_{yy}}{dx}\right)_p = -f_{yy} - (f_{yy'} + S_{yy})p_y \tag{16}$$

Using the third expression in Eq. (12) to eliminate p_y, we obtain finally

$$\left(\frac{dS_{yy}}{dx}\right)_p = -f_{yy} + (f_{yy'} + S_{yy})(f_{y'y'})^{-1}(f_{y'y} + S_{yy}) \tag{17}$$

Hence, S_{yy} satisfies a Ricatti equation along $\mathcal{E}_0$.

At (x_0, y_0) on $\mathcal{E}_0$, S_{yy} becomes unbounded. Equation (17) indicates that near (x_0,y_0)

$$\left(\frac{dS_{yy}}{dx}\right)_p \to S_{yy}(f_{y'y'})^{-1}S_{yy} \tag{18}$$

From expression (18) it follows that the asymptotic behavior of S_{yy} on $\mathcal{E}_0$ near (x_0,y_0) is given by

$$(x - x_0)S_{yy} = -f_{y'y'}(x_0,y_0,p_0) + 0[(x-x_0)] \tag{19}$$

That is, S_{yy} has a first-order pole at x_0. Equation (19) will be used later to demonstrate the necessity of the boundedness of the second partials of $S(x,y)$ along $\mathcal{E}_0$.

Let us next show that a singularity point of S_{yy} does in fact correspond to the conjugate point as usually defined in the calculus of variations. From Eq. (12) it follows that at a singularity point of S_{yy}, $p_y(x,y)$ also becomes unbounded and that the ratio S_{yy}/p_y is defined and given by

$$\lim_{(x,y)\to(x_0,y_0)} \left[\frac{S_{yy}}{p_y}\right] = -f_{y'y'}(x_0,y_0,p_0) \tag{20}$$

We recall that p is the value of the slope of the extremal through (x,y). Let us say that the extremal through (x,y) to (x_2,y_2) is described by the function

$$y = \phi(x,a) \tag{21}$$

where a is the parameter that generates a one-parameter family of extremals passing through (x_2,y_2) and for the particular value a_0 generates the extremal $\mathcal{E}_0$. It may be shown from Eq. (10) that except at corners, ϕ has continuous second partial derivatives in x and a. If we assume that $\phi_a \neq 0$, then Eq. (21) defines a unique a for each (x,y) and we may write $a=a(x,y)$. It then follows that the slope function $p(x,y)$ is defined by the relation

$$p(x,y) = \phi'(x,a(x,y)) \tag{22}$$

and $p_y(x,y)$ on $\mathcal{E}_0$ is given by

$$p_y(x,y) = \lim_{\delta a\to 0}\left[\frac{\phi'(x,a_0+\delta a) - \phi'(x,a_0)}{\phi(x,a_0+\delta a) - \phi(x,a_0)}\right] = \frac{\phi'_a(x,a_0)}{\phi_a(x,a_0)} \tag{23}$$

Now $\phi(x,a)$ satisfies the Euler equation in Eq. (10). Let us substitute $\phi(x,a)$ into Eq. (10) and differentiate with respect to a and interchange differentiations, which is permissible since the appropriate partial derivatives exist. We obtain

$$\frac{d}{dx}(f_{y'y'}\phi_a) - \left(f_{yy} - \frac{d}{dx}f_{y'y'}\right)\phi_a = 0 \tag{24}$$

which is recognized as the Jacobi equation. In a sense, this equation provides the link between a dynamic-programming formulation of conjugate-point phenomena and a calculus-of-variations formulation where the zero crossings of ϕ_a are the object of study. A zero crossing of a nontrivial solution of the Jacobi equation in Eq. (24) satisfying the boundary condition $\phi_a(x_2,a_0) = 0$ is defined as a conjugate point, or more precisely a point conjugate to the terminal point (x_2,y_2). From Eq. (23) and (24) it follows that $p_y(x,y)$ becomes unbounded when and only when $\phi_a(x,a_0)$, a nontrivial solution of Eq. (24), is zero. Hence, it follows from Eq. (20) that S_{yy} is always unbounded at a conjugate point and further, except possibly at the point (x_2,y_2), S_{yy} is bounded at all nonconjugate points of $\mathcal{E}_0$. Hence, the necessity of the boundedness of S_{yy} for $x_1 < x < x_2$ in order that $\mathcal{E}_0$ be a minimizing extremal can be demonstrated from the Jacobi necessary condition in the calculus of variations.

IV. A Dynamic-Programming Proof of the Jacobi Necessary Condition

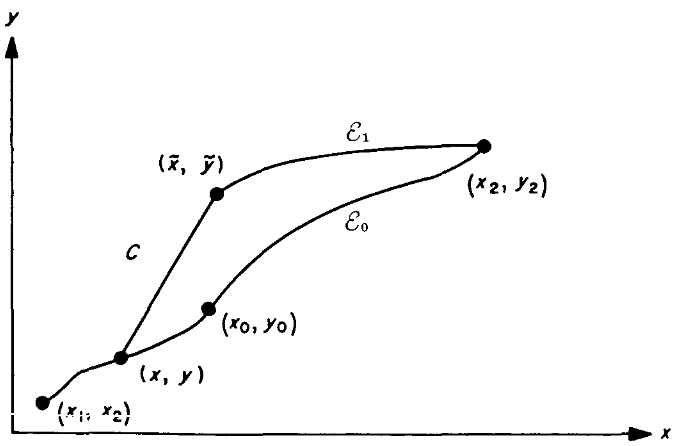
To demonstrate the necessity of the Jacobi condition for a conjugate point (x_0,y_0) lying on the interior of $\mathcal{E}_0$ (i.e., $x_1 < x_0 < x_2$), we compare the value $S(x,y)$ associated with $\mathcal{E}_0$ with the value V obtained from following a path neighboring $\mathcal{E}_0$ from (x,y) to (x_2,y_2). We wish to show that the difference $V - S(x,y)$ can be made negative for some neighboring path when (x_0,y_0) is an interior point of $\mathcal{E}_0$. Such a neighboring path is indicated in Sketch 3, where an arbitrary curve C is followed from (x,y) to $(\tilde{x},\tilde{y})$ and an extremizing curve $\mathcal{E}_1$ is then followed from $(\tilde{x},\tilde{y})$ to (x_2,y_2). Hence, V is given by

$$V = S(\tilde{x},\tilde{y}) + \int_{\substack{x \\ C}}^{\tilde{x}} f(x,y,y')dx \tag{25}$$

The value $S(x,y)$ can be written as

$$S(x,y) = S(x_0,y_0) + \int_{\substack{x \\ \mathcal{E}_0}}^{x_0} f(x,y,p)dx \tag{26}$$

where we must be careful to define $S(x_0,y_0)$ to be the limiting value of $S(x,y)$ as (x,y) approaches (x_0,y_0) along $\mathcal{E}_0$ from the (x_2,y_2) side. By expanding the right-hand members of Eq. (25) and (26) in Taylor series about the point $[\tilde{x}, \phi(\tilde{x},a_0)]$ on $\mathcal{E}_0$, one can show, using the asymptotic properties of S_{yy} at (x_0,y_0), that the difference $V - S(x,y)$ can be made negative. This approach is cumbersome and requires care in expanding the S functions.



Sketch 3

An alternate approach, which we follow here, is to show that the second variation of the functional in expression (1) can be made negative. Since $\mathcal{E}_0$ is an extremal, the first variation about $\mathcal{E}_0$ is zero; if we can make the second variation negative, this implies that a neighboring curve arbitrarily close to $\mathcal{E}_0$ in both ordinate and slope can be found for which $V - S(x,y)$ is negative. Hence, we can write

$$V - S(x,y) = \delta^2 I + \text{higher order terms} \tag{27}$$

where $\delta^2 I$ is the second variation of the functional in expression (1) and is given by

$$\delta^2 I = \frac{1}{2}\int_{x_1}^{x_2} [\eta^2 f_{yy} + 2\eta\eta' f_{yy'} + \eta'^2 f_{y'y'}]\, dx \tag{28}$$

Here, η is the difference in ordinate between the neighboring curve and $\mathcal{E}_0$, and η' is the difference in slopes. From Sketch 3 we observe that η and η' are zero over the interval $[x_1, x]$. The functions f_{yy}, $f_{yy'}$ and $f_{y'y'}$ are evaluated on $\mathcal{E}_0$. By making both η and η' arbitrarily small, we can make the second variation term in Eq. (27) dominate the right-hand side and thus control the sign of $V - S(x,y)$. The second variation in Eq. (28) is composed of two parts; namely, a contribution from traveling along C from (x,y) to $(\tilde{x},\tilde{y})$ and a contribution from $\mathcal{E}_1$ between $(\tilde{x},\tilde{y})$ and (x_2,y_2). The extremal $\mathcal{E}_1$ is described by the function $\phi(x,a_1)$. Let us define J to be the component of $\delta^2 I$ that results from following $\mathcal{E}_1$ between $(\tilde{x},\tilde{y})$ *and* $(x_2 y_2)$. Then J can be written

$$J(\tilde{x},\tilde{\eta}) = \frac{1}{2}\int_{\tilde{x}}^{x_2} [\eta^2 f_{yy} + 2\eta\eta' f_{yy} + \eta'^2 f_{y'y'}]\, dx \tag{29}$$

where $\tilde{\eta}$ is $\tilde{y} - \phi(\tilde{x}, a_0)$ and $\eta(x)$ is given by

$$\eta(x) = \phi(x,a_1) - \phi(x,a_0) = \phi_a(x,a_0)(a_1 - a_0) + 0(a_1 - a_0)^2 \tag{30}$$

We shall next prove that

$$J(\tilde{x},\tilde{y}) = \frac{1}{2}(\tilde{\eta}^2 S_{yy})\,|_{\tilde{x}} + 0(\tilde{\eta}^3) \tag{31}$$

where S_{yy} is evaluated on $\mathcal{E}_0$. The identity

$$\frac{d}{dx}(\eta^2 S_{yy})_p = 2\eta\eta' S_{yy} + \eta^2\left(\frac{dS_{yy}}{dx}\right)_p$$

becomes, with the use of Eq. (12) and (16),

$$\frac{d}{dx}(\eta^2 S_{yy})_p = -\eta^2 f_{yy} - 2\eta\eta' f_{yy'} - (2\eta\eta' p_y - \eta^2 p_y^2) f_{y'y'} \tag{32}$$

But from Eq. (23) and (30) it follows that

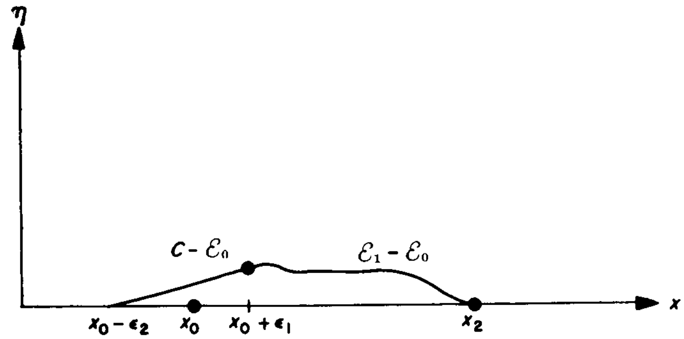
$$p_y(x, y) = \frac{\eta'}{\eta} + 0[(a_1 - a_0)] \tag{33}$$

and hence Eq. (32) becomes

$$\frac{d}{dx}(\eta^2 S_{yy})_p = -\eta^2 f_{yy} - 2\eta\eta' f_{yy'} - \eta'^2 f_{y'y'} + 0[(a_1 - a_0)^3] \tag{34}$$

Integrating Eq. (34) and using Eq. (29) and the fact that $\eta(x_2) = 0$, we obtain the result in Eq. (31).

Let us now find a path neighboring $\mathcal{E}_0$ from (x, y) to (x_2, y_2) for which the second variation is negative. Such a path in ηx-space is indicated in Sketch 4. We set $x = x_0 - \epsilon_2$



Sketch 4

and $\tilde{x} = x_0 + \epsilon_1$ with ϵ_1, $\epsilon_2 > 0$, but arbitrarily small. We choose the path C in Sketch 4 so that $\eta(x)$ is a straight line connecting $(x_0 - \epsilon_2, 0)$ and $(x_0 + \epsilon_1, \tilde{\eta})$. We follow $\mathcal{E}_1$ from $x_0 + \epsilon_1$ to x_2. The second variation for this path is

$$\delta^2 I = \frac{1}{2}\int_{x_0-\epsilon_2}^{x_0+\epsilon_1} [\eta^2 f_{yy} + 2\eta\eta' f_{yy'} + \eta'^2 f_{y'y'}]\, dx + \frac{1}{2}(\tilde{\eta}^2 S_{yy})\Big|_{x_0+\epsilon_1} + 0(\tilde{\eta}^3) \tag{35}$$

Expanding the integral in Eq. (35) about x_0 and noting that

$$\eta(x) = \frac{\tilde{\eta}(x - x_0 + \epsilon_2)}{(\epsilon_1 + \epsilon_2)}, \qquad x_0 - \epsilon_2 \leq x < x_0 + \epsilon_1 \tag{36}$$

we have the result

$$\delta^2 I = \frac{1}{2}\tilde{\eta}^2\left[\frac{\epsilon_1 + \epsilon_2}{3} f_{yy} + \frac{1}{3}(2\epsilon_1 - \epsilon_2)\frac{df_{yy'}}{dx} + f_{yy'} + \frac{1}{2}\left(\frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2}\right)\frac{df_{y'y'}}{dx} + \left(\frac{1}{\epsilon_1 + \epsilon_2}\right) f_{y'y'}\right]\Bigg|_{x_0} + \frac{1}{2}(\tilde{\eta}^2 S_{yy})\Bigg|_{x_0+\epsilon_1}$$
$$+ \tilde{\eta}^2 0(\,|\,\epsilon_1, \epsilon_2\,|\,)^2 + 0(\tilde{\eta}^3) \tag{37}$$

Using the asymptotic property of S_{yy} near x_0 as given by Eq. (19), we have the result

$$(\tilde{\eta}^2 S_{yy})|_{x_0+\epsilon_1} = \frac{-\tilde{\eta}^2}{\epsilon_1}(f_{y'y'})|_{x_0} + 0(\epsilon_0^1) \tag{38}$$

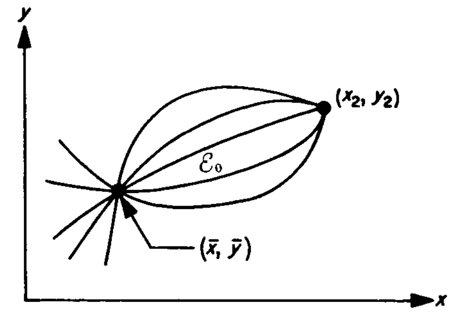
If we retain only the dominant terms, $\delta^2 I$ becomes

$$\delta^2 I = -\frac{1}{2}\tilde{\eta}^2 \frac{\epsilon_2}{\epsilon_1(\epsilon_1 + \epsilon_2)} f_{y'y'}|_{x_0} + \tilde{\eta}^2\, 0(|\epsilon_1, \epsilon_2|^0) + 0(\tilde{\eta}^3) \tag{39}$$

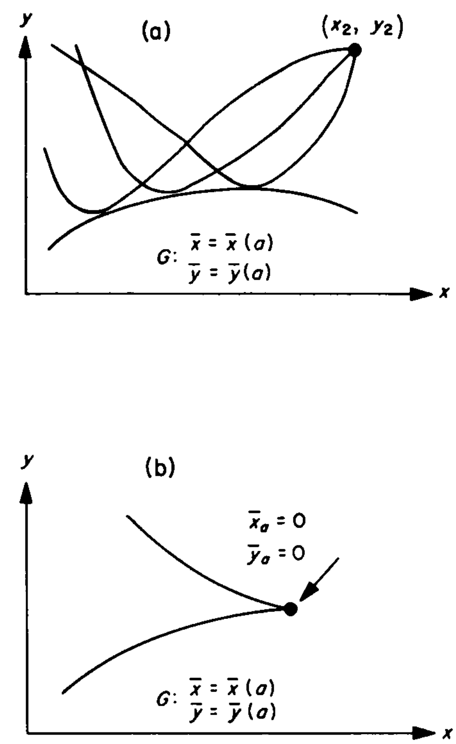
It is clear that by making $\tilde{\eta}$, ϵ_1, ϵ_2 sufficiently small, the sign of $\delta^2 I$ will be negative since $f_{y'y'} > 0$. This demonstrates the necessity of the boundedness of S_{yy} over the open interval (x_1, x_2). When x_0 and x_1 coincide, which corresponds to a conjugate point at the end point, the preceding proof does not apply because ϵ_2 must become zero in this case. For this special case one must consider the geometric nature of the conjugate point—a subject that has received considerable historical attention in the calculus of variations. The remainder of this Report is concerned with topics related to this facet of the conjugate-point problem.

V. The Envelope Theorem

Let us consider the propagation of the conjugate point on neighboring extremals. We have seen from the Jacobi equation in Eq. (24) that the vanishing of $\phi_a(x, a)$, $(x \neq x_2)$ marks a conjugate point and that, in particular, $\phi_a(x_0, a_0) = 0$ defines a conjugate at (x_0, y_0) on $\mathcal{E}_0$. The extremals neighboring $\mathcal{E}_0$ are generated by the parameter a. If we let $\bar{x}$ denote a vanishing point of ϕ_a other than at x_2, then $\bar{x} = \bar{x}(a)$ with $\bar{x} = x_0 = \bar{x}(a_0)$. From the theory of differential equations (see, for example, Ref. 3) and the implicit function theorem, it is known that if there exists a vanishing point (x_0, y_0) of ϕ_a for a given a_0, then for a range of values of a in some neighborhood of a_0, say, $a_1 < a_0 < a_2$, ϕ_a also vanishes for $x = \bar{x}(a)$. Furthermore, $\bar{x}(a)$ is a continuous function, since the coefficients of ϕ''_a, ϕ'_a, and ϕ_a in Eq. (24) are continuous near (x_0, y_0) because of our assumption of a cornerless extremal at (x_0, y_0). So we conclude that if (x_0, y_0) is a conjugate point on $\mathcal{E}_0$, then there is some neighborhood of extremals about $\mathcal{E}_0$ that possesses conjugate points in some region including (x_0, y_0). An investigation of the relationship $y = \phi(x, a)$, $\phi_a(x, a) = 0$ defining $x(a)$ and $y(a)$ will reveal that the behavior of conjugate points on neighboring extremals follows two patterns as indicated in Sketches 5 and 6. Conjugate points represent either a confluence of extremals through the same point as in Sketch 5 or a locus of contact points of the neighboring extremals with an envelope G of the family as described by the curve $\bar{x} = \bar{x}(a)$, $\bar{y} = \bar{y}(a)$ in Sketches 6a and b. The special case of an envelope with a cusp where $\phi_{aa}(\bar{x}, a) = 0$ is shown in Sketch 6b.



Sketch 5



Sketch 6

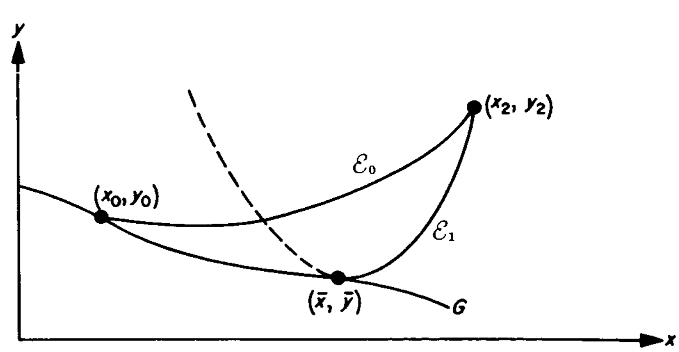
Let us develop here the so-called envelope theorem of the calculus of variations (Ref. 4), which is applicable to all conjugate points where $\phi_{aa} \neq 0$ (i.e., except the confluence type shown in Sketch 5 and at a cusp point as in Sketch 6b). We wish to prove that

$$\int_G^{\bar{x}} f(\bar{x}, \bar{y}, \bar{y}')\, d\bar{x} + S(\bar{x}, \bar{y}) = \text{constant} \tag{40}$$

where the integration is taken along G, the envelope (see Sketch 7). Differentiating Eq. (40) with respect to x we have, if Eq. (40) is true,

$$S_x + S_y \bar{y}' + f(\bar{x}, \bar{y}, \bar{y}') = 0 \tag{41}$$

But at $(\bar{x}, \bar{y})$ the slope of G and the contacting extremal are identical; thus $\bar{y}' = p$. Replacing $\bar{y}'$ with p, we observe that Eq. (41) becomes identical with Eq. (6a).



Sketch 7

Berkovitz and Dreyfus (Ref. 2) have shown that S_x and S_y are continuous along an extremal and their arguments apply as well to our relaxed definition for S. These continuity properties are equivalent to the well-known corner conditions of the calculus of variations. From this it follows that Eq. (6a) holds at a conjugate point also, in spite of the unboundedness of the second partials of S; for as we approach (x_0, y_0) from (x_2, y_2) along $\mathcal{E}_0$, we see that each term in Eq. (6a) is continuous, so that the sum $f + S_x + S_y p$ must remain at zero even at (x_0, y_0). Since this value of p is at least an extremizing value, it also follows that Eq. (6b) holds at (x_0, y_0).

Hence, Eq. (41) reduces to Eq. (6a) and the result in Eq. (40) is established. From Eq. (40) it now follows that

$$S(x_0, y_0) = S(\bar{x}, \bar{y}) + \int_{\substack{x_0 \\ G}}^{\bar{x}} f(\bar{x}, \bar{y}, p)\, dx \tag{42}$$

The result in Eq. (42) can also be obtained through the calculus of variations, using the Hilbert integral (Ref. 4).

VI. The Necessity of the Jacobi Condition for Envelope Contact Points

We wish to show that if (x_0, y_0) occurring at the initial point of $\mathcal{E}_0$ is an envelope contact point, then $\mathcal{E}_0$ is not minimizing. Having demonstrated the result that Eq. (6a) holds at a conjugate point, it is a simple matter using the envelope theorem in Eq. (42), to tie our dynamic-programming formulation to the standard calculus-of-variations proof hinging on the fact that G is not an extremizing curve. Since this is so, there exists a curve arbitrarily near G that produces a first-order reduction in the integral in Eq. (42), and if we were to follow that curve from (x_0, y_0) to $(\bar{x}, \bar{y})$ and then $\mathcal{E}_1$ from $(\bar{x}, \bar{y})$ to (x_2, y_2), we would obtain a value V for this path such that $V < S(x_0, y_0)$. As a matter of fact, letting I be given by

$$I = \int_{\substack{x_0 \\ G}}^{\bar{x}} f(\bar{x}, \bar{y}, p)\, dx \tag{43}$$

one can easily show that for a curve G' neighboring G by an amount $\delta y(x)$ and with $\delta y(x_0) = \delta y(\bar{x}) = 0$, the first-order change in I is given by

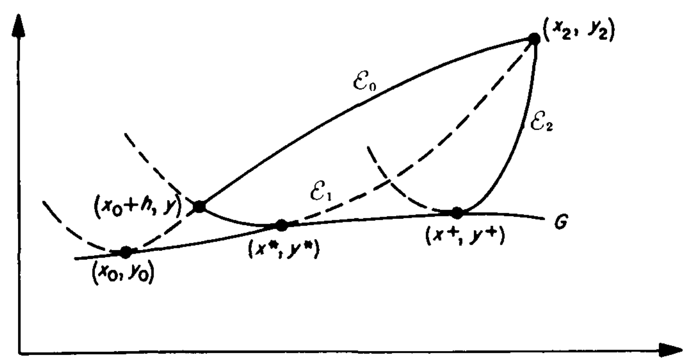
$$\begin{aligned}\delta I &= \int_{\substack{x_0 \\ G}}^{\bar{x}} \left[f_y - \frac{d}{dx}(f_{y'}) \right] \delta y dx + f_{y'}\, \delta y \Big|_{x_0}^{\bar{x}} \\ &= \int_{\substack{x_0 \\ G}}^{\bar{x}} \left[f_{y'y'} \left(\frac{\phi'^2_a}{\phi_{aa}} \right) \right] \delta y\, dx \end{aligned} \tag{44}$$

Since neither $f_{y'y'}$ nor ϕ'_a is zero on G, our contention that G is not an extremal is verified.

We have demonstrated the necessity of the Jacobi condition for the case where the conjugate point is an envelope contact point occurring at the initial point of $\mathcal{E}_0$. If (x_0, y_0) is an interior point of $\mathcal{E}_0$, it is clear that the above arguments still hold.

VII. A Global Property of Extremals Near an Envelope

For the case of envelope contact points, it should be observed that if the initial point (x, y) of $\mathcal{E}_0$ is taken sufficiently close to (x_0, y_0), but excluding this point, then $\mathcal{E}_0$, although a locally minimizing curve, does not provide a global minimum between (x, y) and (x_2, y_2). From Sketch 8, we wish to prove that there exists an alternate path from $(x_0 + h, y)$ to (x_2, y_2) for h sufficiently small, which provides a value V for the integral such that $V < S(x_0 + h, y)$. We will choose the alternate path as follows: From $(x_0 + h, y)$ we travel along $\mathcal{E}_1$, which intersects $\mathcal{E}_0$ at $(x_0 + h, y)$ and has a contact point with the envelope G at (x^*, y^*); then, we travel along a curve G' neighboring G from (x^*, y^*) to (x^+, y^+), and from (x^+, y^+) we travel along the extremal $\mathcal{E}_2$ to (x_2, y_2). The existence of $\mathcal{E}_1$ and $\mathcal{E}_2$ and their intersecting G is guaranteed by the continuity properties of $\phi(x, a)$ and the fact that $\phi_a(\bar{x}(a), a) = 0$ and $\phi_{aa} \neq 0$. In fact, if $a = a_0$



Sketch 8

generates the extremal $\mathcal{E}_0$, then it follows that $\mathcal{E}_1$ is generated by $a_0 + \delta a$, where δa is given by

$$\delta a = -\frac{2\phi'_a h}{\phi_{aa}} + 0(h^2) \qquad (45)$$

where the coefficient of h is evaluated at (x_0, y_0). Furthermore, x^*, the contact point of $\mathcal{E}_1$ with G, is given by

$$x^* = x_0 + 2h + 0(h^2) \qquad (46)$$

The existence of the envelope segment from (x^*, y^*) to (x^+, y^+) is guaranteed by the assumption that $\phi_{aa} \neq 0$.

Now $S(x_0 + h, y)$ is given by

$$S(x_0 + h, y) = S(x_0, y_0) - \int_{x_0 \atop \mathcal{E}_0}^{x_0+h} f\, dx \qquad (47)$$

and V is given by

$$V = \int_{x_0+h \atop \mathcal{E}_1}^{x^*} f\, dx + \int_{x^* \atop G'}^{x^+} f\, dx + S(x^+, y^+) \qquad (48)$$

With the use of the envelope theorem, subtracting V from $S(x_0 + h, y)$ yields

$$\begin{aligned} S(x_0 + h, y) - V = &\left[\int_{x_0 \atop G}^{x_0+h} f\, dx - \int_{x_0 \atop \mathcal{E}_0}^{x_0+h} f\, dx\right] \\ &+ \left[\int_{x_0+h \atop G}^{x^*} f\, dx - \int_{x_0+h \atop \mathcal{E}_1}^{x^*} f\, dx\right] \\ &- \left[\int_{x^* \atop G'}^{x^+} f\, dx - \int_{x^* \atop G}^{x^+} f\, dx\right] \end{aligned} \qquad (49)$$

The first two bracketed terms in Eq. (49) can be shown from the envelope theorem to be $S(x_0 + h, y)$ minus the value of the integral obtained by following $\mathcal{E}_1$ from $(x_0 + h, y)$ *to* (x_2, y_2). Since $\mathcal{E}_1$ has a conjugate point on its interior, this difference must be negative. Therefore, we must show that the third bracketed term in Eq. (49) can be made sufficiently negative to cause the difference $S(x_0 + h, y) - V$ to be positive. It seems clear that this can be accomplished.

To evaluate the negative contribution, we will require several expansions. If we let Δy be the difference between $\mathcal{E}_0$ and G, we have

$$\Delta y = \phi(\bar{x}, a(\bar{x})) - \phi(\bar{x}, a_0) \qquad (50)$$

Expanding Eq. (50) about x_0 and a_0 and using the fact that $\phi_a(\bar{x}, a(\bar{x})) = 0$, we have

$$\Delta y = \frac{1}{2}\left(\frac{\phi'^2_a}{\phi_{aa}}\right)_0 (x - x_0)^2 + 0\,[(x - x_0)^3] \qquad (51)$$

where $(\ )_0$ denotes evaluation at (x_0, y_0). Similarly, the difference between $\mathcal{E}_1$ and G is given by

$$\Delta y = -\frac{1}{2}\left(\frac{\phi'^2_a}{\phi_{aa}}\right)_* (x^* - x)^2 + 0\,[(x^* - x)^3] \qquad (52)$$

Since Δy in Eq. (51) varies in a second-order manner with $(x - x_0)$, it is necessary to expand the difference in integrals between $\mathcal{E}$ and G in Eq. (49) to second-order terms. Thus, setting ΔI_1 to be the expression

$$\Delta I_1 \triangleq \int_{x_0 \atop G}^{x_0+h} f\, dx - \int_{x_0 \atop \mathcal{E}_0}^{x_0+h} f\, dx \qquad (53)$$

we have

$$\begin{aligned} \Delta I_1 = &\int_{x_0 \atop G}^{x_0+h} (f_y \Delta y + f_{y'} \Delta y')\, dx \\ &+ \frac{1}{2}\int_{x_0 \atop G}^{x_0+h} (f_{yy}\Delta y^2 + 2f_{yy'}\Delta y \Delta y' + f_{y'y'}\Delta y'^2)\, dx \\ &+ 0(\,|\,\Delta y, \Delta y'\,|^3) \end{aligned} \qquad (54)$$

A similar expression results for ΔI_2, which is the difference appearing in the second bracketed term in Eq. (49). Expanding the coefficients of Δy and $\Delta y'$ about (x_0, y_0) and using Eq. (44) and (51), we obtain

$$\Delta I_1 + \Delta I_2 = -\frac{2}{3}\left(f_{y'y'} \frac{\phi'^4_a}{\phi^2_{aa}}\right)_0 h^3 + 0(h^4) \qquad (55)$$

Since $\Delta I_1 + \Delta I_2$ goes as h^3 and since G' in Eq. (49) is arbitrary other than the requirement that it pass through (x^*, y^*) and (x^+, y^+), we should easily be able to change the polarity of Eq. (49). As a matter of fact, the difference in integrals along G and G' appearing in Eq. (49) can be written as

$$\int_{\substack{x^* \\ G'}}^{x^+} f\,dx - \int_{\substack{x^* \\ G}}^{x^+} f\,dx = \int_{\substack{x^* \\ G}}^{x^+} f_{y'y'} \frac{\phi'^2_a}{\phi_{aa}} \delta y\,dx + \frac{1}{2}\int_{\substack{x^* \\ G}}^{x^+} (f_{yy}\,\delta y^2 + 2f_{yy'}\,\delta y \delta y' + f_{y'y'}\delta y'^2)\,dx + 0(\,|\,\delta y, \delta y'\,|^3\,) \tag{56}$$

where δy is the ordinate of G' minus the ordinate of G. Here, the first variation appearing in Eq. (44) and the vanishing of δy at the end points have been used. If we let x^+ and δy be given by

$$\left.\begin{aligned} x^+ &= x^* + 4h \\ \delta y &= \frac{1}{2}\left(f_{y'y'}\frac{\phi'^2_a}{\phi_{aa}}\right)_0 (x - x^*)(x - x^+) \end{aligned}\right\} \tag{57}$$

after some algebraic manipulation it follows that Eq. (56) becomes

$$\int_{\substack{x^* \\ G'}}^{x^+} f\,dx - \int_{\substack{x^* \\ G}}^{x^+} f\,dx = -\frac{8}{3}\left(f_{y'y'}\frac{\phi'^4_a}{\phi^2_{aa}}\right)_0 h^3 + 0(h^4) \tag{58}$$

Combining Eq. (55) and (58), we see that for the particular choice in Eq. (57) for the variation δy and interval $(x^+ - x^*)$, Eq. (49) becomes

$$S(x_0 + h, y) - V = 2\left(f_{y'y'}\frac{\phi'^4_a}{\phi^2_{aa}}\right)_0 h^3 + 0(h^4) \tag{59}$$

Hence, for $h > 0$ and sufficiently small, we have from Eq. (59)

$$S(x_0 + h, y) - V > 0 \tag{60}$$

which is the desired inequality. Therefore, for the initial point (x, y) of $\mathcal{E}_0$ sufficiently close to (x_0, y_0), *an envelope contact point*, $\mathcal{E}_0$ is not a globally minimizing curve.

VIII. Examples

We conclude this Report with two examples illustrating various facets of the material that has been covered.

Example A

Consider the problem

$$S(x, y) = \min_{\{y(x)\}} \int_x^1 (y'^2 - \pi^2 y^2)\,dx, \qquad y(1) = 0 \tag{61}$$

The extremals for this Stürm-Liouville-type problem (see Sketch 9) are

$$y = a \sin \pi x \tag{62}$$

where a is constant along each extremal, and for a particular extremal passing through the point (x_1, y_1), a is given by

$$a(x_1, y_1) = \frac{y_1}{\sin \pi x_1} \tag{63}$$

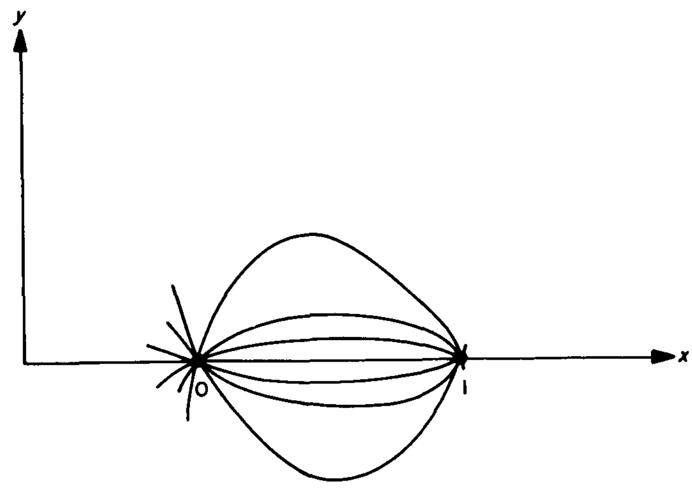
For this example, $S(x, y)$ is

$$S(x, y) = -\pi y^2 \cot \pi x \tag{64}$$

and

$$S_{yy} = -2\pi \cot \pi x \tag{65}$$

Hence, $(0, 0)$ is a conjugate point for all extremals emanating from $(1, 0)$.



Sketch 9

We note from Eq. (65) that near $x = 0$

$$xS_{yy} = -2 + 0(x^2) \tag{66}$$

which is in agreement with the asymptotic result in Eq. (19). Let us compute the second variation of Eq. (61) for the case when $x_1 < 0$, using a path construction as described in Sketch 4 and in Eq. (35). Here, the quantity $J(\tilde{x}, \tilde{\eta})$ becomes

$$J(\tilde{x}, \tilde{\eta}) = -\tilde{\eta}^2 \pi \cot \pi\tilde{x} \tag{67}$$

In Eq. (35) let us set ϵ_1 and $\epsilon_2 = \tilde{\eta}$. For this problem, Eq. (35) becomes

$$\delta^2 I = \frac{1}{2}\tilde{\eta}^2 [1 - \pi\tilde{\eta}(\pi + 2 \cot \pi\tilde{\eta})] + 0(\tilde{\eta}^3) \tag{68}$$

which is easily shown to be negative for some range of positive values of $\tilde{\eta}$. Hence, for boundary conditions that have $(0, 0)$ as an interior point, $\mathcal{E}_0$ is not a minimizing curve. In fact, for this case, one can find curves that make the functional in Eq. (61) arbitrarily negative. For the case where $x_1 = 0$, $y_1 = 0$, $\mathcal{E}_0$ is an absolutely minimizing curve, although it can be verified by substituting Eq. (62) into Eq. (61) that $\mathcal{E}_0$ is not unique.

Example B

Consider the problem

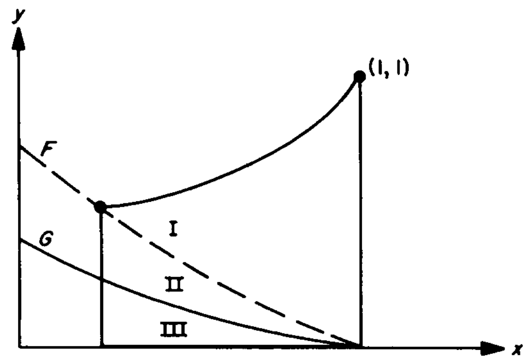
$$S(x, y) = \min_{\{\xi(x)\}} \left[\int_{\tau}^{\tau_2} \sqrt{y(\dot{x}^2 + \dot{y}^2)}\, d\tau \right] \tag{69}$$

where $x(\tau_2) = y(\tau_2) = 1$. The extremal solutions for this problem satisfying the boundary condition are of two types:

$$y = \left(\frac{x - a}{1 - a}\right)^2 + \left(\frac{1 - x}{2}\right)^2 \tag{70a}$$

The straight-line segment solution in Sketch 10

(70b)



Sketch 10

For (x, y) lying in region I, the parabola lying totally within I is a globally minimizing path. For (x, y) in II or III the straight-line segment solution is the better curve; for (x, y) in III there exists no parabolic solution, since the curve G is an envelope to this one-parameter family. From Eq. (70) we see, by setting $\partial y/\partial a = 0$ at a point of contact $(\bar{x}, \bar{y})$ of a parabolic extremal with G, that G satisfies the relation

$$\bar{y} = \frac{1}{4}(1 - \bar{x})^2 \tag{71}$$

It is clear from (70a) that through any point in I or II $(x < 1)$ there pass two extremals generated by two values of a except for (x, y) lying on G where the two extremals collapse to one. We choose the extremal with no contact point on G in its interior, and for this case it can be shown that

$$S(x, y) = \begin{cases} \text{I: } \dfrac{\sqrt{2}}{6}\left\{\left[y + 1 + \sqrt{(x-1)^2 + (y-1)^2}\right]^{3/2} - \left[y + 1 - \sqrt{(x-1)^2 + (y-1)^2}\right]^{3/2}\right\} \\ \text{II and III: } \dfrac{2}{3}(1 + y^{3/2}) \end{cases} \tag{72}$$

On the boundary curve F in Sketch 10 the two forms in Eq. (72) have the same value. In Eq. (72), $S(x, y)$ is strictly the optimal performance function. We observe that S_x and S_y are continuous along extremals, but that they are discontinuous across the boundary curve F. It can also be shown that the second partials of the expression for region I in Eq. (72) become infinite when the parabolic extremal touches the envelope.

IX. Summary

In the foregoing we have worked with the simplest problem in the calculus of variations in developing the conjugate-point necessary conditions from a dynamic-programming point of view. Several obvious generalizations should be noted. These results can be extended to the case where $y(x)$ is an n-dimensional vector. An envelope conjugate point then becomes a contact point with a generally $(n-1)$-dimensional envelope surface. The envelope theorem still applies for a one-parameter family of extremals whose contact points trace out a curve contained in the envelope surface. This approach is also applicable when the fixed end-point condition is relaxed so that the terminal point lies on a specified manifold.

The necessity of the Jacobi condition for the optimal-control problem couched in a dynamic-programming formulation (see Ref. 1, Chap. 4) can also be demonstrated, following proofs analogous to those given here. Basically, near a conjugate point one studies the asymptotic properties of the matrix S_{yy}, which satisfies a certain matrix Ricatti equation analogous to Eq. (17) and which is well known in optimal-control and estimation theory.

Finally, the result obtained in Section VII concerning the nonoptimality in a global sense of a relatively minimizing extremal whose end point is too near an envelope is of some consequence. It has obvious practical implications, particularly for practitioners of the numerical art of generating optimal or even suboptimal (but more practical) paths. It suggests that in certain cases of an end point near an envelope, a fundamentally different path may exist that could produce a substantial reduction in the performance function. This is, of course, implicit in the example presented in Sketch 10.

References

1. Dreyfus, Stuart, E., *Dynamic Programming and the Calculus of Variations,* Report No. R-441-PR, The RAND Corp., Santa Monica, Calif., August 1965.
2. Berkovitz, L. D., and Dreyfus, S. E., *A Dynamic Programming Approach to the Nonparametric Problem in the Calculus of Variations,* Report No. RM-4329-PR, The RAND Corp., Santa Monica, Calif., December 1964.
3. Coddington, E. A., and Levinson, N., *Theory of Ordinary Differential Equations,* McGraw-Hill Book Co., Inc., New York, 1955.
4. Bliss, G. A., *Lectures on the Calculus of Variations,* The University of Chicago Press, Chicago, Ill., 1945.