1960

# An introduction to the calculus of variations

Kenneth Cantwell Bouchelle
*Lehigh University*

# AN INTRODUCTION TO
# THE CALCULUS OF VARIATIONS

by

Kenneth Cantwell Bouchelle

A THESIS

Presented to the Graduate Faculty

of Lehigh University

in Candidacy for the Degree of

Master of Science

Lehigh University

June, 1960

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

MAY 5 1960
(Date)

Clarence A. Shook
Professor in Charge

A. E. Raynor
Head of the Department

ACKNOWLEDGEMENT

The author wishes to express his sincere appreciation for the guidance and assistance of Professor C. A. Shook in the preparation of this thesis, as well as for his inspired teaching of the several courses that the author took under him.

TABLE OF CONTENTS

# INTRODUCTION

The development of the theory of the calculus of variations has from the beginning been closely connected with that of the differential and integral calculus. Many of the methods of the latter are applicable to the calculus of variations. This is illustrated in Chapter 1, where a fairly careful study is made of one of the least complicated problems of the calculus of variations.

The general statement of the simplest type of problems dealt with in the calculus of variations is as follows: We are given a function $f(x,\phi,\phi')$ of three arguments, which in the region of the arguments considered is continuous and has continuous derivatives of the first and second orders. If in the function $f$ we replace $\phi$ by a function $y = \phi(x)$ and $\phi'$ by the derivative $y' = \phi'(x)$, $f$ becomes a function of $x$, and an integral of the form

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

becomes a definite number depending on the behavior of the function $y = \phi(x)$; that is, the integral is a function of the function $\phi(x)$. The fundamental problem of the calculus of variations is now as follows:

Among all the functions which are defined and continuous and possess continuous first and second derivatives in the interval $x_1 \leq x \leq x_2$, and for which the boundary values $y_1 = \phi(x_1)$ and $y_2 = \phi(x_2)$ are prescribed, to find that for which the integral $I(\phi)$ has the least possible value.

This general problem is investigated in some detail in Chapter 2. In that chapter, the fact is established that the solution of Euler's differential equation

$$f_y - \frac{d}{dx}f_{y'} = f_{y'y'}y'' + f_{yy'}y' + f_{xy'} - f_y = 0$$

plays a major role in determining the desired minimizing function $y = \phi(x)$. In general it is not possible to solve Euler's differential equation explicitly in terms of elementary functions. However, in important special cases and in fact in most of the classical examples of the calculus of variations, the equation can be solved by means of integration. We briefly mention here three special cases.

First is the case in which the function f does not contain the derivative $y' = \phi'$ explicitly; that is, $f = f(x,\phi)$. Here, Euler's differential equation is no longer a differential equation. It becomes simply the equation $f_y(x,y) = 0$, which is an implicit definition of the solution $y = y(x)$.

The second special case is that in which the function f does not contain the function $y = \phi(x)$ explicitly; that is, $f = f(x,y')$. Here, Euler's differential equation becomes $\frac{d}{dx}f_{y'} = 0$, which at once gives the result $f_{y'} = c$, where c is an arbitrary constant of integration. This last equation may be used to express y' as a function g(x,c) and we thus have the equation

$$y' = g(x,c)$$

from which we obtain by a simple integration

$$y = \int_0^x g(t,c)\,dt + a.$$

The third special case is that in which f does not contain x explicitly; that is, $f = f(y,y')$. It can be shown that in this case

$$E = f(y,y') - y'f_{y'}(y,y') = c$$

is an integral of Euler's differential equation. The solution of Euler's differential equation, again obtained by integration, will be of the form $x = h(y,c) + a$, which can be solved for y to obtain the function y(x,c,a).

This third and most important special case is illustrated by the problem of Chapter 3; namely, the problem of determining a surface of revolution of minimum area passing through two given points.

# CHAPTER 1

## DETERMINATION OF THE SHORTEST ARC BETWEEN TWO POINTS

The problem of determining the shortest arc between two given points can serve as an illustration of the general theory and methods of the calculus of variations, though it is a specialized case. We shall deal with this problem in this chapter, assuming at the outset that we know nothing about shortest distances or straight lines.

Let the coordinates of the two points to be joined by an arc of minimal length be $(x_1,y_1)$ and $(x_2,y_2)$. These points will be identified by the numbers 1 and 2. The equation of the minimal-length arc, in case such an arc exists, will be supposed to be of the form

$$y = y(x), \qquad x_1 \leq x \leq x_2.$$

The length of any such arc is given by the line integral I

$$I = \int_1^2 ds,$$

where $ds = \sqrt{dx^2 + dy^2}$ and $dy = y'dx$ $(y' = dy/dx)$.

Thus the length of the arc is

$$(1.1) \qquad I = \int_{x_1}^{x_2} f(y')\, dx,$$

where $f(y') = (1 + y'^2)^{1/2}$.

The only restrictions to be placed on the functions $y(x)$ to be considered are that, on the interval $x_1 \leq x \leq x_2$, $y(x)$ must be continuous and it must consist of a finite number of arcs on each of which it is continuously differentiable. That is, the arcs $y = y(x)$ are required to be continuous and to be composed of a finite number of arcs on each of which the tangent to the arc turns continuously. Such an arc will be called an "admissible arc", and the function defining such an arc will be called an "admissible function". The problem of finding the shortest distance between the two points 1 and 2 is therefore the problem of finding that admissible function which satisfies the end conditions and which makes the integral (1.1) a minimum.

Let it be supposed that a particular arc, which we shall denote by $E_{12}$, with the equation $y = g(x)$, is known to be the solution arc. Let $\eta(x)$ be an admissible function satisfying the further condition that $\eta(x_1) = \eta(x_2) = 0$. Then, the equation

$$y = g(x) + a\eta(x),$$

where a is any arbitrary constant, represents a one-parameter family of admissible arcs, each of which passes through the two points 1 and 2, and the soluion arc is included in this family (that is, when a = 0).

For this function $y = g(x) + a\eta(x)$, the value of the integral (1.1) taken along any arc depends on a, and is therefore written

$$(1.2) \qquad I(a) = \int_{x_1}^{x_2} f(g' + a\eta')\,dx.$$

Along the solution arc $E_{12}$ this integral has the value $I(0)$.

From elementary calculus it is known that if a function $I(a)$ has a minimum value at $a = a_1$, then $I'(a_1) = 0$, where $I'(a)$ denotes the derivative of $I(a)$ with respect to the argument a. Hence the assumption that a solution is furnished by $y = g(x)$ implies that in equation (1.2) we must have $I'(0) = 0$. From (1.2) we find that

$$(1.3) \qquad I'(0) = \int_{x_1}^{x_2} f_{y'}(g')\eta'\,dx = 0,$$

where $f_{y'}(g')$ denotes the derivative of $f(y')$ with respect to $y'$, evaluated at $y' = g'$. Since this derivative $f_{y'}(g')$ is a function of x, it will be denoted by $M(x)$.

The implications of the necessity of the condition $I'(0) = 0$ will more easily be seen if we first prove the following lemma, which will be used in later chapters as well as in the current problem.

Fundamental Lemma: Let $M(x)$ be a function either continuous on the interval $x_1 \leq x \leq x_2$ or else such that the interval can be subdivided into a finite number of subintervals on each of which $M(x)$ is continuous. Let

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

for every admissible function $\eta(x)$ satisfying $\eta(x_1) = \eta(x_2) = 0$.

Then, $M(x)$ is necessarily a constant.

To prove the lemma, we note that since

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

for all $\eta(x)$ such that $\eta(x_1) = \eta(x_2) = 0$, then for all of these same functions $\eta(x)$, we must have

$$(1.4) \qquad \int_{x_1}^{x_2} \left[M(x) - C\right]\eta'(x)\,dx = 0,$$

where $C$ is any constant. Now, the function $\eta_1(x)$, given by

$$(1.5) \qquad \eta_1(x) = \left[\int_{x_1}^{x} M(x)\,dx\right] - C_1(x - x_1)$$

is an admissible function which vanishes at $x_1$. It will also vanish at $x_2$, provided we choose $C_1$ to satisfy

$$(1.6) \qquad C_1 = \frac{1}{x_2 - x_1}\int_{x_1}^{x_2} M(x)\,dx.$$

Thus equation (1.4) must hold for $\eta(x) = \eta_1(x)$, and this yields

$$\int_{x_1}^{x_2} \left[M(x) - C_1\right]^2 dx = 0.$$

But this last equation is satisfied only if $M(x) \equiv C_1$. This concludes the proof of the lemma.

Thus, along the solution arc $E_{12}$, direct application of the above lemma to (1.3) yields the requirement that

$$(1.7) \qquad f_{y'}(g') = \frac{g'}{\sqrt{1 + g'^2}} = C,$$

where C is some constant. From this it follows at once that g' must itself equal a constant, which in turn implies that the solution arc $E_{12}$ is a straight line joining the points 1 and 2; that is, a necessary condition for $E_{12}$ to be an arc of minimal length is that $E_{12}$ be a straight line.

It has not yet been shown that the above necessary condition is sufficient; we now show that such is indeed the case.

Let $\phi(x)$ denote the increment that must be added to the ordinate of $E_{12}$ at each value x in order to obtain any arbitrarily selected admissible arc $C_{12}$ joining the points 1 and 2. That is, $C_{12}$ will have as its equation

$$y = g(x) + \emptyset(x), \qquad (x_1 \leq x \leq x_2)$$

where $\emptyset(x_1) = \emptyset(x_2) = 0$. The difference of the lengths of $C_{12}$ and $E_{12}$ can be expressed as

$$I(C_{12}) - I(E_{12}) = \int_{x_1}^{x_2} \left[ f(g' + \emptyset') - f(g') \right] dx$$

or, using Taylor's formula,

$$(1.8) \qquad I(C_{12}) - I(E_{12}) = \int_{x_1}^{x_2} f_{y'}(g')\emptyset' dx + \frac{1}{2}\int_{x_1}^{x_2} f_{y'y'}(g' + \theta\emptyset')\emptyset'^2 dx,$$

where $f_{y'y'}$ is the second derivative of the function f with respect to y', and θ is a suitable value between 0 and 1. In this last equation, recall that $f_{y'}$ is constant, by (1.7), and $\emptyset(x_1) = \emptyset(x_2) = 0$, so that the first integral vanishes. Furthermore, since

$$f_{y'y'} = \frac{1}{(1 + y'^2)^{3/2}}$$

is clearly always positive, it follows that the second integral of (1.8) is never negative. This means that $I(C_{12}) - I(E_{12}) \geq 0$, equality holding only if $\emptyset'(x) \equiv 0$, in which case $\emptyset(x) = C$; but since $\emptyset(x_1) = \emptyset(x_2) = 0$, this means that $C_{12}$ coincides with $E_{12}$.

Thus it has been proved that the shortest arc between two points is necessarily the straight-line segment joining them, and that this segment is actually shorter than any other admissible arc with the same end points. However, the methods of this proof are not in general extendable to less specialized problems of the calculus of variations. A second proof which can be extended to somewhat more general problems is now given.

Consider a straight-line segment $E_{34}$, of variable finite length, which moves so that its end points describe simultaneously two curves C and D, and let the equations of these curves in parametric form be

$$C: \quad x = x_3(t), \qquad y = y_3(t),$$

$$D: \quad x = x_4(t), \qquad y = y_4(t). \quad (t_a \leq t \leq t_b)$$

We suppose that these functions are continuous on the given interval, and that the interval can be subdivided into a finite number of subintervals on each of which the functions have continuous derivatives such that $x'^2 + y'^2 \neq 0$.

The line $E_{34}$ has as length I given by

$$I = \sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2},$$

and using the notation $p = \frac{y_4 - y_3}{x_4 - x_3}$ to denote its slope, the differential of the length of $E_{34}$ is

$$(1.9a)\quad dI(E_{34}) = \frac{(dx_4 - dx_3) + p(dy_4 - dy_3)}{\sqrt{1 + p^2}} ,$$

or, equivalently and more neatly,

$$(1.9b)\qquad dI(E_{34}) = \frac{dx + p\,dy}{\sqrt{1 + p^2}} \Bigg|_3^4 .$$

Now, consider the integral I* defined by the formula

$$I^* = \int \frac{dx + p\,dy}{\sqrt{1 + p^2}} .$$

This integral will be well defined along an arbitrary curve C, when p is a function of x and y, if the integral is calculated by first substituting for x, y, dx, and dy their values in terms of t and dt as obtained from the parametric equations of the curve C.

Now, let $t_3$ and $t_5$ be two parametric values which define points 3 and 5 on curve C, and which simultaneously define the points 4 and 6 respectively on the curve D. Integrating (1.9b) with respect to t from $t_3$ to $t_5$ shows that the difference of the lengths $I(E_{34})$ and $I(E_{56})$ of the moving line segment in the two positions $E_{34}$ and $E_{56}$ is given by

$$(1.10) \quad I(E_{56}) - I(E_{34}) = I^*(D_{46}) - I^*(C_{35}).$$

We now introduce the notion of a field. For our purposes here, a field F is defined as a region of the xy-plane which can be covered with a one-parameter family of straight-line[1] segments each of which intersects the fixed curve D exactly once, and which has the additional property that one and only one of these segments passes through each point (x,y) of the field F.

At every point of F, the straight line of the field through that point has a slope p(x,y), which we call the slope-function of the field. The integral $I^*$ defined above has a definite value along any arc $C_{35}$ in the field F having equations of the form

$$x = x(t), \quad y = y(t). \qquad (t_3 \leq t \leq t_5)$$

Similarly, the integral $I_*$, obtained by replacing the p of $I^*$ by the slope function p(x,y), also has a definite value along any such arc $C_{35}$ so defined.

The integral $I_*$ has two interesting properties. First, its values are the same along all curves $C_{35}$ in the field having the same end points 3 and 5. To prove this, recall that through each point of $C_{35}$ there is by definition of the field exactly one straight line of the

[1] These straight-line segments are for this special case the so-called extremals, defined for the general problem of Chapter 2 on page 19.

field intersecting the curve D. Applying (I.10) to this one-parameter family of straight lines determined by the points of $C_{35}$ gives

$$I_*(C_{35}) = I_*(D_{46}) - I(E_{56}) + I(E_{34})$$

where every term of the right member of the equation is determined by the two points 3 and 5; that is, $I_*(C_{35})$ is independent of the particular curve in F that is chosen.

A second property of the integral $I_*$ is that along each segment of any one of the straight lines of the field F, the value of $I_*$ is equal to the length of the segment. This is clear since along each of these lines the differentials dx and dy satisfy the equation $dy = p\,dx$, so that the integrand of $I_*$ reduces to simply $\sqrt{1 + p^2}\,dx$, which is the integrand of the length integral.

Now, in order to conclude our second proof that a straight-line segment joining a pair of points 1 and 2 is shorter than every other arc joining them, consider the field F' formed by covering the entire xy-plane with the one-parameter family of lines parallel to $E_{12}$. Let $C_{12}$ be an arc in the field F' joining 1 and 2, and let $C_{12}$ be defined by parametric equations of the form

$$x = x(t), \quad y = y(t), \qquad (t_1 \leq t \leq t_2).$$

Thus we have

$$(1.11) \qquad I(E_{12}) = I^*(E_{12}) = I_*(E_{12}) = I_*(C_{12}).$$

If we now call $\theta$ the angle between the tangent to C and the tangent to E, we may write

$$(1.12) \qquad \cos\theta = \frac{x' + py'}{\sqrt{(1 + p^2)(x'^2 + y'^2)}} ,$$

and since $ds = \sqrt{x'^2 + y'^2}\, dt$, then

$$(1.13) \qquad I_* = \int \frac{dx + p\, dy}{\sqrt{1 + p^2}} = \int \cos\theta\, ds ,$$

so that (1.11) may be extended to

$$(1.14) \qquad I(E_{12}) = \int_{s_1}^{s_2} \cos\theta\, ds.$$

The difference between the values of I along $C_{12}$ and $E_{12}$ is therefore

$$(1.15) \quad I(C_{12}) - I(E_{12}) = \int_{s_1}^{s_2} (1 - \cos\theta)\, ds \geq 0$$

where equality holds only if $\cos\theta = 1$ at every point of $C_{12}$; that is, only if $C_{12}$ coincides with $E_{12}$ identically.

We have thus concluded the second proof of the fact that the straight line $E_{12}$ is indeed shorter than any other arc joining the points 1 and 2. This second proof is itself more general, since the class of arcs in parametric form is larger than the "admissible" arcs of the first proof. However, the methods and results of even the second proof are still quite specialized, so we proceed directly to the general problem of the calculus of variations.

# CHAPTER 2

## THE GENERAL PROBLEM OF THE CALCULUS OF VARIATIONS

Many of the problems with which the calculus of variations is concerned involve minimizing integrals which are special cases of the more general integral

$$I = \int_{x_1}^{x_2} f(x,y,y')\,dx \tag{2.1}$$

where for each particular problem the integrand function $f(x,y,y')$ of (2.1) is known. Hence, it is desirable to find characteristic properties of minimizing arcs for this more general integral.

In order to define a class of admissible arcs for the general case above, let us assume that there is a region R of sets of values $(x,y,y')$ in which the function $f(x,y,y')$ is continuous and has continuous derivatives at least up to and including the fourth order. The sets of values $(x,y,y')$ interior to the region R will be called "admissible sets". An arc of the form

$$y = y(x) \qquad (x_1 \leq x \leq x_2) \tag{2.2}$$

will be called an "admissible arc" if it is continuous and is continuously differentiable at all but a finite number of points and all of the sets of values $(x,y,y')$

on it are admissible sets. Thus for an admissible arc $y = y(x)$ the interval from $x_1$ to $x_2$ can always be divided into a finite number of subintervals on each of which $y(x)$ is continuous and has a continuous derivative.

There are in all four conditions which will now be proved to be necessary for an arc to minimize (2.1); they will be numbered with Roman numerals in the development that follows.

Let us now suppose as in the previous chapter that a particular arc $E_{12}$ with equation in the form (2.2) actually furnishes a minimum for (2.1). Let $\eta(x)$ be an admissible function satisfying $\eta(x_1) = \eta(x_2) = 0$. Then the family

$$(2.3) \qquad y = y(x) + a\eta(x), \qquad (x_1 \le x \le x_2)$$

where a is a parameter, contains $E_{12}$ for any $\eta(x)$ by setting $a = 0$. The integral (2.1) now is

$$(2.4) \qquad I(a) = \int_{x_1}^{x_2} f(x, y + a\eta, y' + a\eta')\, dx$$

where the particular value $I(0)$ must be a minimum, thus giving the requirement that $I'(0) = 0$.

Actual differentiation yields

$$(2.5) \qquad I'(0) = \int_{x_1}^{x_2} \left[f_y\eta + f_{y'}\eta'\right] dx$$

where $f_y$ and $f_{y'}$ denote the partial derivatives of $f(x,y,y')$ with respect to y and y' respectively.

Integrating by parts the first member of the integrand of (2.5) gives

$$\int_{x_1}^{x_2} f_y \eta dx = \eta \int_{x_1}^{x} f_y dx \Big]_{x_1}^{x_2} - \int_{x_1}^{x_2} \Big[\eta' \int_{x_1}^{x} f_y\Big] dx$$

or since $\eta(x_1) = \eta(x_2) = 0$, equation (2.5) may be written

$$(2.6) \qquad I'(0) = \int_{x_1}^{x_2} \Big[f_{y'} - \int_{x_1}^{x} f_y dx\Big] \eta' dx.$$

The integral on the right side must vanish for every admissible function $\eta$ which satisfies the end conditions, so we may now apply the fundamental lemma of the first chapter; that is, the bracketed part of the integrand of (2.6) is constant. This leads directly to the first necessary condition.

I For every minimizing arc $E_{12}$ there exists a constant c such that

$$(2.7) \qquad f_{y'} = \int_{x_1}^{x} f_y dx + c$$

holds identically on $E_{12}$.

Furthermore, on each sub-arc of $E_{12}$ on which the tangent turns continuously, we must have

$$(2.8) \qquad \frac{d}{dx} f_{y'} - f_y = 0.$$

This last equation is obtained directly from (2.7) by differentiation and is the familiar Euler's differential equation.

The solutions of equation (2.7) are usually called extremals because they are the only curves which can give the integral I an extreme value. However, we shall define an extremal as a solution of Euler's equation which has a continuously turning tangent (that is, is continuously differentiable) and which has a continuous second derivative $y''(x)$.

Now let 3 be any arbitrary point on the minimizing arc $E_{12}$ and let 4 be another point of this arc near enough to 3 that the portion of the arc between them is continuously differentiable; that is, choose the point 4 so near to 3 that there is no corner between them. Through the point 3 pass an arbitrary curve C with equation given by $y = Y(x)$. The fixed point 4 can now be joined to a moveable point 5 on the curve C (see Figure 1) by a one-parameter family of arcs $E_{54}$ containing the arc $E_{34}$ as one of its members.

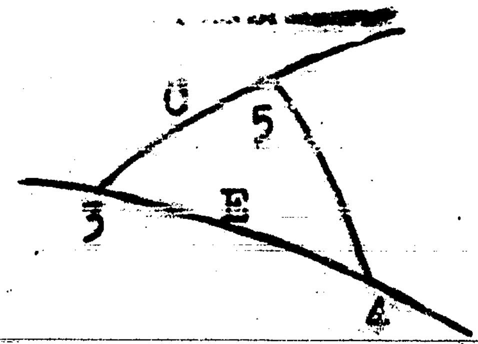


Figure 1

Such a one-parameter family is given by

$$(2.9) \quad y = y(x) + \frac{Y(a) - y(a)}{x_4 - a}(x_4 - x) = y(x,a)$$

where the parameter a is simply $x_5$.

It will be recalled that by hypothesis the arc $E_{12}$ makes the integral $I(E_{12})$ given by (2.1) a minimum. Therefore, as the point 5 moves along C away from the point 3, the integral

$$(2.10) \quad I(C_{35} + E_{54}) = \int_{x_3}^{x_5} f(x,Y,Y')dx + I(E_{54})$$

can not decrease from the value $I(E_{34})$ which it has when the point 5 is at 3, and furthermore at the point 3, the differential of (2.10) with respect to $x_5$ must not be negative.

In order to express this result more neatly, consider now a one-parameter family of extremal arcs

$$(2.11) \qquad y = y(x,b) \qquad x_3 \le x \le x_4$$

satisfying the Euler differential equation of the form

$$(2.12) \qquad \frac{\partial}{\partial x} f_{y'} - f_y = 0$$

where the partial derivative symbol must be used since y is here a function of two variables. Regarding $x_3$, $x_4$, and b all as variables, the value of the integral I along an arc is of the form

$$(2.13) \quad I(x_3,x_4,b) = \int_{x_3}^{x_4} f(x,\ y(x,b),\ y'(x,b))dx.$$

Along an extremal, using (2.12), we have

$$(2.14) \quad \frac{\partial f}{\partial b} = f_y y_b + f_{y'} y'_b = y_b \frac{\partial}{\partial x} f_{y'} + y'_b f_{y'} = $$
$$= \frac{\partial}{\partial x}(f_{y'} y_b).$$

So, the partial derivatives of the function $I(x_3,x_4,b)$ are

$$(2.15) \quad \frac{\partial I}{\partial x_3} = -f\Big|^3, \quad \frac{\partial I}{\partial x_4} = f\Big|^4, \quad \frac{\partial I}{\partial b} = f_{y'} y_b \Big|_3^4$$

where the arguments of f and its derivatives are understood to be the values y, y' belonging to the family of (2.11). Now, suppose the variables $x_3$, $x_4$, and b are functions $x_3(t)$, $x_4(t)$, b(t) respectively of a variable t, so that the end-points 3 and 4 of the extremals (2.11) describe simultaneously two curves C and D (see Figure 2) whose equations are

$$(2.16) \quad \begin{aligned} x_3 &= x_3(t), \quad y_3 = y(x_3(t),b(t)) = y_3(t) \\ x_4 &= x_4(t), \quad y_4 = y(x_4(t),b(t)) = y_4(t) \end{aligned}$$

The differentials along these curves C and D are found by attaching suitable subscripts 3 or 4 to x, dx and dy in

(2.17) $dx = x'(t)dt, \quad dy = y'dx + y_b db.$

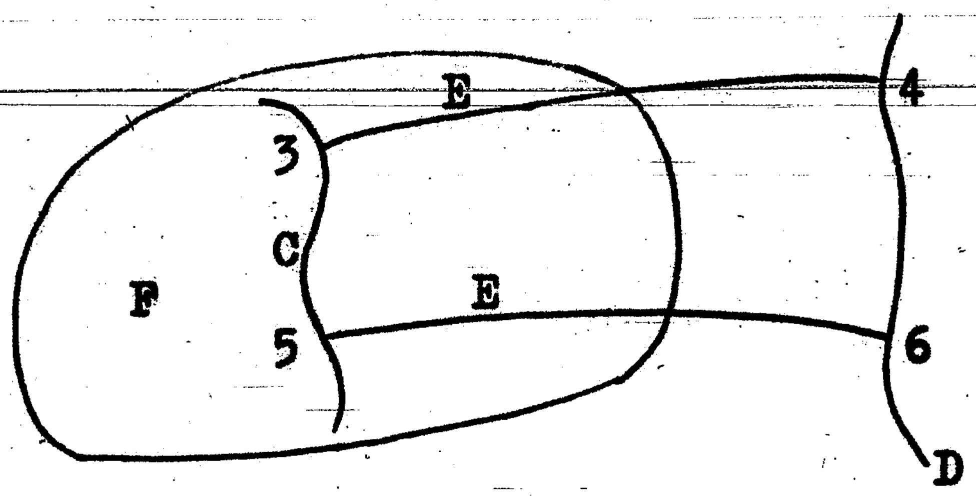


Figure 2

From (2.15) we compute the differential

$$(2.18) \qquad dI = \frac{\partial I}{\partial x_3}dx_3 + \frac{\partial I}{\partial x_4}dx_4 + \frac{\partial I}{\partial b}db =$$

$$f\,dx + f_{y'}y_b db \Big|_3^4.$$

With the help of equations (2.17), this may be written

$$(2.19) \quad dI(E_{34}) = f(x,y,p) + (dy - p\,dx)f_{y'}(x,y,p) \Big|_3^4.$$

This equation holds at every position where $E_{34}$ is an extremal, and at every position where it satisfies (2.8). In (2.19) the differentials dx and dy are those of the curves C and D at the points 3 and 4 of Figure 2, and the values of p are the slope of $E_{34}$ at these two points.

Now, let I** be defined by

$$(2.20) \quad I^{**} = \int \left[ f(x,y,p)dx + (dy - p\,dx)f_{y'}(x,y,p) \right]$$

Since the variable arc $E_{34}$ is assumed to be an extremal, then for $E_{34}$ and $E_{56}$ of Figure 2 we have

$$(2.21) \quad I(E_{56}) - I(E_{34}) = I^{**}(D_{46}) - I^{**}(C_{35}).$$

We again introduce the notion of a field F, as in Chapter 1, except that here we are not limited to straight lines as extremals. That is, we define a region F of the plane to be a field if it has associated with it a one-parameter family of extremals, each intersecting a fixed curve D in exactly one point, and such that through each point (x,y) of the field F there passes one and only one extremal of the family. We again define the slope-function of the field as the slope p(x,y) of the extremal at each point (x,y) of the field.

Substituting this slope function for p in $I^{**}$, and denoting the resulting integral by $I_{**}$, it is clear that the integrand of $I_{**}$ depends only on x, y, dx, and dy. Hence, $I_{**}(C_{35})$ depends only on the points 3 and 5, and is independent of the arc $C_{35}$, since the points 3 and 5 determine the three values $I(E_{56})$, $I(E_{34})$ and $I_{**}(D_{46})$ in equation (2.21), where $I^{**}$ has been replaced by $I_{**}$. Further, along an extremal arc of the field F the value of $I_{**}$ is the same as that of I, since along each extremal we have $dy = p\,dx$, which reduces the integrand of $I_{**}$ to $f(x,y,p)\,dx$.

Now returning to equation (2.10), where (see Figure 1) the arc $D_{46}$ of (2.21) is the fixed point 4, we note that the derivative of the first integral of that equation with respect to $x_5$ is $f(x_5,Y,Y')$. Thus when 5 is at 3, the differential of $I(C_{35} + E_{54})$ is the value at 3 of the quantity

$$(2.22) \qquad f(x,Y,Y')dx - f(x,y,y')dx - (dy - y'dx)f_{y'}(x,y,y'),$$

where dy and dx belong to C, and satisfy $dy = Y'dx$. At the point 3, since the ordinates of C and E are equal, we may write equation (2.22) in the form

$$(2.23) \qquad \Big[f(x,y,Y')-f(x,y,y')-(Y'-y')f_{y'}(x,y,y')\Big]dx\Big|^{3}$$

The function (2.23) is called the Weierstrass E-function and is usually denoted by $E(x,y,y',Y')$.

Recalling now the earlier result on page 20 that the differential of $I(C_{35} + E_{54})$ must not be negative, we now state a second necessary condition that an arc $y = y(x)$ minimize the integral (2.1):

II At every element $(x,y,y')$ of a minimizing arc $E_{12}$ the condition $E(x,y,y',Y') \geq 0$ must be satisfied for every admissible set $(x,y,Y')$ different from $(x,y,y')$.

When Taylor's formula is applied to the Weierstrass E-function, it may be written in the form

$$(2.24)\quad E(x,y,y',Y')=\tfrac{1}{2}(Y'-y')^2 f_{y'y'}(x,y,y'+\theta(Y'-y'))$$

where $0<\theta<1$. Now, let $Y'$ approach $y'$, and we have immediately a third necessary condition:

III At every element $(x,y,y')$ of a minimizing arc $E_{12}$ the condition

$$(2.25)\qquad f_{y'y'}(x,y,y')\geq 0$$

must be satisfied.

For the fourth necessary condition, we first observe that through a fixed point 1, there passes in general a one-parameter family of extremals. If such a family has an envelope G, then the contact point 3 of an extremal arc $E_{12}$ of the family with G will be called a "point 3 conjugate to 1 on $E_{12}$".

Let $E_{14}$ and $E_{13}$ be two extremals of the family through the point 1 touching an envelope G at their end

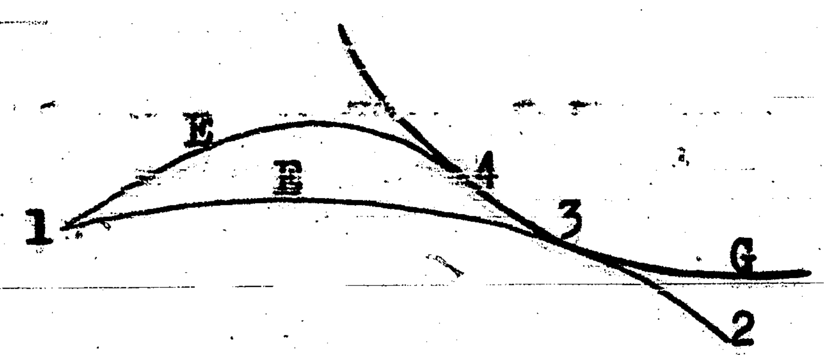


Figure 3

points (see Figure 3). If we replace the arc $C_{35}$ of equation (2.21) by the fixed point 1 and the arc $D_{46}$ by the arc $G_{43}$ we have

$$(2.26) \qquad I(E_{13}) - I(E_{14}) = I_{**}(G_{43}).$$

Furthermore, since each extremal is tangent to G, the slope of the envelope G is at each point the slope-function used to define $I_{**}$ so that $I_{**}(G_{43}) = I(G_{43})$. Hence we have that the values of the integral I along the arcs $E_{14}$, $E_{13}$, and $G_{43}$ must satisfy the relation

$$(2.27) \qquad I(E_{14}) + I(G_{43}) = I(E_{13})$$

for every position of the point 4 preceding point 3 on the envelope G. Thus the value of I along the composite arc $E_{14} + G_{43} + E_{32}$ in Figure 3 is equal to its value along $E_{12}$.

It is now necessary to look at the possibility of $G_{43}$ itself being an extremal. By definition, an extremal has continuous first and second derivatives so that along any extremal the Euler equation (2.8) can be written in the form

$$(2.28) \qquad \frac{d}{dx} f_{y'} - f_y = f_{y'x} + f_{y'y}y' + f_{y'y'}y'' - f_y = 0.$$

A property of a differential equation of this type is that when it can be solved for $y''$ there is one and only one solution of it through any arbitrarily selected point and direction $(x_3, y_3, y_3')$. Clearly it can be solved for $y''$ if we require that $f_{y'y'}$ be different from zero at every point of the extremal $E_{12}$. Under these conditions, it is clear that if $G_{43}$ were an extremal it would necessarily coincide with $E_{12}$ at every point, as it would also with every other extremal arc through the point 1. Then, there would be no one-parameter family such as has been supposed.

Thus, $G_{43}$ can not be an extremal, so it can be replaced by an arc $C_{43}$ giving I a smaller value. Consequently, in every neighborhood of $E_{12}$ there is a composite arc $E_{14} + C_{43} + E_{32}$ giving I a smaller value than $E_{12}$ which contradicts the assumption that $I(E_{12})$ furnished a minimum.

This gives the fourth necessary condition:

IV On a minimizing extremal arc $E_{12}$ with $f_{y'y'} \neq 0$ everywhere on it, there can be no point 3 conjugate to 1 between 1 and 2.

We have thus established four necessary conditions for the integral (2.1) to be a minimum. With relatively moderate changes, they can be made over into conditions which are also sufficient to insure an extreme value for the integral. Before investigating these changes in detal, we first consider a fundamental sufficiency theorem which for some problems of the calculus of variations is all that is needed in the way of sufficient conditions.

In order to establish this theorem it is necessary to be more explicit in stating the properties of the family of extremal arcs of the field F defined earlier in this chapter. It is supposed that the family has the form

$$(2.29) \qquad y = y(x,a) \qquad a_1 \le a \le a_2; \quad x_1(a) \le x \le x_2(a)$$

where the functions $y(x,a)$, $y'(x,a)$ and their partial derivatives up to and including those of the second order as well as the functions $x_1(a)$ and $x_2(a)$ defining the end-points of the extremals are all continuous. It is further supposed that the point of the curve D (see Figure 4) on each extremal is defined by a function $x = h(a)$ which with its first derivatives is continuous on

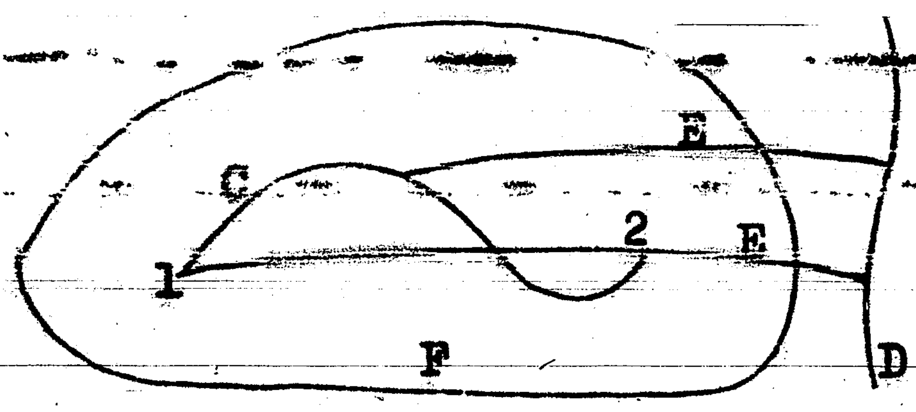


Figure 4

the interval from $a_1$ to $a_2$ and also it is supposed that the partial derivative $y_a$ is everywhere different from zero on the extremal arcs. We are further assuming that to each point $(x,y)$ in F there corresponds a value $a(x,y)$ which defines the unique extremal of the field through that point. Since we require that $y_a$ be different from zero, it can be proven that $a(x,y)$ and its first partial derivatives are continuous in F, which is also true of the slope-function $p(x,y) = y'(x,a(x,y))$ of the field F.

The integral

$$(2.30) \quad I_{**} = \int \left[ f(x,y,p)dx + (dy - p\,dx)f_{y'}(x,y,p) \right],$$

where it will be recalled p is the slope-function of the field F, has a definite value on each admissible arc $C_{12}$ in the field, and indeed this value is the same for all such arcs $C_{12}$ with the same end points. Since it has also been shown earlier that this value is the same as the value of the original integral (2.1) in the case where the points 1 and 2 are end-points of an extremal arc $E_{12}$ of the field, we may write (see Figure 4)

$$(2.31) \quad I(C_{12})-I(E_{12}) = I(C_{12})-I_{**}(E_{12}) = I(C_{12})-I_{**}(C_{12}).$$

Substituting the values of the integrals I and $I_{**}$, we have

$$(2.32)\quad I(C_{12}) - I(E_{12}) = \int_{x_1}^{x_2} E(x,y,p(x,y),y')dx$$

where the integrand function is the Weierstrass E-function of equation (2.23). This gives the desired theorem:

FUNDAMENTAL SUFFICIENCY THEOREM: Let $E_{12}$ be an extremal arc of a field F such that at each point (x,y) of F the inequality

$$(2.33)\qquad E(x,y,p(x,y),y') \geq 0$$

holds for every admissible set (x,y,y') different from (x,y,p).
Then $I(E_{12})$ is a minimum in F, or, equivalently, the inequality $I(E_{12}) \leq I(C_{12})$ is satisfied for every admissible arc $C_{12}$ in F joining the points 1 and 2. Further, if the equality sign is omitted in (2.33), then $I(E_{12}) < I(C_{12})$ unless $C_{12}$ coincides with $E_{12}$.

The theorem follows directly from (2.32), since the hypothesis (2.33) implies at once that $I(E_{12}) \leq I(C_{12})$. If the E-function vanishes only when y' = p then the equality $I(E_{12}) = I(C_{12})$ can hold only if the equation y' = p(x,y) is satisfied at every point of $C_{12}$ in which case, since the differential equation y' = p(x,y) has one and only one solution through the initial point 1, $C_{12}$ must coincide with $E_{12}$.

A useful corollary may now be proved by considering equation (2.24) in the form

$$(2.34)\quad E(x,y,p,y') = \tfrac{1}{2}(y'-p)^2 f_{y'y'}(x,y,p+\theta(y'-p)).$$

In order to state the corollary more succinctly, we now define a regular problem to be one for which the derivative $f_{y'y'}$ has the same sign for all admissible sets $(x,y,y')$ and for which every set $(x,y,y')$ with the $y'$ element satisfying $y_1' < y' < y_2'$ is admissible whenever the sets $(x,y,y_1')$ and $(x,y,y_2')$ have this property. Equation (2.34) shows that for regular problems the hypothesis $E(x,y,p,y') > 0$ when $y' \neq p$ surely holds provided $f_{y'y'}$ is positive. This gives the corollary:

Corollary: If $E_{12}$ is an extremal arc of a field F for a regular problem with $f_{y'y'} > 0$ then the inequality $I(C_{12}) > I(E_{12})$ holds for every admissible arc $C_{12}$ in F different from $E_{12}$ and joining the points 1 and 2.

The problems of both Chapter 1 and Chapter 3, as well as several other classic problems of the calculus of variations, are regular problems with $f_{y'y'} > 0$.

Before constructing sets of sufficiency conditions out of the four necessary conditions we prove the following useful lemma:

Lemma: Every extremal arc $E_{12}$ having $f_{y'y'} \neq 0$ along it and containing no point conjugate to 1 is interior to a field F of which it itself is an extremal arc.

To prove the lemma, we first observe that the arc $E_{12}$ is a member, for $a = 0$, of a one-parameter family of extremals $y = y(x,a)$ having $y_a(x,0)$ different from zero at each point of $E_{12}$. We now proceed to construct such a family.

When $f_{y'y'} \neq 0$, there is a family of extremals

$$y = y(x,a,b) \tag{2.35}$$

containing $E_{12}$ for a special pair of parametric values $a_o$ and $b_o$. Further, this family may be chosen so that the derivative with respect to x of the determinant

$$\Delta(x,x_1) = \begin{vmatrix} y_a(x,a_o,b_o) & y_b(x,a_o,b_o) \\ y_a(x_1,a_o,b_o) & y_b(x_1,a_o,b_o) \end{vmatrix} \tag{2.36}$$

is different from zero at the end-point 1 of $E_{12}$, a condition that may be expressed by $\Delta'(x_1,x_1) \neq 0$. A positive constant $\varepsilon$ may now be chosen so that $\Delta'(x,x_o)$ is not zero for every pair of values $(x,x_o)$ satisfying the conditions

$$x_1 - \varepsilon \leq x_o < x_1, \quad x_1 - \varepsilon \leq x \leq x_1 + \varepsilon, \tag{2.37}$$

since for small $\varepsilon$ the pairs $(x,x_o)$ above are near the pair $(x_1,x_1)$. For every $x_o$ satisfying the first condition of (2.37) the determinant $\Delta(x,x_o)$ vanishes at $x_o$, and its derivative $\Delta'(x,x_o)$ is different from zero everywhere on the interval expressed by the second condition of (2.37). Clearly $\triangle(x,x_0)$ is different from zero on the interval $x_1 \leq x \leq x_1 + \varepsilon$ and if $x_0$ is selected sufficiently near to $x_1$ then $\triangle(x,x_0)$ will also be different from zero on $x_1 + \varepsilon \leq x \leq x_2$. For when $E_{12}$ contains no point conjugate to 1 the determinant $\triangle(x,x_1)$, whose zeros determine the conjugate points, must be different from zero on $x_1 + \varepsilon \leq x \leq x_2$ so that $\triangle(x,x_0)$ will also have this property when $x_0$ is near to $x_1$.

Now let

$$(2.38) \qquad k = y_b(x_0,a_0,b_0), \qquad m = -y_a(x_0,a_0,b_0).$$

Then a one parameter family with the properties listed at the beginning of the last paragraph is given by

$$(2.39) \qquad y = y(x,a_0 + k\alpha, b_0 + m\alpha) = y(x,\alpha)$$

since for $\alpha = 0$ it gives the arc $E_{12}$ and furthermore its derivative

$$(2.40) \quad y_\alpha(x,0) = y_a(x,a_0,b_0)k + y_b(x,a_0,b_0)m = \triangle(x,x_0)$$

is different from zero on the entire interval $x_1 \leq x \leq x_2$.

This one-parameter family (2.39) simply covers a field F in the neighborhood of the arc $E_{12}$ since we may take $\varepsilon$ so small that the derivative $y_\alpha(x,\alpha)$ remains different from zero whenever x and $\alpha$ satisfy the inequalities $x_1 \le x \le x_2$ and $|\alpha| \le \varepsilon$. Such a field is shown in Figure 5.

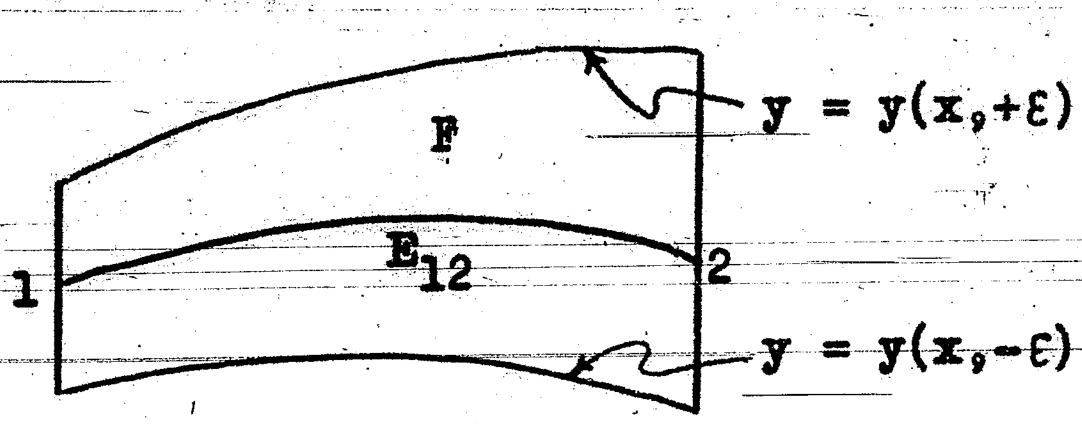


**Figure 5**

On any ordinate of this region F the value $y(x,\alpha)$ varies monotonically as $\alpha$ increases from $-\varepsilon$ to $+\varepsilon$. Thus through each point of F there passes a unique extremal of the family. This means that for each point of F the equation $y = y(x,\alpha)$ has a unique solution $\alpha(x,y)$. It may be shown that this function and its first partial derivatives are continuous in F, from which it follows that the same is true of the slope function $p(x,y) = y'(x,\alpha(x,y))$ of the field. This completes the proof of the lemma.

We are now in a position to prove a theorem involving sufficiency conditions for a weak relative minimum, which is defined as follows: the value $I(E_{12})$ is a weak relative minimum if there is a neighborhood R' of the values $(x,y,y')$ on $E_{12}$ such that the inequality $I(E_{12}) \le I(C_{12})$ is true for all admissible arcs $C_{12}$ whose elements $(x,y,y')$ lie in R'.

Theorem: Let $E_{12}$ be an arc without corners having the properties:

1) it is an extremal,

2) $f_{y'y'} > 0$ at every set of values $(x,y,y')$ on it,

3) it contains no point 3 conjugate to 1.

Then $I(E_{12})$ is surely a weak relative minimum; that is, the inequality $I(E_{12}) < I(C_{12})$ holds for every admissible arc $C_{12}$ distinct from $E_{12}$, joining 1 with 2, and having its elements $(x,y,y')$ all in a sufficiently small neighborhood R' of those elements on $E_{12}$.

We note that the three numbered hypotheses of the theorem are directly related respectively to the three necessary conditions I, III, and IV. To prove the theorem, choose a neighborhood R' of the values $(x,y,y')$ on $E_{12}$ so small that all elements $(x,y,y')$ in R' have their points $(x,y)$ in a field F. The existence of this field which has the arc $E_{12}$ as one of its extremals is guaranteed by the hypotheses of the theorem and the lemma of page 31. The neighborhood R' is also to be chosen so small that for the slope-function $p(x,y)$ of F the elements $(x,y,p+\theta(y'-p))$ having $0 \leq \theta \leq 1$ are all admissible and make $f_{y'y'} \neq 0$. Then the function

$$(2.41) \quad E(x,y,p(x,y),y') = \tfrac{1}{2}(y'-p)^2 f_{y'y'}(x,y,p+\theta(y'-p))$$

is positive for all elements $(x,y,y')$ in $R'$ with $y'$ not equal to $p$. The proof of this theorem is now completed by applying the fundamental sufficiency theorem of page 30 with $R$ replaced by $R'$ in the definition of admissible sets.

A second sufficiency theorem can now be readily proved which involves sufficiency conditions for a strong relative minimum; the value $I(E_{12})$ will be called a strong relative minimum if there is a neighborhood $R'$ of the values $(x,y,y')$ on $E_{12}$ such that $I(E_{12}) \leq I(C_{12})$ for all admissible arcs $C_{12}$ whose elements $(x,y,y')$ have their points $(x,y)$ in a small neighborhood $F$ of those on $E_{12}$.

Theorem: Let $E_{12}$ be an arc without corners having the properties 1), 2) and 3) of the previous theorem and also the additional property

4) at every element $(x,y,y')$ in a neighborhood $R'$ of the elements on $E_{12}$ the condition $E(x,y,y',Y') > 0$ is satisfied for every admissible set $(x,y,Y')$ with $Y' \neq y'$.

Then $I(E_{12})$ is a strong relative minimum; that is, the inequality $I(E_{12}) < I(C_{12})$ holds for every admissible arc $C_{12}$ distinct from $E_{12}$, joining 1 with 2, and having its points $(x,y)$ all in a sufficiently small neighborhood F of those on $E_{12}$.

As in the previous theorem, properties 1),2), and 3) again insure the existence of a field F having $E_{12}$ as one of its extremals. If we take the field so small that all the elements $(x,y,p(x,y))$ belonging to it are in the neighborhood R' of property 4) then according to that hypothesis $E(x,y,p(x,y),y') > 0$ must hold for every element $(x,y,y')$ in F distinct from $(x,y,p(x,y))$, so that again the fundamental sufficiency theorem gives at once the conclusion of this theorem.

The preceding theorems are by no means the only theorems involving sufficiency conditions, but they are general enough to cover a large portion of the problems encountered in the calculus of variations. It will be noted that there are problems for which the latter two theorems do not apply; for example, a problem where the minimizing arc is permitted to have corners. However, since most of the applications of the theory of the calculus of variations are to the class of regular problems defined on page 31 we content ourselves for the moment ~~with one more theorem concerning such problems.~~

For regular problems, it can be shown that a minimizing arc $E_{12}$ can have no corners and that it has a continuous second derivative and is therefore an extremal. Since the derivative $f_{y'y'}$ never vanishes in R the necessary condition III may be strengthened to hold for all admissible elements (x,y,y') having their points (x,y) in a neighborhood of those on $E_{12}$, and it may be further strengthened by excluding the equality sign. If the envelope of the one-parameter family of extremals through the point 1 has a branch projecting backward from the conjugate point 3 then the proof of the necessary condition IV shows that the point 3 can lie neither between 1 and 2 nor at 2 on $E_{12}$ so that this latter augmented version of condition IV is now necessary for a minimum. This gives the following theorem:

Theorem: A minimizing arc $E_{12}$ for a regular problem must be an extremal on which $f_{y'y'}$ is everywhere greater than zero. If the envelope of the one-parameter family of extremals through the point 1 has a branch projecting backward toward 1 from the point 3 conjugate to 1 on $E_{12}$, then the point 3 can be neither between 1 and 2 nor at 2. Furthermore an arc $E_{12}$ with these properties surely furnishes a strong relative minimum.

To complete the discussion of relative minima for regular problems, it remains to consider what happens when the envelope does not have the branch of the above theorem projecting backward toward the point 1. In this case, the earlier proof of the necessary condition can not be applied, but the condition IV is nevertheless necessary, as can be shown by a second proof, which is omitted here, that uses the second derivative I"(0) of the integral (2.4).

Further extensions of the general theory of this chapter are both possible and numerous. For instance, the problem where one or both of the end points of the arc are variable has been given considerable attention in the literature of the calculus of variations. However, an investigation into these further extensions would require more space than is available here. Instead, we now proceed to a relatively thorough investigation of one of the classical problems of the calculus of variations.

# CHAPTER 3

## SURFACES OF REVOLUTION OF MINIMUM AREA

The problem of determining a surface of revolution of minimum area is in many respects the most satisfactory illustration which we have of the principles of the general theory of the calculus of variations. If a wire circle is dipped into a soap solution and withdrawn and a second smaller circle is first made to touch the circular disk of soap film bounded by the first wire circle and then drawn away, the two circles will then be joined by a surface of soap film. In the case where the circles are parallel and have their centers on the same axis perpendicular to their planes, the surface of soap film is a surface of revolution. It is provable by the principles of mechanics, and also intuitively true, that a surface so formed must be one of minimum area. The determination of the shape of this film by analytic means is the subject and purpose of this chapter.

For convenience, let the x-axis be the common axis of the two circles and let the points where the circles intersect the upper half on an xy-plane through their axis be called 1 and 2. If the intersection of the surface with this plane is in the form $y = y(x)$ then from ordinary calculus we know that the area of the surface of revolution is $2\pi$ times the value of the integral

$$(3.1) \qquad I = \int_{x_1}^{x_2} f(y,y')dx$$

where $f(y,y')$ has the value

$$(3.2) \qquad f(y,y') = y\sqrt{1+y'^2}\,.$$

Without loss of generality, we may immediately make the restriction that for all arcs $y = y(x)$ which we shall consider, it is true that

$$(3.3) \qquad y \geq 0,$$

since if an arc does have a portion or portions below the x-axis, the surface generated when this arc is rotated about the x-axis is the same as if the portion(s) below the axis had first been reflected above the axis.

Thus for this chapter an admissible arc $y = y(x)$ will be one which in the interval $[x_1, x_2]$ is continuous and has a tangent which turns continuously except perhaps at a finite number of points and which in addition satisfies condition (3.3).

Our problem now is to determine among all admissible arcs joining two given points 1 and 2 that one which minimizes the integral I of (3.1).

The necessary condition I of Chapter 2 stated that the minimizing arcs must be solutions of the equation

$$(3.4) \qquad f_{y'} = \int_{x_1}^{x} f_y dx + c.$$

For the problem of the minimum surface of revolution, this takes the form

$$(3.5) \qquad \frac{yy'}{\sqrt{1+y'^2}} = \int_{x_1}^{x_2} \sqrt{1+y'^2}\, dx + c = s + c$$

where s is the length of the minimizing arc measured from the point 1 to the point whose abscissa is x. At a point of the arc where $y \neq 0$ equation (3.5) can be written in the form

$$(3.6) \qquad y' = \frac{s+c}{\sqrt{y^2-(s+c)^2}}.$$

Since y and s are both continuous, then so is y' at such a point. But if y' is continuous, then both y and s have continuous derivatives so that equation (3.6) shows that y' must also have a continuous derivative. Thus at all points above the x-axis the minimizing arc has continuous curvature and no corners.

When it is known that along a minimizing arc there is a continuous derivative y" then Euler's equation (2.8) may be written

$$(3.7) \qquad \frac{d}{dx} f_{y'} - f_y = f_{y'y}y' + f_{y'y'}y'' - f_y = 0$$

from which it follows that along a minimizing arc with second derivative y" we have

$$(3.8)\quad \frac{d}{dx}(f-y'f_{y'}) = y'(f_y - f_{y'y}y' - f_{y'y'}y'') = 0$$

and hence

$$(3.9)\quad f - y'f_{y'} = b$$

where b is a constant of integration.

In the problem of the minimum surface of revolution, equation (3.9) becomes

$$(3.10)\quad \frac{y}{\sqrt{1+y'^2}} = b.$$

Solving this last equation for y' gives

$$(3.11)\quad \frac{dy}{\sqrt{\left(\frac{y}{b}\right)^2 - 1}} = dx$$

which upon integration yields the result

$$(3.12)\quad b \log\left[\frac{y}{b} + \sqrt{\left(\frac{y}{b}\right)^2 - 1}\right] = x - a.$$

It follows readily by solving this last equation for y that the extremals of our current problem are the arcs

$$(3.13)\quad y = \frac{b}{2}\left[e^{\frac{x-a}{b}} + e^{-\frac{x-a}{b}}\right] = b \cosh \frac{x-a}{b}.$$

Arcs of this type are called catenaries and their shape is that assumed by a chain attached to the two points 1 and 2.

We see at once that a minimizing arc $y = y(x)$ with corners is impossible since as was indicated earlier such corners would have to be on the x-axis and the parts of the minimizing arc between these corners and above the x-axis would have to be segments of catenaries which have no points in common with the x-axis. That is, we have established the following lemma:

Lemma: If 1 and 2 are two points in the half plane $y \geq 0$ then an admissible arc $y = y(x)$ joining them and generating a surface of revolution of minimum area must be a single arc without corners of one of the catenaries of equation (3.13).

Now we are ready to investigate the number and the character of the catenaries (3.13) which pass through the two given points 1 and 2. The condition that a catenary of the form (3.13) pass through the point 1 is given by the equation

$$(3.14) \qquad y_1 = b \cosh \frac{x_1 - a}{b} .$$

It is now convenient to express the parameters a and b in terms of a new parameter $\alpha$ by means of the equations

$$(3.15) \qquad a = x_1 - y_1\left(\frac{\alpha}{\cosh \alpha}\right), \qquad b = \frac{y_1}{\cosh \alpha} .$$

Now the one-parameter family of catenaries through the point 1 is given by

$$(3.16)\quad y = \frac{y_1}{\cosh\alpha}\cosh\left(\alpha + \frac{x-x_1}{y_1}\cosh\alpha\right) = y(x,\alpha).$$

Denoting differentiation with respect to x and $\alpha$ by primes and subscripts respectively, we obtain

$$(3.17)\quad y' = \sinh\left(\alpha + \frac{x-x_1}{y_1}\cosh\alpha\right),\quad y_1' = \sinh\alpha,$$

$$(3.18)\quad y_\alpha = \frac{y'y_1'}{\cosh\alpha}\left(x - \frac{y}{y'} - x_1 + \frac{y_1}{y_1'}\right).$$

The tangents to the catenary at the points $(x_1,y_1)$ and $(x,y)$ are given in terms of the running coordinates $(X,Y)$ by the equations

$$(3.19)\quad Y - y_1 = y_1'(X - x_1),\quad Y - y = y'(X - x).$$

Eliminating X from these equations, we obtain

$$(3.20)\quad Y = \frac{y'y_1'}{y'-y_1'}\left(x - \frac{y}{y'} - x_1 + \frac{y_1}{y_1'}\right)$$

so that, from equation (3.18), the derivative $y_\alpha$ can be written

$$(3.21)\quad y_\alpha = \frac{y'-y_1'}{\cosh\alpha}\,Y.$$

This last equation enables an interesting construction for the point conjugate to 1 on a catenary. At the point of tangency P of a curve $y = y(x,\alpha)$ with the envelope G of the family, the derivative must vanish, which means that the coordinate Y must be zero. Thus, the tangents to the catenary at the point 1 and at the point P must intersect on the x-axis as shown in Figure 6.

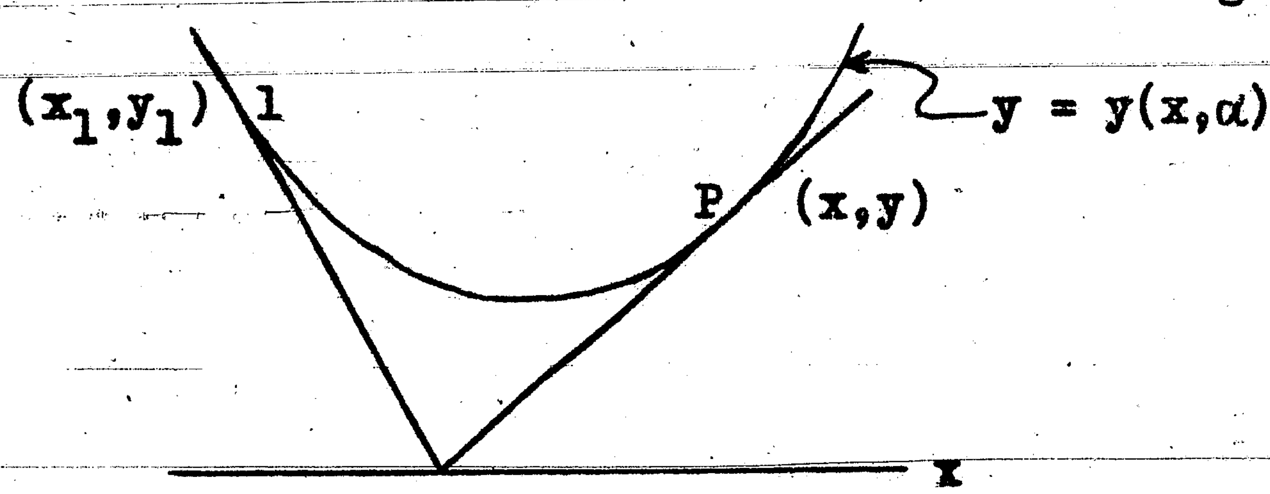


Figure 6

From Figure 6 and the formula (3.21) it is clear that as the point $(x,y)$ moves from the point 1 to the right along the catenary $y = y(x,\alpha)$ the value of $y_\alpha$ is at first positive, and it becomes negative only if $(x,y)$ passes a point P conjugate to 1. In other words, $y_\alpha > 0$ at a value $x > x_1$ implies that there is no conjugate point between 1 and $(x,y)$, while $y_\alpha < 0$ implies that there is such a point. We also note from Figure 6 that for points $(x,y)$ which are conjugate to 1 and have $x > x_1$, then $y_1' < 0$ and $y' > 0$. Using equations (3.16), (3.17), (3.21) and the equation $y_\alpha = 0$, we find

$$y'' = \frac{y}{y_1^2} \cosh^2\alpha > 0, \qquad (3.22)$$

$$(3.23)\qquad y'_\alpha = \frac{y^2 y'_1}{y' y_1^2}\cosh\alpha < 0,$$

$$(3.24)\qquad y_{\alpha\alpha} = \frac{y' y_1'^2}{y_1^2}\left[\left(\frac{y}{y'}\right)^3 - \left(\frac{y_1}{y'_1}\right)^3\right] > 0.$$

We now look at the changes in the ordinates of the catenary $y = y(x,\alpha)$ when $x > x_1$ is kept fixed, and $\alpha$ varies. Using equation (3.16) this ordinate may be represented in the form

$$(3.25)\qquad y(x,\alpha) = \left(\frac{\alpha}{\cosh\alpha} + \frac{x-x_1}{y_1}\right)\left[\frac{y_1\cosh\left[\left(\frac{\alpha}{\cosh\alpha} + \frac{x-x_1}{y_1}\right)\cosh\alpha\right]}{\left(\frac{\alpha}{\cosh\alpha} + \frac{x-x_1}{y_1}\right)\cosh\alpha}\right].$$

As $\alpha$ approaches either plus or minus infinity we may use l'Hospital's rule for evaluating the indeterminate form; namely

$$(3.26)\qquad \lim_{u\to\pm\infty} \frac{u}{\cosh u} = \lim_{u\to\pm\infty} \frac{1}{\sinh u} = 0$$

and so we see that the ordinate expressed by (3.25) approaches plus infinity when $\alpha$ approaches either plus or minus infinity. Furthermore since by equation (3.24) we know that $y_{\alpha\alpha} > 0$ then the derivative $y_\alpha$ changes from negative to positive whenever it vanishes, and hence can vanish only once.

Thus when $x > x_1$ is fixed and $\alpha$ varies from minus infinity to plus infinity the ordinate $y(x,\alpha)$ diminishes from plus infinity to a minimum and then increases to plus infinity again. If we denote this minimum for a fixed x by g(x) then the equation y = g(x) defines a curve which can be shown to be the envelope G of the family of catenaries through the point 1, as shown in Figure 7.

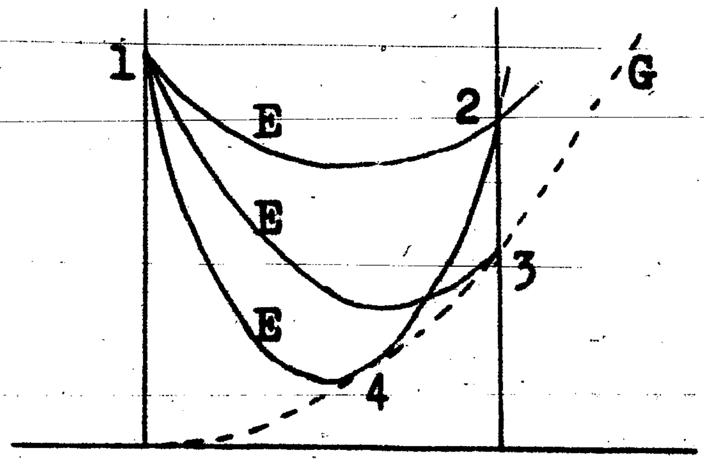


Figure 7.

From the above argument it follows that through a point 2 above the envelope G there pass two catenaries on which the derivative $y_\alpha$ has opposite signs at 2, which as previously noted means that one of the catenaries has a point conjugate to 1 between 1 and 2 while the other catenary has none. Hence we have established the following theorem:

Theorem: A point 2 above the envelope G (as in Figure 7) is joined to the point 1 by two catenaries of the family $y = b \cosh\left[(x-a)/b\right]$. On one of these is a point 4 conjugate to 1,and on the other there is no such conjugate point. A point 3 on the envelope G is joined to 1 by a single catenary on which 3 is conjugate to 1. A point below G is joined to 1 by no catenary of the family.

We are now ready to use some of the results of Chapter 2 to establish for the one-parameter family of catenaries through the point 1 another theorem which Figure 8 below will help to make clear.

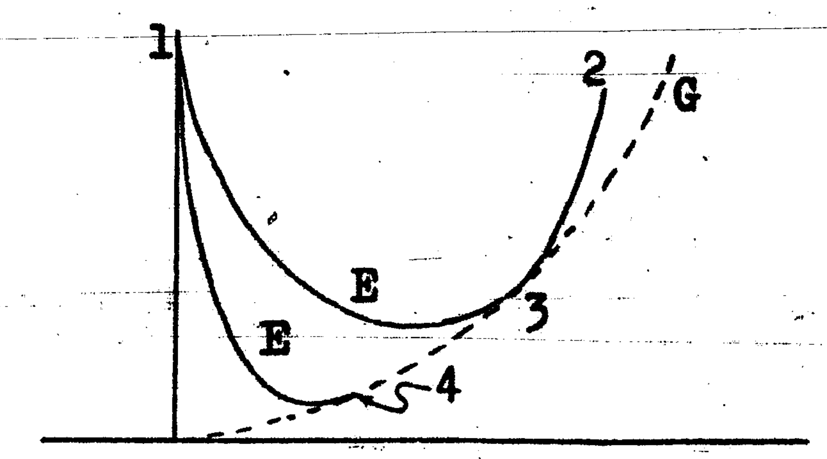


Figure 8

Theorem: If two catenaries $E_{14}$ and $E_{13}$ of the family through the point 1 touch the envelope G at the points 4 and 3, as in Figure 8, then the areas of the surfaces of revolution generated by the arcs $E_{14}$ + $G_{43}$ and $E_{12}$ are equal. This may be expressed by the equation

$$(3.27) \qquad I(E_{14}) + I(G_{43}) = I(E_{13}).$$

The proof is simple with the help of formula (2.21) of the preceding chapter where the curve C of that formula is now the fixed point 1 and the curve D is the envelope G. That is, the formula (2.21) may now be written

$$(3.28) \qquad I(E_{13}) - I(E_{14}) = I^{**}(G_{43}).$$

Since at every point of the arc G we have $dy = p\,dx$, where p is the slope of the catenary through that point, we have

$$(3.29) \qquad I^{**}(G_{43}) = \int_{x_4}^{x_3} \{f(y,p)dx + (dy - pdx)f_{y'}(y,p)\}$$

$$= \int_{x_4}^{x_3} f(y,p)dx = I(G_{43}),$$

and the theorem is established.

Thus the necessary condition IV of Chapter 2 may for this problem be stated as follows:

Theorem: If a catenary arc $E_{12}$ is to generate a surface of revolution of minimum area then the contact point 3 (shown in Figure 8) of the catenary with the envelope G of the one-parameter family of extremals through the point 1 must not lie on $E_{12}$.

We see this is true because the arc $E_{14} + G_{43} + E_{32}$ of Figure 8 generates the same surface area as $E_{12}$, and the arc $G_{43}$ can always be replaced by an arc $C_{43}$ which will generate a smaller area since $G_{43}$ is not an arc of a catenary of the family (3.13). That $G_{43}$ can never be such a catenary is clear when we note that at each point of it the equation (3.10) defines a value b which is the same as the one belonging to the catenary tangent to G at that point, but these values b vary from point to point on G, as shown by the second equation of (3.15) while on the catenaries they must be constant.

We now might want to conclude that the surface of revolution generated by a catenary joining 1 and 2 and having on it no conjugate point is smaller than that furnished by every other arc $y = y(x)$ joining 1 and 2. This conclusion need not be valid, for while such a catenary minimizes I with respect to other curves lying sufficiently near to it, there may be in some cases other curves not so near that give I a smaller value. That is, we have so far shown that this catenary with no conjugate point is a weak relative minimum, as explained in the theorem of the preceding chapter on page 35.

In order to obtain a more complete sufficiency proof, we now construct a field F of extremals for this problem. Suppose $E_{12}$ is a catenary arc having on it no point conjugate to 1 and with the equation

$$(3.30) \qquad y = b_0 \cosh \frac{x-a_0}{b_0} .$$

We take a point '0' on the catenary so near to 1 that the point 3 conjugate to 0 is to the right of 2, as in Figure 9.

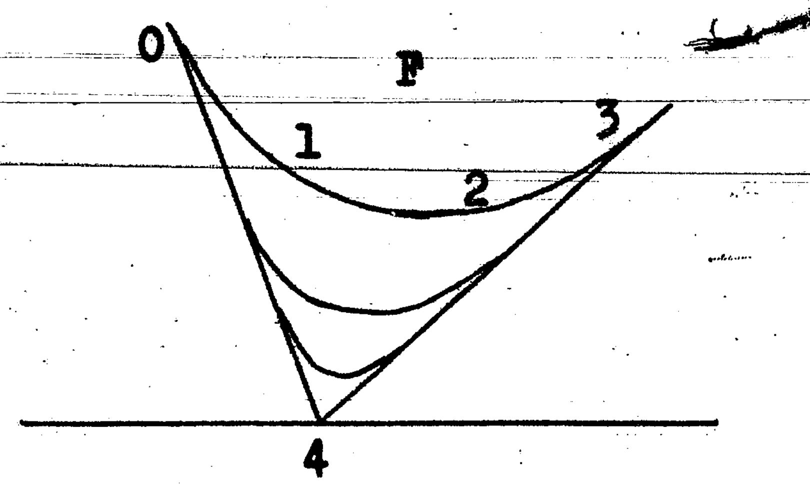


Figure 9

As before, the tangents to the catenary (3.30) meet on the x-axis at a point 4. We now make use of the transformation

$$(3.31) \qquad x - x_4 = \frac{b_o}{b}(X - x_4), \qquad y = \frac{b_o}{b} Y,$$

which stretches the plane along the radii through the point 4 in such a way that every point (x,y) is replaced by a point (X,Y). Using (3.31) we see that the points (x,y) on the catenary (3.30) are transformed into the points (X,Y) which satisfy the equation

$$(3.32) \qquad y = b \cosh \frac{1}{b}\left[x - x_4 + \frac{b}{b_o}(x_4 - a_o)\right] = y(x,b).$$

This is a catenary of the family $y = b \cosh \frac{x-a}{b}$ with the parameter $a = x_4 - b(x_4 - a_o)/b_o$. Now if we treat b as a variable we obtain a one-parameter family of catenary arcs containing the original catenary $E_{12}$ for the special value $b = b_o$. Furthermore, each of these arcs is tangent to the two tangent lines joining the point 4 to the points 0 and 3. It is clear that through each point (x,y) of the V-shaped region F bounded by the two radii joining the point 4 to the points 0 and 3 there passes a unique extremal of the family (3.22) which justifies calling this region F a field of extremals.

This means that for each point (x,y) in F the equation $y = y(x,b)$ has a unique solution $b(x,y)$ which can be shown to be continuous within and on the boundary of the V-shaped field F except at the point 4, and to have continuous derivatives in the interior of the field. The same properties hold for the slope-function $p(x,y)$ of the field; this slope-function can be expressed in the form

$$p(x,y) = y'(x,b(x,y)). \tag{3.33}$$

The extremal arc $E_{12}$ given by equation (3.30) around which the field F has been constructed generates a surface of revolution of smaller area than that generated by every other arc $C_{12}$ in the field F joining the points 1 and 2 and having equations

$$x = x(t), \quad y = y(t). \qquad (t_1 \leq t \leq t_2) \tag{3.34}$$

This is true when we recall from Chapter 2 that on an extremal (in this case a catenary) the values I and I** are equal, and also that I** is invariant; in other words, we have

$$(3.35) \qquad I(E_{12}) = I^{**}(E_{12}) = I^{**}(C_{12}),$$

and hence

$$(3.36) \qquad I(C_{12}) - I(E_{12}) = I(C_{12}) - I^{**}(C_{12}) \geq 0,$$

where the equality sign holds only if $C_{12}$ coincides with $E_{12}$.

These results can be summarized in the following theorem:

Theorem: An admissible arc $y = y(x)$ $(x_1 \leq x \leq x_2)$ in the half plane $y \geq 0$, joining two given points 1 and 2 and generating a surface of revolution of minimum area when rotated about the x-axis, must have the properties:

1) It is a single arc without corners of one of the catenaries $y = b \cosh \frac{x-a}{b}$.
2) It has on it no point of contact with the envelope G of the one-parameter family of these catenaries through the point 1.

If $E_{12}$ is an arc having these properties, and if F is one of the V-shaped regions shown in Figure 9 containing $E_{12}$ in its interior and bounded by two tangents to the catenary E which meet on the x-axis, then the area of the surface of revolution generated by $E_{12}$ is smaller than the area generated by every other arc $C_{12}$ of the type (3.34) in the region F and joining the points 1 and 2.

It is at once evident that the V-shaped field F in which the catenary $E_{12}$ furnishes a minimum in the above theorem is not unique, for the two points 0 and 3 can be chosen at slightly different locations and the field will retain its properties. In fact, though the proof will not be given here, the following theorem can now be established:

Theorem: If $E_{12}$ is a catenary of the family (3.13) having on it no point conjugate to 1 except possibly at 2, then the surface of revolution which it generates is smaller in area than that generated by every other arc $C_{12}$ with equations of the type (3.34) joining 1 with 2 and, except possibly at 2, lying entirely above the envelope G of the one-parameter family of catenaries through the point 1.

We still need to determine what arc will provide a minimum surface of revolution when the point 2 is below the envelope G; also, we do not yet know that the catenary joining 1 with 2 will provide the minimum surface of revolution when we include for comparison all arcs $C_{12}$ joining 1 with 2 regardless of whether or not they are within the V-shaped region. We also need to know the minimizing arc when the point 2 is directly beneath the point 1; this latter case will be investigated next.

Consider a segment $E_{12}$ of the vertical line through the point 1, and also an arc $C_{13}$ with length $\ell$ equal to that of $E_{12}$, as in Figure 10.

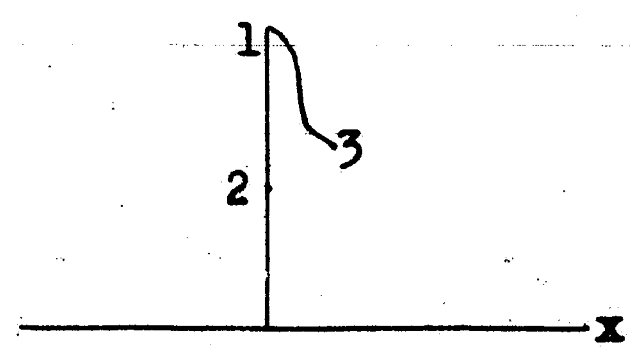


Figure 10

Let the points at a distance s from 1 on $E_{12}$ and $C_{13}$ respectively have the ordinates y and Y. If $C_{13}$ has a single point Y distinct from the corresponding point y of $E_{12}$, we must have $Y > y$; equality holds therefore only if $C_{13}$ coincides with $E_{12}$. Thus, the difference of the areas of the surfaces of revolution generated by the two arcs is $2\pi$ times the difference

$$(3.37) \quad I(C_{13}) - I(E_{12}) = \int_0^\ell Y\,ds - \int_0^\ell y\,ds =$$

$$= \int_0^l (Y - y)ds \geq 0.$$

This may be stated in a theorem as follows:

Theorem: If a vertical straight line has its upper end point 1 in common with an arc $C_{13}$ of the same length, as in Figure 10, then the area of the surface of revolution generated by rotating $E_{12}$ about the x-axis is always less than that generated by $C_{13}$ unless $C_{13}$ is coincident with $E_{12}$.

Thus we conclude that if the points 1 and 2 are in the same vertical line then the straight line joining them generates a smaller surface of revolution than that generated by every other arc joining the same two points.

This last theorem leads directly to another result for the case when the points 1 and 2 are not in the same vertical line. Let 3 and 4 be points on the x-axis directly beneath 1 and 2 respectively, as in Figure 11.

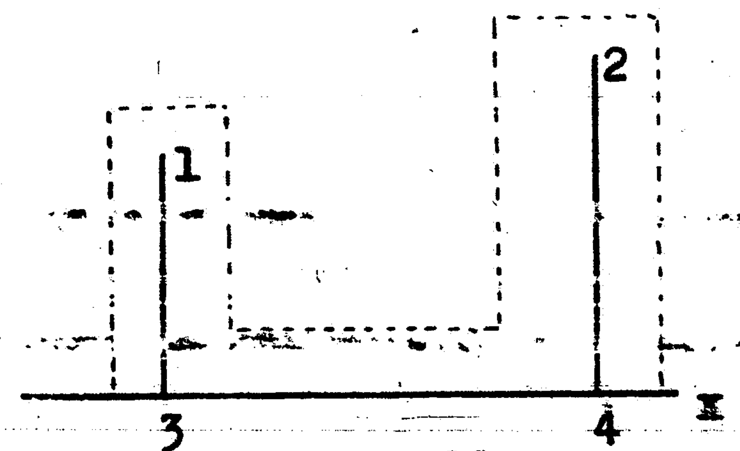


Figure 11

The theorem most recently proved shows that if an arc $C_{12}$ in the half plane $y \geq 0$ has length greater than $y_1 + y_2$, then the area of the surface of revolution generated by $C_{12}$ is greater than or equal to that generated by the broken line $L_{1342}$. Thus if we take a neighborhood of $L_{1342}$ such as is represented by the dotted line and the x-axis in Figure 11 so close to $L_{1342}$ that any arc in it joining 1 and 2 is necessarily longer than $y_1 + y_2$, then in this neighborhood the line $L_{1342}$ is a minimizing arc for the problem of determining a curve joining 1 and 2 and generating a surface of revolution of minimum area. This is the so-called Goldschmidt discontinuous solution.

We have shown that the Goldschmidt discontinuous solution $L_{1342}$ and the catenary arc $E_{12}$ without contact with the envelope G both furnish minima with respect to curves lying near them. We now ask whether or not one of them furnishes an absolute minimum when compared with every arc $C_{12}$ joining the points 1 and 2 and lying in the half plane $y \geq 0$.

To answer this question, let $C_{12}$ be any arc joining 1 and 2 and distinct from $L_{1342}$ and intersecting the envelope G for the first time at a point 5, as shown in Figure 12. When the point 7 on G is near enough to the point 3, the length of $E_{17} + G_{75}$ is greater than $y_1 + y_5$

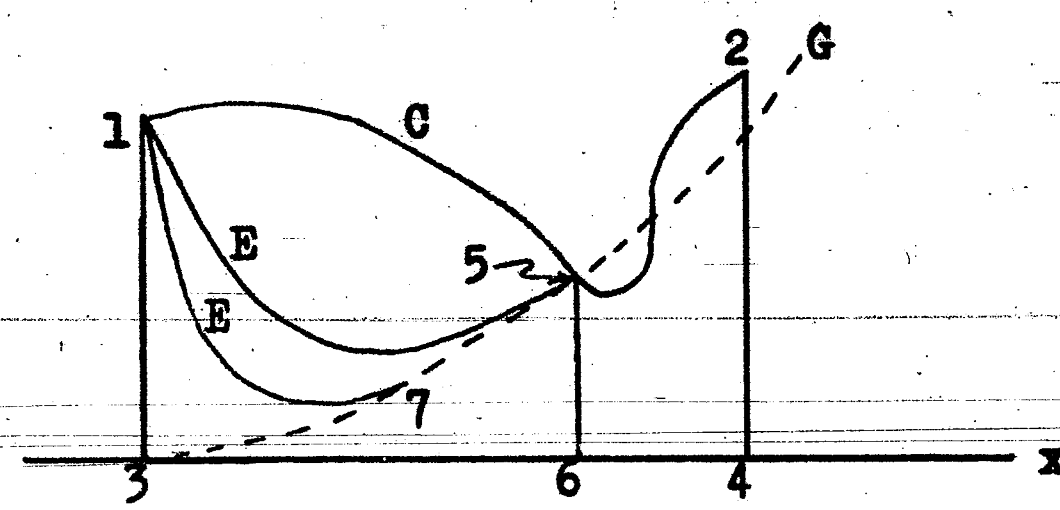


Figure 12

and hence from our previous theorem we have

$$(3.38) \quad I(C_{15}) \geq I(E_{15}) = I(E_{17} + G_{75}) \geq I(L_{1365}),$$

where equality holds only when 5 is at 3 and $C_{15}$ therefore coincides with $L_{13}$. Furthermore, if we allow 5 to move along $C_{12}$ toward the point 2, the difference

$$(3.39) \quad I(C_{15}) - I(L_{1365}) = \int_0^{s_5} y\, ds - \frac{1}{2}(y_1^2 + y_5^2),$$

where s is arc length measured from 1 toward 5 on $C_{15}$, has its derivative

$$(3.40) \qquad y_5\left(1 - \frac{dy_5}{ds_5}\right)$$

with respect to $s_5$ always positive or zero, since the maximum absolute value of the ratio $dy_5/ds_5$ is unity. Hence the difference (3.39) is never decreasing as 5 moves toward 2 on $C_{12}$ and since it starts with a positive or zero value when the point 5 is on G, as shown by (3.38), then when 5 reaches 2 we have

(3.41) $$I(C_{12}) - I(L_{1342}) \geq 0.$$

It can be shown that equality holds only when $C_{12}$ and $L_{1342}$ are coincident.

Clearly when the point 2 is on or below the envelope G, every curve joining 1 and 2 must intersect G. That is, we have established the following fact:

When there are fewer than two catenaries joining the points 1 and 2 the Goldschmidt discontinuous solution always furnishes an absolute minimum.

Thus we have shown that when the point 2 is above G the catenary $E_{12}$ having no contact with G, and the Goldschmidt discontinuous solution $L_{1342}$, both furnish minima in sufficiently small neighborhoods, and the one which generates a smaller area than the other surely provides an absolute minimum. That is, in the case when $I(E_{12}) < I(L_{1342})$ then $I(E_{12})$ is smaller than all the values of I on arcs $C_{12}$ above the envelope G by a previous theorem and also is smaller than the values $I(C_{12})$ for curves meeting G since for such curves $I(E_{12}) < I(L_{1342}) \leq I(C_{12})$. A similar argument holds for $I(L_{1342}) < I(E_{12})$, and when the two integrals are equal, then each of the arcs $E_{12}$ and $L_{1342}$ generates a smaller surface of revolution than other arcs with the same end points.

In order to have some criterion for distinguishing which of the values $I(E_{12})$ and $I(L_{1342})$ is the smaller, we present a geometric argument. The difference of these values is given by

$$(3.42)\quad I(E_{12}) - I(L_{1342}) = \int_0^{s_5} y\,ds - \frac{1}{2}(y_1^2 + y_2^2).$$

As the point 2 moves from 1 along a fixed catenary E the derivative $y_2(1 - \frac{dy_2}{ds_2})$ of this difference is always positive since the tangent to the catenary is never vertical. Further, the absolute value of the ratio $dy_2/ds_2$ is never as great as unity. Since $I(E_{12}) = 0$ when the point 2 is at 1, the difference $I(E_{12}) - I(L_{1342})$ is then negative and when 2 is on G the difference is positive since $I(L_{1342})$ is smaller than the value of I on any other arc intersecting G. Thus $I(E_{12}) = I(L_{1342})$ for only one position of the point 2 on the catenary between these extremes.

In terms of the parameter

$$(3.43)\qquad u = \alpha + (x - x_1)\frac{\cosh \alpha}{y_1}$$

the equations of the family (3.16) of catenaries through the point 1 may be written

$$(3.44)\quad x = x_1 + \frac{y_1}{\cosh \alpha}(u - \alpha), \quad y = y_1 \frac{\cosh u}{\cosh \alpha}$$

and the values of the two integrals in terms of the parameter u of the point 2 are found to be

$$(3.45) \qquad I(E_{12}) = \int_{\alpha}^{u} y\sqrt{x_u^2 + y_u^2}\, du = \frac{1}{2}\left(\frac{y_1}{\cosh\alpha}\right)^2 \Big[u + (\sinh u)(\cosh u)\Big]\Big|_{\alpha}^{u},$$

$$(3.46) \qquad I(L_{1342}) = \frac{1}{2}(y_1^2 + y_2^2) = \frac{1}{2}\left(\frac{y_1}{\cosh\alpha}\right)^2 \Big[\cosh^2\alpha + \cosh^2 u\Big].$$

This means that $I(E_{12}) = I(L_{1342})$ when

$$(3.47) \qquad u + \sinh u \cosh u - \cosh^2 u = \alpha + \sinh\alpha \cosh\alpha + \cosh^2\alpha .$$

This last equation and the equation (3.43) define the locus H of the points where $I(E_{12}) = I(L_{1342})$. It can be shown that the shape of the curve H is similar to that of G, as indicated in Figure 13.

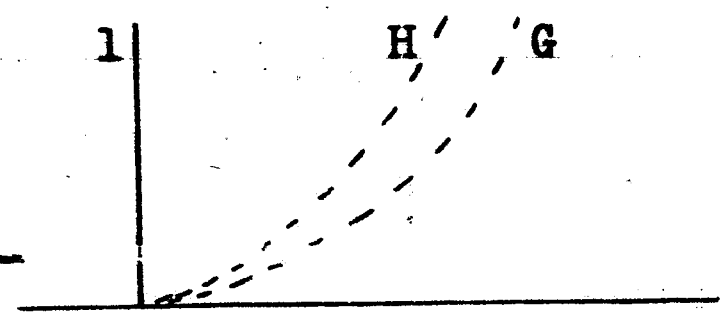


Figure 13

We now conclude our discussion of surfaces of revolution of minimum area with the following summary:

SUMMARY: For a point 2 above the curve H in Figure 13 the Goldschmidt discontinuous solution $L_{1342}$ joining 1 to 2 generates a minimum surface of revolution relative to those generated by other arcs and lying in a sufficiently small neighborhood of $L_{1342}$; but the smallest surface of all, the absolute minimum, is in this case furnished by the unique catenary arc $E_{12}$ joining 1 with 2 and having on it no point of contact with the envelope G.

When 2 is on H the surfaces generated by $L_{1342}$ and $E_{12}$ are equal in area and smaller than those generated by other arcs joining these two points 1 and 2.

When 2 is between H and G, the catenary arc $E_{12}$ furnishes a relative minimum and the Goldschmidt discontinuous solution furnishes the absolute minimum.

When 2 is on or below G the Goldschmidt solution is the only minimizing arc joining 1 and 2, and it furnishes an absolute minimum.

## BIBLIOGRAPHY


1. G.A.Bliss, Calculus of Variations, The Open Court Publishing Company, (1925).
2. G.A.Bliss, Lectures on the Calculus of Variations, University of Chicago Press, (1946).
3. R.S.Burington and C.C.Torrance, Higher Mathematics. McGraw-Hill Book Company, Inc.,(1939).
4. R. Courant,(translated by J.E.McShane),Differential and Integral Calculus,Vol.II,Blackie & Son,(1936).
5. G.Joos,Theoretical Physics, Chapter IV, G.E.Stechert & Co., (1934).
6. R.Weinstock, Calculus of Variations, McGraw-Hill Book Company Inc., (1952).
7. __________,Contributions to the Calculus of Variations, The University of Chicago Press, (1937).

# VITA

Kenneth C. Bouchelle, son of Mr. and Mrs. Henry Wirt Bouchelle, Sr. was born in Elkton, Maryland on August 19, 1927. He completed elementary school in 1940 and high school in 1944, both in Elkton. He received the Bachelor of Arts Degree, with a major in mathematics and a minor in physics, from Western Maryland College, Westminster, Maryland in 1948. The next six years he taught in Calvert Junior and Senior High School, Cecil County, Maryland, until he was drafted into the United States Army in September, 1954. He served for two years in the Signal Corps at the Army Electronic Proving Grounds, Fort Huachuca, Arizona, assigned to Combat Development Department as a mathematician in the Scientific and Professional Program. Upon discharge from the Army, he taught for one year at Lindley Junior High School, Greensboro, North Carolina. In September of 1957, he came to Lehigh University as a Graduate Assistant in the Mathematics Department. He was recently appointed Instructor, a position he is currently holding.