CALCULUS OF VARIATIONS

(with supplementary notes and exercises)

R. Courant

1945 - 1946

Revised and amended by J. Moser, 1962

Supplementary notes by Martin Kruskal and Hanan Rubin

Courant Institute of Mathematical Sciences
New York University

The Courant Institute publishes a number of sets of lecture notes. A list of titles currently available will be sent upon request.

Courant Institute of Mathematical Sciences
251 Mercer Street, New York, New York 10012

Contents

INTRODUCTION

The Calculus of Variations has assumed an increasingly important role in modern developments in analysis, geometry, and physics. Originating as a study of certain maximum and minimum problems not treatable by the methods of elementary calculus, variational calculus in its present form provides powerful methods for the treatment of differential equations, the theory of invariants, existence theorems in geometric function theory, variational principles in mechanics. Also important are the applications to boundary value problems in partial differential equations and in the numerical calculation of many types of problems which can be stated in variational form. No literature representing these diverging viewpoints is to be found among standard texts on calculus of variations, and in this course an attempt will be made to do justice to this variety of problems.

The subject matter with which calculus of variations is concerned is a class of extremum (i.e. maximum or minimum) problems which can be considered an extension of the familiar class of extremum problems dealt with by elementary differential calculus. In the elementary problems one seeks extremal values of a function of one or more (but in any case a finite number) real variables. In the more general problems considered by calculus of variations, the functions to be extremized, sometimes called functionals, have functions as independent variables. The area $A(f)$ below a curve $y = f(x)$, for example, is a functional since its value depends upon a whole function f. (It is possible to treat a functional as a function of an enumerable set of Fourier coefficients, but this attack usually leads to almost insuperable difficulties.)

One of the earliest problems of this type was the isoperimetric problem considered by the ancient Greeks. This is to find, among all closed curves of a given length, the one which encloses the maximum area. It is intuitively evident that the solution is a circle, but this fact has been

satisfactorily proved only in recent times, and the corresponding theorem concerning the sphere is even more difficult.

The modern development of calculus of variations, however, began in 1696 with the formulation of the brachistochrone problem by John Bernoulli. This problem is to find, among all curves connecting two given points, that one which has the property that a particle sliding along it under the action of gravity alone falls from one point to the other in the least time. This problem excited great interest among the mathematicians of that day, and gave rise to a train of research which is still continuing.

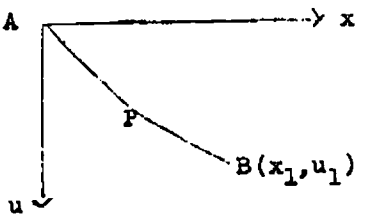
Subsequent developments in classical calculus of variations were the derivation of necessary conditions for an extremum (corresponding to the conditions grad $f(x_1,x_2,\ldots x_n)=0$ for a function f of n variables) by Euler and more rigorously by Lagrange; and the development of sufficient conditions (corresponding to the consideration of the quadratic form in second derivatives of $f(x_1,x_2,\ldots x_n)$ at a stationary point) by Hamilton, Jacobi, and others; culminating in the completion of this theory by Weierstrass.

The broader aspects of physical variational principles were first set forth by Maupertius, and were given a firmer foundation by the work of Euler, Hamilton, Jacobi and Gauss.

We will now consider the mathematical formulation of several problems:

a) The Brachistochrone

A particle P slides under the influence of gravity along a curve connecting A and B. The velocity v at any point is given by



$$v = \frac{ds}{dt} = \sqrt{2gu}\ ,$$

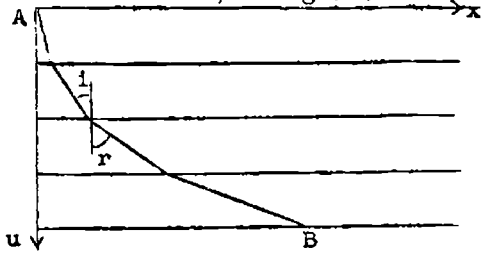
so that the time of fall T is

$$T = \int_{(A)}^{(B)} dt = \int_{(A)}^{(B)} \frac{ds}{\sqrt{2gu}}$$

Suppose the curve is given by $u = f(x)$, where $f(0) = 0$, $f(x_1) = u_1$, and $f(x)$ is assumed to be piecewise differentiable. Then $ds = \sqrt{1 + u'^2}\,dx$. Hence the solution of the problem can be obtained by finding the function $u = f(x)$ which minimizes the integral T (a functional)

$$T = \frac{1}{\sqrt{2g}} \int_0^{x_1} \sqrt{\frac{1+u'^2}{u}}\, dx.$$

Bernoulli obtained the solution to this problem using an entirely different line of reasoning. He approximated the path $u = f(x)$ by a series of line segments dividing the distance fallen into equal parts, the particle velocity being assumed constant throughout each segment. It is an elementary exercise in calculus to derive Snell's law of refraction

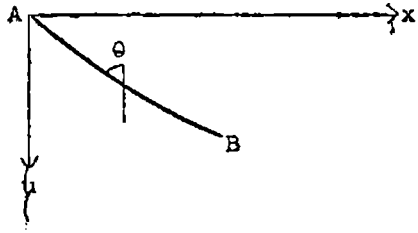


$$\frac{\sin i}{v_i} = \frac{\sin r}{v_r} = \text{constant}$$

as the condition for the path of minimum time across a discontinuity. Taking the limit as the segments are made smaller, Bernoulli argued that the curve would be given by

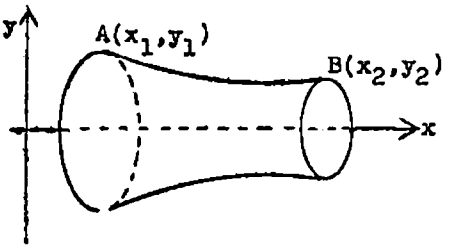
$$\frac{\sin \theta}{\sqrt{2gu}} = \text{constant}$$

which is indeed the correct answer, characterizing the cycloid. Of course, Bernoulli's solution is only an indication rather than a proof, since he neither justified the limiting process, nor showed that his solution was unique.



b) Minimum Area of a Surface of Revolution

Consider the surface generated by revolving the curve AB about the x-axis. If the equation of this curve is $y = f(x)$, where $f(x_1) = y_2$, and f is piecewise differentiable, then the area of the surface is given by the functional



$$I(f) = 2\pi \int_{x_1}^{x_2} f \sqrt{1 + f'^2}\, dx.$$

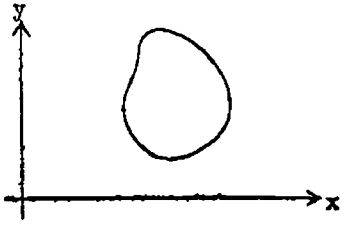
The problem, then, is to determine f so that I(f) is a minimum. The problem can be "solved" physically by stretching a soap film between the two circles (made of wire) at A and B. Surface tension in the film will minimize the area.

c) Geodesics

The curve of shortest length connecting two points in a plane is a straight line. This need not be taken as an axiom, but can be proved. Similarly, on the surface of a sphere, the curve of least length is the arc of a great circle. In general, on any surface, the curves of least length connecting pairs of points are called geodesics and their determination leads to problems in calculus of variations. In case the surface is developable (i.e. one which can be deformed into a plane without altering length -- e.g. a cone) the geodesics are given by the corresponding straight lines in the plane.

d) The Isoperimetric Problem

Consider a plane closed curve given in parametric form by $x = f(t)$, $y = g(t)$ where f and g have piecewise continuous derivatives, are of period 2π in t,



and the curve has a given length L,

$$L = \int_0^{2\pi} \sqrt{\dot{x}^2 + \dot{y}^2}\, dt \ .$$

The problem is to find, among all f and g satisfying these conditions, the pair that maximizes the area A

$$A = \frac{1}{2} \int_0^{2\pi} (x\dot{y} - y\dot{x})\, dt \ .$$

This problem is different from the foregoing three problems in that we seek to extremize a functional A of two variables (the functions f and g) subject to a prescribed condition, L = constant. All such problems in calculus of variations are called isoperimetric problems; the analogy with the corresponding elementary problem of extremizing a function F(x,y) of two real variables subject to an auxiliary condition, say G(x,y) = 0, is evident.

We will now prove that the circle, i.e.

$$f(t) = a_1 \sin t + a_2 \cos t$$

$$g(t) = a_2 \sin t - a_1 \cos t$$

maximizes A, subject to L = constant. Consider the expression

$$I = \frac{L^2}{4\pi} - A \ .$$

For the circle, I = 0. We then wish to show that I > 0 for all other curves. Let $t = 2\pi \frac{s}{L}$, where s is arc length. Then

$$\dot{x}^2 + \dot{y}^2 = \dot{s}^2 = L^2/4\pi^2 \ ,$$

and

$$\frac{L^2}{4\pi} = \frac{1}{2} \int_0^{2\pi} (\dot{x}^2 + \dot{y}^2)\, dt \ .$$

Hence

$$I = \frac{1}{2} \int_0^{2\pi} [(\dot{x}^2 + \dot{y}^2) - (x\dot{y} - y\dot{x})]\, dt$$

$$= \frac{1}{4} \int_0^{2\pi} [(\dot{x} + y)^2 + (\dot{y} - x)^2 + (\dot{x}^2 - x^2) + (\dot{y}^2 - y^2)]\, dt \ .$$

Since $(\dot{x} + y)^2 + (\dot{y} - x)^2 \geq 0$, we will consider

$$I_1 = \int_0^{2\pi} (\dot{x}^2 - x^2)dt + \int_0^{2\pi} (\dot{y}^2 - y^2)\, dt.$$

Under the conditions imposed, we may expand x and y in Fourier series

$$x \sim \sum_{n=0}^{\infty} (a_n \cos nt + b_n \sin nt)$$

$$y \sim \sum_{n=0}^{\infty} (a_n' \cos nt + b_n' \sin nt) \quad .$$

By taking the center of gravity of the curve as the origin (i.e. translating the axes $x' = x + x_0$, $y' = y + y_0$) so that

$$\int_0^{2\pi} x'\, dt = \int_0^{2\pi} y'\, dt = 0$$

we have

$$a_0 = a_0' = 0.$$

Then, dropping the primes,

$$\int_0^{2\pi} (\dot{x}^2 - x^2)dt = \pi \sum_{n=1}^{\infty} [n^2(a_n^2 + b_n^2) - (a_n^2 + b_n^2)]$$

which is positive unless $a_n = b_n = 0$ for $n > 1$, i.e. unless

$$x = a_1 \cos t + b_1 \sin t.$$

Similarly

$$\int_0^{2\pi} (\dot{y}^2 - y^2)\, dt > 0$$

unless $y = a_1' \cos t + b_1' \sin t.$

But in the case that x and y are both of this form, we have

$$I = \frac{1}{4} \int_0^{2\pi} [(\dot{x} + y)^2 + (\dot{y} - x)^2]\, dt$$

$$= \frac{\pi}{4} [(a_1 + a_1' + b_1' - b_1)^2 + (a_1 - a_1' + b_1 + b_1')^2]$$

which is zero only if $a_1 = b_1'$, $a_1' = b_1$. Hence $I > 0$ unless

$$x = a_1 \cos t + b_1 \sin t$$

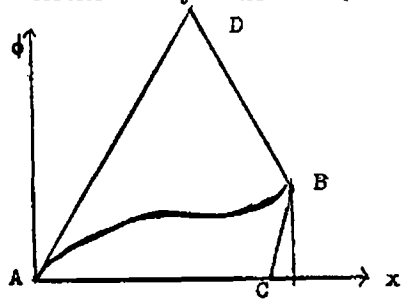
$$y = b_1 \cos t - a_1 \sin t$$

which are the parametric equations of a circle.

In all the problems that we have considered so far we have tacitly assumed that they make sense, i.e. that a solution exists. However, this is by no means always the case. For example, consider the integral

$$I(\phi) = \int_0^1 \frac{dx}{1 + [\phi'(x)]^2}$$

$$\phi(0) = 0, \quad \phi(1) = 1$$



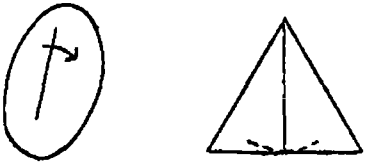
where ϕ is subject to the condition that it pass through the two points A and B, and let us try to find a continuous- and piecewise differentiable function ϕ which either maximizes or minimizes $I(\phi)$. By inspection we see that

$$0 < I(\phi) < 1$$

since the integrand is positive and always less than one. However from the figure it is easily seen that by picking point C very close to $x = 1$ we can make I take values as close to unity as we please for the curve A C B, and by taking the ordinate of D large enough we can make I as small as we please for the curve A D B. Since there is no admissible curve ϕ which will make $I(\phi)$ take on the values 0 or 1, there is no solution to either the minimum or maximum problem.

Let us now consider a problem in which the existence of a solution depends on the class of admissible curves. We look for a closed curve of minimum area, within which a line of given length can turn through a complete revolution. If we limit ourselves to convex curves, the solution is given by the equilateral triangle having the given line as

altitude. However, if we remove the restriction of convexity, it can be shown that the area can be made as small as we please, and since no solution exists having zero area, there is no solution to the problem.



We therefore see that in order for a variational problem to be solvable, some care must be taken in the choice of admissible functions.

Problems

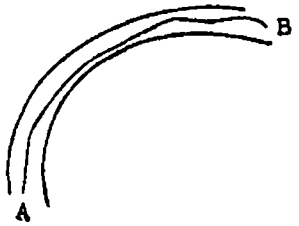
1) Verify that the property

$$\frac{\sin \theta}{\sqrt{u}} = \text{constant}$$

characterizes the cycloid.

2) Characterize the larger great circle arc between two points on a sphere as the solution of a minimax problem. Generalize to a great circle arc which winds around n times.

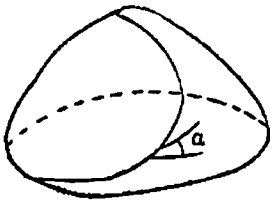
3) Find the shortest path between two points A and B lying between two convex curves as boundaries. This is an example of a general type of problem in which auxiliary conditions in the form of inequalities are imposed (in contradistinction to isoperimetric problems where auxiliary <u>equations</u> must be satisfied).



4) Find the path which requires the least time in ascending a rotationally symmetric mountain using a car with its velocity a given function f of the angle of inclination α, such that

$f(0) = v_0$, $f(\frac{\pi}{2}) = 0$,

$f(\alpha)$ and $f'(\alpha)$ monotonic.



5) Show that any admissible function $\phi(x)$ can be approximated by admissible functions $\phi_1(x)$ and $\phi_2(x)$ such that $I(\phi)$ can be made as small as we please and $I(\phi)_2$ as close to unity as we please, where

$$I(\phi) = \int_0^1 \frac{dx}{1 + [\phi'(x)]^2} \ ,$$

I. FORMALISM OF CALCULUS OF VARIATIONS

1. The Euler Equation. The simplest type of problem in the calculus of variations is to extremize a functional $I(\phi)$ of only one independent variable, the function $\phi(x)$. In practice, the functional is usually an integral of a given form, and we will henceforth restrict our discussion to functionals which are integrals. In general terms, then, the simplest type of problem is to extremize

$$I(\phi) = \int_a^b F(x,\phi(x),\phi'(x))dx \tag{1}$$

where F is a given function which we will assume has continuous first partial derivatives and piecewise continuous second partial derivatives. The function $\phi(x)$ will be restricted to the class of admissible functions satisfying the conditions

$$\begin{aligned} &\phi(a) = A, \quad \phi(b) = B \\ &\phi(x) \text{ continuous} \\ &\phi'(x) \text{ piecewise continuous} \end{aligned} \tag{2}$$

The brachistochrone is an example of this type of problem.

Assuming that an admissible function $u(x)$ exists for which $I(u)$ is an extremum, we first wish to find a necessary condition which this function must satisfy.

Consider a function $\phi(x,t)$ such that

$$\begin{aligned} &\phi(x,t) \text{ is admissible for all } t \\ &\phi(x,t) \text{ and } \phi_t(x,t) \text{ are continuous} \\ &\phi_{xt}(x,t) \text{ is piecewise continuous} \\ &\phi(x,0) = u(x) \end{aligned} \tag{3}$$

For example, we may choose $\phi(x,t) = u(x) + t\zeta(x)$; however, any function satisfying (3) will suffice. If we define

$$G(t) = I(\phi(x,t)) \quad , \tag{4}$$

then G(t) has a stationary point at t = 0. Accordingly

$$(5) \qquad \left.\frac{dG}{dt}\right|_{t=0} = \frac{d}{dt}\int_a^b F(x,\phi,\phi')dx \,\Big|_{t=0} = 0 \ .$$

Differentiating under the integral sign, we have

$$(6) \qquad \int_a^b [F_u\zeta + F_{u'}\zeta']dx = 0 \ ,$$

where*

$$(7) \qquad \zeta(x) = \phi_t(x,0) \ .$$

According to (2), (3), and (7) we see that

(8) ζ is continuous
ζ' piecewise continuous
$\zeta(a) = \zeta(b) = 0$,

the last equation being true since $\phi(a,t) = A$ and $\phi(b,t) = B$.

If we modify (6) by integrating by parts, the analogy with the corresponding necessary condition for an extremum of a function of n variables is revealed.

$$(9) \qquad \int_a^b [F_u\zeta + F_{u'}\zeta']dx = F_{u'}\zeta\Big|_a^b + \int_a^b [F_u - \frac{d}{dx}F_{u'}]\zeta dx$$

$$= \int_a^b [F_u - \frac{d}{dx}F_{u'}]\zeta dx = 0 \ .$$

In the case of a function $f(x_1,\ldots,x_n)$ of n variables, we may derive the necessary conditions for an extremum by considering $g(t) = f(x_1(t), x_2(t),\ldots x_n(t))$ where the equations, $x_i = x_i(t)$ define a curve in n-dimensional space. If $(x_1(0),x_2(0),\ldots x_n(0)$ is to be an extremum, then

$$(10) \qquad \left.\frac{dg}{dt}\right|_{t=0} = 0 \ .$$

---

* In the literature ζ and ζ' are usually called the variation of u and u' respectively, and written δu, $\delta u'$.

In other words,

$$\sum_{i=1}^{n} f_{x_i} \dot{x}_i = 0 . \tag{11}$$

In vector notation, this is the inner product

$$\text{grad } f \cdot \vec{V} = 0 , \tag{12}$$

where $\vec{V}$ is the "velocity" vector with components $(\dot{x}_1, \dot{x}_2, \ldots \dot{x}_n)$. This relation must hold for arbitrary V, from which we may conclude that

$$\text{grad } f = 0 , \tag{13}$$

which is the desired necessary condition for an extremum. Referring to (9), we see that we have an equation similar to (11), the discrete sum being replaced by an integral, x_i and f_{x_i} by ζ and $[F_u - \frac{d}{dx} F_{u'}]$ respectively. By analogy, the necessary condition corresponding to (13) would be

$$F_u - \frac{d}{dx} F_{u'} = 0. \tag{14}$$

This, indeed, is the well known Euler equation for the extremizing function u. We observe, however, that

$$\frac{d}{dx} F_{u'} = F_{u'u'} u'' + F_{uu'} u' + F_{u'x} \tag{15}$$

does not exist for all admissible functions. Hence the Euler equation (14) does not constitute an a priori satisfactory formulation of a necessary condition since it is not clear in advance that for an <u>extremizing</u> function u(x), the quantity $\frac{d}{dx} F_{u'}$ exists. This difficulty may be avoided by integrating (6) by parts in the following way:

$$\int_a^b [F_u \zeta + F_{u'} \zeta'] dx = \zeta \int^x F_u dx \Big|_a^b + \int_a^b \zeta' (F_{u'} - \int^x F_u dx) dx \tag{16}$$

$$= \int_a^b \zeta' (F_{u'} - \int^x F_u dx) dx = 0 .$$

Equation (16) must hold for all ζ such that

(17) $\zeta(a) = \zeta(b) = 0$

ζ continuous

ζ' piecewise continuous.

For our purposes it will be convenient to prove the following Fundamental Lemma: if

$$\int_a^b \zeta'(x)f(x)dx = 0$$

for all ζ satisfying (17), and $f(x)$ is piecewise continuous then $f(x)$ is a constant.

Proof: Since

$$\int_a^b \zeta' dx = \zeta\Big|_a^b = 0 ,$$

it follows that, for any constant C,

$$\int_a^b \zeta'(f - C)dx = 0 . \tag{18}$$

In particular, we may choose C so that

$$\int_a^b (f - C)dx = 0 ; \quad \text{i.e. } C = \int_a^b f dx/(b - a) . \tag{19}$$

With this choice of C, it follows that the function

$$\int_a^x (f - C)dx$$

satisfies (17); hence (18), which must hold for all functions ζ which satisfy (17), must hold in particular for

$$\zeta = \int_a^x (f - C)dx ,$$

i.e. for $\zeta' = f - C$. If we substitute this function in (18)

we obtain

$$(20) \qquad \int_a^b (f - C)^2 dx = 0 .$$

Since $f(x)$ is piecewise continuous, this implies that $f - C \equiv 0$ which proves the lemma.

Applying the lemma to (16) we conclude that

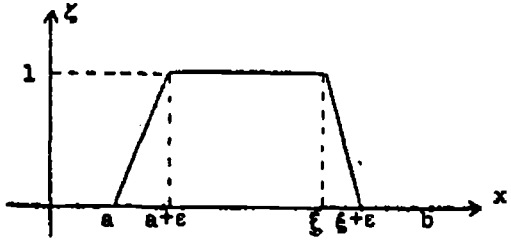
$$(21) \qquad F_{u'} - \int^x F_u dx = C .$$

This is the desired necessary condition which an extremizing function u must satisfy. It is, in fact, the first integral of the Euler equation (14). Since F_u is continuous, we may differentiate (21) and conclude that if u is an extremizing function, $\frac{d}{dx} F_{u'}$ exists and is in fact, continuous, being equal to F_u. Referring to (15) we observe that u" is continuous at every point when u' is continuous, provided that $F_{u'u'} \neq 0$ along the extremals. If we exclude the finitely many points of discontinuity of u' we see that the extremum function has not only a continuous first but also a continuous second derivative. For these points Euler's equation (14) is a satisfactory necessary condition.

It is important, however, to realize that this derivation of (21) is needlessly elaborate. Since all we are seeking is a <u>necessary</u> condition, it follows that any condition which an extremizing function u must satisfy for any particular class of ζ's will be a necessary condition. Now (21) is a relation which must hold for every value of x between a and b, i.e. it is a one parameter set of relations. It therefore seems reasonable that if we choose almost any one parameter family of ζ's, we should be able to derive (21) .

For example, let

$$(22)\qquad \zeta(x) = \begin{cases} \frac{1}{\varepsilon}(x - a) & a \le x \le a + \varepsilon \\ 1 & a + \varepsilon \le x \le \xi \\ 1 - \frac{1}{\varepsilon}(x - \xi) & \xi \le x \le \xi + \varepsilon \\ 0 & \xi + \varepsilon \le x \le b \end{cases}$$



If we substitute (22) into (6), we obtain

$$\int_a^{a+\varepsilon} \frac{1}{\varepsilon}(x - a)F_u\,dx + \int_{a+\varepsilon}^{\xi} F_u\,dx + \int_{\xi}^{\xi+\varepsilon} [1 - \frac{1}{\varepsilon}(x - \xi)]F_u\,dx$$

$$+ \frac{1}{\varepsilon}\int_a^{a+\varepsilon} F_{u'}\,dx - \frac{1}{\varepsilon}\int_{\xi}^{\xi+\varepsilon} F_{u'}\,dx = 0.$$

But

$$\left| \int_a^{a+\varepsilon} \frac{1}{\varepsilon}(x - a)F_u\,dx \right| \le \int_a^{a+\varepsilon} |F_u|\,dx \to 0, \text{ as } \varepsilon \to 0$$

and similarly for the third term. Also, by the Mean Value Theorem

$$\frac{1}{\varepsilon}\int_a^{a+\varepsilon} F_{u'}\,dx = F_{u'}(a + \alpha\varepsilon), \quad (0 \le \alpha \le 1)$$

and

$$\frac{1}{\varepsilon}\int_{\xi}^{\xi+\varepsilon} F_{u'}\,dx = F_{u'}(\xi + \beta\varepsilon), \quad (0 \le \beta \le 1),$$

so that letting ε approach zero, we obtain

$$\int_a^{\xi} F_u\,dx + F_{u'}\Big|_{\xi}^{a} = 0$$

which is equivalent to (21). Many other specific one parameter families can also furnish a derivation of (21).

Problems

1) Give a direct proof that if

$$\int_a^b \zeta(x)f(x)dx = 0$$

for all ζ satisfying (17) where $f(x)$ is piecewise continuous then $f(x)$ is identically zero.

2) Repeat problem (1) where ζ is restricted to the class of functions having continuous second derivatives.

3) Prove the following generalization of the Fundamental Lemma:

Any piecewise continuous function $f(x)$, for which

$$\int_a^b \zeta^{[k]}(x)f(x)dx = 0$$

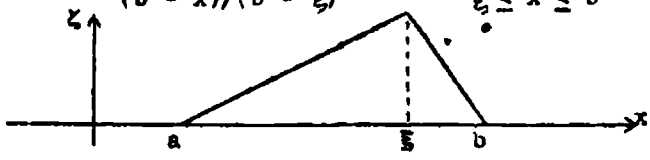
for all $\zeta(x)$ such that ζ, ζ', $\zeta^{[k]}$ are continuous and

$$\zeta^{[n]}(a) = \zeta^{[n]}(b) = 0$$

for $n = 0,1,\ldots k-1$; is a polynomial of degree $k-1$.

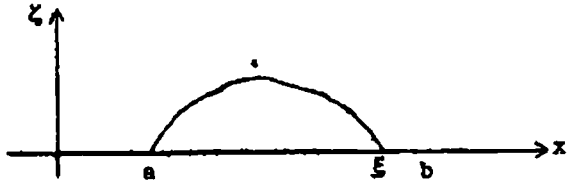
4) Derive Euler's equation using the special variation

$$\zeta(x,\xi) = \begin{cases} (x - a)/(\xi - a) & a \le x < \xi \\ (b - x)/(b - \xi) & \xi \le x \le b \end{cases}$$



5) Derive Euler's equation using the special variation

$$\zeta(x,\xi) = \begin{cases} (x - a)(\xi - x) & a \le x \le \xi \\ 0 & \xi < x \le b \end{cases}$$



6) Show that if $F = F(x,u')$, (i.e. F is not a function of u explicitly), the Euler equation may be solved by quadratures in the form

$$u = \int g(x,C_1)dx + C_2$$

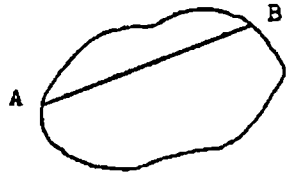
where $u' = g(x,C_1)$ is the explicit solution of the implicit equation

$$F_{u'}(x,u') = C_1 \quad .$$

7) Derive a similar solution of the Euler equation for the case $F = F(u,u')$.

8) Use the Euler equation to solve the brachistochrone problem, the problem of minimum surface of revolution, and the problem of the shortest distance between two points of a plane.

9) The isoperimetric problem may be reduced to the Euler equation in the following way. Consider two fixed points A and B on a closed curve. Then, assuming that the closed curve (of given length) encloses maximum area, the area enclosed by the arc $\widehat{AB}$ and the chord AB must certainly be a maximum for all arcs of the length of $\widehat{AB}$, say L.



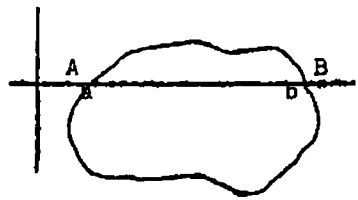
Let the curve AB be given by

$x = \phi(s)$

$y = \psi(s)$

s = arc length

A:(s=0) and B:(s=L) .



Then $\dot{\phi}^2 + \dot{\psi}^2 = 1$, and the area is

$$A = \int_a^b y\,dx = \int_0^L \psi\dot{\phi}\,ds = \int_0^L \psi \sqrt{1 - \dot{\psi}^2}\,ds$$

where $\psi(0) = \psi(L) = 0$.

To extremize $A(\psi)$ is now a simple exercise.

Obtain ψ, and ϕ by use of the Euler equation, and then show that if b is allowed to vary, but L is kept constant, the maximum of A will be for a semi-circle.

2. Generalizations of the Euler Equation. The Euler equation (14) may be generalized along three lines by altering the form of F to contain

a) more than one dependent variable

b) more than one independent variable

c) higher than first derivatives

or any combination of these.

a) Suppose F contains two dependent variables, so that

$$(23) \qquad I = I(\phi,\psi) = \int_a^b F(x,\phi,\phi',\psi,\psi')dx \ ,$$

where ϕ and ψ are admissible functions. In order that $I(u,v)$ be an extremum, it is necessary that $I(u,v)$ be an extremum with respect to u and v considered independently. In other words, a necessary condition is the pair of Euler equations

$$(24) \qquad F_u - \frac{d}{dx} F_{u'} = 0$$

$$F_v - \frac{d}{dx} F_{v'} = 0 \ .$$

In general, if F contains n dependent variables there will be n Euler equations.

An example of such a system is found in the Hamilton Principle of Least Action. Briefly, this says that if a mechanical system be described by n independent coordinates $q_1,\ldots q_n$ (n degrees of freedom), then the motion of the system -- i.e. the determination of $q_i = q_i(t)$ -- will be such as to minimize the "action" I,

$$(25) \qquad I = \int_{t_o}^{t_1} [T(q_1,\ldots q_n,\dot{q}_1,\ldots \dot{q}_n) - U(q_1,\ldots q_n)]dt \ ,$$

where T and U are given functions, the kinetic and potential energy respectively. (Actually, the integral is not an extremum, but only stationary.) The Euler equations which must be satisfied to make I stationary are

$$(26) \qquad \frac{\partial(T - U)}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) = 0 \qquad i = 1,2,\ldots n \ .$$

(These are, in fact, the Lagrangian equations of motion. (For one degree of freedom, this reduces to the familiar Newton equation F = ma.)

b) Next let us suppose that F contains two independent variables x, y so that

$$I(\phi) = \iint_R F(x, y, \phi, \phi_x, \phi_y)dxdy \ , \tag{27}$$

where R is a given closed connected region in the x,y plane, bounded by the curve C. The function will be restricted to the class of admissible functions

$$\begin{aligned} &\phi \text{ continuous in } R \\ &\phi_x, \ \phi_y \text{ piecewise continuous in } R \\ &\phi \text{ takes given values on } C \end{aligned} \tag{28}$$

Suppose that there is an admissible function u(x,y) for which I(u) is an extremum. If we write

$$\phi(x,y,t) = u(x,y) + t\zeta(x,y) \tag{29}$$

where ζ is admissible except that $\zeta = 0$ on C, then ϕ is admissible for all values of t, and

$$G(t) = I(\phi(x,y,t)) \tag{30}$$

is an extremum at $t = 0$. Accordingly,

$$\left.\frac{dG}{dt}\right|_{t=0} = \iint_R [\zeta F_u + \zeta_x F_{u_x} + \zeta_y F_{u_y}]dxdy = 0 \ . \tag{31}$$

Since $\zeta = 0$ on C, this reduces upon integration by parts to

$$\iint_R \zeta[F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y}]dxdy = 0 \ , \tag{32}^*$$

for all admissible ζ. We conclude that the Euler equation in this case is

$$F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} = 0 \ . \tag{33}$$

We observe, however, that, as for the simple Euler equation

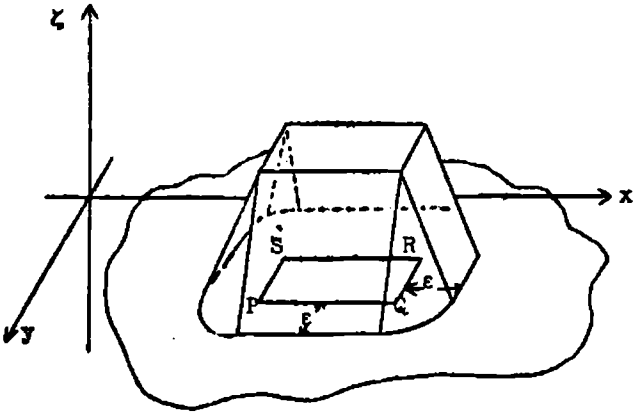
---

* $\partial G/\partial x$ is taken to mean the derivative with respect to x, holding y, but not u or its derivatives constant:

$$\frac{\partial G}{\partial x} = G_x + G_u u_x + G_{u_y} u_{xy} + G_{u_x} u_{xx} \ .$$

(14) of the preceding section, the derivation of (33) implies the existence of second partial derivatives of u, and hence is inapplicable to some of the admissible functions. In the earlier case we were able to integrate by parts in a different order and show that for an extremizing function u" must exist. We are unable to do the corresponding thing here and hence (33) is a necessary condition only if it can be shown that an extremizing function possesses second partial derivatives. We may, however, derive an equation corresponding to the first integral (21) of the Euler equation (14). To do this we investigate (31) directly by considering the following special variation $\zeta(x,y)$; let $\overline{PQRS}$ be an arbitrary rectangle with sides parallel to the coordinate axes and contained inside R, and let $\mu(x,y)$ be the distance from (x,y) to $\overline{PQRS}$; then we define

(36)
$$\zeta(x,y) = \begin{cases} 1 & \text{if } x,y \text{ in } \overline{PQRS} \\ 1 - \mu/\varepsilon & \text{if } 0 \le \mu \le \varepsilon \\ 0 & \text{if } \mu > \varepsilon \end{cases}$$



Then ζ is admissible for any $\varepsilon > 0$, and

$$\zeta_x = 0 \text{ on } \overline{PQ},\ \overline{RS}$$
$$\zeta_y = 0 \text{ on } \overline{PS},\ \overline{RQ}$$
$$\zeta_x = \varepsilon \text{ on } \overline{PS},\ = -\varepsilon \text{ on } \overline{QR}$$
$$\zeta_y = \varepsilon \text{ on } \overline{RS},\ = -\varepsilon \text{ on } \overline{PQ}$$

If we substitute this function ζ into (31), use the theorem of the mean, and let $\varepsilon \longrightarrow 0$ we obtain

$$\iint_{\square} F_u dxdy + \int_P^S F_{u_x} dy - \int_Q^R F_{u_x} dy + \int_P^R F_{u_y} dx - \int_S^R F_{u_y} dx = 0 .$$

This may be conveniently written

$$\iint_{\square} F_u dxdy = \oint_{\square} (F_{u_x} dy - F_{u_y} dx) . \tag{37}$$

Equation (37) corresponds in our present case to the first integral (21) of the Euler equation. It no longer contains an arbitrary function ζ, but now contains an arbitrary rectangle. We may approximate any region by rectangles and hence conclude that for any rectifiable closed curve C' bounding a region R' in R and for any extremizing u,

$$\iint_{R'} F_u dxdy = \oint_{C'} (F_{u_x} dy - F_{u_y} dx) . \tag{38}$$

This is known as Haar's lemma.

The chief value of this lemma lies in the fact that it is applicable to any admissible function since it does not presuppose (as does the Euler equation (33)) the existence of second partial derivatives. If we know that the extremizing function u has continuous second partial derivatives, then the Euler equation (33) will follow directly from (38) by use of the Green's formula

$$\oint_{C'} Pdx + Qdy = \iint_{R'} (Q_x - P_y)dxdy \tag{39}$$

In case F contains n independent variables $x_1, \ldots x_n$, the equation corresponding to (33) is

$$F_u - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} F_{u_{x_i}} = 0 \tag{40}$$

where here, as before, the condition is necessary only if the extremizing function possesses second partial derivatives.

Two important examples of this type of problem are the Dirichlet and Plateau problems, namely to minimize

$$\text{a)} \iint_R (u_x^2 + u_y^2)\,dxdy \tag{41}$$

$$\text{b)} \iint_R \sqrt{1 + u_x^2 + u_y^2}\; dxdy \quad ;$$

the first integral representing, say, the potential energy of an electrostatic field, and the second the area of a surface projecting on R. With given boundary values on C, the latter problem becomes that of minimizing the area of a surface having a given closed space curve as its boundary. The Euler equations corresponding to (41) are

$$\text{a)} \quad u_{xx} + u_{yy} = 0 \tag{42}$$

$$\text{b)} \quad (1 + u_y^2)u_{xx} + 2u_x u_y u_{xy} + (1 + u_x^2)u_{yy} = 0$$

The first equation is the well known Laplace equation, whose solutions are harmonic functions. If we treat the Dirichlet problem from the point of view of Haar's lemma, we get the condition

$$\oint_{C'} (u_x dy - u_y dx) = 0 \tag{43}$$

for an extremizing function. This must hold for all closed rectifiable C's in R and hence $u_x dy - u_y dx$ must be an exact differential. It follows that another function v exists such that

$$u_x = v_y \; ; \quad u_y = -v_x \quad . \tag{44}$$

These are the Cauchy-Riemann equations which a harmonic function must satisfy. (In particular, assuming that second derivatives of u and v exist, the Laplace equation (42a) follows from (44) by elimination of v.)*

c) Finally, let us suppose that F contains a second derivative so that

$$I(\phi) = \int_a^b F(x,\phi,\phi',\phi'')dx \tag{45}$$

the class of admissible functions ϕ now being such that

$$\begin{aligned} &\phi,\ \phi' \text{ continuous} \\ &\phi'' \text{ piecewise continuous} \\ &\phi(a) = A_1,\quad \phi(b) = B_1 \\ &\phi'(a) = A_2,\quad \phi'(b) = B_2 \end{aligned} \tag{46}$$

We again suppose, that for the function u, I(u) is an extremum and set $\phi = u + t\zeta$, where ζ is admissible but $\zeta = \zeta' = 0$ at a and b. Then, if

$$G(t) = I(u + t\zeta)\ ,$$

it follows that G(0) is an extremum and so

$$\left.\frac{dG}{dt}\right|_{t=0} = \int_a^b [\zeta F_u + \zeta' F_{u'} + \zeta'' F_{u''}]dx = 0\ . \tag{47}$$

Integrating by parts, we have

$$\int_a^b \zeta[F_u - \frac{d}{dx} F_{u'} + \frac{d^2}{dx^2} F_{u''}]dx = 0 \tag{48}$$

Hence the Euler equation in this case is

$$F_u - \frac{d}{dx} F_{u'} + \frac{d^2}{dx^2} F_{u''} = 0\ . \tag{49}$$

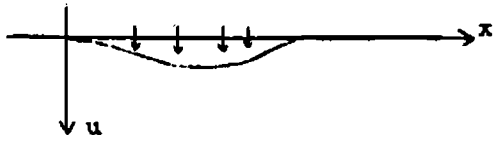
Expanding $(d^2/dx^2)F_{u''}$ we observe that (49) is an equation of

---

* Actually, if (44) holds, continuity of the first derivatives alone will insure (42a)--as well as existence of all higher derivatives.

the fourth order, and implies the existence of u"". As in the first section (page 12) it can be shown, for example by considering a special variation ζ, that <u>for an extremizing function</u> u"" must exist and hence that (49) must hold.

As an example of equation (49) we consider the problem of determining the deflection of a loaded elastic beam. Suppose the beam is clamped rigidly at 0 and 1, and is loaded with a weight per unit length which varies along the x axis, $W = f(x)$.



Then if u(x) is the deflection of the beam at a point x, the potential energy of the system is given by

$$(50) \qquad P.E. = \int_0^1 [\tfrac{1}{2} \alpha u''^2 - uf(x)]dx ,$$

where α is a constant determined by the physical properties of the beam. Since the beam is rigidly clamped at the ends, we have $u = u' = 0$ at 0 and 1.

It follows from a basic principle of mechanics that the equilibrium deflection of the beam will be such as to minimize the potential energy. Hence, if u is the deflection, it must satisfy the Euler equation for minimizing (50), namely

$$(51) \qquad \alpha u'''' - f(x) = 0 .$$

This is a fourth order equation whose integration introduces four constants. These may be determined from the four end conditions $u = u' = 0$ at 0 and 1.

In general, if F contains the n'th derivative $\phi^{[n]}(x)$, so that

$$(52) \qquad I(\phi) = \int_a^b F(x, \phi, \phi', \ldots \phi^{[n]})dx ,$$

then an extremizing function u will satisfy the Euler equation

$$(53) \qquad F_u - \frac{d}{dx} F_{u'} + \frac{d^2}{dx^2} F_{u''} + \ldots (-1)^n \frac{d^n}{dx^n} F_{u^{[n]}} = 0 \ .$$

Problems

1) Derive the Euler equation (49) assuming that only u" is piecewise continuous, using integration by parts and the results of problem 3, page 16 .

2) Do problem (1) using the special variation

$$\zeta(x) = \begin{cases} (x-a)^2(x-\xi)^2 & a \le x \le \xi \\ 0 & \xi < x \le b \end{cases}$$

3) Derive the Euler equation (33) from Haar's lemma, assuming that u possesses all second partial derivatives.

3. Natural Boundary Conditions In section 1, we considered the stationary values of

$$I(\phi) = \int_a^b F(x, \phi, \phi') \, dx$$

where ϕ was required to be a continuous function with a piecewise continuous derivative and such that

$$\phi(a) = A, \quad \phi(b) = B \ .$$

Suppose we now drop this last condition leaving the values of $\phi(a)$ and $\phi(b)$ open. Considering

$$G(t) = I(u + t\zeta)$$

in the usual way; we conclude that

$$\left.\frac{dG}{dt}\right|_{t=0} = \int_a^b [F_u \zeta + F_{u'} \zeta'] dx = 0 \ .$$

However, we may no longer require that $\zeta(a) = \zeta(b) = 0$. Hence, the integration by parts gives, for every ζ,

$$F_{u'}\zeta\Big|_a^b + \int_a^b [F_u - \frac{d}{dx} F_{u'}]\zeta dx = 0 .$$

Consideration of a family of ζ's such that $\zeta(a) = \zeta(b) = 0$ yields the Euler equation as before. There remains

$$F_{u'}\zeta\Big|_a^b = 0 ,$$

and since $\zeta(a)$ and $\zeta(b)$ are arbitrary, we conclude that

$$F_{u'}\Big|_{x=a} = F_{u'}\Big|_{x=b} = 0 . \tag{54}$$

Thus, if we do not prescribe any end point values for the extremizing function u, we find that such a function must automatically satisfy a relation at each end anyway--this is the so-called natural boundary condition. It is evident a priori, that the Euler equation is a necessary condition whether or not boundary conditions are imposed, since any extremum, if it exists, will have definite boundary values.

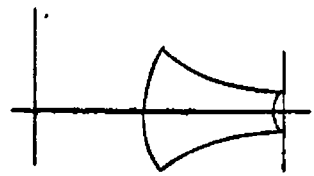
For example, in the problem of the minimum area of revolution, $F = u\sqrt{1 + u'^2}$ and

$$F_{u'} = \frac{uu'}{\sqrt{1 + u'^2}} .$$

If we prescribe $u(a) = A$, i.e. fix one end, but leave $u(b)$ free, we obtain the natural boundary condition

$$u(b)\cdot u'(b) = 0 .$$

Hence the curve which gives the minimum area of revolution will (if b-a is small) be that one which satisfies the Euler equation and is horizontal at $x = b$. If we think of the soap film analogy of this problem, we see that the natural boundary condition is that the film be perpendicular to a wall at b.

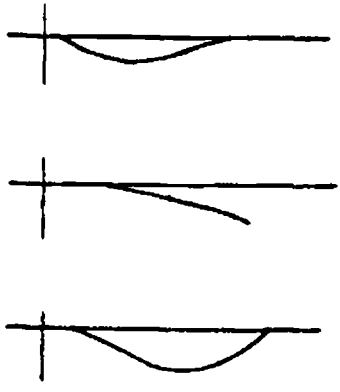


In case F contains higher derivatives the same phenomenon occurs. Instead of deriving general conditions, we will consider the specific example of determining the deflection of a loaded beam which is clamped at one end and at the other end is respectively clamped, free, or supported. Referring to the discussion on page 25, we know that the equilibrium deflection u(x) of the beam will be such as to minimize

$$(55) \qquad P.E. = \int_0^1 [\tfrac{1}{2} a(u'')^2 - u\, f(x)]dx \ .$$

Here the prescribed end conditions are $u = u' = 0$ at $x = 0$, and

i) $u = u' = 0$ at $x = 1$ (clamped)
ii) no condition at $x = 1$ (free)
iii) $u = 0$ at $x = 1$ (supported)



Case i) has already been discussed (see page 24). It leads to the Euler equation of fourth order

$$(56) \qquad au'''' - f(x) = 0 \ ;$$

the four end conditions necessary to determine a specific solution being prescribed.

We consider case ii) directly, i.e. let $\phi = u + t\zeta$,

where $\zeta = \zeta' = 0$ at $x = 0$, but ζ and ζ' are left free at $x = 1$. Then if u minimizes (55) we will have

$$\frac{d}{dt}\Big|_{t=0} \int_0^1 [\tfrac{a}{2}\phi''^2 - \phi f(x)]dx = \int_0^1 (au''\zeta'' - f\zeta)dx = 0 .$$

Integrating by parts, and making use of the condition $\zeta(0) = \zeta'(0) = 0$, we have

$$a[u''\zeta'\Big|_{x=1} - u''\zeta\Big|_{x=1}] + \int_0^1 [au'''' - f]\zeta dx = 0 .$$

Hence, the Euler equation (56) must still hold. But in addition we must have, for all ζ,

$$u''\zeta' = u'''\zeta = 0 \quad \text{at } x = 1 .$$

In other words, if we prescribe no end conditions at $x = 1$--leave it free--we are automatically led to the "natural" boundary conditions

(57) $$u'' = u''' = 0 \quad \text{at } x = 1$$

We still have the necessary four boundary conditions to determine the specific integral of (56).

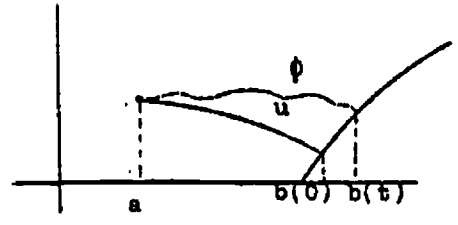
In case iii), where we prescribe $u(1) = 0$ but leave $u'(1)$ free, we must prescribe $\zeta(1) = 0$, but leave $\zeta'(1)$ free. Then it follows that

$$u'' = 0 \quad \text{at } x = 1$$

but no longer that $u''' = 0$ there. Again we have four boundary conditions--three prescribed and one "natural".

A second type of free end condition is to leave the coordinates a and b themselves free, prescribing say only that ϕ have end points on a given curve. Suppose, for simplicity, that we fix one end point a, $\phi(a) = A$, but require only that $\phi(b) = g(b)$ is a given function.

Supposing that $u(x)$ is an extremizing function which intersects $g(x)$ at $x = b(0)$, we let $\phi = u + t\zeta$ and $b = b(t)$. Then if



$$G(t) = \int_a^{b(t)} F(x, \phi, \phi')dx ,$$

we have

$$(58) \quad \left.\frac{dG}{dt}\right|_{t=0} = \int_a^{b(0)} [F_u \zeta + F_{u'} \zeta']dx + F(x,u,u') \Big|_{x=b(0)} \cdot \dot{b}(0) = 0$$

Since

$$(59) \quad u(b(t)) + t\zeta(b(t)) = g(b(t)) ,$$

differentiating and setting $t = 0$, we conclude that

$$(60) \quad \dot{b}(0) = \frac{\zeta(b(0))}{g'(b(0)) - u'(b(0))} .$$

Substituting this in (58) and integrating by parts we obtain, for all ζ

$$(61) \quad \int_a^{b(0)} [F_u - \frac{d}{dx} F_{u'}]\zeta dx + [F_{u'} + \frac{F}{g' - u'}]\zeta \Big|_{x=b(0)} = 0 .$$

Hence an extremizing function u must satisfy the usual Euler equation, a fixed end condition $u(a) = A$, plus a condition

$$(62) \quad F_{u'} + \frac{F}{g' - u'} = 0 \quad \text{at } x = b .$$

Condition (62) is called the Transversality Condition. Together with $u(b) = g(b)$, $u(a) = A$, and the Euler equation, it determines the point b and the solution u. As before leaving the end condition free results in an automatic end condition--in this case a relation between u and u'. The transversality condition (62) reduces to the previously derived natural boundary condition (54) for the case of b fixed, i.e. the fixed curve is $x = b$; and $g'(b) = \infty$, so that the second term in (62) drops out leaving only

$$F_{u'} \Big|_{x=b} = 0 .$$

In many of the specific examples considered so far, F was of the form

$$F = P(x,u)\sqrt{1 + u'^2}$$

In this case the transversality condition (62) reduces to

$$u'g' = -1 \ ,$$

in other words the extremizing curve u must be orthogonal to the given curve g.

Problems

1) Show that the natural boundary condition at a free boundary for $F = F(x,\phi,\phi', \phi'')$ is

$$F_{u'} - \frac{d}{dx} F_{u''} = 0$$

$$F_{u''} = 0 \ .$$

2) The condition that ϕ have its end point, $x = b$, on a fixed curve g may be reduced to the free end condition for fixed b by transforming the x,y plane so that g becomes a vertical line. Derive the transversality condition in this way.

4. Degenerate Euler Equation. The Euler equation for the simplest problem is

$$F_u - \frac{d}{dx} F_{u'} = 0 \ . \tag{14}$$

In case F is linear in u', i.e. of the form

$$F(x,u,u') = A(x,u)u' + B(x,u) \ , \tag{63}$$

the Euler equation reduces to

$$A_x - B_u = 0 \ . \tag{64}$$

This is no longer a differential equation but is, in fact, an implicit relation which in general will define u as a function of x. It follows that u(a) and u(b) may not be arbitrarily prescribed in general.

The converse is also true, i.e. if the Euler equation degenerates from a differential equation into an ordinary equation then F must be of the form indicated in (63).

For, expanding (14) we have

$$F_u - F_{u'x} - F_{u'u}\cdot u' - F_{u'u'}\cdot u'' = 0 \ . \tag{65}$$

If (65) is not a differential equation, then the term containing u" (there is only one) must disappear, i.e.

$$F_{u'u'} = 0 \ . \tag{66}$$

But (66) implies that F is linear in u', i.e. is of the form indicated in (63). Hence, a necessary and sufficient condition that the Euler equation degenerate is that F be linear in u'.

An important special case of this degeneration is when (63) is satisfied, but also A(x,u) and B(x,u) satisfy (64) identically. Then <u>any</u> u will satisfy the degenerate Euler equation and hence will extremize I(u). In this case

$$I(u) = \int_a^b Adu + Bdx \ ,$$

is a line integral whose value is independent of the path of integration (i.e. of u) since we are supposing that (64) is satisfied identically. In other words, I(u) is a constant, and so has no proper extremum.

In case F contains a second derivative, the Euler equation is of fourth order

$$u''''F_{u''u''} + u'''[\ldots] + \ldots = 0 \ . \tag{67}$$

Hence it will degenerate to a lower order if

$$F_{u''u''} = 0 \ .$$

i.e. if F is linear in u"

$$F(x,u,u',u'') = A(x,u,u')u'' + B(x,u,u') \ . \tag{68}$$

In two dimensions following the same reasoning, we conclude that the Euler equation (33) will degenerate if and only if F is linear in u_x and u_y,

$$F(x,y,u,u_x u_y) = A(x,y,u) + B(x,y,u)u_x + C(x,y,u)u_y \tag{69}$$

If (69) holds, the Euler equation becomes

$$A_u - B_x - C_y = 0 , \tag{70}$$

which in general defines u implicitly as a function of x and y.

Problems

1) Show that (67) is of at most second order if (63) holds.
2) Under what circumstances will (67) degenerate even more (i.e. be of less than second order)?
3) Prove that if and only if F is of the form*

$$F = \frac{\partial L}{\partial x}(x,y,u) + \frac{\partial M}{\partial y}(x,y,u) ,$$

then the degenerate Euler equation (70) becomes an identity. As before this means that I(u) is a constant.
4) Show that the Euler equation for two independent variables and second derivatives will degenerate if and only if F is of the form

$$F(x,y,u,u_x,u_y,u_{xx},u_{xy},u_{yy}) =$$

$$P(x,y,u,u_x,u_y)[u_{xx}u_{yy} - u_{xy}^2] + Q(x,y,u,u_x,u_y) .$$

---

* See footnote, page 20.

5. Isoperimetric Problems. Isoperimetric problems in calculus of variations are concerned with extremizing integrals subject to some sort of auxiliary conditions on the dependent functions ϕ, ψ, etc. These conditions may be given in the form of integrals, functional relationships, or differential equations:

$$\text{a)} \quad \int_a^b G(x, \phi, \phi', \psi, \psi', \ldots)dx = 0$$

$$\text{b)} \quad G(x, \phi, \psi, \ldots) = 0$$

$$\text{c)} \quad G(x, \phi, \phi', \psi, \psi', \ldots) = 0 \ .$$

For the corresponding problems in elementary calculus, a necessary condition for an extremum is given by the Euler-Lagrange rule (sometimes called the method of Lagrangian multipliers). We recall that according to this rule, if, among all points (x,y) such that $g(x,y) = 0$, the point (x_o,y_o) extremizes $f(x,y)$, then-- providing $g_x(x_o,y_o)$ and $g_y(x_o,y_o)$ do not both vanish--it follows that (x_o,y_o) is found among the stationary points of $f(x,y) + \lambda g(x,y)$ considered as a function of two independent variables, where the constant λ is determined such that $g(x_o,y_o) = 0$. In other words x_o, y_o, and λ must satisfy the three equations

$$\text{(72)} \qquad \begin{aligned} f_x(x_o,y_o) + \lambda g_x(x_o,y_o) &= 0 \\ f_y(x_o,y_o) + \lambda g_y(x_o,y_o) &= 0 \\ g(x_o,y_o) &= 0 \ . \end{aligned}$$

We shall generalize this rule to cover the present problems in calculus of variations.

a) We first consider the simplest problem, with an integral side condition. Stated explicitly, we are required to extremize

$$I(\phi) = \int_a^b F(x,\phi,\phi')dx$$

among all admissible ϕ (see page 10), which satisfy the additional restriction that

$$(74) \qquad K(\phi) = \int_a^b G(x,\phi,\phi')dx = K_o \text{ (constant)} .$$

We suppose that there is such a function u, extremizing (73) subject to (74). Let $\phi = u(x) + t_1\zeta_1(x) + t_2\zeta_2(x)$ where ζ_1 and ζ_2 are admissible but vanish at a and b. Then

$$(75) \qquad \begin{aligned} I &= I(t_1,t_2) \\ K &= K(t_1,t_2) . \end{aligned}$$

Since we require that ϕ always satisfy (74), the parameters t_1 and t_2 are not independent, but must satisfy $K(t_1,t_2) = K_o$. From our hypothesis that u is an extremizing function it follows that $I(0,0)$ is an extremum of $I(t_1,t_2)$ for all t_1, t_2 which satisfy $K(t_1,t_2) = K_o$. Hence we may apply the Euler-Lagrange rule for ordinary functions to conclude that--providing K_{t_1} and K_{t_2} do not both vanish at $(0,0)$-- there must be a number λ so that

$$(76) \qquad \frac{\partial}{\partial t_i}[I + \lambda K] = \int_a^b \zeta_i[F + \lambda G]_u dx = 0$$

$$i = 1,2.$$

where we define the Euler operator

$$(77) \qquad [H]_u = H_u - \frac{d}{dx} H_{u'} .$$

Since (76) must hold for all ζ_1 and ζ_2, we conclude that

$$(78) \qquad [F + \lambda G]_u = 0 .$$

The restriction that K_{t_1} and K_{t_2} do not both vanish at $(0,0)$

means that

$$K_{t_i}\Big|_0 = \int_a^b \zeta_i [G]_u dx \neq 0 , \quad i = 1 \text{ or } 2 .$$

or

$$[G]_u \neq 0 . \tag{79}$$

The required generalization of the Euler-Lagrange rule is hence the following: if $I(u)$ is an extremum subject to $K(u) = K_o$, and $[G]_u \neq 0$, then there is a constant λ such that $[F + \lambda G]_u = 0$. In general (78) is a second order differential equation containing a parameter λ. Its solution will be in the form

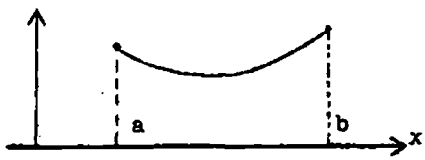
$$u = u(x,\lambda)$$

since the two integration constants may be determined from the (fixed or natural) boundary conditions. The value of λ is then determined so that

$$K(\lambda) = \int_a^b G(x,u(x,\lambda),u'(x,\lambda))dx = K_o .$$

As an illustration, consider the problem of determining the shape of a chain of given length L hanging under gravity.

If $u(x)$ is the shape taken by the chain,



then u is such as to minimize the potential energy

$$PE = \int_a^b u ds = \int_a^b u\sqrt{1 + u'^2}\, dx ,$$

subject to the usual admissibility conditions, plus the additional restriction that the length

$$\int_a^b \sqrt{1 + u'^2}\, dx = L$$

is a given constant. Observe that the integral to be minimized is the same as the one for the problem of minimum area of revolution, for which the solution was seen to be a catenary. In this problem $F = u\sqrt{1 + u'^2}$, $G = \sqrt{1 + u'^2}$, so the generalized Euler-Lagrange rule states that there is a constant λ such that

$$(80) \qquad [F + \lambda G]_u = \sqrt{1 + u'^2} - \frac{d}{dx} \cdot \frac{u'(u + \lambda)}{\sqrt{1 + u'^2}} = 0 .$$

The solutions of this equation have the form $u = v(x) - \lambda$ where $v(x)$ is a solution for $F = 0$, hence represents a catenary. Solving (80), with $u(a) = u_0$, $u(b) = u_1$, we would get $u = u(x,\lambda)$. Then λ would be determined so that

$$\int_a^b \sqrt{1 + u'^2(x,\lambda)}\, dx = L ,$$

and the result would be a catenary of length L between the points a and b. The side condition is therefore seen to be merely a restriction on the class of admissible functions.

In case F contains more dependent variables, or there are more side conditions we have: if $u_1, \ldots u_k$ extremize $I(u_1, \ldots u_k)$ subject to the n conditions

$$K_i = \int_a^b G_i(x, u_1, u_1', \ldots u_k, u_k')dx = 0 , \qquad i = 1,2,\ldots n ,$$

then, in general, there are n constants $\lambda_1, \ldots \lambda_n$ such that

$$[F + \lambda_1 G_1 + \ldots \lambda_n G_n]_{u_j} = 0 \qquad j = 1,2,\ldots k .$$

b) We next consider the problem of extremizing

$$(81) \qquad I(\phi,\psi) = \int_a^b F(x,\phi,\phi',\psi,\psi')dx$$

where ϕ and ψ are subject to the side condition

$$(82) \qquad G(x,\phi,\psi) = 0 \ .$$

It will be shown that the generalized Euler-Lagrange rule still applies here, except that λ is no longer a constant, but instead is, in general, a function of x. A proof of this may be obtained directly by eliminating, say, ψ from (82)--i.e. obtain $\psi = \psi(x,\phi)$--and so reducing the problem to that of the simplest case with no side conditions. We prefer to consider the variation of I.

Suppose, then, that $I(u,v)$ is an extremum under the restriction that $G(x,u,v) = 0$. Let $\phi(x,t) = u(x) + t\zeta_1(x)$ where $\zeta_1(a) = \zeta_1(b) = 0$, and $\psi(x,t) = v(x) + \zeta_2(x,t)$ where ζ_2 is determined, once ζ_1 is given, by the relation

$$G(x,u + t\zeta_1, v + \zeta_2) = 0$$

automatically insuring that $\zeta_2(x,0) = 0$. We know that $I(\phi,\psi)$ is an extremum for $t = 0$, and $G(x,\phi,\psi) \equiv 0$ (in t) so that for any function $\lambda(x)$

$$\frac{d}{dt} \int_a^b [F + \lambda G]dx\Big|_{t=0} = 0$$

Accordingly, we have

$$(83) \qquad \int_a^b \left\{\zeta_1[F + \lambda G]_u + \frac{d\zeta_2}{dt}\Big|_{t=0} \cdot [F + \lambda G]_{v'}\right\} dx = 0$$

for all admissible ζ_1 (ζ_2 is not arbitrary). Since (83) holds for all $\lambda(x)$, we will try to find a particular $\lambda(x)$ so that $[F + \lambda G]_{v'} = 0$. Expanding, we have, since G does not contain v'

$$(84) \qquad [F + \lambda G]_{v'} = F_{v'} + \lambda G_{v'} - \frac{d}{dx} F_{v'} = 0 \ .$$

Hence, we may solve for λ if $G_v \neq 0$. If $G_v = 0$, but $G_u \neq 0$, then the roles of ζ_1 and ζ_2 may be interchanged. If both vanish, the procedure breaks down. We note that since $G_{v'} = 0$, the condition $G_v \neq 0$ may be written $[G]_v \neq 0$. This choice of λ leaves only the first term in (83),

$$\int_a^b \zeta_1 [F + \lambda G]_u dx = 0 .$$

Therefore, we conclude that

$$[F + \lambda G]_u = 0 .$$

Hence, the Euler-Lagrange rule still holds, except that λ is now a function of x: i.e. we have shown that if u and v extremize $I(\phi,\psi)$ subject to $G(x,\phi,\psi) = 0$, then there is a function $\lambda(x)$, such that

$$[F + \lambda G]_u = [F + \lambda G]_v = 0 ,$$

provided that $[G]_u$ and $[G]_v$ do not both vanish. In general the Euler equations in u and v (if there are more dependent variables, there will be a corresponding number of Euler equations) are second order differential equations with fixed or natural end conditions; their solution will be in the form

$$u = u(x,\lambda(x))$$
$$v = v(x,\lambda(x)) .$$

The function $\lambda(x)$ is then determined by solving the equation $G(x,u,v) = C$ for $\lambda(x)$.

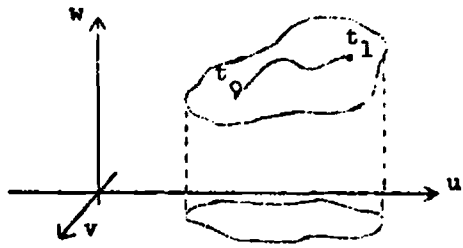
An example of this type of problem is that of geodesics on a surface. Suppose (u,v,w) are the rectangular coordinates of a point on a surface determined by the equation

(85) $$G(u,v,w) = 0 .$$

Then a curve on this surface may be given in the form

$$u = u(t), \quad v = v(t), \quad w = w(t),$$

where u, v, and w are not independent, but must satisfy (85). The arc length of the curve, which is defined to be a minimum between two fixed points for the geodesics is given by



$$(86) \qquad \int_{t_0}^{t_1} F\,dt = \int_{t_0}^{t_1} \sqrt{u^2 + v^2 + w^2}\; dt \ .$$

Hence, we seek to minimize (86) subject to (85). The Euler equations are

$$\frac{d}{dt} \frac{\dot{u}}{\sqrt{\dot{u}^2 + \dot{v}^2 + \dot{w}^2}} - \lambda G_u = 0 \ ,$$

and similarly for v and w. Since the parameter t is arbitrary we may choose t = sa, where s is the arc length; then

$$\sqrt{\dot{u}^2 + \dot{v}^2 + \dot{w}^2} = \frac{1}{a} \ ,$$

and the equations become (with a different λ)

$$u'' - \lambda G_u = 0$$
$$v'' - \lambda G_v = 0$$
$$w'' - \lambda G_w = 0 \ .$$

As soon as a specific G is given, these together with G = 0 may be solved for u, v, w, and λ. We observe, however, that for any G, the geodesics must be such that the directions

$$G_u : G_v : G_w \ , \qquad u'', v'', w''$$

are parallel. Since these are the directions respectively, of the normal to the surface and the principal normal to the curve, we conclude that the principal normal to a geodesic at every

point coincides with the normal to the surface.

c) We finally consider the isoperimetric problem having a differential equation as a side condition. We seek to extremize

$$I(\phi,\psi) = \int_a^b F(x,\phi,\psi,\phi',\psi')\,dx$$

subject to the side condition

$$G(x,\phi,\phi',\psi,\psi') = 0 \quad (87)$$

This type of problem is encountered, for example, in non-holonomic dynamical systems, in which the number of independent coordinates is not equal to the number of degrees of freedom. The Euler-Lagrange rule still holds for this case, i.e. if u and v extremize $I(\phi,\psi)$ subject to (87) and if $[G]_u$, $[G]_v$ are not both zero, then there is a function $\lambda(x)$ such that

$$[F + \lambda G]_u = [F + \lambda G]_v = 0$$

and $G(x,u,v,u',v') = 0$.

The proof is more involved here because not even one of the dependent functions, say ϕ, can be subjected to an arbitrary admissible variation. For, if we attempt this, ψ is determined as the solution of a first order differential equation, but must satisfy two boundary conditions, which is in general impossible. This difficulty may be overcome by considering variations of the form $\phi = u + t_1\zeta_1 + t_2\zeta_2$, $\psi = v + (t_1,t_2,x)$ where a relation between t_1 and t_2 is fixed in order to satisfy the second boundary condition.*

As an example, we consider the simplest problem of the first section only expressed in a slightly different manner.

---

* For details, see Bolza, "Variationsrechnung", p. 558 or Hilbert, "Zur Variationrechnung", Math. Annal. Vol. LXII, No.3.

We seek to extremize

$$I = \int_a^b F(x,v,u)dx$$

subject to the condition

$$u = v' .$$

Here $F = F(x,v,u)$ and $G = u - v'$. Then, according to the rule, there should be a function $\lambda(x)$ so that $u - v' = 0$, and

$$[F + \lambda G]_u = [F + \lambda G]_v = 0 .$$

We have, respectively

$$F_u + \lambda = 0$$

$$F_v - \frac{d}{dx}\lambda(-1) = F_v + \lambda' = 0$$

Hence, $\lambda = -F_u$. Eliminating λ, we have

$$F_v - \frac{d}{dx} F_u = 0 ,$$

or, since $u = v'$

$$F_v - \frac{d}{dx} F_{v'} = 0 ,$$

the usual Euler equation. Of course, this problem avoids the difficulty inherent in the general problems since only v (and not u) must satisfy boundary conditions, so that in this case arbitrary variations of v are permissible.

The problem of differential equation side conditions for the case of more than one independent variable is even more difficult, and has only been solved in special instances.

Problems

1) By use of the Euler-Lagrange rule, solve the isoperimetric problem of the circle--i.e. minimize

$$L(u) = \int_a^b \sqrt{1 + u^2}\, dx$$

subject to the condition that

$$A(u) = \int_a^b u\,dx = A_0, \text{ a constant}$$

2) Use the Euler-Lagrange rule to find the shape of a hanging chain (see page 35).

e) Find the geodesics on the sphere

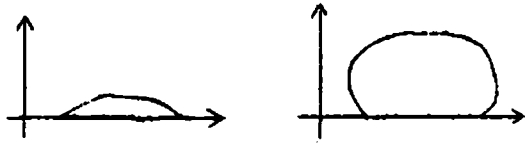
$$G(u,v,w) = u^2 + v^2 + w^2 - 1 = 0 .$$

See page 39.

6. Parametric Form of the Theory. For many types of problems having physical origins the class of admissible functions considered up to now is too restricted. For example, in the isoperimetric problem we seek a curve of given length which maximizes the area enclosed between the curve and the straight line joining its endpoints. If the given length is sufficiently small,

$$L \leq \frac{\pi(b-a)}{2}$$

the maximizing arc may be written in the form $y = u(x)$, where $a \leq x \leq b$. However if $L > \pi(b-a)/2$, the problem still has a solution, but it is no longer expressible in the form of a function $y = u(x)$, $a \leq x \leq b$. This artificial restriction may be removed by expressing F in parametric form, as follows. Consider the functional



$$I = \int_{t_0}^{t_1} H(x,y,\dot{x},\dot{y})\,dt \quad , \tag{88}$$

where $x = x(t)$, $y = y(t)$ are parametric equations of an admissible curve. We require that for an admissible curve

$$x(t_o) = a, \quad x(t_1) = b$$
$$y(t_o) = A, \quad y(t_1) = B$$

and that $x(t)$ and $y(t)$ are continuous functions of t, with piecewise continuous derivatives such that $\dot{x}^2 + \dot{y}^2 \neq 0$. The important point in this case is that for the problem to make sense H cannot be arbitrary but must be a homogeneous function of the first degree in $\dot{x}$ and $\dot{y}$. This follows from the requirement that I sould depend only upon the curve joining the fixed end points and not upon the particular parametric representation used to describe that curve. Hence, if we replace the parameter t by another parameter $\tau = \tau(t)$ in a one to one way ($\dot{\tau} > 0$), I should not change. We therefore have the following equation:

$$\int_{t_o}^{t} H(x,y,\dot{x},\dot{y})dt = \int_{\tau_o}^{\tau} H(x,y,\dot{\tau}\frac{dx}{d\tau},\dot{\tau}\frac{dy}{d\tau})\frac{d\tau}{\dot{\tau}}$$

$$= \int_{\tau_o}^{\tau} H(x,y,\frac{dx}{d\tau},\frac{dy}{d\tau})d\tau \ .$$

This requires that, for $\dot{\tau} > 0$

$$H(x,y,\dot{\tau}\frac{dx}{d\tau},\dot{\tau}\frac{dy}{d\tau}) = \dot{\tau}\, H(x,y,\frac{dx}{d\tau},\frac{dy}{d\tau}) \quad . \tag{89}$$

In particular if $\dot{\tau} = k$, a positive constant we have

$$H(x,y,k\dot{x},k\dot{y}) = kH(x,y,\dot{x}.\dot{y}) \ ,$$

which is the definition of a homogeneous function of the first degree in $\dot{x}$ and $\dot{y}$. Thus, if a problem be expressed in parametric form, the integrand must be homogeneous of first degree in $\dot{x}$ and $\dot{y}$.

If $\dot{x} > 0$ throughout the interval (t_0, t_1), we may take $\dot{\tau} = 1/\dot{x}$ in (89), and setting $y' = \dot{y}/\dot{x}$ (89) becomes

$$H(x,y,\dot{x},\dot{y}) = \dot{x}H(x,y,1,y') = \frac{dx}{dt} F(x,y,y') ,$$

and we may write

$$(90) \qquad I = \int_{t_0}^{t_1} H(x,y,\dot{x},\dot{y})dt = \int_a^b F(x,y,y')dx .$$

Thus if x always increases as the curve is traversed (which is the case if $\dot{x} > 0$), the homogeneous problem may be transformed back into the inhomogeneous problem. Of course everything which has been done for $\dot{\tau} > 0$ and $\dot{x} > 0$ applies equally well if the direction of the inequality is reversed.

The integral (88) may be treated exactly as in the other problems we have considered. We obtain, as a necessary condition satisfied by an extremal $x = x(t)$, $y = y(t)$, the pair of simultaneous equations

$$(91) \qquad \begin{aligned} H_x - \frac{d}{dt} H_{\dot{x}} &= 0 \\ H_y - \frac{d}{dt} H_{\dot{y}} &= 0 . \end{aligned}$$

These equations must hold for the extremizing curve independently of the choice of parameter in terms of which it is defined. The choice of parameter is thus left to expediency in any actual problem under consideration.

As an example, we consider the problem of determining the geodesics on a surface, phrased in parametric form rather than isoperimetrically (cf. page 40). Suppose the surface $G(u,v,w) = 0$ is defined parametrically as $u = u(x,y)$, $v = v(x,y)$, $w = w(x,y)$. That is, we have a correspondence between a region in the x,y plane and the surface G in the u,v,w space. A curve on G will then be represented by a curve in the x,y plane. If this curve be given parametrically in the form $x = x(t)$, $y = y(t)$, then the curve on the surface $G = 0$ will take the form $u = u(x(t),y(t))$, $v = v(x(t),y(t))$,

$w = w(x(t),y(t))$. It has for element of arc length

$$ds = \sqrt{E\dot{x}^2 + 2F\dot{x}\dot{y} + G\dot{y}^2}\ dt$$

where

$$E = u_x^2 + v_x^2 + w_x^2$$
$$F = u_x u_y + v_x v_y + w_x w_y$$
$$G = u_y^2 + v_y^2 + w_y^2$$

are given functions of x and y, and

$$\Delta = EG - F^2 > 0\ .$$

Denoting the radical by H, the integral to be minimized then becomes

$$I = \int_{t_0}^{t_1} H dt$$

Euler's equations (91) for the minimizing curve become

$$(92) \qquad \frac{E_x\dot{x}^2 + 2F_x\dot{x}\dot{y} + G_x\dot{y}^2}{2H} - \frac{d}{dt}\frac{E\dot{x} + F\dot{y}}{H} = 0$$

$$\frac{E_y\dot{x}^2 + 2F_y\dot{x}\dot{y} + G_y\dot{y}^2}{2H} - \frac{d}{dt}\frac{F\dot{x} + G\dot{y}}{H} = 0\ .$$

In general (92) constitutes a very unwieldy system of equations. However, we recall that any parameter may be used in place of t. In particular, if we use a parameter, s, proportional to the arc length, H is a constant and the equations (92) reduce to

$$(93) \qquad E_x x'^2 + 2F_x x'y' + G_x y'^2 = 2\frac{d}{ds}(Ex' + Fy')$$
$$E_y x'^2 + 2F_y x'y' + G_y y'^2 = 2\frac{d}{ds}(Fx' + Gy')$$

where $x' = \frac{dx}{ds}$, $y' = \frac{dy}{ds}$. These equations are then the appropriate system to use in determining the geodesics for any particular surface determined by the functions u,v,w.

We remark that although the parametric formulation of the problem results in two Euler equations (91), these equations are not independent because of the requirement that H be homogeneous of first degree in $\dot{x}$ and $\dot{y}$. (This, of course, is to be expected, since an equivalent inhomogeneous problem would have only one Euler equation.) In fact, H must satisfy the Euler homogeneity identity

$$H = \dot{x}H_{\dot{x}} + \dot{y}H_{\dot{y}} \tag{94}$$

and combining (91) with (94) we obtain the single equation

$$H_{x\dot{y}} - H_{y\dot{x}} + \bar{H}(\dot{x}\dot{y} - \dot{x}\dot{y}) = 0 , \tag{95}$$

where

$$\bar{H} = \frac{H_{\dot{x}\dot{x}}}{\dot{y}^2} = \frac{H_{\dot{x}\dot{y}}}{-\dot{x}\dot{y}} = \frac{H_{\dot{y}\dot{y}}}{\dot{x}^2} , \quad *$$

which is equivalent to the system (91).

The homogeneous formulation may be carried out for the case of more than one independent variable in a straightforward way, the only formal change being that the Euler equations will be partial differential equations. We consider as an example the Plateau problem of minimal surfaces already discussed on page 22. We will now formulate this problem parametrically.

---

* This relation is obtained from (94) by differentiating with respect to $\dot{x}$ and $\dot{y}$:

$$H_{\dot{x}} = H_{\dot{x}} + \dot{x}H_{\dot{x}\dot{x}} + \dot{y}H_{\dot{x}\dot{y}}$$

$$H_{\dot{y}} = \dot{x}H_{\dot{x}\dot{y}} + H_{\dot{y}} + \dot{y}H_{\dot{y}\dot{y}}$$

hence

$$\dot{x}H_{\dot{x}\dot{x}} + \dot{y}H_{\dot{x}\dot{y}} = \dot{x}H_{\dot{x}\dot{y}} + \dot{y}H_{\dot{y}\dot{y}} = 0 ,$$

which leads immediately to the stated equation.

Let $x = x(t_1,t_2)$, $y = y(t_1,t_2)$, $z = z(t_1,t_2)$ be the parametric equations of a surface such that as t_1,t_2 traces out the curve g in the t_1,t_2 plane x,y,z traces out the fixed space curve Γ in the x,y,z space thus satisfying the boundary condition. We wish to find the functions x,y,z for which the area of the surface enclosed by Γ is a minimum--i.e. to minimize the integral

$$I = \iint_g \sqrt{EF - G^2}\, dt_1 dt_2 = \iint_g \sqrt{W}\, dt_1 dt_2$$

where

$$E = x_{t_1}^2 + y_{t_1}^2 + z_{t_1}^2$$

$$F = x_{t_1} \cdot x_{t_2} + y_{t_1} \cdot y_{t_2} + z_{t_1} \cdot z_{t_2}$$

$$G = x_{t_2}^2 + y_{t_2}^2 + z_{t_2}^2 .$$

The resulting Euler equations are

$$\frac{\partial}{\partial t_1}\left(\frac{W_{x_{t_1}}}{\sqrt{W}}\right) + \frac{\partial}{\partial t_2}\left(\frac{W_{x_{t_2}}}{\sqrt{W}}\right) = 0 \tag{96}$$

$$\frac{\partial}{\partial t_1}\left(\frac{W_{x_{t_1}}}{\sqrt{W}}\right) + \frac{\partial}{\partial t_2}\left(\frac{W_{x_{t_2}}}{\sqrt{W}}\right) = 0$$

$$\frac{\partial}{\partial t_1}\left(\frac{W_{x_{t_1}}}{\sqrt{W}}\right) + \frac{\partial}{\partial t_2}\left(\frac{W_{x_{t_2}}}{\sqrt{W}}\right) = 0 .$$

Here as before the equations can be greatly simplified by a proper choice of parameters. We can always choose t_1 and t_2 so that $F = 0$ and $E = G$, and with this choice, equations (95) reduce to very elegant form

$$x_{uu} + x_{vv} = 0 \tag{97}$$

$$y_{uu} + y_{vv} = 0$$

$$z_{uu} + z_{vv} = 0$$

Problems

1) Verify that (97) results from (96) on setting $F = 0$, $E = G$. Compare with the previous method used, page 22.

2) Use (93) to find the geodesics on a sphere, representing the sphere parametrically by

$$u = \sin x \sin y$$
$$v = \sin x \cos y$$
$$w = \cos x .$$

Compare with the previous methods mentioned, i.e. in simplest form (page 11) and in isoperimetric form (page 38).

3) Derive (95) from (91) and (94).

7. Invariance of the Euler Equation. On page 13 we mentioned that the Euler Equation may be thought of as a generalization of the vector equation grad $f = 0$, where $f = f(x_1, \ldots x_n)$. We recall that if the independent variables $x_1, \ldots x_n$ are transformed into new variables $\xi_1(x_1, \ldots x_n), \ldots \xi_n(x_1, \ldots x_n)$ in such a way that the Jacobian

$$\frac{\partial(x_1, \ldots x_n)}{\partial(\xi_1, \ldots \xi_n)} \neq 0 \text{ or } \infty ,$$

then if $\text{grad}_{(x)} f = 0$ at a point $(\bar{x}_1, \ldots \bar{x}_n)$ it will follow that $\text{grad}_{(\xi)} f = 0$ at the corresponding point $(\bar{\xi}_1, \ldots \bar{\xi}_n)$. We say that the equation grad $f = 0$ is invariant under transformation of the coordinate system. We would then expect that the Euler equation should also be inveriant under transformation of the coordinate system. That this is so may be verified by considering the result of replacing x by $\xi = \xi(x)$ in

$$I = \int_{x_0}^{x_1} F(x, u, u_x)dx = \int_{\xi_0}^{\xi_1} F[x(\xi), u(x(\xi)), u_\xi(x(\xi))\xi_x]x_\xi d\xi$$

$$= \int_{\xi_0}^{\xi_1} H(\xi, v, v_\xi)d\xi .$$

where $v(\xi) = u(x(\xi))$. A one parameter family $u(x) + t\zeta(x)$ will correspond to $v(\xi) + t\eta(\xi)$ where $\zeta(x(\xi)) = \eta(\xi)$; substituting these expressions into the above integrals, differentiating with respect to t and setting $t = 0$, we have

$$\int_{x_0}^{x_1} \zeta [F]_{u(x)} dx = \int_{\xi_0}^{\xi_1} \eta [H]_{v(\xi)} d\xi = 0 .$$

Since $d\xi = \xi_x dx$ and $\zeta(x) = \eta(\xi)$, we conclude that

$$\int_{x_0}^{x_1} \zeta \left\{ [F]_u - [H]_v \xi_x \right\} dx = 0$$

or

$$[F]_u = [H]_v \xi_x . \tag{98}$$

In the case of more than one independent variable, a change of coordinate system from $(x_1, \ldots x_n)$ to $(\xi_1, \ldots \xi_n)$ requires that the Euler operator in the new system be multiplied by the Jacobian of the transformation. We have

$$[F(x_1, \ldots x_n;\ u, u_{x_1}, \ldots u_{x_n})]_u$$

$$= [H(\xi_1, \ldots \xi_n;\ v, v_{\xi_1}, \ldots v_{\xi_n})]_v \frac{\partial(\xi_1, \ldots \xi_n)}{\partial(x_1, \ldots x_n)}$$

where

$$v(\xi_1, \ldots \xi_n) = u(x_1(\xi_1, \ldots \xi_n), \ldots x_n(\xi_1, \ldots \xi_n))$$

and

$$H = F[x_1(\xi_1, \ldots \xi_n), \ldots x_n(\xi_1, \ldots \xi_n);\ v(\xi_1, \ldots \xi_n);\ u_{x_1}(\xi_1, \ldots \xi_n),$$

$$\ldots u_{x_n}(\xi_1, \ldots \xi_n)] \frac{\partial(x_1, \ldots x_n)}{\partial(\xi_1, \ldots \xi_n)}$$

with

$$u_{x_i} = v_{\xi_1} \frac{\partial \xi_1}{\partial x_i} + \ldots + v_{\xi_n} \frac{\partial \xi_n}{\partial x_i} .$$

In particular, if for a given u, $[F]_u = 0$, then $[H]_v = 0$, so that the Euler equation is invariant. The property that a curve be an extremal (i.e. a solution of Euler's equation) remains unaltered by a transformation of the independent variables.

The invariance of the Euler expression is a useful principle in actually computing the transformed differential expression. One of the most important differential operators is the Laplacian $\triangle u = u_{xx} + u_{yy} = -\frac{1}{2}[F]_u$, where $F = u_x^2 + u_y^2$. If we wish to find $\triangle u$ in polar coordinates $x = r\cos\theta$, $y = r\sin\theta$, we may calculate $\frac{\partial(x,y)}{\partial(r,\theta)} = r$, $H = r(v_r^2 + \frac{1}{r^2}v_\theta^2)$, and conclude that

$$[F]_u = [H]_v \cdot \frac{1}{r} = \frac{1}{r}(H_v - \frac{\partial}{\partial r}H_{v_r} - \frac{\partial}{\partial\theta}H_{v_\theta}$$

$$= -\frac{2}{r}(rv_{rr} + v_r + \frac{1}{r}v_{\theta\theta}) .$$

Hence, from $\triangle u = -\frac{1}{2}[F]_u$

$$\triangle v = v_{rr} + \frac{1}{r}v_r + \frac{1}{r^2}v_{\theta\theta} ,$$

a result obtained without computation of second derivatives.

Problems

1) Find $\triangle u$ for spherical coordinates $x = r\cos\theta\sin\psi$, $y = r\sin\theta\sin\psi$, $z = r\cos\psi$. Generalize to n dimensions.

2) Finc $\triangle u$ for $\xi = \xi(x,y)$, $\eta = \eta(x,y)$.

3) In n-dimensions,

$$\triangle u = \sum_{i=1}^{n} u_{x_i x_i} .$$

Let $x_i = x_i(\xi_1, \ldots \xi_n)$ and $\xi_i = \xi_i(x_1, \ldots x_n)$, and define

$$g^{ik} = \sum_{j=1}^{n} \frac{\partial\xi_i}{\partial x_j} \cdot \frac{\partial\xi_k}{\partial x_j} .$$

Then show that

(a) $$D = \begin{vmatrix} g^{11} & g^{12} & \cdots & g^{1n} \\ g^{21} & & & \vdots \\ \vdots & & & \vdots \\ g^{n1} & \cdots & \cdots & g^{nn} \end{vmatrix} = \left[\frac{(\xi_1, \ldots \xi_n)}{(x_1, \ldots x_n)} \right]^2$$

(b) $$\Delta u = \sqrt{D} \cdot \sum_j \frac{\partial}{\partial \xi_j} \left(\frac{1}{\sqrt{D}} \sum_k g^{jk} u_{\xi_k} \right) .$$

8. The Legendre Condition. In the theory of extrema of functions of a single variable, a necessary condition for a minimum, besides $f'(x) = 0$, is that $f'' \geq 0$ (if it exists). A condition somewhat analogous to this will be seen to hold in Calculus of Variations.

Let us suppose that there is an admissible function u, for which

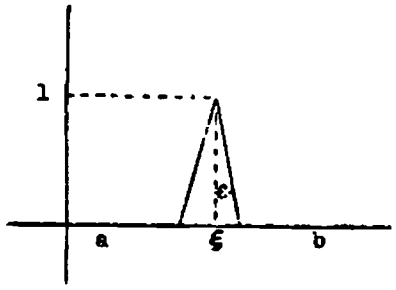
$$I(u) = \int_a^b F(x,u,u')dx$$

is a minimum. Then $G(t) = I(u + t\zeta)$ has a minimum at $t = 0$; accordingly $G'(0) = 0$ (from which follows the Euler equation $[F]_u = 0$) and also $G''(0) \geq 0$, assuming its existence. Hence for every admissible ζ,

(99) $$G''(0) = \int_a^b [F_{uu}\zeta^2 + F_{uu'}\zeta\zeta' + F_{u'u'}\zeta^2]dx \geq 0 .$$

We choose a special variation

$$\zeta(x) = \begin{cases} 0 & a \leq x \leq \xi - \varepsilon \\ 1 + (x - \xi)/\varepsilon & \xi - \varepsilon \leq x \leq \xi \\ 1 - (x - \xi)/\varepsilon & \xi \leq x \leq \xi + \varepsilon \\ 0 & \xi + \varepsilon \leq x \leq b . \end{cases}$$



If we substitute this function in (99) and let $\varepsilon \to 0$, the term

$$\frac{1}{\varepsilon^2} \int_{\xi-\varepsilon}^{\xi+\varepsilon} F_{u'u'}\,dx$$

will dominate the left side of the inequality (99) and determine its sign. Thus the sign of $F_{u'u'}$ determines the sign of $\Omega''(0)$, and for a minimum, the weak Legendre Condition

(100) $$F_{u'u'} \geq 0$$

must hold for all points on the curve u. We have seen (page 31) that $F_{u'u'} \neq 0$ is essential in order that the Euler equation should not degenerate. In many problems the strong Legendre condition $F_{u'u'} > 0$ holds, but this is still not sufficient to guarantee a minimum. (We will see in the next chapter that if $F_{\phi'\phi'} > 0$ for all admissible ϕ, then a solution of the Euler equation is a minimum.)

Problems frequently take the form $F = g(\phi)\sqrt{1 + \phi'^2}$ for which $F_{\phi'\phi'} = g/(1 + \phi'^2)$ and $F_{\phi'\phi'} > 0$ if $g(\phi) > 0$, so that for such a function Legendre's condition is sufficient.

In case F contains more dependent functions, the Legendre condition is that the matrix

$$(F_{u_i u_j})$$

be positive definite, that is

$$\sum_I \sum_J \lambda_i \lambda_j F_{u_i' u_j'} \geq 0$$

for all λ_i, λ_j .

Problems

1) Verify the last statement above.

II HAMILTON JACOBI THEORY - SUFFICIENT CONDITIONS

1. The Legendre Transformation. Transformation from point to line coordinates is frequently useful in the theory of differential equations. We consider first the one dimensional case of a curve $u = f(x)$, which can be considered as the envelope of its tangent lines. The tangent line at any point x,u is given by

$$U + xf'(x) - f(x) = f'(x)X \,, \tag{1}$$

U and X being coordinates along the line. The line (1) is determined by its coordinates (i.e. slope and intercept)

$$\begin{cases} \xi = f'(x) \\ W = xf'(x) - u \end{cases} \tag{2}$$

where a unique value of (ξ, w) is assigned to each point (x,u). The curve $u = f(x)$ can then be represented as $w = W(\xi)$ on elimination of x between the equations (2), which can always be done if $f''(x) \neq 0$.* Between corresponding sets of coordinates (x,u) and (ξ,w) (i.e. referring to the same point on the curve) there exists the symmetric relation

$$u + w = x\xi \tag{3}$$

in virtue of (1).

The equations (2) allow a transformation from point to line coordinates. To reverse this procedure we suppose $w = W(\xi)$ is given and find the envelope of the one parameter family of lines

---

*If $f''(x) \neq 0$, $\xi = f'(x)$ can be solved for x, and substitution in $w = xf'(x) - f(x)$ gives the required relation. Inversion of $w = xf'(x) - f(x)$ instead of $\xi = f'(x)$ requires $xf''(x) \neq 0$ so nothing is gained. The points $f''(x) = 0$ (inflection points) are singularities in line coordinates, and $f''(x) \equiv 0$ represents a single straight line. Similarly $W''(\xi) = 0$ is a cusp in point coordinates, and $W''(\xi) \equiv 0$ represents a pencil of lines through a fixed point (no envelope). Duality is observed as points on a line ($f''(x) \not\equiv 0$) and lines through a point ($W''(\xi) = 0$).

(4) $$U + W(\xi) = \xi X \ .$$

To do this we differentiate with respect to the parameter ξ and combine with (4), setting $U = u$ and $X = x$ (coordinates along the envelope), obtaining

(5) $$\begin{cases} x = W'(\xi) \\ u = \xi W'(\xi) - W \ . \end{cases}$$

The duality of the transformation is evident on comparison of (2) and (5). Elimination of ξ in (5) to obtain $u = f(x)$ is possible if $W''(\xi) \neq 0$.

Another more formal way of deriving (5) by inversion of (2) is to differentiate equation (3) with respect to ξ treating u,w, and x as functions of ξ. We have

$$f'(x)\frac{dx}{d\xi} + W'(\xi) = \frac{dx}{d\xi} + x \ .$$

But, since $\xi = f'(x)$ from (2), this reduces to

$$x = W'(\xi) \ ,$$

and the other part of (5) is obtained by substitution in (3).

The foregoing can be easily extended to the case of n independent variables. We have $u = f(x_1, \ldots x_n)$ and introduce as coordinates the direction numbers and intercept of the tangent plane to this surface

(6) $$\begin{cases} \xi_i = f_{x_i} \\ w = \sum_{j=1}^{n} x_j f_{x_j} - f \ . \end{cases}$$

The paramters x_i can be eliminated if the Hessian $|f_{x_i x_j}| \neq 0$,* (i.e. the Jacobian of $\xi_i = f_{x_i}$ does not vanish) yielding

*The condition $|f_{x_i x_j}| = 0$ means that the family of tangent planes is a smaller than n parameter family. For example, with n=2 there is a one parameter family, i.e. a developable, or a zero parameter family, a plane.

$$w = W(\xi_1, \ldots \xi_n) \ .$$

As before, the symmetric relation

$$u + w = \sum_{i=1}^{n} x_i \xi_i \tag{7}$$

holds, and differentiation with respect to ξ_j gives

$$\sum_{i=1}^{n} f_{x_i} \frac{\partial x_i}{\partial \xi_j} + W_{\xi_j} = x_j + \sum_{i=1}^{n} \frac{\partial x_i}{\partial \xi_j} \xi_i$$

which, on using $\xi_i = f_{x_i}$, and then substituting in (7) results in

$$\begin{cases} x_j = W_{\xi_j} \\ u = \sum_{i=1}^{n} \xi_i W_{\xi_i} - W \ . \end{cases} \tag{8}$$

The Legendre transformation is most frequently useful in transforming partial differential equations in which derivatives appear in a more complicated form than the independent variables

2. <u>The Distance Function - Reduction to Canonical Form.</u> In this chapter we will be concerned with the variational problem represented by the integral

$$I(u_1, \ldots u_n) = \int_t^{t'} F(s;\ u_1, \ldots u_n;\ u_1', \ldots u_n')\, ds \tag{9}$$

of which the Euler equations are

$$F_{u_i} - \frac{d}{ds} F_{u_i'} = 0 \ . \tag{10}$$

The solutions of (10) (extremals) represent families of curves $u_i(s)$ in the n+1 dimensional space $(s, u_1, \ldots u_n)$, it being possible to impose on each curve 2n conditions (if we ignore degeneracy). Thus we can consider a single curve passing

through each point (τ,k) of the space (i.e. $s = \tau$, $u_1 = k_1$) and satisfying there the intial conditions $u_i' = k_i'$, or consider an extremal to connect every pair of points (τ,k) (t,q) in some neighborhood (boundary value problem), or we can consider an n parameter family of extremals passing through the point (τ,k), leaving open the remaining n conditions. If we restrict the path of integration of (9) to be always taken along extremals then a unique "distance"* I is associated with every pair of points (τ,k), (t,q) which can be connected by an extremal, thus defining I as a function of the $2n + 2$ variables $(\tau; k_1,\ldots k_n; t; q_1,\ldots q_n)$:

$$(11) \qquad I(\tau,k,t,q) = \int_{\tau}^{t} F(s,u,u')ds \ .$$

An essential concept in the theory of sufficient conditions is that of a field which is defined as follows: a family of curves in a neighborhood is said to form a field if one and only one curve of the family passes through each point of the neighborhood. In particular an n parameter family of extremals through a point can form a field, and in general to form a field in n+1 dimensional space an n parameter family of curves is required. If we have a neighborhood covered by a field of extremals through a point, then by use of (11) it is possible to define a single valued function I over the neighborhood. In this case it will be convenient to work in the n+2 dimensional (I,s,u) space.

As an example consider

$$(12) \qquad I = \int_{\tau}^{t} \sqrt{1 + (\frac{du}{ds})^2}\ ds \ ,$$

*If F is a quadratic form in u_i' then the extremals are geodesics and I is actually distance. More generally F can be considered a metric in a "Finsler Space" with Finsler distance I.

for the length of a curve u(s), for which the Euler equation is $\frac{d^2u}{ds^2} = 0$. Here n = 1 and we have a two parameter family of straight lines as extremals, joining every pair of points in the (s,u) plane. If (τ,k) is fixed we have a field consisting of all rays through (τ,k) and $I(t,q)$ is a cone with vertex at $(0,\tau,k)$ in the three dimensional (I,t,q) space.

If in (11) we hold (τ,k) fast, I is defined as a function of (t,q) in a neighborhood of (τ,k) and this function can then be considered to depend on the n+1 parameters (τ,k). By the usual method of elimination of parameters a partial differential equation satisfied by this n+1 parameter family of functions can be found, namely:

$$(13) \qquad \begin{cases} I_t = F(t,q,q') - \sum_{i=1}^{n} q_i' F_{q_i'}(t,q,q') \\ I_{q_i} = F_{q_i'}(t,q,q') \end{cases}$$

which is a relation between the n+1 derivatives I_t, I_q expressed implicitly in terms of n paramters $q_1', \ldots q_n'$. We observe that the formalism for eliminating the parameters q' is exactly that of the Legendre Transformation (cf. (6)). Rather than proceeding in this way, we shall impose the Legendre transformation at the outset, in (9), thereby obtaining the reduction of (9) and (10) to canonical form. We observe, however, that the proper Legendre transformation is applied in terms of u' (i.e. q') which would necessitate subjecting u to a very complicated transformation. In order to avoid this difficulty we introduce the functions $p_i = u_i'$ and consider the variational problem

$$(14) \qquad I(u_1,\ldots u_n; p_1,\ldots p_n) = \int_{\tau}^{t} F(s; u_1,\ldots u_n; p_1,\ldots p_n)ds$$

subject to the side conditions

$$p_i = u_i' \quad .$$

This suggests the use of the Lagrange multiplier rule by which (14) becomes

$$I(u,p) = \int_{\tau}^{t} [F(s,u,p) + \sum_{i=1}^{n} \lambda_i (u_i' - p_i)]ds$$

for which the Euler equations are

$$F_{u_i} - \frac{d}{ds} \lambda_i = 0, \quad F_{p_i} - \lambda_i = 0 ,$$

and elimination of λ_i finally leads us to consider

$$(15) \qquad I(u,p) = \int_{\tau}^{t} [F(s,u,p) + \sum_{i=1}^{n} F_{p_i} (u_i' - p_i)]ds ,$$

together with the necessary restriction on F that $|F_{p_i p_j}| \neq 0$. Since the Lagrange multiplier rule was not proved for differential equations as side conditions, we accept (15) and show independently that it is equivalent to (9). By equivalence we mean that, although the function space over which I(u,p) is defined is much wider (2n independent functions u and p rather than n functions u), the extremals for both integrals are the same. Writing down the Euler equations of (15)

$$F_{u_j} + \sum_{i=1}^{n} F_{p_i u_j} (u_i' - p_i) - \frac{d}{ds} F_{p_j} = 0$$

$$F_{p_j} + \sum_{i=1}^{n} F_{p_i p_j} (u_i' - p_i) - F_{p_j} = \sum_{i=1}^{n} F_{p_i p_j} (u_i' - p_i) = 0 .$$

But, since $|F_{p_i p_j}| \neq 0$, the second equation implies $u_i' - p_i = 0$, and substitution into the first yields

$$F_{u_j} - \frac{d}{ds} F_{u_j'} = 0 ,$$

which is the set of Euler equations for (9). It is interesting to note that the variational problem (15) is degenerate, 2n

first order equations replacing n of the second order.

In (15) the u's and p's are independent, so the Legendre transformation

$$(16)\qquad \begin{cases} v_i = F_{p_i}(s,u,p) \\ L(s,u,v) = \sum_{i=1}^{n} p_i v_i - F(s,u,p) \end{cases}$$

can be applied treating u and s as parameters. The dependence of L on s,u, and v is given by eliminating p from (16). The condition for elimination is satisfied in virtue of our assumption $|F_{p_i p_j}| \neq 0$. We now have

$$(17)\qquad I(u,v) = \int_{\tau}^{t} [\sum_{i=1}^{n} u_i v_i - L(s,u,v)]ds$$

which is again a degenerate variational problem in the 2n functions $(u_1,\ldots u_n;\ v_1,\ldots v_n)$. The Euler equations of (17) take the canonical form

$$(18)\qquad \begin{cases} \frac{d}{ds} v_i + L_{u_i} = 0 \\ \frac{d}{ds} u_i - L_{v_i} = 0 \ . \end{cases}$$

These transformations have their origin in classical mechanics, the equations (10) being Lagrange's equations of motion for a conservative system in the generalized coordinates u_i and momenta p_i (or v_i), F being the difference between kinetic and potential energies, and L their sum. The canonical form (18) can of course be derived directly from (10) without reference to a variational integral.

3. <u>The Hamilton-Jacobi Partial Differential Equation.</u> We will now consider (17) in a neighborhood in which it is assumed that a unique extremal exists joining any two points (τ,k) (t,q), the conjugate function v taking on the values 1 and λ

(i.e. $v_i = l_i$ at $s = t$, $v_i = \lambda_i$ at $s = \tau$). The extremal passing through these points can be represented by

$$(19) \qquad \begin{cases} u_i = f_i(s,t,q,\tau,k) \\ v_i = g_i(s,t,q,\tau,k) \end{cases}$$

where

$$(20) \qquad \begin{cases} f_i(t,t,q,\tau,k) = q_i, & f_i(\tau,t,q,\tau,k) = k_i \\ g_i(t,t,q,\tau,k) = l_i, & g_i(\tau,t,q,\tau,k) = \lambda_i \end{cases}.$$

In order to find the partial differential equation satisfied by the function $I(t,q,\tau,k)$ we must calculate its partial derivatives, I_t, I_{q_i}, etc. and this can be done most concisely by use of the variational symbol δ. Let the independent variables be taken as functions of a parameter ε; viz. $t(\varepsilon)$, $\tau(\varepsilon)$, $q_i(\varepsilon)$, $k_i(\varepsilon)$, $l_i(\varepsilon)$, $\lambda_i(\varepsilon)$; I then becomes a function of ε. We have

$$\frac{dI}{d\varepsilon} = I_t \frac{dt}{d\varepsilon} + I_\tau \frac{d\tau}{d\varepsilon} + \sum_{i=1}^{n} I_{q_i} \frac{dq_i}{d\varepsilon} + \sum_{i=1}^{n} I_{k_i} \frac{dk_i}{d\varepsilon},$$

and letting $\left.\frac{dI}{d\varepsilon}\right|_{\varepsilon=0} = \delta I$, $\left.\frac{dt}{d\varepsilon}\right|_{\varepsilon=0} = \delta t$, etc. we have

$$(21) \qquad \delta I = I_t \delta t + I_\tau \delta t + \sum_{i=1}^{n} I_{q_i} \delta q_i + \sum_{i=1}^{n} I_{k_i} \delta k_i .$$

Performing the variation in

$$I(\varepsilon) = \int_{\tau(\varepsilon)}^{t(\varepsilon)} [\Sigma u_i'(\varepsilon) v_i(\varepsilon) - L(s,u(\varepsilon),v(\varepsilon))]ds$$

where

$$u_i(\varepsilon) = f_i(s,t(\varepsilon),q(\varepsilon),\tau(\varepsilon),k(\varepsilon))$$
$$v_i(\varepsilon) = g_i(s,t(\varepsilon),q(\varepsilon),\tau(\varepsilon),k(\varepsilon))$$

we get

$$\delta l = [\Sigma q_i' l_i - L(t,q,l)]\delta t - [\Sigma k_i' \lambda_i - L(\tau,k,\lambda)]\delta\tau$$
$$+ \int_\tau^t \sum_{i=1}^n (u_i' \delta v_i + v_i \delta u_i' - L_{u_i} \delta u_i - L_{v_i} \delta v_i) ds .$$

The terms $\Sigma(u_i' - L_{v_i})\delta v_i$ in the integral vanish by (18) (u and v are extremals), and on integrating by parts the remaining terms in the integral become

$$[\Sigma v_i \delta u_i]_\tau^t - \int_\tau^t \Sigma(v_i' + L_{u_i})\delta u_i ds = \sum_{i=1}^n [v_i \delta u_i]_\tau^t .$$

We now evaluate δu at $s = t$ and $s = \tau$. This is not equal to δq or δk because in δq, the variation is performed after s is set equal to t, that is after s has become a function of ε. Differentiating (19) we have

$$\delta u = f_t \delta t + f_q \delta q + f_\tau \delta\tau + f_k \delta k .* \tag{22}$$

However, differentiating the first relation in (20),

$$\delta q = f' \delta t + f_t \delta t + f_q \delta q + f_\tau \delta\tau + f_k \delta k , \tag{23}$$

where f' is the derivative of f with respect to its first argument, evaluated at $s = t$; i.e. $f' = q'$. By equating coefficients of the various δ's in (23) we see that $f_t = -q'$, $f_q = 1$, $f_\tau = f_k = 0$ evaluated at $s = t$. Similarly $f_\tau = -k'$, $f_k = 1$, $f_t = f_q = 0$ at $s = \tau$, and (22) reduces to

$$\delta u|_{s=t} = \delta q - q' \delta t$$
$$\delta u|_{s=\tau} = \delta k - k' \delta\tau ,$$

from which we get

$$\sum_{i=1}^n [v_i \delta u_i]_\tau^t = \Sigma(l_i \delta q_i - l_i q_i' \delta t - \lambda_i \delta k_i + \lambda_i k_i' \delta\tau)$$

*The subscript i is omitted.

and finally

$$(24) \qquad \delta I = \sum_{i=1}^{n} l_i \delta q_i - L(t,q,l)\delta t - \sum_{i=1}^{n} \lambda_i \delta k_i + L(\tau,k,\lambda)\delta\tau \ .$$

We now read off

$$(25) \qquad \begin{cases} I_t = -L(t,q,l) \ , & I_{q_i} = l_i \\ I_\tau = L(\tau,k,\lambda) \ , & I_{k_i} = -\lambda_i \ . \end{cases}$$

In particular if we assume the endpoint (τ,k) is fixed, I as a function of t and q satisfies the first order Hamilton-Jacobi partial differential equation

$$(26) \qquad I_t + L(t; q_1,\ldots q_n; I_{q_1},\ldots I_{q_n}) = 0$$

obtained by elimination of the l_i from (25).

Problem

Evaluate the partial derivatives (25) by direct differentiation rather than variation.

The equation (26) was derived from a function I having n+1 parameters, but its importance lies in the existence of other more general solutions, depending on arbitrary functions rather than on a number of parameters. Those solutions will be shown to represent distances measured from an initial surface rather than from an initial point as above. In order to motivate the manipulation to follow, we will first discuss briefly the general theory of partial differential equations of the first order, proofs being given only when necessary for our application.

First of all, it is clear that if from an n parameter family of solutions of any partial differential equation we form an n-1 parameter family by considering one of the parameters a_n as a function of the others, $a_n = f(a_1,\ldots a_{n-1})$, the envelope of this solution with respect to its n-1 parameters is also a

solution (since it has at every point the same tangent plane as a known solution), and in fact depends on an arbitrary function, namely f.

Now consider the simple case of a first order quasi-linear partial differential equation in two independent variables

$$au_x + bu_y = c \ , \tag{27}$$

a, b, and c being functions of x, y, and u. Geometrically (27) states that at every point of a solution u(x,y) the normal to the surface $u = u(x,y)$ (direction components $p = u_x$, $q = u_y$, -1) is perpendicular to the line element having components (a,b,c) at this particular point. In other words the element of surface contains the line element (a,b,c). If we integrate the ordinary differential equations

$$\frac{dx}{ds} = a(x,y,u), \quad \frac{dy}{ds} = b(x,y,u), \quad \frac{dz}{ds} = c(x,y,u) \tag{28}$$

we get a two parameter field of characteristic curves in the x,y,u space. Any solution of (27) is a one parameter family of characteristics, and conversely. Given a non-characteristic initial curve $x = x(s)$, $y = y(s)$, $u = u(s)$, the family of characteristics intersecting this line traces out an integral surface. In this way we have reduced the problem of the solution of the partial differential equation (27) to that of the integration of the three ordinary differential equations (28). It is clear that two integral surfaces can intersect only along characteristics, since to any common point belongs the whole characteristic through that point, and conversely this property completely characterizes the characteristic curves.

Let us now consider the more general equation

$$F(x,y,u,p,q) = 0 \ . \tag{29}$$

We note that for a linear equation (a,b,c in (27) functions of x and y alone), the characteristics are independent of u and form a one parameter family in the x,y plane. From this we would expect that the characteristics in the general case (29)

would depend not only on u but also on p and q, and would not be fixed curves in the x,y,u space. In order to reduce this problem to that of ordinary differential equations it would therefore seem reasonable to consider curves x(s), y(s), u(s), p(s), q(s) in five dimensions. In the x,y,u space such a curve represents a space curve x(s), y(s), u(s) together with a surface element p(s), q(s) associated with each point. At a given point (x,y,u) (29) allows a one parameter relation between p and q. Geometrically this means the surface element (tangent plane) traces out a one parameter family passing through the point, and envelopes a cone (in the quasilinear case the cone degenerates into a line, with a one parameter family of planes through the line). The line elements of these cones are the characteristic directions and are given by the differential equations*

$$(30) \qquad \frac{dx}{ds} = F_p \ , \quad \frac{dy}{ds} = F_q \ , \quad \frac{du}{ds} = pF_p + qF_q \ .$$

A curve which is a solution of the system (29) (subject to (30)) has a characteristic direction at every point, but a family of such curves is not necessarily an integral of (29). In order that these curves lie on an integral surface p and q must satisfy the two further relations

$$(31) \qquad \frac{dp}{ds} = -(pF_u + F_x) \qquad \frac{dq}{ds} = -(qF_u + F_y) \ .$$

Using (30) and (31) we have five ordinary differential equations in the five variables x(s), y(s), u(s), p(s), q(s). In general a solution curve will exist passing through an arbitrary initial point x_o, y_o, u_o, p_o, q_o. There is a five parameter family of solutions, but since s admits of an arbitrary translation, and (29) is a relation between the initial values, the solutions reduce to three parameters. Interpreted in the (x,y,u) space, an initial value is a point (x,y,u) and a surface element (p,q),

*For proofs of the following, see Courant-Hilbert, Vol. II, Chapter I, par. 3.

where (p,q) is not arbitrary but must satisfy (29) (i.e. must be tangent to a cone). Given such an initial value a characteristic curve x(s), y(s), u(s) is determined, and with it tangent elements p(s), q(s) comprising a characteristic strip. As in the quasilinear case integral surfaces are composed of one parameter families of characteristic curves or strips. Given an initial curve x(s), y(s), u(s) we use the relation

$$\frac{du}{ds} = \frac{du}{dx}\frac{dx}{ds} + \frac{du}{dy}\frac{dy}{ds} = p\frac{dx}{ds} + q\frac{dy}{ds}$$

together with (29) to determine the remaining initial values p(s) and q(s), the integral surface being given then by use of (30) and (31). Here as before the characteristics represent intersections of integral surfaces.

The general partial differential equation in n independent variables

$$F(x_1,\cdots,x_n; u; p_1,\cdots,p_n) \tag{32}$$

is somewhat more general than equation (26) with which we are concerned, since the dependent variable, I, is absent in (26)*, but this difference is unessential for, if instead of looking for a solution of (32) having the form $u(x_1,\cdots x_n)$ we consider it to be given implicitly as

$$\phi(u; x_1,\cdots x_n) = 0 ,$$

we can rewrite (32) as a differential equation in the n+1 independent variables u, $x_1,\cdots x_n$. We have

$$\sum_{i=1}^{n} \phi_{x_i} dx_i + \phi_u du = 0$$

and

$$du = \sum_{i=1}^{n} u_{x_i} dx_i = \sum_{i=1}^{n} p_i dx_i$$

---

*Here u corresponds to I and $(x_1,x_2,\cdots,x_n)$ to $(t,q_1,\cdots,q_n)$.

from which we obtain immediately

$$p_i = -\frac{\phi_{x_i}}{\phi_u} ,$$

equation (32) becoming

$$(33) \qquad F(x_1, \ldots x_n;\ u;\ -\frac{\phi_{x_1}}{\phi_u}, \ldots \frac{\phi_{x_n}}{\phi_u}) = 0 ,$$

in which the dependent variable ϕ no longer appears. If we consider the implicit relation (33) solved for ϕ_u and replace ϕ by I, u by t, and x_i by u_i we get exactly

$$(26) \qquad I_t + L(t;\ u_1, \ldots u_n;\ I_{u_1}, \ldots I_{u_n}) = 0 .$$

The characteristic differential equations for (32) are

$$(34) \qquad \frac{dx_i}{ds} = F_{p_i} , \quad \frac{du}{ds} = \sum_{i=1}^{n} p_i F_{p_i} , \quad \frac{dp_i}{ds} = -(F_{x_i} + F_u p_i)$$

where we have 2n+1 differential equations in the 2n+1 variables $x_i(s)$, $u(s)$, $p_i(s)$. The corresponding characteristic differential equations for (26) take the form

$$(35) \qquad \frac{du_i}{dt} = L_{v_i} , \quad \frac{dv_i}{dt} = -L_{u_i}$$

$$\frac{dI}{dt} = \sum_{i=1}^{n} v_i L_{v_i} - L , \quad \frac{dI_t}{dt} = -L_t$$

where we have set $v_i = I_{u_i}$ and taken t for the parameter along the characteristic curves. Here there are 2n+2 differential equations for the 2n+2 variables $u_i(t)$, $I(t)$, $I_{u_i}(t) = v_i(t)$, and $I_t(t)$. However, we see that since I and I_t do not appear in L, the 2n equations

$$(18) \qquad \frac{du_i}{dt} = L_{v_i} \qquad\qquad \frac{dv_i}{dt} = -L_{u_i}$$

can be treated as a system in itself, independent of the other two. But these are exactly the equations for the extremals of our variational problem. We thus see that the extremals of the integrals (14) and (15) are given by the projection in the (u,t) space of the characteristics of the corresponding Hamilton-Jacobi partial differential equation (which are in general curves in the (I,u,t) space). In the example

$$(12) \qquad I = \int_{\tau}^{t} \sqrt{1 + \left(\frac{du}{ds}\right)^2}\, ds$$

the Hamilton-Jacobi equation is

$$I_t^2 + I_u^2 = 1 ,$$

of which the characteristics are straight lines in the t,u,I space making angles of 45° with the t,u plane, and the projections (extremals) are the previously found straight lines in the t,u plane.

We now come to our chief result that from a complete solution of the Hamilton-Jacobi equation we can construct all the extremals of our variational problem. By a complete solution is meant a solution

$$(36) \qquad I(t; u_1,\ldots u_n; a_1,\ldots a_n) - a \qquad |I_{u_i a_i}| \neq 0$$

depending on the n+1 parameters a_i,a (the n+1 st parameter a is additive since only derivatives of I appear in the differential equation). The envelope of the n parameter family resulting from setting $a = f(a_1,\ldots a_n)$ is obtained from

$$(37) \qquad I_{a_i}(t; u_1,\ldots u_n; a_1,\ldots a_n) - f_{a_i}(a_1,\ldots a_n) = 0$$

$$(i=1,\ldots n) ,$$

by eliminating the paramters a_i in (36) and (37). For each value of the a_i the intersection of (36) and (37) is a characteristic, (the integral surface given by the envelope is traced

out by allowing the a_i to take all their values) so the projection of this characteristic in the t,u space namely (37), is the equation of an extremal (represented in (37) as the intersection of n surfaces). Since the function f is arbitrary the quantities $f_{a_i} = b_i$ can be given arbitrary values, and since $|I_{a_i u_i}| \neq 0$ the n equations (37) can be solved for u_i giving

$$u_i = u_i(t; a_1, \ldots a_n; b_1, \ldots b_n) \tag{38}$$

which is the required 2n parameter family of extremals. The theory of characteristics has been used only for motivation, and we will now prove this statement independently.

If from a complete solution $I(t; u_1, \ldots u_n; a_1, \ldots a_n)$, $|I_{u_i a_j}| \neq 0$, of the equation

$$I_t + L(t, u, I_u) = 0 , \tag{26}$$

we define the functions $u_i(t; a_1, \ldots a_n; b_1, \ldots b_n)$ implicitly by

$$I_{a_i}(t; u_1, \ldots u_n; a_1, \ldots a_n) = b_i, \quad i=1, \ldots n \tag{39}$$

and the conjugate functions $v_i(t; a_1, \ldots a_n; b_1, \ldots b_n)$ by

$$v_i = I_{u_i}(t; u_1, \ldots u_n; a_1, \ldots a_n) , \tag{40}$$

the u_i here to be replaced by their values from (39), we get a 2n parameter family of extremals satisfying the canonical equations $du_i/dt = L_{v_i}$, $dv_i/dt = -L_{u_i}$.

Considering I as a function of t, u_i, and a_i, and with u_i and v_i functions of t, a_i, and b_i we differentiate (26) with respect to a_i (t and u_i constant) and (39) with respect to t (a_i and b_i constant) obtaining

$$I_{ta_i} + \sum_{j=1}^{n} L_{v_j} I_{u_j a_i} = 0$$

$$I_{a_i t} + \sum_{j=1}^{n} I_{a_i u_j} \frac{\partial u_j}{\partial t} = 0$$

Subtracting and remembering that $|I_{a_i u_j}| \neq 0$, we have $du_j/dt = L_{v_j}$, the total derivative du_j/dt implying that a_i and b_i are held constant. Similarly differentiating (40) with respect to t and (26) with respect to u_i we have

$$\frac{dv_i}{dt} = I_{u_i t} + \sum_{j=1}^{n} I_{u_i u_j} \frac{du_j}{dt}$$

$$I_{tu_i} + L_{u_i} + \sum_{j=1}^{n} L_{v_j} I_{u_j u_i} = 0 .$$

Subtracting we obtain

$$\frac{dv_i}{dt} + L_{u_i} = \sum_{j=1}^{n} I_{u_j u_i} \left(\frac{du_j}{dt} - L_{v_j}\right) = 0 ,$$

which completes the proof.

We will now proceed to show that while the Hamilton-Jacobi equation was derived by considering distance measured from a point in the (t,u) space, an integral of the equation in general represents distance measured from a surface in the (t,u) space. This becomes clear geometrically if we take for our complete solution (36) the integral $I(t,u,\tau,k)$ representing distance from the point (τ,k), the (τ,k) being n+1 parameters. An envelope 1' is constructed by assuming some relation $f(\tau; k_1,..,k_n) = 0$, and since $I = 0$ for every point (τ,k) satisfying $f = 0$, also $I' = 0$ on the surface $f = 0$. Further the envelope of "spheres" $I = C$ is the surface $I' = C$, so that the latter locus represents those points (t,u) which are at a distance C from the initial surface $f(\tau,k) = 0$.

To make this more precise we define the distance from a given point (t,q) to a surface $T(\tau,k) = 0$, $T_\tau \neq 0$, to be the minimum distance measured along all extremals through the point which intersect the surface. This is essentially the problem of the free boundary, one endpoint (t,q) being fixed and the other lying on the surface $T(\tau,k) = 0$. The condition that the distance I from the point (t,q) to an arbitrary point (τ,k) on

$T = 0$ be stationary with respect to variations of (τ,k) is given by

$$\delta I = L(\tau,k,\lambda)\delta\tau - \sum_{i=1}^{n} \lambda_i \delta k_i = 0 \tag{41}$$

using (24), while we have

$$T_\tau \delta\tau + \sum_{i=1}^{n} T_{k_i} \delta k_i = 0 \tag{42}$$

since (τ,k) is constrained to lie on $T = 0$. Eliminating $\delta\tau$ between (41) and (42) (using $T_\tau \neq 0$) we have

$$\delta I = -\sum_{i=1}^{n} \left(\frac{L(\tau,k,\lambda)}{T_\tau} T_{k_i} - \lambda_i\right)\delta k_i = 0 ,$$

and since the δk_i are independent,

$$\frac{\lambda_i(\tau,k)}{T_{k_i}(\tau,k)} = \frac{L(\tau,k,\lambda)}{T_\tau(\delta,k)} \quad . \tag{43}$$

The n <u>transversality</u> conditions (43) in general serve to select one or more extremals from the point (t,q) to the surface $T = 0$, and express a relation connecting the tangent plane to the surface (through the T_{k_i}) and the slope of the extremal where it intersects the surface (through the λ_i, which are equivalent to the k_i). Equation (43) is a direct generalization of equation (62) on p. 29 of Chapter I, and in the case of geodesics where I is actually distance, reduces to orthogonality. Dispensing with the external point (t,q), the n conditions (43) in general serve to determine a unique extremal from the n parameter family through each point on the surface $T = 0$, thereby forming a <u>field of extremals transverse to the surface</u> $T = 0$,* at least

*Since only derivatives of T appear in (43), transversality for a surface $T = 0$ implies transversality for the family of surfaces $T =$ constant.

in some neighborhood of the (t,u) space surrounding a region of the surface. In such a neighborhood we can define a single valued function $\bar{I}(t,u)$ taking the value zero on $T = 0$. If we consider a curve $t = t(\varepsilon)$, $u_i = u_i(\varepsilon)$ lying in this field, we have an extremal intersecting the curve at each point, with initial values $\tau(\varepsilon)$, $k_i(\varepsilon)$ lying on $T = 0$. The corresponding distance function $I(t,u,\tau,k)$ as a function of $2n+2$ variables satisfies equation (24). The variation δI (which is equal to $\delta\bar{I}$ for the particular variation we have chosen) can be considered the sum of two variations, one with (t,u) fixed, and the other with (τ,k) fixed. The former vanishes in virtue of the transversality condition (41), so that we have

$$\delta\bar{I} = \sum_{i=1}^{n} l_i \delta u_i - L(t,u,l)\delta t \tag{44}$$

from which follows $\bar{I}_t = -L(t,u,l)$ and $\bar{I}_{u_i} = l_i$, and by elimination of l_i, $\bar{I}_t + L(t,u,\bar{I}_u) = 0$, so that the function $\bar{I}$ which we have constructed as the distance from an arbitrary surface $T(\tau,k) = 0$ satisfies the Hamilton-Jacobi equation.

The converse fact that any solution of the Hamilton-Jacobi equation represents the distance from a surface (or, as a degeneration, from a point) is readily verified. From a given integral $l(s,u)$ we construct the n ordinary differential equations

$$\frac{du_i}{ds} = L_{v_i}(s,u_i,v_i) \tag{45}$$

where we have used the notation

$$v_i = I_{u_i}(s,u) \, . \tag{46}$$

The solution of (45) subject to the initial conditions $u_i = q_i$ at $s = t$ are

$$\begin{cases} u_i = u_i(s;\, q_1,\ldots q_n) \, , \\ v_i = v_i(s;\, q_1,\ldots q_n) \, , \end{cases} \tag{47}$$

where the values of u_i have been substituted in (46) to get v_i. Differentiating (46) with respect to s, q being constant

$$\frac{dv_i}{ds} = I_{u_i s} + \sum_{j=1}^{n} I_{u_i u_j} \frac{du_j}{ds} = I_{u_i s} + \sum_{j=1}^{n} I_{u_i u_j} L_{v_j} .$$

Differentiating $I_s + L(s,u,v) = 0$ with respect to u_i we have

$$I_{su_i} + L_{u_i} + \sum_{j=1}^{n} L_{v_j} \frac{\partial v_j}{\partial u_i} = I_{su_i} + L_{u_i} + \sum_{j=1}^{n} L_{v_j} I_{u_j u_i} = 0 ,$$

and subtracting from the previous result

$$\frac{dv_i}{ds} - L_{u_i} = \sum_{j=1}^{n} I_{u_i u_j} \left(\frac{du_j}{ds} - L_{v_j}\right) = 0 .$$

The constructed functions u and v are therefore an n parameter family of extremals. Furthermore, they are transversal to the surface I = const. For from

$$I_s + L(s,u,v) = 0 \quad \text{and} \quad V_i = I_{u_i}(s,u)$$

we have immediately

$$\frac{v_i(s,u)}{I_{u_i}(s,u)} = \frac{L(s,u,v)}{I_s(s,u)}$$

which are exactly the transversality conditions (43) for the surface $l(s,u) = \text{const}$. It remains to be shown that I is actually the distance function along this family of transverse extremals. Taking $I(\tau,k) = 0$, we have

$$I = \int_{\tau}^{t} \frac{dI}{ds} ds = \int_{\tau}^{t} \left(I_s + \sum_{i=1}^{n} I_{u_i} \frac{du_i}{ds}\right) ds = \int_{\tau}^{t} \left(\sum_{i=1}^{n} v_i u_i' - L(s,u,v)\right) ds$$

which is exactly (17).

The foregoing concepts can be interpreted in terms of the propagation of light through a medium having a variable index of refraction. The light rays (extremals) are given as paths of least time (I is a minimum). The construction of solutions

as envelopes is exactly Huygens's Principle for the construction of wave fronts.

4. The two body problem. We consider the problem of determining the motion of two bodies of mass m_1 and m_2 acted on only by the Newtonian gravitational force between them,

$$F = \frac{Gm_1m_2}{(r_1 + r_2)^2} ,$$

where r_1 and r_2 are the distances of m_1 and m_2 from the center of mass C which we may consider fixed. From the relation $m_1r_1 = m_2r_2$, we have

$$F = -\frac{Gm_1m_2'}{r_1^2} \quad \text{where } m_2' = \frac{m_2^3}{(m_1 + m_2)^2}$$

so that the problem is reduced to that of a fixed mass m_2' attracting a mass m_1 at a distance r_1. With C as origin we take coordinates x and y in the plane determined by the initial position and velocity of m_1. The motion is described as making the integral (T-V)dt stationary, where T and V are the kinetic and potential energies respectively. Taking $m_1 = 1$ we have

$$T = (\dot{x}^2 + \dot{y}^2)/2$$
$$V = -k^2/\sqrt{x^2 + y^2}$$
$$F = T - V = (\dot{x}^2 + \dot{y}^2)/2 + k^2/\sqrt{x^2 + y^2} .$$

Reducing the canonical form we use the notation $\dot{x} = p$, $\dot{y} = q$ and have $F_{\dot{x}} = p$, $F_{\dot{y}} = q$ yielding

$$L(t,x,y,p,q) = \tfrac{1}{2}(p^2 + q^2) - k^2/\sqrt{x^2 + y^2}$$

for the Hamilton function and

$$\phi_t + \tfrac{1}{2}(\phi_x^2 + \phi_y^2) - k^2/\sqrt{x^2 + y^2} = 0$$

for the Hamilton-Jacobi equation. Changing to polar coordinates, $x = r \cos \theta$, $y = r \sin \theta$, this reduces to

$$\phi_t + \frac{1}{2} \left(\phi_r^2 + \frac{1}{r^2} \phi_\theta^2\right) = \frac{k^2}{r} \quad .$$

A two paramter solution can be found by writing ϕ as the sum of three functions of t, r, and θ respectively, in particular

$$\phi = \alpha t + \beta\theta + R(r) \quad .$$

Substituting in the differential equation we solve for R obtaining

$$\phi = \alpha t + \beta\theta + \int_{r_o}^{r} \sqrt{\frac{2k^2}{\rho} - \frac{\beta^2}{\rho^2} - 2\alpha} \; d\rho \quad .$$

To solve for the extremals we differentiate with respect to the parameters obtaining

$$t - \int_{r_o}^{r} \frac{d\rho}{\sqrt{\frac{2k^2}{\rho} - \frac{\beta^2}{\rho^2} - 2\alpha}} = t_o$$

$$\theta - \beta \int_{r_o}^{r} \frac{d\rho}{\rho^2 \sqrt{\frac{2k^2}{\rho} - \frac{\beta^2}{\rho^2} - 2\alpha}} = \theta_o \quad ,$$

t_o and θ being arbitrary constants. The second equation gives θ as a function of r, i.e. the particle path, while the first gives r as a function of time. The second equation can be integrated using the substitution $\rho = 1/\sigma$ giving

$$\theta = \theta_o - \arcsin \frac{\frac{\beta^2}{k^2 r} - 1}{\sqrt{1 + \frac{2\alpha\beta^2}{k^4}}}$$

Writing $\delta = \frac{\beta^2}{k^2}$ and $\varepsilon^2 = \sqrt{1 + \frac{2a\beta^2}{k^4}}$ we have the conic of eccentricity ε

$$r = \frac{\delta}{1 - \varepsilon^2 \sin(\theta - \theta_0)} \quad .$$

5. The Homogeneous Case - Geodesics. Up to now it has been assumed that $\left|F_{u_i' u_j'}\right| \neq 0$, but this condition excludes the important case where F is homogeneous of first order in u_i'. For this case we have $\left|F_{u_i' u_j'}\right| \equiv 0$, since if F is homogeneous of first order, $F_{u_i'}$ is homogeneous of zero-th order, and applying Euler's homogeneity condition

$$\sum_{j=1}^{n} u_j' F_{u_i' u_j'} = 0 \ ,$$

we see the determinant $\left|F_{u_i' u_j'}\right| \equiv 0$. If in addition to being homogeneous of first order in u_i', F is independent of s (which is the case for variational problems in parametric form), the problem can be reduced to the form already discussed by taking u_n for the variable of integration s. We then have

$$(48) \qquad J = \int_{E}^{n} F\left(u_1, \ldots u_n; \frac{du_1}{du_n}, \ldots \frac{du_{n-1}}{du_n}, 1\right) du_n \ ,$$

as our integral, and the Hamilton-Jacobi equation becomes

$$(49) \qquad J_{u_n} + L\left(u_1, \ldots u_{n-1}; J_{u_1}, \ldots J_{u_{n-1}}\right) = 0$$

where

$$L\left(u_1, \ldots u_{n-1}; v_1, \ldots v_{n-1}\right) = \sum_{i=1}^{n-1} p_i v_i - F\left(u_1, \ldots u_{n-1}; p_1, \ldots p_{n-1}\right)$$

with $p_i = \frac{du_i}{du_n}$, $v_i = F_{p_i}$. The equation (49) is nothing more than the homogeneity condition, since from

$$\sum_{i=1}^{n} \bar{p}_i F_{\bar{p}_i} - F = 0 , \qquad \bar{p}_i = \frac{du_i}{dt} ,$$

substituting $p_i = \bar{p}_i / \bar{p}_n$ we have

$$(50) \quad F_{\bar{p}_n} + \sum_{i=1}^{n-1} p_i F_{\bar{p}_i} - F(u_1, \ldots u_n; p_1, \ldots p_{n-1}, 1) = 0 .$$

The expressions (25) for the derivatives J_t, J_{u_i} are true even with $\left| F_{u_i' u_j'} \right| = 0$,* and if we substitute J_{u_i} for $F_{\bar{p}_i} = v_i$ (i.e. 1_i) in (50) we get exactly (49).

Let us illustrate another method of attack with the case of geodesics on an n dimensional manifold. We have

$$(51) \qquad I = \int_{\tau}^{t} \sqrt{Q} \, ds ,$$

where $Q = \sum_{i,k=1}^{n} g_{ik} u_i' u_k'$, the g_{ik} being functions of the u_i. From (25) we have

$$(52) \qquad \begin{cases} I_t = 0 \\ I_{q_i} = F_{u_i'} \Big|_{s=t} = \frac{1}{\sqrt{Q}} \sum_{k=1}^{n} g_{ik} q_k' . \end{cases}$$

If we denote by (g^{ik}) the matrix reciprocal to (g_{ik}) (i.e. $\sum_{j=1}^{n} g_{ij} g^{jk} = \delta_i^k$), solving (52) for q_k' gives

$$(53) \qquad \frac{1}{\sqrt{Q}} q_k' = \sum_{i=1}^{n} g^{ik} I_{q_i} .$$

---

*This condition does not prevent the use of the Legendre transformation, but only invalidates its inversion; in other words the extremals of (9) are included in those of (17), and this fact is all that is needed in deriving (25).

From the homogeneity relation

$$\sum_{k=1}^{n} q_k' F_{q_k'} = \sum_{k=1}^{n} q_k' I_{q_k} = \sqrt{Q} \quad ,$$

so that multiplying through (53) by I_{q_k} and summing with respect to k we obtain

$$\sum_{i,k=1}^{n} g^{ik} I_{q_i} I_{q_k} = 1 \tag{54}$$

which is a partial differential equation in the I_{q_i} taking the place of the Hamilton-Jacobi equation. The connection between (54) and the Hamilton-Jacobi equation is clarified by a consideration of the Euler equations of (51). We have

$$\frac{d}{ds}\left(\frac{\partial\sqrt{Q}}{\partial u_i'}\right) - \frac{\partial\sqrt{Q}}{\partial u_i} = 0$$

or

$$\frac{d}{ds}\left(\frac{\partial Q}{\partial u_i'} \cdot \frac{1}{\sqrt{Q}}\right) - \frac{\partial Q}{\partial u_i} \cdot \frac{1}{\sqrt{Q}} = 0 \ .$$

If for any admitted curve (and in particular for the extremals) we take the parameter s to be proportional to the arc length then Q is a constant, and (55) reduces to

$$\frac{d}{ds}\left(\frac{\partial Q}{\partial u_i'}\right) - \frac{\partial Q}{\partial u_i} = 0 \tag{56}$$

for the function $Q(u_1, \ldots u_n; u_1', \ldots u_n') = C$. Now, (56) suggests that we look for the extremals of the integral

$$J = \int_{\tau}^{\sharp} Q ds \tag{57}$$

subject to the condition $Q = C$, where Q is homogeneous but of <u>second</u> order, so the Hamilton-Jacobi theory can be applied directly. We have

$$Q = \sum_{i,k=1}^{n} g_{ik} p_i p_k \quad , \qquad p_i = u_i'$$

$$v_i = Q_{p_i} = 2 \sum_{k=1}^{n} g_{ik} p_k \quad .$$

Solving for p_k,

$$p_k = \frac{1}{2} \sum_{i=1}^{n} g^{ik} v_i \quad ,$$

$$L = \sum_{i=1}^{n} p_i v_i - Q$$

$$= Q$$

$$= \sum_{i,k=1}^{n} g_{ik} p_i p_k$$

$$= \frac{1}{4} \sum_{i,k=1}^{n} g^{ik} v_i v_k \tag{58}$$

and the differential equation is

$$J_t + \frac{1}{4} \sum_{i,k=1}^{n} g^{ik} J_{u_i} J_{u_k} = 0 \quad . \tag{59}$$

In order to find an integral of (59) containing n parameters we try a solution in the form $J(t; u_1, \ldots u_n) = f(t) + J^*(u_1, \ldots u_n)$ and it follows immediately that $f'(t)$ must be independent of t, i.e. $f(t) = \alpha t$. J^* then satisfies the equation

$$\alpha + \frac{1}{4} \sum_{i,k=1}^{n} g^{ik} J_{u_i} J_{u_k} = 0 \quad .$$

This differential equation in general allows an n parameter solution, so α may be given the specific value $-1/4$, and we have

$$(60) \qquad \sum_{i,k=1}^{n} g^{ik} J_{u_i} J_{u_k} = 1 \quad ,$$

which is (54) again.

From (58) it is apparent that Q is constant (viz. $Q = 1/4$) for any solution of (60), so that any n parameter integral of (60) can be used to obtain the extremals of (51) as well as those of (57). Using the methods outlined here it is possible to obtain the geodesics on an ellipsoid. The surfaces

$$(61) \qquad \frac{x^2}{a^2 + s} + \frac{y^2}{b^2 + s} + \frac{z^2}{c^2 + s} = 1$$

in the paramter s, where $a > b > c$ are given constants, represent confocal ellipsoids, one-sheeted hyperboloids, and two-sheeted hyperboloids as $-c^2 < s$, $-b^2 < s < -c^2$, $-a^2 < s < -b^2$ respectively. At a given point (x,y,z) (61) is a cubic in s with three real roots (s_1, s_2, s_3) representing three mutually orthogonal surfaces one of each type, through the point (x,y,z).* The parameters (s_1, s_2, s_3) can be taken as a new coordinate system, and in particular, taking our given ellipsoid as $s_1 = 0$, the remaining two coordinates can be used as parameters on the ellipse, with the transformation formulas (setting $-s_1 = u$, $-s_2 = v$)

$$x = \sqrt{\frac{a(a-u)(a-v)}{(a-b)(a-c)}}$$

$$y = \sqrt{\frac{b(b-u)(v-b)}{(b-c)(a-b)}}$$

$$z = \sqrt{\frac{c(u-c)(v-c)}{(a-c)(b-c)}} \quad .$$

The values of

$$g_{11} = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2 \quad ,$$

$$g_{12} = g_{21} = \left(\frac{\partial x}{\partial u}\right)\left(\frac{\partial x}{\partial v}\right) + \left(\frac{\partial y}{\partial u}\right)\left(\frac{\partial y}{\partial v}\right) + \left(\frac{\partial z}{\partial u}\right)\left(\frac{\partial z}{\partial v}\right) \quad ,$$

---

*For details see Courant-Hilbert, Vol. I, p.195.

$$g_{22} = \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2 \quad ,$$

become

$$g_{11} = (u-v)A(u)$$
$$g_{12} = 0$$
$$g_{22} = (v-u)A(v) \quad ,$$

where

$$A(w) = \frac{1}{4}\frac{w}{(a-w)(b-w)(c-w)} \quad .$$

Also

$$g^{11} = \frac{1}{(u-v)A(u)}$$
$$g^{12} = 0$$
$$g^{22} = \frac{1}{(v-u)(A(v)} \quad .$$

Equation (60) takes the form

$$A(v)J_u^2 - A(u)J_v^2 = (u-v)A(u)A(v) \quad . \tag{62}$$

To get a one* parameter solution of (62) we try $J(u,v) = \phi(u) + \psi(v)$, and separation of variables gives

$$\frac{\phi'(u)^2}{A(u)} - u = \frac{\psi'(v)^2}{A(v)} - v = \alpha$$

so that

$$J(u,v,\alpha) = \int \sqrt{(u+\alpha)A(u)}\, du + \int \sqrt{(v+\alpha)A(v)}\, dv \quad .$$

Differentiating with respect to α we get the two parameter family of geodesics

$$\int \sqrt{\frac{A(u)}{u+\alpha}}\, du + \int \sqrt{\frac{A(v)}{v+\alpha}}\, dv = \beta \tag{63}$$

This can be solved for u or v in terms of elliptic functions.

---

*Because of the homogeneity this problem has essentially one dependent variable. This is obvious if we take v, say, for the parameter t.

6. Sufficient conditions. Analogous to the vanishing of the first derivative of a function $f(x)$ of a single variable we have found the vanishing of the first variation (leading to the Euler equations) as a necessary condition for the minimum of a variational problem. Corresponding to the sufficient condition $f''(x) > 0$ we might look for sufficient conditions in a variational problem by investigating the second variation. Although such considerations can lead to further necessary conditions (e.g. the Legendre condition $F_{u'u'} \geq 0$, p. 65) they can never lead to a sufficient condition. The reason for this is that in order to derive a sufficient condition we must consider all possible admissible variations, i.e. $\phi(x,\varepsilon) = u(x) + \zeta(x,\varepsilon)$ where $\zeta(x,\varepsilon)$ is an arbitrary admissible function with zero boundary values, vanishing for $\varepsilon = 0$. However, it is easy to construct an admissible variation (e.g. $\zeta = (1-x)\varepsilon \sin x/\varepsilon^2$, $0 \leq x \leq 1$) for which $\zeta_x(x,\varepsilon)$ does not approach zero for $\varepsilon \to 0$. In this case the varied integral

$$I(\varepsilon) = \int_{x_0}^{x_1} F(x,\phi(x,\varepsilon),\ \phi_x(x,\varepsilon))dx$$

does not converge to the desired integral

$$I = \int_{x_0}^{x_1} F(x,u,u')dx$$

as $\varepsilon \to 0$, and the variational problem does not reduce to that of minimizing a function of the single variable ε. A variation $\zeta(x,\varepsilon)$ which satisfies both conditions $\zeta(x,\varepsilon) \to 0$, $\zeta_x(x,\varepsilon) \to 0$ as $\varepsilon \to 0$ is called a weak variation, and geometrically means that the curve $u(x)$ is compared with curves that approximate $u(x)$ in slope as well as position. A curve $u(x)$ which minimizes an integral with respect to all weak variations is called a weak relative minimum (relative referring to the fact that $u(x)$

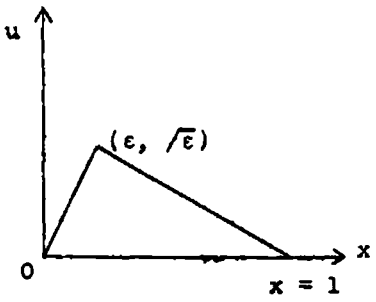
is a minimum only with respect to curves in its neighborhood). For example consider

$$I = \int_0^1 [u'(x)^2 - u'(x)^3]dx$$

subject to $u(0) = u(1) = 0$. The extremals are straight lines, and there is a unique extremal $u = 0$ satisfying the boundary conditions. The value of I is zero for $u = 0$, and is positive for all curves satisfying the condition $u'(x) < 1$, so $u = 0$ minimizes I with respect to this restricted class of neighboring curves. However, by taking the admissible curve

$$u = \frac{x}{\sqrt{\varepsilon}} \qquad (0 \leq x \leq \varepsilon)$$

$$u = \frac{\sqrt{\varepsilon}\,(1-x)}{1-\varepsilon} \qquad (\varepsilon \leq x \leq 1)$$



which approaches the extremal $u = 0$ uniformly we can make I negative.

We shall now proceed to establish Weierstrass' sufficient condition for an extremal to be a strong relative minimum for the integral

$$I(u) = \int_\tau^t F(s,u,u')ds \tag{64}$$

subject to the boundary conditions

$$A: \ (s = \tau, \ u = k) \qquad B: \ (s = t, \ u = q) \ .$$

In order to do this we must compare the values of the integral (64) over the extremal between A and B, (I_e), and an arbitrary curve C in its neighborhood, (I_c). By expressing I_e as an integral along the path C, we shall reduce this problem to a comparison of the integrands alone. Assuming that the extremal in

question can be imbedded in a field of extremals*, we define a slope function $p(u,s)$ as the slope of the extremal through the point $P = (u,s)$. Further we define the single valued distance function $\bar{I}(P)$ (see p.60ff)** which has the property that $I_e = \bar{I}(B) - \bar{I}(A)$. The differential $d\bar{I} = \bar{I}_s ds + \bar{I}_u du$ is exact and we have

$$(65) \qquad I_e = \int_A^B (\bar{I}_s ds + \bar{I}_u du) = \int_\tau^t (\bar{I}_s + u'\bar{I}_u) ds$$

where the integration is taken over any curve joining A to B, and in particular over the curve C. Using (25) and (16) we have $\bar{I}_s = F(s,u,p) - pF_p(s,u,p)$ and $\bar{I}_u = F_p(s,u,p)$ so that (65) reduces to

$$I_e = \int_\tau^t [F(s,u,p) + (u' - p)F_p(s,u,p)]ds. \quad ***$$

and finally we have

$$(66) \quad \Delta I = I_c - I_e = \int_\tau^t [F(s,u,u') - F(s,u,p) - (u'-p)F_p(s,u,p)]ds$$

where the integration is taken along the path C, u' representing the slope of the curve C and p the slope of the field at each point. The integrand of (66)

$$(67) \quad E(s,u,p,u') = F(s,u,u') - F(s,u,p) - (u'-p)F_p(s,u,p)$$

is known as Weierstrass' E function. If the condition $E \geq 0$ is satisfied at each point s, u, p in the field for all values of u', then from (66) $\Delta I \geq 0$ for all curves C in the field; in

---

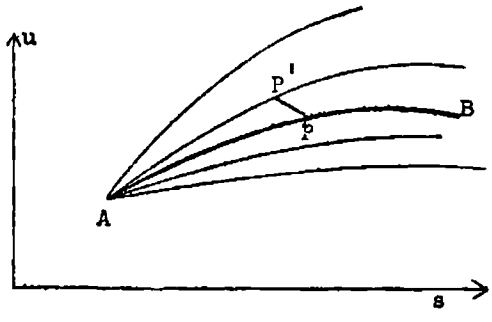
*This condition will be investigated in section 7.

**For the existence of $\bar{I}$ we must in general have a field of extremals transverse to some surface, but for the case of only one dependent variable u this condition is automatically satisfied for any field.

***This expression is known as Hilbert's Invariant Integral.

other words $I_c \geq I_e$ and the extremal joining A and B is actually a minimum. We therefore conclude that a sufficient condition for an extremal joining two points to be a minimum is that it be possible to imbed it in a field throughout which the condition $E(s,u,p,u') \geq 0$ holds for all values of u'. Obviously if $E > 0$ for all $u' \neq p$, then we have a proper minimum with $I_c > I_e$ for C not the extremal in question.

We can show that the weaker condition $E \geq 0$ along the extremal for all values of u' is necessary for a minimum. For, let us assume that the extremal between A and B is a minimum and at a point P on it $E < 0$ for some value u'. The extremal can be imbedded in a field emerging from A (that this condition is necessary will be shown in the next section). By continuity $E < 0$ on a small line segment PP' of slope u' where P' can be connected to A by an extremal of the field. Now taking the path AP'PB for the curve C in (66) we have $E = 0$ along AP' and PB, with $E < 0$ on P'P, thereby reaching the contradiction $I_c < I_e$. Since P was any point and u' any value, we conclude that $E \geq 0$ along an extremal for all u' is a necessary condition for a minimum.



If from $E > 0$ for all u' along an extremal we could conclude the sufficiont condition that $E \geq 0$ in a neighborhood for all u' then tho analogy with functions of a single variable would be complete (i.e. $f''(x) \geq 0$ necessary and $f''(x) > 0$ sufficient for a minimum). Unfortunately this inference is not true unless we impose the restriction that u' be uniformly bounded in the neighborhood (see Bolza, Lectures on the Caculus of Variations, p. 99).

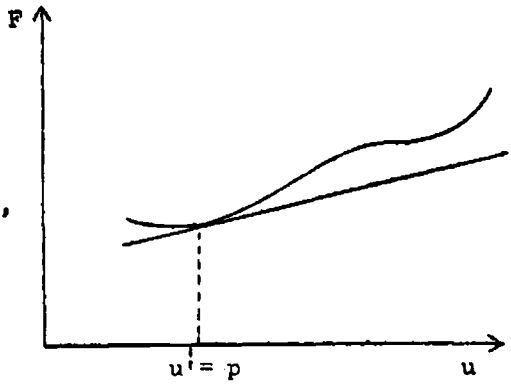
The significance of the Legendre condition can be appreciated by comparison with the E function. Using the theorem of the mean we have

$$F(s,u,u') = F(s,u,p)+(u'-p)F_p(s,u,p) + \frac{1}{2}(u'-\bar{p})^2 F_{pp}(s,u,\bar{p})$$

where $\bar{p} = p + \theta(u'-p)$, $0 < \theta < 1$, so (67) becomes

$$E(s,u,p,u') = \frac{1}{2}(u' - \bar{p})^2 F_{pp}(s,u,\bar{p}) \tag{68}$$

$\bar{p}$ being some value between p and u'. From (68) we see that if $F_{u'u'}(s,u,u') \geq 0$ at all points of the field for arbitrary values of u', then also $E \geq 0$, so this condition is sufficient. A problem for which $F_{u'u'}(s,u,u') > 0$ for all values of the quantities s, u, u' is said to be regular, and for such a problem the existence of a field of extremals guarantees a proper minimum. The connection between the Legendre condition and Weierstrass' E function may be interpreted geometrically in the following way. Considering F(s,u,u') as a function of the direction u' at a fixed point of the field (s,u) (thereby fixing p(s,u)), E(s,u,p,u') = 0 is the equation of the tangent to F at the point u' = p, and $E \geq 0$ for all u' states that F(s,u,u') lies above the tangent line. The condition $F_{u'u'}(s,u,u') \geq 0$ for all u' means that the curve is convex, and therefore lies above the tangent line. If this is true for all points (s,u), then the same statement can be made of E. The necessary condition $E \geq 0$ along an extremal includes the weak Legendre $F_{pp}(s,u,p) \geq 0$, since $E \geq 0$ for all u' implies convexity at the tangent point u' = p. Although the strong Legendre condition $F_{pp}(s,u,p) > 0$ is neither necessary nor sufficient for a strong minimum it is sufficient for a weak minimum. This is clear since from $F_{pp}(s,u,p) > 0$ on the extremal we can conclude that



$F_{u'u'}(s,u,u') > 0$ for (s,u) in some neighborhood of the extremal and u' in the neighborhood of p. This means, however, that $E \geq 0$ for weak variations (by (68)) which is sufficient for a minimum. Although $F_{u'u'}(s,u,u') > 0$ along an extremal for all u' is somewhat stronger than $E > 0$ along an extremal, it is still not sufficient for a strong minimum.

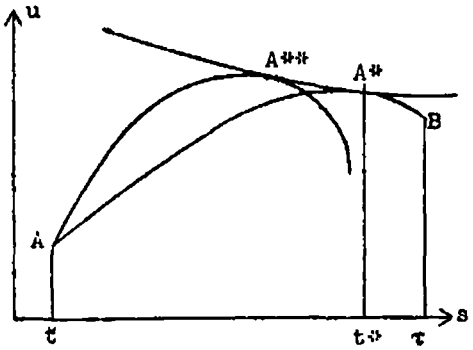
7. Construction of a Field - The Conjugate Point. We have observed that an essential point in the theory of sufficient conditions is the possibility of imbedding a given extremal in a field. We shall now see that this can always be done if the endpoints of the extremal are not too far apart. In general a one parameter family of extremals through a point will constitute a field up to its envelope.

But first of all it is important to note that if an extremal can be imbedded in a field H of extremals so that $E(s,u,p,u') \geq 0$ over the field, then the condition $E(s,u,p^*,u') \geq 0$ will also hold for any other field H* over their common (s,u) domain. For, supposing that $E < 0$ for some u' at a point (s,u,p^*), we can construct a curve C as on p. 85 for which $I_c < I_e$. However, this is impossible since the curve C lies in the field H in which I_e is proved to be a minimum. We therefore need consider the construction of but a single imbedding field. Moreover, it can be shown that if any imbedding field exists, then the family of extremals through one endpoint is such a field (see Bolza, p. 57). We therefore consider a family of extremals $u(s,a)$ through the endpoint A: (t,q). The first point at which $u(s,a_0)$ is intersected again by a neighboring extremal $u(s, a_0+\varepsilon)$ is given by the simultaneous solution of

$$\begin{cases} u = u(s,a_0) \\ 0 = u_a(s,a_0) \end{cases} . \tag{69}$$

This first point of intersection A*: (t^*,q^*), is called the conjugate point to A. If an envelope exists than A* is the

intersection of $u(s,a_0)$ with the envelope. Supplementing our previous necessary and sufficient conditions we now have the necessary condition $t^* \leq \tau$ and the sufficient condition $t^* < \tau$ for a minimum. The necessity of $t^* \leq \tau$ may be indicated by the following geometrical argument. If the conjugate point A* (i.e. intersection with the envelope) lies before B, then taking another extremal with conjugate point A** we can construct the admissible curve AA**A*B, where the arcs AA** and A*B are extremals and A**A* is along the envelope. The envelope has the same slope as the field at every point, so E = 0 along it. Using (66) we have



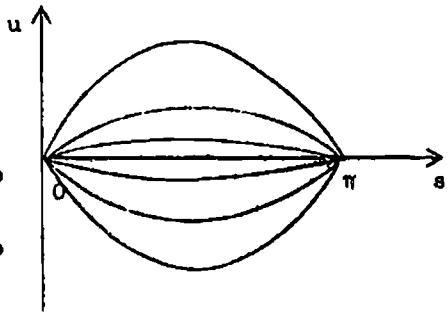
$$I_{AA^{**}} + I_{A^{**}A^*} = I_{AA^*} .$$

However, since the envelope is not an extremal in general, by connecting A** and A* by an extremal the value of $I_{AA^{**}A^*}$ can be reduced, so I_{AA^*} is not a minimum. This necessary condition can be established more rigorously by consideration of the second variation (cf. Bolza, p. 57). The sufficiency of $t^* > \tau$ for the existence of a field of extremals can be seen as follows. We have $u_a(s,a_0) \neq 0$ for $t < s \leq \tau < t^*$, and assuming u_a is continuous we may take $u_a > 0$ in this interval. By continuity we also have $u_a(s, a_0+\varepsilon) > 0$ for $|\varepsilon|$ small enough, so that at a fixed point s, u is a monotonic increasing function of a, covering a neighborhood of $\varepsilon = 0$ once and only once.

Let us take the example

$$I = \int_0^{t^*} (u'^2 - u^2)ds \qquad u(0) = 0 .$$

The extremals are
$u = a \sin s + b \cos s$,
and through the point $(0,0)$ there is the field $u = a \sin s$. We have $F_{u'u'} = 2$, $p = u \cot s$, $E = (u' - p)^2$, and the conjugate point to $(0,0)$ is $(0,\pi)$. The problem is a regular one, so the extremal joining $(0,0)$ to any point (t,u) is a minimum of $t < \pi$.



Summarizing the results of this and the previous section we have as necessary conditions for a strong minimum the Euler equations, the weak Legendre condition $F_{pp}(s,u,p) \geq 0$ which is included in Weierstrass' condition $E(s,u,p,u') \geq 0$ along the extremal, and the conjugate point condition $t^* \geq \tau$. As sufficient conditions we have the Euler equations together with either $F_{u'u'}(s,u,u') \geq 0$ or $E(s,u,p,u') \geq 0$ in a neighborhood, and $t^* > \tau$. For a weak minimum we have as necessary conditions the Euler equations, $F_{pp}(s,u,p) \geq 0$, and $t^* \geq \tau$, which become sufficient on dropping the equality signs.

Problem: Show that for $F = u\sqrt{1 + (u')^2}$ the tangents to an extremal at a point A and at its conjugate point A* intersect on the s axis.

SUPPLEMENTARY NOTES TO CHAPTER II

1. The Hamilton-Jacobi Equation. In this paragraph we give a different and more direct derivation of the Hamilton-Jacobi equations.

We consider a variational problem

$$\int_{t_1}^{t_2} F(t, u, u')dt$$

where $u = (u_1,\ldots,u_n)$, $u' = \frac{du}{dt}$. The fundamental concept in the Hamilton-Jacobi theory is the distance function, or the extremal integral, i.e. the above integral taken over the extremals of the variational problem. The extremals are the solutions of the Euler equations

$$\frac{d}{dt}(F_{u_k'}) = F_{u_k} \quad .$$

We consider a domain D in the (t,u) space in which two points P, Q lying in D can be joined by an extremal in D in exactly one way. Then the integral

$$I(P,Q) = \int_P^Q F(t,u,u')dt$$

taken over the extremal from P to Q is a function of the end-points only. This integral is the so-called extremal integral, or distance function.

In this function we first consider the point P as fixed and let Q vary. We denote the coordinates of Q by the capital letters T, $U_1,\ldots,U_n$. Then

$$I(P,Q) = J(T,U) \quad .$$

The aim is to compute the partial derivatives J_T, J_{U_k}.

If $u(t,\alpha)$ denotes a family of extremals depending on a parameter α we obtain from $I(\alpha) = \int_{t_1}^{t_2} F(t, u(t,\alpha), u'(t,\alpha))dt$,

$$\frac{dI}{d\alpha} = \int_{t_1}^{t_2} \sum_{k=1}^{n} (F_{u_k} u_{k\alpha} + F_{u_k'} u_{k\alpha}')dt$$

$$= \int_{t_1}^{t_2} \sum_{k=1}^{n} (F_{u_k} - \frac{d}{dt} F_{u_k'}) u_{k\alpha}\, dt + \sum_{k=1}^{n} u_{k\alpha} F_{u_k'} \Big|_{t_1}^{t_2} .$$

Hence

$$\frac{dI}{d\alpha} = u_{k\alpha} F_{u_k'} \Big|_{t_1}^{t_2} .$$

This is the standard computation by which the Euler equation was derived.

Now using $U_k = \alpha$ as parameter, since $\frac{\partial u_k}{\partial U_\ell} = 0$ at P and

$$\frac{\partial u_k}{\partial U_\ell} \begin{cases} = 0 & \text{for } t = T, \quad k \neq \ell \\ = 1 & \text{for } t = T, \quad k = \ell . \end{cases}$$

we obtain

(1) $$J_{U_k} = F_{u_k'} \Big|_{t = T} .$$

In order to compute J_T we consider the extremal integral from P to a variable point $(t,u(t))$, lying on a fixed extremal through P. This integral is given by $J(t,u(t))$. Then

$$\frac{d}{dt} J(t,u(t)) = J_t + \sum_{k=1}^{n} J_{u_k} u_k' = F(t,u,u') .$$

Now inserting $t = T$ we obtain

$$J_T = F - \sum_{k=1}^{n} J_{U_k} u_k'(T)$$

and with (1)

$$(2) \qquad J_T = F - \sum_{k=1}^{n} F_{u_k'} u_k'(T) \quad .$$

Thus we have derived the fundamental equations: (1), (2).

If P also is considered variable it is easy to get the derivations of the extremal integral. Denote the coordinates of P and Q by (t,u) and (T,U) respectively. Then

$$I(P,Q) = J(t,u;\ T,U) \ .$$

By definition, the partial derivatives J_T and J_{U_k} are obtained by keeping t and u, i.e. P, fixed. Therefore according to (1) and (2)

$$(3) \qquad J_{U_k} = F_{u_k'} \Big|_T \ ; \quad J_T = F - \sum_{k=1}^{n} u_k' F_{u_k'} \Big|_T \ .$$

Here $u_k'(T)$ has to be computed as the derivative of the extremal u_k connecting P and Q.

To derive similar formulas for J_t, J_{u_k} one only has to exchange the role of P and Q. Observing that $I(Q,P) = -I(P,Q)$ we find

$$(3') \qquad J_{u_k} = -F_{u_k'} \Big|_t \ , \quad J_t = -F + \sum_{k=1}^{n} u_k' F_{u_k'} \Big|_t \ .$$

One can write the equations (3) and (3') in another form. Here the variation of the extremal integral is expressed by the variation of the endpoints only:

$$\delta J = \sum_k J_{U_k} \delta U_k + J_T \delta T + \sum_k J_{u_k} \delta u_k + J_t \delta t$$

$$(4) \qquad \begin{cases} \delta J = \sum_k F_{u_k'} \Big|_T \delta U_k + \left(F - \sum_k F_{u_k'} u_k'\right) \Big|_T \delta T \\ \qquad - \sum F_{u_k'} \Big|_t \delta u_k - \left(F - \sum_k F_{u_k'} u_k'\right) \Big|_t \delta t \end{cases}$$

From equation (3) one easily is led to the Hamilton-Jacobi equation. We consider P fixed. Formula (3) consists of n+1 equations for the unknown functions $J(U,T)$ and $u_k'(T)$. In order to separate the unknown functions we introduce new variables $v = (v_1,\ldots,v_n)$ in the place of $u' = (u_1',\ldots,u_n')$ by setting

$$v_k = F_{u_k'}(t,u,u') \quad . \tag{5}$$

We have to assume here that

$$\det (F_{u_k'u_i'}) \neq 0 \quad .$$

Then the unknown functions u_k' are transformed into the unknown functions v_k. Introduce the Hamilton function $H(t,u,v)$ by eliminating u' in

$$H = \sum u_k' v_k - F(t,u,u') \quad .$$

Then the equations (3) are transformed into

$$J_{U_k} = v_k(T)$$

$$J_T = - H(T,U,v(T)) \quad .$$

If we insert $v_k(T) = J_{U_k}$ in the last equation we have in

$$J_T + H(T,U,J_{U_k}) = 0 \tag{6}$$

an equation which only involves the unknown function J. If it is solved we find the n functions $v_k(T)$ from $v_k = J_{U_k}$. Thus by the Legendre transformation (5) we have reduced the (n+1) equations (3) to the one equation (6). This is the Hamilton-Jacobi equation. It is a partial differential equation of first order, with the simplification that the unknown function, J, does not itself appear. We have just shown that the extremal integral $J(T,U)$ satisfies the Hamilton-Jacobi equation, if one considers the endpoint Q as variable and keeps the initial point P fixed.

2. Correspondence between solutions of the Hamilton-Jacobi equation and Fields. In the preceding paragraph we showed that the extremal integral from a fixed point P to a variable one is a solution of the Hamilton-Jacobi equation. We now want to find all solutions of this equation. First we will show that the extremal integral taken over the extremals from a fixed surface C to a variable point also is a solution of the Hamilton-Jacobi equation. Later on we will show that these actually give all solutions of (6).

For the proof we have to construct extremals from C to a variable point Q with coordinates (T,U). This is a problem with a free boundary point on the left end. A necessary condition for an extremal is the transversality condition. Let us denote by P the point where the extremal from Q cuts C; the coordinates of P being (t,u). Then the first variation of the extremal integral J for fixed Q is , according to (4)

$$\delta J = - \sum F_{u_k'} \, \delta u_k - (F - \sum_k F_{u_k'} \, u_k') \, \delta t \quad .$$

Here δu_k, δt are variations on the surface C. This expression has to vanish. Observe that the resulting equation is the transversality condition. Thus if we consider Q variable we obtain from (4)

$$(7) \quad \delta J = \sum_k F_{u_k'} \Big|_T \delta U_k + (F - \sum_k F_{u_k'} \, u_k') \Big|_T \delta T \quad .$$

The contribution at P dropping out because of the transversality condition. Since δU_k, δT are independent variations we derive the equations (3) from (7). Applying the Legendre transformation to (3) we see that J(T,U) satisfies the Hamilton-Jacobi equation (6).

The extremal integral has the additional property of vanishing on C. Later on we will see that all solutions of the Hamilton-Jacobi equation which vanish on a surface C are identical with the extremal integral taken from C to a variable

point. As a matter of fact, for this purpose we only have to prove that the solution of (6) with the initial condition $J = 0$ on C is unique. Then the solution has to coincide with the extremal integral.

The extremal integral has the following geometric meaning. From the surface C construct all extremals which cut C transversally. Then in a sufficiently small region, where the extremals cover the (t,u) space simply, the integral J(T,U) is defined as $\int \mathcal{L}\, dt$ taken over the extremal through $Q = (T,U)$ which cuts C transversally.

The family of extremals which cut C transversally is called a "field".

Now it is clear that the surface $J = a$, where a is a small constant also cuts the above field transversally. Namely on $J = a$ we have

$$0 = \delta J = \sum_k F_{u'_k} \Big|_T \delta U_k + (F - \sum_k F_{u'_k} u'_k) \Big|_T \delta T$$

which is the transversality condition. Denote the surface $J = a$ by C_a, so that $C_0 = C$. Then all these surfaces C_a cut the field constructed above transversally. In other words: The fields constructed transversally through C_0 and through C_a are identical. These surfaces C_a play an important role in optics where they are called wavefronts. The extremals correspond to the light rays and the fields transversal to the C_a are called light congruences.

We conclude: To every field of transversals through C there corresponds a solution of the Hamilton-Jacobi equation, namely the extremal integral taken from C along an extremal of the field. Conversely, to every solution J^* of the Hamilton-Jacobi equation there corresponds a field. Consider the surface $J^* = a$ and construct the field transversal to it. Then the extremal integral J taken along this field from the surface $J^* = a$ is also a solution of the Hamilton-Jacobi equation. Now J+a is another solution such that J^* and J+a both are equal on

the surface $J^* = a$. If we invoke the uniqueness of the solutions of the Hamilton-Jacobi equation we see that $J + a = J^*$. Thus the one to one correspondence between the solutions of the Hamilton-Jacobi equation and the fields of extremals is established.

3. Application to differential geometry. Consider the curves on a 2-dimensional surface. The element of length is given by

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2 \quad .$$

when the E, F, G are functions of u,v satisfying $EG - F^2 > 0$.

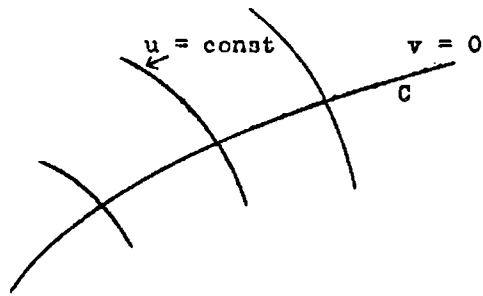
If we consider the variational problem

$$\int ds = \int \sqrt{E + 2F\,v' + G\,v'^2}\,du$$

the extremals coincide with the geodesics and transversality corresponds to orthogonality.

We consider the neighborhood of a curve C which we take as $v = 0$ and try to introduce the coordinates u,v in such a way that E, F, G attain a simple form. The coordinates we are aiming at are the Gauss' coordinates.

For this purpose draw the geodesics transversal to C. We label these geodesics by $u = \text{const}$ where u is taken as the arc length from some point on C. Now we introduce v as the length along $u = \text{constant}$. Then C is given by $v = 0$. We claim this coordinate system is orthogonal, i.e. the lines $u = \text{constant}$ and $v = \text{constant}$ cut each other transversally. This is an immediate consequence of our general considerations.



The geodesics u = const form a field. The extremal integral taken from C is J = v by definition. The surfaces, here curves, J = v = constant cut the field u = constant transversally as we wanted to show.

In the coordinates u,v the expressions E,F,G have a simplified form. Since the coordinate system is orthogonal one easily finds F = 0. Since for u = constant

$$ds = dv$$

by definition of v we have G = 1. Thus in these coordinates the line element has the simple form

$$ds^2 = E(u,v)\, du^2 + dv^2 \quad .$$

Since n was the arc length along C we have

$$E(u,0) = 1 \quad .$$

If one takes for C a geodesic one finds that

$$E_v(u,0) = 0 \quad .$$

4. Analytical representation of a field. In the preceding paragraph we defined a field as the family of all extremals cutting a surface C transversally. Therefore the extremals of the family will depend on n parameters. We represent the extremals of a field in canonical coordinates as $u = u(t,\alpha)$, $v = v(t,\alpha)$ where $\alpha = \alpha_1,\ldots,\alpha_n$ stands for n parameters. We assume C is given by $t = \phi(u)$. Since the $u(t,\alpha)$ should give all the extremals, at least in a certain neighborhood, we require

$$\det\left(\frac{\partial u_k}{\partial \alpha_\ell}\right) \neq 0 \quad . \tag{8}$$

We now ask the question whether every family of extremals $u(t,\alpha)$ depending on n parameters and satisfying (8) form a field, i.e. cuts some surface C transversally. The answer is no.

In what follows we will derive the characterizing properties of a field.

For this purpose we use the characteristic theory derived in Chap II and try to construct the solution $J(t,u)$ of the Hamilton-Jacobi equation

$$J_t + H(t,u,J_u) = 0 \tag{9}$$

which is built up by the extremals of the given family. The equations for the characteristics are, if we write $w = J$, $v_k = J_{u_k}$:

$$\begin{cases} u_k' = H_{v_k} \\ v_k' = -H_{v_k} \qquad (k = 1,\dots,n) \\ w' = \sum_k v_k H_{v_k} - H \quad . \end{cases} \tag{10}$$

The general solution is obtained by taking an n parameter family of solutions $u = u(t,a)$, $v = v(t,a)$, $w = w(t,a)$ where $a = (a_1,\dots,a_n)$ and

$$\det \left(\frac{\partial u_k}{\partial a_l}\right) \neq 0 \quad .$$

Then expressing the a by the u we obtain $w(t,a) = J(t,u)$ which is the desired solution.

Now we take for $u = u(t,a)$, $v = v(t,a)$ the given family of extremals and try to determine $w(t,a)$. Since $u(t,a)$, $v(t,a)$ are extremals the first 2n equations in (10) are satisfied. Thus there remains

$$w' = \sum v_k H_{v_k} - H = g(t,a)$$

where $g(t,a)$ is a given function. Thus we only have to give w for some value of t. Because of $w(t,a) = J(t,u)$ we also have

$$w_{a_k} = \sum_{r=1}^{n} J_{u_r} u_{ra_k} = \sum_r v_r u_{ra_k}$$

which is given. Thus in order to determine w for $t = t_0$, we have to integrate the n equations (11). This is only possible if the integrability condition

$$\frac{\partial}{\partial a_l} \sum_r v_r \, u_{ra_k} = \frac{\partial}{\partial a_k} \left(\sum_r v_r \, u_{ra_l} \right)$$

is satisfied. These equations can be written in the form

$$[a_k, a_l] = 0 \tag{12}$$

where the bracket is an abbreviation for

$$[a_k, a_l] = \sum_r \left(u_{ra_k} v_{ra_l} - u_{ra_l} v_{ra_k} \right) .$$

This expression plays an important role in mechanics and optics. It is called the "Lagrange bracket".

If the condition (12) is satisfied for $t = t_0$ we can determine w up to a constant for $t = t_0$ from (11). Then from the last equation of (10) it is determined for small $t - t_0$ up to an additive constant. By the way this implies: If (12) is satisfied for $t = t_0$ it is satisfied for all t for which u,v are defined. One can check this statement directly by showing that $\frac{d}{dt} [a_k, a_l] = 0$ for the solutions of the Hamilton equations. This property is even characteristic for Hamiltonian systems.

If in the function $w(t,a)$ we express the a by u we find in $J(t,u) = w(t,a)$ a solution of (9). This function J is determined up to an additive constant. This constant is irrelevant for the family of surfaces $J = \text{const.}$ which cut the given family of extremals transversally.

Thus we have seen: A "field of extremals" consists of a family of extremals $u(t,a)$, $v(t,a)$, depending on n parameters, $a = (a_1, \ldots, a_n)$ such that

$$\det \left(\frac{\partial u_k}{\partial a_l} \right) \neq 0 \quad , \qquad [a_k, a_l] = 0 \quad .$$

The last condition is empty if $n = 1$, since $[a,a] = 0$ holds by definition of the Lagrange bracket.

§ 5. Conjugate Points. In Chapter II necessary conditions for an extremal to give a minimum were derived, e.g. the Legendre and the Weierstrass criteria. Both these conditions refer to extremals which are sufficiently short. In order to make this point clear we repeat the argument for the Weierstrass condition.

Let

$$I(C) = \int_C F(t,u,u')dt$$

be the integral considered and let C_0 be an extremal from P to Q. We introduced the extremal integral $I^*(C)$ which was defined as the integral $\int F(t,u,u')dt$ taken from P along an extremal to a variable point. In other words the extremal integral is defined on the field of extremals which pass through P. We can write these extremals with n parameters $\alpha = (\alpha_1,\ldots,\alpha_n)$ in the form $u(t,\alpha)$, where $u(t,0) = v(t)$ corresponds to the extremal from P to Q. Now we consider a region in the neighborhood of $u(t,0)$ which is covered simply by the $u(t,\alpha)$. In other words through every point of this (t,u) neighborhood there should pass exactly one extremal $u(t,\alpha)$. Such a neighborhood exists for all t such that

$$\Delta(t) = \det\left(\frac{\partial u_k(t,0)}{\partial \alpha_\ell}\right) \neq 0 \tag{13}$$

This statement holds since for any point $(\hat{t},\hat{u})$ one can for sufficiently small $|\hat{u}_k - v_k(\hat{t})|$ solve the equations $u_k(\hat{t},\alpha) = \hat{u}_k$ by the implicit function theorem. Therefore we assume that in the interval of integration $t_1 < t \leq t_2$ condition (1) is satisfied (t_1,t_2 are the t coordinates of P,Q). We have to exclude $t = t_1$ since $u(t_1,\alpha) = v(t_1)$.

Under this assumption we can write according to (66) in Chap. II

$$I(C) - I(C_0) = I(C) - I^*(C) = \int_C E(t,u,p,u')dt$$

where C is any curve lying in the above neighborhood which connects P and Q. Here we made use of the fact that $I^*(C)$ depends only on the endpoints of C.

Thus we see: If in $t_1 < t \leq t_2$ the condition (13) and the Weierstrass condition $E \geq 0$ show that C_o is a minimal. This condition (13) is certainly fulfilled if $|t_2-t_1|$ is sufficiently small since the parameters a can be chosen in such a way that (13) holds for those $t > t_1$ which are close to t_1.

We determine the first zero of $\Delta(t)$ to the right of t_1 and denote it by τ, so that $\Delta(\tau) = 0$ and $\Delta(t) \neq 0$ for $t_1 < t < \tau$. Such a τ need not exist in which case the condition (13) is always satisfied. We now assume there is such a τ. We call $(\tau, v(\tau))$ the "conjugate point" of P on the extremal $u = v(t)$.

If now the interval of integration (t_1, t_2) contains τ the Weierstrass condition does not ensure that C_o is a minimal. And, in fact, we are going to prove the result of Jacobi: If $t_1 < \tau < t_2$ then the extremal $u = v(t)$ is not a minimal, provided $\det (F_{u_k' u_\ell'}) \neq 0$ along the extremal. In the special case $\tau = t_2$ no general statement can be made.

We shall give a proof which is essentially due to Bliss in this section.

Let $u = v(t) + \varepsilon w(t)$ where $w(t)$ is a function that vanishes at t_1 and t_2. Then

$$I(\varepsilon) = \int_{t_1}^{t_2} F(t,u,u')\, dt$$

is a function of ε and the second variation will be

$$I''(0) = \int_{t_1}^{t_2} \Phi(t,w,w')\, dt \tag{14}$$

where

$$\Phi(t,w,w') = \sum_{k,\ell} (w_k' w_\ell' F_{w_k' w_\ell'} + 2 w_k w_\ell' F_{w_k w_\ell'} + w_k w_\ell F_{w_k w_\ell}) .$$

Here the argument in the derivatives of F are t, v, v'. If we now find a function w such that $I''(0) < 0$, then for sufficiently small ε we have

$$I(\varepsilon) < I(0) .$$

Now $\int_{t_1}^{t_2} \Phi \, dt$ represents a new variational problem with respect to the variable functions w. The extremal corresponds to w = 0. It is easy to see that the conjugate point of $(t_1,0)$ along the extremal w = 0 is $(\tau,0)$. Namely if u(t,a) is a family of extremals such that u(t,0) = v(t) then

$$\frac{d}{dt} F_{u_k'}(t,u,u') = F_{u_k}(t,u,u') \quad .$$

Differentiating these equations with respect to a and putting a = 0 gives

$$\frac{d}{dt} \Phi_{w_k'}(t,w,w') = \Phi_{w_k}(t,w,w') \tag{15}$$

as one caneasily check. Here $w = u_a(t,0)$. If we apply this consideration to the above family u(t,a) which depends on the n parameters $a = (a_1,\ldots,a_n)$ we obtain n solutions

$$w^{(\nu)}(t) = \frac{\partial u(t,0)}{\partial a_\nu} \quad .$$

The equations (15) are the Euler equations of the variational problem (14). They are linear differential equations since Φ is quadratic in w,w'. We found n solutions $w^{(\nu)}$ of (15) which are independent since

$$\det (w_k^{(\ell)}) = \Delta(t) \neq 0 \quad . \tag{16}$$

Thus we have reduced the statements of Jacobi to the simplified problem (14) which sometimes is called the accessory problem of the given problem.

Now let us assume that τ is the first zero of $\Delta(t)$ behind t_1. Because of (16) it follows that there exist constants c_ℓ $(\ell = 1,\ldots,n)$ which are not all zero such that

$$W(t) = \sum c_\ell \, w^{(\ell)}(t)$$

is also a solution of (15) and $W(\tau) = 0$. Furthermore $W(t) \neq 0$ for $t_1 < t < \tau$ since not all the c_ℓ are zero. Hence the

problem (14) has an extremal solution which is not identically zero and which passes through $(t_1,0)$ and $(\tau,0)$. For this extremal we have

$$\int_{t_1}^{\tau} \Phi(t,W,W')\, dt = 0 \quad ,$$

as we are going to show now.

Since Φ is quadratic in w and w' we have for any constant λ

$$\Phi(t,\lambda w,\lambda w') = \lambda^2\, \Phi(t,w,w') \quad .$$

Differentiating with respect to λ and putting $\lambda = 1$ gives Euler's identity

$$\sum_k (\Phi_{w_k}\, w_k + \Phi_{w_k'}\, w_k') = 2\Phi$$

Therefore we have

$$2\int_{t_1}^{t_2} \Phi\, dt = \sum_k \int_{t_1}^{t_2} (\Phi_{w_k}\, w_k + \Phi_{w_k'}\, w_k')\, dt$$

$$= \sum_k w_k\, \Phi_{w_k'}\Big|_{t_1}^{t_2} + \sum_k \int_{t_1}^{t_2} (\Phi_{w_k} - \frac{d}{dt}\, \Phi_{w_k'}) w_k\, dt \quad .$$

For every extremal we have

$$\int_{t_1}^{t_2} \Phi\, dt = \frac{1}{2} \sum_k w_k\, \Phi_{w_k'}\Big|_{t_1}^{t_2} \quad .$$

Hence

$$\int_{t_1}^{\tau} \Phi\,(t,W,W')dt = \frac{1}{2} \sum w_k\, \Phi_{w_k'}\,(t,W,W')\Big|_{t_1}^{\tau} = 0 \quad .$$

If we now construct the function $\phi(t)$ by

$$\phi(t) = \begin{cases} W(t) & \text{for } t_1 \le t \le \tau \\ 0 & \text{for } \tau \le t \le t_2 \end{cases}$$

it is continuous but has a corner at $t = \tau$ and it is

$$\int_{t_1}^{t_2} \Phi(t,\phi,\phi')dt = \int_{t_1}^{\tau} \Phi(t,W,W')dt = 0 \quad .$$

We can find such a curve in every neighborhood of w=0 since $\phi = \varepsilon\, \phi(t)$ will satisfy the same relations as ϕ since Φ is a homogeneous function in w,w'. The function $\phi(t)$, however, is not an extremal and can be replaced by a function which even makes (14) negative. Namely if this would not be possible the function $\phi(t)$ would have to satisfy the Weierstrass-Erdmann condition at $t = \tau$, i.e. $\Phi_{w_k'}$ would have to be continuous at $t = \tau$. For $t > \tau$ this expression is 0 since $w = \phi = 0$. The expressions $\Phi_{w_k'}(t,W,W')$ are not all zero for $t = \tau$. Otherwise we would have for $w = W$, $t = \tau$

$$\Phi_{w_k'} = \sum_{\ell} F_{w_k'w_\ell'}\, w_\ell' + F_{w_k'w_\ell}\, w_\ell$$
$$= \sum_{\ell} F_{w_k'w_\ell'}\, W_\ell' = 0 \qquad \text{for } k = 1,\ldots,n \quad .$$

Since we assumed $\det (F_{w_k'w_\ell'}) \neq 0$ it would follow $W'(\tau) = 0$. From the uniqueness theorem for systems of ordinary differential equations it would follow that $W(t) \equiv C$, which is a contradiction.

Therefore we can construct a function $\psi(t)$ with $\psi(t_1) = \psi(t_2) = 0$ lying in an arbitrary neighborhood of $w = 0$, such that

$$\int_{t_1}^{t_2} \Phi(t,\psi,\psi')dt < 0 \quad .$$

This completes the proof of the necessity of Jacobi's condition for a minimum.

For the computation of the conjugate point it is convenient to use the accessory variational problem (14) since the Euler equations are linear and in general easier to solve than the original Euler equation.

Example: Conjugate points of a geodesic on a surface: Using the Gauss coordinates which were introduced in §3 we have

$$\int F(u,v,v')\,du = \int \sqrt{E(u,v) + v'^2}\,du \quad .$$

Here $v = v(u)$ is the variable function and u plays the role of t. The geodesic considered is $v = 0$. Furthermore we have $E(u,0) = 1$, $E_v(u,0) = 0$. The function $-\frac{1}{2}E_{vv}(u,0) = K(u)$ is called the Gauss' curvature of the surface at the point $(u,0)$. We compute

$$\Phi(u,w,w') = \tfrac{1}{2}(w'^2 + \tfrac{1}{2}E_{vv}\,w^2) = \tfrac{1}{2}(w'^2 - Kw^2) \quad .$$

The linear Euler equation is

$$\frac{d}{dt}\Phi_{w'} - \Phi_w = w'' + K(u)\,w = 0 \quad . \tag{17}$$

This is called the Jacobi equation. In order to determine the conjugate point of $(u_1,0)$ one has to construct a solution $w(t)$ of (17) which vanishes at $u = u_1$ but does not vanish identically. The next zero of $w(t)$ is the conjugate point. For K = constant, we have $w = \sin\sqrt{K}\,(u-u_0)$, if $K > 0$. Then the distance between a point and its conjugate points is $\pi/\sqrt{K}$. For $K \leq 0$ there is no conjugate point.

6. Application to Sturm's theory. We consider the quadratic function ($n=1$).

$$F(t,u,u') = \tfrac{1}{2}(u'^2 - q(t)u^2) \quad ,$$

so that the Euler equation is

$$u'' + q(t)u = 0 \quad . \tag{18}$$

The conjugate points with respect to the extremal $u = 0$ are easily interpreted in this case. Let $w(t)$ be a solution of the Euler equation which vanishes for $t = t_1$, but not identically. Then $u(t,\alpha) = \alpha w(t)$ is a family of solutions through the point

$(t_1,0)$. Furthermore $\frac{\partial u}{\partial \alpha} = w(t)$. We obtain the conjugate point of $(t_1,0)$ as the next zero of $w(t)$, if it exists.

According to Jacobi's criterion one can interpret the interval between t_1 and τ as the longest interval in which $u = 0$ is a minimal for $\int F\,dt$. Observe that $F_{u'u'} = 1$ and therefore all conditions of the criterion are satisfied. From this interpretation follows immediately Sturm's Separation Theorem: If t_1,t_2 are consecutive zeros of a solution $w(t)$ of (18) then in $t_1 \leq t \leq t_2$ there is exactly one zero of any other solution.

Proof: We prove first there is at most one zero of any other solution. If there were two consecutive zeros τ_1,τ_2 in the above interval then (τ_1,τ_2) would be the longest interval in which $u = 0$ is a minimal, i.e. if τ is any number greater than τ_2 (we can choose it $< t_2$) then $u = 0$ is not a minimal in (τ_1,τ). But the interval (τ_1,τ) is contained in (t_1,t_2) and in every interval (t_1,τ) with $\tau < t_2$, $u = 0$ is a minimal. This is a contradiction.

If, however, there is a solution w_1 which has no zero in $t_1 \leq t < t_2$ then $u = 0$ would be a minimal in a larger interval than (t_1,t_2) which again leads to a contradiction.

This theorem can be proved more directly for the equation (18). But this argument can be generalized to more general equations and for $n > 1$.

Denote the sequence of zeros of a solution $w(t)$ by t_ν such that $t_\nu < t_{\nu+1}$. Let ν run over all integers, or only some of them, or even none of them if there are no zeros. Let $\omega(t)$ be another non-vanishing solution of (18) and let τ_ν with $\tau_\nu < \tau_{\nu+1}$ be the sequence of zeros of ω. We choose ν such that $t_1 \leq \tau_1 < t_2$. According to Sturm's Separation Theorem there exists such a zero τ_1 of ω_1. Then it follows that for all ν:

$$t_\nu \leq \tau_\nu < t_{\nu+1}$$

The zeros of the $w(t)$ and $\omega(t)$ "separate" each other.

One proves easily: If $t_1 = \tau_1$ then $t_\nu = \tau_\nu$ and this

occurs only if $w(t)$ and $\omega(t)$ are proportional to each other (linearly dependent).

In general the equation (18) can not be solved explicitly and therefore the t_ν cannot be determined explicitly. That is why it is valuable to give bounds for the t_ν. This is done by Sturm's Comparison Theorem: If $q(t) \leq Q(t)$ and

$$w''(t) + q(t)\, w(t) = 0$$
$$W''(t) + Q(t)\, W(t) = 0$$

and $T_1 < T_2$ are consecutive zeros of $W(t)$ then $w(t)$ has at most one zero in $T_1 \leq t < T_2$, provided $w(t)$ is not identically zero.

Proof: Assume there are two zeros $t_1 < t_2$ of w in $T_1 \leq t < T_2$. Then determine a τ in $t_2 < \tau < T_2$. According to Jacobi's condition $\frac{1}{2}\int_{t_1}^{\tau} (w'^2 - q(t)w^2)dt$ has not a minimum for $w = 0$. In other words there is a function $u(t)$ such that $u(t_1) = u(\tau) = 0$ and

$$\int_{t_1}^{\tau} (u'^2 - q(t)u^2)dt < 0 \quad .$$

If we now define $u(t) = 0$ outside (t_1,τ) we find

$$\int_{T_1}^{\tau} (u'^2 - Q(t)u^2)dt = \int_{t_1}^{\tau} (u'^2 - Q(t)u^2)dt$$

$$\leq \int_{t_1}^{\tau} (u'^2 - q(t)u^2)dt < 0 \quad .$$

Hence $u = 0$ is not a minimum for $\int(u'^2 - Qu^2)dt$ in (T_1,τ) which is a subinterval of (T_1,T_2). This is a contradiction.

This theorem implies that the zeros of w grow closer together as $q(t)$ increases. If, for instance,

$$m^2 \leq q(t) \leq M^2$$

where m,M are positive constants, it follows that

$$\frac{\pi}{M} \leq |t_{\nu+1} - t_\nu| \leq \frac{\pi}{m} \quad .$$

Namely for $Q(t) = M$ we have

$$W(t) = c \sin (Mt + \alpha)$$

and $T_{\nu+1} - T_\nu = \frac{\pi}{M}$. Similarly for $q(t) = m$. The comparison theorem gives the above result.

III. DIRECT METHODS IN THE CALCULUS OF VARIATIONS

Introduction

The so-called direct methods in the calculus of variations represent a relatively modern trend which has established the calculus of variations in a dominating position in mathematical analysis.

Two general points of view in the calculus of variations are relevant for various domains of mathematics, namely the formation of invariants and covariants in function spaces, and the characterization of mathematical entities by extremum properties. We shall concentrate on the second topic. Such a characterization is useful in many fields of mathematics, and often serves to simplify more involved deductions. For example, in the theory of numbers the greatest common divisor of two integers, a and b, can be characterized as the minimum of the expression $|ax + by|$, where for x and y all integers are admitted "to competition". In this course we shall confine ourselves to the field of mathematical analysis.

In the mathematical treatment of physical phenomena it is often expedient to use formulations by means of which the quantities under consideration appear as extrema. An example of that is Fermat's Principle in optics. In mechanics the principle of stable equilibrium has a basic importance: a system is in stable equilibrium if, and only if, the potential energy is a minimum. For elementary mechanics equilibrium conditions are expressed by certain local conditions (vanishing of the sum of all forces and moments). With the help of the calculus of variations it becomes possible to characterize a state of equilibrium by one value only, the extremum of a functional. The variational equations furnish, then, the local conditions.

The classical methods of the calculus of variations can be considered as indirect methods, in contrast to the modern direct methods. The distinction is not absolute, for many

"modern" ideas appear already at the beginning of the calculus of variations, although without the degree of precision now attained.

Generally speaking, the direct methods aim at solving boundary value problems of differential equations by reducing them to an equivalent extremum problem of the calculus of variations, and then by attacking this problem directly, a procedure which is, so to speak, the reverse of classical calculus of variations.

The most notable example of the direct approach goes back to Gauss and William Thompson. They considered the boundary value problem of the harmonic equation $\triangle u = 0$ for a domain G in the xy-plane, under the condition that the function u be regular in G and attain prescribed continuous boundary values at the boundary.

The classical formalism of the calculus of variations for the integral

$$D[\phi] = \iint_G (\phi_x^2 + \phi_y^2)\, dxdy$$

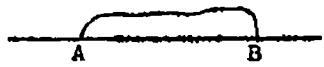
shows:

If u(x,y) furnishes the minimum of the integral when for ϕ all functions are admitted to competition which are continuous in G and on its boundary, attain the prescribed boundary values, and possess continuous first and second derivatives in G, then u(x,y) is the solution of the boundary value problem for D[u] = 0 in G.

Gauss and Thompson thought that, since the integral $D[\phi]$ is always positive, it must have a minimum: hence the existence of a solution of the boundary value problem appeared established. This reasoning was later resumed by Dirichlet, and a decisive use of it, under the name of Dirichlet's Principle, was made by Bernhard Riemann in his epoch-making investigations on the theory of functions. However, it was soon observed by Weierstrass that the reasoning suffered from a very serious gap. A

set of non-negative numbers necessarily has a greatest lower bound, but this lower bound need not be a minimum actually attained in the set. To make Dirichlet's Principle a cogent proof, the existence of a minimum, rather than a greatest lower bound, has to be established. That this is not a trivial matter can be seen from many simple examples of extremum problems apparently "reasonable".

(a) Find the shortest curve from A to B with the condition that it should be perpendicular to AB at A and B.



The length of the admissible curves has a greatest lower bound, namely $\overline{AB}$; however, no shortest curve exists.

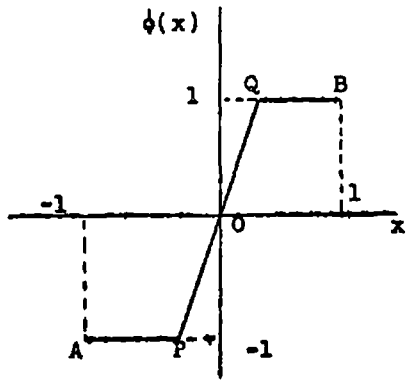
(b) Find a function $\phi(x)$ continuous, having a piecewise continuous derivative, for which the integral

$$I = \int_{-1}^{1} x^2 \, \phi'(x)^2 \, dx \quad ,$$

attains the smallest possible value, with the boundary conditions:

$$\phi(-1) = -1 \qquad \phi(1) = 1 \quad .$$

The integral is always positive and has a greatest lower bound, namely, 0. However, we can make it as small as we want, by taking for $\phi(x)$ the function represented by APQB. If ϵ is the distance from P to the point (0,-1),



$$\phi'(x) = \frac{1}{\varepsilon} \quad \text{for } -\varepsilon < x < \varepsilon \ ,$$

$$\phi'(x) = 0 \quad \text{for } -1 < x < -\varepsilon \quad \text{or} \quad \varepsilon < x < 1 \ .$$

Hence

$$I = \int_{-\varepsilon}^{\varepsilon} \frac{x^2}{\varepsilon^2}\, dx = \frac{2}{3}\,\varepsilon \ ,$$

and it can be made as small as we want. But the only function for which $I = 0$ is $\phi(x) = c$, and it does not correspond to a curve through the given end-points.

Weierstrass' criticism was met only much later by Hilbert, when he succeeded in 1900 in establishing the existence of a minimum for problems involving the integral $D[\phi]$, and thus opened the way for broad developments in the calculus of variations.

The direct methods thus inaugurated marked a great progress in pure and applied analysis, all of it based upon the reduction of boundary value problems to minimum problems. Three related goals are envisaged by such methods:

a) Existence proofs for solutions of boundary value problems,

b) Analysis of the properties of these solutions,

c) Numerical procedures for calculating the solutions.

This last point of view has been stressed by Rayleigh and, in a broader way, by Walter Ritz, who developed powerful numerical methods of great importance for physics and engineering.

A few examples will show how certain results concerning minimum problems can be attained directly.

By the formal approach the isoperimetric property of the circle (or of the sphere) is reduced to a differential equation, supplemented by certain sufficient conditions, for instance, Weierstrass' conditions. However, the direct approach leads to the result in a straightforward way.

On the basis of Steiner's proof * or of the classical calculus of variations, we may assume as proved that the circle is

* See this proof in Courant-Robbins, What is Mathematics, p.375.

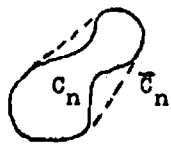
the solution, provided a solution is assumed to exist. Hence only the existence of a solution needs to be proved.

We consider the following problem: Among all continuous, closed curves C having a given length L, find one which makes the enclosed area A(C) a maximum.

Since any admissible curve C can be completely enclosed in a circle of radius L/2, $A(C) \leq \pi L^2/4$; hence a least upper bound M exists for all the areas, and a maximizing sequence $C_1,\ldots,C_n,\ldots$ of admissible curves exists such that

$$A_n(C_n) \rightarrow M \qquad \text{for } n \rightarrow \infty \ .$$

Each C_n can be assumed to be a convex curve, for if not, it could be replaced by a convex admissible curve of larger area: C_n is first replaced by its "convex hull" $\overline{C}_n$; then $\overline{C}_n$ is magnified into a similar admissible curve of length L, $\overline{\overline{C}}_n$, and



$$A(C_n) < A(\overline{C}_n) < A(\overline{\overline{C}}_n) \ .$$

(This reasoning is not valid in 3 dimensions. Let us consider a sphere with a long spine. It can be replaced by a convex surface enclosing a larger volume: but this surface will also have a larger area, which prevents us from extending the argument to more than 2 dimensions.)

We now make use of the following theorem: In a sequence of convex curves lying in a closed domain there is a subsequence which converges to a closed convex curve.

Hence a subsequence of curves C_n converges to a convex curve C: since the area of a sequence of convex curves depends continuously on the curves, and the areas A_n of C_n converging to M, A(C) = M.

We now make use of a very important fact, the lower semi-continuity of length:

If C_n converges to C, then

$$\underline{\lim}\ L(C_n) \geq L(C) \quad .$$

In the present case, we then have

$$L(C) \leq L \quad .$$

The equality sign, however, must hold, since, if $L(C) < L$, C would be magnified into a curve of length L, whose area would then exceed M. Thus the existence of a curve of maximum area and length L is established.

Lower semi-continuity of length is only an example of a property which occurs in all "reasonable" variational problems and is of great importance for the direct methods.

Consider a function $I[\phi]$, where the independent function ϕ ranges over a specific "function space". Consider a sequence of admissible functions ϕ_n which tend to a limiting function u in this function space. Then $I[\phi]$ is called lower semi-continuous at the place u if

$$\lim I[\phi_n] \geq I[u] \quad ,$$

no matter what sequence ϕ_n tending to u is considered.

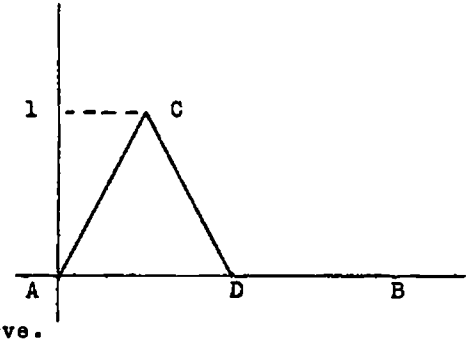
Compactness in Function Space, Arzela's Theorem and Applications

In the ordinary theory of maxima or minima, the existence of a greatest or smallest value of a function in a closed domain is assured by the Bolzano-Weierstrass convergence theorem: a bounded set of points always contains a convergent subsequence. This fact, together with the continuity of the function, serves to secure the existence of an extreme value.

As seen before, in the calculus of variations the continuity of the function space often has to be replaced by a weaker property, semi-continuity; for, when ϕ_n converges to u,

generally $I(\phi_n) \neq I(u)$. Another difficulty in the calculus of variations arises from the fact that the Bolzano-Weierstrass convergence theorem does not hold if the elements of the set are no longer points on a line or in a n-dimensional space, but functions, curves or surfaces.

As an example, we can consider the curves ACDB in a closed interval where AD = 2/n and where C remains at the distance 1 from AB. These curves are continuous, but in the sequence there is no subsequence which converges to an admissible continuous curve.



Fortunately, however, there exists a remedy which very often proves sufficient in the direct methods of the calculus of variations. By a suitable restrictive condition imposed on the functions of a set, one can again obtain a theorem analogous to the Bolzano-Weierstrass theorem. This condition is that of equicontinuity.

A set of functions $f_1(P), \ldots f_n(P)$ is <u>equicontinuous</u> in a domain B if, given any $\varepsilon > 0$, there exists a δ, depending on ε alone, not on the particular $f_n(P)$, such that

$$|f_n(Q) - f_n(P)| < \varepsilon \quad \text{for } |Q - P| < \delta(\varepsilon)$$

uniformly for all P in B and all functions $f_n(P)$. This condition implies, of course, that each function separately be continuous.

As an example, let us consider the set of functions

$$\int_0^1 e^{xy}\, g(y)\, dy = f(x) \quad ,$$

where g(y) is any piecewise continuous function such that $|g(y)| < 1$ and where x has an upper bound. Then

$$|f(x) - f(\xi)| \leq \int_0^1 |g(y)|\ |e^{xy} - e^{\xi y}|\ dy \quad ;$$

as

$$|e^{xy} - e^{\xi y}| < \varepsilon$$

for

$$|x - \xi| < \delta(\varepsilon) \quad ,$$

then

$$|f(x) - f(\xi)| < \varepsilon$$

for $|x-\xi| < \delta(\varepsilon)$, hence $f(x)$ is equicontinuous for all $g(y)$. Of course, e^{xy} can be replaced by any continuous function $K(x,y)$, and the statement of equicontinuity remains true.

As another example, let us consider a sequence $\{f_n(x)\}$ in an interval $(x_0 x_1)$ for which

$$\int_{x_0}^{x_1} [f_n'(x)]^2\ dx \leq M \quad ,$$

where M is a fixed constant. $\{f_n(x)\}$ is equicontinuous, since the expression

$$|f_n(x+h) - f_n(x)|^2 = |\int_x^{x+h} f'(x)\ dx|^2$$

by Schwartz' inequality *, does not exceed

$$h \int_x^{x+h} [f'(x)]^2\ dx \leq Mh \quad ,$$

and we have merely to take, for a given ε,

$$h < \delta(\varepsilon) = \frac{\varepsilon^2}{M} \quad .$$

---

* Schwartz' inequality for integrals states that $[\int_a^b fg]^2 \leq \int_a^b f^2 \cdot \int_a^b g^2$, and is proved from the relation $\int_a^b (f-g)^2 \geq 0$.

Arzela's convergence theorem then states: Given in a closed domain B a set of functions $\{f(P)\}$ which are uniformly bounded (i.e. $|f(P)| < M$) and equicontinuous, then there exists a subsequence $\{f_n(P)\}$ which converges uniformly.

We cover the domain B with a net $\mathcal{H}_1$ of lattice points, and let $\frac{1}{2}$ be the common distance between two consecutive points. The values of f(P) at the first lattice point form a bounded, infinite set of numbers, and hence, according to the Bolzano-Weierstrass theorem, there exists a sequence of these values which converges at the first lattice point. From this sequence we can choose a subsequence of functions which converges at the second lattice point (the lattice points can be ordered), and so forth for all lattice points (there is a finite number of them). We thus obtain a subsequence of functions converging at all the lattice points:

$$S_1(P): \quad f_{1,1}(P), \; f_{2,1}(P), \; \ldots$$

Let us now take the middle points between the lattice points and consider the new set of lattice points thus obtained, $\mathcal{H}_2$. We similarly have a subsequence of functions of $S_1(P)$ converging at all these lattice points:

$$S_2(P): \quad f_{1,2}(P), \; f_{2,2}(P), \; \ldots$$

Continuing this process, we obtain the following sequences:

$$\begin{aligned}
&S_1(P): \quad f_{1,1}(P), \; f_{2,1}(P), \ldots, \; f_{n,1}(P), \ldots \\
&S_2(P): \quad f_{1,2}(P), \; f_{2,2}(P), \ldots, \; f_{n,2}(P), \ldots \\
&\ldots\ldots \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&S_i(P): \quad f_{1,i}(P), \; f_{2,i}(P), \ldots, \; f_{n,i}(P), \ldots
\end{aligned}$$

Each sequence $S_i(P)$ is a subsequence of all the preceding, and it converges at all the lattice points of the net $\mathcal{H}_i$.

We now choose the diagonal sequence $\{f_n(P)\}$:

$$f_{1,1}(P), \; f_{2,2}(P), \ldots, \; f_{n,n}(P), \ldots$$

which, except for a finite number of its terms, is an infinite subsequence of every $S_i(P)$, and hence $\{f_n(P)\}$ converges at all lattice points.

There remains to show that the diagonal subsequence $\{f_n(P)\}$ converges uniformly at any arbitrary point Q of B. ε being given, since the diagonal sequence $\{f_n(P)\}$ is equicontinuous, there exists, uniformly for every k, an integer $N(\varepsilon)$ such that

$$|f_k(Q) - f_k(L)| < \frac{\varepsilon}{3}$$

for $|Q-L| < 2^{-N(\varepsilon)}$, L being a lattice point of the net $\mathfrak{N}_N$.

Since $\{f_n(P)\}$ converges at each of the lattice points, we can find a $\ell(\varepsilon)$ such that

$$|f_n(L) - f_m(L)| < \frac{\varepsilon}{3}$$

for $m,n > \ell(\varepsilon)$.

Then, for any point P

$$\begin{aligned} |f_n(P) - f_m(P)| \leq\ & |f_n(P) - f_n(L)| \\ & + |f_m(L) - f_m(P)| \\ & + |f_n(L) - f_m(L)| \end{aligned}$$

or

$$|f_n(P) - f_m(P)| \leq \varepsilon \quad .$$

Therefore $\{f_n(P)\}$ converges uniformly.

Problem

Prove that, if a set of functions $f_n(P)$ is equicontinuous in a closed domain, and if the $f_n(P)$ are bounded at least at one point, then the functions f_n are bounded everywhere in the domain.

Application to geodesics: Lipschitz's condition: A set of functions are said to satisfy Lipschitz's condition if

$$\left| \frac{f(x_2) - f(x_1)}{x_2 - x_1} \right|$$

is uniformly bounded for all functions $f(x)$, provided that x_1 and x_2 are in a closed interval.

According to Hilbert, the preceding concept immediately permits proof of the existence of a shortest connection between two points A and B on a given surface.

Let $[x(t), y(t), z(t)]$ be the parametric representation of a set of curves through two points A and B, t being the arc length or proportional to the arc length. We also assume that the length of the curves has an upper bound, then we can take $0 \leq t \leq 1$. Then the functions $x(t)$, $y(t)$, $z(t)$ are equicontinuous.

If we assume that the functions possess piecewise continuous derivatives, $\dot{x}(t)$, $\dot{y}(t)$, $\dot{z}(t)$, we have, t being the arc length or proportional to the arc length,

$$\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} = C \quad ,$$

s and σ being two values of t,

$$|x(s) - x(\sigma)| \leq \int_\sigma^s |\dot{x}(t)|dt = \int_\sigma^s \sqrt{[\dot{x}(t)]^2}dt \quad ,$$

which implies

$$|x(s) - x(\sigma)| \leq \int_\sigma^s \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt \quad ;$$

or

$$|x(s) - x(\sigma)| \leq C|s - \sigma| \quad ,$$

which means that, under the assumption made for t, the curves are equicontinuous.

But, even if the derivatives do not exist, the conclusion remains the same, as it can easily be seen by representing the length as the lowest bound of the sum

$$\sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2} \quad ,$$

where the least upper bound is taken with respect to modes of inserting the intermediate points (x_i, y_i, z_i) between A and B on the curve, and for all n.

As the length of the curves has an upper bound, x(t), y(t), and z(t) also have an upper bound, and, as they are equicontinuous, there is, by Arzela's theorem, a subsequence $x_m(t)$ such that $x_m(t)$ converges uniformly. Similar subsequences $y_m(t)$ and $z_m(t)$ also converge. Therefore we can construct a common subsequence $x_m(t)$, $y_m(t)$, $z_m(t)$, for which all three components converge. Hence there is a limit curve C joining A and B.

By the semi-continuity property of length,

$$L(C) \leq \underline{\lim}\, L(C_m) = d \quad ,$$

where d is the greatest lower bound of the length of the curves joining A and B on the surface. Hence the < sign is impossible and

$$L(C) = d \quad .$$

There remains to show that the minimizing curve C has piecewise continuous first and second derivatives and satisfy Euler's equation, provided that the given surface is sufficiently smooth. But this follows from the fact that C must give the shortest length between any two of its points. If B is sufficiently near A, the arc AB of C must be regular and satisfy Euler's equation on the basis of the classical theory of the calculus of variations (See first part of these notes.) Hence this property follows for the entire arc of C between the given end-points.

Problem

1) Prove the theorem used above that: In a sequance of convex curves of bounded length lying in a closed domain there is a subsequence which converges to a closed convex curve.

2) Prove by the method above that there exists on a torus a shortest geodesic topologically equivalent to any prescribed closed circuit on the torus.

Direct variational methods in the theory of integral equations: Following Homgren, we shall indicate briefly how direct variational methods can be used for the treatment of Fredholm and Hilbert integral equations. In the general theory of integral equations, which was inaugurated by Fredholm, Hilbert emphasized the importance of the eigenvalue theory.

Let $K(s,t)$ be a continuous and symmetric function called the "kernel", with $0 \leq s \leq 1$ and $0 \leq t \leq 1$. The eigenvalue problem is to find a function $u(t)$ such that

$$\int_0^1 K(s,t)\, u(t) = \mu u(s) \quad .$$

Such functions u are called eigenfunctions and the $\lambda = 1/\mu$ eigenvalues.

We shall now prove the existence of one eigenvalue. We consider the integral

$$I(\phi,\phi) = \int\int K(s,t)\phi(s)\phi(t)dsdt \quad ,$$

where ϕ is piecewise continuous in the interval $(0,1)$. We assume that $|K(s,t)|$ has an upper bound. (Integrals without indication of limits shall be taken over the domain defined by $0 \leq s \leq 1$ and $0 \leq t \leq 1$.)

We assume that the kernel $K(s,t)$ is such that $I(\phi,\phi)$ is sometimes positive; in other words, that $K(s,t)$ is not a so-called "negative definite" kernel. We write

$$\int [\phi(s)]^2\, ds = (\phi,\phi) \quad .$$

Then

$$\frac{I(\phi,\phi)}{(\phi,\phi)}$$

has a least upper bound μ which is positive, and there exists a maximizing sequence $\phi_1,\ldots,\phi_n,\ldots$, for which

$$\frac{I(\phi,\phi)}{(\phi,\phi)} \to \mu \quad .$$

The function ϕ could be normed, that is $(\phi,\phi) = 1$, but this is not essential.

We now consider the new sequence $\phi_n + \varepsilon\zeta_n$, and we have

$$I(\phi_n + \varepsilon\zeta_n,\ \phi_n + \varepsilon\zeta_n) \leq \mu(\phi_n + \varepsilon\zeta_n,\ \phi_n + \varepsilon\zeta_n)$$

or

$$I(\phi_n,\phi_n) + 2\varepsilon I(\phi_n,\zeta_n) + \varepsilon^2 I(\zeta_n,\zeta_n) \leq \mu[(\phi_n,\phi_n) + 2\varepsilon(\phi_n,\zeta_n) + \varepsilon^2(\zeta_n,\zeta_n)] \quad ,$$

or

$$I(\phi_n,\phi_n) - \mu(\phi_n,\phi_n) + 2\varepsilon[I(\phi_n,\zeta_n) - \mu(\phi_n,\zeta_n)] + \varepsilon^2[I(\zeta_n,\zeta_n) - \mu(\zeta_n,\zeta_n)] \leq 0 \quad .$$

The quadratic form in ε will always be negative if

$$[I(\phi_n,\zeta_n)-\mu(\phi_n,\zeta_n)]^2-[I(\phi_n,\phi_n)-\mu(\phi_n,\phi_n)][I(\zeta_n,\zeta_n)-\mu(\zeta_n,\zeta_n)] \leq 0.$$

We assume that (ζ_n,ζ_n) remains bounded; hence, as $I(\phi_n,\phi_n) - \mu(\phi_n,\phi_n) \to 0$

$$I(\phi_n,\zeta_n) - \mu(\phi_n,\zeta_n) \to 0 \quad ,$$

or

$$\int \zeta_n(s)\left[\int K(s,t)\phi_n(t)dt - \mu\phi_n(s)\right]ds \to 0 \quad .$$

If we write

$$\int K(s,t)\phi_n(t)dt = \eta_n(s) \quad ,$$

the functions $\eta_n(s)$ are equicontinuous and equibounded; hence, by Arzela's convergence theorem, there exists a subsequence, which we shall also call $\eta_n(s)$, of these functions converging uniformly to a continuous function $u(s)$.

$u(s)$ is not identically 0. If it were,

$$\int K(s,t)\phi_n(t)dt$$

would tend to 0. $\phi_n(t)$ can be assumed to be normed, then the upper bound μ would not be positive, which is contrary to the hypothesis; hence $u(s)$ is not identically 0.

With the new notation,

$$\int \zeta_n(s)[\int K(s,t)\phi_n(t)dt - \mu\phi_n(s)]\,ds$$

becomes

$$\int \zeta_n(s)[\eta_n(s) - \mu\phi_n(s)]\,ds \quad ,$$

and it tends to 0. We consider a special variation for ζ:

$$\zeta_n(s) = K(s,r) \quad ,$$

where r is an arbitrary parameter such that

$$0 \leq r \leq 1 \quad .$$

Then

$$\int K(s,r)\eta_n(s)ds - \mu\eta_n(r)$$

tends to 0. But $\eta_n(r)$ converges uniformly to $u(r)$, hence

$$\int K(s,r)u(s)ds = \mu u(r) \quad .$$

If we multiply both sides by $u(r)$ and integrate, we verify immediately that $u(r)$ is a solution.

By the same method we can solve the inhomogeneous Fredholm equation

$$\int K(s,t)u(t)dt - \nu u(s) - g(s) = 0 \quad ,$$

where ν is a constant greater than μ, and $g(s)$ a given function.

Solving this problem is equivalent to solving the following variational problem: Find the function u for which the integral

$$I(\phi) = \iint K(s,t)\phi(s)\phi(t)dsdt - \nu\int[\phi(s)]^2ds - 2\int g(s)\phi(s)ds$$

is a maximum.

The proof is left to the student.

Dirichlet's Principle

Dirichlet's integral. Let G be a domain of the xy-plane, the boundary of which, γ, is a Jordan curve, i.e., a continuous curve without double points. $\phi(x,y)$ is continuous in $G + \gamma$, ϕ_x and ϕ_y being piecewise continuous in G. $\phi = \overline{g}$ on γ, $\overline{g}$ being continuous on γ. Functions ϕ satisfying these conditions form a class of admissible functions. Dirichlet's integral for ϕ is defined as:

$$D(\phi) = \iint_G (\phi_x^2 + \phi_y^2)dxdy = \iint_G (\phi_r^2 + \frac{1}{r^2}\phi_\theta^2)r\,dr d\theta \quad .$$

$D(\phi)$ has a greatest lower bound d, and

$$D(\phi) \geq d \quad .$$

Minimizing sequences. For a variational integral $I(\phi)$ a sequence of admissible functions $\phi_1, \phi_2, \phi_3, \ldots$ such that the values of $I(\phi_1), I(\phi_2), I(\phi_3), \ldots$ tend to the greater lower bound d of $I(\phi)$ is called a minimizing sequence. Wherever the set of possible values of $I(\phi)$ is bounded from below, the existence of a minimizing sequence is insured, even though we need not specify a definite construction of the minimizing sequence. (Such a construction will be the main point in the task (c) of computing numerical values.) For example, Dirichlet's integral having a greatest lower bound d, there is a minimizing sequence of functions $\phi_n(x,y)$ for which $D(\phi_n) \to d$ as $n \to \infty$. However, a minimizing sequence need not converge; even if it converges the limit function need not be admissible.

Explcit expression of Dirichlet's integral for a circle. Hadamard's objection. The difficulties just mentioned are clearly shown by a fact first discovered by Hadamard: not only is the solvability of Dirichlet's variational problem not obvious, but Dirichlet's minimum problem actually is unsolvable in some cases in which the boundary value problem for the differential equation $\triangle u = 0$ can be solved. Thus the idea of reducing the latter to the former seemed even more discredited.

Let K be the unit circle, and introduce polar coordinates r,θ. On the circumference k of K continuous boundary values $\bar{g} = \bar{g}(\theta)$ are given. Consider the (not necessarily convergent) Fourier series of $\bar{g}$:

$$\frac{a_0}{2} + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta) \quad .$$

Then, for $r < 1$, the solution of $\triangle u = 0$ satisfying the boundary condition $\bar{u} = \bar{g}$ on k is given by the convergent series:

$$u(r,\theta) = \frac{a_0}{2} + \sum_{\nu=1}^{\infty} r^\nu (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta) \quad .$$

If we represent by $D_\rho(u)$ Dirichlet's integral for the circle of radius $\rho < 1$ about the origin,

$$D_\rho(u) = \pi \sum_{\nu=1}^{\infty} \nu(a_\nu^2 + b_\nu^2)\rho^{2\nu} \quad .$$

This implies for every N

$$\pi \sum_{\nu=1}^{N} (a_\nu^2 + b_\nu^2)\rho^{2\nu} \leq \pi \sum_{\nu=1}^{\infty} \nu(a_\nu^2 + b_\nu^2) \quad ,$$

where the right-hand side may be a divergent series. By letting ρ tend to 1, we infer immediately: Dirichlet's integral for the harmonic function

$$u(r,\theta) = \frac{a_0}{2} + \sum_{\nu=1}^{\infty} r^\nu (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta)$$

over the unit circle is given by the series

$$D(u) = \pi \sum_{\nu=1}^{\infty} \nu(a_\nu^2 + b_\nu^2)$$

and exists if, and only if, this series converges.

Now, as pointed out by Hadamard, there exist continuous functions $\bar{g}(\theta)$ for which this series diverges; e.g. if $\bar{g}(\theta)$ is given by the uniformly convergent Fourier expansion

$$\bar{g} = \sum_{\mu=1}^{\infty} \frac{\sin \mu 1\theta}{\mu^2} \quad .$$

Then

$$D(u) = \pi \sum_{\mu=1}^{\infty} \frac{\mu 1}{\mu^4} \quad ,$$

which does not converge. With boundary values such as this $\bar{g}$, the boundary value problem of $\triangle u = 0$ can therefore certainly not be reduced to a variational problem for Dirichlet's integral, and Dirichlet's principle is invalid. No full equivalence between the variational problem and the boundary value problem exists.

The Correct Formulation of Dirichlet's Principle. The last difficulty can be avoided by restricting the prescribed boundary value in such a manner as not to exclude from the outset the solvability of the variational problem. While for the boundary value problem such conditions are not necessary, they are essential to make the variational problem meaningful. Accordingly, it will be assumed that the prescribed boundary values $\bar{g}$ are the values on γ of a function g in $G + \gamma$ for which $D[g]$ is finite. In other words, we explicitly assume that there exists at least one admissible function with a finite Dirichlet integral. Thus one is led to the following formulation:

Dirichlet's Principle. Given a domain G whose boundary γ is a Jordan curve, let g be a function continuous in $G + \gamma$, piecewise smooth in G, and with a finite Dirichlet integral $D(g)$. Let ϕ be the class of all functions continuous in $G + \gamma$, piecewise smooth in G, and with the same boundary values as g. Then the problem of finding a function for which $D(\phi) = \text{minimum} = d$, has a unique solution $\phi = u$. This function u is the solution of the boundary value problem of $\triangle u = 0$ with the values $\bar{g}$ on γ.

Lower semi-continuity of Dirichlet's integral for harmonic functions. If a sequence of harmonic functions u_n converges to a harmonic function u uniformly in every closed subdomain of G, then

$$D_G(u) \leq \underline{\lim}\, D_G(u_n) \quad .$$

Proof: For any closed subdomain G' of G the assumed uniform convergence of the u_n implies by Harnack's theorem * the uniform convergence of the derivatives of u_n to those of u. Hence

$$D_{G'}(u) = \lim_{n\to\infty} D_{G'}(u_n) \leq \underline{\lim}\, D_G(u_n)$$

and, by letting G' tend to G,

$$D_G(u) \leq \underline{\lim}\, D_G(u_n) \quad .$$

Proof of Dirichlet's Principle for the Circle. Let the domain G be the unit circle and consider the Fourier series of the given boundary function $\overline{g} = g(r,\vartheta)$

$$\frac{a_o}{2} + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu\vartheta + b_\nu \sin \nu\vartheta) \quad .$$

This series need not converge, but for $r < 1$ the series

$$u = \frac{a_o}{2} + \sum_{\nu=1}^{\infty} r^2(a_\nu \cos \nu\vartheta + b_\nu \sin \nu\vartheta)$$

does converge, and u is harmonic.

Let v be any other admissible function for which D(v) is finite, and $\zeta = u - v$. On the boundary $\zeta = 0$.

$$D(v) = D(u - \zeta) = D(u) + D(\zeta) - 2D(u,\zeta) \quad ,$$

where $D(u,\zeta)$ represents the so-called bilinear form

$$D(u,\zeta) = \iint_G (u_x\zeta_x + u_y\zeta_y)\, dxdy \quad .$$

By Green's formula

$$D(u,\zeta) = -\iint_G \zeta\Delta u\, dxdy + \int_\gamma \zeta \frac{\partial u}{\partial r}\, ds \quad .$$

---

* Harnack's theorem states: If a sequence of harmonic functions converges uniformly in a domain, then their derivatives converge uniformly in every closed subdomain and the limit function is again harmonic.
For a proof, see, e.g., Foundations of Potential Theory by O. D. Kellogg, p. 248.

The first term of the right side vanishes because $\triangle u = 0$, and the last term vanishes too because $\zeta = 0$ on γ. Hence

$$D(u,\zeta) = 0$$

and, as

$$D(\zeta) > 0 \quad ,$$

$$D(v) \geq D(u) \quad ,$$

which means that Dirichlet's integral is minimum for u.

This reasoning, however, has a gap, for Green's formula is not applicable to u in the whole domain. In order to fill this gap, consider the "harmonic polynomials"

$$u_n = \frac{a_0}{2} + \sum_{\nu=1}^{n} r^2(a_\nu \cos \nu\theta + b_\nu \sin \nu\theta)$$

and write $\zeta_n = u_n - v$. Then

$$D(v) = D(u_n) + D(\zeta_n) - 2D(u_n,\zeta_n) \quad .$$

By Green's formula, applicable to the polynomials u_n,

$$D[u_n,\zeta_n] = -\iint_G \zeta_n \triangle u_n dxdy + \int_0^{2\pi} \zeta_n \frac{\partial u_n}{\partial r} d\theta$$
$$r=1$$

The first term of the right side vanishes because $\triangle u_n = 0$. It can be seen that the second term vanishes too by substituting for u_n its explicit expression, and by observing that the first 2n+1 Fourier coefficients of ζ_n are zero, so that

$$\int_0^{2\pi} \zeta_n d\theta = 0 \ , \quad \int_0^{2\pi} \zeta_n \cos \nu\theta d\theta = 0, \quad \int_0^{2\pi} \zeta_n \sin \nu\theta d\theta = 0,$$

for $\nu = 1,2,\ldots,n$. Hence

$$D(u_n,\zeta_n) = 0 \quad ,$$

and

$$D(v) = D(u_n) + D(\zeta_n)$$

or $$D(v) \geq D(u_n)$$

In G the polynomials u_n tend to the harmonic function u with the boundary value $\bar{g}$, i.e., $\lim_{n\to\infty} u_n = u$, and the convergence is uniform in every closed subdomain of G. The semi-continuity of the Dirichlet integral for harmonic functions leads to

$$D(u) \leq \underline{\lim}\, D(u_n) \quad .$$

This and the preceding inequality imply

$$D(u) \leq D(v) \quad ,$$

which proves the minimum property of D(u).

To prove the uniqueness, let ζ be a function for which $D(\zeta)$ is finite and $\bar{\zeta} = 0$. Then, because of the minimum property of u,

$$D(u+\varepsilon\zeta) = D(u) + 2\varepsilon D(u,\zeta) + \varepsilon^2 D(\zeta) \geq d \quad ,$$

for ε positive or negative, which implies that

$$D(u,\zeta) = 0 \quad .$$

Let v be any function other than u, and $\zeta = v - u$, then

$$D(v) = D(u) + D(v-u)$$

Hence

$$D(u) < D(v) \quad ,$$

except for $D(v-u) = 0$, i.e., $v - u = \text{const}$. But u and v having the same boundary values

$$v - u = 0 \quad .$$

<u>"Distance" in function space</u>. <u>Triangle inequalities</u>. Let ϕ and ψ be two admissible functions, and write

$$D(\phi,\psi) = \iint_G (\phi_x\psi_x + \phi_y\psi_y)\, dxdy \quad .$$

With this notation

$$D(\phi) = D(\phi,\phi)$$

and

$$D(\phi + \psi) = D(\phi) + D(\psi) + 2D(\phi,\psi) \quad .$$

As $D(\phi)$ has a greatest lower bound, there exists a minimizing sequence $\phi_1,\phi_2,\ldots$ Consider the new sequence $\phi_n + \varepsilon\zeta_n$. These functions will be admissible if ζ_n and its derivatives satisfy certain conditions of continuity and if ζ_n vanishes at the boundary.

$$D(\phi_n + \varepsilon\zeta_n) = D(\phi_n) + 2\varepsilon D(\phi_n,\zeta_n) + \varepsilon^2 D(\zeta_n) \geq d \quad .$$

We assume that $D(\zeta_n)$ has an upper bound R. If we write $D(\phi_n) = d_n$ and $D(\phi_n,\zeta_n) = V_n$, the above inequality can be written

$$D(\phi_n + \varepsilon\zeta_n) = d_n + 2\varepsilon V_n + \varepsilon^2 D(\zeta_n) \geq d$$

or, a fortiori,

$$d_n + 2\varepsilon V_n + \varepsilon^2 R \geq d$$

or

$$\varepsilon^2 R + 2\varepsilon V_n + d_n - d \geq 0 \quad .$$

This inequality is satisfied for any ε, if the quadratic form is positive definate, that is if

$$V_n^2 - R(d_n - d) \leq 0 \quad ,$$

which implies that

$$V_n = D(\phi_n,\zeta_n) \text{ tends uniformly to } 0 \text{ as } n \to \infty \quad .$$

We now take a special function $\zeta_n = \phi_n - \phi_m$, where m is fixed. Then

$$|D(\phi_n,\ \phi_n - \phi_m)| \leq \sqrt{R(d_n - d)}$$

and also

$$|D(\phi_m, \phi_m - \phi_n)| \leq \sqrt{R(d_m - d)} \quad .$$

By subtracting the left-hand sides,

$$D(\phi_m - \phi_n) \leq \sqrt{R(d_m - d)} + \sqrt{R(d_n - d)} \quad .$$

Hence

$$D(\phi_m - \phi_n) \to 0 \quad \text{as} \quad m,n \to \infty \quad .$$

$D(\phi - \psi)$ can be considered as the "distance" between two functions ϕ and ψ in the function space under consideration. For these "distances" the following so-called triangle inequalities hold:

$$\sqrt{D(\phi)} + \sqrt{D(\psi)} \geq \sqrt{D(\phi + \psi)}$$

$$\sqrt{D(\phi)} - \sqrt{D(\psi)} \leq \sqrt{D(\phi - \psi)}$$

The fact that $D(\phi - \phi_n) \to 0$ as $n \to \infty$ does not imply that ϕ_n converges to ϕ. If we take $\phi = 0$, the following example will show that, although $D(\phi_n) \to 0$ as $n \to \infty$, ϕ_n need not tend to 0.

Consider the minimum problem for $D(\phi)$ in a circle of radius 1, when the admissible functions are to vanish on the boundary. This minimum problem is solved by $\phi \equiv 0$ and by no other function; $d = 0$ is the minimum value, not merely the greatest lower bound.

Now we define in polar coordinates r,θ a sequence of admissible functions ϕ_n by

$$\phi_n = \begin{cases} 1 & , \quad \text{for } r \leq \rho_n^2 \\ \dfrac{\log r}{\log \rho} - 1 & , \quad \text{for } \rho_n^2 \leq r \leq \rho_n \\ 0 & , \quad \text{for } \rho_n \leq r \leq 1 \end{cases} ,$$

Then

$$D(\phi_n) = \int_0^{2\pi}\int_{\rho_n^2}^{\rho_n} (\frac{\partial \phi_n}{\partial r})^2 r \, dr d\theta = 2\pi \int_{\rho_n^2}^{\rho_n} (\frac{1}{r \log \rho_n})^2 r \, dr = - \frac{2\pi}{\log \rho_n} .$$

We now take $\rho_n = \frac{1}{n}$. Then

$$D(\phi_n) \to 0 \quad \text{as} \quad n \to \infty ,$$

but ϕ_n does not tend to 0 at the center of the circle, where its value is 1.

Thus a minimizing sequence cannot in general be expected to yield the solution of the problem by a mere passage to the limit. The essential point in the "direct variational methods" is to introduce an appropriate sequence that will guarantee convergence.

Construction of a harmonic function u by a "smoothing process". Consider a minimizing sequence ϕ_n of admissible functions in G. As seen by the example above, the functions ϕ_n need not converge.

We consider a circle K in G, and we replace a function ϕ_n by a function ω_n such that $\omega_n = \phi_n$ outside of K and ω_n is harmonic in K. By this operation we are "smooting out" ϕ_n. Dirichlet's principle having been proved for the circle,

$$D_K(\omega_n) \leq D_K(\phi_n)$$

and therefore

$$D(\omega_n) \leq D(\phi_n) .$$

The functions ω_n are admissible functions and form a new minimizing sequence. Hence

$$D(\omega_n - \omega_m) \to 0 \quad \text{as} \quad n \text{ and } m \to \infty .$$

The aim will be to prove that $\omega_n - \omega_m$ converges in the interior of K, while $\phi_n - \phi_m$ need not converge. We will use the limit function in order to construct the solution of the Dirichlet problem. Before carrying out this program we need several lemmas about harmonic functions.

Preparations: Let $u(x,y)$ be a continuous function with continuous first and second derivatives, satisfying the differential equation

$$u_{xx} + u_{yy} = 0$$

then

$$u(P) = \frac{1}{2\pi} \int_C u(\overline{P})\, d\theta$$

where the integral is taken along a circle C about P.

By Gauss's theorem

$$\int_C (u_{xx} + u_{yy})dxdy = \int_C (u_x \frac{\partial x}{\partial r} + u_y \frac{\partial y}{\partial r})ds = 0$$

or

$$\int_C \frac{\partial u(\overline{P})}{\partial r}\, rd\theta = 0 \quad ;$$

or, because of the continuity of the first derivatives of u,

$$\frac{d}{dr} \int_C u(P)\, d\theta = 0 \quad ,$$

which means that

$$\frac{1}{2\pi} \int_C u(P)d\theta = \text{const.} \quad \text{with respect to } r \quad .$$

If we let r tend to 0, we see that the value of the constant is $u(P)$, hence

$$u(P) = \frac{1}{2\pi} \int_C u(\overline{P})\, d\theta \quad .$$

We shall now prove the converse, namely, that if a function, which is assumed to be continuous, satisfies the mean value equality for any radius, then the function is harmonic.

$$u(P) = \frac{1}{2\pi} \int_0^{2\pi} u(\overline{P})\, d\theta$$

for every r.

We multiply both sides by r and integrate with respect to r. Hence

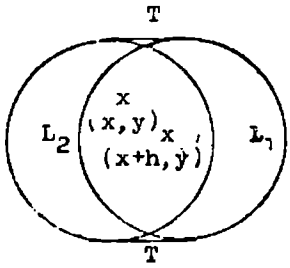
$$u(P) = \frac{1}{\pi r^2} \iint u(\bar{P})\, dxdy \quad ,$$

which can be considered as the mean value theorem for the area (the double integral is taken over the circle of radius r about P).

In order to prove that u is differentiable, we consider the expression:

$$\pi r^2 \frac{u(x+h,y) - u(x,y)}{h} = \frac{1}{h} \iint_{L_1} u(\bar{P})dxdy - \frac{1}{h} \iint_{L_2} u(\bar{P})dxdy \quad ,$$

where the first integral is taken over the lune L_1 and the second over the lune L_2 between the circles of radius r about the points (x,y) and (x+h,y).



Denoting the two triangular regions (see figure) by T we have

$$\frac{1}{h} \iint_{L_1+T} u(\bar{P})dxdy = \frac{1}{h} \int_{-r}^{r} dy \int_{0}^{h} u(\bar{P})dx = \int_{-r}^{r} u(\tilde{P})dy \quad ,$$

where $\tilde{P}$ is some point in L_1+T on the segment parallel to the x-axis. Because of the continuity of the function u,

$$\int_{-r}^{r} u(P)dy \to \int_{-r}^{r} u(\bar{P})dy \qquad \text{as } h \to 0 \quad ,$$

and similarly for the other lune. Hence, as $h \to 0$,

$$\frac{1}{h} \iint_{L_1} u(\bar{P})dxdy - \frac{1}{h} \iint_{L_2} u(\bar{P})dxdy = \frac{1}{h} \iint_{L_1+T} u(\bar{P})dxdy - \frac{1}{h} \iint_{L_2+T} u(\bar{P})dxdy$$

$$\to \int_{C} u(\bar{P})dy \quad ,$$

where the integral on the right-hand side is taken along the circle C of radius r about the point (x,y). Hence

$$\pi r^2 \frac{u(x+h,y) - u(x,y)}{h} \to \int_C u(\bar{P})dy \quad \text{as } h \to 0 \ ,$$

and u possesses a derivative u_x given by the integral

$$\frac{1}{\pi r^2} \int_C u(\bar{P})dy \ .$$

By Gauss's theorem

$$u_x = \frac{1}{\pi r^2} \int_C u(\bar{P})dy = \frac{1}{\pi r^2} \iint_C u_x \, dxdy \ ,$$

which means that u_x possesses the same mean value property as u. Of course, the same argument holds for u_y. Hence the function u has continuous derivatives of every order.

We shall now prove that u is harmonic. Since u has continuous derivatives, we can write

$$u(x + h\cos\theta, y + h\sin\theta) = u(x,y) + h\cos\theta u_x + h\sin\theta u_y$$
$$+ h^2(u_{xx}\cos^2\theta + 2u_{xy}\sin\theta\cos\theta + u_{yy}\sin^2\theta) + h^3R \ ,$$

where R is bounded.

Integrating both sides with respect to θ along the circle C of radius h about P, we obtain

$$\int_C u(\bar{P})d\theta = 2\pi\, u(x,y) + \pi h^2(u_{xx} + u_{yy}) + h^3M \ ,$$

where M is bounded. As the mean value property has been assumed for the function u,

$$\pi h^2(u_{xx} + u_{yy}) + h^3M = 0 \ ,$$

or

$$(u_{xx} + u_{yy}) + hR = 0 \ ,$$

and, by letting $h \to 0$,

$$u_{xx} + u_{yy} = 0 \quad ,$$

which proves that u is harmonic.

As an application, we can prove that if a sequence of harmonic functions u_n converges uniformly to u, then the limit function u is harmonic. (We have made use of that theorem previously.)

$$u = \lim_{n\to\infty} u_n = \lim_{n\to\infty} \frac{1}{2\pi} \int_0^{2\pi} u_n(\bar{P})d\theta \quad ;$$

but, because of the uniform convergence, we can interchange the order of the operations of integration and passing to the limit, and we obtain

$$u = \frac{1}{2\pi} \int_0^{2\pi} \lim_{n\to\infty} u_n(\bar{P})d\theta = \frac{1}{2\pi} \int_0^{2\pi} u(\bar{P})d\theta \quad ,$$

hence u is harmonic.

Lemma: Let u be harmonic in K and let K' be a closed subdomain of K. Then there exists a constant C depending on K,K' only, such that

$$|u(P)| \leq |u(P_0)| + C\sqrt{D_K(u)} \quad \text{for} \quad P,P_0 \quad \text{in } K' \quad .$$

Consequence: Let u_n be a sequence of harmonic functions with $D_K(u_n) \to 0$ for $n \to \infty$. In order to prove that u_n converges to zero uniformly in a closed subdomain K' of K it is sufficient to prove $u_n(P_n) \to 0$ for some point P_n in K'. Namely then from the lemma

$$|u_n(P)| \leq u_n(P_n) + C\sqrt{D_K(u_n)}$$

follows $u_n(P) \to 0$.

Proof of the Lemma: Since u is harmonic in K, the same holds for u_x, u_y. Therefore

$$u_x(P) = \frac{1}{\pi h^2} \iint_{C_h} u_x(\bar{P})\, dxdy$$

where C_h is a circle of radius h about P. We take $P \in K'$ and choose h so small that C_h is contained in K. h depends on K,K' only. Schwarz' inequality gives

$$u_x^2(P) \leq \frac{1}{\pi h^2} \iint_{C_h} u_x^2 \, dxdy$$

and

$$u_x^2(P) + u_y^2(P) \leq \frac{1}{\pi h^2} D_{C_h}(u) \leq \frac{1}{\pi h^2} D_K(u) \quad .$$

Let P_o be any other point in K'. We have

$$u(P) = u(P_o) + \int_{P_o}^{P} u_x \, dx + u_y \, dy \quad .$$

Hence if L is an upper bound for the distance of P, P_o in K', we obtain

$$|u(P)| \leq |u(P_o)| + \frac{L}{\sqrt{\pi} h} \sqrt{D_K(u)} \quad .$$

Convergence of ω_n: Let ϕ_n be a minimizing sequence of admissible functions in G. In order to obtain a convergent sequence from the ϕ_n we considered a circle K in G and replaced the function ϕ_n by

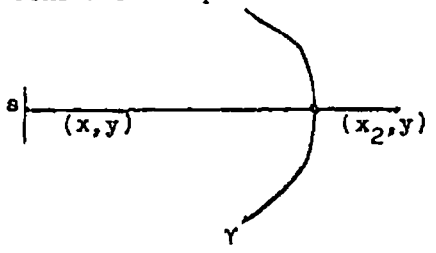
$$\omega_n = \begin{cases} \phi_n & \text{in G, outside K} \\ \text{harmonic in K} & . \end{cases}$$

This requires that one can solve the Dirichlet problem for a circle, which we have done already. Then the ω_n also form a minimizing sequence since $D(\omega_n) \leq D(\phi_n)$. Furthermore with $\sigma_{nm} = \omega_n - \omega_m$

$$D(\sigma_{nm}) \to 0 \qquad \text{as } n,m \to \infty \quad .$$

From this we want to derive that $\sigma_{nm} \to 0$ uniformly in any closed subdomain K' of K. According to the Lemma just proved it is sufficient to find a point P_{nm} in K' for which $\sigma_{nm}(P_{nm}) \to 0$ as $n,m \to \infty$.

Through a point (x_1,y) in K' we consider the parallel to the x-axis, and let (x_2,y) be its intersection with the boundary γ of G. As $\sigma_{nm}(x_2,y) = 0$, we can write



$$\sigma_{nm}(x_1,y) = \sigma_{nm}(x_1,y) - \sigma_{nm}(x_2,y) = \int_{x_2}^{x_1} \sigma_{nm_x}(x,y)\, dx \quad ,$$

which implies by Schwarz's inequality

$$\sigma^2_{nm}(x_1,y) \leq |x_1 - x_2| \int_{x_2}^{x_1} \sigma^2_{nm_x}(x,y)\, dx \quad ,$$

or, as G is bounded

$$\sigma^2_{nm}(x_1,y) \leq L \int_{x_2}^{x_1} \sigma^2_{nm_x}(x,y)\, dx \quad ,$$

where L is a given length. By integrating between two values of y, y_1 and y_2, in K',

$$\int_{y_2}^{y_1} \sigma^2_{nm}(x_1,y)\, dy \leq L \int_{y_2}^{y_1}\int_{x_2}^{x_1} \sigma^2_{nm_x}(x,y)\, dxdy$$

or, a fortiori,

$$\int_{y_2}^{y_1} \sigma^2_{mn}(x,y)\, dy \leq L\, D(\sigma_{mn}) \quad .$$

The integral of the left-hand side is taken along a segment of the straight line s in K' parallel to the y-axis. Hence at some point of s

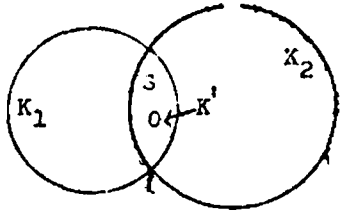
$$\sigma_{nm}^2 \leq \frac{L\, D\,(\sigma_{mn})}{|y_2 - y_1|}$$

Hence $\sigma_{nm} = \omega_n - \omega_m$ tends uniformly to 0, and ω_n tends uniformly to a function u. u is a harmonic function because of the theorem: If in a circle a sequence of harmonic functions converges to a limit, the limit is a harmonic function.

If we let now K' tend to K, we have defined a harmonic function u in K.

The smoothing process can be applied to any circle in G; it leads in every such circle to the definition of a certain harmonic function u. We assert that this construction defines a uniquely determined function in the whole domain G. For the proof we need only show that the functions u_1 and u_2 resulting from the smoothing in two overlapping circles K_1 and K_2 are identical in the common part S of these two circles.

Let ω_n^1 and ω_n^2 be the minimizing sequences originating from sequence ϕ_n be smoothing in the circle K_1 and K_2 respectively. Then the mixed sequence $\omega_1^1, \omega_1^2, \omega_2^1, \omega_2^2, \ldots$ is also a minimizing sequence, and therefore $D_s(\omega_n^1 - \omega_n^2) \to 0$ as $n \to \infty$. If K' is a circle in S, a fortiori, $D_{K'}(\omega_n^1 - \omega_n^2) \to 0$. ω_n^1 and ω_n^2 are harmonic in K' and converge to u_1 and u_2 respectively. From the preceding argument it follows that the mixed sequence also converges to a harmonic function u' in K'. Therefore u_1 and u_2 are identical with u' in K'.



Proof that $D(u) = d$. Let G' be a closed subdomain of G. It can be covered by a finite number of circles K_1, and by taking the circles small enough, we can be sure that they all lie entirely in G. Then

$$G' \leq \sum_1 K_1 \subset G$$

and hence

$$D_{G'}(\phi_n - u) \leq \sum_1 D_{K_1}(\phi_n - u) \quad .$$

By the triangle inequality

$$\sqrt{D_{G'}(u)} \leq \sqrt{D_{G'}(\phi_n - u)} + \sqrt{D_{G'}(\phi_n)} \leq \sqrt{D_{G'}(\phi_n - u)} + \sqrt{D(\phi_n)} .$$

The left side of this inequality is independent of n, therefore

$$\sqrt{D_{G'}(u)} \leq \text{g.l.b.} \left(\sqrt{D_{G'}(\phi_n - u)} + \sqrt{D(\phi_n)} \right) \quad ,$$

the g.l.b. being taken over the values of n. But

$$D_{G'}(\phi_n - u) \to 0 \quad \text{and} \quad D(\phi_n) \to d \quad \text{as} \quad n \to \infty \quad ,$$

therefore

$$\sqrt{D_{G'}(u)} \leq \sqrt{d} \quad , \quad \text{i.e.} \quad D_{G'}(u) \leq d \quad .$$

Consequently, since

$$D(u) = \lim_{G' \to G} D_{G'}(u) \quad ,$$

$$D(u) \leq d \quad ,$$

However, u being an admissible function,

$$D(u) \geq d \quad ,$$

hence

$$D(u) = d \quad .$$

Proof that the function u attains the prescribed boundary values. The function u, so far defined in G only, can be extended continuously to γ and assumes the prescribed boundary values $\bar{g}$ on γ. More precisely: there is a quantity $\varepsilon(\delta)$ tending to 0 with δ such that for all points P in G we have $|u(P) - g(R)| < \varepsilon$ if the distance PR from P to a point R on γ is less than δ. This statement expresses the uniform convergence of u(P) to the prescribed boundary values.

For the proof we assume that R is one of the points on γ nearest to P, at the distance $PR = 5h < \delta$.

We now consider the circle of radius 10h about R. This circle defines in ω a certain subregion L.

We shall show first that h can be taken so small that $D_L(\phi_n)$ is smaller than any prescribed positive number σ_1

$$D_L(\phi_n) < \sigma_1^2 \qquad \text{for } \underline{\text{all}}\ n \text{ and for } h < h(\sigma_1),$$

As $D(\phi_n - \phi_m) \to 0$, we can choose N so large that

$$D_L(\phi_n - \phi_m) \le D(\phi_n - \phi_m) < \frac{\sigma_1^2}{4}$$

for $n,m \ge N$; having fixed N we can choose h so small that

$$D_L(\phi_\nu) < \frac{\sigma_1^2}{4} \qquad \text{for } \nu \le N \ .$$

By the triangle inequality

$$\sqrt{D_L(\phi_n)} \le \sqrt{D_L(\phi_n - \phi_N)} + \sqrt{D_L(\phi_N)} \le \frac{\sigma_1^2}{4} + \frac{\sigma_1^2}{4} < \sigma_1^2 \ .$$

This proves the statement. Hence this is a positive function $\sigma_1(h)$, tending to zero for $h \to 0$, independent of n, such that

$$D_L(\phi_n) \le \sigma_1^2(h) \ .$$

We now make use of the following

Lemma: There is a circle $r = \tilde{r}$ about P, with $3h \le \tilde{r} \le 4h$, such that for two points M_1 and M_2 on each connected arc of this circle the inequality

$$|\phi(M_2) - \phi(M_1)|^2 \le \frac{2\pi\tilde{r}\sigma_1^2}{h} \le 8\pi\sigma_1^2 \ ,$$

where σ_1^2 is an upper bound of $D_L(\phi_n)$, holds.

In polar coordinates this last fact can be written

$$\iint_L (\phi_r^2 + \frac{1}{r^2}\phi_\vartheta^2) r \, d\vartheta dr \le \sigma_1^2 \ ,$$

or, by writing $r\theta = s$, and dropping the integral over ϕ_r^2

$$\int_{3h}^{4h} dr \int \phi_s^2 \, ds \leq \sigma_1^2 \quad .$$

By the mean value theorem for integrals there will be a value $r = \tilde{r}$, with $3h \leq \tilde{r} \leq 4h$, for which

$$\int_0^{2\pi r} \phi_s^2 \, ds \leq \frac{\sigma_1^2}{h} \quad .$$

On this circle $r = \tilde{r}$ we have for the oscillation of ϕ between two points M_1 and M_2

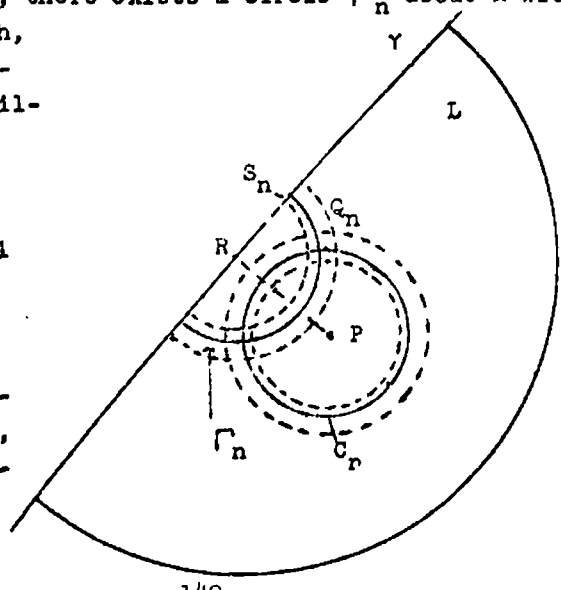
$$|\phi(M_2) - \phi(M_1)|^2 \leq 2\pi\tilde{r} \frac{\sigma_1^2}{h} \leq 8\pi\sigma_1^2 \quad ,$$

which proves the lemma. This can be written

$$\phi_n(M_2) - \phi_n(M_1) \leq 2\sqrt{2\pi}\, \sigma_1(h)$$

where $\sigma_1(h)$ is independent of n. M_1, M_2 are two points on a circle C_n which now depends on n. If its radius is denoted by $\tilde{r}_n$ we have $3h \leq \tilde{r}_n \leq 4h$.

Similarly, there exists a circle Γ_n about R with radius ρ_n, $3h < \rho_n < 4h$, on whose circumference the oscillation of ϕ_n is again less than $2\sqrt{2\pi}\sigma_1(h)$. This circle will intersect the circle C_n about P at a point Q_n and if n is sufficiently small, it will also intersect γ at a



point S_n so that Q_n and S_n are connected by an arc of this circle in G. Therefore, since $\overline{\phi}_n = \overline{g}$,

$$|\phi_n(Q_n) - g(S_n)| \leq 2\sqrt{2\pi}\,\sigma_1(h) \quad .$$

Let u_n be the harmonic function in the circle C_n about P obtained by smoothing ϕ_n; then the value $u_n(P)$ coincides with some value of ϕ_n on the circle C_n * and hence cannot differ from $\phi_n(Q_n)$ by more than $2\sqrt{2\pi}\sigma_1(h)$:

$$|u_n(P) - \phi_n(Q_n)| \leq 2\sqrt{2\pi}\sigma_1(h) \quad .$$

Finally, since g is continuous on γ, h can be taken such that

$$|g(S_n) - g(R)| \leq \sigma_2(h) \quad ,$$

where $\sigma_2(h) \to 0$ as $h \to 0$. By combining the three inequalities

$$|u_n(P) - g(R)| \leq 4\sqrt{2\pi}\sigma_1(h) + \sigma_2(h) \quad ,$$

which proves that u attains the boundary values.

Thus u is recognized as the solution of the boundary value problem. Since it has been proved that $D(u) = d$ and since u is admissible in the variational problem, it solves the variational problem. The proof of the uniqueness of this solution of the variational problem is now exactly the same as in the case of the circle. Hence Dirichlet's principle, as we stated it above, is established.

Alternative proof of Dirichlet's principle: We make the same assumption as at the beginning of the first proof, and we consider a minimizing sequence of admissible functions $\phi_1, \phi_2, \ldots, \phi_n, \ldots$

At a point P we consider the circle C of radius h about P. This circle will lie entirely in G for all points P whose distance from the boundary is greater than h. At P we consider the function

---

* by the mean value theorem.

$$\omega_n(P) = \frac{1}{\pi h^2} \iint_c \phi_n(\bar{P})\, dxdy \quad .$$

The functions ϕ_n being continuous, the functions ω_n are differentiable.

We shall now prove that, for a fixed h, $\omega_n(P)$ tends uniformly to u(P).

For the admissible functions

$$D(\phi_n - \phi_m) \to 0 \qquad \text{as } n,m \to \infty$$

and

$$D(\phi_n, \zeta_n) \to 0 \qquad \text{uniformly as } n \to \infty$$

provided that $D(\zeta_n)$ remains bounded, the functions ζ_n have piecewise continuous first derivatives, and $\bar{\zeta}_n = 0$.

Let us write

$$H(\zeta) = \iint_G \zeta^2\, dxdy \quad .$$

We will prove that

$$H(\zeta) \leq CD(\zeta) \quad , \qquad \text{where C is a positive constant.}$$

We remark that we need to prove the proposition only for a square. G being bounded, we can surround it by a square and consider a function ζ^* such that

$$\zeta^* = \zeta \quad \text{in G}$$

and

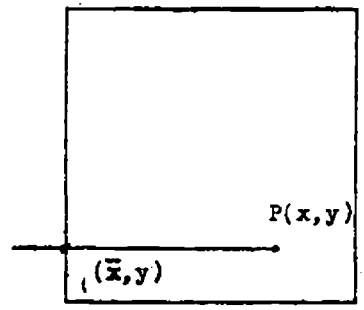
$$\zeta^* = 0 \quad \text{outside of G but inside the square surrounding G.}$$

Let P(x,y) be a point in a square, we take x and y-axis parallel to the sides of the square and let $\bar{x}$ be the point where the parallel to the x-axis through P intersects the side of the square.

Then, as $\zeta(\bar{x},y) = 0$,

$$\zeta(x,y) = \int_{\bar{x}}^{x} \zeta_x\, dx \quad .$$

(We remark that the parallel to the x-axis through P must cut the boundary of G at a finite number of points at most, in order that the discontinuity of the derivative ζ occurs at a finite number of points at most. This is the only condition imposed upon the boundary of G.)



By Schwarz's inequality

$$\zeta(P) = \int_{\bar{x}}^{x} \zeta_x \, dx$$

implies that

$$\zeta^2(P) \leq \ell \int_L \zeta_x^2 \, dx \quad ,$$

where L is the side of the square or, integrating with respect to y,

$$\int_L \zeta^2(P) \, dy \leq \ell D(\zeta)$$

and, integrating now with respect to x,

$$\iint \zeta^2 \, dxdy \leq \ell^2 \, D(\zeta) \quad ,$$

which is the proposition we wanted to prove.

Let R be a point on the boundary γ and consider a circle of radius ℓ about R. We can always take ℓ small enough, so that the subdomain L of G limited by the circle is a simply connected part of G.

We shall prove that

$$H_L(\zeta) \leq 4\pi^2 \ell^2 \, D_L(\zeta) \quad .$$

Consider a circle C_r of radius r, $r \leq \ell$, about R. As $\xi = 0$,

$$\zeta(P) = \int \zeta_s \, ds \quad ,$$

where P is a point on C_r and where the integral is taken from the boundary to P along C_r.

Hence, by Schwarz's inequality

$$\zeta^2(P) \leq 2\pi \ell \int \zeta_s^2 \, ds \quad ,$$

where the integral is taken along the arc of C_r in L, or a fortiori, integrating both sides with respect to s,

$$\int \zeta^2(P) \, ds \leq 4\pi^2 \ell^2 \int \zeta_s^2 \, ds \quad ,$$

and, integrating now with respect to r,

$$H_L(\zeta) \leq 4\pi^2 \ell^2 \, D_L(\zeta) \quad .$$

For the functions ω_n defined above,

$$\pi h^2 [\omega_n(P) - \omega_m(P)] = \iint_C [\phi_n(\overline{P}) - \phi_m(\overline{P})] \, dxdy \quad ,$$

where the double integral is taken over a circle C of radius h about P. But this last equality implies that

$$\left\{\pi h^2 [\omega_n(P) - \omega_m(P)]\right\}^2 \leq \pi h^2 H_C(\phi_n - \phi_m) \quad .$$

But

$$H_C(\phi_n - \phi_m) \leq H_G(\phi_n - \phi_m)$$

and, as $\phi_n - \phi_m = 0$ on the boundary γ,

$$H_G(\phi_n - \phi_m) \leq \mathfrak{D} D(\phi_n - \phi_m)$$

by the lemma we have proved above.

ϕ_n and ϕ_m being admissible functions,

$$D(\phi_n - \phi_m) \to 0 \qquad \text{as } n,m \to \infty \quad ,$$

hence

$H_G(\phi_n - \phi_m)$ and, a fortiori, $H_C(\phi_n - \phi_m)$ tend to 0,

which implies that

$\omega_n(P) - \omega_m(P)$ also tends to 0.

Hence, for a fixed h, ω_n uniformly tends to a limit function u(P).

We shall now prove that, for another h, we obtain the same function u(P). We consider the function

$$\psi(x,y;\,h) = \log\frac{r}{h} + \frac{1}{2}\left(1 - \frac{r^2}{h^2}\right) \qquad \text{for } r < h$$

$$= 0 \qquad \text{for } r > h \ .$$

$$\psi = \psi_r = 0 \qquad \text{for } r = h \ .$$

Let H be the circle of radius h about P, and K the circle of radius k. For ζ we take the function

$$\psi(h) - \psi(k) \ .$$

Using Green's formula,

$$D(\phi_n,\psi) = \frac{2}{k^2}\iint_H \phi \, dxdy \ .$$

And $D(\phi_n,\zeta) = D(\phi_n,\psi(h)) - (\phi_n,\psi(k)) \to 0$ as $n \to \infty$. Hence

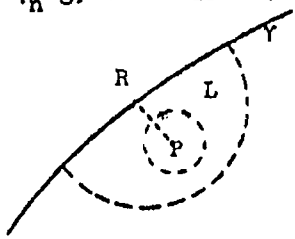
$$\frac{1}{h^2}\iint_H \phi_n \, dxdy - \frac{1}{k^2}\iint_K \phi_n \, dxdy \to 0 \qquad \text{as } n \to \infty \ ,$$

which shows that

$$u_h = u_k \ .$$

Hence u is defined independently of h in any open domain.

We shall show now that u has the prescribed boundary values. For ζ we take the function ϕ_n-g, which is 0 on the boundary.



Let R be the point nearest to P on the boundary, and PR = 2h. The circle of radius 4h about R will determine a subdomain L in G. Let K be a circle of radius h about P. We have seen before that we can take h so small that, for all n,

$$D_L(\phi_n - g) \leq \sigma(h) \quad ,$$

where $\sigma(h)$ tends to 0 with h. By the lemma proved above

$$H_L(\phi_n - g) \leq 4\pi^2(4h)^2 D_L(\phi_n - g)$$

or

$$H_L(\phi_n - g) \leq 64\pi^2 h^2 \sigma(h) \; .$$

But

$$H_K(\phi_n - g) \leq H_L(\phi_n - g) \quad .$$

Hence

$$\frac{1}{\pi h^2} H_K(\phi_n - g) = \frac{1}{\pi h^2} \iint_K (\phi_n - g)^2 \, dxdy \leq 64\pi\sigma(h) \quad .$$

By Schwarz's inequality for integrals,

$$\left[\frac{1}{\pi h^2} \iint_K (\phi_n - g) \, dxdy\right]^2 \leq \frac{1}{\pi h^2} \iint_K (\phi_n - g)^2 \, dxdy \leq 64\pi\sigma(h) \quad .$$

But the left-hand side is

$$[\omega_n(P) - \tilde{g}(P)]^2 \quad ,$$

where $\tilde{g}(P)$ is the mean value of g over K.

This is true for any n, hence, if $n \to \infty$, we obtain

$$[u(P) - \tilde{g}(P)]^2 \quad ,$$

and this expression tends to 0 with h. But when h tends to 0, $\tilde{g}(P)$ tends to g(P). Hence u(P) solves the boundary value condition.

We have to show now that u solves the minimum problem.

We consider a subdomain G' in G. We can cover G', except for an arbitrarily small area ε, with non-overlapping circles K_ν of radius ρ_ν, the largest radius being less than any preassigned value. If f is a function having an upper bound M in G',

$$\iint_{G'} f \, dxdy = \sum_\nu \pi\rho_\nu^2 f(P_\nu) + \delta \quad ,$$

where P_ν is the center of the circle K_ν, and where

$$|\delta| < \varepsilon M \quad .$$

If we take $f = u_x^2 + u_y^2$, then

$$D_{G'}(u) = \sum_\nu \pi\rho_\nu^2 \, [u_x^2(P_\nu) + u_y^2(P_\nu)] + \delta$$

But

$$u_x(P_\nu) = \lim_{n\to\infty} \frac{1}{\pi\rho_\nu^2} \iint_{K_\nu} \phi_{n_x} \, dxdy$$

and, by Schwarz's inequality,

$$u_x^2(P_\nu) \le \frac{1}{\pi\rho_\nu^2} \iint_{K_\nu} \phi_{n_x}^2 \, dxdy \quad .$$

Hence

$$D_{G'}(u) \le \lim_{n\to\infty} D(\phi_n) = d \quad ,$$

or

$$D_G(u) = \lim_{G'\to G} D_{G'}(u) \le d \quad .$$

But, u being admissible,

$$D_G(u) = d \quad .$$

Finally we shall show that u is harmonic using

$$D(u,\zeta) = 0 \quad .$$

We consider the circle of radius a and the circle of radius b about P, and for ζ we take the following function:

$$\zeta = \log \frac{r}{a} \qquad \text{for } b < r < a$$

$$\zeta = 0 \qquad \text{for } r > a$$

and $$\zeta = \log \frac{b}{a} \qquad \text{for } r < b \quad .$$

Then

$$D(u,\zeta) = \frac{1}{a}\int_a u\,ds - \frac{1}{b}\int_b u\,ds = 0 \quad ,$$

where the integrals are taken along the circles of radius a and b. The last equality shows that u is harmonic.

Thus the second proof of Dirichlet's principle is completed.

Problems

1. Prove the triangle inequalities in the function space.
2. Prove the inequality

$$H(\zeta) \leq CD(\zeta) \qquad (C > 0) \quad ,$$

where ζ is assumed to be 0 only on an arc of the boundary.

Numerical Procedures

The Ritz method. In giving an existence proof for a solution to a variational problem one merely requires the existence of minimizing sequences, with suitable convergence properties. In practical applications or for purposes of computation there still remains the problem of actually constructing a minimizing sequence, and, furthermore, one which converges with a fair degree of rapidity. The method described below was first introduced by W. Ritz, who applied it to problems concerning elastic plates.*

We consider a variational integral $I(\Phi)$ defined over a function class R of admissible functions. An enumerable sequence of functions $w_1,\ldots,w_n,\ldots$ contained in the class R is said to be complete if every function Φ in R can be approximated by a finite linear combination

---

* Ritz, Walter, "Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik," Journal für die reine und angewandte Mathematik, Vol. 135 (1908); "Theorie der Transversalschwingungen einer quadratischen Platte mit freien Rändern", Annalen der Physik, Vol. 38 (1900).

$$W_n = a_1 w_1 + a_2 w_2 + \ldots + a_n w_n$$

of functions belonging to the sequence $\{w_n\}$ with preassinged accuracy. The approximation can be understood in several senses. Given any Φ in R and any ε, we want a W_n such that

a) $$|I(\Phi) - I(W_n)| < \varepsilon$$

b) $$\iint_G (\Phi - W_n)^2 \, dxdy < \varepsilon$$

c) $$|\Phi - W_n| < \varepsilon \quad .$$

In the following we shall take the approximation in the sense a).

For example, we know from the theory of Fourier series that the sequence of functions

$$\sin n\pi x \qquad (n = 1,2,\ldots)$$

forms a complete system for all functions $\Phi(x)$ which are continuous, have a piecewise continuous derivative, and vanish at 0 and 1.

Except for the trigonometric functions, the most important and most useful complete system is given by the integral powers of x, or, in two dimensions, $x^n y^m$. The linear combinations of such functions are polynomials. Weierstrass proved the following important theorem.

If f(x) is an arbitrary continuous function in a closed interval, then it may be approximated in this interval to any desired degree of accuracy by a polynomial $P_n(x)$, provided that n is taken s ufficiently large.

This theorem is valid for higher dimensions as well.

Returning to the given variational integral $I(\Phi)$, we suppose, in order that the problem make sense, that the integral has a greatest lower bound d. From this follows immediately the existence of minimizing sequences Φ_n such that $I(\Phi_n) \to d$.

The Ritz method consists in setting up a minimizing sequence by means of a series of auxiliary minimum problems.

We consider, for a fixed n, the integral

$$I(W_n) = I(a_1 w_1 + \ldots + a_n w_n) \quad ,$$

where $w_1, \ldots, w_n$ are the first n numbers of a complete system $\{w_n\}$ for the admissible function class R. The integral then becomes a function of the n coefficients $a_1, \ldots, a_n$ varying independently. We next consider the problem: to find the set of coefficients $a_1, \ldots a_n$ which makes $I(W_n)$ a minimum. Since I has a lower bound and depends continuously on the n parameters $a_1, \ldots, a_n$, it must attain a minimum; according to the ordinary theory of maxima and minima, the system of n equations

$$\frac{\partial}{\partial a_i} I(W_n) = 0$$

serves to determine the particular values $a_i = c_i$ which give the minimum. We denote the minimizing function by $u_n = c_1 w_1 + \ldots + c_n w_n$. The essence of the Ritz method is then contained in the following theorem:

The sequence of functions $u_1, \ldots, u_n$, which are the solutions to the successive minimum problems $I(W_n)$ formed for each n, are a minimizing sequence to the original variational problem.

First, it is seen that $I(u_n)$ is a monotonically decreasing function of n, since we may regard every function W_{n-1} admissible in the (n-1)st minimum problem as an admissible function for the n-th minimum problem with the additional side condition $a_n = 0$. Therefore

$$I(u_n) \rightarrow \delta \geq d \quad .$$

Next, the existence of a minimizing sequence $\{\Phi_n\}$ to the variational problem implies that, for some sufficiently large k,

$$I(\Phi_k) < d + \frac{\epsilon}{2} \quad .$$

Since the system $w_1, \ldots, w_n, \ldots$ is complete, there exists a suitable function $W_n = a_1 w_1 + \ldots + a_n w_n$ such that

$$I(W_n) < I(\Phi_k) + \frac{\varepsilon}{2} .$$

But, by definition of u_n,

$$I(u_n) \leq I(W_n) ,$$

hence

$$I(u_n) < d + \varepsilon ,$$

which establishes the convergence of $I(u_n)$ to d.

The process of constructing the minimizing sequence $\{u_n\}$ depends on solving the system of n equations:

$$\frac{\partial}{\partial a_i} I(W_n) = 0 .$$

The process is considerably simplified if the given integral is quadratic, since in that case we have a system of linear equations in the a's.

As an example, let us consider the case where on the boundary $\Phi = g$ is a polynomial, the boundary being given by $B(x,y) = 0$. We take for functions Φ the functions

$$\Phi = g + B(x,y)(a + bx + cy + \ldots) .$$

This sequence of functions Φ is a minimizing sequence, and $I(\Phi)$ is a function $Q(a,b,c,\ldots)$ of the coefficients a,b,c,..., and the problem is reduced to finding the minimum of Q with respect to a,b,c,... .

However, we shall now come to a method by which the functions Φ are determined directly, without going through polynomials.

Method of Finite Differences

The fundamental idea of this method is to replace the differential equation by a "difference equation" (an equation involving finite differences), thereby reducing the problem to a simple system of linear algebraic equations in a finite number of unknowns.

We begin by covering the xy-plane with a quadratic mesh consisting of squares of side h. To do this we draw in the plane the two sets of parallel lines

$$x = mh \qquad (m = 0,1,2,\ldots)$$
$$y = nh \qquad (n = 0,1,2,\ldots) \quad .$$

These two families of lines intersect in points which we call the net, or lattice, points of the mesh.

Now, instead of considering functions of the continuous variables x and y, we consider functions which are defined only at the lattice points of the above mesh. That is, the functions are to be defined solely for the arguments x = mh, y = nh, where h is some fixed number. In any bounded domain only a finite number of lattice points will be present and hence each function will take on only a finite number of values. It is impossible to speak of the derivatives of such functions. Instead we define what we call the difference quotients of these discrete valued functions

Let u(x,y) be a function defined at the lattice points of the xy-plane. Then the forward difference quotient of u with respect to x at a lattice point (x,y) is defined to be

$$u_x(x,y) = \frac{u(x+h,y) - u(x,y)}{h}$$

and the backward difference quotient with respect to x

$$u_{\bar{x}}(x,y) = \frac{u(x,y) - u(x-h,y)}{h} \quad .$$

In general, these two difference quotients are not equal.

In a manner similar to the above we may define the second difference quotients of a function, i.e., the difference quotients of the first difference quotients. The forward second difference quotient is given by

$$u_{\bar{x}x}(x,y) = \frac{u_{\bar{x}}(x+h,y) - u_{\bar{x}}(x,y)}{h} = \frac{u(x+h,y) - 2u(x,y) + u(x-h,y)}{h^2} \quad .$$

Also

$$u_{x\bar{x}}(x,y) = \frac{u_x(x,y) - u_x(x-h,y)}{h} = \frac{u(x+h,y) - 2u(x,y) + u(x-h,y)}{h^2}$$

whence

$$u_{\bar{x}x} = u_{x\bar{x}} \quad .$$

We could, if we wished, consider the second difference quotients u_{xx} and $u_{\bar{x}\bar{x}}$. However, the use of $u_{\bar{x}x}$ makes for greater symmetry.

We now replace the Laplace operator Δ by the difference operator, which we denote by Δ_h, to apply to functions defined only at the lattice points. Thus

$$\Delta_h u = u_{x\bar{x}} + u_{y\bar{y}}$$

or

$$\Delta_h u = \frac{1}{h^2}\,[u(x+h,y) + u(x,y+h) + u(x-h,y) + u(x,y-h) - 4u(x,y)].$$

The significance of this operator becomes clear if we consider a net point P_0 and its four neighboring net points P_1, P_2, P_3, P_4. (Two net points are called neighbors if the distance between them is h.) Hence

$$\Delta_h u = \frac{1}{h^2}\,[u(P_1) + u(P_2) + u(P_3) + u(P_4) - 4u(P_0)] \quad .$$

That is, the value of $\Delta_h u$ at P_0 is four times the excess of the arithmetic mean of the four neighboring values over $u(P_0)$, this excess being divided by the area h^2 of the mesh.

The equation $\Delta u = 0$ corresponds to the difference equation

$$u(P_0) = \frac{u(P_1) + (P_2) + u(P_3) + u(P_4)}{4} \quad .$$

Hence, in a quadratic net, the Laplace equation states that the value of u at a lattice point P is the arithmetic mean of the values of u at the four neighbors of P. We have seen before that solutions of $\Delta u = 0$ possess this remarkable mean value property, where, in that case, the neighbors of P are the points of the circumference of a circle about P as a center.

Boundary value problem in a net. We cover the xy-plane with a quadratic net. Let G be any bounded domain in the xy-plane with a piecewise smooth boundary γ. The net domain G_h corresponding to the domain G consists of all the net points which lie in G. A net point is said to be a boundary point of G_h if not all of its four neighbors are in G; if all four neighboring points are in G, the point is said to be an interior point of G_h. The boundary γ_h of G_h is defined as the set of all boundary points of G_h.

To solve, with any specified degree of accuracy, the boundary value problem of the differential equation $\triangle u = 0$ for the domain G, we replace the differential equation by the difference equation $\triangle_h u = 0$ and the domain G by the corresponding net domain G_h. If $u = g(x,y)$ is the boundary function prescribed on γ, then the boundary values at the points of γ_h are chosen as follows. If a net point P of γ_h lies on γ, then the value of of $g(x,y)$ at P is taken as the value of u at P; if P does not lie on γ, we take as the value of u at P the value of $g(x,y)$ at a point of γ near to P.

We may now solve the boundary value problem for the net domain. Let N be the number of interior points of the net domain G_h. We may set up for each of the interior points the difference equation $\triangle_h u = 0$. In each case this is a linear equation involving five values of u. Some of these equations contain known quantities, i.e., those for points that are neighbors of boundary points. The other equations are homogeneous. Altogether we obtain a system of N linear equations in N unknowns, i.e., the values of u at the N interior points.

Existence and uniqueness of the solution. First we see that the maximum and minimum values of u certainly are attained on the boundary γ_h of G_h. For, if the maximum were attained at an interior point P, then the value of u at one of the four neighbors, say a, would be at least as large as at P, because of the mean value property. If the value at Q is larger than that at

P, we have a contradiction and the statement is proved. If the two values are equal, by continuing such a process about Q, we see that u must be constant in G_h, and again the statement is true. The minimum property is proved similarly.

It follows that the boundary value problems

$$\begin{aligned} \Delta_h u &= 0 \qquad \text{in } G_h \\ u &= 0 \qquad \text{on } \gamma_h \end{aligned}$$

has a unique solution

$$u = 0 \qquad \text{in } G_h \quad .$$

For this problem we have a system of N homogenous linear equations. For the general boundary value problem

$$\begin{aligned} \Delta_h u &= 0 \qquad \text{in } G_h \\ u &= g \qquad \text{on } \gamma_h \quad , \end{aligned}$$

we have the same N linear equations with the addition of a constant in some of them, i.e., a non-homogeneous system. From the theory of systems of linear equations the existence of the unique solution 0 for the homogeneous system implies the existence of a unique solution for the non-homogeneous system.

Practical methods. Various practical methods have been devised to solve quickly the system of N linear equations. When the domain has symmetric or special shapes, shortcuts may be found. When this is impossible general procedures can be used. These methods consist of processes of repeated manipulations which may be performed mechanically.

We begin by assuming for u(x,y) at the interior net points of G_h any values whatsoever. It is desirable, however, to make this "first approximation" (which we denote by u_1) in such a way that the assumed values lie between the maximum and minimum boundary values. We now consider two procedures.

a) Order the interior net points of G_h in some arbitrary manner, $P_1, P_2, \ldots P_N$. Then replace $u_1(P_1)$ (our assumed "first

approximation" at P_1) by the arithmetic mean of the assumed values u_1 at the four neighbors of P_1. Using this value, do the same for $u_1(P_2)$. Using the new values at P_1 and P_2, repeat the process for $u_1(P_3)$. Continue this process until the values of u_1 at all N of the interior points have been "corrected". We denote the corrected values by $u_2(P)$. They give a "second approximation" to the final solution. Again we start out with P_1 and proceed exactly as before, to determine a "third approximation" $u_3(P)$. This process is to be continued as long as notable differences occur between a value and its replaced value. When this is no longer the case we consider these values to be a good approximation to the actual solution of the boundary value problem.

b) Instead of proceeding as previously, where consecutive replacements were made, we obtain a second approximation immediately by replacing $u_1(P)$ by the mean value of the first approximation u_1 at the four neighbors of P. Thus the second approximation is obtained directly from the first one. In the same manner we obtain a third approximation u_3 directly from u_2, and so on. After a while the values will no longer change noticeably and a satisfactory solution is thus obtained.

A remarkable merit of these two methods is that, in the long run, they tend to correct any numerical errors which may have been made during the replacements, for we continue to take arithmetic means as long as marked changes occur.

Convergence of the difference equation to a differential equation. If the mesh h of G_h tends to zero, then $G_h \to G$ and $\gamma_h \to \gamma$, and, furthermore, the difference equation $\Delta_h u = 0$ tends to the differential equation $\Delta u = 0$ in G. Likewise, the boundary values along γ_h approach the boundary values along γ, and the solution of the boundary value problem of the difference equation tends to the solution of the corresponding boundary value problem for the differential equation.*

* For a proof see R. Courant, K. Friedrichs and H. Lewy, "Über die partiellen Differenzengleichungen der mathematischen Physik", Mathematische Annalen, Bd. 100 (1928). See especially pp.47-54.

Method of gradients.* Besides the method of finite differences, still another alternative to the Ritz procedure should be mentioned. This is the method of gradients which goes back to a paper published by Hadamard in 1908.** Highly suggestive as is Hadamard's attempt, difficulties of convergence were encountered. However, recent developments in the theory of conformal mapping and in Plateau's problem throw new light on Hadamard's idea, so that it seems justified now to expect from it not only purely mathematical existence proofs but also a basis for numerical treatment in suitable cases.

The principle of the method may be understood from the elementary geometric concept of a vector gradient. Let $u = f(x_1,\ldots,x_n)$ be a non-negative function of the n variables x_1, or, as we might say, of the position vector $X = (x_1,\ldots,x_n)$, and let us seek to determine a vector $X = X_o$ for which u is at least stationary. We then proceed as follows: on the surface $u = f(x)$ we move a point $(x_1,\ldots x_n,u)$ so that $x_1(t)$ and $u(t)$ become functions of a time-parameter t. Then the velocity of ascent or descent along the line $X = X(t), u = u(t)$ on the surface is

$$\frac{du}{dt} = \dot{u} = \sum_{i=1}^{n} \dot{x}_1 \, f_{x_1} = \dot{X} \text{ grad } f.$$

We now choose the line along which the motion proceeds so that the descent is as steep as possible (lines of steepest descent).

(1) $$\dot{X} = - \text{grad } f \quad ,$$

so that

$$\dot{u} = - (\text{grad } f)^2$$

---

* On this question see R. Courant, "Variational methods for the solution of problems of equilibrium and vibrations," Bulletin of the American Mathematical Society, Vol. 49, No. 1, January 1943.

** Hadamard, J., Mémoire sur le problème d'analyse relatif à l'équilibre des plaques élastiques encastrées, Mémoires presentés par divers savants éstrangers à l'Académie des Sciences de l'Institute de France (2) Vol. 33, (1908).

Hence the position vector X moves according to the system of ordinary differential equations (1) along the lines of steepest descent with respect to the function f. Under very general assumptions, it is clear that X, starting from an arbitrary initial position, will, for $t \to \infty$, approach a position for which grad f = 0, and therefore for which f is stationary and possibly a minimum. However, instead of using the continuous procedure given by the differential equation (1), we may proceed stepwise, correcting a set of approximations x to the solutions of the equations grad f = 0 by corrections proportional to the respective components of - grad f.

This elementary idea can be generalized to variational problems. If we wish to determine a function u(x,y) defined in G and having prescribed boundary values such that u is the solution of a variational problem

$$(2) \qquad I(v) = \iint_G F(x,y,v,v_x,v_y)\, dxdy = \min. \quad ,$$

then we interpret the desired function u as the limit for $t \to \infty$ of a function v(x,y,t), whose values may be chosen arbitrarily for t = 0 and for all t thereafter are determined in such a way that the expression I(v), considered as a function I(t) of t, decreases as rapidly as possible towards its minimal value. Of course the boundary values of v(x,y,t) are the same as those for u(x,y), so that v_t must vanish at the boundary. If we choose v = v(x,y,t), we find

$$(3) \qquad I(t) = - \iint_G v_t\, E(v)\, dxdy \quad ,$$

where E(v) is the Euler expression corresponding to (2).

To consider a concrete example, we suppose that

$$I(v) = \iint_G (v_x^2 + v_y^2)\, dxdy \quad ,$$

so that our minimum problem amounts to determining the equilibrium of a membrane with given boundary deflections g(s). Then

$E(v) = -2\Delta v$. Incidentally (3) displays an analogy between the Euler expression and the gradient of a function $F(x_1,\ldots,x_n)$ of n independent variables. The variation or "velocity" of $I(v)$ is expressed as an "inner product" of the velocity of the "independent function" v with the Euler expression $E(v)$, the gradient of a functional in function space.

We now assure ourselves of a steady descent or decrease of $I(t)$ by choosing v_t in accordance with the differential equation

$$v_t = -k\,E(v) \quad , \tag{4}$$

where k is a positive arbitrary function of x,y. (3) then becomes

$$\dot{I}(t) = -\iint_G k[L(v)]^2\,dxdy \quad ,$$

and again we can infer that, for $t \to \infty$, $v(x,y,t)$ will tend to the solution $u(x,y)$ of the corresponding boundary value problem $E(u) = 0$.

For the case of the membrane the differential equation (4) becomes

$$v_t = \Delta v \quad , \tag{5}$$

the equation of heat transfer. In our interpretation this equation describes a rapid approach to a stationary state along the "lines of steepest descent". While for the equations (4) and (5) the convergence of v for $t \to \infty$ can be proved, serious difficulty arises if we want to replace our continuous process by a process of stepwise corrections as would be required for numerical applications. Each step means a correction proportional to Δv, thus introducing higher and higher derivatives of the initial function v. Another great difficulty is presented by rigid boundary values.

Yet there do exist classes of problems where such difficulties can be overcome if the method is extended properly. First of all we may observe that it is not necessary to select

the steepest descent along the gradient; it suffices to secure a safe descent at a suitably fast rate. Furthermore, if we consider problems for which the boundary value problem of the differential equations presents no difficulty for the domain G, but for which a degree of freedom in the boundary values is left, then the problem reduces to one for finding those boundary values, and now all our difficulties disappear.

Application of the calculus of variations to the eigenvalue problems

Extremum properties of eigenvalues. Let $u_1,\dots,u_n$ be the components of a vector u, and

$$Q(u,u) = Q(u) = \sum_{\substack{i=1\\k=1}}^{n} a_{ik}u_i u_k$$

be a symmetric quadratic form, with

$$a_{ik} = a_{ki} \quad .$$

The so-called "mixed form" is

$$Q(u,w) = \sum_{\substack{i=1\\k=1}}^{n} a_{ik}u_i w_k \quad .$$

Let us write

$$H(u,u) = H(u) = \sum_{i=1}^{n} u_i^2 \quad ;$$

the orthogonality condition for two vectors u and w is:

$$H(u,w) = 0 \quad .$$

In the theory of quadratic forms it is proved that by an orthogonal transformation the vector u can be transformed into a vector v such that the form Q(u) is transformed into $\sum \lambda_i v_i^2$, with H(u) being transformed into $\sum v_i^2$. The λ_i are called eigenvalues or characteristic values.

These numbers can be considered as the solutions of a sequence of minimum problems.

The first problem is to make $Q(u)$ a minimum, with the subsidiary condition $H(u) = 1$. The minimum, λ_1, will be attained for a certain vector u, namely e^1, and $Q(e^1) = \lambda_1$, $H(e^1) = 1$.

The second minimum problem will be to make $Q(u)$ a minimum, with the two subsidiary conditions $H(u) = 1$, $H(u,e^1) = 0$. The minimum, λ_2, will be attained for a certain vector u, namely e^2, and $H(e^1, e^2) = 0$, $Q(e^1,e^2) = 0$, $Q(e^2) = \lambda_2$.

The k-th minimum problem will be to make $Q(u)$ a minimum, with the k subsidiary conditions $H(u) = 1$, $H(u,e^1) = 0,\ldots,$ $H(u,e^{k-1}) = 0$. The minimum, λ_k, will be attained for a certain vector u, namely e^k, and $H(e^i, e^k) = \delta_{ik}$, $Q(e^i,e^k) = \lambda_k\delta_{ik}$, where $\delta_{ik} = 1$ if $i = k$ and $= 0$ if $i \neq k$.

We can obtain similar results for quadratic functionals. We consider a self-adjoint partial differential equation of the second order

$$(1)\quad L(u) + \lambda\rho u = (pu_x)_x + (pu_y)_y - qu + \lambda\rho u = 0\ ,\ (p>0,\ \rho>0)$$

where u is a function of two independent variables, x and y, defined over a domain G, of which the boundary, Γ, is a continuous curve with a piecewise continuous tangent. The boundary condition is $u = 0$ or, more generally, $\partial u/\partial n + \sigma u = 0$, where σ is a piecewise continuous function of the arc length on Γ and $\partial/\partial n$ denotes differentiation along the normal to Γ. For the variational problems equivalent to these eigenvalue problems the following quadratic functionals are to be considered:

$$E(\Phi) = D(\Phi) + \int_\Gamma p\sigma\Phi^2\,ds\ ,$$

$$D(\Phi) = \iint_G p(\Phi_x^2 + \Phi_y^2)\,dxdy + \iint_G q\Phi^2\,dxdy\ ,$$

$$H(\Phi) = \iint_G \rho\Phi^2\,dxdy\ ,$$

and the corresponding "mixed" forms

$$E(\Phi,\psi) = D(\Phi,\psi) + \int_{\Gamma} p\,\Phi\psi\,ds$$

$$D(\Phi,\psi) = \iint_{G} p(\Phi_x\psi_x + \Phi_y\psi_y)\,dxdy + \iint_{G} q\Phi\psi\,dxdy$$

$$H(\Phi,\psi) = \iint_{G} \rho\Phi\psi\,dxdy \quad .$$

For these expressions we have the relations:

$$E(\Phi + \psi) = E(\Phi) + 2E(\Phi,\psi) + E(\psi) \quad ,$$

$$H(\Phi + \psi) = H(\Phi) + 2H(\Phi,\psi) + H(\psi) \quad .$$

A function Φ is admissible if continuous in $G + \Gamma$ and possessing piecewise continuous first derivatives.

We can obtain the eigenvalues λ_γ and the corresponding eigenfunctions u of the differential equation (1) through the following minimum properties:

> Among all admissible functions the function for which the expression $E(\Phi)$ is a minimum with the subsidiary condition $H(\Phi) = 1$, is an eigenfunction u^1 of the differential equation (1) with the natural boundary condition $\partial\Phi/\partial n + \sigma\Phi = 0$. The minimum value of $E(\Phi)$ is the corresponding eigenvalue. If, to the condition $H(\Phi) = 1$, we add the new subsidiary condition $H(\Phi,u^1) = 0$, then the solution of this new minimum problem is an eigenfunction u^2 of (1) with the same boundary condition, and the minimum value $E(u^2) = \lambda_2$ is the corresponding eigenvalue. Generally the variational problem $E(\Phi)$ = minimum, with the subsidiary conditions $H(\Phi) = 1$ and $H(\Phi,u^i) = 0$ $(i = 1,2,\ldots k-1)$ has a solution u^k which is an eigenfunction of (1) with the boundary condition $\partial\Phi/\partial n + \sigma\Phi = 0$, and the minimum value $E(u^k)$ is the corresponding eigenvalue λ_n.

Instead of making $E(\Phi)$ a minimum with the condition $H(\Phi) = 1$, we can abandon this condition and make the quotient $E(\Phi)/H(\Phi)$ a minimum; the solution is then given with an arbitrary factor of proportionality.

We shall assume here that the minimum problems have a solution, and show that their solutions are the eigenfunctions of partial differential equation (1).

We consider the solution u^1 of the first variational problem, $H(u^1) = 1$. Let ζ be an admissible function and ε an arbitrary constant, then $u^1 + \varepsilon\zeta$ is also an admissible function, and

$$E(u^1 + \varepsilon\zeta) \geq \lambda_1 H(u^1 + \varepsilon\zeta)$$

or

$$2\varepsilon[E(u^1,\zeta) - \lambda_1 H(u^1,\zeta)] + \varepsilon^2[E(\zeta) - \lambda_1 H(\zeta)] \geq 0$$

for every ε, which implies

$$E(u^1,\zeta) - \lambda_1 H(u^1,\zeta) = 0 \quad .$$

Because of Green's formula

$$E(u^1,\zeta) = -\iint_G \zeta L(u^1)\, dxdy + \int_\Gamma p\, \zeta u^1\, ds \quad ,$$

and, as the function ζ is arbitrary, we obtain the equation (1) for $u = u^1$ and $\lambda = \lambda_1$.

In the second minimum problem, let η be an admissible function; then $\zeta = \eta + tu^1$ is an admissible function, and we determine t so that $H(\zeta,u^1) = 0$; we obtain $t = -H(u^1,\eta)$. Substituting in

$$E(u^2,\zeta) - \lambda_2 H(u^2,\zeta) = 0 \quad ,$$

we obtain

$$E(u^2,\eta) - \lambda_2 H(u^2,\eta) + t[E(u^1,u^2) - \lambda_2 H(u^1,u^2)] = 0 \ .$$

The last term is 0, and we get

$$E(u^2,\eta) - \lambda_2 H(u^2,\eta) = 0 \quad ,$$

which is the same equation as in the first case. Hence u^2 is an eigenfunction of (1) and λ_2 the corresponding eigenvalue.

In the same way, it is seen in the general case that the equation

$$E(u^i,\eta) - \lambda_i H(u^i,\eta) = 0$$

holds, η being an arbitrary admissible function. For the normal solutions of the successive minimum problems we have the relations

$$\begin{aligned} E(u^i) &= \lambda_i \quad , \quad & E(u^i,u^k) &= 0 \\ H(u^i) &= 1 \quad , \quad & H(u^i,u^k) &= 0 \end{aligned} \qquad (i \neq k) .$$

The eigenvalues satisfy the inequality

$$\lambda_{n-1} \leq \lambda_n \quad ,$$

for in the n-th minimum problem the domain of the functions Φ admitted to competition is not larger than in the (n-1)th problem. Hence the minimum λ_n is not smaller than the minimum λ_{n-1}.

We shall simply mention here that other eigenvalue problems can be treated with the help of the calculus of variations. The integral may be simple or multiple, and the differential equations may be of the second order or higher.

The maximum-minimum property of the eigenvalues. We can replace the recurrence definition of the n-th eigenvalue and the corresponding eigenfunction by a definition in which the n-th eigenvalue and the n-th eigenfunction are determined without knowing the preceding ones.

We consider the same variational problems as before, but we replace the conditions $H(\Phi,u^i) = 0$ $(i = 1,2,\ldots n-1)$ by the n-1 new conditions $H(\Phi,v^i) = 0$ $(i = 1,2,\ldots,n-1)$, where $v^1,v^2,\ldots v^{n-1}$ are any functions piecewise continuous in G. It is not decided whether the new problem has a solution. However, the expressions $D(\Phi)$ and $E(\Phi)$ have a lower bound, which depends on the functions $v^1,v^2,\ldots v^{n-1}$ and which we shall call $d(v^1,v^2,\ldots v^{n-1})$.

Given n-1 functions $v^1, v^2, \ldots v^{n-1}$, piecewise continuous in G, and $d(v^1, v^2, \ldots v^{n-1})$ being the minimum or the lower bound of all the values that the expression $E(\Phi)/H(\Phi)$ can take, when Φ is any admissible function which satisfies the conditions $H(\Phi, v^i) = 0$ $(i = 1,2,\ldots,n-1)$, then λ_n is equal to the greatest value that $d(v^1, v^2, \ldots, v^{n-1})$ can take when all sets of admissible functions are taken for $v^1, v^2, \ldots v^{n-1}$. This maximum-minimum is attained for $u = u^n$ and $v^1 = u^1$, $v^2 = u^2, \ldots, v^{n-1} = u^{n-1}$.

To prove this statement, we remark first that for $v^i = u^i$ $(1 \leq i \leq n-1)$, by definition, $d(v^1, v^2, \ldots, v^{n-1}) = \lambda_n$. Then, we shall show that for arbitrarily chosen $v^1, \ldots, v^{n-1}$, $d(v^1, v^2, \ldots v^{n-1}) \leq \lambda_n$. We simply have to show that there is one function Φ, satisfying the conditions $H(\Phi, v^i) = 0$ $(i=1,2,\ldots n-1)$ for which $E(\Phi) \leq \lambda_n$ and $H(\Phi) = 1$. We consider a linear combination of the first n eigenvalues, $\Phi = \sum_{i=1}^{n} c_i u^i$. The n-1 relations $H(\Phi, v^i) = 0$ give n-1 linear homogeneous equations for determining the n constants $c_1, c_2, \ldots, c_n$; they can always be solved. The condition $H(\Phi) = \sum_{i=1}^{n} c_i^2 = 1$ gives the factor of proportionality. Now

$$E(\Phi) = \sum_{i,k=1}^{n} c_i c_k E(u_i, u_k) \quad ,$$

but

$$E(u_i, u_k) = 0 \qquad (i \neq k)$$

and

$$E(u_i) = \lambda_i \quad .$$

Hence

$$E(\Phi) = \sum_{i=1}^{n} c_i^2 \lambda_i \quad ;$$

because of $\sum_{i=1}^{n} c_i^2 = 1$ and $\lambda_n \geq \lambda_i$ $(i = 1, \ldots, n)$,

$$E(\Phi) \leq \lambda_n .$$

Hence the minimum $d(v^1,..,v^{n-1})$ is not greater than λ_n, and λ_n is the greatest value that the minimum can take.

The maximum-minimum property of eigenvalues is extremely useful in many physical problems. It leads immediately to two principles which we shall simply state here:

1. By strengthening the conditions in a minimum problem, the value of the minimum is not decreased, and conversely by weakening the conditions the minimum decreases or does not increase.

2. Consider two minimum problems for the same domain of admissible functions $\underline{\Phi}$. If, for each function $\underline{\Phi}$, the expression to be minimized is not smaller in the first problem than in the second, then the minimum in the first problem is not smaller than in the second.

Physically, the first principle can be stated:

When a vibratory system is forced to vibrate under certain imposed conditions, then the fundamental frequency and each harmonic can only increase. Conversely, when the conditions under which the system vibrates are weakened, the fundamental frequency and each harmonic can only decrease.

For further discussion of the extremum properties of eigenvalues we refer to Courant, R., and Hilbert, D., Methoden der Mathematischen Physik, vol. 1, chapter VI.

Supplementary Notes and Exercises

1. References.

Two of the earlier works treating the Calculus of Variations are as follows:

Moigno-Lindelöf, Calcul des Variations (1861)
Jordan, Cours d'Analyse, Vol. III (1887)

Some later works on the Calculus of Variations, which reflect the very important influence of the lectures of Weierstrass, are as follows:

Kneser, Lehrbuch der Variationsrechnung (1900)
Bolza, Lectures on the Calculus of Variations (1904)
Bolza, Vorlesungen über Variationsrechnung (1909)
Hadamard, Lecons sur le calcul des variations (1910)
Tonelli, Fondamenti di calcolo delle variazioni (1921)
Goursat, Cours d'Analyse mathématique, Vol. III (1923)
Bliss, Calculus of Variations (1925)
Courant-Hilbert, Methoden der mathematischen Physik, Vol. I (1931)

2. Notes on the Brachistochrone problem.

Exercise: Imagine a vertical plane with two given points A and B. Find the time required for a mass point to slide without friction under the influence of gravity from A to B along each of the following curves:

a) a straight line segment

b) a circular arc starting vertically

c) a vertical straight line segment followed by another straight line segment, as follows:

Johann Bernoulli solved the Brachistochrone problem by dividing the plane into horizontal strips of width ω and assuming the speed of the particle to be constant in each strip. He then applied Snell's Law of Refraction (as explained in the notes on pp. 2-3), passed to the limit as $\omega \to 0$, and concluded that the required curve is characterized by the property

$$\frac{\sin \alpha}{\sqrt{u}} = \text{constant}$$

where α is the angle between the tangent to the curve and the vertical. It can easily be seen that this condition is equivalent to

$$x = \int_0^u \frac{d\xi}{\sqrt{\frac{k^2}{\xi} - 1}}$$

where k is a constant.

Exercise: Evaluate the above integral and identify the curve.

A critical evaluation of Johann Bernoulli's procedure from the point of view of modern rigor reveals two serious gaps in the reasoning. Conclusions should follow from assumptions by logical reasoning. However, two very important points in

Bernoulli's argument were not logically justified. One of these points will become apparent upon examining Bernoulli's reasoning in the following way: The solution S to a problem P was being sought. In order to solve P, problems P_ω (corresponding to dividing the plane into strips of width ω) were considered, and their solutions S_ω were found. The problems P_ω were chosen so that they resemble P with the P_ω resembling P more and more closely as ω is chosen smaller and smaller. It was then assumed without logical justification that the S_ω resemble S more and more closely as ω is chosen smaller and smaller. In other words, it was assumed without foundation that

$$S_\omega \rightarrow S$$

as

$$P_\omega \rightarrow P \quad .$$

Although this procedure gave the correct result when applied to the Brachistochrone problem, the same procedure may very well give incorrect results for other problems.

The second gap in Bernoulli's reasoning was his tacit assumption of the existence of a solution to the problem. Since the problem sounded reasonable, he did not doubt for a moment the existence of a solution. However, as is well known today, even problems which sound very reasonable may not have solutions. Bernoulli's procedure merely indicated that if there is a solution to the minimum problem, then it must be the solution which he found, the cycloid. And, as pointed out above, even this was not shown in a logical way. Thus, Bernoulli's procedure, although highly significant for its historical interest, actually proved nothing at all.

3. The road of quickest ascent to the top of a mountain.

Consider the problem of how to build a road up to the top of a mountain in such a way as to minimize the length of time it takes to get to the top for a car whose speed is a given function of the angle of inclination. We assume that the moun-

tain is symmetric about the vertical axis through its apex, which we take to be the z-axis. If r represents the distance from this axis, the equation of the surface of the mountain may be written as $z = g(r)$. The surface may be differentiable at the apex, in which case $g_r(0) = 0$ (as in the diagram), but this need not be assumed; it may instead behave conically there.

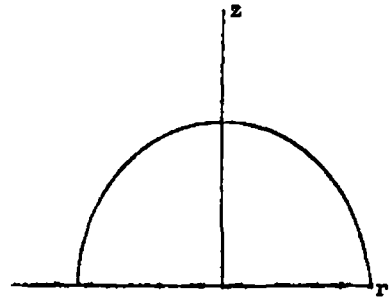


Diagram: The curve $z = g(r)$.

A road on the mountain is represented by a curve, starting say at a distance R from the z-axis and ending at the apex. Let t be time, α the angle of inclination of the curve (the road) to the horizontal, ds an element of arclength along the curve, and the dr and dz the corresponding elements of length along the radial and vertical axes respectively. Clearly

$$g_r(r)dr = dz = \sin \alpha \cdot ds$$

is the projection of ds on the z-axis. If the prescribed speed v is given by $f(\alpha)$, then

$$\frac{ds}{dt} = f(\alpha) \quad , \qquad dt = \frac{ds}{f(\alpha)} = \frac{g_r(r)}{f(\alpha) \sin \alpha} dr \quad , \quad \text{and}$$

$$T = \int_R^0 \frac{g_r(r)}{f(\alpha) \sin \alpha} dr = \int_0^R \frac{-g_r(r)}{f(\alpha) \sin \alpha} dr$$

is the length of time required for the ascent.

Taking $g_r(r) < 0$ (except perhaps at $r = 0$) to insure that the ground is always rising as the apex is approached, it is clear that T will be minimized if α is chosen always so as to maximize $f(\alpha) \sin \alpha$. Assuming, as is reasonable for the interpretation of the problem, that $f(0) = v_0$, $f(\frac{\pi}{2}) = 0$, and $f(\alpha)$ is monotonic it follows that $f(\alpha) \sin \alpha$ will have a positive

maximum for some $\alpha = \bar{\alpha}$, which we assume unique, satisfying $0 < \bar{\alpha} < \frac{\pi}{2}$. However it is necessary to notice here that at the distance r from the z-axis α is subject to the condition $\alpha \leq \arctan(-g_r(r))$, since the angle of inclination of the road cannot exceed the maximal angle of inclination of the ground. Assuming for simplicity that $f(\alpha) \sin \alpha$ is monotonic for $0 \leq \alpha \leq \bar{\alpha}$, the maximum value of $f(\alpha) \sin \alpha$ occurs for

$$\alpha = \min [\bar{\alpha}, -\arctan g_r(r)] \quad .$$

It remains to be shown that there exists a curve on the surface with this prescribed angle of inclination. At each point of the horizontal plane z = 0 consider the direction or directions which are the vertical projections of directions in the surface with angle of inclination as prescribed. There will be one or two of them according to whether $-\arctan g_r(r)$ is equal to or greater than the prescribed angle. In the first case the direction will be toward the z-axis, i.e. radial; in the second case the two directions will form an angle bisected by the radial direction. If in the second case we always choose the counter-clockwise direction, say, then we have a unique direction at each point (excluding r = 0), and this direction field will be continuous if $g_r(r)$ is. An integral curve of this direction field, starting from the initial point at distance R, will reach the center with finite arclength if $g_r(r)$ is bounded. Then the curve on the surface, of which this integral curve is the vertical projection, will have the prescribed angle of

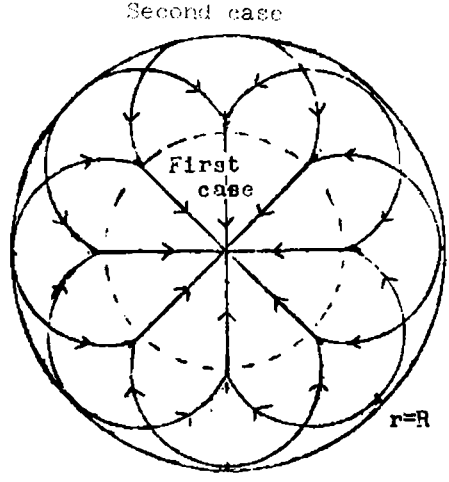


Diagram: Directions in the plane z = 0.

inclination at each point, (will have finite arclength,) and will obviously require the shortest time for ascent, since for each value of r it minimizes the integrand of T.

The solution may, in short, be described as a curve which has the fixed angle of inclination $\bar{\alpha}$ (the value of α which maximizes $f(\alpha) \sin \alpha$) wherever possible, and elsewhere is directed toward the z-axis. If the surface is differentiable at the apex, this latter phenomenon (case 1) must occur in the neighborhood of it. If it does not occur, i.e. if the curve has always the angle of inclination $\bar{\alpha}$, then

$$T = \frac{f(0) - g(R)}{f(\bar{\alpha}) \sin \bar{\alpha}}$$

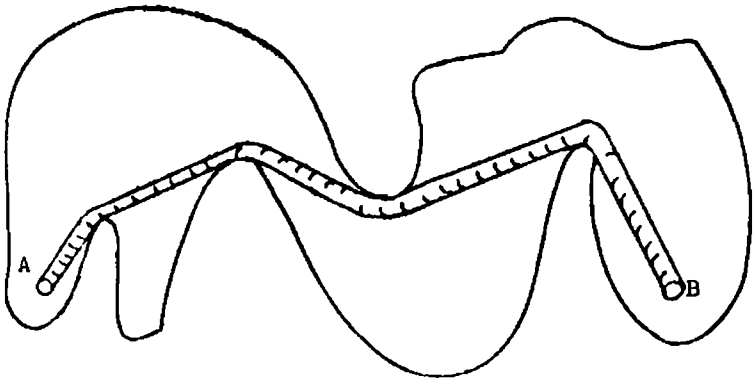
In any case it is obvious that the solution is unique except for the choice (which may occur more than once) as to which way to spiral around the mountain, clockwise or counter-clockwise, since for any other curve the integrand in T will be larger for some value of r (and therefore in a neighborhood) and no smaller anywhere else.

4. The shortest connection between two points in a closed simply-connected region.

An interesting problem is to find the shortest possible path along which to walk from one point to another on a winding street bounded by well-behaved curves (where it is permissible to walk along the boundary curves). This problem may be formulated in the following way: Describe a method for finding the shortest connection between a point A and a point B in a closed simply-connected region bounded by a smooth curve (i.e., a continuous curve with a continuously turning tangent).

The solution to the problem can be visualized by imagining a string attached to the point A, pulled taut, and attached to the point B. It is intuitively clear that the string would follow a path such as indicated in Figure 1, but it would be desirable to describe a mathematical method for finding this path.

Figure 1. An indication of the shortest path from A to B furnished by a string.



It will be assumed here that there is a solution to the problem; i.e., that a shortest connection really exists. Let this solution be the curve C.

It is immediately apparent that any (connected) portion of C consisting of interior points of the region must be a straight line segment. For, if an interior arc of C were not a straight line segment, a point P on this arc could be selected for which an ε-neighborhood lies in the region. Then a point P^* on the arc with its distance from P less than ε could also be selected, and the chord PP^* could replace the portion of C between P and P^*. This replacement would shorten the connection between A and B, in contradiction to the assumption that C provides the shortest connection. Thus, it follows that C consists only of straight line segments in the interior of the region and arcs of the boundary curve.

Moreover, it can be seen that the interior straight line segments can meet the boundary only along tangent lines to the boundary. In order to show this, let a portion of C consisting of an interior straight line segment be denoted by C_1, and let an adjacent portion of C consisting of an arc of the boundary

be denoted by C_b. Then, let α be the "interior angle" * between C_i and C_b at the point where they

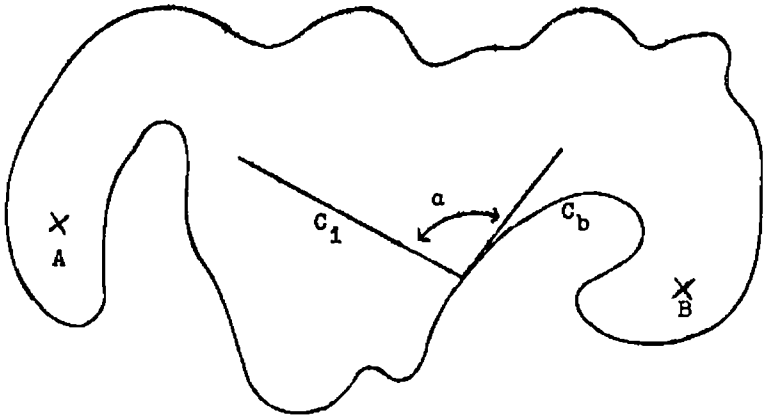


Figure 2. The interior angle between C_i and C_b.

meet. It follows immediately that

(1) $$\alpha \geq \pi \quad ,$$

for otherwise C could be shortened by drawing a straight line segment between a point on C_i and a point on C_b, both points being chosen sufficiently close to the vertex so that the straight line segment should lie entirely within the region. From the assumption that the boundary curve is smooth and, therefore, cannot have corners, it follows that

(2) $$\alpha \leq \pi \quad .$$

---

* Clearly there are two angles α and β between C_i and C_b where $\alpha + \beta = 2\pi$. One of the angles has the property that there are points within it exterior to the region arbitrarily close to the vertex. The other angle has the property that all points within it sufficiently close to the vertex are interior points of the region. Let the latter angle be called the interior angle between C_i and C_b. (See Figure 2.)

Then, from (1) and (2), it follows that $\alpha = \pi$, which means that C_1 meets the boundary along a tangent line. Thus, it has been established that C can consist only of arcs of the boundary curve and interior straight line segments which at both endpoints terminate as tangents to the boundary (or terminate in A or B).

The solution to the problem is now close at hand. For, all that has to be done is consider the totality of straight line segments with the property specified above for their endpoints. Since the boundary curve was assumed to be smooth, it follows that the totality of all such straight line segments constitutes a finite number. As a consequence, it is possible to connect A with B in only a finite number of ways by paths consisting of such straight line segments and arcs of the boundary curve. Thus, the minimum problem has been reduced to selecting the shortest path from a finite number. This constitutes a solution to the problem. Each possible path can be given a name and listed on a piece of paper. The length of each path can be written next to the name of the path. Then, it is a trivial matter to go down the list and select the shortest path. In fact, any maximum or minimum problem is completely solved when it is reduced to a selection from a finite number of objects.

In the event that the boundary curve is continuous but only piecewise smooth, the problem can be solved in the same way with the exception that supporting lines at the corners have to be considered in addition to the tangents. This situation is illustrated at the point S in Figure 3.

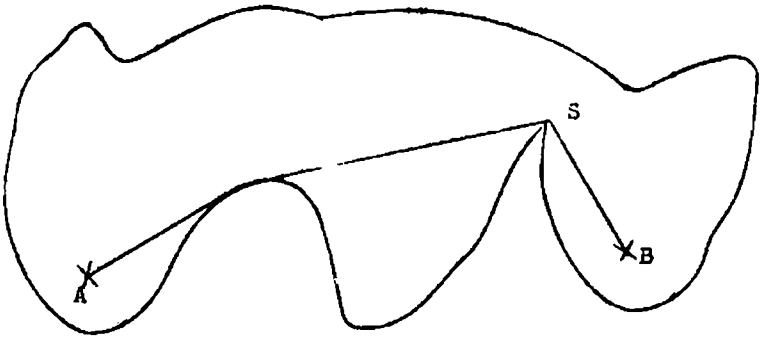


Figure 3. A region bounded by a piecewise smooth curve with the shortest path containing supporting lines to the boundary at the point S.

5. The shortest connection in a plane between two points.

It is well known that the straight line segment is the shortest plane curve connecting two points. This intuitively evident property of the straight line can be proved in a very simple way.

Specifically, it will now be proved that the straight line between two points is shorter than the connection provided by any other curve capable of being represented parametrically by

$$x = f(t)$$
$$y = g(t)$$

for $\alpha \leq t \leq \beta$, where f and g are continuously differentiable functions of t.

The definition of the length L of such a curve will be taken as

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

(where $\dot{x}$ and $\dot{y}$ represent $\frac{dx}{dt}$ and $\frac{dy}{dt}$ respectively). It can easily be verified that length defined in this way is invariant under translations and rotations; and, therefore, the proof can be simplified by placing the two points in question at (0,0) and

(a,0), where a is the length of the straight line segment joining the two points.

Then, since

$$\sqrt{\dot{x}^2 + \dot{y}^2} \geq \dot{x}$$

it follows that

$$\int_\alpha^\beta \sqrt{\dot{x}^2 + \dot{y}^2}\, dt \geq \int_\alpha^\beta \dot{x}\, dt = \int_0^a dx = a$$

or

$$L \geq a \ .$$

Moreover, since $\dot{y}$ vanishes identically only for the straight line, it follows that

$$L > a$$

for all curves other than the straight line which are capable of being represented parametrically as specified above. Thus, the length of any such curve is greater than the length of the straight line segment.

An alternate proof can be given with the advantage that it does not assume the proof of the invariance of length under translations and rotations. For this proof, let the two points in question be (x_0,y_0) and (x_1,y_1). Then, the following inequality is to be established.

$$\sqrt{(x_1-x_0)^2 + (y_1-y_0)^2} \leq \int_\alpha^\beta \sqrt{\dot{x}^2 + \dot{y}^2}\, dt \ .$$

This follows directly from the following application of Schwarz's inequality:

$$(3) \quad (x_1-x_0)\dot{x} + (y_1-y_0)\dot{y} \leq \sqrt{(x_1-x_0)^2 + (y_1-y_0)^2}\,\sqrt{\dot{x}^2 + \dot{y}^2} \ .$$

This inequality, upon integration, implies:

$$(4)\quad \int_\alpha^\beta [(x_1-x_0)\dot{x} + (y_1-y_0)\dot{y}]dt \le \sqrt{(x_1-x_0)^2 + (y_1-y_0)^2}\int_\alpha^\beta \sqrt{\dot{x}^2+\dot{y}^2}\,dt$$

However, the integral on the left can be simplified as follows:

$$\int_\alpha^\beta [(x_1-x_0)\dot{x} + (y_1-y_0)\dot{y}]dt = \int_{x_0}^{x_1} (x_1-x_0)dx + \int_{y_0}^{y_1} (y_1-y_0)dy$$

$$= (x_1-x_0)^2 + (y_1-y_0)^2 \quad .$$

This simplification together with (4) implies:

$$(x_1-x_0)^2 + (y_1-y_0)^2 \le \sqrt{(x_1-x_0)^2 + (y_1-y_0)^2}\int_\alpha^\beta \sqrt{\dot{x}^2+\dot{y}^2}\,dt$$

or, dividing by $\sqrt{(x_1-x_0)^2 + (y_1-y_0)^2}$:

$$\sqrt{(x_1-x_0)^2 + (y_1-y_0)^2} \le \int_\alpha^\beta \sqrt{\dot{x}^2+\dot{y}^2}\,dt \quad .$$

The uniqueness of the solution can be shown from this proof by noting when the equality holds in Schwarz's inequality. The equality sign holds in (3) only if

$$\dot{x} : \dot{y} = (x_1 - x_0) : (y_1 - y_0) \quad ,$$

that is, only if the ratio of $\dot{x}$ to $\dot{y}$ is constant. Since this property is enjoyed only by the straight line, it follows that the straight line is the unique minimizing curve.

6. Problems for which no solutions exist.

In Section 2, when critically evaluating Johann Bernoulli's method for solving the brachistochrone problem, it was mentioned that entirely reasonable-sounding maximum or minimum problems may possess no solutions. By way of illustration, several problems of this nature will now be discussed.

An example of a very simple problem without a solution is the following: It is required to find the shortest plane curve

connecting two points A and B, where the further restriction is imposed that all curves admitted to the competition must be perpendicular at A to straight line AB. Let the length of the straight line segment joining A and B be denoted by d. The straight line segment itself, of course, does not satisfy the perpendicularity requirement. Thus the length of every curve which does satisfy this requirement must be greater than d. Consequently, if there were such a curve of minimum length, its length would be some number, say d_{min}, greater than d. However, it is clear that it is possible to find curves satisfying the perpendicularity requirement with lengths arbitrarily close to d. In particular, such a curve can be found with its length equal to some number d_o where $d_{min} > d_o > d$. This obviously constitutes a contradiction since d_{min} was supposed to represent the length of the shortest curve satisfying the perpendicularity requirement. Thus, the minimum problem cannot have a solution.

In other words, the greatest lower bound of the lengths of all curves admitted to the competition is d. Yet, the only curve of length d joining A and B is the straight line, which is not admissible as a solution to the problem. As a consequence, there is no admissible curve of minimum length.

An example will now be given of a two-dimensional problem having a solution, which, when generalized to three dimensions in a very natural and reasonable way, no longer has a solution.

The two-dimensional problem is to find the shortest connection in a plane between two points A and B where the connecting curve is required to pass through a point C which is not on straight line AB. The solution obviously consists of straight line segments AC and CB.

A corresponding problem in three dimensions would be to find the surface of least area spanning a circle (or some other non-self-intersecting closed plane curve) where it is required that the surface pass through a point P not in the plane of the circle. Clearly, the area of any such surface is greater than the area of the disk spanning the circle. On the other hand,

the area of the circular disk will now be shown to be the greatest lower bound of the areas of all surfaces passing through P and spanning the circle.

In order to see that the area of the disk is the greatest lower bound, consider surfaces obtained in the following way. Make a hole in the disk by removing a small circular area from its interior. Then consider the cone determined by the point P and the boundary of this small circular area. The surface consisting of the disk with the hole and the lateral surface of the cone is then admissible for the problem. (See Figure 4.) Moreover, by choosing the hole small enough, the area of this surface can evidently be made as close as desired to the area of the disk.*

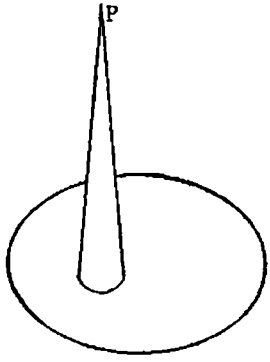


Figure 4. The admissible surface with area close to that of the disk.

Thus, the greatest lower bound of all areas of admissible surfaces is not the area of any admissible surface, but is rather the area of the circular disk, which, of course, does not

* That the surface areas can be made arbitrarily close to the area of the disk can be seen very easily by considering a special case with no essential loss of generality. Let the disk be bounded by a unit circle, let the point P be such that the foot of the perpendicular from it to the disk is an interior point of the disk, and let the distance from P to the disk be unity. Then, make the circular hole in the disk with the center at the foot of the perpendicular from P and with radius equal to ϵ. The area of the disk with the hole is then $\pi(1-\epsilon^2)$, and the lateral area of the cone is $\pi\epsilon\sqrt{1+\epsilon^2}$. The area of the admissible surface is then $\pi(1-\epsilon^2+\epsilon\sqrt{1+\epsilon^2})$. This area can certainly be made arbitrarily close to π by choosing ϵ sufficiently small.

pass through P. Yet, the greatest lower bound would have to be the area of the admissible surface of least area if there were such a surface. Therefore, there can be no admissible surface of minimum area.

Another minimum problem without a solution is the following: Find the function $u = u(x)$ which minimizes the integral

$$\int_{-1}^{1} x^2 u'^2 \, dx$$

where u is required to be continuous with a piecewise continuous first derivative and is required to satisfy the conditions $u(1) = 1$ and $u(-1) = -1$. If there were a minimizing function, the value of the integral for this function would have to be positive since the integral is never negative and is different from zero unless $u' \not\equiv 0$ (except possible for isolated values of x). But $u' \not\equiv 0$ together with the requirement that u be continuous implies that u is constant. Clearly this is incompatible with the conditions $u(1) = 1$ and $u(-1) = -1$, and the value of the integral for a minimizing function would have to be positive as stated above. On the other hand, the greatest lower bound of values of the integral for admissible functions can be seen to be zero by considering the following one-parameter family of admissible functions (See Figure 5):

$$u^{\epsilon}(x) = \begin{cases} -1 & \text{for} \quad -1 \leq x \leq -\epsilon \\ \frac{x}{\epsilon} & \text{for} \quad -\epsilon < x \leq \epsilon \\ 1 & \text{for} \quad \epsilon < x \leq 1 \end{cases}$$

The value of the integral is then:

$$\int_{-\epsilon}^{\epsilon} x^2 \frac{1}{\epsilon^2} = \frac{2\epsilon}{3} \quad .$$

Thus, the integral can be made arbitrarily close to zero for an admissible function by choosing ϵ sufficiently small, and the minimum value of the integral could not be greater than zero; i.e., a minimizing function does not exist.

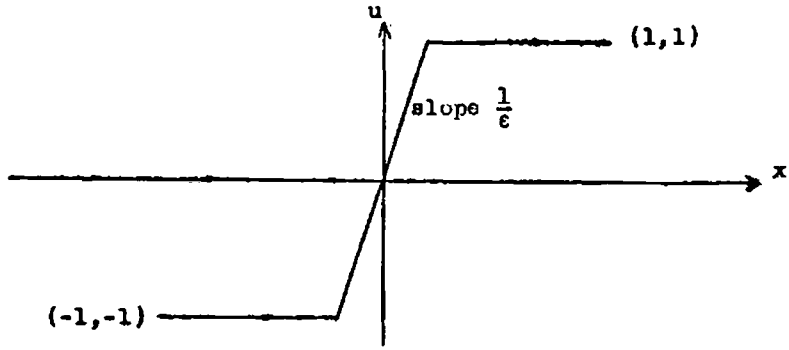


Figure 5. The functions $u^{\epsilon}(x)$ for which $\int_{-1}^{1} x^2 u'^2 dx \sim 0$.

A maximum and minimum problem which leads to a very interesting situation is the following: Find a maximizing function and also a minimizing function for the integral

$$\int_0^1 \frac{1}{1+u'^2}\, dx$$

where u is required to be continuous with a piecewise continuous first derivative and is required to satisfy the conditions $u(0) = 0$ and $u(1) = 1$. Since the integrand is always greater than zero and never exceeds unity, and the interval of integration is of unit length, it follows that the value of the integral for admissible functions cannot even equal unity since this value could only occur for $u' \equiv 0$. It will now be shown that the greatest lower bound is zero and the least upper bound is unity for the values of the integral for admissible functions. This, of course, would imply that neither a maximizing function nor a minimizing function exists.

A one-parameter family of admissible functions for which the integral assumes values arbitrarily close to zero is the following (See Figure 6):

$$u_1^\varepsilon(x) = \begin{cases} \frac{x}{\varepsilon} & \text{for} \quad 0 \le x \le \frac{1+\varepsilon}{2} \\ -\frac{x}{\varepsilon} + \frac{1+\varepsilon}{\varepsilon} & \text{for} \quad \frac{1+\varepsilon}{2} < x \le 1 \end{cases} .$$

For these functions the integral becomes

$$\int_0^{\frac{1+\varepsilon}{2}} \frac{1}{1 + (\frac{1}{\varepsilon})^2}\, dx + \int_{\frac{1+\varepsilon}{2}}^1 \frac{1}{1 + (-\frac{1}{\varepsilon})^2}\, dx = \frac{\varepsilon^2}{\varepsilon^2 + 1} .$$

Similarly, a one-parameter family of admissible functions

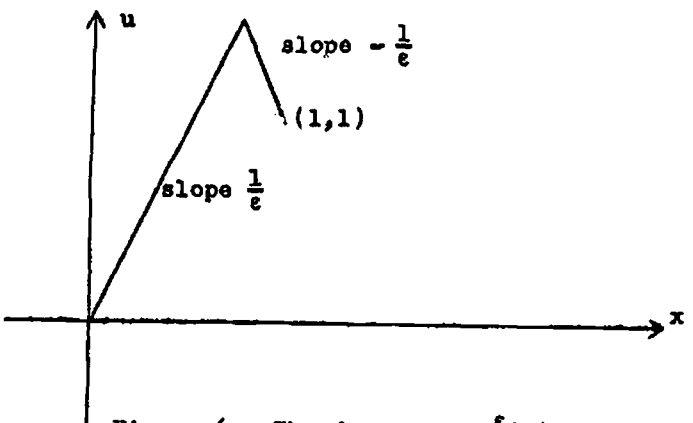


Figure 6. The functions $u_1^\varepsilon(x)$ for which $\int_0^1 \frac{1}{1 + u'^2}\, dx \sim 0$.

for which the integral assumes values arbitrarily close to unity is the following (See Figure 7):

$$u_2^\varepsilon = \begin{cases} \frac{x}{\varepsilon} & \text{for} \quad 0 \le x \le \varepsilon \\ 1 & \text{for} \quad \varepsilon < x \le 1 \end{cases} .$$

For these functions, the integral becomes

$$\int_0^\varepsilon \frac{1}{1 + \frac{1}{\varepsilon^2}}\, dx + \int_\varepsilon^1 dx = \frac{\varepsilon^3}{\varepsilon^2 + 1} + 1 - \varepsilon .$$

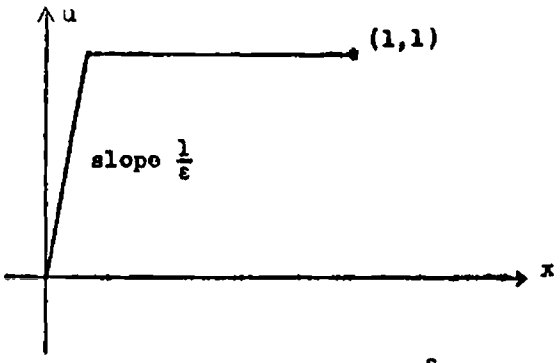


Figure 7. The functions $u_2^{\varepsilon}(x)$ for which $\int_0^1 \frac{1}{1 + u'^2}\, dx \sim 1$.

However, the interesting feature of this maximum and minimum problem is not the mere non-existence of a solution. A much more striking property can be demonstrated for the integral in question. It can easily be seen that any admissible function can be approximated to any prescribed degree of accuracy by admissible functions for which the integral is arbitrarily close to zero and also by admissible functions for which the integral is arbitrarily close to unity. The approximating functions for which the integral is close to zero consist of polygonal paths made up of straight line segments of slope $\frac{1}{\varepsilon}$ and $-\frac{1}{\varepsilon}$. (See Figure 8.)

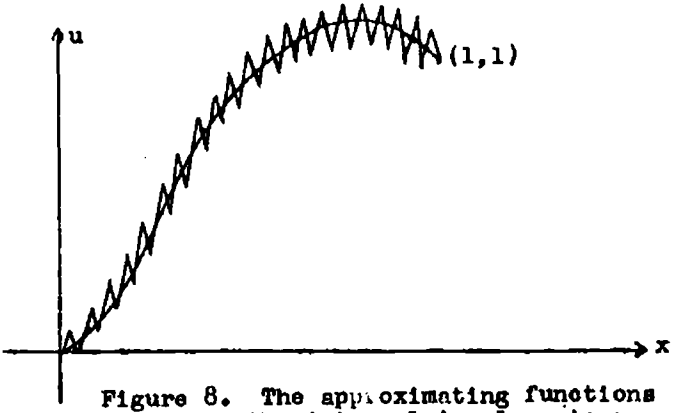


Figure 8. The approximating functions for which the integral is close to zero.

For these functions, the integral has exactly the same value as for the $u_1^\epsilon(x)$ since the integrand depends only on u'^2 and is therefore constant. In the same manner, the approximating functions for which the integral is close to unity consist of polygonal paths made up of straight line segments of slope $\pm \frac{1}{\epsilon}$ and zero. (See Figure 9.) For these functions, the integral has

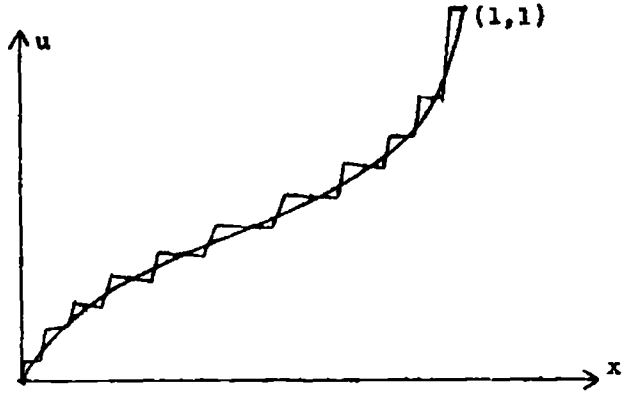


Figure 9. The approximating functions for which the integral is close to unity.

evidently the same value as for the $u_2^\epsilon(x)$ if the given admissible function is monotonic, since the integrand is piecewise constant and the total length of interval for which u' equals $\frac{1}{\epsilon}$ is the same in both cases.*

Thus, the rather surprising fact appears that it is possible to have a sequence of admissible functions converging to an admissible limit function with the values of the integral converging to the least upper bound or greatest lower bound of all possible values, and yet have the value of the integral for the limit function itself different from the least upper bound or

* The result holds even if the given admissible function is not monotonic, but the construction is then slightly more complicated.

greatest lower bound. In fact, in this problem, such a sequence can be found to converge to any admissible function.

7. Semi-continuity.

Let us examine the familiar proof of Weierstrass' theorem that every continuous function $F(u_1,\ldots,u_n) = F(P)$ in a bounded closed domain D of n-dimensional space possesses a minimum (and a maximum) in D. From the nature of F and D we know that the set of values of F(P) for P in D must be bounded, and hence have a greatest lower bound d. Then there is a sequence P_n of points of D such that $F(P_n) \to d$. Because D is bounded we can select a convergent subsequence P_n' of the sequence P_n, and $\overline{P} = \lim P_n'$ is in D because D is closed. Then

$$F(\overline{P}) = F(\lim P_n') = \lim F(P_n') = d$$

since F is continuous, and so F actually achieves the minimum d for the point $\overline{P}$ in D.

In this proof we have not made essential use of the full strength of the continuity of F(P). For if we knew merely that $F(\lim P_n') \leq \lim F(P_n')$, we would obtain by the same argument that $F(\overline{P}) \leq d$, which together with the fact that $F(\overline{P}) \geq d$ (since d = g.l.b. F(P) and $\overline{P}$ is in D) would again give that $F(\overline{P}) = d$. Hence if we define a lower semi-continuous function F(P) to be one such that

$$F(\lim P_n) \leq \underline{\lim}\, F(P_n)$$

for every convergent sequence P_n, we may state the following improved form of Weierstrass' theorem: Every lower semi-continuous function F(P) in a bounded closed domain D possesses a minimum in D. If we define upper semi-continuity analogously, it can easily be seen that a function is continuous if and only if it is both upper and lower semi-continuous.

We may define semi-continuity also for functionals; $F\{u\}$ is lower semi-continuous if, for every sequence $u_n(x)$ of admissible functions which converges uniformly to an admissible function,

$$F\{\lim u_n\} \leq \underline{\lim}\, F\{u_n\} .$$

One of the reasons for the importance of this concept in the Calculus of Variations is that very few functionals that occur in applications are continuous, whereas many of them are semi-continuous.

A few examples follow. The functional $G\{u\} = \int_{x_0}^{x_1} F(u(x))dx$, where F is a continuous function, is continuous, since if $u_n(x)$ converges uniformly to $u(x)$, then $G\{u_n\} \to G\{u\}$. The length $L\{C\}$ of a curve C is a lower semi-continuous functional but not an upper semi-continuous one, since if the curves C_n approximate (converge to) C then their lengths $L\{C_n\}$ must at least be approximately $L\{C\}$ but may be much larger (if the curves oscillate enough). The functional $I\{u\} = \int_0^1 \frac{1}{1 + u'^2}\, dx$ is neither upper nor lower semi-continuous, as the previous discussion of it shows.

8. Generalization of the Fundamental Lemma.

Suppose that $f(x)$ is a piecewise continuous function for $a \leq x \leq b$ and that

$$\int_a^b \zeta^{[k]}(x) f(x)dx = 0$$

for every function $\zeta(x)$ such that $\zeta, \zeta', \ldots, \zeta^{[k-1]}$ vanish at a and b and are continuous and $\zeta^{[k]}$ is piecewise continuous. It can then be proved that $f(x)$ must be a polynomial of degree at most k-1. In the special case k=1 this is just the statement of the Fundamental Lemma. The proof for the general case is analogous to the proof for k=1; to illustrate we shall carry it out for k=2.

If p and q are constants, we have for any permissible $\zeta(x)$, integrating by parts,

$$\int_a^b \zeta''(x)\{px + q\}dx = [\zeta'(x)\{px + q\}]_a^b - [\zeta(x)p]_a^b = 0 \quad ,$$

since ζ and ζ' vanish at a and b. Therefore

$$\int_a^b \zeta''(x)\{f(x) - px - q\}\, dx = 0 \quad .$$

We now try to determine p and q such that $f(x) - px - q$ is the second derivative of a permissible function $\zeta(x)$. Suppose that this were the case: $\zeta''(x) = f(x) - px - q$. Then since $\zeta(a) = \zeta'(a) = 0$,

$$\zeta'(x) = \int_a^x \{f(\xi) - p\xi - q\}\, d\xi = \int_a^x f(\xi)d\xi - \frac{p}{2}(x^2 - a^2) - q(x-a) \quad ,$$

$$\zeta(x) = \int_a^x \int_a^\eta f(\xi)d\xi d\eta - \frac{p}{2}\left[\frac{x^3 - a^2}{3} - a^2(x-a)\right] - q\left[\frac{x^2 - a^2}{2} - a(x-a)\right] \quad .$$

The requirement $\zeta(b) = \zeta'(b) = 0$ provides for the determination of p and q the two simultaneous equations

$$\frac{(b-a)(b+a)}{2}\, p + (b-a)q = \int_a^b f(\xi)d\xi \quad ,$$

$$\frac{(b-a)^2(b+2a)}{6}\, p + \frac{(b-a)^2}{2}\, q = \int_a^b \int_a^\eta f(\xi)d\xi d\eta \quad .$$

The determinant of the coefficients of p and q is given by

$$\frac{(b-a)(b+a)}{2} \cdot \frac{(b-a)^2}{2} - (b-a)\,\frac{(b-a)^2(b+2a)}{6} = \frac{(b-a)^4}{12} \neq 0 \quad ,$$

hence there exist constants p and q with the prescribed properties. Finally, the above constructed function $\zeta(x)$ is permissible since ζ'' equals $f(x) - px - q$ and is consequently piecewise continuous. Therefore

$$0 = \int_a^b \zeta''(x)\{f(x) - px - q\}\, dx = \int_a^b \{f(x) - px - q\}^2\, dx \quad ,$$

whence $f(x) \equiv px + q$, which is the desired conclusion for k=2.

9. Derivation of Euler's equation by special variations.

We wish to prove, by special choice of the function $\zeta(x)$, that if

$$\int_a^b \left\{\zeta(x)F_u(x) + \zeta'(x)F_{u'}(x)\right\} dx = 0$$

for all continuous functions $\zeta(x)$ which vanish at a and b and have piecewise continuous first derivatives, then $F_u(x) - \frac{d}{dx} F_{u'}(x) = 0$.

We illustrate the method by choosing

$$\zeta(x) = \begin{cases} (x-a)(\xi-x) & , \quad \text{for} \quad a \leq x \leq \xi \quad , \\ 0 & , \quad \text{for} \quad \xi < x \leq b \quad , \end{cases}$$

where ξ is arbitrary between a and b. Then

$$0 = \int_a^{\xi} \left\{(x-a)(\xi-x)F_u(x) + [(\xi-x) - (x-a)]F_{u'}(x)\right\} dx \quad .$$

Differentiation with respect to ξ gives

$$0 = -(\xi-a)F_{u'}(\xi) + \int_a^{\xi} \left\{(x-a)F_u(x) + F_{u'}(x)\right\} dx \quad .$$

Clearly $F_{u'}(\xi)$ must be differentiable at least for $\xi \neq a$, hence we obtain

$$0 = -F_{u'}(\xi) - (\xi-a)\frac{d}{d\xi} F_{u'}(\xi) + (\xi-a)F_u(\xi) + F_{u'}(\xi)$$

$$= (\xi-a)[F_u(\xi) - \frac{d}{d\xi} F_{u'}(\xi)] \quad .$$

Dividing by $(\xi-a)$ and then replacing the arbitrary ξ by x, we have the desired result. In a similar way Euler's equation can be derived by choosing other specific one-parameter families of functions $\zeta(x)$.

10. Characterization of the longer great circle arc.

We know that if A and B are two points on a sphere, not diametrically opposed, then the shorter great circle arc between

them is the unique solution to the problem of finding the curve of shortest length joining them on the sphere. We wish to characterize the longer great circle arc between them in a similar way, in this case as the solution of a minimax problem.

Let δ be a fixed great circle which separates A and B, let D be a point on δ, let $\Gamma(D)$ be the set of curves on the sphere which go from A to B and pass through D, and let $L\{\gamma\}$ be the length of any curve γ. Then

$$\min_{\gamma \text{ in } \Gamma(D)} L\{\gamma\}$$

is achieved for a unique curve $\gamma(D)$, consisting of the shorter great circle arcs from A to D and from D to B, and is a continuous function of D. Hence

$$\max_{D \text{ on } \delta} L\{\gamma(D)\} = \max_{D \text{ on } \delta} \min_{\gamma \text{ in } \Gamma(D)} L\{\gamma\}$$

is achieved for some $D = \bar{D}$ and the corresponding $\gamma(\bar{D})$. It is easily seen that this maximizing point $\bar{D}$ is unique, that it is the intersection of δ with the longer great circle arc between A and B, and that therefore $\gamma(\bar{D})$ is that longer great circle arc, as desired.

The curve from A to B which lies on the great circle determined by them and goes once around the sphere, covering the shorter arc twice, can be characterized in a similar way, in this case using two fixed great circles neither of which separates A and B which do not intersect on the great circle determined by A and B. The great circle arc from A to B which winds around the sphere any specified number of times can likewise be characterized as the solution of a minimax problem, by using a corresponding number of fixed great circles.

11. Integration of Euler's equation in special cases.

Euler's equation can be solved by means of quadratures in the cases that $F(x,u,u')$ does not contain u or x explicitly and

can be solved trivially in the case that $F(x,u,u')$ does not contain u' explicitly.

If $F(x,u,u')$ does not depend explicitly on its third argument, Euler's equation

$$F_u - \frac{d}{dx} F_{u'} = 0$$

reduces simply to

$$F_u = 0 \quad .$$

This equation can be regarded as defining u implicitly as a function of x, and Euler's equation is thus solved.

If u does not appear explicitly in $F(x,u,u')$, then Euler's equation reduces to

$$\frac{d}{dx} F_{u'} = 0$$

or

$$F_{u'} = k$$

where k is an arbitrary constant. This equation can be used to find u' as a function of x (and k), and u can subsequently be obtained by means of a quadrature.

Perhaps the most significant case is that in which x does not occur explicitly in $F(x,u,u')$. In this case, the Euler equation does not at first seem to simplify. Yet, x and u play similar geometric roles in the variational problem, and the absence of x ought to lead to a simplification just as the absence of u does. In order to find this simplification, the absence of x can be thought of as a condition which permits F to be differentiated as follows:

$$\frac{dF}{dx} = u'F_u + u''F_{u'}$$

or

$$\frac{dF}{dx} - u''F_{u'} = u'F_u \quad .$$

Subtracting $u' \frac{d}{dx} F_{u'}$ from both sides gives

$$\frac{dF}{dx} - u''F_{u'} - u' \frac{d}{dx} F_{u'} = u'F_u - u' \frac{d}{dx} F_{u'}$$

or

$$\frac{d}{dx}\,(F - u'F_{u'}) = u'(F_u - \frac{d}{dx}\,F_{u'}) \quad .$$

However, the right side vanishes since u is assumed to satisfy Euler's equation, and the conclusion is that

$$F - u'F_{u'} = k$$

where k is an arbitrary constant. Thus, the above equation must be satisfied by any function u which satisfies Euler's equation in the case that F *does not* explicitly contain x. This equation can be used for finding u' as a function of u (and k),* which means that $\frac{dx}{du}$ can be found as a function of u. Then, x as a function of u can be found by means of a quadrature.

It is very important to realize, however, that although the simplified condition

$$F - u'F_{u'} = k$$

and Euler's equation

$$F_u - \frac{d}{dx}\,F_{u'} = 0$$

are both necessary conditions, they are not equivalent. The exact situation, in view of the relation

$$\frac{d}{dx}\,(F - u'F_{u'}) = u'(F_u - \frac{d}{dx}\,F_{u'}) \quad ,$$

is clearly that every solution of the Euler equation satisfies the simplified condition (for some value of k), but that the simplified condition is satisfied both by

---

* Actually u' may be found to be a multi-valued function of u. This function, for example, frequently contains a square root where both the positive and negative determinations must be used. In fact, it can be seen that if u' is a single-valued function of u, then the continuous function u must necessarily be a monotone function of x. And there are many problems where F does not explicitly contain x, and yet the extremizing curve is not monotone. An illustration of this point is afforded, for example, by the problem of finding the surface of revolution of least area or by the Brachistochrone problem. These problems will be discussed later in Section 13.

$$u' \neq 0$$

and by the solutions of Euler's equation and by no other functions. It is evident that

$$u = c$$

where c is any constant always satisfies the simplified condition (for some value of k).

Since

$$u = c$$

always satisfies the simplified condition it would be desirable to have a simple criterion for determining whether or not it satisfies the Euler equation in its original form. It is clear that $F_{u'}$ reduces to a constant when u is set equal to c and u' is set equal to zero since F (and consequently $F_{u'}$) does not contain x explicitly. Therefore, the Euler equation

$$F_u - \frac{d}{dx} F_{u'} = 0$$

reduces simply to

$$F_u = 0$$

for functions of the form

$$u = c \quad .$$

Clearly, a necessary and sufficient condition for a function of this form to satisfy Euler's equation is

$$F_u \Big|_{\substack{u = c \\ u' = 0}} = 0 \quad .$$

The above results can be summarized in the following theorem:

Theorem 1: If the integrand function F does not depend explicitly on x, then the Euler equation

$$F_u - \frac{d}{dx} F_{u'} = 0$$

and the "simplified condition"

$$F - u'F_{u'} = k$$

are related as follows:

a) Every solution of the Euler equation is a solution of the simplified condition for some value of k;

b) For any value of k, every solution of the simplified condition no arc of which is of the form

$$u = c$$

(where c is a constant) is a solution of the Euler equation;

c) The function

$$u = c$$

satisfies the Euler equation if and only if

$$F_u \Big|_{\substack{u = c \\ u' = 0}} = 0$$

and always satisfies the simplified condition for some value of k; or, expressed in other words (in a slightly weaker form):

The set of integral curves of the Euler equation is precisely the set of integral curves of the simplified condition no arc of which is of the form

$$u = c$$

where

$$F_u \Big|_{\substack{u = c \\ u' = 0}} \neq 0 \quad .$$

12. The shortest connection between two points on a sphere.

That the shorter great circle arc connecting two points on a sphere has minimum length among all connections on the sphere can be proved in a quite elementary geometric way. This geometric proof will now be described, and then the same result will be obtained from an analytic formulation of the problem.

The geometric proof is based on the theorem that the sum of the lengths of two sides of a spherical triangle (where each side is the shorter great circle arc connecting two of the vertices) is greater than the length of the third side. If the vertices of a spherical triangle are joined by straight line segments to the center of the sphere, it is clear that this theorem is equivalent to the theorem that the sum of two face angles of a trihedral angle is greater than the third face angle. But this theorem about trihedral angles can be proved in a very simple way by using the fact that the sum of the lengths of any two sides of a plane triangle is greater than the length of the third side.

As an immediate consequence of the theorem about spherical triangles it follows that the length of any spherical polygonal path (made up of shorter great circle arcs) is greater than the length of the shorter great circle arc joining the end-points. This result can be seen as follows: Shorten the spherical polygonal path by joining any pair of alternate vertices (i.e., vertices with exactly one other vertex between them along the spherical polygonal path) by the shorter of the two possible great circle arcs. The theorem about spherical triangles guarantees that this shortens the path. Then, apply the same process to the shortened path. Repeat the process until, after a finite number of steps, the result will be simply a great circle arc joining the two points. Thus, the shorter great circle arc is certainly shorter than the spherical polygonal path.

The result for spherical polygonal paths, however, implies the desired result for any curve. All that remains to be shown is that the length of any curve on the sphere can be approximated to any prescribed degree of accuracy by the length of an inscribed spherical polygonal path. It would then follow that no curve can provide a shorter connection than the great circle arc.

In order to prove that the length of any curve on the sphere can be approximated by the length of an inscribed spherical polygonal path, the elementary geometric definition of the

length of a curve will be recalled. Consider a sequence of n+1 points P_n^0, P_n^1,...,P_n^n on the curve in that order where P_n^0 and P_n^n are the end-points. Join consecutive points by straight line segments, giving an inscribed space polygonal path π_n joining the end-points of the curve. Let L_n denote the length of π_n. Allow n to tend to infinity, choosing the points P_n^i (i=0,1,...,n) so that the length of the longest side of the space polygonal path should tend to zero. Then, $\lim_{n\to\infty} L_n = L$ is, by definition, the length of the curve.*

It will now be shown that the length of the spherical polygonal path having the same vertices as the space polygonal path π_n approaches a limit as $n \to \infty$ and that this limit is, in fact, equal to L, the length of the curve. Let the length of the side of the space polygonal path joining P_n^{i-1} and P_n^i be denoted by $\lambda_n^{(i)}$ and the length of the corresponding side of the spherical polygonal path by $\lambda_n^{(i)*}$. Thus, the length L_n of the space polygonal path π_n is given by

$$L_n = \sum_{\nu=1}^{n} \lambda_n^{(\nu)}$$

and the length L_n^* of the corresponding spherical polygonal path is given by

$$L_n^* = \sum_{\nu=1}^{n} \lambda_n^{(\nu)*} \quad .$$

It remains to be proved that L_n^* converges and that

$$\lim_{n\to\infty} L_n^* = L \quad .$$

Since a product of limits equals the limit of the product, and since L_n^* is the product of L_n and L_n^*/L_n, it is enough to prove that L_n^*/L_n converges and that

---

* Curves for which the limit exists (independent of the way in which the points P_n^i are chosen) are called rectifiable curves. Only such curves are of interest in the present discussion.

$$\lim_{n\to\infty} \frac{L_n^*}{L_n} = 1 \quad .$$

It is immediately obvious that

$$L_n^* > L_n$$

or

$$\frac{L_n^*}{L_n} > 1 \tag{5}$$

since arcs of circles are longer than the subtended chords. On the other hand, the points $P_n^{(i)}$ were chosen so that the length of the longest chord tends to zero as n tends to infinity, and the ratio of arc length to chord length in a circle of constant radius tends to unity as the chord length tends to zero. This implies that

$$\lim_{n\to\infty} \frac{\lambda_n^{(\nu)*}}{\lambda_n^{(\nu)}} = 1$$

where the convergence is uniform with respect to ν since the sequence of longest chords converges more slowly than any other sequence of chords. In other words, for every $\varepsilon > 0$, there exists a number $N = N(\varepsilon)$ such that

$$\left| \frac{\lambda_n^{(\nu)*}}{\lambda_n^{(\nu)}} - 1 \right| < \varepsilon \qquad \text{for } n > N$$

or, in view of (5):

$$\frac{\lambda_n^{(\nu)*}}{\lambda_n^{(\nu)}} < 1 + \varepsilon$$

or

$$\lambda_n^{(\nu)*} < \lambda_n^{(\nu)}(1 + \varepsilon) \quad . \tag{6}$$

An estimate for L_n^*/L_n can now be made as follows in view of (6):

$$\frac{L_n^*}{L_n} = \frac{\sum_{\nu=1}^{n} \lambda_n^{(\nu)*}}{\sum_{\nu=1}^{n} \lambda_n^{(\nu)}} < \frac{\sum_{\nu=1}^{n} \lambda_n^{(\nu)}(1+\epsilon)}{\sum_{\nu=1}^{n} \lambda_n^{(\nu)}} = 1 + \epsilon$$

or

$$\frac{L_n^*}{L_n} < 1 + \epsilon \qquad \text{for } n > N \quad . \tag{7}$$

However, (7) together with (5) clearly implies that L_n^*/L_n converges and that

$$\lim_{n\to\infty} \frac{L_n^*}{L_n} = 1 \quad .$$

This in turn implies that

$$\lim_{n\to\infty} L_n^* = L \quad ,$$

which means that the length of any curve on the sphere can be approximated by the lengths of inscribed spherical polygonal paths. And, as already indicated, this implies that the shorter great circle arc joining two points on a sphere provides a connection of minimum length.

The same result will now be obtained analytically by writing parametric equations for a sphere of radius r with its center at the origin, as follows:

$$\begin{aligned} x &= r \cos\phi \sin\theta \\ y &= r \sin\phi \sin\theta \\ z &= r \cos\theta \end{aligned} \tag{8}$$

for $0 \le \phi \le 2\pi$ and $0 \le \theta \le \pi$ where ϕ and θ can be identified in the usual way as the angles of spherical (polar) coordinates. Any curve on the sphere can be represented by letting ϕ and θ be functions of a parameter t, as follows:

$$\begin{aligned} \phi &= f(t) \\ \theta &= g(t) \quad . \end{aligned}$$

In order to calculate arc length for such a curve, $(ds/dt)^2$ will now be calculated as follows:

$$\left(\frac{ds}{dt}\right)^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2 = r^2 \sin^2\theta\, \phi'^2 + r^2\theta'^2$$

where $\phi' = d\phi/dt$ and $\theta' = d\theta/dt$. Thus, the length L of a curve joining the points in question is given by

$$L = r\int_{t_o}^{t_1} \sqrt{\sin^2\theta\, \phi'^2 + \theta'^2}\, dt \quad .$$

It thus appears that the problem reduces to finding a curve which minimizes the following integral I (omitting the unessential r):

$$I = \int_{t_o}^{t_1} \sqrt{\sin^2\theta\, \phi'^2 + \theta'^2}\, dt \quad .$$

The solution to the problem becomes immediately apparent when the symmetry of the sphere is taken advantage of by placing the two points in a special position. Place the two points on the same line of longitude and make the reasonable assumption that θ can be used as a parameter for the minimizing curve. In other words, assume a solution of the form

$$\phi = G(\theta) \quad .$$

Then, the integral to be minimized is

$$I^* = \int_{\theta_o}^{\theta_1} \sqrt{\sin^2\theta\, \phi'^2 + 1}\, d\theta$$

where $\phi' = d\phi/d\theta$. The integrand and hence the integral is clearly minimized for $\phi' = 0$, and the minimizing curve must be the shorter arc of the line of longitude or great circle through the two points.

It would be desirable, however, to reach the same conclusion without placing the points in any special position and without making any use of the geometric interpretation of ϕ and θ. This will now be done with the aid of Euler's equation.

The integrand function F of the integral I for this variational problem is

$$F(t,\phi,\theta,\phi',\theta') = \sqrt{\sin^2\theta\ \phi'^2 + \theta'^2} \quad .$$

The Euler equation for the two functions ϕ and θ are

$$F_\phi - \frac{d}{dt} F_{\phi'} = 0$$

$$F_\theta - \frac{d}{dt} F_{\theta'} = 0$$

or

$$-\frac{d}{dt} \frac{\sin^2\theta\ \phi'}{\sqrt{\sin^2\theta\ \phi'^2 + \theta'^2}} = 0$$

$$\frac{\sin\theta\cos\theta\ \phi'^2}{\sqrt{\sin^2\theta\ \phi'^2 + \theta'^2}} - \frac{d}{dt} \frac{\theta'}{\sqrt{\sin^2\theta\ \phi'^2 + \theta'^2}} = 0 \quad .$$

In order to simplify the Euler equations, the parameter t will be renamed s and will be identified with arc length on the curve represented by

$$\phi = f(s)$$
$$\theta = g(s) \quad .$$

It then follows that

$$(\frac{ds}{ds})^2 \sin^2\theta\ \dot{\phi}^2 + \dot{\theta}^2 = 1 \tag{9}$$

where $\dot{\phi} = d\phi/ds$ and $\dot{\theta} = d\theta/ds$. The Euler equations can then be written as follows:

$$\frac{d}{ds}(\sin^2\theta\ \dot{\phi}) = 0$$
$$\frac{d}{ds}\dot{\theta} = \sin\theta\cos\theta\ \dot{\phi}^2 \quad . \tag{10}$$

These equations represent a necessary condition for a minimizing curve, and it will now be shown that any curve satisfying these equations is a great circle arc.

Cartesian parametric equations for the curve represented by the functions ϕ and θ can be obtained by substituting ϕ and θ as functions of s into (8). This gives x, y, and z as functions of s. To show that this curve is a great circle arc, it must be shown that (x,y,z) lies on the same plane through the

origin for all values of x. In other words, it must be shown that there exists constants A, B, and C such that

$$Ax + By + Cz \equiv 0$$

where at least one of the constants is different from zero. That is, it must be shown that x, y, and z as functions of s are linearly dependent. This linear dependency will be demonstrated by showing that the three functions x, y, and z all satisfy the following differential equation:

$$\ddot{u} + u = 0 \quad .$$

This second order linear differential equation (whose solutions, of course, are linear combinations of sin s and cos s) cannot have more than two independent solutions according to the theory of differential equations, and it would follow that x, y, and z are linearly dependent. Therefore, it should be shown that

$$\begin{aligned} \ddot{x} &= -x \\ \ddot{y} &= -y \\ \ddot{z} &= -z \quad . \end{aligned}$$

As an illustration, $\ddot{x}$ will now be calculated. The function x is given by

$$x = r \cos \phi \sin \theta \quad .$$

Its derivative $\dot{x}$ is given by

$$\dot{x} = r(-\sin \phi \sin \theta \dot{\phi} + \cos \phi \cos \theta \dot{\theta})$$

or

$$\dot{x} = r(-\frac{\sin \phi}{\sin \theta} \sin^2 \theta \dot{\phi} + \cos \phi \cos \theta \dot{\theta}) \quad .$$

Differentiating once more and using the Euler equations (10), it follows that $\ddot{x}$ is given by

$$\ddot{x} = r[-\dot{\phi}(\cos \phi \sin \theta \dot{\phi} - \sin \phi \cos \theta \dot{\theta}) + \cos \phi \sin \theta \cos^2 \theta \dot{\phi}^2 + \dot{\theta}(-\sin \phi \cos \theta \dot{\phi} - \cos \phi \sin \theta \dot{\theta})] \quad .$$

The terms in $\dot{\phi}\dot{\theta}$ drop out and the terms in $\dot{\phi}^2$ and $\dot{\theta}^2$ can be collected separately while factoring out $-\cos \phi \sin \theta$, giving:

$$\ddot{x} = -r \cos \phi \sin \theta \left[(1 - \cos^2\theta)\dot{\phi}^2 + \dot{\theta}^2\right]$$

or

$$\ddot{x} = -r \cos \phi \sin \theta \left(\sin^2\theta\, \dot{\phi}^2 + \dot{\theta}^2\right) \quad .$$

In view of (9), it follows that

$$\ddot{x} = -r \cos \phi \sin \theta = -x \quad .$$

It can be shown in exactly the same way that $\ddot{y} = -y$, and the calculation for showing that $\ddot{z} = -z$ is even simpler.

It has now been shown that the three functions x, y, and z satisfy the same second order linear differential equation. Consequently, x, y, and z as functions of s are linearly dependent, and any curve satisfying the Euler equations (10) must lie in a plane through the origin. In other words, all curves on the sphere satisfying the Euler equations are great circle arcs. And, since the Euler equations constitute a necessary condition for a minimum, it follows that, if a minimizing curve is assumed to exist, then a great circle arc must furnish the minimum.

13. Application of Euler's equation to classical problems.

Three classical problems will now be solved with the aid of Euler's equation; namely, the Brachistochrone problem, the problem of finding the surface of revolution of least area, and the Isoperimetric problem. Of course, in each case the existence of a solution will be assumed and the only possible solution will be found.

On pages 170-171 of the notes, the Brachistochrone problem is formulated and it is shown that the total time of descent T of a particle sliding along the curve

$$u = f(x)$$

from the point (0,0) to the point (x_1, u_1) is given by

$$T = \frac{1}{\sqrt{2g}} \int_0^{x_1} \frac{\sqrt{1 + u'^2}}{\sqrt{u}}\, dx$$

(where the positive u-axis is chosen in the downward direction). Thus, the Brachistochrone problem amounts to minimizing the integral

$$I = \int_0^{x_1} \frac{\sqrt{1 + u'^2}}{\sqrt{u}}\, dx \quad .$$

In this case, the function

$$F(x,u,u') = \frac{\sqrt{1 + u'^2}}{\sqrt{u}}$$

does not depend explicitly on x, and the following simplified form of Euler's equation (as found in Section 8) can be used:

$$F - u'F_{u'} = k$$

where k is an arbitrary constant. For the particular function F of this problem, this equation becomes

$$\frac{\sqrt{1 + u'^2}}{\sqrt{u}} - \frac{u'^2}{\sqrt{u}\sqrt{1 + u'^2}} = k$$

or

$$\frac{1}{\sqrt{u}\sqrt{1 + u'^2}} = k \quad .$$

In view of Theorem 1 of Section 8, it would be advisable to determine first of all whether the solutions of the above simplified condition of the form

$$u = c$$

satisfy the Euler equation in its original form. For the integrand function F of this problem, it is clear that

$$F_u \Big|_{\substack{u = c \\ u' = 0}} = -\frac{1}{2c^{3/2}} \neq 0$$

and, therefore, the set of integral curves of the Euler equation is precisely the set of integral curves of the simplified condition no arc of which is of the form

$$u = c \quad .$$

It is most convenient to find the solutions of the simplified condition in parametric form. For this purpose, it will first be shown that u' can be used as a parameter on every curve satisfying the Euler equation of this particular problem; that is, it will be shown that u" vanishes on no solution curve. All that has to be shown is that the vanishing of u" at any point is incompatible with the Euler equation. If the function F of this problem is substituted into the Euler equation and then u" is set equal to zero, the result turns out to be

$$\frac{-1}{2u^{3/2}\sqrt{1 + u'^2}} = 0 \quad ,$$

which clearly cannot hold at any point where u' exists. Thus u" cannot vanish and u' can be used as a parameter.

In order to find the integral curves of

$$\frac{1}{\sqrt{u}\sqrt{1 + u'^2}} = k \quad ,$$

solving for u immediately expresses u in terms of the parameter u' as follows:

$$u = \frac{1}{k^2(1 + u'^2)} \quad .$$

However, instead of proceeding to obtain x as a function of the parameter u', a new parameter θ will be introduced by the relation

$$u' = \cot \frac{\theta}{2} \quad ,$$

where the monotone character of the cotangent function insures that θ can be used as a parameter since u' can.

Expressing u in terms of the new parameter θ gives

$$u = \frac{1}{k^2(1 + \cot^2(\frac{\theta}{2}))} = \frac{1}{k^2} \sin^2(\frac{\theta}{2}) = \frac{1}{2k^2}(1 - \cos \theta) \quad .$$

Then, x as a function of the parameter θ can be found by first finding $dx/d\theta$ with the chain rule of differentiation, as follows:

$$\frac{dx}{d\theta} = \frac{dx}{du}\frac{du}{d\theta} = \frac{1}{u'}\frac{du}{d\theta} = \frac{1}{\cot\frac{\theta}{2}}\frac{du}{d\theta} \quad .$$

Byt du/dθ can be calculated from the relation found above expressing u in terms of θ. The result of calculating du/dθ and substituting is

$$\frac{dx}{d\theta} = \frac{1}{\cot\frac{\theta}{2}}\frac{1}{2k^2}\sin\theta = \frac{1}{2k^2}(1 - \cos\theta)$$

since

$$\cot\frac{\theta}{2} \equiv \frac{\sin\theta}{1 - \cos\theta} \quad .$$

It then follows that

$$x = \frac{1}{2k^2}(\theta - \sin\theta)$$

where the constant of integration has been set equal to zero, which means that it has been decided that θ = 0 should correspond to the point (0,0).

Thus, all solutions of the Euler equation are arcs of a cycloid* of the form

$$x = \frac{1}{2k^2}(\theta - \sin\theta)$$

$$u = \frac{1}{2k^2}(1 - \cos\theta)$$

or

$$x = a(\theta - \sin\theta)$$
$$u = a(1 - \sin\theta)$$

where $1/2k^2$ is renamed a. That is, the solution of the

---

* It is unnecessary to consider the possibility of the minimizing curve being more than a complete arch of a cycloid (in which case u' could not be used as a parameter) since the time of descent could be shortened by rounding off the cusp at the top, thereby shortening the path and increasing the speeds.

Brachistochrone problem, if it exists, must be an arc of a cycloid.*

It might be remarked that the reason that the simplified equation was solved using u' as a parameter rather than by solving for u' as a function of u was that u' is not a single-valued function of u. If the simplified equation is solved for u', the result is

$$u' = \pm\sqrt{\frac{1}{uk^2} - 1}$$

and the plus-or-minus sign must be taken into consideration when solving the equation. Of course, this means that integral curves must be found which agree in direction at every point with either of the direction fields D^+ and D^- (See Figure 10) given by:

$$D^+: \quad u' = \sqrt{\frac{1}{uk^2} - 1}$$

$$D^-: \quad u' = -\sqrt{\frac{1}{uk^2} - 1} \quad .$$

However, instead of following through with this approach, it was preferred to solve for u as a single-valued function of u' regarded as a parameter.

It might further be of interest to note that for certain positions of the end-point (x_1, u_1), it can be seen that the path of quickest descent takes the sliding particle lower than the end-point and then up to the end-point (in which case the

---

* One slight refinement might be made in the reasoning given above. Since the integrand function F is infinite at the origin, it would be better to consider problems where the particle is given an initial velocity v_o and then to allow v_o to tend to zero. It turns out that all that has to be done if the initial velocity is v_o is replace u by $u + (v_o^2/2g)$ (where g is the constant representing gravitational acceleration) in the integrand function F. The solution in this case is also a cycloid, and this cycloid approaches the cycloid found above when v_o tends to zero.

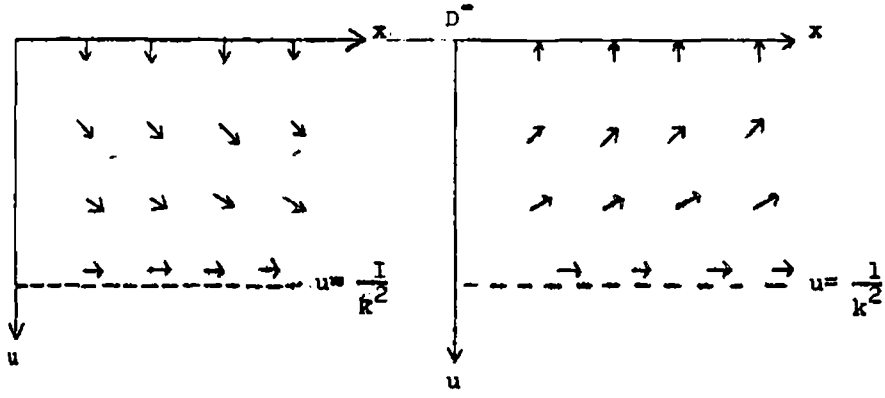


Figure 10. The direction in fields D^+ and D^-.

particle switches from the D^+ direction field to the D^- direction field at the line $u = 1/k^2$). On the other hand, for other positions of the end-point (x_1, u_1), it can be seen that the particle slides along a monotone arc for quickest descent.

The problem of finding the surface of revolution of minimum area can be formulated as follows: Join two given points (on the same side of the x-axis) in the x-u plane by a curve such that the surface obtained by rotating the curve about the x-axis should have minimum surface area. This area A of the surface obtained in this way by rotating the curve

$$u = f(x)$$

around the x-axis is given by

$$A = 2\pi \int_{x_0}^{x_1} u \sqrt{1 + u'^2}\, dx$$

and the integral to be minimized is

$$I = \int_{x_0}^{x_1} u \sqrt{1 + u'^2}\, dx \quad .$$

Clearly the function

$$F(x,u,u') = u\sqrt{1 + u'^2}$$

does not explicitly contain x and the simplified form of Euler's equation

$$F - u'F_{u'} = k$$

can be used. Thus, a necessary condition for a minimum is

$$u\sqrt{1 + u'^2} - \frac{uu'^2}{\sqrt{1 + u'^2}} = k$$

or

$$\frac{u}{\sqrt{1 + u'^2}} = k$$

or, solved for u'

$$u' = \pm\sqrt{\frac{u^2}{k^2} - 1} \quad .$$

However, instead of solving the simplified equation when it is solved for u' in which case the plus-or-minus sign would have to be considered, the simplified equation will be solved by essentially the same method that was used for the Brachistochrone problem; i.e., by obtaining the integral curves in parametric form.

It can be seen in the same way as for the Brachistochrone problem that all solutions of the simplified equation of the form

$$u = c$$

can be disregarded since they do not satisfy the Euler equation in its original form*, and it can also be seen that u' can be used as a parameter. Then, solving the simplified equation for u expresses u in terms of the parameter u' as follows:

$$u = k\sqrt{1 + u'^2} \quad .$$

---

* In this case, $F_u \Big|_{\substack{u = c \\ u' = 0}} = 1 \neq 0$.

However, a new parameter θ given by

$$u' = \sinh \theta$$

will be introduced, where the monotone character of the hyperbolic sine guarantees that θ can be used as a parameter since u' can.

In terms of the new parameter θ, u is given by

$$u = k\sqrt{1 + \sinh^2\theta} = k \cos \theta \quad .$$

Then, according to the chain rule of differentiation:

$$\frac{dx}{d\theta} = \frac{dx}{du}\frac{du}{d\theta} = \frac{1}{u'}\frac{du}{d\theta} = \frac{1}{\sinh \theta}\frac{du}{d\theta} \quad . \tag{11}$$

But

$$\frac{du}{d\theta} = k \sinh \theta$$

and, therefore, (in view of (11)):

$$\frac{dx}{d\theta} = k$$

and

$$x - a = k\theta \quad .$$

The integral curves are therefore:

$$x - a = k\theta$$
$$u = k \cosh \theta$$

or, upon eliminating the parameter θ:

$$u = k \cos \frac{x - a}{k} \quad ,$$

which is a catenary. Again, of course, the assumption has been made that a minimizing curve exists.

As explained on page 187 of the notes, the classical Isoperimetric problem can be reduced to finding the arc of given length L with both end-points on the x-axis such that the area bounded by the arc and the x-axis is maximum. Let parametric equations for the arc with the arc lengths used as parameter be

$$x = \phi(s)$$
$$u = \psi(s)$$

with

$$\dot{x}^2 + \dot{u}^2 = 1$$

or, (assuming that x is an increasing function of s and therefore, taking the positive determination of the square root):

$$\dot{x} = \sqrt{1 - \dot{u}^2} \quad , \tag{12}$$

where the dots denote differentiation with respect to s. The area in question is given by

$$A = \int_{x_0}^{x_1} u\,dx = \int_0^L u\dot{x}\,ds = \int_0^L u\sqrt{1 - \dot{u}^2}\,ds$$

and the function

$$F(x,u,\dot{u}) = u\sqrt{1 - \dot{u}^2}$$

for this variational problem does not depend explicitly on s. The simplified form of Euler's equation for this case is

$$F - \dot{u}F_{\dot{u}} = k$$

or

$$u\sqrt{1 - \dot{u}^2} + \frac{u\dot{u}^2}{\sqrt{1 - \dot{u}^2}} = k$$

or

$$\frac{u}{\sqrt{1 - \dot{u}^2}} = k \quad .$$

It can be seen for this problem, just as for the last two problems, that solutions of the form

$$u = c$$

can be disregarded and that $\dot{u}$ can be used as a parameter.

Solving the simplified equation for u in terms of the parameter $\dot{u}$ gives

$$u = k\sqrt{1 - \dot{u}^2} \quad . \tag{13}$$

It is clear from (13) that the parameter $\dot{u}$ never exceeds unity in absolute value and that a new parameter θ can, therefore, be introduced by the relation

$$\dot{u} = \cos \theta \quad ,$$

where the monotone character of the cosine function in an appropriate interval insures that θ can be used as a parameter since $\dot{u}$ can.

In terms of the new parameter θ, u is given by

$$u = k\sqrt{1 - \cos^2\theta} = k \sin \theta$$

In order to find s in terms of θ, $ds/d\theta$ will first be found as follows:

$$\frac{ds}{d\theta} = \frac{ds}{du}\frac{du}{d\theta} = \frac{1}{\dot{u}}\frac{du}{d\theta} = \frac{1}{\cos \theta}\frac{du}{d\theta} \quad .$$

Since

$$\frac{du}{d\theta} = k \cos \theta$$

it follows that

$$\frac{ds}{d\theta} = k$$

and

$$s - a = k\theta \quad .$$

Thus, the solution (in the su-plane) is given by

$$s - a = k\theta$$
$$u = k \sin \theta$$

or, upon eliminating the parameter θ,

$$u = k \sin \frac{s - a}{k}$$

or, since u equals zero for the initial point,

$$u = k \sin \frac{s}{k} \quad .$$

In order to obtain x as a function of s, it will be noted that (12) and (13) together imply

$$\dot{x} = \frac{u}{k}$$

and, therefore,

$$\dot{x} = \sin \frac{s}{k}$$

and

$$x - b = -k \cos \frac{s}{k} \quad .$$

Thus, the solution is

$$x - b = -k \cos \frac{s}{k}$$

$$u = k \sin \frac{s}{k}$$

which is clearly a semi-circle. This means that the semi-circle is the arc of given length with both end-points on the x-axis such that the area bounded by the arc and the x-axis is maximum. It follows that the circle encloses maximum area among all non-self-intersecting closed plane curves of the same perimeter provided that such a maximizing curve exists.

14. Invariance and the Euler Expression.

The notion of invariance plays a major role in the mathematica and physics of the twentieth century. In particular, relativistic physics is grounded on principles of invariance. This concept will now be discussed, and a very important rule of invariance for Euler's equation will be brought to light.

The principle of invariance in a limited form held a fundamental position in the seventeenth century mechanics of Galileo and Newton. This restricted concept of invariance was embodied in the classical principle of relativity, which may be formulated in the following way: If two coordinate systems K and K^* are so related that K^* is in a state of uniform motion of translation with respect to K (i.e., if K^* moves without rotation at a constant velocity with respect to K), then the general laws of mechanics have the same formulation in the K^* system as in the K system. In other words, the laws of mechanics are "invariant under all transformations" of coordinates such that the original and transformed systems are related by uniform translatory motion. For example, if the Newtonian formula

holds with respect to an observer standing on the ground, then it also holds with respect to an observer on a train traveling in a straight line at constant speed. This is an expression of the invariant way in which natural phenomena run their course without depending on the coordinate system used to describe them.

In everyday life, a very similar principle of invariance would be taken for granted. If a man should change his name from Jack to George, this change would not necessarily influence his character and other general traits. In other words, the essence of any object is invariant under all changes of names which are merely used to label the object. And, since coordinates are simply mathematical labels, it is to be expected that the laws of physics should be invariant under changes in the coordinate system used for describing physical phenomena. It is just the old story of a rose smelling as sweet by any other name.

The modern principle of relativity is based on a notion of invariance including all transformations of coordinates. The transformed system may be in any state of accelerated motion with respect to the original system. Einstein envisaged the need for physics of using mathematical entities invariant under all such coordinate transformations. If these invariant entities are used in writing the formulas of physics, the formulas would then be invariant under general coordinate transformations. It is therefore of great interest in mathematics to investigate invariance and invariant mathematical entities.

An extremely simple invariant mathematical entity is the vector. In fact, vectors are useful for providing insight into geometrical and physical situations just because of their invariant nature; i.e., just because they depend only on intrinsic properties of the particular situation. Thus, it would be in keeping with the point of the present discussion to explain precisely what is meant by saying that a vector is invariant. Then, it will be shown later that the left side of Euler's equation enjoys this property of invariance in a very analogous sense.

For the sake of simplicity, the meaning of the invariance of the vector will be explained first only for transformations from one rectangular Cartesian coordinate system to another of the same type. Thus, the transformations to be considered are combined translations and rotations. These transformations, of course, constitute linear transformations

$$\bar{x}_i = \sum_{k=1}^{n} a_{ik}x_k + \beta_i \qquad (i = 1,2,\ldots,n) \tag{14}$$

where the matrix of the a_{ik} is orthogonal; i.e., where

$$\sum_{i=1}^{n} a_{ik}a_{im} = \delta_k^m = \begin{cases} 0 & \text{if } k \neq m \\ 1 & \text{if } k = m \end{cases} . \tag{15}$$

An n-dimensional vector is defined as an entity which can be realized in every n-dimensional coordinate system by giving an ordered n-tuple of numbers (called the components of the vector in that coordinate system) where the n-tuples corresponding to different coordinate systems are so inter-related that they possess the following property called invariance: If the coordinates transform according to (14), then the components $(u_1,u_2,\ldots,u_n)$ of the vector transform according to the corresponding homogeneous equations

$$\bar{u}_i = \sum_{k=1}^{n} a_{ik}u_k \qquad (i = 1,2,\ldots,n) \quad . \tag{16}$$

In other words, the components of the vector transform like coordinates or, more precisely, like differences of coordinates. (In fact, the components of an ordinary displacement vector are actually the differences of the coordinates of the end-points of the displacement.) Thus, saying that a vector is invariant means, by definition, that its components transform according to (16) when the coordinates transform according to (14).

As a consequence of the invariance of a vector as defined above, vector equations are invariant in the sense that they undergo no change at all when the coordinates are transformed.

For example, if the inner product (scalar product) of the vectors U and V equals the number λ for a particular choice of coordinate system; i.e., if

$$U \cdot V = \lambda \quad ,$$

then the same equation is valid no matter which coordinates are used. This follows at once from the invariance of vectors in the following way: The inner product in the transformed coordinate system, say $\bar{\lambda}$, is given by:

$$\bar{\lambda} = \sum_{i=1}^{n} \bar{u}_i \bar{v}_i \quad .$$

However, in view of (16) it follows that

$$\bar{\lambda} = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} u_k \sum_{m=1}^{n} a_{im} v_m \right)$$

or

$$\bar{\lambda} = \sum_{i,k,m=1}^{n} a_{ik} u_k a_{im} v_m$$

or

$$\bar{\lambda} = \sum_{k,m=1}^{n} \left(\sum_{i=1}^{n} a_{ik} a_{im} \right) u_k v_m \quad .$$

After having changed the order of summation in this way, it follows in view of (15) that

$$\bar{\lambda} = \sum_{r=1}^{n} u_r v_r = \lambda \quad .$$

Therefore

$$U \cdot V = \lambda$$

independent of the choice of coordinate system. As an example, work in mechanics is given for all coordinate systems by the inner product of the vector force involved and the vector displacement.

In general, saying that a mathematical entity is invariant means that it undergoes coordinate transformations according to

a certain law which holds for all coordinate transformations. Moreover, in order for such a law of transformation to be said to express the invariance of the mathematical entity, it must follow as a consequence of the law that "inner products" are invariant in the sense that they undergo no change at all when the coordinate system is transformed.

Actually vectors are defined so as to be invariant not only with respect to orthogonal linear transformations, but also with respect to the general transformation of coordinates

$$\bar{x}^i = F^i(x^1, x^2, \ldots, x^n) \qquad (i = 1,2,\ldots,n)$$

which, it will be assumed, has an inverse. In this case, two different kinds of vectors are defined, contravariant vectors and covariant vectors. A contravariant vector* $(u^1, u^2, \ldots, u^n)$ is defined so as to transform like differences of coordinates in the small; that is, like infinitesimal displacements.** On the other hand, the law of transformation for a covariant vector $(v_1, v_2, \ldots, v_n)$ is expressly designed so that the inner product of a contravariant and covariant vector should be a scalar; i.e., independent of the coordinate system. That is, covariant vectors transform in such a way that

---

* It is customary to use superscripts for the components of a contravariant vector and subscripts for the components of a covariant vector. In fact, when the distinction is made between the two types of vectors, it is proper to use superscripts for the coordinates themselves since they are akin to the components of a contravariant vector.

** Precisely, contravariant vectors are defined so that their components transform according to

$$\bar{u}^i = \sum_{k=1}^{n} \frac{\partial \bar{x}^i}{\partial x^k} u^k \quad ,$$

whereas infinitesimal displacements transform according to the chain rule of differentiation:

$$d\bar{x}^i = \sum_{k=1}^{n} \frac{\partial \bar{x}^i}{\partial x^k} dx^k \quad .$$

$$\sum_{i=1}^{n} \bar{u}^i \bar{v}_i = \sum_{r=1}^{n} u^r v_r$$

Examples of contravariant vectors are displacements and velocities, while examples of covariant vectors are gradients and forces.

To sum up, there are two kinds of vectors; namely, those which behave like infinitesimal displacements and those which behave like gradients. Both types of vectors are defined so as to be invariant; that is, so as to transform in very special ways. As a result, the inner product of a vector like an infinitesimal displacement and a vector like a gradient is independent of the coordinate system.

Before proceeding to explain and demonstrate the invariant nature of the Euler expression (i.e., the left side of Euler's equation), it will be of interest to show first that the Euler expression is very similar to another invariant entity, the ordinary gradient. Just as the gradient of a function is a vector (whose components are simply the first partial derivatives of the function), it will be seen that the Euler expression is analogous to this type of vector and might well be called the gradient of its associated functional. In order to uncover this gradient-like nature of the Euler expression, the extremizing problem for a functional will now be compared with the ordinary extremizing problem for a function of n variables (as is done on pages 12-13 of the notes).

The extremizing problem for a function of n variables which will be considered here is formulated as follows: Find a necessary condition for a point $(u_1^o, u_2^o, \ldots, u_n^o)$ in ordinary n-space to extremize the function of n variables

$$I = I(u_1, u_2, \ldots, u_n)$$

The necessary condition will be found by displacing the point $(u_1^o, u_2^o, \ldots, u_n^o)$ in n-space. That is, let the coordinates u_i be arbitrary functions of a parameter t,

$$u_i = u_i(t) \qquad (i = 1,2,\ldots,n)$$

where $t = 0$ corresponds to the original point; to wit:

$$u_i^o = u_i(0) \quad .$$

Thus, the point is displaced along some curve in n-space; or, in other words, the point is imbedded in a family of points. Certainly, if the point $(u_1^o,u_2^o,\ldots,u_n^o)$ extremizes the function among all points (in a neighborhood), then it also extremizes the function among all points of the family. This means that any necessary condition derived by considering the family of points will also be a necessary condition for the original problem. In other words, it is good enough to find a necessary condition for $t = 0$ to extremize the function I when the coordinates u_i are substituted as functions of t into the function I. That is to say, a necessary condition will be found for the function of one variable t

$$I(t) = I(x_1(t),x_2(t),\ldots,x_n(t))$$

to take an extreme value for $t = 0$. As shown in the theory of differential calculus, a necessary condition is

$$\dot{I} = \frac{d}{dt}I(t) \Big|_{t=0} = 0 \quad .$$

The quantity $\dot{I}$ in this case is given by the chain rule of differentiation as

$$\dot{I} = \sum_{i=1}^{n} \frac{\partial I}{\partial x_i} \frac{dx_i}{dt} \Big|_{t=0}$$

or

$$\dot{I} = V \cdot \text{grad } I \quad ,$$

where the velocity vector V is simply

$$V = \left(\frac{dx_1}{dt},\frac{dx_2}{dt},\ldots,\frac{dx_n}{dt}\right) \Big|_{t=0}$$

Thus, the quantity $\dot{I}$ has been expressed as the inner product of the gradient of I and an arbitrary infinitesimal displacement vector (an infinitesimal displacement vector being essentially the same as a velocity vector). The necessary condition is the vanishing of $\dot{I}$ for arbitrary V, which implies the vanishing of grad I (i.e., the vanishing of all the first partial derivatives of I). But the important point for the present analogy is that $\dot{I}$ is an inner product of grad I and an arbitrary infinitesimal displacement vector.

A necessary condition for a function u(x) to extremize the functional

$$I[\phi] = \int_a^b F(x,\phi(x),\phi'(x))\, dx$$

when substituted for $\phi(x)$ was found on pp. 11-17 of the notes in an exactly analogous manner. In this case, the argument is not a point in n-space depending on a coordinate for each of n-subscripts, but is a function depending on the continuous range of its independent variable. The function u(x) was "displaced" or imbedded in an arbitrary family of admissible functions

$$\phi = \phi(x,t) \quad ,$$

where t = 0 corresponds to the original function; that is

$$u(x) = \phi(x,0) \quad .$$

Certainly, if the function u(x) extremizes the functional among all admissible functions, then it also extremizes the functional among all functions of the family. Thus, a necessary condition was found for t = 0 to extremize the integral

$$I[\phi(x,t)] \equiv I(t) = \int_a^b F(x,\phi(x,t),\phi_x(x,t))\, dx$$

where the integral is simply a function of the one variable t. The necessary condition used was, of course,

$$\dot{I} \equiv \frac{d}{dt} I(t) \Big|_{t=0} = 0$$

The quantity $\dot{I}$ (called the first variation of I and sometimes denoted by $\delta(I)$ was found in this case by differentiating under the integral sign and subsequently integrating by parts, giving:

$$\dot{I} = \int_a^b \zeta(F_u - \frac{d}{dx} F_{u'}) \, dx \tag{17}$$

where the "infinitesimal displacement" ζ is simply

$$\zeta(x) = \phi_t(x,0)$$

The function $\zeta(x)$ (sometimes denoted by δu or $\dot{u}$) corresponds for any fixed value of x to a component of the infinitesimal displacement V for the extremum problem for a function of n variables. That is, ζ or δu or $\dot{u}$ is of the nature of dx (or $\frac{dx}{dt}$). The expression for $\dot{I}$ given by (17) can be rewritten as an inner product

$$\dot{I} = (\zeta, [F]_u)$$

where the Euler expression

$$F_u - \frac{d}{dx} F_{u'}$$

is abbreviated by $[F]_u$ and the inner product of two functions $f(x)$ and $g(x)$ is taken to be

$$(f,g) = \int_a^b f(x)g(x) \, dx \quad .$$

This definition of the inner product is a very natural generalization of the definition of the inner product of two vectors as a sum very similar to the above integral.*

Thus, the first variation $\dot{I}$ has been expressed as an inner product of an arbitrary infinitesimal displacement ζ (arbitrary

---

* For an abstract definition of an inner product over any linear space as a bilinear form, consult any work on linear algebra.

except for the requirement that it vanish at the end-points of the interval) and the Euler expression $[F]_u$. By analogy with the corresponding result for extremizing a function of n variables, it appears that the Euler expression $[F]_u$ behaves like a gradient and might therefore be called the gradient of the functional I; i.e.

$$[F]_u = \text{grad } I \quad .$$

It is therefore to be expected that the Euler expression will possess the property of invariance in the same way that an ordinary gradient vector does and that the invariance will stem from the fact that the inner product of the infinitesimal displacement and the gradient of the functional should be independent of the coordinate system. However, before exhibiting this property of invariance, the concept of the gradient of a functional will be illustrated by two further examples.

In order to find the gradient of a functional, the first variation of the functional has to be expressed as the inner product of an arbitrary infinitesimal displacement and some other expression. Then, this other expression is called the gradient of the functional.

As an example, the gradient will now be calculated for the functional

$$I[u] = \iint_{\square} K(x,y)u(x)u(y)\,dxdy \tag{18}$$

with $K(x,y) = K(y,x)$ where the integration is over the unit square $0 \leq x \leq 1$, $0 \leq y \leq 1$. The function $u(x)$ is defined in the unit interval and function $K(x,y)$ is defined in the unit square. If $u(x)$ is imbedded in the family of functions $\phi(x,t)$ where

$$u(x) = \phi(x,0)$$

it is clear that the first variation of the functional is given by

$$\dot{I} = \iint_{\square} K(x,y)[\dot{u}(x)u(y) + u(x)\dot{u}(y)]\,dxdy$$

where

$$\dot{u}(x) = \frac{d}{dt}\,\phi(x,t)\,\Big|_{t=0} \quad .$$

In view of the symmetry of $K(x,y)$, $\dot{I}$ is then given by

$$\dot{I} = 2\iint_{\square} K(x,y)\dot{u}(x)u(y)\,dxdy$$

or

$$\dot{I} = \int_0^1 \dot{u}(x)\left[2\int_0^1 K(x,y)u(y)\,dy\right]dx$$

or

$$\dot{I} = \left(\dot{u}(x),\ 2\int_0^1 K(x,y)u(y)\,dy\right)$$

Thus

$$\text{grad } I = 2\int_0^1 K(x,y)u(y)\,dy \tag{19}$$

As another example, the gradient will be calculated for the functional

$$I[u] = \iint_{\square} F(x,y,u(x),u(y),u'(x),u'(y))\,dxdy \tag{20}$$

where the integration is over the unit square, where

$$u(0) = u(1) = 0 \quad ,$$

and where F is symmetrical in x and y, in u(x) and u(y), and also in u'(x) and u'(y). Upon imbedding u(x) in a family of functions depending on t, (displacing u(x) into a family of functions), it immediately follows that the first variation is given by

$$\dot{I} = \iint_{\square} \left[F_3\dot{u}(x) + F_4\dot{u}(y) + F_5\frac{d}{dx}\,\dot{u}(x) + F_6\frac{d}{dy}\,\dot{u}(y)\right]dxdy$$

where F_i denotes the partial derivative of F with respect to its i-th argument. In view of the symmetries of F, $\dot{I}$ is given by

$$\dot{I} = 2\iint_{\square} \left[F_3\dot{u}(x) + F_5\frac{d}{dx}\,\dot{u}(x)\right]dxdy$$

or, upon integration by parts

$$\dot{I} = 2\iint_{\square} \dot{u}(x)[F_3 - \frac{\partial}{\partial x} F_5]\, dxdy$$

or

$$\dot{I} = \int_0^1 \dot{u}(x)\, [2\int_0^1 (F_3 - \frac{\partial}{\partial x} F_5)dy]dx$$

or

$$\dot{I} = (\dot{u}(x),\ 2\int_0^1 (F_3 - \frac{\partial}{\partial x} F_5)dy)$$

Thus

$$\text{grad } I = 2\int_0^1 (F_3 - \frac{\partial}{\partial x} F_5)dy \tag{21}$$

It might be remarked that the functional in (18) is a special case of the functional in (20), and the gradient in (21), therefore, reduces to the gradient in (19) upon setting

$$F_3 = K(x,y)u(y)$$
$$F_5 = 0$$

Thus, illustrations have been given to clarify what is meant by calling the Euler expression $[F]_u$ the gradient of its associated functional, and it has been indicated that gradient-like behavior ought to imply some kind of invariance. In short, it would seem that the stage has been set for the demonstration of the invariance of the Euler expression. However, just one more preliminary discussion will be made before showing this invariance.

In order to prepare for displaying the invariant nature of the Euler expression, a few remarks about notation for transformed functions will now be made. Specifically, the notation to be used for a function $a(x)$ after it has undergone a transformation

$$x = f(\xi)$$

will be discussed. The transformed function $\alpha(f(\xi))$ is sometimes denoted simply by $\alpha(\xi)$ and is sometimes denoted by some new name, say $\beta(\xi)$ or $\alpha^*(\xi)$. Which of these two notations is to be used depends on the point-of-view of the particular discussion in the following way:

If the viewpoint to be emphasized is the dependency of some quantity on other quantities rather than functional relationships in the mathematical sense; i.e., if α is not regarded as the name of a certain function but rather the name of a certain quantity, then writing $\alpha(x)$ expresses the fact that the quantity $\alpha(x)$ depends on the quantity x. In this case, the trasformation means that the quantity x depends on the quantity ξ, and, therefore, the quantity α depends on the quantity ξ. This fact can be expressed by writing the symbol $\alpha(\xi)$, which simply means that the quantity α depends on the quantity ξ.

On the other hand, if functional connections are being stressed, then a new name must be used to denote the transformed function. That is to say, if $\alpha(x)$ is regarded as denoting a functional relationship in the mathematical sense (such as $2x + \sin x$), then $\alpha(f(\xi))$ (i.e., $2f(\xi) + \sin f(\xi)$) is a function of ξ which cannot be denoted by $\alpha(\xi)$ (i.e., $2\xi + \sin \xi$) but must be denoted by some new name, say $\alpha^*(\xi)$.

For example, if the quantity called the volume V of a certain system should depend on the temperature Θ, a physicist interested primarily in quantities might write

$$V = V(\Theta) \quad .$$

And, if the temperature Θ should in turn depend on the time t, which means that the volume depends on the time t, this dependency might be expressed by

$$V = V(t) \quad .$$

Here, of course, V really stands for a quantity, the volume, and not for a functional relationship. To be more precise and

emphasize the functional relationships involved, the same physical situation might be described mathematically by

$$V = q(\Theta)$$

and

$$\Theta = r(t)$$

and, therefore

$$V = q(r(t)) = q^*(t)$$

The notation emphasizing quantities, which is widely used by physicists, is rather simple and is good enough for many purposes. However, this simplified notation leads to a difficulty, for example, in writing symbols for partial derivatives. As an illustration, let U denote the quantity called the internal energy of, say, a chemical system. The state of such a system is characterized by the three thermodynamic coordinates P, V, and Θ, where P represents pressure, V represents volume, and Θ represents (absolute) temperature. However, only two of these thermodynamic coordinates are independent since the three are related (for states of equilibrium) by the equation of state which every chemical system possesses. For example, the equation of state for an ideal gas is essentially

(22) $$V = \frac{\Theta}{P}$$.

Thus, the internal energy U may be regarded as depending on any two of the quantities V, Θ, and P. In other words, for any particular chemical system

$$U = U(V,\Theta)$$

expresses the fact that U depends on V and Θ, while

$$U = U(P,\Theta)$$

expresses the fact that U can also be regarded as depending on P and Θ. This is quite simple, but the difficulty arises when the partial derivatives of U are considered. The partial

derivative of U with respect to Θ when U is regarded as a function of V and Θ and the partial derivative of U with respect to Θ when U is regarded as a function of P and Θ are partial derivatives of different functions and require different symbols. Physicists denote the former partial derivative by $(\frac{\partial U}{\partial \Theta})_V$ and the latter by $(\frac{\partial U}{\partial \Theta})_P$. The partial derivative $(\frac{\partial U}{\partial \Theta})_V$ is called the partial derivative of U with respect to Θ "when V is held constant", and $(\frac{\partial U}{\partial \Theta})_P$ is called the partial derivative of U with respect to Θ "when P is held constant". It is very important to realize that the two partial derivatives are in general quite different. It should be understood as stated for example, on page 57 of Heat and Thermodynamics by M.W. Zemansky that "the two partial derivatives $(\frac{\partial U}{\partial \Theta})_V$ and $(\frac{\partial U}{\partial \Theta})_P$ are not equal. The first is a function of Θ and V, and the second a function of Θ and P. They are different mathematically and also have a different physical meaning". Perhaps it can be seen most directly that the two partial derivatives are really different by considering a simple hypothetical example. For instance, let U be given by

$$U = V + \Theta$$

where V, Θ, and P are related according to (22). Then, U is also given by

$$U = \frac{\Theta}{P} + \Theta$$

Then, the two partial derivatives are given by

$$(\frac{\partial U}{\partial \Theta})_V = 1$$

$$(\frac{\partial U}{\partial \Theta})_P = \frac{1}{P} + 1 \quad ,$$

The entire difficulty with the notation for partial derivatives completely disappears, however, if the more precise notation is used:

$$U = f(V,\Theta)$$

$$U = f(K(P,\Theta) = g(P,\Theta)$$

where

$$V = K(P,Q) \quad .$$

In this case, the two partial derivatives are simply $\frac{\partial f}{\partial Q}$ and $\frac{\partial g}{\partial Q}$.

Thus, two notations are available for transformed functions. Which notation to use is governed by the nature of the particular discussion. It is frequently desirable when the meaning is clear to use the simplified notation in order to keep the notation from getting unnecessarily complicated and pedantic. On the other hand, the simplified notation in some cases may lead to confusion and it would then pay to denote transformed functions by new names.

The above discussion of notation constitutes the last preliminary step, and it is now time for the coup de grace, the demonstration of the invariance (more precisely, covariance) of the Euler expression. In other words, it will now be shown that the Euler expression undergoes all transformations according to one and the same rule. This rule will be obtained by considering the most general change of independent variable in the integral

$$I[u] = \int_a^b F(x,u,u')\, dx \quad ,$$

the first variation of which of course, is

$$\dot{I} = \int_a^b \dot{u}[F]_{u(x)}\, dx \tag{23}$$

where

$$[F]_{u(x)} = F_u - \frac{d}{dx} F_{u'} \quad .$$

Let the transformation or change of variable be given by

$$x = f(\xi) \tag{24}$$

with an inverse everywhere in the interval $a \leq x \leq b$, and let a^* and b^* be the numbers for which

$$a = f(a^*)$$
$$b = f(b^*)$$

Using the simplified notation, let the function $u(x)$ after the transformation be denoted by $u(\xi)$ and let $\frac{d}{d\xi}\, u(\xi)$ be denoted simply by $u'(\xi)$. But, to be slightly more precise for the function F, let the function $F(x,u,u')$ after the transformation be denoted by $G(\xi,u,u')$; i.e., let

$$F(f(\xi),u(\xi),\tfrac{1}{f'}\, u'(\xi)) = G(\xi,u(\xi),u'(\xi)) \quad .$$

Thus, the transformed functional is

$$I[u] = \int_{a^*}^{b^*} G(\xi,u,u')f'd\xi$$

and its first variation is

$$\dot{I} = \int_{a^*}^{b^*} \dot{u}[Gf']_{u(\xi)}\, d\xi \tag{25}$$

where

$$[Gf']_{u(\xi)} \equiv G_u f' - \frac{d}{d\xi}\,(G_{u'}f') \quad .$$

Comparison of (23) and (25) implies that

$$\int_a^b \dot{u}[F]_{u(x)}\, dx \equiv \int_{a^*}^{b^*} \dot{u}[Gf']_{u(\xi)}\, d\xi$$

for arbitrary $\dot{u}$; or making a change of variable in the integral on the right in order to have the same limits of integration for both integrals, it follows that

$$\int_a^b \dot{u}[F]_{u(x)}\, dx \equiv \int_a^b \dot{u}[Gf']_{u(\xi)}\, \tfrac{1}{f'}\, dx$$

for arbitrary $\dot{u}$. Since $\dot{u}$ is arbitrary, the following identity is implied:

$$[F]_{u(x)} = [Gf']_{u(\xi)}\, \tfrac{1}{f'} \quad .$$

or

$$f'[F]_{u(x)} = [Gf']_{u(\xi)} \quad .$$

And this identity expresses the invariance of the Euler expression or, in particular, the independence from choice of coordinate system of the inner product of $\dot{u}$ and $[F]_u$. This is the rule according to which the Euler expression undergoes all transformations.

In the derivation given above for the property of invariance, the simplified notation made it possible to avoid the writing of unduly complicated expressions, but at the same time certain steps in the reasoning might not be quite clear at first. For example, the comparison of (23) and (25) might be questioned somewhat. Therefore, in order to show that the entire derivation can be justified step by step, the same derivation will now be repeated without using the simplified notation; i.e., transformed functions will be given new names. Once this has been done, a return can be made to the simplified notation since any possible misunderstandings will have been cleared up.

In order to have a somewhat uniform notation for this derivation, the following scheme will be followed in naming functions: If a function $m(x)$ is transformed according to (24), the transformed function will be denoted by $m^*(\xi)$; and, if a function $n(\xi)$ is transformed according to the inverse of (24), then the transformed function will be called $n^\circ(x)$.

Following the above convention, let the function $u(x)$ after the transformation be denoted by $u^*(\xi)$; i.e., let

$$u(f(\xi)) = u^*(\xi)$$

and let the function $F(x,u,u')$ after the transformation be denoted by $G(\xi,u^*,u^{*\prime})$; i.e., let

$$F(f(\xi),u^*(\xi),\frac{1}{f'(\xi)}\frac{du^*}{d\xi}) = G(\xi,u^*,\frac{du^*}{d\xi}) \quad .$$

Thus, the transformed functional I^* is given by

$$I^*[u^*] = \int_{a^*}^{b^*} G(\xi,u^*,u^{*\prime})f'(\xi)\,d\xi \quad .$$

The first variation of this functional is, of course,

$$\dot{I}^* = \int_{a^*}^{b^*} \dot{u}^*(\xi)[Gf']_{u^*(\xi)}\, d\xi \tag{26}$$

where

$$[Gf']_{u^*(\xi)} \equiv G_{u^*}(\xi,u^*,u^{*\prime})f'(\xi) - \frac{d}{d\xi}\left[G_{u^{*\prime}}(\xi,u^*,u^{*\prime})f'(\xi)\right] .$$

However, in order to compare $\dot{I}^*$ with $\dot{I}$, x will be introduced as the variable of integration in $\dot{I}^*$, giving

$$\dot{I}^* = \int_a^b \dot{u}(x)\left\{[Gf']_{u^*(\xi)}\right\}^o \frac{1}{f'^o}\, dx \tag{27}$$

where the meaning of $\left\{[Gf']_{u^*(\xi)}\right\}^o$ is as follows: The expression $[Gf']_{u^*(\xi)}$ is some function of ξ, say

$$[Gf']_{u^*(\xi)} = \psi(\xi) .$$

Then, $\left\{[Gf']_{u^*(\xi)}\right\}^o$ is given by

$$\left\{[Gf']_{u^*(\xi)}\right\}^o = \psi(f^{-1}(x))$$

where f^{-1} is the inverse function of f. Similarly, the function f'^o is to be taken as

$$f'^o = f'(f^{-1}(x)) .$$

Now, for the purpose of comparing $\dot{I}$ and $\dot{I}^*$, I[u] and I*[u*] will be compared first. It is clear that

$$I[u] \equiv I^*[u^*]$$

is an identity with respect to the function u since making a change of variable does not affect the value of an integral. Thus, if u is imbedded in the family of functions $\phi(x,t)$ and u* is imbedded in the family $\phi^*(\xi,t)$ where

$$\phi^*(\xi,t) \equiv \phi(f(\xi),t) ,$$

it follows that

$$I[\phi(x,t)] \equiv I^*[\phi^*(\xi,t)]$$

and

$$\frac{d}{dt} I[\phi(x,t)]\Big|^{t=0} \equiv \frac{d}{dt} I^*[\phi^*(\xi,t)]\Big|^{t=0}$$

or, in other words,

$$\dot{I} \equiv \dot{I}^* \quad .$$

Using this last identity and expressing $\dot{I}$ and $\dot{I}^*$ as given in (23) and (27), it follows that

$$\int_a^b \dot{u}(x)[F]_{u(x)}dx \equiv \int_a^b \dot{u}(x)\left\{[Gf']_{u^*(\xi)}\right\}^0 \frac{1}{f'^0}\,dx$$

for arbitrary $\dot{u}(x)$ (vanishing, of course, at the end-points). In view of the arbitrariness of $\dot{u}(x)$, it follows that

$$[F]_{u(x)} = \left\{[Gf']_{u^*(\xi)}\right\}^0 \frac{1}{f'^0}$$

or

$$f'^0[F]_{u(x)} = \left\{[Gf']_{u^*(\xi)}\right\}^0 \tag{28}$$

where both sides are functions of x, or

$$f'\left\{[F]_{u(x)}\right\}^* = [Gf']_{u^*(\xi)} \tag{29}$$

where both sides are functions of ξ. In other words, (29) states that calculating the Euler expression of F and then changing to new coordinates and then multiplying by the factor f' gives the same result as first changing F to new coordinates and then multiplying by the factor f' and then finding the Euler expression.

However, the notation used above is really unnecessarily pedantic for the simple situation at hand. The function $u^*(\xi)$ might simply be called $u(\xi)$, and similarly the other superscripts of stars (*) and naughts (^{0}) might be omitted. Thus, the relation (28) or (29) showing the invariance of the Euler expression can be written in the same form as previously found, as follows:

$$f'[F]_{u(x)} = [Gf']_{u(\xi)} \quad .$$

This completes the second and more detailed derivation of the rule according to which the Euler expression undergoes transformations of coordinates; i.e., the rule which expresses the invariance of the Euler expression.

Thus, it appears that the formalism of the Calculus of Variations insofar as the Euler expression is concerned is the same in all coordinate systems except for the necessity of using as a factor the derivative f' of the transformation relating any two particular coordinate systems.

The analogous rule of transformation for the Euler expression when u is a function of two (or several) variables will now be found. In view of the above detailed discussion, it will be assumed that the meaning is clear when the simplified notation is used. The functional to be considered is

$$I[u] = \iint_R F(x,y,u,p,q)\, dxdy$$

where

$$p = u_x$$
$$q = u_y$$

and R is a region of the xy-plane. The first variation is, of course

$$\dot{I} = \iint_R \dot{u}[F]_{u(x,y)}\, dxdy \tag{30}$$

where

$$[F]_{u(x,y)} = F_u - \frac{\partial}{\partial x} F_p - \frac{\partial}{\partial y} F_q$$

Let the transformation to be made be given by

$$x = f(\xi,\eta)$$
$$y = g(\xi,\eta)$$

with an inverse everywhere in the region R of the xy-plane, and let R* be the region of the $\xi\eta$-plane which is mapped onto the

region R of the xy-plane. Also, let the function $F(x,y,u,p,q)$ after the transformation be denoted by $G(\xi,\eta,u,\pi,\kappa)$ where

$$\pi = u_\xi$$
$$\kappa = u_\eta$$

The transformed functional is then

$$I[u] = \iint_{R^*} G(\xi,\eta,u,\pi,\kappa)\, j \, d\xi d\eta$$

where j represents the Jacobian of the transformation; that is

$$j = \frac{\partial(x,y)}{\partial(\xi,\eta)} \quad .$$

The first variation is clearly given by

$$\dot{I} = \iint_{R^*} \dot{u}[Gj]_{u(\xi,\eta)} \, d\xi d\eta$$

or

$$\dot{I} = \iint_{R} \dot{u}[Gj]_{u(\xi,\eta)} \, \frac{1}{j} \, dxdy \tag{31}$$

where

$$[Gj]_{u(\xi,\eta)} = G_u j - \frac{\partial}{\partial\xi}(G_p j) - \frac{\partial}{\partial\eta}(G_q j) \quad .$$

In view of the arbitrariness of $\dot{u}$, comparison of (30) and (31) now yields the identity:

$$j[F]_{u(x,y)} = [Gj]_{u(\xi,\eta)} \quad .$$

This identity, of course, expresses the invariance of the Euler expression, and it once again appears that the formalism is the same for all coordinate systems except for the Jacobian j in this case.

It might further be noted that if u is a function of n variables $(x_1,x_2,\ldots,x_n)$, reasoning exactly parallel to that used above shows that the Euler expression transformes according to

$$j[F]_{u(x_1,x_2,\ldots,x_n)} = [Gj]_{u(\xi_1,\xi_2,\ldots,\xi_n)} \quad .$$

A very elementary application of the invariance of the Euler expression is its use in transforming the Laplacian

$$\triangle u = u_{xx} + u_{yy}$$

to polar coordinates (r,θ) given by

$$\begin{aligned} x &= r \cos \theta \\ y &= r \sin \theta \end{aligned} \quad .$$

The transformation can, of course, be made directly, but the transforming of the second derivatives requires considerable calculation. With the help of Euler's expression, however, the result will now be obtained by transforming only first derivatives. The main point is to observe that the Laplacian is essentially the Euler expression for the Dirichlet integral (as indicated on pp. 23-24 of the notes). In other words, if the function F is given by

$$-2F = u_x^2 + u_y^2 \quad ,$$

then the Euler expression is simply

$$[F]_{u(x,y)} = \triangle u \quad . \tag{32}$$

In order to take advantage of the invariance of the Euler expression, the Jacobian j is needed and can be seen to be given by

$$j \equiv \frac{\partial(x,y)}{\partial(r,\theta)} = r \quad .$$

In order to find the transformed integrand function G, u_x and u_y will have to be transformed to polar coordinates, as follows:

$$\begin{aligned} u_x &= u_r r_x + u_\theta \theta_x \\ u_y &= u_r r_y + u_\theta \theta_y \end{aligned}$$

where

$$r_x = \frac{y_\theta}{j} = \cos \theta$$

$$\theta_x = -\frac{y_r}{j} = -\frac{\sin\theta}{r}$$

$$r_y = -\frac{x_\theta}{j} = \sin\theta$$

$$\theta_y = \frac{x_r}{j} = \frac{\cos\theta}{r} \quad .$$

Thus, u_x and u_y are given by

$$u_x = u_r \cos\theta - u_\theta \frac{\sin\theta}{r}$$

$$u_y = u_r \sin\theta + u_\theta \frac{\cos\theta}{r} \quad ,$$

G is given by

$$-2G = u_x^2 + u_y^2 = u_r^2 + \frac{1}{r^2} u_\theta^2 \quad ,$$

and Gj is given by

$$-2Gj = ru_r^2 + \frac{1}{r} u_\theta^2 \quad .$$

The Euler expression for Gj is then

$$[Gj]_{u(r,\theta)} = \frac{\partial}{\partial r}(ru_r) + \frac{1}{r} u_{\theta\theta} \quad . \tag{33}$$

Substitution of (32) and (33) into

$$j[F]_{u(x,y)} = [Gj]_{u(r,\theta)}$$

gives

$$r \triangle u = \frac{\partial}{\partial r}(ru_r) + \frac{1}{r} u_{\theta\theta}$$

and, therefore

$$\triangle u = \frac{1}{r}\frac{\partial}{\partial r}(ru_r) + \frac{1}{r^2} u_{\theta\theta} \quad .$$

Thus, the Laplacian has been transformed to polar coordinates simply by transforming first derivatives and using the invariance of the Euler expression.

As a further application of the invariance of the Euler expression, the Laplacian will now be transformed to 3-dimensional polar coordinates (spherical coordinates) and also to n-dimensional polar coordinates. However, the transformation will

be accomplished in an even more efficient manner than was used above for 2-dimensional polar coordinates. The method used above would be too complicated for the following reason: In order to transform $u_x^2 + u_y^2$ to polar coordinates, u_x was so transformed and also u_y was transformed and then $u_x^2 + u_y^2$ was calculated. If the attempt is made to carry out the analogous computation for more than two dimensions, the amount of calculation involved is almost prohibitive. Therefore, before transforming the Laplacian in 3 or n dimensions, a method will first be developed for transforming the sums of the squares of the first partial derivatives of u into any curvilinear coordinate system without explicitly transforming any of the partial derivatives themselves.

For this purpose, let the rectangular Cartesian coordinates be $(x_1, x_2, \ldots, x_n)$ and let the curvilinear coordinates $(\xi_1, \xi_2, \ldots, \xi_n)$ be given by

$$x_i = f_i(\xi_1, \xi_2, \ldots, \xi_n) \qquad (i = 1,2,\ldots,n)$$

where the transformation is assumed to have an inverse. Also, let the matrix of the Jacobian j be denoted by J; i.e.

$$J = \begin{pmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_2}{\partial \xi_1} & \cdots & \frac{\partial x_n}{\partial \xi_1} \\ \frac{\partial x_1}{\partial \xi_2} & \frac{\partial x_2}{\partial \xi_2} & \cdots & \frac{\partial x_n}{\partial \xi_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_1}{\partial \xi_n} & \frac{\partial x_2}{\partial \xi_n} & \cdots & \frac{\partial x_n}{\partial \xi_n} \end{pmatrix}$$

where

$$\det J = j \quad .$$

The chain rule of differentiation gives

$$\frac{\partial u}{\partial \xi_i} = \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial \xi_i} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial \xi_i} + \ldots + \frac{\partial u}{\partial x_n}\frac{\partial x_n}{\partial \xi_i} \qquad (i = 1,2,\ldots,n)$$

or, in matrix notation,

$$\begin{pmatrix} \frac{\partial u}{\partial \xi_1} \\ \frac{\partial u}{\partial \xi_2} \\ \vdots \\ \frac{\partial u}{\partial \xi_n} \end{pmatrix} = J \begin{pmatrix} \frac{\partial u}{\partial x_1} \\ \frac{\partial u}{\partial x_2} \\ \vdots \\ \frac{\partial u}{\partial x_n} \end{pmatrix}$$

Since it has been assumed that the transformation has an inverse, the matrix J is non-singular and has an inverse J^{-1}, so that

$$\begin{pmatrix} \frac{\partial u}{\partial x_1} \\ \frac{\partial u}{\partial x_2} \\ \vdots \\ \frac{\partial u}{\partial x_n} \end{pmatrix} = J^{-1} \begin{pmatrix} \frac{\partial u}{\partial \xi_1} \\ \frac{\partial u}{\partial \xi_2} \\ \vdots \\ \frac{\partial u}{\partial \xi_n} \end{pmatrix} \tag{34}$$

If the transpose of each side is taken and the transpose of J^{-1} is denoted by $(J^{-1})'$, it follows that

$$(\frac{\partial u}{\partial x_1} \frac{\partial u}{\partial x_2} \cdots \frac{\partial u}{\partial x_n}) = (\frac{\partial u}{\partial \xi_1} \frac{\partial u}{\partial \xi_2} \cdots \frac{\partial u}{\partial \xi_n})(J^{-1})' \quad . \tag{35}$$

Multiplying (34) on the left by (35) gives

$$(\frac{\partial u}{\partial x_1} \frac{\partial u}{\partial x_2} \cdots \frac{\partial u}{\partial x_n}) \begin{pmatrix} \frac{\partial u}{\partial x_1} \\ \frac{\partial u}{\partial x_2} \\ \vdots \\ \frac{\partial u}{\partial x_n} \end{pmatrix} = (\frac{\partial u}{\partial \xi_1} \frac{\partial u}{\partial \xi_2} \cdots \frac{\partial u}{\partial \xi_n})(J^{-1})' J^{-1} \begin{pmatrix} \frac{\partial u}{\partial \xi_1} \\ \frac{\partial u}{\partial \xi_2} \\ \vdots \\ \frac{\partial u}{\partial \xi_n} \end{pmatrix} . \tag{36}$$

However,

$$\left(\frac{\partial u}{\partial x_1}\ \frac{\partial u}{\partial x_2}\ \cdots\ \frac{\partial u}{\partial x_n}\right)\begin{pmatrix}\frac{\partial u}{\partial x_1}\\ \frac{\partial u}{\partial x_2}\\ \vdots\\ \frac{\partial u}{\partial x_n}\end{pmatrix} = \sum_{i=1}^{n}\left(\frac{\partial u}{\partial x_i}\right)^2$$

and

$$(J^{-1})'J^{-1} = (J')^{-1}J^{-1} = (JJ')^{-1} \quad .$$

Thus, (36) can be written in the form

$$(37)\qquad \sum_{i=1}^{n}\left(\frac{\partial u}{\partial x_i}\right)^2 = \left(\frac{\partial u}{\partial \xi_1}\ \frac{\partial u}{\partial \xi_2}\ \cdots\ \frac{\partial u}{\partial \xi_n}\right)(JJ')^{-1}\begin{pmatrix}\frac{\partial u}{\partial \xi_1}\\ \frac{\partial u}{\partial \xi_2}\\ \vdots\\ \frac{\partial u}{\partial \xi}\end{pmatrix}$$

or, for 3-dimensional polar coordinates (r,ϕ,θ):

$$(38)\qquad u_x^2 + u_y^2 + u_z^2 = (u_r\ u_\phi\ u_\theta)(JJ')^{-1}\begin{pmatrix}u_r\\ u_\phi\\ u_\theta\end{pmatrix} \quad .$$

Thus, the sum of the squares of the partial derivatives of u can be calculated simply according to (37) (or (38)). Also, the absolute value of the Jacobian j can be calculated from the matrix JJ' according to

$$j = \det J = \pm\sqrt{\det (JJ')}$$

that is,

$$(39)\qquad |j| = \left|\sqrt{\det (JJ')}\right| \quad ,$$

where the vertical bars denote absolute value. (The reason that it is preferable to have the computation depend as much as possible on JJ' rather than on J is that JJ' turns out to be a much simpler matrix in significant examples.)

Now that the above rules have been prepared, the 3-dimensional Laplacian

$$\triangle u \equiv u_{xx} + u_{yy} + u_{zz}$$

will be transformed to polar coordinates (r,ϕ,θ) where

$$\begin{aligned} x &= r \cos \phi \sin \theta \\ y &= r \sin \phi \sin \theta \\ z &= r \cos \theta \end{aligned}$$

In this case, if the integrand function F for a variational problem is given by

$$-2F = u_x^2 + u_y^2 + u_z^2$$

then the Euler expression is simply

$$[F]_{u(x,y,z)} = \triangle u$$

In order to transform F according to (38), the matrix $(JJ')^{-1}$ will be calculated. In this case, the matrix J is given by

$$J = \begin{pmatrix} \cos \phi \sin \theta & \sin \phi \cos \theta & \cos \theta \\ -r \sin \phi \sin \theta & r \cos \phi \sin \theta & 0 \\ r \cos \phi \cos \theta & r \sin \phi \cos \theta & -r \sin \theta \end{pmatrix}$$

and, it is a simple matter to see that the product of J and its transpose J' is given by

$$JJ' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 \sin^2 \theta & 0 \\ 0 & 0 & r^2 \end{pmatrix}$$

and, therefore

$$(JJ')^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2 \sin^2 \theta} & 0 \\ 0 & 0 & \frac{1}{r^2} \end{pmatrix}$$

Then, according to (38), it follows that

$$u_x^2 + u_y^2 + u_z^2 = u_r^2 + \frac{1}{r^2\sin^2\theta}\, u_\phi^2 + \frac{1}{r^2}\, u_\theta^2$$

and that the transformed integrand function G is given by

$$-2G = u_r^2 + \frac{1}{r^2\sin^2\theta}\, u_\phi^2 + \frac{1}{r^2}\, u_\theta^2 \quad .$$

The Jacobian according to (39) is simply given by

$$|j| = |\sqrt{\det(JJ^T)}| = |r^2\sin\theta|$$

where there is no need to determine whether the plus or minus sign is to be used when the absolute bars are removed since replacing j by -j causes no change in

$$(40) \qquad [F]_{u(x,y,z)} = \frac{1}{j}\,[Gj]_{u(r,\phi,\theta)} \quad .$$

It is then clear that Gj is given (except for a possible minus sign) by

$$-2Gj = r^2\sin\theta\, u_r^2 + \frac{u_\phi^2}{\sin\theta} + \sin\theta\, u_\theta^2$$

and

$$[Gj]_{u(r,\phi,\theta)} = \frac{\partial}{\partial r}(r^2\sin\theta\, u_r) + \frac{\partial}{\partial\phi}\left(\frac{u_\phi}{\sin\theta}\right) + \frac{\partial}{\partial\theta}(\sin\theta\, u_\theta).$$

Substitution into (40) immediately gives the result

$$\triangle u = \frac{1}{r^2\sin\theta}\left\{\frac{\partial}{\partial r}(r^2\sin\theta\, u_r) + \frac{\partial}{\partial\phi}\left(\frac{u_\phi}{\sin\theta}\right) + \frac{\partial}{\partial\theta}(\sin\theta\, u_\theta)\right\} .$$

The Laplacian in n dimensions will now be transformed to polar coordinates in the same way. As for the definition of n-dimensional polar coordinates, they will be defined recursively by generalizing (n-1)-dimensional polar coordinates in the same way that 3-dimensional polar coordinates are obtained by generalizing 2-dimensional polar coordinates. It would, therefore, be natural first of all to examine the way in which 2-dimensional polar coordinates are generalized to 3-dimensional polar coordinates. For this purpose, let n-dimensional rectangular

Cartesian coordinates be denoted by $(x_1^{(n)}, x_2^{(n)}, \ldots, x_n^{(n)})$ and let n-dimensional polar coordinates be denoted by $(r, \theta_2, \theta_3, \ldots, \theta_n)$ (where the first polar coordinate has been called r rather than θ_1 because it plays a geometric role different from that of the other polar coordinates; namely, r represents a distance and the θ's represent angles). The equations defining 2-dimensional polar coordinates are then

$$x_1^{(2)} = r \cos \theta_2$$
$$x_2^{(2)} = r \sin \theta_2$$

where θ_2 is usually called θ. Similarly, the equations defining 3-dimensional polar coordinates are

$$x_1^{(3)} = r \cos \theta_2 \sin \theta_3$$
$$x_2^{(3)} = r \sin \theta_2 \sin \theta_3$$
$$x_3^{(3)} = r \cos \theta_3$$

where θ_2 is usually called ϕ and θ_3 is usually called θ. It is clear when the equations are written in this way that two principles are adhered to in the generalization. First of all, the geometric meaning of r as the distance from the origin is retained. And secondly, the coordinates are defined so that the 3-dimensional polar coordinates applied to the 2-dimensional subspace orthogonal to the x_3-axis at the origin should be identical with the 2-dimensional polar coordinates in this subspace. This is clearly seen by setting θ_3 equal to $\pi/2$. In order to follow the same two principles, n-dimensional polar coordinates will be defined in general for $n > 2$ as follows:

$$x_i^{(n)} = x_i^{(n-1)} \sin \theta_n \qquad (i = 1,2,\ldots,n-1)$$
$$x_n^{(n)} = r \cos \theta_n$$

Here, it can easily be seen that r represents distance from the origin; and it is also clear that setting θ_n equal to $\pi/2$ yields

(n-1)-dimensional polar coordinates; i.e., the n-dimensional polar coordinates reduce to (n-1)-dimensional polar coordinates for the (n-1)-dimensional subspace orthogonal to the x_n-axis.

Now that n-dimensional polar coordinates have been defined, a return will be made to the original purpose of transforming the Laplacian. If the integrand function F is given by

$$-2F = \sum_{i=1}^{n} \left(\frac{\partial u}{\partial x_i}\right)^2$$

then the Euler expression is given by

$$[F]_{u(x)} = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} \equiv \Delta u$$

Let the Jacobian for the n-dimensional case be denoted by $J^{(n)}$ and its matrix by $J^{(n)}$. Then, in order to use (37), $(J^{(n)}J^{(n)'})$ will have to be calculated. Let it be observed first that the matrix $J^{(n)}$ for $n > 2$ is obtained by adding one row and one column to the matrix $J^{(n-1)}$ sin θ_n as follows:

$$(41) \quad J^{(n)} = \begin{pmatrix} & & & \cos\theta_n \\ & J^{(n-1)} \sin\theta_n & & 0 \\ & & & 0 \\ & & & \vdots \\ x_1^{(n-1)}\cos\theta_n & x_2^{(n-1)}\cos\theta_n & \cdots\ x^{(n-1)}\cos\theta_n & -r\sin\theta_n \end{pmatrix}$$

It can be seen rather easily that the product of this matrix and its transpose is a matrix in diagonal form if the matrix $J^{(n-1)}J^{(n-1)'}$ is in diagonal form.* However, it has already appeared that this property is enjoyed by the matrix $J^{(3)}J^{(3)'}$, and therefore $J^{(n)}J^{(n)'}$ is given by

* Actually, saying that $J^{(n)}J^{(n)'}$ is in diagonal form amounts to saying that n-dimensional polar coordinates constitute an orthogonal coordinate system.

$$J^{(n)}J^{(n)\prime} = \begin{pmatrix} g_{11}^{(n)} & 0 & \dots & 0 \\ 0 & g_{22}^{(n)} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & g_{nn}^{(n)} \end{pmatrix}$$

where the $g_{ii}^{(n)}$ have to be determined. However, in view of (41), it is clear that

$$g_{11}^{(n)} = 1$$

$$g_{ii}^{(n)} = g_{ii}^{(n-1)}\sin^2\theta_n \qquad (i \neq 1,n)$$

$$g_{nn}^{(n)} = r^2$$

and in view of the 3-dimensional case already considered, the above recursion formula for $g_{ii}^{(n)}$ for $i \neq 1,n$ gives

$$g_{ii}^{(n)} = g_{ii}^{(n-1)}\sin^2\theta_n = r^2 \prod_{\nu=i+1}^{n} \sin^2\theta_\nu \qquad (i \neq i,n)$$

Clearly, $(J^{(n)}J^{(n)\prime})^{-1}$ is given by

$$(J^{(n)}J^{(n)\prime})^{-1} = \begin{pmatrix} g_{(n)}^{11} & 0 & \dots & 0 \\ 0 & g_{(n)}^{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & g_{(n)}^{nn} \end{pmatrix}$$

where

$$g_{(n)}^{ii} = \frac{1}{g_{ii}^{(n)}} \qquad (i = 1,2,\dots,n) \quad .$$

The Jacobian $j^{(n)}$ is given (except possibly for an unessential minus sign) by

$$j^{(n)} = \sqrt{\det\,(J^{(n)}J^{(n)\prime})} = \sqrt{r^{2(n-1)} \prod_{i=2}^{n-1} \left(\prod_{\nu=i+1}^{n} \sin^2\theta_\nu \right)}$$

$$= r^{n-1} \prod_{i=2}^{n-1} \left(\prod_{\nu=i+1}^{n} \sin\theta_\nu \right)$$

or, if the products are combined,

$$j^{(n)} = r^{n-1} \prod_{\nu=3}^{n} \sin^{\nu-2} \theta_\nu \quad .$$

The transformed integrand function G is given according to (37) by

$$-2G = u_r^2 + \sum_{i=2}^{n-1} \frac{(\frac{\partial u}{\partial \theta_i})^2}{r^2 \prod_{\nu=i+1}^{n} \sin^2\theta_\nu} + \frac{1}{r^2} (\frac{\partial u}{\partial \theta_n})^2$$

and $Gj^{(n)}$ is given by

$$-2Gj^{(n)} = \left[u_r^2 + \sum_{i=2}^{n-1} \frac{(\frac{\partial u}{\partial \theta_i})^2}{r^2 \prod_{\nu=i+1}^{n} \sin^2\theta_\nu} + \frac{1}{r^2} (\frac{\partial u}{\partial \theta_n})^2 \right] r^{n-1} \prod_{\nu=3}^{n} \sin^{\nu-2}\theta_\nu .$$

The Euler expression for Gj is clearly given by

$$[Gj^{(n)}]_{u(r,\theta)} = \frac{\partial}{\partial r}(u_r r^{n-1} \prod_{\nu=3}^{n} \sin^{\nu-2}\theta_\nu) + \sum_{i=2}^{n-1} \frac{\partial}{\partial \theta_i} \left(\frac{\frac{\partial u}{\partial \theta_i} r^{n-3} \prod_{\nu=3}^{n} \sin^{\nu-2}\theta_\nu}{\prod_{\nu=i+1}^{n} \sin^2\theta_\nu} \right)$$

$$+ \frac{\partial}{\partial \theta_n} (\frac{\partial u}{\partial \theta_n} r^{n-3} \prod_{\nu=3}^{n} \sin^{\nu-2} \theta_\nu) .$$

Substituting into

$$[F]_{u(x)} = \frac{1}{j^{(n)}} [Gj^{(n)}]_{u(r,\theta)}$$

yields for $n > 2$:

$$\Delta u = \frac{1}{r^{n-1} \prod_{\nu=3}^{n} \sin^{\nu-2}\theta_\nu} \left\{ \frac{\partial}{\partial r} (u_r r^{n-1} \prod_{\nu=3}^{n} \sin^{\nu-2}\theta_\nu) \right.$$

$$\left. + \sum_{i=2}^{n-1} \frac{\partial}{\partial \theta_i} \left(\frac{\frac{\partial u}{\partial \theta_i} r^{n-3} \prod_{\nu=3}^{n} \sin^{\nu-2}\theta_\nu}{\prod_{\nu=i+1}^{n} \sin^2\theta_\nu} \right) + \frac{\partial}{\partial \theta_n} (\frac{\partial u}{\partial \theta_n} r^{n-3} \prod_{\nu=3}^{n} \sin^{\nu-2} \theta_\nu) \right\} .$$

Thus, the Laplacian in n dimensions has been transformed to polar coordinates. It might be noted that for n = 3, upon setting

$$\theta_2 = \phi$$
$$\theta_3 = \theta$$

the same result follows that was found previously for the 3-dimensional case.

It might further be observed that the n-dimensional Laplacian can be expressed in arbitrary curvilinear coordinates $(\xi_1, \xi_2, \ldots, \xi_n)$ in terms of the matrix JJ'. The result can easily be seen to be

$$\Delta u = \frac{1}{2\sqrt{\det(JJ')}} \sum_{i=1}^{n} \frac{\partial}{\partial \xi_1} \left\{ \sqrt{\det(JJ')} \frac{\partial}{\partial u_i} \left[(u_1 u_2 \ldots u_n)(JJ')^{-1} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \right] \right\}$$

where

$$u_i = \frac{\partial}{\partial \xi_i} \quad .$$

However, the fact that $(JJ')^{-1}$ is a symmetric matrix for any J implies that

$$\frac{\partial}{\partial u_i} \left[(u_1\ u_2 \ldots u_n)(JJ')^{-1} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \right] = 2(\delta_1^i\ \delta_2^i \ldots \delta_n^i)(JJ')^{-1} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

and, therefore, it follows that

$$\Delta u = \frac{1}{\sqrt{\det(JJ')}} \sum_{i=1}^{n} \frac{\partial}{\partial \xi_i} \left\{ \sqrt{\det(JJ')} (\delta_1^i\ \delta_2^i \ldots \delta_n^i)(JJ')^{-1} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \right\}.$$

In a notation often used, the elements of the matrix $(JJ')^{-1}$ are denoted by g^{ik} and $\det(JJ')$ may be denoted, for example, by a. Using this notation, the result can clearly be written as follows:

$$\Delta u = \frac{1}{\sqrt{a}} \sum_{i=1}^{n} \frac{\partial}{\partial \xi_i} \left(\sqrt{a} \sum_{k=1}^{n} g^{ik} \frac{\partial u}{\partial \xi_k} \right) .$$

Thus, it has been shown that the invariant nature of the Euler expression provides a highly powerful tool for transforming the Laplacian in n dimensions into curvilinear coordinates. However, this application constitutes merely one phase of the general significance of the Euler expression's invariance. The notion of invariance holds a predominant position in modern science and, therefore, the invariance of the Euler expression is of major importance in itself. Moreover, it is to be expected that the Euler expression ought to be instrumental in the construction of other invariant mathematical entities.

15. Transformation Theory.

a) Canonical Transformations.

The extremal of a variational problem

$$\int_a^b F(x,u,u') \, dx$$

satisfies the Euler equation

$$\frac{d}{dt} F_{u'_\nu} = F_{u_\nu}$$

which also can be written in the Hamiltonian form as derived in Chapter II. For this purpose we applied the Legendre transformation

$$v_\nu = F_{u'_\nu}$$

assuming $\det(F_{u'_\nu u'_\mu}) \neq 0$ and introduced

$$H(x,u,v) = \sum u'_\nu v_\nu - F .$$

Then the Euler equations are transformed into the symmetric form

$$\begin{cases} u_\nu' = H_{v_\nu} \\ v_\nu' = -H_{u_\nu} \end{cases} \tag{42}$$

This form of the differential equation is used frequently in mechanics where H represents the energy. The unusual feature of this system of differential equations is that it is determined by <u>one</u> function, namely H, only. There arises the question in which respect these Hamiltonian systems are special. This question will be answered in this Section.

If one introduces new coordinates in place of u,v it turns out that the transformed system in general will not be of the Hamiltonian form any more. This leads to the problem of characterizing those coordinate transformations which transform Hamiltonian systems into Hamiltonian systems. These transformations are called "canonical" and the aim of the transformation theory is the theory of the canonical transformations.

Before going into this question it is useful to observe that the Hamilton equations can be considered as the Euler equation of the variational problem

$$\int \Big(\sum_{\nu=1}^{n} u_\nu' v_\nu - H(x,u,v) \Big)\, dt \tag{43}$$

which is obtained from the old one by replacing the F by $\sum u_\nu' v_\nu - H(x,u,v)$. The Euler equation gives:

$$\frac{d}{dx} v_\nu = -H_{u_\nu}$$

$$0 = \frac{\partial}{\partial v_\nu} \Big(\sum_{\nu=1}^{n} u_\nu' v_\nu - H \Big) = u_\nu' - H_{v_\nu}$$

which are the Hamiltonian equations. This remark will be useful for the following consideration.

Let first $u_\nu = \phi_\nu(x,p)$ be a transformation of the coordinates u_ν alone into new coordinates p_ν. Such transformations

are of importance in applications where one tries to simplify the differential equation by taking into account symmetry properties, like polar coordinates are appropriate for rotationally symmetric systems. Is it possible to find functions $\psi_\nu(p,q,x)$ such that the transformation

$$u_\nu = \phi_\nu(x,p)$$
$$v_\nu = \psi_\nu(x,p,q)$$

transforms (42) into a Hamiltonian system

$$p'_\nu = G_{q_\nu}$$
$$q'_\nu = -G_{p_\nu} \quad ? \tag{44}$$

Furthermore we require that the Jacobian

$$\det\left(\frac{\partial(\phi_\nu,\psi_\nu)}{\partial(p_\kappa,q_\lambda)}\right) = \det\begin{pmatrix} \frac{\partial\phi}{\partial p} & , & 0 \\ \frac{\partial\psi}{\partial p} & , & \frac{\partial\psi}{\partial q} \end{pmatrix} = \det\left(\frac{\partial\phi_\nu}{\partial p_\mu}\right) \cdot \det\left(\frac{\partial\psi_\nu}{\partial q_\mu}\right)$$

should not vanish.

The answer to this question is yes as can be seen by the following simple argument. The equations (42) are equivalent to the Euler equations of

$$\int F(x,u,u')dx = \int F(x,\phi(p,x),\sum \phi_{p_\nu} p'_\nu + \phi_x)dx = \int \Phi(x,p,p')dx$$

where

$$\Phi(x,p,p') = F(x,\phi(p,x),\sum \phi_{p_\nu} p'_\nu + \phi_x) \quad .$$

Therefore they are equivalent to the Euler equations of this problem $\frac{d}{dx}\Phi_{p'_\nu} = \Phi_{p_\nu}$. Hence introducing

$$q_\nu = \Phi_{p'_\nu}$$

and the corresponding Hamilton function

$$G(x,p,q) = \sum p'_\nu q_\nu - \Phi$$

we obtain a Hamiltonian system as the transformed system.

In order to compute the functions $\psi_\nu(x,p,q)$ just compute

$$q_\nu = \Phi_{p'_\nu} = \sum_\mu F_{u'_\mu} \phi_{\mu p_\nu} = \sum_\mu v_\mu \phi_{\mu p_\nu}(x,p) \ .$$

We assume $\det(\phi_{\mu p_\nu}) \neq 0$. *Therefore one can solve these equations* for v_μ and obtain for $v_\nu = \psi_\nu(x,p,q)$ even linear functions in q. One easily checks that $\det(\partial\psi_\nu/\partial q_\mu) \neq 0$.

Now we want to treat the general question of under which conditions a transformation

$$\begin{cases} u_\nu = \phi_\nu(x,p,q) \\ v_\nu = \psi_\nu(x,p,q) \end{cases}$$

transforms a Hamiltonian system into a Hamiltonian system. We make use of the fact that the Hamilton equations (42) are the Euler equation of (43):

$$\int (\sum_{\nu=1}^{n} u'_\nu v_\nu - H(x,u,v))dx = \int \Phi(x,p,q,p',q')dx$$

where

$$\Phi = \sum_{\nu=1}^{n} v_\nu(\phi_{\nu x} + \sum_{\mu=1}^{n} \phi_{\nu p_\mu} p'_\mu + \phi_{\nu q_\mu} q'_\mu) - H(x,\phi,\psi) \ .$$

We want to find sufficient conditions under which this variational problem has the same extremals as $\int (\sum_{\nu=1}^{n} p'_\nu q_\nu - G(x,p,q))dx$ with some function $G(x,p,q)$. We certainly achieve this aim if we assume that the functions Φ and $\sum p'_\nu q_\nu - G$ differ only by the total derivative $\frac{d}{dx} W(x,p,q)$ of some function $W(x,p,q)$, because then

$$\int_a^b (\Phi - \sum p'_\nu q_\nu + G)dx = W(x,p,q) \Big|_a^b$$

does not depend on the curve of integration.

We have to investigate under which conditions there exist two functions $G(x,p,q)$, $W(x,p,q)$ such that

$$(45)\quad \sum_{\nu=1}^{n} u'_\nu v_\nu - H(x,u,v) = \sum_{\nu=1}^{n} p'_\nu q_\nu - G(x,p,q) + \frac{d}{dx} W(x,p,q) \ .$$

In this equation we have to replace u_ν, v_ν by ϕ_ν, ψ_ν. Equation (45) should hold identically with respect to the variables p,q,p',q',x. If we compare the coefficients of p'_μ we find

$$\sum_{\nu=1}^{n} \phi_{\nu p_\mu} \psi_\nu = q_\mu + W_{p_\mu} \ ,$$

the coefficient of q'_μ gives

$$\sum_{\nu=1}^{n} \phi_{\nu q_\mu} \psi_\nu = W_{q_\mu}$$

and finally

$$\sum_{\nu=1}^{n} \phi_{\nu x} \psi_x - H = -G + W_x \ .$$

The last equation can be satisfied by appropriate choice of $G(x,p,q)$. The other 2n equations should be satisfied after an appropriate choice of W. They determine all derivatives of W:

$$W_{p_\mu} = -q_\mu + \sum_{\kappa=1}^{n} \phi_{\kappa p_\mu} \psi_\nu = f_\mu(x,p,q)$$

$$W_{q_\mu} = \sum_{\kappa=1}^{n} \phi_{\kappa q_\mu} \psi_\nu = g_\mu(x,p,q)$$

and they can be solved for W if and only if the integrability conditions

$$f_{\mu p_\nu} = f_{\nu p_\mu} \ , \quad f_{\mu q_\nu} = g_{\nu p_\mu} \ , \quad g_{\mu q_\nu} = g_{\nu q_\mu}$$

are fulfilled. These conditions lead to the following relations

$$(46)\quad \begin{cases} f_{\mu p_\nu} - f_{\nu p_\mu} = \sum_{\kappa=1}^{n} (\phi_{\kappa p_\mu} \psi_{\kappa p_\nu} - \phi_{\kappa p_\nu} \psi_{\kappa p_\mu}) = 0 \\ f_{\mu q_\nu} - g_{\nu p_\mu} = -\delta_{\nu\mu} + \sum_{\kappa=1}^{n} (\phi_{\kappa p_\mu} \psi_{\kappa q_\nu} - \phi_{\kappa q_\nu} \psi_{\kappa p_\mu}) = 0 \\ g_{\mu q_\nu} - g_{\nu q_\mu} = \sum_{\kappa=1}^{n} (\phi_{\kappa q_\mu} \psi_{\nu q_\nu} - \phi_{\kappa q_\nu} \psi_{\kappa q_\mu}) = 0 \ . \end{cases}$$

Here $\delta_{\nu\mu}$ denotes the Kronecker symbol: $\delta_{\nu\mu} = 0$ for $\nu \neq \mu$, and $\delta_{\nu\nu} = 1$.

To simplify the notation we introduce the Lagrange bracket symbols

$$[s,t] = \sum_{k=1}^{n} (\phi_{ks}\psi_{kt} - \phi_{kt}\psi_{ks})$$

where s,t are two of the variables p_ν, q_μ. Then we can say

$$(46^*) \qquad \begin{cases} [p_\nu, p_\mu] = [q_\nu, q_\mu] = 0 \\ [p_\nu, q_\mu] = 0 \quad \text{for } \nu \neq \mu\,; \qquad [p_\nu, q_\nu] = 1 \end{cases}$$

are conditions which guarantee the existence of a function W and G, such that (45) holds. Hence under these conditions the Hamiltonian system (42) is transformed into the Hamiltonian system (44).

Definition: A transformation $u_\nu = \phi_\nu(x,p,q)$, $v_\nu = \psi_\nu(x,p,q)$ which satisfies the equations (46) or (46*) is called a canonical transformation.

We just have shown that every canonical transformation preserves the Hamiltonian character of a system of differential equations. It is interesting that the conditions (46) do not depend on the special Hamiltonian function.

Using matrix notation we can write the conditions (46) in a very simple form: Introduce the 2n by 2n matrices

$$M = \begin{pmatrix} \phi_{\nu p_\mu} & \phi_{\nu q_\mu} \\ \psi_{\nu p_\mu} & \psi_{\nu q_\mu} \end{pmatrix}\,; \qquad J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

where I_n is the $n \times n$ unit matrix. Then the conditions (46) can be written in the form

$$(47) \qquad M'JM = J$$

where M' denotes the transposed of M.

A matrix which satisfies the relation (47) is called symplectic. These matrices play a role in different parts of mathematics. All these symplectic matrices form a group, which is

called the symplectic group - one of the elementary groups in the theory of Lie groups. To establish the group character we have to show that M^{-1} also is symplectic. For this purpose observe that from (47) follows $\det M' \cdot \det M = (\det M)^2 = 1$ and therefore $\det M = \pm 1$. Hence we can form M^{-1}. Multiplying (47) from the right by M^{-1} and from the left by $(M')^{-1}$ one obtains

$$J = (M^{-1})' J M^{-1}$$

which proves that M^{-1} is symplectic. Similarly one shows that the product of two symplectic matrices is symplectic. From the group property of the symplectic matrices follows immediately that the canonical transformations form a group where multiplication means applying one transformation after the other.

Although this characterization of the canonical transformation is very general and algebraically satisfactory we want to discuss another property of canonical transformations which plays an important role in mechanics and optics. For this purpose we go back to equation (45) but consider u,p as independent variables, instead of p,q. In other words we write

$$p_\nu = \alpha_\nu(x,u,p)$$

$$q_\nu = \beta_\nu(x,u,p)$$

and

$$W = V(x,u,p) \quad .$$

Then we get from (45)

$$\sum u'_\nu \alpha_\nu - H = \sum p'_\nu \beta_\nu - G + \frac{d}{dx} V \quad .$$

If we now compare the coefficients of u'_ν, p'_ν we obtain

$$\begin{cases} v_\nu = \alpha_\nu = V_{u_\nu} \\ q_\nu = \beta_\nu = -V_{p_\nu} \end{cases}$$

and the remaining terms give

$$-H = -G + V_x \quad .$$

Therefore we have a representation of a canonical transformation by one function V(x,u,p) which is called the "generating function". In this representation, however, the old and the new variables are mixed up and one has carefully to observe the order:

$$\begin{aligned} v_\nu &= V_{u_\nu}(x,u,p) \\ q_\nu &= -V_{p_\nu}(x,u,p) \end{aligned} \qquad (48)$$

If we assume $\det(V_{p_\nu u_\mu}) \neq 0$, we can solve the last equation for u_ν in terms of p,q. Inserting the result into the first equation we get v_ν in terms of p,q. This must give a canonical transformation according to the derivation. Furthermore we have an explicit representation of the new Hamiltonian

$$G = H + V_x \qquad (48*)$$

Observe that in general the transformation of a system of differential equations is an involved process while here the new system is obtained directly from the Hamiltonian G. However, one still has to invert the equation (48) in the way indicated above. In particular, if the transformation does not depend explicitly on x, we have $V_x = 0$ and $G = H$. In other words G is obtained from H by substituting the new variables.

There is another representation of a canonical system with a generating function. In this case we pick u,q as independent variables. Furthermore we introduce

$$U(q) = W + \sum_\nu p_\nu q_\nu \quad .$$

Then (45) goes over into

$$\sum u'_\nu v_\nu - H = -\sum p_\nu q'_\nu - G + \frac{d}{dx} U(x,u,q) \quad .$$

Comparing coefficients of u'_ν, q'_ν we find

$$\begin{cases} v_\nu = U_{u_\nu}(x,u,q) \\ p_\nu = U_{q_\nu}(x,u,q) \end{cases} \qquad (49)$$

and

(49*) $$G = H + U_x$$

In order to obtain the coordinates u,v in terms of p,q we have to assume det $(U_{q_\nu u_\mu}) \neq 0$ and to solve the second equation (49) for u_ν. Notice that the function U(x,u,q) is not completely determined by the transformation. If one adds any function of x alone the new function will generate the same transformation.

The transformations obtained from (48) and (49) are canonical transformations according to their derivation. But not every canonical transformation can be brought into the form (48) or the form (49). One can prove that the necessary and sufficient condition for a canonical transformation

$$u_\nu = \phi_\nu(x,p,q)$$
$$v_\nu = \psi_\nu(x,p,q)$$

to be equivalent to (48) is that

$$\det (\psi_{\nu p_\mu}) \neq 0 \quad .$$

Similarly the transformation is equivalent to (49) if and only if $\det (\psi_{\nu q_\mu}) \neq 0$. For instance the identity transformation $v_\nu = q_\nu$, $u_\nu = p_\nu$ cannot be represented in the form (48) but in the form (49) with $U = \sum_{\nu=1}^{n} u_\nu q_\nu$.

Using the generating function we can easily answer the question, how to extend a transformation

$$p_\nu = \phi_\nu^*(x,u)$$

of the u into p to a canonical transformation of the u,v variables into p,q. We assume that $\det (\phi_\nu^*) \neq 0$. We try to construct a generating function U(x,u,q) such that

$$U_{q_\nu} = \phi_\nu^*(x,u) \quad .$$

For example

$$U = \sum_{\nu=1}^{n} q_\nu \phi_\nu^*(x,u)$$

has this property and

$$\det\left(U_{q_\nu u_\mu}\right) = \det\left(\phi^*_{\nu u_\mu}\right) \neq 0 \quad .$$

Therefore formula (49) shows how to transform the v_ν:

$$v_\nu = U_{u_\nu} = \sum q_\mu \phi^*_{\mu u_\nu}$$

which is a linear transformation between the v and the q.

Examples:

1) The transformation of rectangular coordinates into polar coordinates

$$p = \frac{1}{2}(u^2 + v^2)$$

$$q = \arctan \frac{v}{u}$$

or

$$u = \sqrt{2p} \cos q$$
$$v = \sqrt{2p} \sin q$$

is canonical.

2) An interesting example is the transformation which one obtains by extending the inversion on a circle to a canonical transformation. The inversion can be written in the form

$$p_\nu = \frac{u_\nu}{u_1^2 + u_2^2 + \ldots + u_n^2} \qquad (\nu = 1,2,\ldots,n)$$

or

$$u_\nu = \frac{p_\nu}{p_1^2 + p_2^2 + \ldots + p_n^2} \tag{50}$$

so that

$$(p_1^2 + \ldots + p_n^2)(u_1^2 + \ldots + u_n^2) = 1$$

This transformation T is called an involution because it is identical with its inverse. This property should be preserved for the extended canonical transformation.

We construct a generating function $U(x,u,q)$ and have according to (49)

$$U_{q_\nu} = \frac{u_\nu}{u_1^2 + u_2^2 + \ldots + u_n^2}$$

and we choose

$$U = \frac{\sum_{\nu=1}^{n} u_\nu q_\nu}{\sum_{\nu=1}^{n} u_\nu^2} \quad .$$

One checks easily that $\det (U_{u_\nu q_\mu}) \neq 0$. For simplicity we write u,v for the vectors $(u_\nu),(v_\nu)$ and

$$|u|^2 = \sum u_\nu^2 \quad , \qquad (u,v) = \sum u_\nu v_\nu \quad .$$

Similarly for p,q. Then from the first equation of (49)

$$v_\nu = U_{u_\nu} = \frac{q_\nu}{|u|^2} - 2\,\frac{(u,q)u_\nu}{|u|^4} \quad ,$$

and since $|u|^2|p|^2 = 1$, and $u_\nu/|u|^2 = p_\nu$ we obtain the formula

$$v_\nu = |p|^2 q_\nu - 2(p,q)p_\nu \tag{50'}$$

which gives the transformation of the v_ν.

This transformation plays a role in celestial mechanics in the problem to describe the motion of particles in the 3-body problem for a collision. A remarkable property of this transformation is that its inverse also is rational in the coordinates, and as a matter of fact, the inverse has the same form as the original transformation. It is an involution. To check this we compute

$$(u,p) = \left(\frac{p}{|p|^2}, p\right) = 1$$

$$(u,q) = \left(\frac{p}{|p|^2}, q\right) = \frac{(p,q)}{|p|^2}$$

$$(u,v) = |p|^2(u,q) - 2(p,q)(u,p)$$
$$- (p,q) - 2(p,q) = -(p,q) \quad .$$

Therefore we find from (50') by solving for q:

$$q_\nu = \frac{1}{|p|^2} v_\nu + \frac{2(p,q)}{|p|^2} p_\nu = |u|^2 v_\nu + 2(p,q)u_\nu = |u|^2 v_\nu - 2(u,v)u_\nu .$$

Hence the inverse mapping

$$\begin{cases} p_\nu = \dfrac{u_\nu}{|u|^2} \\ q_\nu = |u|^2 v_\nu - 2(u,v)u_\nu \end{cases}$$

has exactly the same form as (50), (50').

As an application we solve the Hamilton equations for

$$H = \sum_{\nu=1}^{n} u_\nu^2 \sqrt{\sum_{\nu=1}^{n} v_\nu^2} = |u|^2 |v| \quad .$$

We make use of the above transformation (50), (50') and observe that it does not depend explicitly on x. Therefore we obtain the new Hamiltonian just by expressing $|u|^2|v|$ in the new variables p,q. For this purpose we compute from (50'):

$$|v|^2 = \sum_{\nu=1}^{n} v_\nu^2 = |p|^4|q|^2 - 4|p|^2(p,q)^2 + 4(p,q)^2|p|^2$$
$$= |p|^4|q|^2$$

or because of $|u|^2|p|^2 = 1$

$$|u|^2|v|^2 = |p|^2|q|^2 \quad .$$

Then

$$|u|^2|v| = |u|\,|p|\,|q| = |q|$$

and we have

$$H(x,u,v) = |u|^2|v| = |q| = G(x,p,q) \quad .$$

Therefore the new Hamiltonian does not depend on the p variables at all. The Hamilton equations

$$\begin{cases} p_\nu' = G_{q_\nu} \\ q_\nu' = -G_{p_\nu} = 0 \end{cases}$$

can be solved explicitly:

$$q_\nu(x) = q_\nu(0)$$
$$p_\nu(x) = G_{q_\nu} x + p_\nu(0) = \frac{q_\nu}{q_\nu^3} x + p_\nu(0)$$

The solutions are straight lines. The above transformation (50),(50') represents the solutions in the old coordinates u,v explicitly, if one inserts for p_ν, q_ν the obtained functions of x.

b) Infinitesimal Canonical Transformation.

In the preceding section we introduced the canonical transformations by means of their property of preserving the Hamiltonian character of a system of differential equations. There is another connection which we develop now: If we consider a system of differential equations

$$\frac{dw_\nu}{dx} = f_\nu(x,w) \qquad (\nu = 1,2,\ldots,m)$$

they define in a natural way a transformation. If the solution of this system with the initial values $w_\nu = w_\nu^0$ are denoted by $w_\nu = g_\nu(x,w^0)$, then we have for every sufficiently small $|x|$ a transformation of the w_ν^0 into the w_ν. For $x = 0$ this is the identity transformation and the Jacobian equals one. Therefore the Jacobian is different from zero for sufficiently small $|x|$. We want to call $w_\nu = g_\nu(x,w^0)$ the transformations "generated" by the systems $w_\nu' = f_\nu(x,w)$.

We are going to prove the following statement. The systems of differential equations for which the transformations generated by it are canonical are exactly the Hamiltonian systems. This theorem shows the close connection between the Hamiltonian systems and the canonical transformations.

Before proving this we consider transformations

$$u_\nu = \phi_\nu(x,p,q)$$
(51)
$$v_\nu = \psi_\nu(x,p,q)$$

depending on a parameter x and denote it symbolically by $T(x)$. For small x we can write formally $T(x) = T(0) + xT_x(0) + \dots$. We ask the question what is the condition on $T_x(0)$ and $T(0)$ for $T(x)$ to be canonical? Of course $T(0)$ has to be canonical as one suspects by letting $x = 0$. What is the condition for $T_x(0)$? The transformation $T_x(0)$ is obtained from (51) as

$$\begin{cases} u_\nu = \phi_{\nu x}(0,p,q) \\ v_\nu = \psi_{\nu x}(0,p,q) \end{cases}$$

and is called the infinitesimal transformation of (51).

We obtain such a condition immediately from (47) by introducing the matrix

$$M(x) = \begin{pmatrix} \phi_{\nu p_\mu} & \phi_{\nu q_\mu} \\ \psi_{\nu p_\mu} & \psi_{\nu q_\mu} \end{pmatrix} \tag{52}$$

If (51) is canonical we have

$$M'JM = J$$

and differentiating with respect to x:

$$M'_x JM + M'JM_x = 0$$

Therefore $M'JM_x$ is a symmetric matrix, since $(M'JM_x)' = M'_x J'M = -M'_x JM = M'JM_x$. This is the condition for the infinitesimal transformation to be canonical. It can be expressed in another form namely that JM_xM^{-1} be symmetric, which is equivalent to the old condition since

$$M'JM_x = M'(JM_xM^{-1})M \quad .$$

This also shows that JM_xM^{-1} is symmetric if and only if $M'JM_x$ is symmetric. Defining $S = -JM_xM^{-1}$ a symmetric matrix we have with $J^{-1} = -J$

$$M_x = JSM \tag{53}$$

as the condition for the infinitesimal transformation to be canonical.

We apply this result to a family of transformations which are generated by a system of differential equations

$$w_\nu' = f_\nu(x,w) \qquad (\nu = 1,\ldots,m) \tag{54}$$

The transformation generated by (54) we write in the form

$$w_\nu = \chi_\nu(x,r)$$

where $r_\nu = w_\nu(0)$. We define $M(x)$ by

$$M = \left(\frac{\partial \chi_\nu}{\partial r_\mu}\right) \qquad (\nu,\mu = 1,\ldots,m) \quad .$$

From (53) we have

$$\chi_{\nu x} = f_\nu(x,\chi)$$

and differentiating with respect to r_μ

$$\chi_{\nu r_\mu x} = \sum_{k=1}^{m} f_{\nu w_k} \chi_{k r_\mu} \quad .$$

These equations can be combined into the matrix equation

$$M_x = (f_{\nu w_\mu})M \quad .$$

Comparing this equation with (53) we see that the transformations generated by (54) are canonical if and only if

$$(f_{\nu w_\mu}) = JS$$

where S is a symmetric matrix. This implies m is even, so that we can write $m = 2n$. To discuss this condition we define $g_1,\ldots,g_{2n}$ by

$$g_\nu = -f_{\nu+n} \quad , \quad g_{\nu+n} = f_\nu \quad \text{for } \nu = 1,2,\ldots,n \quad .$$

Then

$$-J(f_\nu) = g_\nu$$

and

$$S = J^{-1}(f_{\nu w_\mu}) = (g_{\nu w_\mu}) \quad .$$

This matrix has to be symmetric. This condition is identical with the integrability condition for a function $H(x,w)$, such that

$$g_\nu = H_{w_\nu} \quad .$$

Therefore $(f_\nu) = J(g_\nu) = J(H_{w_\nu})$ and the system (54) has the form.

$$w' = J(H_{w_\nu}) \quad .$$

If we now define $w_\nu = u_\nu$, $w_{\nu+n} = v_\nu$ for $\nu = 1,2,\ldots,n$ we have

$$\begin{aligned} u'_\nu &= H_{v_\nu} \\ v'_\nu &= -H_{u_\nu} \end{aligned} \tag{55}$$

which shows the Hamiltonian character of the system.

Therefore the Hamiltonian systems can be characterized by the property that the transformations generated by them are canonical.

As a consequence of this result we obtain a simple proof for the existence of a complete solution of the Hamilton equations: Denote the transformation generated by (55) by

$$\begin{aligned} u_\nu &= \phi_\nu(x,p,q) \\ v_\nu &= \psi_\nu(x,p,q) \end{aligned}$$

where $p_\nu = u_\nu(0)$, $q_\nu = v_\nu(0)$. Then this transformation is canonical and for sufficiently small $|x|$ we have $\det(\psi_{\nu q_\mu}) \neq 0$. Therefore we can write the transformation in the form (49) with a generating function $U(x,u,q)$. The new Hamiltonian is given by

$$G(x,p,q) = H(x,u,v) + U_x \quad .$$

Since, however, p,q, does not depend on x, we have $0 = p' = G_q$, $0 = q' = -G_p$, so that $G = G(x)$ does not depend on p or q. If we now form $S(x,u,q) = U(x,u,q) - \int_0^x G(x)dx$, we can take S as a generating function of the same transformation. Then we find that the new Hamiltonian is

$$0 = H(x,u,v) + S_x$$

and the equation for S is

$$H(x,u,S_u) + S_x = 0$$

which is just the Hamilton Jacobi equation. The function $S(x,u,q)$ is a complete solution since

$$\det (S_{u_\nu q_\mu}) = \det (U_{u_\nu q_\mu}) \neq 0 \quad .$$

Thus we have established the existence of a complete solution of the Hamilton Jacobi equation.

This approach allows a new interpretation of the Hamilton Jacobi equation. The solution of the Hamilton Jacobi equation $S(x,u,q)$ is a generating function of a canonical transformation for which the new Hamiltonian is $G \equiv 0$. Then the new system is $p' = q' = 0$ which is easily integrated.

Frequently it is too complicated a task to solve the Hamilton Jacobi equation but one succeeds in simplifying the Hamiltonian. For instance, in the example of the last section

$$H = \sum_{\nu=1}^{n} u_\nu^2 \sqrt{\sum v_\nu^2}$$

we found a canonical transformation such that the new Hamiltonian

$$G = \sqrt{\sum_{\nu=1}^{n} q_\nu^2}$$

did not depend on the p_ν. Therefore $q_\nu' = -G_{p_\nu} = 0$ and q_ν are independent of x. In this way the transformation theory is used for solving the Hamilton equation.

More important, however, is the application of the transformation theory in the perturbation theory of classical mechanics.* Consider a Hamiltonian $H = H_0 + \varepsilon H_1$ which is close to a

---

* To distinguish it from the perturbation theory of quantum mechanics.

Hamiltonian of a system which can be integrated. For example in Astronomy one is interested in the motion of a planet which moves almost on an ellipse and H_0 would represent the Hamiltonian which describes the attraction of the sun. The other planets, however, cause some perturbation of this system which is to be considered small. On would like to get some approximate solution for the "perturbed" system. For this purpose assume that the Hamilton Jacobi equation for the unperturbed system

$$H_0(x,u,S_u) + S_x = 0$$

can be solved. Considering $S(x,u,q)$ as the generating function of a canonical transformation we obtain for the new Hamiltonian

$$G = H_0(x,u,S_u) + \varepsilon H_1(x,u,S_u) + S_x = \varepsilon H_1(x,u,S_u)$$

so that we can write $G = \varepsilon G_1(x,p,q)$. The Hamiltonian system is transformed into

$$p_\nu' = \varepsilon G_{1q_\nu}$$

$$q_\nu' = -\varepsilon G_{1p_\nu}$$

which shows that the coordinates p_ν, q_ν only change slowly. By approximate integration of this simplified system one can find an approximation to the solution.

c) Application of Canonical Transformation in Optics.

In optics one investigates the effect of lenses or other optical instruments on light rays. The basic law that describes these processes is Fermat's principle which states that the light rays are extremals of a variational problem: A ray starting at P and ending at Q has the property of minimizing the time required to travel along a curve from P to Q. For a homogeneous medium the light rays are straight lines and the statement of Fermat's principle is obvious. This principle, however, holds also for rays passing through optical instruments provided one considers the frequency of light as extremely large. This is the idealization of geometrical optics.

Therefore we can consider the light rays as extremals of a variational problem. If we consider isotropic media only this variational problem can be written in the form

$$\int_{t_1}^{t_2} n(x) \sqrt{\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2}\, dt$$

where $x = (x_1, x_2, x_3)$ and $n(x)$ is the refraction index. This index of refraction will be discontinuous on the surface of a lense which will make the theory as we have developed it not directly applicable. We can avoid this difficulty by replacing $n(x)$ by a continuous function which is a good approximation to the discontinuous function.

We want to describe the effect of a lense on light rays. We choose the coordinate x_1 such that it passes through the lense. The x_1, x_2, x_3 axis are assumed to be mutually orthogonal. Since we are only interested in rays passing through the instrument we can rewrite the variational problem in the inhomogeneous form

$$\int_a^b n(x) \sqrt{1 + x_2'^2 + x_3'^2}\, dx_1$$

where $x_2' = dx_2/dx_1$; $x_3' = dx_3/dx_1$. The light rays are the extremals of this variational problem. Introducing the canonical variables

$$y_2 = \frac{\partial}{\partial x_2'} \left(n(x)\sqrt{1 + x_2'^2 + x_3'^2}\right) = n(x) \frac{x_2'}{\sqrt{1 + x_2'^2 + x_3'^2}}$$

(56)

$$y_3 = \frac{\partial}{\partial x_3'} \left(n(x)\sqrt{1 + x_2'^2 + x_3'^2}\right) = n(x) \frac{x_3'}{\sqrt{1 + x_2'^2 + x_3'^2}}$$

and the Hamiltonian function

$$H(x,y) = x_2'y_2 + x_3'y_3 - n(x)\sqrt{1 + x_2'^2 + x_3'^2} = -\sqrt{n^2 - y_2^2 - y_3^2}$$

we can write the differential equations for the extremals in the Hamiltonian form

$$(57)\qquad \begin{aligned} x_\nu' &= H_{y_\nu} \\ y_\nu' &= -H_{x_\nu} \end{aligned} \qquad (\nu = 2,3).$$

The variables y_2, y_3 determine the direction of the ray. According to the result of the last section this Hamiltonian system generates a canonical transformation which will be characteristic for the optical instrument. We denote the initial values of x_ν, y_ν $(\nu = 2,3)$ for $x_1 = 0$ by ξ_ν, η_ν and the solutions of (57) with these initial values by

$$\begin{cases} x_\nu = \phi_\nu(x_1, \xi, \eta) \\ y_\nu = \psi_\nu(x_1, \xi, \eta) \end{cases} \qquad (\nu = 2,3)$$

where $\xi = (\xi_2, \xi_3)$; $\eta = (\eta_2, \eta_3)$. Then we know that the transformation of the $\xi_\nu \eta_\nu$ into the x_ν, y_ν for every x_1 are a canonical transformation.

We now assume that the whole optical instrument lies completely in the interval $0 < x < 1$. In other words we want to assume that $n(x) = \text{const} = n$ for $x_1 \leq 0$ and $x_1 \geq 0$. In those regions $n(x)$ represents the refraction index of the air.

We want to study the transformation of the rays starting in a plane $x_1 = 0$ on one side of the instrument into the corresponding rays passing through a plane $x_1 = 1$ on the other side. Then this transformation is given by

$$(58)\qquad \begin{aligned} x_\nu &= \phi_\nu(1, \xi, \eta) \\ y_\nu &= \psi_\nu(1, \xi, \eta) \end{aligned} \qquad (\nu = 2,3)$$

which is a canonical transformation of the variables ξ_ν, η_ν into the variables x_ν, y_ν $(\nu = 2,3)$.

If we now assume

$$\det(\phi_{\nu\xi_\mu}) \neq 0 \qquad (\nu,\mu = 2,3)$$

we can find a generating function $V(x,\xi)$ such that

$$(59) \qquad \begin{cases} y_\nu = V_{x_\nu} \\ \eta_\nu = - V_{\xi_\nu} \end{cases} \qquad (\nu = 2,3)$$

This function plays an important role in optics and is called the "Eikonal". All the optical properties of the instruments are determined by this function. The property (59) of the function $V(x,\xi)$ can be interpreted as follows. If $S(x,\xi)$ is the extremal integral taken along an extremal from ξ to x, then according to our calculations in Chapter II we have $S_{x_\nu} = y_\nu$ and $S_{\xi_\nu} = -\eta_\nu$ for the direction of the corresponding extremals. Therefore the function $V(x,\xi)$ can be identified with the extremal integral from ξ to x. In other words $V(x,\xi)$ represents the time in which the light travels from ξ to x.*

In optics one is interested in transformations which give a sharp picture of an object; in other words transformations that transform all rays which start in a point into rays which pass through the image point. A transformation of this kind is called "stigmatic". An optical instrument that produces a transformation of this kind is called an "absolute" instrument.

It is a very interesting result of Maxwell, that every stigmatic mapping is trivial in the sense that object and image have the same size. In other words no enlargement is possible with a sharp picture. An enlarging lens system can only give approximately sharp pictures in the neighborhood of the axis. As a matter of fact the mirror is the only optical instrument that is known to produce a stigmatic transformation.

We want to prove Maxwell's theorem. We assume we have an isotropic medium, i.e. the variational problem is of the form

$$\int_{t_1}^{t_2} n(x)\sqrt{\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2}\, dt = \int_{t_1}^{t_2} F(x,\dot{x})\, dt$$

* For further details see: R.K. Luneburg, Mathematical Theory of Optics, Brown University, 1944, esp. Chapter II.

Assume there exists a stigmatic transformation of $x = (x_1, x_2, x_3)$ into a point $\xi = (\xi_1, \xi_2, \xi_3)$:

$$\xi_\nu = \phi_\nu(x) \qquad \text{or} \qquad x_\nu = \psi_\nu(\xi) \qquad (\nu = 1,2,3)$$

According to the assumption all extremals starting at x intersect each other at ξ, provided they can be continued that far. This means x and ξ are conjugate points though not necessarily consecutive ones. One proves, as for conjugate points, that the integral

$$\int_x^\xi F(x,\dot{x})\, dt$$

has the same value for all extremals connecting x and ξ. Therefore it is a function of $x = (x_1, x_2, x_3)$ along:

$$V(x) = \int_x^\xi F(x,\dot{x})\, dt$$

Secondly we want a connection between the differentials dx and $d\xi$. For this purpose consider an extremal $\mathcal{E}$ connecting x and ξ and a point y on it close to x. Let $\eta = \phi_\nu(y)$ be its image point. It is clear that η lies on the continuation of the extremal $\mathcal{E}$ since in η all extremals starting at y intersect, especially $\mathcal{E}$. Then

$$\int_x^y F\, dt - \int_\xi^\eta F\, dt = \int_x^\xi F\, dt - \int_y^\eta F\, dt = V(x) - V(y) \quad .$$

If we now pass to the derivative by letting y approach x we have

$$(60)\quad n(x)\sqrt{(dx_1)^2 + (dx_2)^2 + (dx_3)^2}$$
$$- n(\phi(x)\sqrt{(d\xi_1)^2 + (d\xi_2)^2 + (d\xi_3)^2}$$
$$= V_{x_1} dx_1 + V_{x_2} dx_2 + V_{x_3} dx_3 \ .$$

Here $d\xi_\mu = \phi_{\nu x_\mu} dx_\mu$ is the image of dx and this relation holds for all dx which lie on an extremal connecting x and ξ. But since both sides are analytic functions of dx this equation holds identically for all dx. This can be interpreted as follows: If we integrate the left hand side over a curve g and denote its image curve by γ then

$$\int_g n(x)\sqrt{\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2}\, dt - \int_\gamma n(\xi)(\dot{\xi}_1^2 + \dot{\xi}_2^2 + \dot{\xi}_3^2)dt = V(b) - V(a)$$

where a,b are the initial and end points of g. In particular, if g is a closed curve the right hand side vanishes. In other words the integral $\int F\, dt$ taken over a closed curve has the same value if it is taken over the image curve.

The aim of the proof is to show that V is a constant. Then the above statement will hold for any curve, not only closed ones. For this purpose we observe that (60) holds for any vector dx. Replace dx by λ dx where λ is any real number. Then the left hand side is replaced by $|\lambda|$ times the left hand side. But the right hand side is multiplied by λ. So if we take $\lambda = -1$ the left hand side does not change but the right hand side changes sign. Therefore

$$V_{x_1} dx_1 + V_{x_2} dx_2 + V_{x_3} dx_3 = 0$$

or V = const.

Therefore we have the following result: For every stigmatic transformation the integral $\int F\, dt$ taken from a point to its image point is a constant. Furthermore

$$\int_g n(x)\sqrt{\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2}\, dt = \int_\gamma n(\xi)\sqrt{\dot{\xi}_1^2 + \dot{\xi}_2^2 + \dot{\xi}_3^2}\, dt$$

which says that the "optical length" of a curve and its image curve are the same.

If in particular the refraction index along g and along γ are the same constant, as would be the case for a mapping through

a lens system with air on both sides, then the actual length of g and γ would coincide. This proves Maxwell's theorem.

16. An Approximation Theorem for Minimum Problems with Side Conditions.

In order to solve a minimum problem in which objects entering into the competition are required to satisfy certain side conditions or boundary conditions, it may be possible to approximate the given problem by others without side conditions. In other words, it seems reasonable to attempt to obtain the solution to a minimum problem with side conditions as the limit of solutions to a sequence of problems without side conditions.

For example, this strategy can be used to find the point closest to the origin on the straight line

$$Ax + By + C = 0 \quad ;$$

that is, the point which minimizes $f(x,y)$ where

$$f(x,y) \equiv x^2 + y^2$$

subject to the side condition

$$g(x,y) = 0$$

where

$$g(x,y) \equiv Ax + By + C \quad .$$

The solution can be obtained by first finding the point (x_n,y_n) which minimizes $f_n(x,y)$ given by

$$f_n(x,y) \equiv f(x,y) + n[g(x,y)]^2$$

or, in particular,

$$f_n(x,y) \equiv x^2 + y^2 + n(Ax + By + C)^2$$

when no side conditions are imposed. It seems plausible that $|g(x_n,y_n)|$ will be small for large values of n since (x_n,y_n)

minimizes $f_n'(x,y)$ and the non-negative term $n[g(x,y)]^2$, therefore, could not have a very large value. Thus, it might be expected that, as n tends to infinity, (x_n,y_n) will approach a point (x_∞,y_∞) such that

$$g(x_\infty,y_\infty) = 0$$

and (x_∞,y_∞) minimizes $f(x,y)$ among all points satisfying the side condition. In fact, it can be verified that this is precisely the case since solving the equations

$$\frac{\partial}{\partial x} f_n\Big|^{x=x_n} = 2x_n + 2An(Ax_n + By_n + C) = 0$$

$$\frac{\partial}{\partial y} f_n\Big|^{y=y_n} = 2y_n + 2Bn(Ax_n + By_n + C) = 0$$

yields

$$x_n = \frac{-ACn}{1 + (A^2 + B^2)n}$$

$$y_n = \frac{-BCn}{1 + (A^2 + B^2)n} \quad .$$

Clearly, as n tends to infinity the point $(-\frac{AC}{A^2+B^2},-\frac{BC}{A^2+B^2})$ is approached, and it can be checked by elementary methods that this is the closest point to the origin on the straight line.

Of course, in order to be able to use such a method in other problems, some theorem is needed guaranteeing that the solution is produced by the limiting process. An appropriate theorem will now be proved, and the theorem will then be used to justify the passage to the limit in a significant example. Specifically, the theorem will be applied in proving a Lagrange multiplier rule (Euler-Lagrange rule) for functions of two variables. In addition, it will be indicated how the theorem might be used to justify a procedure for obtaining the solution to minimum problems with artificial boundary conditions as the limit of solutions of sequences of problems in which no boundary conditions are imposed in advance.

The underlying theorem for the approximation procedure is the following:

AN APPROXIMATION THEOREM:

Part I:

If:

a) $\Phi(P)$ and $\psi(P)$ are lower semi-continuous real-valued functions on a convergence space* S;

b) $\psi(P) \geq 0$ for all P in S and there exist points in S for which $\psi(P) = 0$;

c) A denotes the problem: Find a point satisfying the side condition

$$\psi(P) = 0$$

at which $\Phi(P)$ takes on its least value for all P in S satisfying the side condition;

d) A_t denotes the problem: Find a point for which $\Phi(P) + t\psi(P)$ takes on its least value for all P in S; and

e) There exist a sequence $\{t_n\}$ of positive real numbers, a sequence $\{P_n\}$ of points in S, and a point P_∞ in S such that $t_n \to \infty$ as $n \to \infty$, P_n solves A_{t_n}, and $P_n \to P_\infty$ as $n \to \infty$;

then:

P_∞ solves A.

Part II:

If S is compact, then condition e) of Part I is automatically satisfied. Moreover, every sequence $\{\tau_n\}$ of positive numbers such that $\tau_n \to \infty$ as $n \to \infty$ possess a subsequence which can be taken as the sequence $\{t_n\}$ of condition e) of Part I.

The proof of Part I of the theorem depends primarily on an inequality relating the least value d_n for the function minimized in problem A_{t_n} and the greatest lower bound d of the values of $\Phi(P)$ for points satisfying the side condition imposed in Problem A. Specifically, it will now be shown that if d_n is

* The term convergence space here denotes any space in which a notion of convergence is defined.

given by

$$d_n = \underline{\Phi}(P_n) + t_n\psi(P_n)$$

and if d is given by

$$d = \underset{[\psi(P)=0]}{\text{g.l.b.}}\ \underline{\Phi}(P) \quad ,$$

then d is finite and

$$d_n \leq d \quad .$$

This follows immediately from the fact that, if P is any point in S satisfying the side condition

$$\psi(P) = 0 \quad ,$$

then

$$\underline{\Phi}(P) = \underline{\Phi}(P) + t_n\psi(P) \geq d_n \quad ,$$

where the inequality holds since d_n is the least value of $\underline{\Phi}(P) + t_n\psi(P)$ for all P in S. Thus,

$$d_n \leq \underline{\Phi}(P)$$

for all P satisfying the side condition, and, consequently

$$d_n \leq \underset{[\psi(P)=0]}{\text{g.l.b.}}\ \underline{\Phi}(P) \quad .$$

That is,

$$d_n \leq d \quad ,$$

or

$$\underline{\Phi}(P_n) + t_n\psi(P_n) \leq d \quad . \tag{61}$$

This last inequality can be weakened by dropping the non-negative term $t_n\psi(P_n)$, giving the inequality

$$\underline{\Phi}(P_n) \leq d \quad ,$$

which, together with the relation

$$\underline{\Phi}(P_\infty) \leq \underline{\lim}\ \underline{\Phi}(P_n)$$

expressing the lower semi-continuity of $\underline{\Phi}(P)$, implies that $\underline{\lim}\ \underline{\Phi}(P_n)$ is finite and that

$$\underline{\Phi}(P_\infty) \leq \underline{\lim}\, \underline{\Phi}(P_n) \leq d \quad ,$$

and, of course,

$$\underline{\Phi}(P_\infty) \leq d \quad . \tag{62}$$

Moreover, the inequality expressing the lower semi-continuity of $\underline{\Phi}(P)$ can be slightly weakened and written as a strict inequality as follows:

$$\underline{\Phi}(P_\infty) - 1 < \underline{\lim}\, \underline{\Phi}(P_n) \quad .$$

But, since $\underline{\Phi}(P_\infty) - 1$ is less than the lower limit of the set of numbers $\underline{\Phi}(P_n)$, it follows that there are at most a finite number of values of n for which

$$\underline{\Phi}(P_n) < \underline{\Phi}(P_\infty) - 1 \quad ,$$

or, in other words, there exists a positive number N such that

$$\underline{\Phi}(P_n) \geq \underline{\Phi}(P_\infty) - 1 \qquad \text{for } n > N \quad .$$

In view of (61), it then follows that

$$\underline{\Phi}(P_\infty) - 1 + t_n \psi(P_n) \leq d \quad \text{for } n > N$$

or

$$t_n \psi(P_n) \leq d - \underline{\Phi}(P_\infty) + 1 \quad \text{for } n > N \quad .$$

In other words, the sequence of non-negative numbers $t_n \psi(P_n)$ is bounded; and since $t_n \to \infty$ as $n \to \infty$, it follows that $\psi(P_n) \to 0$ as $n \to \infty$, and, clearly

$$\underline{\lim}\, \psi(P_n) = \lim_{n \to \infty} \psi(P_n) = 0 \quad .$$

However, the lower semi-continuity of $\psi(P)$ implies that

$$\psi(P_\infty) \leq \underline{\lim}\, \psi(P_n) = 0 \quad ,$$

while condition b) of the theorem insures that

$$\psi(P_\infty) \geq 0 \quad ,$$

from which it follows that

$$\psi(P_\infty) = 0 \quad .$$

Thus, P_∞ satisfies the side condition for problem A.

Since P_∞ satisfies the side condition for problem A, and since, of course,

$$\underline{\Phi}(P) \geq d$$

for all points satisfying the side condition, it follows, in particular, that

$$\underline{\Phi}(P_\infty) \geq d \quad .$$

But this relation together with (62) implies that

$$\underline{\Phi}(P_\infty) = d \quad .$$

Finally, this equality and the fact that P_∞ satisfies the side condition of problem A exactly expresses the fact that P_∞ solves A, completing the proof of Part I.

The proof of Part II is quite simple in view of the theorem which states that a lower semi-continuous function on a compact space possesses a least value. Let $\{\tau_n\}$ be any sequence of positive numbers such that $\tau_n \to \infty$ as $n \to \infty$. Then, since $\underline{\Phi}(P)$ and $\psi(P)$ are both lower semi-continuous, it follows that $\underline{\Phi}(P) + \tau_n\psi(P)$ is lower semi-continuous and therefore takes on its least value when P ranges over the compact space S; that is, there exists at least one point, say π_n, which solves A_{τ_n}. The compactness of S insures that the sequence of points π_n has a subsequence of points $\pi_{m(n)}$ which converges to a point in S, say P_∞. Let the point P_n be defined by

$$P_n = \pi_{m(n)} \quad ,$$

and let t_n be defined by

$$t_n = \tau_{m(n)} \quad .$$

It is clear that $t_n \to \infty$ as $n \to \infty$, P_n solves A_{t_n}, and $P_n \to P_\infty$ as $n \to \infty$. This completes the proof of Part II and the entire approximation theorem.

The approximation theorem can be used to prove the following form of the Lagrange multiplier rule:

THEOREM:

If $f(x,y)$ and $g(x,y)$ are continuous and possess continuous derivatives f_x, f_y, g_x, and g_y in an open region R containing the point (x_o,y_o) with

$$[g_x(x_o,y_o)]^2 + [g_y(x_o,y_o)]^2 > 0 \quad ,$$

and, moreover, if (x_o,y_o) furnishes a relative minimum for $f(x,y)$ among points in R satisfying the side condition

$$g(x,y) = 0 \quad ;$$

then there exists a number λ such that

$$f_x(x_o,y_o) + \lambda g_x(x_o,y_o) = 0$$

and

$$f_y(x_o,y_o) + \lambda g_y(x_o,y_o) = 0 \quad .$$

Before proving the theorem, it might first be remarked that it is good enough to show that the Jacobian $\frac{\partial(f,g)}{\partial(x,y)}$ vanishes at the point (x_o,y_o). The vanishing of this Jacobian together with the condition

$$[g_x(x_o,y_o)]^2 + [g_y(x_o,y_o)]^2 > 0$$

is clearly tantamount to the existence of a number λ with the required properties.

The proof of the vanishing of the Jacobian will be divided into two cases. The first and more important case is that in which there exists in the open region R an open sub-region G containing the point (x_o,y_o) such that

$$f(x,y) > f(x_o,y_o)$$

for all points in G other than (x_o,y_o) satisfying the side condition

$$g(x,y) = 0 \quad .$$

The essential feature for Case I is that (x_o,y_o) is the unique point in the open region G at which $f(x,y)$ takes on its least value among its values at points satisfying the side condition.

In order to apply the approximation theorem, let any closed bounded region G* contained in G with (x_o,y_o) as an interior point be identified with the space S. In this case, of course, S is compact and Part II of the approximation theorem guarantees the existence of numbers t_n and points P_n and P_∞ as specified in condition e) of Part I. Moreover, let $f(x,y)$ be identified with $\mathfrak{T}(P)$ and $[g(x,y)]^2$ with $\psi(P)$, the side condition

$$g(x,y) = 0$$

being equivalent to the side condition

$$[g(x,y)]^2 = 0 \quad .$$

Since problem A has a unique solution (x_o,y_o) in this case, it follows that P_∞ is the point (x_o,y_o). Further, let (x_n,y_n) represent the point P_n; that is, let (x_n,y_n) be a point at which $f(x,y) + t_n[g(x,y)]^2$ takes on its least value in G*.

Since $(x_n,y_n) \to (x_o,y_o)$ as $n \to \infty$ and (x_o,y_o) is an interior point of G*, it follows that there exists a positive number M such that (x_n,y_n) is an interior point of G* for $n > M$. But the fact that $f(x,y) + t_n[g(x,y)]^2$ takes on its least value at an interior point (x_n,y_n) implies that the point (x_n,y_n) is a stationary point. It follows that

$$f_x(x_n,y_n) + 2t_n g(x_n,y_n)g_x(x_n,y_n) = 0 \qquad \text{for } n > M$$

and

$$f_y(x_n,y_n) + 2t_n g(x_n,y_n)g_y(x_n,y_n) = 0 \qquad \text{for } n > M \quad ;$$

and, consequently,

$$\left.\frac{\partial(f,g)}{\partial(x,y)}\right|_{\substack{x=x_n \\ y=y_n}} = 0 \qquad \text{for } n > M$$

But the Jacobian is a continuous function since the partial derivatives of $f(x,y)$ and $g(x,y)$ have been assumed to be continuous. Therefore, letting n tend to infinity yields

$$\left.\frac{\partial(f,g)}{\partial(x,y)}\right|_{\substack{x=x_o \\ y=y_o}} = 0 \quad ,$$

completing the proof for Case I.

Case II is that in which there does not exist an open region G containing (x_o,y_o) such that

$$f(x,y) > f(x_o,y_o)$$

for all points in G other than (x_o,y_o) satisfying the side condition. However, the fact that (x_o,y_o) is a relative minimum point insures the existence of a neighborhood of (x_o,y_o) such that

$$f(x,y) \geq f(x_o,y_o)$$

for all points in the neighborhood satisfying the side condition. It follows that in Case II there exists in every neighborhood of (x_o,y_o) at least one point other than (x_o,y_o) which satisfies the side condition and at which

$$f(x,y) = f(x_o,y_o) \quad .$$

That is, (x_o,y_o) is an accumulation point of a set of points at every point of which

$$f(x,y) = f(x_o,y_o)$$

and

$$g(x,y) = 0 \quad .$$

Clearly, the mapping

$$u = f(x,y)$$
$$v = g(x,y)$$

is not one-to-one for neighborhoods of (x_0,y_0) and the inverse-function theorem insures that

$$\left.\frac{\partial(f,g)}{\partial(x,y)}\right|_{\substack{x=x_0\\ y=y_0}} = 0$$

Thus, the vanishing of the Jacobian for Case II has been proved without even using the approximation theorem, and the proof of the multiplier rule has been completed.

Another application of the approximation theorem is in justifying a procedure which adapts boundary value problems to the Rayleigh-Ritz method. For example, if admissible functions are required to vanish identically on the boundary of a region G, then the problem of minimizing the functional

$$I = \iint_G F(x,y,u,u_x,u_y)\ dxdy$$

can be attacked with the Rayleigh-Ritz method only if approximating functions which vanish on the boundary can be explicitly found. However, an approximation to the solution might be obtained by minimizing the functional

$$I_t = \iint_G F(x,y,u,u_x,u_y)\ dxdy + t\oint \bar{u}^2\ ds$$

for a sufficiently large positive value of t with no boundary conditions prescribed (where $\bar{u}$ denotes the values of u on the boundary of G). This problem can easily be handled with the Rayleigh-Ritz method since the approximating functions are not required to satisfy any boundary conditions. However, some guarantee is needed to insure that the solution to this problem approximates the solution to the original problem. In certain cases, the approximation theorem provides this guarantee. If the class of admissible functions should happen to be compact and if the minimum problem should happen to have a unique solution, then the approximation theorem can be applied directly.

The function $\Phi(P)$ can be identified with $\iint_G F(x,y,u,u_x,u_y)\,dxdy$ and $\psi(P)$ can be identified with $\int \bar{u}^2\,ds$. Then, the approximation theorem insures the existence of sequences of values for t approaching infinity such that the solution to the original problem is approached as the free boundary problem is solved for these values of t.

In addition, it might be attempted to use the approximation theorem to derive Lagrange multiplier rules for the isoperimetric problems of the Calculus of Variations, particularly with partial differential equation side conditions. The main object is, of course, to derive necessary conditions for an object to solve a minimum problem with side conditions by using only the approximation theorem and known necessary conditions for minimum problems of the same type without side conditions.