The Problem of Lagrange in the Calculus of Variations.

By Gilbert Ames Bliss.

TABLE OF CONTENTS.

INTRODUCTION.

The problem of the calculus of variations principally considered in this paper is that of finding in a class of arcs

$$(1) \qquad y_i = y_i(x) \qquad (x_1 \leqq x \leqq x_2;\ i = 1, \cdots, n)$$

satisfying a set of differential equations

$$(2) \qquad \phi_\alpha(x, y_1, \cdots, y_n, y_1', \cdots, y_n') = 0 \qquad (\alpha = 1, \cdots, m < n)$$

and joining two fixed points in the space of points $(x, y_1, \cdots, y_n)$, one which minimizes an integral of the form

$$(3) \qquad I = \int_{x_1}^{x_2} f(x, y_1, \cdots, y_n, y_1', \cdots, y_n')\, dx.$$

A number of paragraphs are also devoted to the similar problem for which the end-points are variable.

The problem seems to have been first formulated by Lagrange for the general case here studied, though somewhat less precisely than in the statement above. He also gave the multiplier rule described in Section 5 below which had been previously deduced by Euler and himself for a number of more special cases. Important additions to the theory have been made by Clebsch, A. Mayer, Kneser, Hilbert, von Escherich, Hahn, Bolza, and many others. Comprehensive treatments of the problem have been given by Bolza [3] * and Hadamard [4], that of Bolza being the more complete. In Chapter V below a brief sketch of the history of the problem is given with a bibliography of the more important papers on which the text of this paper is based.

Since the literature of the problem is extensive and widely scattered, and since recent developments make possible important simplifications, even as compared with the excellent treatments of Bolza and Hadamard, it seemed justifiable to the author of this paper to attempt anew the presentation of those parts of the theory leading to the necessary conditions for a minimum, and to those sufficient to insure a minimum. The paper is a record of lectures which the author has given at intervals for some years past at the University of Chicago.

Some special features of the methods used may perhaps be mentioned. The deduction of the Euler-Lagrange multiplier rule in Sections 3-5 is based upon suggestions in papers by Hahn [13, p. 271] and the author [16, pp. 307, 312], but is different from the proofs hitherto given. The definition of

* The figures in the square brackets refer to the bibliographical list at the end of the paper.

normal arcs in Sections 7 and 8 is that of Bolza [19, p. 440]. A new application of the definition, in Section 15, makes it possible to deduce without the use of special methods the multiplier rule for the case when the functions ϕ_α contain none of the derivatives y_i', as a corollary to the rule deduced in Section 5. The discussions of the necessary conditions of Weierstrass and Clebsch, and of the envelope theorem with the associated deduction of the necessary condition of Mayer, are essentially those of Hahn [21] and Bolza [3, pp. 603-10], but are greatly simplified by the use of the auxiliary formulas of Section 21. The analytic proof of the necessary condition of Mayer in Section 26, by means of the minimum problem associated with the second variation, was suggested by the author for simpler cases [27] and applied to the problem of Lagrange by D. M. Smith [28]. By means of the theory of the minimum problem of the second variation the very elaborate theories of that variation due to Clebsch [29], von Escherich [31], Hahn [33], and others, can be much simplified, as the author has shown [35]. The applications important for this paper are in Sections 26 and 32. The theory of Mayer fields in Sections 28 and 29, and the proofs of the sufficiency theorems in Sections 30 and 31, have been simplified as far as seemed possible.

An effort has been made in each theorem to state clearly the underlying hypotheses. The proof of the multiplier rule in Section 5, for example, is independent of the assumption that the determinant R of page 11 is different from zero. In many of the succeeding theorems, however, this assumption is either made explicitly or else is a consequence of the property III′ which appears frequently.

CHAPTER I.

THE EULER-LAGRANGE MULTIPLIER RULE.

1. *Hypotheses.* In this first chapter the famous multiplier rule of Euler and Lagrange, describing the differential equations satisfied by a minimizing arc for the problem of Lagrange stated in the introduction, is to be deduced. For convenience in the following pages the set $(x, y_1, \cdots, y_n, y_1, \cdots, y_n')$ will be represented by (x, y, y').

As usual we concentrate attention on a particular arc E_{12} with the equations (1) and inquire what properties it must have if it is to be a minimizing arc. The analysis is based upon the following hypotheses:

(a) the functions $y_i(x)$ defining E_{12} are continuous on the interval x_1x_2 and this interval can be subdivided into a finite number of parts on each of which the functions have continuous derivatives;

(b) in a neighborhood $\mathfrak{R}$ of the values (x, y, y') on the arc E_{12} the functions f, ϕ_α have continuous derivatives up to and including those of the **fourth order;**

(c) at every element (x, y, y') on E_{12} the $m \times n$-dimensional matrix $\| \phi_{\alpha y_i'} \|$ has rank m.

The subscript y_i' here indicates the partial derivative of ϕ_α with respect to y_i'. In the following pages literal subscripts, following the indices of functions and elsewhere, will be frequently used to indicate partial derivatives. The hypothesis (c) implies that the equations $\phi_\alpha = 0$ are all independent near E_{12} when regarded as functions of the variables y_i'.

2. *Examples.* A common example of a Lagrange problem is that of the brachistochrone in a resisting medium [3, p. 5]. The differential equation of the motion [5, p. 44] becomes for this case

$$dv/dt = d^2s/dt^2 = g(dy/ds) - R(v),$$

where $R(v)$ is the retardation on the particle per unit mass due to the resistance of the medium. Multiplying by $ds/dx = (ds/dt)(dt/dx) = vdt/dx$ we find the equation

$$vv' = gy' - R(v)s' = gy' - R(v)(1 + y'^2)^{1/2} \tag{4}$$

where the primes denote derivatives with respect to x. The problem is then to find among the pairs of functions $y(x)$, $v(x)$ which have the end-values

$$y(x_1) = y_1, \quad v(x_1) = v_1, \quad y(x_2) = y_2$$

and satisfy equation (4) one which minimizes the time integral

$$I = \int_{s_1}^{s_2} (ds/v) = \int_{x_1}^{x_2} (1/v)(1 + y'^2)^{1/2} dx.$$

It should be noted that this problem is not precisely like that stated in section 1 since the value of v is not prescribed at $x = x_2$. It is in fact a problem of Lagrange with second end-point variable.

The so-called isoperimetric problems form a very large class, and all of them may be stated as Lagrange problems. For example we may seek to find among the arcs $y = y(x)$ $(x_1 \leqq x \leqq x_2)$, joining two given points and having a given length, one which has its center of gravity the lowest. This is the problem of determining the form of a hanging chain suspended between two pegs at its ends. Analytically the problem is to find among the functions $y(x)$ $(x_1 \leqq x \leqq x_2)$ satisfying the conditions

$$y(x_1) = y_1, \quad y(x_2) = y_2, \quad \int_{x_1}^{x_2} (1 + y'^2)^{1/2} dx = l$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} y(1 + y'^2)^{1/2} dx. \tag{5}$$

This problem may be made over into one of the Lagrange type by introducing the new variable

$$z(x) = \int_{x_1}^{x} (1 + y'^2)^{1/2} dx$$

satisfying the differential equation $z' = (1 + y'^2)^{1/2}$. The problem is then to find among the pairs $y(x)$, $z(x)$ satisfying $y(x_1) = y_1$, $z(x_1) = 0$, $y(x_2) = y_2$, $z(x_2) = l$, $z' = (1 + y'^2)^{1/2}$ one which minimizes the integral (5).

More generally suppose we wish to find among the functions $y(x)$ satisfying

$$y(x_1) = y_1, \quad y(x_2) = y_2$$

$$\int_{x_1}^{x_2} g(x, y, y') dx = k, \quad \int_{x_1}^{x_2} h(x, y, y') dx = l$$

one which minimizes

$$I = \int_{x_1}^{x_2} f(x, y, y') dx. \tag{6}$$

The problem is equivalent to that of finding among the sets of functions $y(x)$, $u(x)$, $v(x)$ satisfying

$$\begin{aligned} &y(x_1) = y_1, \quad u(x_1) = 0, \quad v(x_1) = 0, \\ &y(x_2) = y_2, \quad u(x_2) = k, \quad v(x_2) = l, \\ &u' = g(x, y, y'), \qquad v' = h(x, y, y'), \end{aligned}$$

one which minimizes the integral (6). Evidently a similar transformation of the problem could be made no matter how many isoperimetric integrals were to have prescribed constant values.

These illustrations suffice to show the wide applicability of the Lagrange problem.

3. *Admissible arcs and variations.* An *admissible arc*

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2) \tag{7}$$

is one with the continuity properties (a) of Section 1, whose elements (x, y, y') all lie in the region $\mathfrak{R}$, and which satisfies the equations $\phi_\alpha = 0$. If a one-parameter family of admissible arcs

$$y_i = y_i(x, b) \qquad (i = 1, \cdots, n) \tag{8}$$

containing a particular admissible arc E_{12} for the parameter value $b = b_0$ is given, the functions

$$\eta_i(x) = y_{ib}(x, b_0) \qquad (i = 1, \cdots, n),$$

where the subscript b indicates as usual a partial derivative of $y_i(x, b)$, are called variations of the family along E_{12}.

In the tensor analysis it is agreed that a product $G_{ik}H_k$ shall stand for the sum $\Sigma_k G_{ik}H_k$. In other words, when an index k occurs twice in the same term it is understood that the term really represents the sum of n terms of the same type. The index with respect to which the sum is taken is called an umbral index.

With this convention in mind we may define for the arc E_{12} mentioned above the so-called *equations of variation* by the formula

$$\Phi_\alpha(x, \eta, \eta') = \phi_{\alpha y_i}\eta_i + \phi_{\alpha y_i'}\eta_i' = 0 \qquad (\alpha = 1, \cdots, m) \tag{9}$$

in which i is an umbral index with the range $1, \cdots, n$, and the coefficients $\phi_{\alpha y_i}$, $\phi_{\alpha y_i'}$ are supposed to have as arguments the functions $y_i(x)$ belonging to E_{12}. These equations are satisfied by the variations $\eta_i(x)$ along E_{12} as we may readily see by substituting the functions (8) in the equations $\phi_\alpha = 0$, differentiating for b, and setting $b = b_0$. A set of functions $\eta_i(x)$ with the continuity properties described in (a) of Section 1 and satisfying the equations of variation (9) is called a *set of admissible variations*, a nomenclature which is justified by the following very important theorem:

For every set of admissible variations $\eta_i(x)$ along the admissible arc E_{12} there exists a one parameter family (8) of admissible arcs containing E_{12} for the value $b = 0$ and having the functions $\eta_i(x)$ as its variations along E_{12}. For this family the functions $y_i(x, b)$ are continuous and have continuous derivatives with respect to b for all values (x, b) near those defining E_{12}, and the derivatives $y_{ix}(x, b)$ have the same property except possibly at the values of x defining corners of E_{12}.

To prove this theorem we enlarge the system $\phi_\alpha = 0$ to have the form

$$\phi_1 = 0, \cdots, \phi_m = 0, \quad \phi_{m+1} = z_{m+1}, \cdots, \phi_n = z_n \tag{10}$$

where $z_{m+1}, \cdots, z_n$ are new variables and $\phi_{m+1}, \cdots, \phi_n$ are new functions of x, y, y' such that the functional determinant $| \partial\phi_i/\partial y_k' |$ is different from zero along E_{12}.* By means of the last $n - m$ of these equations the functions

* For a proof of the possibility of this adjunction see Bliss [16, pp. 307, 312].

$y_i(x)$ belonging to E_{12} define a set of functions $z_r(x)$ $(r = m+1, \cdots, n)$. We have a corresponding system of equations of variation

$$(11) \qquad \Phi_1 = 0, \cdots, \Phi_m = 0, \quad \Phi_{m+1} = \zeta_{m+1}, \cdots, \Phi_n = \zeta_n$$

along E_{12}, the last $n - m$ of which define a set $\zeta_r(x)$ $(r = m+1, \cdots, n)$ corresponding to every set of admissible variations $\eta_i(x)$.

Suppose now that the set $\eta_i(x)$ is an admissible set of variations for E_{12} defining a set $\zeta_r(x)$ by means of equations (11). Since the functional determinant $|\partial\phi_i/\partial y_k'|$ is different from zero along E_{12} the existence theorems for differential equations * tell us that the system

$$(12) \quad \phi_\alpha = 0, \quad \phi_r = z_r(x) + b\zeta_r(x) \quad (\alpha = 1, \cdots, m;\ r = m+1, \cdots, n)$$

determines uniquely a one-parameter family of solutions $y_i = y_i(x, b)$ with the initial values $y_i(x_1) + b\eta_i(x_1)$ at $x = x_1$. This family contains E_{12} for $b = 0$ and has variations which have the initial values $\eta_i(x_1)$ at $x = x_1$ and which satisfy the equations (11) with the functions $\zeta_r(x)$. The variations of the family are therefore identical with the functions $\eta_i(x)$ originally prescribed, since when the $\zeta_r(x)$ are given there is only one set of solutions of equations (11) with given initial values $\eta_i(x_1)$ at $x = x_1$.

Some slight modifications in the existence theorems referred to are required in order to prove the continuity properties of the family $y_i = y_i(x, b)$ described in the theorem. These are due to the fact that the functions $z_i(x)$ defined by the arc E_{12} are continuous but not necessarily differentiable. The results described can be derived without difficulty, however, when the arc E_{12} has no corners. If the arc E_{12} has corners the existence theorems must be applied successively to the x-intervals between the corner-values of x with initial conditions at the beginning of each interval so chosen that the functions $y_i(x, b)$ are continuous.

COROLLARY. *If a matrix*

$$\begin{Vmatrix} \eta_{11} & \cdots & \eta_{1\mu} \\ \cdot & \cdots & \cdot \\ \eta_{n1} & \cdots & \eta_{n\mu} \end{Vmatrix}$$

whose columns are μ sets of admissible variations along an admissible arc E_{12}, is given, then there exists a μ-parameter family of admissible arcs $y_i = y_i(x, b_1, \cdots, b_\mu)$ containing E_{12} for the values $b_1 = \cdots = b_\mu = 0$ and having the functions η_{is} $(i = 1, \cdots, n)$ as its variations with respect to b_s along E_{12}. The continuity properties of the family are similar to those described in the preceding theorem.

---

* Bolza [3, pp. 168 ff.]; Bliss [14, 15].

This is proved as above with the equations

$$\phi_a = 0, \quad \phi_r = z_r(x) + b_1\zeta_{r1} + \cdots + b_\mu\zeta_{r\mu}$$
$$(\alpha = 1, \cdots, m; \; r = m+1, \cdots, n)$$

replacing equations (12).

4. *The first variation of I.* If the functions $y_i(x, b)$ defining a one-parameter family of admissible arcs containing E_{12} for $b = 0$ are substituted in I then I becomes the function of b defined by the formula

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

The derivative of this function with respect to b at the value $b = 0$ is the expression

$$(13) \qquad I_1(\eta) = \int_{x_1}^{x_2} (f_{y_i}\eta_i + f_{y_i'}\eta_i')\,dx$$

where i is as agreed an umbral symbol and the arguments of the derivatives of f are the functions $y_i(x)$ defining E_{12}.

The expression $I_1(\eta)$ is called the *first variation* of I along the arc E_{12}. For the proofs of the succeeding sections it is desirable to have another form of it. Let λ_0 be a constant and $\lambda_i(x)$ $(i = 1, \cdots, n)$ functions of x on the interval x_1x_2, and let F be defined by the equation

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_1\phi_1 + \cdots + \lambda_n\phi_n.$$

Since the variations η, ζ satisfy the equations (11) the value of $\lambda_0 I_1(\eta)$ is not altered if we add the sum $\lambda_a\Phi_a + \lambda_r(\Phi_r - \zeta_r)$ to its integrand. Then we have

$$(14) \qquad \lambda_0 I_1(\eta) = \int_{x_1}^{x_2} (F_{y_i}\eta_i + F_{y_i'}\eta_i' - \lambda_r\zeta_r)\,dx.$$

So far the functions $\lambda_i(x)$ have been entirely arbitrary. We now determine them so that the equations

$$(15) \qquad F_{y_i'} = \int_{x_1}^{x} F_{y_i}\,dx + c_i \qquad (i = 1, \cdots, n)$$

are satisfied for an arbitrarily selected set of constants λ_0, c_i. This is possible since if we introduce the new variables

$$(16) \qquad v_i = F_{y_i'} = \lambda_0 f_{y_i'} + \lambda_1\phi_{1y_i'} + \cdots + \lambda_n\phi_{ny_i'} \qquad (i = 1, \cdots, n)$$

the equations (15) are equivalent to the equations and initial conditions

$$(17) \qquad dv_i/dx = F_{y_i} = A_{i1}v_1 + \cdots + A_{in}v_n + B_i, \quad v_i(x_1) = c_i \qquad (i = 1, \cdots, n)$$

the coefficients A, B being found by solving the equations (16) for $\lambda_1, \cdots, \lambda_n$ and substituting in F_{y_i}. The equations (17) have unique solutions $v_i(x)$ which are continuous on the interval x_1x_2 and which have continuous derivatives except possibly at the values of x defining the corners of E_{12} where the coefficients A, B may be discontinuous. Equations (16) then determine uniquely the functions $\lambda_i(x)$ continuous except possibly at the corner values of x.

With the help of equations (15) the expression (14) for $\lambda_0 I_1(\eta)$ now takes the form

$$(18) \qquad \lambda_0 I_1(\eta) = -\int_{x_1}^{x_2} \lambda_r \zeta_r dx - c_i \eta_i(x_1) + \eta_i(x_2) F_{y_i'}(x_2)$$

where $F_{y_i'}(x_2)$ represents the value of $F_{y_i'}$ at $x = x_2$. This auxiliary formula will be useful in the next section.

5. *The Euler-Lagrange multiplier rule.* We are now in a position to deduce the famous multiplier rule giving the differential equations which must be satisfied by a minimizing arc E_{12} for the Lagrange problem. The rule was discussed for a special case by Euler in 1744, and generalized by Lagrange whose proof was exceedingly faulty. One difficulty with Lagrange's proof was overcome by Mayer in 1886, and the proof was finally completed when Kneser in 1900 and Hilbert in 1905 removed the last serious defects.* The proof given here is quite different in some respects from those in the literature and is an extension of them.

Suppose that a matrix whose columns are $2n + 1$ sets of admissible variations

$$(19) \qquad \left\| \begin{matrix} \eta_{11} & \cdots & \eta_{1,2n+1} \\ \cdot & \cdots & \cdot \\ \eta_{n1} & \cdots & \eta_{n,2n+1} \end{matrix} \right\|$$

is given. We have seen above that there is a $(2n+1)$-parameter family $y_i(x, b_1, \cdots, b_{2n+1})$ of admissible arcs containing E_{12} for $b_1 = \cdots = b_{2n+1} = 0$ and having the columns of the matrix above as its variations. When the functions defining this family are inserted in the integral I that integral becomes a function $I(b_1, \cdots, b_{2n+1})$ which for $b_1 = \cdots = b_{2n+1} = 0$ takes the value I_0 of the integral along the arc E_{12}. If we let $(x_1, y_{11}, \cdots, y_{n1})$ and $(x_2, y_{12}, \cdots, y_{n2})$ represent the two end-points on the arc E_{12} then the equations

* For the details of the objections to Lagrange's proof and an excellent historical sketch see Bolza [3, p. 566].

$$
\begin{aligned}
&I(b_1, \cdots, b_{2n+1}) = I_0 + u, \\
(20) \qquad &y_i(x_1, b_1, \cdots, b_{2n+1}) = y_{i1}, \\
&y_i(x_2, b_1, \cdots, b_{2n+1}) = y_{i2}, \\
&\qquad (i = 1, \cdots, n)
\end{aligned}
$$

in the variables $u, b_1, \cdots, b_{2n+1}$ have the initial solution $(u, b_1, \cdots, b_{2n+1}) = (0, 0, \cdots, 0)$. If the functional determinant of the first members of these equations with respect to $b_1, \cdots, b_{2n+1}$ is different from zero at this solution, then well-known implicit function theorems tell us that the equations (20) have solutions not only for $u = 0$ but also for every value of u near $u = 0$. There are therefore arcs in the family $y_i(x, b_1, \cdots, b_{2n+1})$ joining the end-points 1 and 2 of E_{12} and giving I values $I_0 + u$ greater than I_0 when u is positive, and similar arcs giving it values less than I_0 when u is negative, which is impossible if E_{12} is a minimizing arc. Hence the functional determinant of the equations (20) must be zero at $(u, b_1, \cdots, b_{2n+1}) = (0, 0, \cdots, 0)$.

The value of this functional determinant is

$$
(21) \qquad \begin{Vmatrix}
I_1(\eta_1) & \cdots & I_1(\eta_{2n+1}) \\
\eta_{11}(x_1) & \cdots & \eta_{1,2n+1}(x_1) \\
\cdot & \cdots & \cdot \\
\eta_{n1}(x_1) & \cdots & \eta_{n,2n+1}(x_1) \\
\eta_{11}(x_2) & \cdots & \eta_{1,2n+1}(x_2) \\
\cdot & \cdots & \cdot \\
\eta_{n1}(x_2) & \cdots & \eta_{n,2n+1}(x_2)
\end{Vmatrix}
$$

where in the first row only the second subscripts of the η's are indicated. It must vanish for every choice of the matrix (19) of admissible variations. Suppose $p < 2n + 1$ the highest rank attainable for (21) and suppose the matrix (19) chosen so that this rank is actually attained. Let λ_0, c_i, d_i $(i = 1, \cdots, n)$ be a set of constants not all zero satisfying the linear equations whose coefficients are the columns of the determinant (21). Normally the constant λ_0 will be different from zero, but in Section 7 the case $\lambda_0 = 0$ is discussed in more detail. In both cases the equation

$$
\lambda_0 I_1(\eta) + c_i \eta_i(x_1) + d_i \eta_i(x_2) = 0
$$

must be satisfied for every set of admissible variations $\eta_i(x)$ whatsoever, since otherwise by deleting a suitable one of the columns of the determinant (21) and replacing it by a set $I_1(\eta)$, $\eta_i(x_1)$, $\eta_i(x_2)$ which does not satisfy the last equation, the determinant could be made to have the rank $p + 1$. If the first term of the last equation is replaced by its value (18) the equation takes the form

$$-\int_{x_1}^{x_2} \lambda_r \zeta_r dx + \eta_i(x_2)[d_i + F_{y_i'}(x_2)] = 0$$

and it must be satisfied for every choice of the admissible variations $\eta_i(x)$, i. e. for every choice of the functions $\zeta_r(x)$ and the end values $\eta_i(x_2)$, since for every such choice there is a set of admissible variations defined by the equations (11). It follows readily that the conditions

$$(22) \qquad \lambda_r(x) \equiv 0, \quad d_i = -F_{y_i'}(x_2) \qquad (r = m+1, \cdots, n;\ i = 1, \cdots, n)$$

must be satisfied. For the set of multipliers λ_0, $\lambda_i(x)$ $(i = 1, \cdots, n)$ for which the equations (15) are satisfied it is evident then that all are identically zero except the first $m+1$. The first $m+1$ of them are not all identically zero, however, since otherwise F would vanish identically and equations (15) and (22) would require the constants c_i, d_i all to be zero as well as λ_0, which we know not to be the case. Hence we have the following theorem:

For every minimizing arc E_{12} there exists a set of constants c_i $(i = 1, \cdots, n)$ and a function

$$(23) \qquad F(x, y, y', \lambda) = \lambda_0 f + \lambda_1(x)\phi_1 + \cdots + \lambda_m(x)\phi_m$$

such that the equations

$$(24) \qquad F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i$$

are satisfied at every point of E_{12}. The constant λ_0 and the functions $\lambda_\alpha(x)$ $(\alpha = 1, \cdots, m)$ are not all identically zero on x_1x_2 and are continuous except possibly at values of x defining corners of E_{12}.

This is a modification of the Euler-Lagrange multiplier rule. We get the rule in its classical form by differentiating the equations (24). The two following corollaries are immediate:

COROLLARY I. THE EULER-LAGRANGE MULTIPLIER RULE. *On every sub-arc between corners of a minimizing arc E_{12} the differential equations*

$$(25) \quad \phi_\alpha(x, y, y') = 0, \quad (d/dx)F_{y_i'} = F_{y_i} \quad (\alpha = 1, \cdots, m;\ i = 1, \cdots, n)$$

must be satisfied, where F is the function (23).

COROLLARY II. THE CORNER CONDITION. *At every corner of a minimizing arc E_{12} the conditions*

$$(26) \quad F_{y_i'}[x, y, y'(x-0), \lambda(x-0)] = F_{y_i'}[x, y, y'(x+0), \lambda(x+0)] \qquad (i = 1, \cdots, n)$$

must be satisfied.

Condition (26) is a consequence of the fact that the second member of (24) is continuous at a corner as well as elsewhere.

There is a third consequence of the equations (24) which is also important. If the functions and multipliers belonging to E_{12} are $y_i(x)$, λ_0, $\lambda_a(x)$ then the $n+m$ equations

$$F_{y_i'}[x, y(x), z, \mu] = \int_{x_1}^{x} F_{y_i}[x, y(x), y'(x), \lambda(x)]dx + c_i,$$

$$\phi_a[x, y(x), z] = 0 \qquad (i = 1, \cdots, n;\ \alpha = 1, \cdots, m)$$

have as solutions the $n+m$ functions $z_i = y_i'(x)$, $\mu_a = \lambda_a(x)$. If the functional determinant

$$R = \begin{vmatrix} F_{y_i' y_k'} & \phi_{a y_i'} \\ \phi_{a y_k'} & 0 \end{vmatrix}$$

of the first members of these equations with respect to the variables z_i, μ_a is different from zero at a point of E_{12} then the existence theorems for implicit functions tell us that the solutions $z_i = y_i'(x)$, $\mu_a = \lambda_a(x)$ of the equations have continuous derivatives of as many orders as the equations themselves have continuous partial derivatives in the variables x, z_i, μ_a. Between corners this is at least one, and we have the following third corollary:

COROLLARY III. THE DIFFERENTIABILITY CONDITION. *Near a point of a minimizing arc E_{12} at which the determinant R is different from zero the functions $y_i(x)$ defining E_{12} have continuous second derivatives and the multipliers $\lambda_a(x)$ have continuous first derivatives.*

The proof given above for the Euler-Lagrange multiplier rule is an extension of the ones ordinarily given because the hypothesis (c) Section 1 is less restrictive than usual. The unsymmetrical assumption commonly made is that a particular one of the determinants of the matrix $\| \phi_{a y_i'} \|$ stays different from zero at every point of E_{12}. The enlargement of the system $\phi_a = 0$ to the system (10) is the device which permits the generalization here made. Equations (24) are recent developments which were unknown to Euler and Lagrange and which are not always deduced even in modern presentations of the subject. They justify the useful Corollaries II and III besides the multiplier rule.

6. *The extremals.* An admissible arc and set of multipliers

$$(27) \qquad y_i = y_i(x), \quad \lambda_0, \quad \lambda_a = \lambda_a(x) \qquad (i = 1, \cdots, n;\ \alpha = 1, \cdots, m;\ x_1 \leqq x \leqq x_2)$$

is called an *extremal* if it has continuous derivatives $y_i'(x)$, $y_i''(x)$, $\lambda_a'(x)$

on the interval x_1x_2, and if furthermore it satisfies the Euler-Lagrange equations (25). The minimizing curves for applications of the theory of the calculus of variations are found among the extremals and it is highly desirable, therefore, that we should examine more thoroughly the differential equations defining these curves and determine how large a family the extremals really form. A minimizing curve must always be a solution of the equations (25), even if it has corners or is without the derivatives $y_i''(x)$, $\lambda_\alpha'(x)$ mentioned above, but such minimizing curves are relatively rare.

The most direct way to characterize the family of extremals satisfying equations (25) is to replace these equations by the equivalent system

$$
\begin{aligned}
&(d/dx)F_{y_i'} - F_{y_i} = F_{y_i'x} + F_{y_i'y_k}y_k' + F_{y_i'y_k'}y_k'' + F_{y_i'\lambda_\beta}\lambda_\beta' - F_{y_i} = 0,\\
(28)\qquad &(d/dx)\phi_\alpha = \phi_{\alpha x} + \phi_{\alpha y_k}y_k' + \phi_{\alpha y_k'}y_k'' = 0,\\
&\phi_\alpha\,[x_1, y(x_1), y'(x_1)] = 0.
\end{aligned}
$$

The first two of these equations are linear in the variables y_k'', λ_β' and the determinant of the coefficients of these variables is precisely the determinant R of page 684. Near an extremal E_{12} on which R is different from zero these two equations can therefore be solved for y_k'', λ_β' and they are readily seen to be equivalent to a system

$$(29)\qquad dy_k/dx = y_k', \quad dy_k'/dx = G_k(x, y, y', \lambda), \quad d\lambda_\beta/dx = H_\beta(x, y, y', \lambda)$$

in the so-called normal form.* Known existence theorems for differential equations now tell us that an extremal E_{12} along which R is different from zero is a member of a family of solutions of equations (29) depending upon $2n + m$ arbitrary constants, since the number of dependent variables y_k, y_k', λ_β in these equations is $2n + m$. If we impose further the m relations in the third row of equations (28) then m of these constants will be determined as function of the $2n$ others, so that the final result is that an extremal along which R is different from zero is a member of a $2n$-parameter family of extremals satisfying equations (25).

For theoretical purposes the properties of the $2n$-parameter family of extremals may be determined most conveniently by a second method.† For the purpose of introducing n new variables v_i and eliminating the $n + m$ variables y_i', λ_α let us consider the system of $n + m$ equations

$$(30)\qquad F_{y_i'}(x, y, y', \lambda) = v_i, \quad \phi_\alpha(x, y, y') = 0.$$

The functional determinant of the first members of these equations with respect

* Bolza [3, p. 589].

† Bolza [3, p. 590].

to the variables y_k', λ_β is again the determinant R of page 684. Known theorems on implicit functions tell us then that near an extremal E_{12} on which R is different from zero the equations (30) have solutions

$$y_k' = \Psi_k(x, y, v), \quad \lambda_\beta = \Pi_\beta(x, y, v) \tag{31}$$

possessing continuous partial derivatives of the first three orders since the first members of equations (30) have such derivatives. The system of equations (25) is now equivalent to the system in normal form

$$dy_k/dx = \Psi_k(x, y, v), \quad dv_k/dx = F_{y_i}[x, y, \Psi(x, y, v), \quad \Pi(x, y, v)] \tag{32}$$

in the variables x, y_k, v_k. Evidently every solution $y_k(x)$, $\lambda_\beta(x)$ of equations (25) defines a set of functions $v_k(x)$ satisfying equations (30) and (31), and therefore also the system (32). Conversely every solution $y_k(x)$, $v_k(x)$ of equations (32) defines a set of functions $\lambda_\beta(x)$ by means of equations (31) with which it satisfies equations (30); and therefore also the original system (25).

Through every initial element

$$(x_0, y_0, v_0) = (x_0, y_{10}, \cdots, y_{n0}, v_{10}, \cdots, v_{n0})$$

in a neighborhood of the set of values (x, y, v) on the extremal E_{12} there passes a unique solution

$$y_i = y_i(x, x_0, y_0, v_0), \quad v_i = v_i(x, x_0, y_0, v_0) \tag{33}$$

of the equations (32) for which the functions y_i, y_{ix}, v_i, v_{ix} have continuous partial derivatives of the first three orders since the second members of equations (32) have such derivatives. The equations expressing the fact that the solutions (33) passes through (x_0, y_0, v_0) are

$$y_{i0} = y_i(x_0, x_0, y_0, v_0), \quad v_{i0} = v_i(x_0, x_0, y_0, v_0),$$

and from them we find

$$\begin{aligned} \delta_{ik} &= (\partial/\partial y_{k0})\, y_i(x_0, x_0, y_0, v_0), & 0 &= (\partial/\partial v_{k0})\, y_i(x_0, x_0, y_0, v_0), \\ 0 &= (\partial/\partial y_{k0})\, v_i(x_0, x_0, y_0, v_0), & \delta_{ik} &= (\partial/\partial v_{k0})\, v_i(x_0, x_0, y_0, v_0), \end{aligned} \tag{34}$$

where δ_{ik} is 1 or 0 when $k = i$ or $k \neq i$, respectively. Since every curve of this system (33) has on it an initial element for which $x = x_1$ we lose none of the curves if we replace x_0 by the fixed value x_1. Let us for convenience rename the constants y_{i0}, v_{i0} and call them a_i, b_i respectively. Then the family (33) takes the form

$$y_i = y_i(x, a, b), \quad v_i = v_i(x, a, b) \tag{35}$$

and it follows readily from equations (34) that the determinant

$$(36) \qquad \begin{vmatrix} \dfrac{\partial y_i}{\partial a_k} & \dfrac{\partial y_i}{\partial b_k} \\ \dfrac{\partial v_i}{\partial a_k} & \dfrac{\partial v_i}{\partial b_k} \end{vmatrix}$$

has the value 1 at $x = x_1$. When we substitute the functions (35) in equations (31) a set of functions $\lambda_\alpha(x, a, b)$ is determined, and we have the final result:

Every extremal E_{12} along which the determinant R is different from zero is a member of a $2n$-parameter family of extremals

$$(37) \qquad y_i = y_i(x, a, b), \quad \lambda_\alpha = \lambda_\alpha(x, a, b)$$

for special values a_0, b_0 of the parameters. The functions y_i, y_{ix}, v_i, v_{ix}, λ_β have continuous partial derivatives of the first three orders in a neighborhood of the values (x, a, b) defining E_{12}, and at the special values (x_1, a_0, b_0) the determinant (36) *is different from zero.*

Thus again we have established the existence of a family of extremals containing $2n$ arbitrary constants.

7. *Normal admissible arcs.* An admissible arc $y_i = y_i(x)$ $(x_1 \leqq x \leqq x_2)$ is said to be *normal* if there exist for it $2n$ sets of admissible variations for which the determinant

$$(38) \qquad \begin{vmatrix} \eta_{11}(x_1) & \cdot & \cdot & \cdot & \eta_{1,2n}(x_1) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1}(x_1) & \cdot & \cdot & \cdot & \eta_{n,2n}(x_1) \\ \eta_{11}(x_2) & \cdot & \cdot & \cdot & \eta_{1,2n}(x_2) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1}(x_2) & \cdot & \cdot & \cdot & \eta_{n,2n}(x_2) \end{vmatrix}$$

is different from zero. It is *normal on a sub-interval* $\xi_1\xi_2$ of x_1x_2 if there exist $2n$ sets of admissible variations for which the last determinant is different from zero when x_1 is replaced by ξ_1 and x_2 by ξ_2. In the sequel we shall frequently need to restrict our proofs to arcs which are *normal on every sub-interval of* x_1x_2.

These definitions doubtless seem at first sight somewhat artificial. If an admissible arc E_{12} is not normal, however, it is in general true that no other admissible arcs near it pass through the end points 1 and 2 of E_{12}, and hence that near E_{12} the class of arcs in which we seek to minimize the integral I has in it only E_{12} itself. The minimum problem in such a case would not be of interest. We shall presently see that there are always an infinity of admissible arcs through the ends of E_{12} when E_{12} is normal.

A necessary and sufficient condition that an admissible arc be normal is that there exists for it no set of multipliers λ_0, $\lambda_\alpha(x)$ having $\lambda_0 = 0$ with which it satisfies the equations.

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i.$$

For a normal extremal arc multipliers in the form $\lambda_0 = 1$, $\lambda_\alpha(x)$ always exist and in this form they are unique.

The processes of Section 5 show that an admissible arc which is not normal has surely a set of multipliers with $\lambda_0 = 0$, since the linear equations whose coefficients are the columns of the determinant (21) have for such an arc a set of solutions λ_0, c_i, d_i with $\lambda_0 = 0$. The first sentence of the theorem will then be justified if we can show that a normal admissible arc has no set of multipliers with $\lambda_0 = 0$.

Suppose that there were a normal admissible arc with a set of multipliers having $\lambda_0 = 0$. Its function F would have the form

$$F = \lambda_1 \phi_1 + \cdot \cdot \cdot \lambda_m \phi_m$$

and every set of admissible variations along it would satisfy the equation

$$0 = \int_{x_1}^{x_2} \lambda_\alpha \Phi_\alpha dx = \int_{x_1}^{x_2} (F_{y_i} \eta_i + F_{y_i'} \eta_i') dx = F_{y_i'}(x_2) \eta_i(x_2) - F_{y_i'}(x_1) \eta_i(x_1)$$

on account of the equations of variations (9) and the equations of the theorem above. Since there is a determinant (38) different from zero it follows that the derivatives $F_{y_i'}$ would all vanish at x_1 and x_2 on our extremal. If we define the variables v_i again by equations (30), or by equations (16) with $\lambda_0 = \lambda_{m+1} = \cdot \cdot \cdot = \lambda_n = 0$, then in equations (17) the coefficients B_i and the initial values $v_i(x_1) = F_{y_i'}(x_1)$ would all vanish. The only continuous solutions of equations (17) under these circumstances are the functions $v_i(x) \equiv 0$, and equations (16) then imply that the multipliers $\lambda_\alpha(x)$ would all vanish identically, which is not the case. Hence a normal admissible arc can not have a set of multipliers with constant multiplier λ_0 equal to zero.

When an extremal arc has multipliers with $\lambda_0 \neq 0$ the multipliers can evidently all be divided by λ_0 to obtain a set of the form $\lambda_0 = 1$, $\lambda_\alpha(x)$. If there were a second set $\lambda_0 = 1$, $\Lambda_\alpha(x)$ the differences 0, $\Lambda_\alpha - \lambda_\alpha$ would also be a set of multipliers for E_{12} with the constant multiplier zero. We have just seen that this is impossible for a normal extremal unless $\Lambda_\alpha - \lambda_\alpha \equiv 0$, so that the multipliers $\lambda_0 = 1$, $\lambda_\alpha(x)$ of a normal extremal E_{12} are unique.

In every neighborhood of a normal admissible arc E_{12} there are an infinity of admissible arcs with the same end-points 1 *and* 2.

To prove this consider the set of $2n$ admissible variations for E_{12} appearing in the determinant (38) and an additional set $\eta_i(x)$. From the results of Section 3 we know that there is a family of admissible arcs $y_i = Y_i(x, b, b_1, b_2, \cdots, b_{2n})$ containing E_{12} when $b = b_1 = \cdots = b_{2n} = 0$ and having the sets $\eta_i(x)$, $\eta_{is}(x)$ $(s = 1, \cdots, 2n)$ as its variations. The $2n$ equations

$$(39) \qquad Y_i(x_1, b, b_1, \cdots, b_{2n}) = y_{i1}, \quad Y_i(x_2, b, b_1, \cdots, b_{2n}) = y_{i2}$$

have the initial solution $(b, b_1, \cdots, b_{2n}) = (0, 0, \cdots, 0)$ at which the functional determinant of their first members with respect to $b_1, \cdots, b_{2n}$ is the determinant (38) and different from zero. Hence by the usual implicit function theorems these equations have solutions $b_s = B_s(b)$ $(s = 1, \cdots, 2n)$ with initial values $B_s(0) = 0$, and the one parameter family of admissible arcs

$$(40) \qquad y_i = Y_i[x, b, B_1(b), \cdots, B_{2n}(b)] = y_i(x, b)$$

defined by them contains the extremal E_{12} for $b = 0$ and has all its curves passing through the points 1 and 2.

COROLLARY. *If each function $\eta_i(x)$ of a set of admissible variations for a normal admissible arc E_{12} vanishes at x_1 and x_2 then there is a one-parameter family of admissible arcs $y_i = y_i(x, b)$ passing through the points* 1 *and* 2, *containing E_{12} for the parameter value $b = 0$, and having the set $\eta_i(x)$ as its variations along E_{12}.*

Let us suppose that in the construction of the family (40) the set $\eta_i(x)$ of the Corollary has been used. Since these functions all vanish at x_1 and x_2 we find from equations (39), by differentiating with respect to b and setting $b = 0$, that

$$\eta_{is}(x_1) B_s'(0) = 0, \quad \eta_{is}(x_2) B_s'(0) = 0.$$

Since the determinant (38) is different from zero these imply that all the derivatives $B_s'(0)$ vanish. Hence the family (40) has the variations

$$y_{ib}(x, 0) = \eta_i(x) + Y_{ib_s} B_s'(0) = \eta_i(x).$$

We know already that the family contains E_{12} for $b = 0$ and has all of its curves passing through 1 and 2.

8. *Problems with variable end-points.** It happens that a number of important applications of the theory of the Lagrange problem are of a slightly

* See Bliss [16].

different type from that described in Section 1. In order to include them as special cases we must permit variable end-points for the curves of the class in which we are seeking a minimum for I. We shall endeavor to find among the arcs

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2)$$

satisfying the system of equations

$$\phi_\alpha(x, y, y') = 0 \qquad (\alpha = 1, \cdots, m < n)$$

and having end-points satisfying the equations

$$\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \qquad (\mu = 1, \cdots, p \leqq 2n + 2) \tag{41}$$

one which minimizes the integral I. The number p must not exceed the number $2n + 2$ of end values x_1, y_{i1}, x_2, y_{i2} since otherwise equations (41) would in general have no solutions. The problem of Section 1 is a special case of this one with the system (41) having the special form

$$x_1 - \alpha_1 = y_{i1} - \beta_{i1} = x_2 - \alpha_2 = y_{i2} - \beta_{i2} = 0$$

for which p has exactly the value $2n + 2$.

Suppose now that E_{12} is a minimizing arc for the new problem with end values $(x_1, y_{i1}, x_2, y_{i2})$. We add to the hypotheses (a), (b), (c) of Section 1 the assumption

(d) the functions ψ_μ have continuous derivatives up to and including those of the fourth order near the end-values $(x_1, y_{i1}, x_2, y_{i2})$ of E_{12}, and at these values the $p \times (2n + 2)$-dimensional matrix

$$\| \psi_{\mu x_1} \quad \psi_{\mu y_{i1}} \quad \psi_{\mu x_2} \quad \psi_{\mu y_{i2}} \| \tag{42}$$

has rank p.

The last part of this assumption implies that the equations $\psi_\mu = 0$ are all independent.

It is evident that the arc E_{12} must minimize I in the class of admissible arcs having the same end-values, and we can infer at once that it must have a system of multipliers with which it satisfies the necessary conditions deduced in Section 5. But it is important that we should analyse the situation somewhat more closely. Let

$$y_i = y_i(x, b) \qquad [x_1(b) \leqq x \leqq x_2(b)] \tag{43}$$

be a one-parameter family of admissible arcs containing E_{12} for $b = 0$ whose end-values satisfy the equations

$$\psi_\mu\{x_1(b), y_i[x_1(b), b], x_2(b), y_i[x_2(b), b]\} = 0.$$

If we use the notations $x_{1b}(0) = \xi_1$, $x_{2b}(0) = \xi_2$ the derivatives of these equations with respect to b for $b = 0$ are the system

$$\begin{aligned}\Psi_\mu(\xi, \eta) = (\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1})\xi_1 + \psi_{\mu y_{i1}}\eta_i(x_1) \\ + (\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2})\xi_2 + \psi_{\mu y_{i2}}\eta_i(x_2).\end{aligned} \tag{44}$$

These are the *equations of variation* on E_{12} for the functions ψ_μ. When the family (43) is substituted in the integral I we find for the first variation the formula

$$I_1(\xi, \eta) = \int_{x_1}^{x_2} (f_{y_i}\eta_i + f_{y_i'}\eta_i')\,dx + f(x_2)\xi_2 - f(x_1)\xi_1$$

where $f(x_1)$ and $f(x_2)$ are the values of f at the points 1 and 2 on E_{12}. With the help of the expression (18) we may also write

$$\begin{aligned}\lambda_0 I_1(\xi, \eta) = -\int_{x_1}^{x_2} \lambda_r \zeta_r\,dx - \lambda_0 f(x_1)\xi_1 \\ - c_i\eta_i(x_1) + \lambda_0 f(x_2)\xi_2 + \eta_i(x_2) F_{y_i'}(x_2)\end{aligned} \tag{45}$$

where the constants c_i may be arbitrarily chosen.

A set of admissible variations for the present problem is a set ξ_1, ξ_2, $\eta_i(x)$ in which ξ_1 and ξ_2 are arbitrary constants and the functions $\eta_i(x)$ form a set of admissible variations in the sense of Section 3. For a matrix

$$\left\| \begin{matrix} \xi_{11} & \cdots & \xi_{1,p+1} \\ \xi_{21} & \cdots & \xi_{2,p+1} \\ \eta_{11} & \cdots & \eta_{1,p+1} \\ \cdot & \cdots & \cdot \\ \eta_{n1} & \cdots & \eta_{n,p+1} \end{matrix} \right\|$$

whose columns are sets of admissible variations there exists a family

$$\begin{aligned} y_i = y_i(x, b_1, \cdots, b_{p+1}) \\ x_1(b_1, \cdots, b_{p+1}) \leqq x \leqq x_2(b_1, \cdots, b_{p+1}) \end{aligned} \tag{46}$$

containing E_{12} for $(b_1, \cdots, b_{p+1}) = (0, \cdots, 0)$ and having the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}(x)$ $(\sigma = 1, \cdots, p+1)$ as its variations along E_{12} with respect to the parameters b_σ. Such a family is that of the Corollary on page 679 with the functions

$$x_\rho(b_1, \cdots, b_{p+1}) = x_\rho + b_\sigma \xi_{\rho\sigma}, \qquad (\rho = 1, 2)$$

adjoined. When the equations of the family (46) are substituted in the integral I and the functions ψ_μ, these become functions of $b_1, \cdots, b_{p+1}$. The first members of the equations

$$I(b_1, \cdots, b_{p+1}) = I_0 + u,$$
$$\psi_\mu(b_1, \cdots, b_{p+1}) = 0$$

must have their functional determinant equal to zero for $(b_1, \cdots, b_{p+1}) = (0, \cdots, 0)$ by the same argument as that on page 682. This determinant is

$$(47) \qquad \begin{vmatrix} I_1(\xi_1, \eta_1) & \cdot & \cdot & \cdot & I_1(\xi_{p+1}, \eta_{p+1}) \\ \Psi_1(\xi_1, \eta_1) & \cdot & \cdot & \cdot & \Psi_1(\xi_{p+1}, \eta_{p+1}) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \Psi_p(\xi_1, \eta_1) & \cdot & \cdot & \cdot & \Psi_p(\xi_{p+1}, \eta_{p+1}) \end{vmatrix}$$

in which only the second subscripts of the sets $\xi_{1\sigma}, \xi_{2\sigma}, \eta_{i\sigma}$ have been indicated. From its vanishing we argue as on page 682 that there exists a set of constants $\lambda_0, d_1, \cdots, d_p$ not all zero such that the equation

$$\lambda_0 I_1(\xi, \eta) + d_\mu \Psi_\mu(\xi, \eta) = 0$$

must hold for every set of admissible variations $\xi_1, \xi_2, \eta_i(x)$. With the help of formulas (44) and (45) this becomes

$$\begin{aligned} -\int_{x_1}^{x_2} \lambda_r \zeta_r dx &+ [-\lambda_0 f(x_1) + d_\mu(\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1})]\, \xi_1 \\ &+ [\lambda_0 f(x_2) + d_\mu(\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2})]\, \xi_2 \\ &+ [-c_i + d_\mu \psi_{\mu y_{i1}}]\, \eta_i(x_1) \\ &+ [F_{y_i'}(x_2) + d_\mu \psi_{\mu y_{i2}}]\, \eta_i(x_2) = 0. \end{aligned}$$

After the arbitrary constants c_i in (45) have been so chosen that the coefficients of the terms in $\eta_i(x_1)$ in the last expression all vanish it follows by an argument like that of page 683 that $\lambda_{\mu+1} \equiv \cdots \equiv \lambda_n \equiv 0$ and that the coefficients of $\xi_1, \xi_2, \eta_i(x_2)$ also vanish. This result is equivalent to saying that all the determinants of order $p+1$ of the matrix

$$\begin{Vmatrix} -\lambda_0 f(x_1) & -F_{y_i'}(x_1) & \lambda_0 f(x_2) & F_{y_i'}(x_2) \\ \psi_{\mu x_1} + \psi_{\mu y_{i1}} y_{i1}' & \psi_{\mu y_{i1}} & \psi_{\mu x_2} + \psi_{\mu y_{i2}} y_{i2}' & \psi_{\mu y_{i2}} \end{Vmatrix}$$

are zero, since the constants c_i are from equations (15) the values $F_{y_i'}(x_1)$, and since the multipliers $1, d_1, \cdots, d_p$ satisfy all the linear equations whose coefficients are columns of the matrix. The rank of the last matrix is unchanged when one column is multiplied by a factor and added to another, and $\lambda_0 f = F$ on the admissible arc E_{12}, so that these results can be formulated as follows:

For every minimizing arc for the problem of Lagrange with variable end-points there exists a set of constants c_i $(i = 1, \cdots, n)$ *and a function*

$$F(x, y, y', \lambda) = \lambda_0 f + \lambda_1(x)\phi_1 + \cdots + \lambda_m(x)\phi_m$$

such that the equations

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i$$

are satisfied at every point of E_{12}. *The constant* λ_0 *and the functions* $\lambda_\alpha(x)$ $(\alpha = 1, \cdots, m)$ *are not all identically zero on* x_1x_2 *and are continuous except possibly at values of* x *defining corners of* E_{12}. *Furthermore the end-values of* E_{12} *must be such that all the determinants of order* $p + 1$ *of the matrix*

$$(48) \quad \left\| \begin{matrix} -F(x_1) + y_{i1}'F_{y_i'}(x_1) & -F_{y_i'}(x_1) & F(x_2) - y_{i2}'F_{y_i'}(x_2) & F_{y_i'}(x_2) \\ \psi_{\mu x_1} & \psi_{\mu y_{i1}} & \psi_{\mu x_2} & \psi_{\mu y_{i2}} \end{matrix} \right\|$$

are zero. These last conditions are the so-called transversality conditions.

It is clear that the multipliers λ_0, $\lambda_\alpha(x)$ can not all vanish identically on x_1x_2. Otherwise the constants $d_1, \cdots, d_p$ would have to satisfy the linear equations whose coefficients are the columns of the matrix (42) which has rank p. The constants $\lambda_0, d_1, \cdots, d_p$ would then all be zero which is not the case.

9. *Normal admissible arcs for problems with variable end-points.* A *normal admissible arc* for the problem of Lagrange with variable end-points is one for which there exist p sets of admissible variations $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ $(\mu = 1, \cdots, p)$ such that the matrix

$$(49) \quad \left| \begin{matrix} \Psi_1(\xi_1, \eta_1) & \cdots & \Psi_1(\xi_p, \eta_p) \\ \cdot & \cdots & \cdot \\ \Psi_p(\xi_1, \eta_1) & \cdots & \Psi_p(\xi_p, \eta_p) \end{matrix} \right|$$

is different from zero. In the elements of the matrix only the second subscripts of the sets $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ are indicated.

A necessary and sufficient condition that an admissible arc for the problem of Lagrange with variable end-points be normal is that there exists for it no set of multipliers λ_0, $\lambda_\alpha(x)$ *having* $\lambda_0 = 0$ *with which it satisfies the conditions of the last theorem. For a normal extremal arc satisfying the conditions of the last theorem multipliers in the form* $\lambda_0 = 1$, $\lambda_\alpha(x)$ *always exist and in this form they are unique.*

The proof of Section 8 shows that an admissible arc which is not normal has surely a set of multipliers with $\lambda_0 = 0$, since the linear equations whose coefficients are the columns of the determinant (47) have for such an arc solutions $\lambda_0, d_1, \cdots, d_p$ with $\lambda_0 = 0$.

Suppose now that there were a normal admissible arc satisfying the conditions of the theorem of Section 8 and having $\lambda_0 = 0$. Since the matrix preceding (48) is of rank less than $p+1$ we should then have constants d_μ $(\mu = 1, \cdots, p)$ such that

$$
(50) \qquad \begin{aligned}
F(x_1) &= d_\mu(\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1}), \\
F_{y_i'}(x_1) &= d_\mu \psi_{\mu y_{i1}}, \\
-F(x_2) &= d_\mu(\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2}), \\
-F_{y_i'}(x_2) &= d_\mu \psi_{\mu y_{i2}}.
\end{aligned}
$$

The numbers $F(x_1)$, $F(x_2)$ would be zero since $\lambda_0 = 0$ and along an admissible arc $F = \lambda_0 f$. After multiplying these equations respectively by ξ_1, $\eta_i(x_1)$, ξ_2, $\eta_i(x_2)$ and adding we should have

$$\eta_i(x_1) F_{y_i'}(x_1) - \eta_i(x_2) F_{y_i'}(x_2) = d_\mu \Psi_\mu(\xi, \eta).$$

The first member of this equation would vanish for every set of admissible variations $\eta_i(x)$, as was proved in Section 7, page 688, and the second member would necessarily have the same property. Since there is a determinant (49) different from zero we should then have $d_\mu = 0$ for every μ, and equations (50) show that $F_{y_i'}(x_1)$ and $F_{y_i'}(x_2)$ would all vanish. As in Section 7, page 688, this would necessitate the vanishing of λ_0, $\lambda_a(x)$ which is impossible. The proof of the uniqueness of the multipliers $\lambda_0 = 1$, $\lambda_a(x)$ is precisely that of Section 7.

In every neighborhood of a normal admissible arc E_{12} for the Lagrange problem with variable end-points there is an infinity of admissible arcs satisfying the end conditions $\psi_\mu = 0$.

The proof is similar to that of the corresponding theorem in Section 7. Select arbitrarily an admissible set of variations ξ_1, ξ_2, $\eta_i(x)$ and p other such sets $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ with determinant (49) different from zero. There is a $p+1$-parameter family

$$
(51) \qquad \begin{gathered}
y_i = Y_i(x, b, b_1, \cdots, b_p) \\
X_1(b, b_1, \cdots b_p) \leqq x \leqq X_2(b, b_1, \cdots, b_p)
\end{gathered}
$$

of admissible arcs containing E_{12} for $(b, b_1, \cdots, b_p) = (0, 0, \cdots, 0)$ and having the sets ξ_1, ξ_2, $\eta_i(x)$ and $\xi_{1\mu}$, $\xi_{2\mu}$, $\eta_{i\mu}(x)$ as its variations along E_{12}. The existence of the functions Y_i is a consequence of the corollary of Section 3 above, and we may take $X_\rho = x_\rho + b\xi_\rho + b_\mu \xi_{\rho\mu}$ $(\rho = 1, 2)$. Each function ψ_μ becomes a function $\psi_\mu(b, b_1, \cdots, b_p)$ when the functions (51) defining these arcs are substituted. The equations

$$(52) \qquad \psi_\mu(b, b_1, \cdots, b_p) = 0$$

have the initial solution $(b, b_1, \cdots b_\mu) = (0, 0, \cdots, 0)$ at which the functional determinant of their first members with respect to $b_1, \cdots, b_\mu$ is the determinant (49) different from zero. Hence these equations have p solutions $b_\mu = B_\mu(b)$ with initial values $B_\mu(0) = 0$. The one-parameter family

$$(53) \qquad y_i = Y_i[x, b, B_1(b), \cdots, B_p(b)] = y_i(x, b) \qquad x_1(b) \leqq x \leqq x_2(b)$$

where

$$x_\rho(b) = X_\rho[b, B_1(b), \cdots, B_p(b)] \qquad (\rho = 1, 2)$$

contains E_{12} for $b = 0$ and satisfies the equations $\psi_\mu = 0$.

COROLLARY. *If a set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ for a normal admissible arc E_{12} for the Lagrange problem with variable end-points satisfies the equations $\Psi_\mu(\xi, \eta) = 0$, then there exists a one parameter family*

$$y_i = y_i(x, b), \quad x_1(b) \leq x \leq x_2(b)$$

of admissible arcs satisfying the end-conditions $\psi_\mu = 0$, containing E_{12} for the parameter value $b = 0$, and having the set ξ_1, ξ_2, $\eta_i(x)$ as its variations along E_{12}.

If the set ξ_1, ξ_2, $\eta_i(x)$ of the Corollary is used in the construction of the family (53) then we find, by differentiating equations (52) with respect to b and setting $b = 0$, that

$$\Psi_\mu(\xi, \eta) + \Psi_\mu(\xi_\nu, \eta_\nu) B_\nu'(0) = 0.$$

But since the first terms in these equations vanish, and since the determinant (49) is different from zero, it follows that $B_\mu'(0) = 0$ for every μ. Hence the variations of the family (53) are the functions

$$y_{ib}(x, 0) = \eta_i(x) + Y_{ib_\mu} B_\mu'(0) = \eta_i(x),$$
$$x_{\rho b}(0) = \xi_\rho + X_{\rho b_\mu} B_\mu'(0) = \xi_\rho, \qquad (\rho = 1, 2)$$

as required in the Corollary.

CHAPTER II.

APPLICATIONS OF THE EULER-LAGRANGE MULTIPLIER RULE.

10. *The brachistochrone in a resisting medium.* Analytically the problem of the brachistochrone in a plane and in a resisting medium is, as we have seen in Section 2, that of finding among the arcs

$$y = y(x), \quad v = v(x) \quad (x_1 \leq x \leq x_2)$$

satisfying the conditions

$$vv' - gy' + R(v)(1 + y'^2)^{1/2} = 0,$$

(54) $$x_1 - \alpha_1 = y_1 - \beta_1 = v_1 - \gamma = x_2 - \alpha_2 = y_2 - \beta_2 = 0,$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} (1/v)(1 + y'^2)^{1/2} dx.$$

In these expressions primes denote derivatives with respect to x. To apply the Euler-Lagrange rule and the transversality conditions of Section 8 we construct the function

$$\begin{aligned} F &= (1/v)(1 + y'^2)^{1/2} + \lambda[vv' - gy' + R(1 + y'^2)^{1/2}] \\ &= H(1 + y'^2)^{1/2} + \lambda(vv' - gy') \end{aligned}$$

where H is a convenient symbol * for the expression

(55) $$H = (1/v) + \lambda R(v).$$

The differential equations of the normal extremals are then easily found to be

(56) $$H(dy/ds) = \lambda g + a, \quad v(d\lambda/ds) = H_v, \quad v(dv/ds) = g(dy/ds) - R$$

where s is the length of arc defined by the equation

$$ds = (1 + y'^2)^{1/2} dx$$

and a is a new constant of integration. By eliminating dy and ds from equations (56) we find

$$H(H_v dv + R d\lambda) = (g\lambda + a) g d\lambda,$$

which gives at once, since $H_\lambda = R$, the relation

(57) $$H^2 = (g\lambda + a)^2 + b^2$$

where b is a second constant of integration. The constant can be taken squared since the first equation (56) shows that H^2 is always greater than $(\lambda g + a)^2$.

Equations (56) and (57) give further

(58) $$\begin{aligned} \frac{dy}{dv} &= \frac{dy}{ds}\frac{ds}{dv} = \frac{v(\lambda g + a)}{g(\lambda g + a) - RH}, \\ \frac{dx}{dv} &= \frac{dx}{ds}\frac{ds}{dv} = \left[1 + \left(\frac{dy}{ds}\right)^2\right]^{1/2} \frac{ds}{dv} = \frac{bv}{g(\lambda g + a) - RH}. \end{aligned}$$

* Bolza [3, p. 577].

Equation (57) is quadratic in λ and when its solution $\lambda = \lambda(v, a, b)$ is substituted in the last equations the values of x and y may be found by quadratures in the form

$$x = \phi(v, a, b) + c, \quad y = \psi(v, a, b) + d, \tag{59}$$

where c and d are again constants of integration. These are the equations of the minimizing arc in parametric form.

It is very easy to set up the matrix (48) for our function F and the five end conditions. It is a square matrix with six rows and columns and its vanishing prescribes the single condition $\lambda(x_2)v(x_2) = 0$. From the equation (57) multiplied by v^2 and equation (55) we then find at $x = x_2$ that $v_2{}^2(a^2 + b^2) = 1$. For the determination of v_2 and the four constants of integration in equations (59) we have therefore in accordance with conditions (54) the five equations

$$\begin{gathered} \phi(v_1, a, b) + c = \alpha_1, \qquad \phi(v_2, a, b) + c = \alpha_2, \\ \psi(v_1, a, b) + d = \beta_1, \qquad \psi(v_2, a, b) + d = \beta_2, \\ v_2{}^2(a^2 + b^2) = 1. \end{gathered} \tag{60}$$

If the resistance function $R(v)$ were known we should now have in equations (57), (56), and (60) the mathematical mechanism for determining possible normal minimizing curves. The adjective possible is used here because the conditions deduced so far have only been shown to be necessary for a normal minimizing arc. They have not been proved to be sufficient to insure a minimum.

11. *Parametric problems in space.* Let us now consider space curves whose equations are given in the parametric form

$$x = x(s), \quad y = y(s), \quad z = z(s) \qquad (s_1 \leq s \leq s_2). \tag{61}$$

The problem to be studied is that of finding among the arcs of this type which satisfy the equation

$$x'^2 + y'^2 + z'^2 - 1 = 0 \tag{62}$$

and join two given points 1 and 2 in xyz-space, one which minimizes an integral of the form

$$I = \int_{s_1}^{s_2} f(x, y, z, x', y', z')\, ds.$$

Primes now denote differentiation with respect to s. Equation (62) restricts the parameter s to be the length of arc measured along the curve (61). If we agree to measure this length always from the point 1 then the conditions for the curve (61) to pass through 1 and 2 are

$$s_1 = x_1 - \alpha_1 = y_1 - \beta_1 = z_1 - \gamma_1 = x_2 - \alpha_2 = y_2 - \beta_2 = z_2 - \gamma_2 = 0$$

where $(\alpha_1, \beta_1, \gamma_1)$ and $(\alpha_2, \beta_2, \gamma_2)$ are the coördinates of these points. Evidently our problem is one with a variable end-point in $sxyz$-space since s_2 is undetermined.

The function F for normal minimizing arcs is

$$F = f + (\lambda/2)(x'^2 + y'^2 + z'^2 - 1)$$

and the differential equations determining such arcs are

$$(63)\qquad \begin{aligned} &f_x - (d/ds) f_{x'} - \lambda' x' - \lambda x'' = 0, \\ &f_y - (d/ds) f_{y'} - \lambda' y' - \lambda y'' = 0, \\ &f_z - (d/ds) f_{z'} - \lambda' z' - \lambda z'' = 0, \\ &x'^2 + y'^2 + z'^2 = 1. \end{aligned}$$

The sum of the first three of these multiplied, respectively, by x', y', z' gives, with the help of the last one,

$$(64)\qquad (d/ds)(f - x' f_{x'} - y' f_{y'} - z' f_{z'} - \lambda) = 0.$$

The matrix (48) for this problem has eight rows and columns and the vanishing of its determinant demands that at the value s_2

$$(65)\qquad \lambda = f - x' f_{x'} - y' f_{y'} - z' f_{z'}.$$

On account of equation (64) this must be an identity in s.

A very important case is the one for which the function f is positively homogeneous and of the first order in x', y', z', i. e. the one for which the equation

$$(66)\qquad f(x, y, z. kx', ky', kz') = kf(x, y, z, x', y', z')$$

is an identity in its arguments for all $k > 0$. The integral I then has the same value for all parametric representations of the arc (61). The integrands of the length integral and of many other integrals important in the applications of the theory of the Lagrange problem satisfy this condition. When equation (66) is differentiated for k, and the substitution $k = 1$ afterward made, we find the identity

$$(67)\qquad x' f_{x'} + y' f_{y'} + z' f_{z'} = f.$$

From equation (65) it is evident that in this case $\lambda = 0$ and equations (63) become

$$(68)\qquad f_x-(d/ds)f_{x'}=0,\quad f_y-(d/ds)f_{y'}=0,\quad f_z-(d/ds)f_{z'}=0,$$

$$(69)\qquad x'^2+y'^2+z'^2-1=0.$$

Only three of these can be independent, since one finds readily that

$$x'P+y'Q+z'R=(d/ds)(f-x'f_{x'}-y'f_{y'}-z'f_{z'})=0$$

where P, Q, R are symbols for the first members of equations (68).

12. *Isoperimetric problems.* Suppose that we seek to find in the class of arcs

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

joining two given points and satisfying relations of the form

$$(70)\qquad \int_{x_1}^{x_2} g_i(x,y,y')\,dx=l_i\qquad (i=1,\cdots,n)$$

one which minimizes an integral

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

We can transform such a problem into a Lagrange problem by introducing new variables

$$(71)\qquad z_i(x)=\int_{x_1}^{x} g_i(x,y,y')\,dx.$$

The problem just stated is then equivalent to that of finding in the class of arcs

$$y=y(x),\quad z_i=z_i(x)\qquad (i=1,\cdots,n;\ x_1\leq x\leq x_2)$$

satisfying the conditions

$$(72)\qquad \begin{aligned} &g_i(x,y,y')-z_i'=0,\\ &y(x_1)=y_1,\quad y(x_2)=y_2,\\ &z_i(x_1)=0,\quad z_i(x_2)=l_i, \end{aligned}\qquad (i=1,\cdots,n)$$

one which minimizes I.

The function F for a normal minimizing arc for this problem has the form

$$(73)\qquad F=f+\lambda_i(g_i-z_i')$$

and the differential equations determining such an arc are

$$(74)\qquad F_y-(d/dx)F_{y'}=0$$

and the n equations

$$F_{z_i} - (d/dx) F_{z_i'} = (d\lambda_i/dx) = 0$$

which show that the multipliers λ_i are in this case all constants. The solutions of equations (74) form a family of the type

$$y = y(x, a, b, \lambda_1, \cdot\cdot\cdot, \lambda_n).$$

It contains $n+2$ arbitrary constants, and that is precisely the number of relations which the end-conditions (72) impose upon them as one readily verifies. It is evident that the equation (74) is unaltered if we think of the function F in it as defined by the equation

$$(75) \qquad F = f + \lambda_i g_i$$

instead of equation (73).

For a minimizing arc which is not normal there would be a function F defined by equation (75) without the first term. It is clear that the equation (74) would then be defining the minimizing arcs for the problem of minimizing one of the integrals (70), say the first one, in the class of curves joining 1 with 2 and keeping the others constant. An arc E_{12} satisfying equations (74) and these conditions would in general be a minimizing arc for this problem, and it is evident that in that case there could be no other arc near E_{12} giving the first integral its minimum value l_1. Hence in a neighborhood of E_{12} the class of arcs joining 1 with 2 and satisfying conditions (70) would consist of E_{12} alone, and the original minimum problem would be a very trivial one in that neighborhood. Evidently the normal minimizing arcs are by far the most important ones. A similar but somewhat more complicated argument justifies the definition of normal minimizing arcs for the general Lagrange problem given in the preceding sections.

13. *The hanging chain.* It is a principle of mechanics that a chain suspended on two pegs will hang so that its center of gravity is as low as possible. In Section 2 it was seen that the form of the chain is therefore that of a minimizing arc for the problem in which we seek among the arcs $y = y(x)$ $(x_1 \leq x \leq x_2)$ satisfying the conditions

$$(76) \qquad y(x_1) = y_1, \quad y(x_2) = y_2, \quad \int_{x_1}^{x_2} (1 + y'^2)^{1/2} dx = l,$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} y(1 + y'^2)^{1/2} dx.$$

The function F for a minimizing arc has the form

$$F = (y + \lambda)(1 + y'^2)^{1/2}$$

and since λ is now constant the differential equation (74) is equivalent to

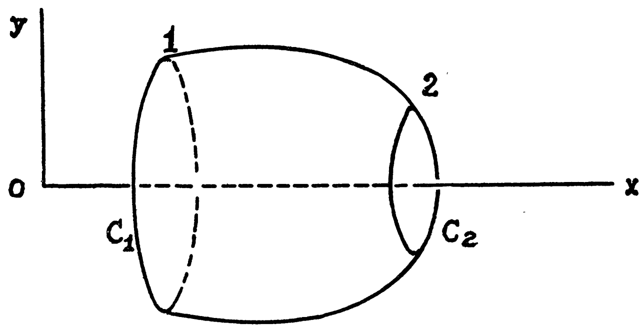
$$F - y'F_{y'} = (y + \lambda)/(1 + y'^2)^{1/2} = b.$$

The integration of this equation has been many times discussed * and its solutions are the catenaries

$$y + \lambda = b\, ch[(x - a)/b].$$

This is a larger family than that of the catenaries for the problem of finding a minimum surface of revolution since it contains an arbitrary constant λ besides a and b. The extra constant is needed, however, for the problem of the hanging chain since there are three conditions (76) to be satisfied for that problem instead of the first two only.

14. *Soap films enclosing a given volume.* Let C_1 and C_2 be two circular discs with a common axis whose edges are joined by a soap film. It is well known that when the volume of air inclosed by the discs and the film is a



fixed constant k the form of the film surface will be that of a surface of revolution enclosing the volume k and having a minimum surface area. To determine the shape of the film we must seek therefore among the arcs $y = y(x)$ $(x_1 \leq x \leq x_2)$ satisfying the conditions

$$y(x_1) = y_1, \quad y(x_2) = y_2, \quad \int_{x_1}^{x_2} y^2 dx = k/\pi$$

one which minimizes the integral

$$I = \int_{x_1}^{x_2} y(1 + y'^2)^{1/2} dx.$$

* See, for example, Bliss [5, p. 91].

The function F is $F = y(1 + y'^2)^{1/2} + \lambda y^2$ and the equation (74) is equivalent to

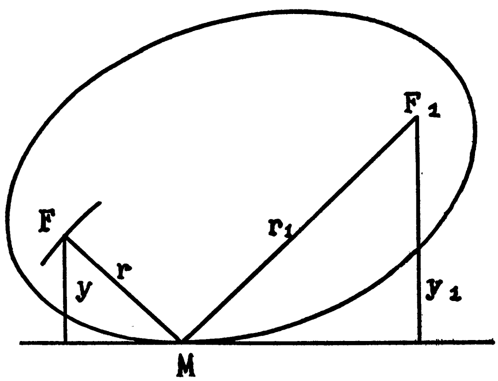
$$(77) \qquad F - y'F_{y'} = y/(1 + y')^{1/2} - \lambda y^2 = c.$$

If we solve this equation for y' and separate the variables we find the solution in the form

$$x = \int \{(c - \lambda y^2)/[y^2 - (c - \lambda y^2)^2]^{1/2}\} dy + d.$$

The integral here is an elliptic integral which can be treated by well known methods.

The solutions of equations (77) can be characterized geometrically in an interesting fashion.* If an ellipse rolls on a straight line, as in the accompanying figure,



its focus F describes a curve whose tangent is at every point perpendicular to FM. The coördinates (x, y) of F, and (x_1, y_1) of F_1, therefore satisfy the equations

$$y = r(dx/ds), \quad y_1 = r_1(dx/ds)$$

since by a well known property of the ellipse the angles made by r and r_1 with the tangent at M are equal. The equations

$$r + r_1 = 2a, \quad yy_1 = b^2$$

express two further well known properties of an ellipse, and elimination of r, r_1, y_1 from these and the preceding ones gives the differential equation

$$y^2 - 2ay(dx/ds) + b^2 = 0$$

* See, for example, Moigno-Lindelöf [6, p. 220].

for the locus of the point F. Equation (77) is identical with this if we set $\lambda = -1/2a$, $c = b^2/2a$. It can similarly be shown that for suitable determinations of λ and c equation (77) is also satisfied by the locus of the focus of a parabola or a hyperbola which rolls on the x-axis. The curves generated as described above by the foci of conics rolling on the x-axis are called unduloids and nodoids.

15. *The case when the functions ϕ_α contain no derivatives.* The problem of this section is that of finding among the arcs

$$(78) \qquad y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2)$$

joining the two given points 1 and 2 and satisfying a set of equations of the form

$$\phi_\alpha(x, y_1, \cdots, y_n) = 0 \qquad (\alpha = 1, \cdots, m < n)$$

one which minimizes an integral

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

Let E_{12} be a particular arc whose minimizing properties are to be studied. It is always presupposed that in a neighborhood of the set of elements (x, y, y') on E_{12} the functions f, ϕ_α have continuous partial derivatives, say of the first four orders, and that the matrix $\| \partial\phi_\alpha/\partial y_i \|$ has rank m at every point of E_{12}.

In order to give this problem the usual Lagrange form we replace it by an equivalent one as follows. We may suppose without loss of generality that at the point 2 the determinant $| \partial\phi_\alpha/\partial y_\beta |$ is one of those of the matrix $\| \partial\phi_\alpha/\partial y_i \|$ which is different from zero. Then we seek to find among the arcs (78) satisfying the conditions

$$(79) \qquad d\phi_\alpha/dx = \phi_{\alpha x} + \phi_{\alpha y_i} y_i' = 0,$$

$$(80) \qquad x_1 - \alpha_1 = y_{i1} - \beta_{i1} = x_2 - \alpha_2 = y_{r2} - \beta_{r2} = 0$$
$$(i = 1, \cdots, n;\ r = m + 1, \cdots, n)$$

one which minimizes I. The coördinates (α_1, β_{i1}) and (α_2, β_{i2}) are those of the points 1 and 2 and necessarily satisfy the equations $\phi_\alpha = 0$. The new problem is evidently equivalent to the old one, at least in a neighborhood of E_{12}, since every arc (78) which joins 1 with 2 and satisfies the equations $\phi_\alpha = 0$ also satisfies (79) and (80); and since, conversely, every arc sufficiently near E_{12} and satisfying (79) and (80) will also satisfy the equations

$\phi_a = 0$ and pass through 1 and 2. This follows because the last $n - m + 1$ equations (80) and the equations $\phi_a = 0$ at 2 imply $y_{a2} - \beta_{a2} = 0$.

Every extremal arc for the new problem is necessarily normal. The determinant analogous to (49) for the end-conditions (80) is in fact

$$\begin{vmatrix} \xi_{11} & \cdot & \cdot & \cdot & \xi_{1p} \\ \eta_{11}(x_1) & \cdot & \cdot & \cdot & \eta_{1p}(x_1) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1}(x_1) & \cdot & \cdot & \cdot & \eta_{np}(x_1) \\ \xi_{21} & \cdot & \cdot & \cdot & \xi_{2p} \\ \eta_{m+1,1}(x_2) & \cdot & \cdot & \cdot & \eta_{m+1,p}(x_2) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \eta_{n1}(x_2) & \cdot & \cdot & \cdot & \eta_{np}(x_2) \end{vmatrix}$$

where $p = 2n - m + 2$, and we can prove that the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}(x)$ $(\sigma = 1, \cdot \cdot \cdot, p)$ can be chosen so that this determinant is different from zero. The equations of variation are in fact readily seen to be the equations

$$(d/dx)\phi_{ay_i}\eta_i = 0$$

which are equivalent to the system

$$\phi_{ay_i}(x)\eta_i(x) = \phi_{ay_i}(x_1)\eta_i(x_1). \tag{81}$$

If the end-values $\eta_i(x_1)$, $\eta_r(x_2)$ are selected arbitrarily these equations determine uniquely the end-values $\eta_a(x_2)$ since the determinant $|\partial\phi_a/\partial y_\beta|$ is by hypothesis different from zero at the point 2. Then the equations (81) and

$$\phi_{ry_i}(x)\eta_i(x) = \zeta_r(x), \tag{82}$$

where the auxiliary functions $\phi_r(x, y)$ are chosen so that the determinant $|\partial\phi_i/\partial y_k|$ is different from zero along E_{12}, determine the end-values $\zeta_r(x_1)$, $\zeta_r(x_2)$ uniquely when $\eta_i(x_1)$, $\eta_r(x_2)$ are given. If functions $\zeta_r(x)$ are chosen with the end-values $\zeta_r(x_1)$, $\zeta_r(x_2)$ but otherwise arbitrarily then equations (81) and (82) determine uniquely a corresponding set of variations $\eta_i(x)$ with the arbitrarily prescribed end-values $\eta_i(x_1)$, $\eta_r(x_2)$. Since ξ_1 and ξ_2 are arbitrary it is evident that the sets $\xi_{1\sigma}$, $\xi_{2\sigma}$, $\eta_{i\sigma}(x)$ can be chosen so that the determinant above is different from zero.

The function F for the Euler-Lagrange multiplier rule of the new problem can be taken in the form

$$F = f + \mu_a(\phi_{ax} + \phi_{ay_k}y_k').$$

By a simple calculation the Euler-Lagrange equations are found to be

$$f_{y_i} - (d/dx)f_{y_i'} - \mu_a'\phi_{ay_i} = 0.$$

If we set $\lambda_a = -\mu_a'$ these are equivalent to the Euler-Lagrange equations calculated for the function

$$F = f + \lambda_a \phi_a$$

and we have the following result:

For the problem of finding among the arcs $y_i = y_i(x)$ $(i = 1, \cdots, n;$ $x_1 \leq x \leq x_2)$ *joining two given points and satisfying the equations*

$$\phi_a(x, y) = 0 \qquad (\alpha = 1, \cdots, m < n)$$

one which minimizes the integral

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

the extremal arcs all satisfy $n + m$ *equations of the form*

$$F_{y_i} - (d/dx) F_{y_i'} = 0, \quad \phi_a = 0$$

where F *is a function of the form* $F = f + \lambda_a \phi_a$.

16. *Geodesics on a surface.** The problem of finding the shortest curve joining two given points on a surface is analytically that of finding among the arcs

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

satisfying the equation

$$\phi(x, y, z) = 0 \tag{83}$$

of the surface and joining the two given points, one which minimizes the integral

$$I = \int_{t_1}^{t_2} (x'^2 + y'^2 + z'^2)^{1/2} dt.$$

The function F for this problem, according to the results of the last section, is

$$F = (x'^2 + y'^2 + z'^2)^{1/2} + \lambda\phi$$

and the Euler-Lagrange equations are $\phi = 0$ and

$$(d/dt) F_{x'} - F_x = d/dt[x'/(x'^2 + y'^2 + z'^2)^{1/2}] - \lambda\phi_x = 0,$$
$$(d/dt) F_{y'} - F_y = d/dt[y'/(x'^2 + y'^2 + z'^2)^{1/2}] - \lambda\phi_y = 0,$$
$$(d/dt) F_{z'} - F_z = d/dt[z'/(x'^2 + y'^2 + z'^2)^{1/2}] - \lambda\phi_z = 0.$$

If these are written in the form

* Bolza [3, p. 553].

$$d^2x/ds^2 = \mu\phi_x, \quad d^2y/ds^2 = \mu\phi_y, \quad d^2z/ds^2 = \mu\phi_z, \quad \phi = 0,$$

where s is the length of arc, they express the fact that at each point of a minimizing arc the principal normal of the arc must coincide with the normal to the surface. Curves which have this property are called *geodesic lines* on the surface. Shortest arcs on a surface must always be sought among the geodesics.

For a sphere the equation (83) has the form

$$x^2 + y^2 + z^2 - 1 = 0$$

and the further equations of the geodesics are

$$d^2x/ds^2 = \mu x, \quad d^2y/ds^2 = \mu y, \quad d^2z/ds^2 = \mu z. \tag{84}$$

Let us determine constants a, b, c so that the expression

$$u = ax + by + cz$$

vanishes with its first derivative at one point of a geodesic on the sphere. Then u must be identically zero on the geodesic since the equation $u_{ss} = \mu u$ is a consequence of equations (84), and since the only solution of this last equation which can vanish with its derivative is $u \equiv 0$. It follows readily that the geodesics on a sphere are great circles cut out of the sphere by the planes $u = 0$.

17. *Brachistochrone on a surface.** Consider a particle of mass m moving in a field of force of such nature that when the particle is at the point (x, y, z) the force acting on it has the projections

$$mX = m(\partial U/\partial x), \quad mY = m(\partial U/\partial y), \quad mZ = m(\partial U/\partial z) \tag{85}$$

on the three coördinate axes, where U is a function of the coördinates x, y, z only. A constant gravitational field in the direction of the negative z-axis, for example, would have

$$X = 0, \quad Y = 0, \quad Z = -g, \quad U = -gz.$$

If a particle were constrained to move on a curve in such a field we should have the force in the direction of the tangent expressed in the two forms

$$mv' = m\,[X(dx/ds) + Y(dy/ds) + Z(dz/ds)]$$

where v is the velocity in the tangent direction, s is the length of arc measured along the curve, and the prime denotes a derivative with respect to the time t. Since $v = ds/dt$ this gives

* Moigno-Lindelöf [6, p. 301].

$$vv' = Xx' + Yy' + Zz' = U',\quad (86)$$
$$v^2 = 2U + c = 2(U - U_1) + v_1^2,$$

where U_1 and v_1 are values of U and v at an initial point 1. For a particle started at 1 with the velocity v_1 the velocity v at a point (x, y, z) is evidently a function of x, y, z and the same for all arcs joining 1 with this point. For an arc

$$(87)\qquad x = x(t),\quad y = y(t),\quad z = z(t) \qquad (t_1 \leq t \leq t_2)$$

joining two fixed points 1 and 2 the time of descent of a particle starting at 1 with the velocity v_1 is

$$T = \int_{s_1}^{s_2} ds/v = \int_{t_1}^{t_2} (1/v)(x'^2 + y'^2 + z'^2)^{1/2} dt$$

where v is the function of x, y, z defined in equation (86).

The problem of finding an arc of quickest descent from a point 1 to a point 2 on a surface

$$(88)\qquad \phi(x, y, z) = 0$$

for a particle starting at 1 with a given velocity v_1 is equivalent analytically to that of finding among the arcs (87) joining the two given points and satisfying the equation (88), one which minimizes the integral T.

The function F for this problem is

$$F = (1/v)(x'^2 + y'^2 + z'^2)^{1/2} + \lambda\phi$$

and the Euler-Lagrange equations have the form

$$\frac{d}{dt} F_{x'} - F_x = \frac{d}{dt} \frac{1}{v} \frac{dx}{ds} + \frac{v_x}{v^2} \frac{ds}{dt} - \lambda\phi_x = 0,$$
$$\frac{d}{dt} F_{y'} - F_y = \frac{d}{dt} \frac{1}{v} \frac{dy}{ds} + \frac{v_y}{v^2} \frac{ds}{dt} - \lambda\phi_y = 0,$$
$$\frac{d}{dt} F_{z'} - F_z = \frac{d}{dt} \frac{1}{v} \frac{dz}{ds} + \frac{v_z}{v^2} \frac{ds}{dt} - \lambda\phi_z = 0$$

to which must be adjoined the equation $\phi = 0$. When multiplied through by dt/ds the equations above become

$$-(v_s/v^2)x_s + (1/v)x_{ss} + (v_x/v^2) - \mu\phi_x = 0,$$
$$-(v_s/v^2)y_s + (1/v)y_{ss} + (v_y/v^2) - \mu\phi_y = 0,$$
$$-(v_s/v^2)z_s + (1/x)z_{ss} + (v_z/v^2) - \mu\phi_z = 0.$$

Multiplied respectively by the direction cosines l, m, n of the direction tangent to the surface, perpendicular to the extremal, and making an acute angle with its principal normal, these give

$$(1/v)(lx_{ss}+my_{ss}+nz_{ss})+(1/v^2)(v_x l+v_y m+v_z n)=0$$

from which we can show that

$$(v^2/\rho)\cos\alpha+R\cos\beta=0 \tag{89}$$

where ρ is the radius of curvature of the curve, α the angle between the radius and the direction $l:m:n$, R the total impressed force, and β the angle between the force and $l:m:n$. This result follows immediately since the numbers ρx_{ss}, ρy_{ss}, ρz_{ss} are the three direction cosines of the principal normal to the curve on which the radius ρ lies, and since from equations (86)

$$vv_x=U_x,\quad vv_y=U_y,\quad vv_z=U_z$$

and U_x, U_y, U_z are the projections on the coördinate axes of the force R. The equation (89) justifies the following characteristic property of brachistochrones on a surface:

Consider a surface $\phi(x,y,z)=0$ lying in a field of force whose vector at (x,y,z) has magnitude R and components X, Y, Z defined by a force function $U(x,y,z)$, as indicated in equations (85). *The centrifugal force of a particle moving on a curve is by definition directed in the direction opposite to that of the radius ρ of the first curvature, and has magnitude v^2/ρ where v is the velocity of the particle. Equation* (89) *shows that at each point of a brachistochrone curve on the surface $\phi=0$ the projection of the centrifugal force on the particular normal to the curve which is also tangent to the surface, is equal to the projection on that same line of the impressed force R.*

This is a characteristic property of brachistochrones. Equation (89) shows that the radius of geodesic curvature $\rho_g=\rho\sec\alpha$ is defined by the equation

$$1/\rho_g=-(R/v^2)\cos\beta. \tag{90}$$

On a surface whose equations are in parametric form with parameters u, v the geodesic curvature of an arc defined by an equation $v=v(u)$ is expressed in terms of $v(u)$, $v'(u)$, $v''(u)$ while the quantities in the second members of the last equation involve only $v(u)$ and $v'(u)$. This equation is consequently a differential equation of the second order. Through each point and direction on the surface there passes therefore one and only one extremal arc for the brachistochrone problem. One can readily verify that the equation

(90) is satisfied by the brachistochrones on a plane which are the well-known cycloids.

18. *The curve of equilibrium of a chain hanging on a surface.** Let us accept from the theories of mechanics the statement that the potential energy of a chain of the form

$$(91) \qquad x = x(t), \quad y = y(t), \quad z = z(t) \qquad (t_1 \leq t \leq t_2)$$

in a field of force like the one described in the last section is

$$P = -\int_{s_1}^{s_2} U ds = -\int_{t_1}^{t_2} U(x'^2 + y'^2 + z'^2)^{1/2} dt,$$

and the statement that a chain at rest will be in equilibrium when the potential energy is a minimum. The problem of finding the position of equilibrium of a chain of given length l joining two given points 1 and 2 and lying on a surface

$$\phi(x, y, z) = 0$$

in such a field is then that of finding among the arcs (91) joining 1 with 2 and satisfying the conditions

$$\int_{t_1}^{t_2} (x'^2 + y'^2 + z'^2)^{1/2} dt = l, \quad \phi(x, y, z) = 0$$

one which minimizes the integral P. In a gravitational field the value of U is $-gz$.

This problem is partly of the isoperimetric and partly of the Lagrange type. By methods used above one readily verifies that its function F now has the form

$$F = (U + \lambda)(x'^2 + y'^2 + z'^2)^{1/2} + \mu\phi,$$

where λ is a constant, and that its extremal arcs satisfy $\phi = 0$ and the equations

$$\begin{aligned} d/dt[(U + \lambda)x'/(x'^2 + y'^2 + z'^2)^{1/2}] - U_x(x'^2 + y'^2 + z'^2)^{1/2} - \mu\phi &= 0, \\ d/dt[(U + \lambda)y'/(x'^2 + y'^2 + z'^2)^{1/2}] - U_y(x'^2 + y'^2 + z'^2)^{1/2} - \mu\phi_y &= 0, \\ d/dt[(U + \lambda)z'/(x'^2 + y'^2 + z'^2)^{1/2}] - U_z(x'^2 + y'^2 + z'^2)^{1/2} - \mu\phi_z &= 0. \end{aligned}$$

These are equivalent to

$$\begin{aligned} U_s x_s + (U + \lambda)x_{ss} - U_x - \nu\phi_x &= 0, \\ U_s y_s + (U + \lambda)y_{ss} - U_y - \nu\phi_y &= 0, \\ U_s z_s + (U + \lambda)z_{ss} - U_z - \nu\phi_z &= 0. \end{aligned}$$

* Moigno-Lindelöf [6, p. 313].

Multiplied respectively by the direction cosines l, m, n of the direction tangent to the surface, perpendicular to the extremal, and making an acute angle with its principal normal, these give

$$(U+\lambda)\cos\alpha/\rho = R\cos\beta,$$

or

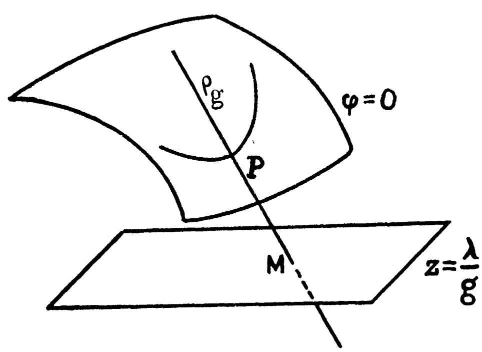
$$\rho_g = (U+\lambda)\sec\beta/R,$$

where ρ, ρ_g, α, β have the significance of the last section. Like the equation (90) this defines a two-parameter family of extremals arcs on the surface $\phi = 0$.

For the particular case of a gravitational field of force $U = -gz$, $R = g$, and β is the angle between the negative z-axis and the direction $l:m:n$ so that $\cos\beta = -n$. Hence in this case

$$\rho_g = [(z-\lambda)/g]/n$$

which says that at each point of a curve of equilibrium the radius of geodesic curvature is equal to the segment PM in the figure, bounded on the line



$l:m:n$ perpendicular to the curve and tangent to the surface $\phi = 0$ by the point P and the plane $z = \lambda/g$. This is a well known property of a catenary $y = c + b\ ch\ [(x-a)/b]$, which is the curve of a hanging chain in a vertical plane. The surface $\phi = 0$ is in this case the xy-plane, the radius ρ_g is the radius of curvature of the catenary, and the plane $z = \lambda/g$ is to be represented by the line $y = c$. The radius of curvature at a point P of the catenary is equal to the intercept on the normal to the catenary at P between the point P and the line $y = c$.

19. *Hamilton's principle.** Suppose that the n particles whose coördinates and masses are x_i, y_i, z_i, m_i $(i = 1, \cdots, n)$ move in a field of force

* Bolza [3, p. 554].

in space such that the force acting at any instant on the i-th particle has components

$$X_i = U_{x_i}, \quad Y_i = U_{y_i}, \quad Z_i = U_{z_i},$$

where U is a function of the time t and the $3n$ coördinates x_i, y_i, z_i. Suppose further that the motions of the particles are restricted by conditions of the form

$$\phi_a = 0 \qquad (\alpha = 1, \cdots, m < 3n),$$

where the functions ϕ_a also depend upon t and the coördinates. The differential equations of motion of the particles, as established in treatises in mechanics, are

$$\begin{aligned} m_i x_i'' &= U_{x_i} + \Sigma_a \lambda_a \phi_{a x_i}, \\ m_i y_i'' &= U_{y_i} + \Sigma_a \lambda_a \phi_{a y_i}, \\ m_i z_i'' &= U_{z_i} + \Sigma_a \lambda_a \phi_{a z_i}, \end{aligned} \tag{92}$$

where α has the range from 1 to m. In this and the following sections of this chapter sums will be indicated as usual and no umbral indices will be used.

Hamilton's principle is simply the statement that the differential equations (92) are the differential equations of the minimizing arcs of the problem of finding in the class of $3n$-dimensional arcs

$$x_i = x_i(t), \quad y_i = y_i(t), \quad z_i = z_i(t) \qquad (t_1 \leq t \leq t_2;\ i = 1, \cdots, n)$$

joining two given points and satisfying the equations $\phi_a = 0$, one which minimizes the integral

$$I = \int_{t_1}^{t_2} (T + U)\, dt$$

where U is the force function and T the so-called kinetic energy

$$T = \tfrac{1}{2} \Sigma_i m_i v_i^2 = \tfrac{1}{2} \Sigma_i m_i (x_i'^2 + y_i'^2 + z_i'^2).$$

It is very easy to show that the equations (92) are the Euler-Lagrange equations for this problem. We have only to set up these equations for the function

$$F = T + U + \Sigma_a \lambda_a \phi_a.$$

An important application of Hamilton's principle is that of determining the equations of motion in terms of the so-called generalized coördinates of Lagrange. The number of coördinates x_i, y_i, z_i is $3n$ and the number of equations $\phi_a = 0$ is m. It is in general possible in an infinity of ways to express these coördinates as functions of t and $3n - m$ arbitrary parameters $q_1, \cdots, q_{3n-m}$ satisfying identically the equations $\phi_a = 0$ and giving all the solutions of these equations. The functions T and U then take the form

$$T = T(t, q, q'), \quad U = U(t, q),$$

and the problem is transformed into that of finding among the arcs $q_r = q_r(t)$ $(r = 1, \cdots, 3n - m)$ joining the two given points one which minimizes the integral I. No adjoined conditions $\phi_a = 0$ are now necessary. The differential equations of the minimizing arcs for the new problem are the equations

$$\frac{d}{dt}\frac{\partial T}{\partial q_r'} - \frac{\partial}{\partial q_r}(T + U) = 0 \qquad (r = 1, \cdots, 3n - m).$$

The important result is that the form of these equations is the same no matter what new coördinates $q_1, \cdots, q_{3n-m}$ with the properties described above are used.

20. *Two forms of the principle of least action.** Let us now consider the somewhat special case where the functions U and ϕ_a of the last section do not contain the time t explicitly. If the equations (92) are multiplied by x_i', y_i', z_i', respectively, added, and integrated we find the well-known relation

$$T = U + h$$

where h is a constant of integration. This is the principle of the conservation of energy which says that the sum of the kinetic energy T and the potential energy $-U$ of a system satisfying equations (92) is always a constant.

Jacobi's form of the principle of least action states that the totality of dynamical trajectories satisfying equations (92) and having a given energy constant h is identical with the totality of extremals for the problem of finding among the arcs

$$x_i = x_i(u), \quad y_i = y_i(u), \quad z_i = z_i(u) \qquad (i = 1, \cdots, n;\ u_1 \leqq u \leqq u_2)$$

joining two given points and satisfying the equations $\phi_a = 0$ one which minimizes the integral

$$I = \int_{u_1}^{u_2} [2(U + h)S]^{1/2} du,$$

where S is simply a notation for the sum

$$S = \Sigma_i m_i (x_{iu}^2 + y_{iu}^2 + z_{iu}^2).$$

The parameter u is not in this case the time, but if at the time t_0 the particles are at the places defined on their trajectories by the parameter value u_0, then it turns out that the time at the place defined by u is

* Bolza [3, pp. 556, 586].

$$t = t_0 + \int_{u_0}^{u} \{S/[2(U+h)]\}^{1/2} du, \tag{93}$$

as one would expect from the relation $S(du/dt)^2 = 2T = 2(U+h)$.

To prove these statements we note that the function F for the minimizing problem just described is

$$F = [2(U+h)S]^{1/2} + \Sigma_a \mu_a \phi_a.$$

A typical one of the Euler-Lagrange equations is

$$(d/du)\{[2(U+h)]/S\}^{1/2} m_i x_{iu} - U_{x_i}\{S/[2(U+h)]\}^{1/2} - \Sigma_a \mu_a \phi_{a x_i} = 0.$$

If we introduce the parameter t along a solution of this equation by means of the formula (93) then the equation itself takes the form

$$m_i x_i'' - U_{x_i} - \Sigma_a \lambda_a \phi_{a x_i} = 0$$

when $\lambda_a = \mu_a\ (du/dt)$, which is the same as the first equation (92).

Lagrange's form of the principle of least action is again a principle for describing those mechanical trajectories which satisfy equations (92) and have a given energy constant h. They are extremals for the problem of finding among the arcs

$$x_i = x_i(t), \quad y_i = y_i(t), \quad z_i = z_i(t) \qquad (i = 1, \cdots, n;\ t_1 \leqq t \leqq t_2)$$

passing through given initial values of the coördinates for a given initial time t_1, passing through given end-values of the coördinates for an unspecified time t_2, and satisfying the equations

$$T - U - h = 0, \quad \phi_a = 0 \tag{94}$$

one which minimizes the integral

$$I = \int_{t_1}^{t_2} T dt.$$

This is a problem with a variable second end-point since t_2 is not specified. The function F for it is

$$F = T + \lambda(T - U - h) + \Sigma_a \mu_a \phi_a$$

and a typical Lagrange equation is

$$(d/dt)(1+\lambda) m_i x_i' + \lambda U_{x_i} - \Sigma_a \mu_a \phi_{a x_i} = 0. \tag{95}$$

When this equation is multiplied by x_i' and added to the other similar ones, it is found with the help of equations (94) that

$$\lambda = (k/2T) - 1/2$$

where k is a constant.

If all the end-values except x_2 are fixed in the theorem of pages 692-3, then the matrix (48) is square and its vanishing requires that

$$F(x_2) - \Sigma_i y_{i2}' F_{y_i'}(x_2) = 0.$$

Interpreted for the function F above this gives $\lambda = -1/2$ at $t = t_2$, with the help of equations (94). It follows that in the formula deduced above for λ we must have $k = 0$ and hence that $\lambda = -1/2$ for all values of t. Equation (95) then takes the form of the first equation (92) when we set $\lambda_a = 2\mu_a$.

CHAPTER III.

FURTHER NECESSARY CONDITIONS FOR A MINIMUM.

In this third chapter three further necessary conditions on a minimizing arc for the Lagrange problem will be developed, analogous to those of Weierstrass, Legendre, and Jacobi for the simpler types of problems of the calculus of variations. The analogue of Legendre's condition was first deduced by Clebsch [20] and the analogue of Jacobi's condition by A. Mayer [24]. For the deduction of these necessary conditions and for a number of other purposes we shall find the auxiliary theorems of the next section convenient.

21. *Two important auxiliary theorems.* Consider a one parameter family of admissible arcs

$$(96) \qquad y_i = y_i(x, b), \quad x_3(b) \leq x \leq x_4(b), \qquad (i = 1, \cdots, n)$$

for which the functions $x_3(b)$, $x_4(b)$, $y_i(x, b)$, $y_i'(x, b)$ are continuous and have continuous derivatives with respect to b in the domain of values (x, b) defined by the inequalities $b' \leq b \leq b''$, $x_3(b) \leq x \leq x_4(b)$, and whose end values describe two arcs C and D. The values of I taken along the arcs (96) are given by the formula

$$I(b) = \int_{x_3}^{x_4} f[x, y(x, b), y'(x, b)]\, dx$$

which has the derivative

$$I'(b) = f x_b \Big|_3^4 + \int_{x_3}^{x_4} \{ f_{y_i} y_{ib} + F_{y_i'} y'_{ib} \}\, dx.$$

The index here is umbral and we shall use umbral indices freely elsewhere in this chapter. Since the arcs (96) are all admissible this result may also be written in the form

$$(97) \qquad \lambda_0 I'(b) = Fx_b\Big|_3^4 + \int_{x_3}^{x_4} \{F_{y_i} y_{ib} + F_{y_i'} y'_{ib}\} dx,$$

where the multipliers λ_0, $\lambda_a(x)$ in the function

$$F = \lambda_0 f + \lambda_a \phi_a$$

are entirely arbitrary. If now a particular arc of the family (96) satisfies the equations

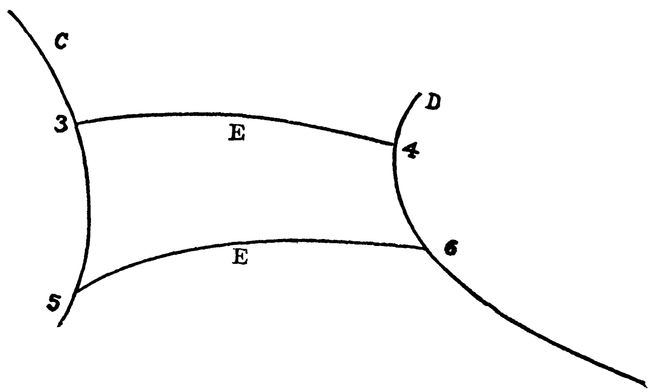
$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i$$

with a set of multipliers λ_0, $\lambda_a(x)$, then the introduction of these multipliers enables us to replace formula (97) by

$$\lambda_0 I'(b) = Fx_b + F_{y_i'} y_{ib}\Big|_3^4$$

where b is the particular value defining that arc. Since the equations of C and D are deduced from

$$x = x(b), \qquad y_i = y_i[x(b), b]$$



by replacing $x(b)$ by $x_3(b)$ and $x_4(b)$, respectively, it follows that along either of these arcs

$$dy_i = y'_i dx + y_{ib} db,$$

and therefore that

$$\lambda_0 dI = F dx + (dy_i - y_i' dx) F_{y_i'}\Big|_3^4.$$

Hence we have the following theorem:

AUXILIARY THEOREM. I. *Let*

$$(98) \qquad y_i = y_i(x, b), \quad x_1(b) \leq x \leq x_2(b), \qquad (i = 1, \cdots, n)$$

be a one-parameter family of admissible arcs without corners whose end-points describe two arcs C and D. If one of the arcs (98) *satisfies the equations*

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i} dx + c_i \tag{99}$$

with a set of multipliers λ_0, $\lambda_\alpha(x)$ *then for the value of* b *defining it the values of* I *along the arcs* (98) *have a differential defined by the equation*

$$\lambda_0 dI = F dx + (dy_i - y_i' dx) F_{y_i'} \Big|_3^4. \tag{100}$$

In this formula the differentials dx, dy_i *at the point 3 are those of* C, *and at the point 4 those of* D.

If the particular arc along which the equation (99) holds is a normal arc then λ_0 can be taken equal to unity in formula (100). If each of the curves (98) has a set of multipliers $\lambda_0(b)$, $\lambda_\alpha(x, b)$ with which it satisfies equations (99), then the formula (100) holds along every arc of the family. We suppose that the functions $\lambda_0(b)$, $\lambda_\alpha(x, b)$ are continuous for $b' \leq b \leq b''$, $x_3(b) \leq x \leq x_4(b)$, and then we have

AUXILIARY THEOREM II. *Suppose that the arcs of the family* (98) *are all extremal arcs with multipliers of the form* $\lambda_0 = 1$, $\lambda_\alpha(x, b)$. *Then the values of* I *on two arcs* E_{34} *and* E_{56} *of the family satisfy the equation*

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

with the values of the integral

$$I^* = \int \{F dx + (dy_i - y_i' dx) F_{y'_i}\}$$

along the corresponding segments C_{35} *and* D_{46} *shown in the last figure.*

This is readily found by integrating both sides of formula (100) with respect to b from the value b' defining simultaneously the points 3 and 4 to the value b'' defining similarly 5 and 6. The integrand of the integral I^* is readily seen to be a continuous function of b on the arcs C_{35} and D_{46} corresponding to the interval $b'b''$, on account of the properties of the functions $x(b)$, $y_i(x, b)$ defining the family (98).

22. *Necessary conditions analogous to those of Weierstrass and Legendre.* Suppose that the equations

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leq x \leq x_2)$$

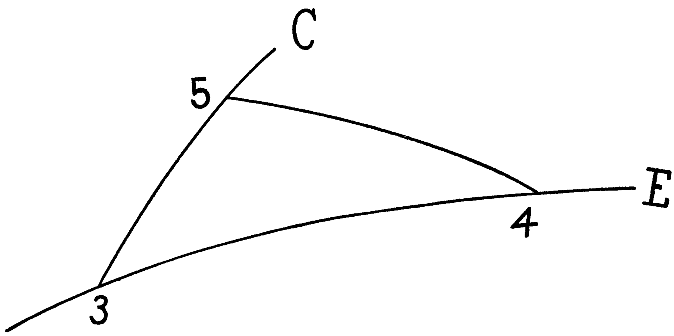
are those of a minimizing arc E_{12} for our problem.

We shall designate a set of values (x, y, y') as admissible if it lies in the neighborhood $\mathfrak{R}$ of page 1, satisfies the equations $\phi_\alpha = 0$, and gives the matrix

$\| \phi_{a y_i'} \|$ the rank m. Let 3 be an arbitrary point on the arc E_{12} and let (x_3, y_{i3}, Y'_{i3}) be an admissible set. There is always an admissible arc

(C) $$y_i = Y_i(x) \qquad (x_3 \leq x \leq x_3 + h)$$

through this set since the equations $\phi_a = 0$ determine uniquely m of the functions $Y_i(x)$ passing with their derivatives through the values prescribed by this initial set when the $n - m$ other functions $Y_i(x)$ have been chosen with initial values of themselves and their derivatives through their corresponding initial values of the set.



Suppose now that the arc E_{12} is normal on every sub-interval, and let 4 be so near to 3 on E_{12} that the arc E_{34} contains no corner. There is a $2n$-parameter family of admissible arcs $y_i = y_i(x, b_1, \cdots, b_{2n})$ containing E_{12} for $(b_1, \cdots, b_{2n}) = (0, \cdots, 0)$ and having $2n$ sets of variations $\eta_i(x)$ for which the determinant (38) with x_1, x_2 replaced by x_3, x_4 is different from zero. The $2n$ equations

$$y_i(x_5, b_1, \cdots, b_{2n}) = Y_i(x_5), \qquad y_i(x_4, b_1, \cdots, b_{2n}) = y_{i4}$$

have the initial solution $(x_5, b_1, \cdots, b_{2n}) = (x_3, 0, \cdots, 0)$ at which their functional determinant for $b_1, \cdots, b_{2n}$ is the determinant (38) with $x_1 = x_3$, $x_2 = x_4$ and different from zero. Hence they determine $2n$ functions $b_\mu = B_\mu(x_5)$ which vanish for $x_5 = x_3$. The family

$$y_i = y_i[x, B_1(x_5), \cdots, B_{2n}(x_5)] = y_i(x, x_5)$$

is now a one-parameter family of arcs joining the curve C of the figure to the point 4. The sum

$$\begin{aligned} \Phi(x_5) &= I(C_{35}) + I(E_{54}) \\ &= \int_{x_3}^{x_5} f(x, Y, Y')\,dx + \int_{x_5}^{x_4} f[x, y(x, x_5), y'(x, x_5)]\,dx \end{aligned}$$

must have its derivative ≥ 0 at x_3 if $I(E_{12})$ is to be a minimum. But with the help of formula (100) this derivative is seen to be

$$\Phi'(x_3) = E(x, y, y', Y', \lambda)\,|^3$$

if we define the E-function by the formula

$$(101)\quad E = F(x, y, Y', \lambda) - F(x, y, y', \lambda) - (Y_i' - y_i')F_{y_i'}(x, y, y', \lambda).$$

The multipliers in F are those associated uniquely with the normal minimizing arc E_{12}. Evidently one may always replace f by F for admissible sets (x, y, y'). We have then the following necessary condition:

ANALOGUE OF WEIERSTRASS NECESSARY CONDITION. *At each element (x, y, y', λ) of a minimizing arc which is normal on every sub-interval the inequality*

$$E(x, y, y', Y', \lambda) \geq 0$$

must be satisfied for every admissible set $(x, y, Y') \neq (x, y, y')$.

The proof just given does not apply to the values x, y, y', λ at the right-hand end of an arc abutting on a corner, but it can be modified easily to be applicable by taking the point 4 at the left of 3, or one can infer the desired result by continuity considerations.

Consider now a set of values π_i satisfying the equations

$$(102)\qquad \phi_{\alpha y_i'}\pi_i = 0$$

at an element (x, y, y') of E_{12}. By means of the equations

$$(103)\qquad \phi_{r y_i'}\pi_i = \kappa_r$$

these define $n - m$ further quantities κ_r. The equations

$$\phi_\alpha(x, y, p) = 0, \qquad \phi_r(x, y, p) = z_r + \epsilon\kappa_r$$

now have the initial solution $(\epsilon, p_1, \cdots, p_n) = (0, y_1', \cdots, y_n')$ and determine uniquely a set of solutions $p_i(\epsilon)$ with initial values $p_i(0) = y_i'$. The derivatives $p_i'(0)$ of these functions satisfy equations (102) and (103) when inserted in place of the numbers π_i and hence must coincide with them. The sets $(x, y, p(\epsilon))$ are now all admissible for sufficiently small values of ϵ, and according to the last theorem must satisfy the condition

$$E(x, y, y', p(\epsilon), \lambda) \geq 0.$$

But we readily verify that this expression vanishes with its first derivative for ϵ at the value $\epsilon = 0$. Its second derivative

$$F_{y_i' y_k'}\pi_i\pi_k$$

at $\epsilon = 0$ must therefore be ≥ 0, from which we infer the

NECESSARY CONDITION OF CLEBSCH. *At every element* (x, y, y', λ) *of a minimizing arc which is normal on every sub-interval the inequality*

$$F_{y_i' y_k'}(x, y, y', \lambda)\pi_i\pi_k \geqq 0$$

must be satisfied by every set $(\pi_i, \cdots, \pi_n) \neq (0, \cdots, 0)$ *which is a solution of the m equations*

$$\phi_{a y_i'}(x, y, y')\pi_i = 0.$$

23. *The envelope theorem.* According to the theorem of page 687, every extremal arc E_{12} along which the determinant R is different from zero is a member of a $2n$-parameter family of extremals of the form

$$y_i = y_i(x, a, b), \qquad \lambda_a = \lambda_a(x, a, b)$$

for special values a_{i0}, b_{i0} of the parameters. The family can be so chosen that the determinant (36) is different from zero at x_1, and we shall see in Section 27, page 727, that this determinant is in fact different from zero everywhere on E_{12}. If the constants a_i, b_i are replaced by functions $a_i(t)$, $b_i(t)$ with the initial values $a_i(0) = a_{i0}$, $b_i(0) = b_{i0}$ a one-parameter family of extremals is defined containing the arc E_{12} for the special parameter value $t = 0$. The arcs of this family will pass through the point 1 for $x = x_1$, and will touch an enveloping curve D at the points defined by a suitably chosen function $x(t)$, if the equations

$$x' = k, \qquad y_{ix}x' + y_{ia_k}a_k' + y_{ib_k}b_k' = ky_{ix},$$
$$y_{i1} = y_i(x_1, a, b)$$

hold identically in t when x, a_i, b_i are replaced by the functions of t described above and the primes denote derivatives with respect to t. The first row of equations imposes the condition that the direction of the tangent to the curve D shall coincide with the direction $1 : y_1' : \cdots : y_n'$ of the tangent to the extremal. In order that these equations may be true it is evidently necessary and sufficient that the equations

$$y_{ia_k}[x(t), a(t), b(t)]\, a_k' + y_{ib_k}[x(t), a(t), b(t)]\, b_k' = 0,$$
$$y_{ia_k}[x_1, a(t), b(t)]\, a_k' + y_{ib_k}[x_1, a(t), b(t)]\, b_k' = 0,$$

hold identically in t. If the derivatives a_k', b_k' are not zero it follows that the determinant

$$(104) \qquad \Delta(x, x_1, a, b) = \begin{vmatrix} y_{ia_k}(x, a, b) & y_{ib_k}(x, a, b) \\ y_{ia_k}(x_1, a, b) & y_{ib_k}(x_1, a, b) \end{vmatrix}$$

vanishes identically in t when $x(t)$, $a_i(t)$, $b_i(t)$ are substituted.

DEFINITION OF A CONJUGATE POINT. A value $x_3 \neq x_1$ is said to define a *point 3 conjugate to 1 on the extremal arc* E_{12} if it is a root of a determinant $\Delta(x, x_1, a_0, b_0)$ belonging to a $2n$-parameter family of extremals $y_i = y_i(x, a, b)$, $\lambda_a = \lambda_a(x, a, b)$ for which the determinant

$$\begin{vmatrix} y_{ia_k} & y_{ib_k} \\ v_{ia_k} & v_{ib_k} \end{vmatrix}$$

is different from zero on E_{12} as described on page 727.

Suppose now that 3 is such a conjugate point, and furthermore one at which the derivative Δ_x does not vanish. It is evident that if $\Delta_x \neq 0$ one at least of the minors of order $2n - 1$ of Δ does not vanish at 3, and that the same property is therefore possessed by one at least of the determinants of order $2n$ of the matrix

$$\left\| \begin{matrix} \Delta_x & \Delta_{a_k} & \Delta_{b_k} \\ 0 & y_{ia_k}(x, a, b) & y_{ib_k}(x, a, b) \\ 0 & y_{ia_k}(x_1, a, b) & y_{ib_k}(x_1, a, b) \end{matrix} \right\|$$

since one at least of these determinants is the product of Δ_x by a non-vanishing minor of Δ. Then the first of the differential equations

$$\begin{aligned} \Delta_x(x, x_1, a, b)\, dx + \Delta_{a_k}(x, x_1, a, b)\, da_k + \Delta_{b_k}(x, x_1, a, b)\, db_k &= 0, \\ y_{ia_k}(x, a, b)\, da_k + \quad y_{ib_k}(x, a, b)\, db_k &= 0, \\ y_{ia_k}(x_1, a, b)\, da_k + \quad y_{ib_k}(x_1, a, b)\, db_k &= 0, \end{aligned} \tag{105}$$

with $2n - 1$ of the others determine functions $x(t)$, $a_k(t)$, $b_k(t)$ with the initial values $x(0) = x_3$, $a_k(0) = a_{k0}$, $b_k(0) = b_{k0}$, and with derivatives x', a_k', b_k' not all zero at $t = 0$. Since $\Delta_x \neq 0$ at 3 it follows further that a_k', b_k' can not all vanish at $t = 0$. Since Δ vanishes at these initial values and has its derivative with respect to t identically zero, it must be itself identically zero in t. One sees readily then that the one remaining equation (105) is a consequence of the others when $x(t)$, $a_k(t)$, $b_k(t)$ are substituted. The following theorem is established:

Let E_{12} be an extremal arc along which the determinant R is different from zero, and let 3 be a point conjugate to 1 on E_{12} at which the derivative Δ_x of the determinant (104) *is different from zero. Then there exists through the point 1 a one-parameter family of extremals*

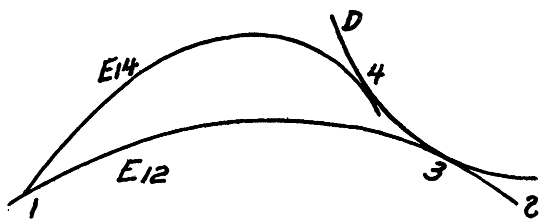
$$y_i = y_i(x, t), \qquad \lambda_a = \lambda_a(x, t) \tag{106}$$

containing E_{12} for the parameter value $t = 0$ and having an envelope D which touches E_{12} at the point 3. The functions y_i, y_{ix}, λ_a and the function $x(t)$

defining D have continuous derivatives in a neighborhood of the values x, t belonging to the arc E_{12}.

The last statement of the theorem is a consequence of the hypothesis (b) of page 676. For as a result of this hypothesis the functions y_i, y_{ix}, λ_a of the theorem on page 687 have continuous derivatives of the second order at least, and the solutions $x(t)$, $a_k(t)$, $b_k(t)$ of the equations (105) must therefore have continuous derivatives of at least the first order.

THE ENVELOPE THEOREM. *If the envelope D of the one-parameter family of extremals* (106) *has a branch projecting backward from 3 toward the*



point 1, as shown in the figure, then for every position of the point 4 on D preceding and near to 3 the arc $E_{14} + D_{43} + E_{32}$ is an admissible arc satisfying the equations $\phi_a = 0$. Furthermore for every such arc

$$I(E_{14} + D_{43} + E_{32}) = I(E_{12}).$$

Expressed in integral form the value of $I(E_{14} + D_{43})$ is

$$I(E_{14} + D_{43}) = \int_{x_1}^{x(t)} f\,[x, y(x,t) y'(x,t)]\,dx + \int_t^0 f\,x'\,dt$$

where the arguments in f in the last integral are $x(t), y[x(t), t], y'[x(t), t]$. The differential of the first integral with respect to t is given by formula (100) of page 716, and that of the second integral is readily found. It follows that

$$dI(E_{14} + D_{43}) = -E(x, y, y', Y', \lambda)\,dx\,|^4$$

where Y' is the slope of D. But this vanishes identically in t since $Y' = y'$ at every point of D, and the final conclusion of the theorem is established. Evidently the envelope D satisfies the equations $\phi_a = 0$ at each point 4 since it is tangent at that point to the extremal arc E_{14}.

24. *The analogue of Jacobi's condition.* The analogue of Jacobi's condition was discovered for the Lagrange problem by A. Mayer. Its statement is as follows:

THE NECESSARY CONDITION OF MAYER. *Let E_{12} be an extremal arc for the Lagrange problem which is normal on every sub-interval of x_1x_2 and has the determinant*

$$R = \begin{vmatrix} F_{y_i' y_k'} & \phi_{a y_i'} \\ \phi_{a y_k'} & 0 \end{vmatrix}$$

different from zero at every point of it. If E_{12} is a minimizing arc for the problem then between 1 and 2 on E_{12} there can be no point 3 conjugate to 1.

The proof of the statement for the case when the envelope has a branch as described in the envelope theorem is not difficult if one accepts the assertion that every extremal arc of a family $y_i(x, a, b)$ whose end-values x_1, x_2 and parameters a, b are sufficiently near to those of a normal extremal arc of the family is also normal. The proof of this assertion depends upon the fact that when the functions $y_i(x, a, b)$ are substituted in the equations of variation, the solutions $\eta_i(x, a, b)$ of those equations are continuous in the parameters a, b as well as x. Hence if there are $2n$ sets of variations η_{is} $(s = 1, \cdots, 2n)$ making the determinant (38) different from zero for the values x_{10}, x_{20}, a_0, b_0 defining the normal extremal, then this determinant will remain different from zero for neighboring values x_1, x_2, a, b.

If the arcs $E_{14} + D_{43} + E_{32}$ of the envelope theorem were all minimizing arcs they would necessarily have continuous multipliers since they have no corners. According to the assertion discussed in the last paragraph those sufficiently near to E_{12} would be normal on the intervals x_1x_4 and x_4x_2 since by hypothesis E_{12} is normal on every sub-interval and hence E_{13} and E_{32} are both normal. It follows readily that the composite arc $E_{14} + D_{43} + E_{32}$ would have the multipliers of the extremal E_{14} along E_{14}, the multipliers of the extremal tangent to D_{43} at each point of that arc, and the multipliers of the extremal E_{12} along E_{32}. Hence on the composite arcs near E_{12} the value of R would be everywhere different from zero as on E_{12}, and by the differentiability condition of page 684, each such arc would necessarily be an extremal. The extremal E_{12} is, however, the only one having its values y_i, v_i at $x = x_2$, or what is the same thing, its values y_i, y_i', λ_a at $x = x_2$. Hence the arcs $E_{14} + D_{43} + E_{32}$ can not all be minimizing arcs since otherwise all of them and the envelope D would necessarily fall upon E_{12} and their multipliers would coincide with those of E_{12}. But this is impossible because the derivatives $a_k'(t)$, $b_k'(t)$ of the family as determined on page 720 do not all vanish.

If an arc $E_{14} + D_{43} + E_{32}$ is not a minimizing arc it is always possible to find a neighboring admissible arc which joins the points 1 and 2 and gives the integral I a smaller value than $I(E_{14} + D_{43} + E_{32})$, that is, a smaller value than $I(E_{12})$, and hence $I(E_{12})$ can not be a minimum.

The preceding proof of the necessary condition of Mayer is a very satisfactory one geometrically because it emphasizes the geometrical interpretation of the conjugate point and the envelope theorem. But it rests upon two restrictive assumptions, namely, the non-vanishing of the derivative Δ_x at the conjugate point 3, and the requirement that the envelope have a branch projecting from 3 toward 1. In the following sections a proof of an entirely different sort is given which is free from these disadvantages.

25. *The second variation for a normal extremal.* It has been proved on page 17 that if the functions $\eta_i(x)$ of a set of admissible variations for a normal extremal arc E_{12} satisfy the relations $\eta_i(x_1) = \eta_i(x_2) = 0$, then there is a one-parameter family of admissible arcs

$$y_i = y_i(x, b) \qquad (x_1 \leq x \leq x_2)$$

joining the points 1 and 2, containing E_{12} for the parameter value $b = 0$, and having the functions $\eta_i(x)$ as its variations along E_{12}. When the various members of the equations

$$I(b) = \int_{x_1}^{x_2} f[x, y(x, b), y'(x, b)]\,dx,$$
$$0 = \phi_\alpha[x, y(x, b), y'(x, b)]$$

are differentiated for b it is found that

$$I'(b) = \int_{x_1}^{x_2} (f_{y_i} y_{ib} + f_{y_i'} y_{ib}')\,dx,$$
$$0 = \phi_{\alpha y_i} y_{ib} + \phi_{\alpha y_i'} y_{ib}',$$

and a second differentiation gives for $b = 0$

$$I''(0) = \int_{x_1}^{x_2} (f_{y_i} y_{ibb} + f_{y_i'} y_{ibb}' + f_{y_i y_k} \eta_i \eta_k + 2 f_{y_i y_k'} \eta_i \eta_k' + f_{y_i' y_k'} \eta_i' \eta_k')\,dx,$$
$$0 = \phi_{\alpha y_i} y_{ibb} + \phi_{\alpha y_i'} y_{ibb}' + \phi_{\alpha y_i y_k} \eta_i \eta_k + 2 \phi_{\alpha y_i y_k'} \eta_i \eta_k' + \phi_{\alpha y_i' y_k'} \eta_i' \eta_k'.$$

When the last equations are multiplied by the factors λ_α, integrated from x_1 to x_2, and added to $I''(0)$ this derivative is found to have the value

$$I''(0) = \int_{x_1}^{x_2} (F_{y_i} y_{ibb} + F_{y_i'} y_{ibb}' + 2\omega)\,dx \tag{107}$$

where

$$2\omega(x, \eta, \eta') = F_{y_i y_k} \eta_i \eta_k + 2 F_{y_i y_k'} \eta_i \eta_k' + F_{y_i' y_k'} \eta_i' \eta_k'. \tag{108}$$

On account of the equations

$$(d/dx) F_{y_i'} = F_{y_i}$$

the first two terms in the integral (107) have the anti-derivative $F_{y_i'}y_{ibb}$ and this vanishes at x_1 and x_2 as one readily sees by differentiating the equations

$$y_{i1} = y_i(x_1, b), \qquad y_{i2} = y_i(x_2, b)$$

twice with respect to b. Hence the following conclusions are justified:

Along a normal extremal arc E_{12} the second variation of the integral I is always expressible in the form

$$I''(0) = \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')\,dx$$

where 2ω is the quadratic form defined by equation (108). *If $I(E_{12})$ is a minimum for the Lagrange problem then this second variation must be $\geqq 0$ for every set of admissible variations $\eta_i(x)$ whose functions satisfy the relations*

(109) $$\eta_i(x_1) = \eta_i(x_2) = 0.$$

Since admissible variations satisfy the differential equations of variations

(110) $$\Phi_a(x, \eta, \eta') = \phi_{ay_i}\eta_i + \phi_{ay_i'}\eta_i' = 0$$

it is clear that these properties of the second variation suggest a minimum problem in $x\eta$-space of the same type as the original Lagrange problem in xy-space. There is an integral $I''(0)$ which must be $\geqq 0$ in the class of arcs $\eta_i = \eta_i(x)$ in $x\eta$-space satisfying the differential equations (110) and passing through the two fixed points $(x, \eta_1, \cdots, \eta_n) = (x_1, 0, \cdots, 0)$ and $(x, \eta_1, \cdots, \eta_n) = (x_2, 0, \cdots, 0)$, as indicated by equations (109). Evidently the minimum of $I''(0)$ in this class of arcs must be $\geqq 0$ if E_{12} is to be a solution of the original Lagrange problem.

The differential equations of the extremal arcs for the problem in $x\eta$-space are the equations

(111) $$(d/dx)\Omega_{\eta_i'} = \Omega_{\eta_i}, \qquad \Phi_a(x, \eta, \eta') = 0$$

where Ω is a function of the form

(112) $$\Omega(x, \eta, \eta', \mu) = \mu_0\omega + \mu_a\Phi_a.$$

These are called by von-Escherich [31, Vol. 107, p. 1236] the *accessory system* of linear differential equations. They are the analogues of the Jacobi differential equation for the simplest problem in the plane. If the arc E_{12} is a normal extremal arc for the original Lagrange problem, then every extremal arc for the new problem in $x\eta$-space has this property, since the equations of variation of the linear equations $\Phi_a = 0$ for the $x\eta$-problem are these equations themselves. Hence it is proper when E_{12} is normal to set $\mu_0 = 1$, the multipliers $\mu_0 = 1$, $\mu_a(x)$ for an extremal arc of the $x\eta$-problem being then unique.

The quadratic form $\Omega(x, \eta, \eta', \mu)$ has the properties

$$2\Omega = \eta_i \Omega_{\eta_i} + \eta_i' \Omega_{\eta_i'} + \mu_a \phi_{\mu_a}, \tag{113}$$

$$u_i \Omega_{v_i} + u_i' \Omega_{v_i'} + \rho_a \Omega_{\sigma_a} = v_i \Omega_{u_i} + v_i' \Omega_{u_i'} + \sigma_a \Omega_{\rho_a}, \tag{114}$$

where the derivatives of Ω are understood to have the arguments (η, η', μ), (u, u', ρ), or (v, v', σ) as indicated by their subscripts. These are well-known formulas for quadratic forms which are readily provable and which will be useful in the following paragraphs.

A final remark concerning the accessory differential equations (111) is also important. These equations are linear and homogeneous in the variables η_i, η_i', η_i'', μ_a, μ_a', and the determinant of coefficients of the variables η_i'', μ_a' is the determinant R which will be assumed different from zero along E_{12}. The arguments of Section 6 therefore tell us at once that the accessory equations have one and but one solution η_i, μ_a taking prescribed values of η_i, $\Omega_{\eta_i'}$ at a given value of x, or, what is the same thing, prescribed values of η_i, η_i', μ_a satisfying the equations of variation. In particular the only solution taking the values $\eta_i = \Omega_{\eta_i'} = 0$, or $\eta_i = \eta_i' = \mu_a = 0$, at a given x is the set of functions $\eta_i(x) \equiv \mu_a(x) \equiv 0$ which one readily sees to be a solution since the accessory equations are linear and homogeneous in η_i, η_i', η_i'', μ_a, μ_a'.

26. *A second proof of the analogue of Jacobi's condition.* Consider now a minimizing arc E_{12} for the original Lagrange problem, which has no corners and along which the determinant R of page 684 is everywhere different from zero. According to the differentiability condition on that same page the arc E_{12} must then be an extremal as defined in section 6. For the developments of the present section the additional assumption will be made that the extremal E_{12} is normal on every sub-interval of x_1x_2.

DEFINITION OF CONJUGATE POINT. A value x_3 is said to define a *point 3 conjugate to 1 on the arc* E_{12} if there exists an extremal $\eta_i = u_i(x)$, $\mu_a = \rho_a(x)$ for the $x\eta$-problem whose functions $u_i(x)$ satisfy the relations $u_i(x_1) = u_i(x_3) = 0$ but are not identically zero on x_1x_3. We shall presently see that the definition of a conjugate point on page 720 is equivalent to the one here given.

With this definition agreed upon the necessary condition of Mayer as stated on page 722 can be proved by showing that if there exists a point 3 conjugate to 1 between 1 and 2 on E_{12} then there exists also an admissible set of variations $\eta_i(x)$ making $I''(0) < 0$. As a first step consider the functions $\eta_i(x)$, $\mu_a(x)$ defined by the equations

$$(115) \qquad \begin{array}{lll} \eta_i(x) \equiv u_i(x), & \mu_a(x) \equiv \rho_a(x) & \text{on } x_1 \leq x \leq x_3, \\ \eta_i(x) \equiv 0, & \mu_a(x) \equiv 0 & \text{on } x_3 \leq x \leq x_2, \end{array}$$

where the functions $u_i(x)$, $\rho_a(x)$ are those indicated in the definition just given for the conjugate point. With the help of the equations (112), (111), (113) it follows readily that for these functions $\eta_i(x)$

$$\begin{aligned} I''(0) &= \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')\, dx = \int_{x_1}^{x_3} 2\Omega(x, u, u', \rho)\, dx \\ &= \int_{x_1}^{x_3} (u_i \Omega_{u_i} + u_i' \Omega_{u_i'} + \rho_a \Omega_{\rho_a})\, dx \\ &= u_i \Omega_{u_i'} \Big|_1^3 = 0. \end{aligned}$$

The functions $\eta_i(x)$ in (115) can not minimize $I''(0)$, however, since, as will be shown in the next paragraph, they do not satisfy the corner conditions

$$(116) \qquad \Omega_{\eta_i'}\,[x, \eta, \eta'(x-0), \mu(x-0)] = \Omega_{\eta_i'}\,[x, \eta, \eta'(x+0), \mu(x+0)]$$

at the point x_3. Hence there must be other admissible variations $\eta_i(x)$ vanishing at x_1 and x_2 and giving $I''(0)$ a value less than zero, and $I(E_{12})$ can not be a minimum.

To show that the corner conditions are not satisfied one may calculate readily the values of the derivatives $\Omega_{\eta_i'}$ for the functions (115) at the left and right of x_3. It is found then that the corner conditions (116) would require that $\Omega_{u_i'} = 0$ at the point x_3 as well as $u_i = 0$, and according to a remark at the end of the preceding section the functions $u_k(x)$, $\rho_a(x)$ would then have to be identically zero, which is not the case. The proof of Mayer's condition is now complete.

27. *The determination of conjugate points.* For a one-parameter family of extremals

$$y_i = y_i(x, b), \qquad \lambda_a = \lambda_a(x, b)$$

the equations

$$(d/dx) F_{y_i'} = F_{y_i}, \qquad \phi_a = 0$$

are identities in x and b. When they are differentiated with respect to b we find

$$\begin{gathered} (d/dx)(F_{y_i' y_k} y_{kb} + F_{y_i' y_k'} y_{kb}' + F_{y_i' \lambda_a} \lambda_{ab}) = F_{y_i y_k} y_{kb} + F_{y_i y_k'} y_{kb}' + F_{y_i \lambda_a} \lambda_{ab}, \\ \phi_{a y_k} y_{kb} + \phi_{a y_k'} y_{kb}' = 0, \end{gathered}$$

and these are precisely the accessory equations with the arguments $\eta_i = y_{ib}$, $\mu_a = \lambda_{ab}$. A $2n$-parameter family of extremals defined by equations similar to equations (35) or (37) on pages 686-7 furnishes by this differentiation process $2n$ solutions

$$(117) \qquad \begin{array}{l} y_{1a_k}, \cdots, y_{na_k}; \ \lambda_{1a_k}, \cdots, \lambda_{ma_k} \\ y_{1b_k}, \cdots, y_{nb_k}; \ \lambda_{1b_k}, \cdots, \lambda_{mb_k} \end{array} \qquad (k = 1, \cdots, n)$$

of the accessory equations. The formulas of most importance here are those for $2n$-parameter families for which the determinant (36) is different from zero at some point, say x_1. We shall see in the next paragraph that it is then different from zero for all values of x.

Since the determinant R is different from zero along E_{12} the equations

$$\zeta_i = \Omega_{\eta_i'}(x, \eta, \eta', \mu), \qquad \Phi_a(x, \eta, \eta') = 0,$$

analogous to equations (30) on page 685, can be solved for η_k', μ_ρ. The solution has the form

$$(118) \qquad \eta_k' = G_k(x, \eta, \zeta), \qquad \mu_\beta = H_\beta(x, \eta, \zeta),$$

and the accessory equations are equivalent to the equations

$$(119) \qquad (d\eta_k/dx) = G_k(x, \eta, \zeta), \ (d\zeta_k/dx) = \Omega_{\eta_k}(x, \eta, G(x, \eta, \zeta), H(x, \eta, \zeta)).$$

All of these equations are linear and homogeneous in the arguments η_i, η_i', μ_a, ζ_i where they occur. For equations of the type (119) it is well known * that $2n$ solutions (η_k, ζ_k) whose determinant is different from zero at a single value of x, will have that determinant different from zero for all values x, and that every other solution is linearly expressible with constant coefficients in terms of $2n$ solutions which have this property. Every solution of the accessory equations is therefore expressible linearly with constant coefficients in terms of the $2n$ corresponding sets (η_k, μ_β) defined by the second of equations (118).*

Since the determinant (36) of page 687 is different from zero at $x = x_1$ it follows that it is different from zero for all values of x. For the $2n$ solutions (117) of the accessory equations define $2n$ solutions (η_k, ζ_k) of equations (119) whose determinant is different from zero. Hence every solution $(\eta_i, \mu_a) = (u_i, \rho_a)$ of the accessory equations is expressible in the form

$$u_i = c_k y_{ia_k} + d_k y_{ib_k}, \quad \rho_a = c_k \lambda_{aa_k} + d_k \lambda_{ab_k}.$$

---

* See, for example, Goursat, *A Course in Mathematical Analysis*, translated by Hedrick and Dunkel, Vol. 2, Part 2, pp. 153-4.

The values x_3 determining conjugate points according to the definition on page 725 are those for which the equations

$$u_i(x_3) = c_k y_{ia_k}(x_3) + d_k y_{ib_k}(x_3) = 0,$$
$$u_i(x_1) = c_k y_{ia_k}(x_1) + d_k y_{ib_k}(x_1) = 0,$$

have solutions c_k, d_k not all zero. But these are precisely the values x_3 for which the determinant $\Delta(x, x_1, a, b)$ vanishes, as indicated in the definition on page 720. We shall see on page 740 that for every ξ on x_1x_2 the zeros of $\Delta(x, \xi, a, b)$ are isolated from ξ when an extension of E_{12} is normal on every sub-interval.

Consider now an n-parameter family of extremals

$$y_i = y_i(x, b_1, \cdots, b_n), \qquad \lambda_a = \lambda_a(x, b_1, \cdots, b_n)$$

all of which pass through the point 1, and such that the functions $v_i = F_{y_i'}$ for the family have their determinant $|v_{ib_k}|$ different from zero at x_1. All of the derivatives y_{ib_k} vanish at x_1 as one may see by differentiating the equations

$$y_{i1} = y_i(x_1, b_1, \cdots, b_n)$$

with respect to b_k. Every solution η_i, μ_a of the accessory equations for which the functions $\eta_i(x)$ all vanish at x_1 is expressible in the form

$$\eta_i = c_k y_{ib_k}, \qquad \mu_a = c_k \lambda_{ab_k},$$

where the coefficients c_k are constants. For such a solution is uniquely determined by its set of values $\eta_i = 0$, $\zeta_i = \Omega_{\eta_i'}$ at $x = x_1$. If the constants c_k are solutions of the equations

$$\zeta_i(x_1) = c_k v_{ib_k}(x_1),$$

which in fact determine them uniquely, then the two solutions η_i, μ_a and $c_k y_{ib_k}$, $c_k \lambda_{ab_k}$ of the accessory equations have the same values $\eta_i = 0$, ζ_i at $x = x_1$ and hence are identical for all values of x. It follows that the points 3 conjugate to 1 on E_{12} are determined by values x_3 for which the equations

$$c_k y_{ib_k}(x_3) = 0$$

have solutions c_k not all zero, that is, by values $x_3 \neq x_1$ which make the determinant $D(x, b) = |y_{ib_k}|$ vanish. These results may be summarized as follows:

Let E_{12} be an extremal arc which is contained in a 2n-parameter family of extremals

$$y_i = y_i(x, a_1, \cdots, a_n, b_1, \cdots, b_n), \quad \lambda_\alpha = \lambda_\alpha(x, a_1, \cdots, a_n, b_1, \cdots, b_n)$$

for special values a_{i0}, b_{i0} *of the parameters. Suppose furthermore that the determinant*

$$\begin{vmatrix} y_{ia_k} & y_{ib_k} \\ v_{ia_k} & v_{ib_k} \end{vmatrix}$$

of the family, where $v_i = F_{y_i'}(x, y, y', \lambda)$, *is different from zero at the point 1 on* E_{12}. *Then the points 3 conjugate to 1 on* E_{12} *are determined by the roots* $x_3 \neq x_1$ *of the function* $\Delta(x, x_1, a_0, b_0)$ *where*

$$\Delta(x, x_1, a, b) = \begin{vmatrix} y_{ia_k}(x, a, b) & y_{ib_k}(x, a, b) \\ y_{ia_k}(x_1, a, b) & y_{ib_k}(x_1, a, b) \end{vmatrix}.$$

If E_{12} *is a member of an n-parameter family of extremals*

$$y_i = y_i(x, b_1, \cdots, b_n), \quad \lambda_\alpha = \lambda_\alpha(x, b_1, \cdots, b_n)$$

all of which pass through the point 1, and such that the determinant $|v_{ib_k}|$ *for the functions* $v_i = F_{y_i'}$ *belonging to the family is different from zero at the point 1 on* E_{12}, *then the points conjugate to 1 on* E_{12} *are determined by the roots* $x_3 \neq x_1$ *of the function* $D(x, b_0)$ *where*

$$D(x, b) = |y_{ib_k}|$$

and the b_{i0} *are the parameter values defining* E_{12}.

CHAPTER IV.

SUFFICIENT CONDITIONS FOR A MINIMUM.

The conditions developed in the preceding chapters are conditions which must be satisfied by every minimizing arc for the Lagrange problem, but they have not been shown to actually insure the minimizing property. In this chapter it is proposed to discuss sets of conditions which are sufficient for a minimum. The methods of proof used are in essence those which Weierstrass applied in similar cases and which have been extended to the Lagrange problem by A. Mayer, Bolza, and others, but they involve important simplifications and improvements.

28. *Mayer fields and the fundamental sufficiency theorem.* The notion of a field has been defined in a number of different ways. The definition given here is not the usual one and is somewhat sophisticated, but it emphasizes properties which are well known for fields of the simplest problem in the

plane, and leads promptly to the theorem which is fundamental for all of the sufficiency proofs. In order to phrase this definition as simply as possible let us agree to call a set of values (x, y, y') *admissible* if it lies interior to the region $\mathfrak{R}$ where the continuity properties of the functions f and ϕ_a have been assumed, and satisfies the equations $\phi_a = 0$, and gives the matrix $\| \phi_{ay_i'} \|$ the rank m.

DEFINITION OF A MAYER FIELD. A *Mayer field* is a region $\mathfrak{F}$ of xy-space containing only interior points and having associated with it a set of functions

$$p_i(x, y), \qquad l_a(x, y)$$

with the following properties:

(a) they have continuous first partial derivatives in $\mathfrak{F}$;
(b) the sets $(x, y, p(x, y))$ defined by the points (x, y) in $\mathfrak{F}$ are all admissible;
(c) the integral

$$I^* = \int \{F(x, y, p, l)\,dx + (dy_i - p_i dx) F_{y_i'}(x, y, p, l)\}$$

formed with these functions is independent of the path in $\mathfrak{F}$.

The integral I^* can also be written in the form

$$I^* = \int \{A dx + B_i dy_i\}$$

where

$$A(x, y) = F(x, y, p, l) - p_k F_{y_k'}(x, y, p, l),$$
$$B_i(x, y) = F_{y_i'}(x, y, p, l).$$

If such an integral is independent of the path every arc is a minimizing arc for it and the Euler-Lagrange differential equations applied to it give the well-known conditions

$$(120) \qquad \partial A/\partial y_i = \partial B_i/\partial x, \quad \partial B_i/\partial y_k = \partial B_k/\partial y_i$$

as necessary conditions for its invariantive property. One may readily prove the identities

$$(121) \quad \partial A/\partial y_i - \partial B_i/\partial x = F_{y_i} - (\partial/\partial x) F_{y_i'} - p_k(\partial/\partial y_k) F_{y_i'} + p_k(\partial B_i/\partial y_k - \partial B_k/\partial y_i) + \phi_a\, \partial l_a/\partial y_i$$

where the partial derivatives indicated by the symbols ∂ are taken with respect to the independent variables x, y_i which occur explicitly and also in the field functions $p_i(x, y)$, $l_a(x, y)$.

From these results it is easy to see that in the field $\mathfrak{F}$ every solution $y_i(x)$ of the equations

$$(122) \qquad dy_i/dx = p_i(x, y)$$

is an extremal with the multipliers $\lambda_\alpha = l_\alpha(x, y(x))$. For in the first place such an arc necessarily satisfies the equations $\phi_\alpha = 0$, since the values (x, y, p) are all admissible; and in the second place the equations (120) and (121) then show that along such an arc

$$F_{y_i} - (d/dx)F_{y_i'} = F_{y_i} - (\partial/\partial x)F_{y_i'} - p_k(\partial/\partial y_k)F_{y_k'} = 0.$$

The arcs satisfying equations (122) are called the *extremals of the field.* Through each point of $\mathfrak{F}$ there passes one and but one such extremal arc since the equations (122) are of the first order. Furthermore the value of I^* along an extremal arc of the field is equal to that of the original integral I, since the equations $dy_i - p_i dx = 0$ are all satisfied along the field extremals.

If E_{12} is an extremal arc of a field $\mathfrak{F}$ then for every admissible arc C_{12} in the field joining the same two points 1 and 2 the formula

$$(123) \qquad I(C_{12}) - I(E_{12}) = \int_{x_1}^{x_2} E[x, y, p(x, y), y', l(x, y)]dx$$

holds, where

$$E = F(x, y, y', l) - F(x, y, p, l) - (y_i' - p_i)F_{y_i'}(x, y, p, l)$$

and the arguments $y(x)$, $y'(x)$ in the integrand are those belonging to C_{12}.

The formula (123) is the analogue of a well-known one of Weierstrass and the proof of it is very simple. For since I^* is independent of the path in $\mathfrak{F}$ and has the same values as I along an extremal of the field it follows that

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

and hence that

$$I(C_{12}) - I(E_{12}) = I(C_{12}) - I^*C_{12}).$$

The last two terms give the integral in the second member of the formula (123) when the integrand f in $I(C_{12})$ is replaced by F. This is evidently permissible since C_{12} is by hypothesis an admissible arc and therefore satisfies the equations $\phi_\alpha = 0$.

With these results in mind it is now possible to prove the following important theorem:

THE FUNDAMENTAL SUFFICIENCY THEOREM. *If E_{12} is an extremal arc of a field $\mathfrak{F}$ and if at each point of the field the condition*

$$E[x, y, p(x, y), y', l(x, y)] > 0$$

holds for every admissible set (x, y, y') *different from* (x, y, p), *then the inequality* $I(C_{12}) > I(E_{12})$ *is true for every admissible arc* C_{12} *in the field and joining the end-points of* E_{12} *but not identical with* E_{12}.

It is evident from formula (123) that the inequality $I(C_{12}) \geqq I(E_{12})$ is necessarily satisfied. The equality sign is appropriate only if the E-function vanishes at every point of C_{12}, that is, only if the equations $y_i' = p_i$ are satisfied at each point of C_{12}. But in that case the arc C_{12} would coincide with E_{12} since the equations $y_i' = p_i$ have only one solution through the point 1 and that is E_{12} itself.

29. *The construction of a field.* The extremal arcs of a field may be regarded as forming an n-parameter family since one of them passes through each point of the field. By analogy with the properties of fields for the simplest problem of the calculus of variations in the plane it might be expected that every n-parameter family of extremals which simply covers a region in xy-space would provide a set of slope functions and multipliers $p_i(x, y)$, $l_\alpha(x, y)$ which would make the integral I^* independent of the path in that region, and hence form a field over the region, but such is not the case. The n-parameter families which can form fields are special in character in somewhat the same way that a two-parameter family of straight lines in xyz-space is special if it is cut orthogonally by a surface. It is well known that not every such family of straight lines has an orthogonal surface.

Let the equations

$$y_i = y_i(x, a_1, \cdots, a_n), \quad \lambda_\alpha = \lambda_\alpha(x, a_1, \cdots, a_n) \tag{124}$$

be an n-parameter family of extremals with the property that the functions y_i, y_{ix}, λ_α have continuous first partial derivatives for all values $(x, a_1, \cdots, a_n)$ satisfying conditions of the form

$$\begin{gathered} \xi_1(a_1, \cdots, a_n) \leqq x \leqq \xi_2(a_1, \cdots, a_n). \\ (a_1, \cdots, a_n) \text{ in a region } A. \end{gathered} \tag{125}$$

Suppose further that there is an n-space

$$x = x_1(a_1, \cdots, a_n), \quad y_i = y_i(x_1(a_1, \cdots, a_n), a_1, \cdots, a_n)$$

cutting the extremals (124) for which the function $x_1(a_1, \cdots, a_n)$ has continuous first partial derivatives in A. The extremals (124) are said to simply cover a field $\mathfrak{F}$ of points (x, y) if to each point of the region there corresponds one and but one set of values x, $a_i(x, y)$ satisfying the first n equations (124) and the conditions (125), and if the functions $a_i(x, y)$ so defined have continuous derivatives in $\mathfrak{F}$. The functions

$$p_i(x, y) = y_{ix}[x, a(x, y)], \quad l_a(x, y) = \lambda_a[x, a(x, y)]$$

are then a set of slope-functions and multipliers for the region $\mathfrak{F}$, and the following theorem can be proved:

Suppose that an n-parameter family of extremals

$$y_i = y_i(x, a_1, \cdots, a_n), \quad \lambda_a = \lambda_a(x, a_1, \cdots, a_n) \tag{126}$$

is intersected by an n-space

$$x = x_1(a_1, \cdots, a_n), \quad y_i = y_i(x_1(a_1, \cdots, a_n), a_1, \cdots, a_n) \tag{127}$$

and simply covers a region $\mathfrak{F}$ of xy-space containing only interior points, in the manner described in the preceding paragraphs. If the parameter values of the extremal through a point (x, y) are denoted by $a_i(x, y)$ then the region $\mathfrak{F}$ is a field with the slope-functions and multipliers

$$p_i(x, y) = y_{ix}[x, a(x, y)], \quad l_a(x, y) = \lambda_a[x, a(x, y)] \tag{128}$$

provided that the integral I^* is independent of the path in the n-space (127).

The proof may be made with the help of the Auxiliary Theorem II of page 716. For an arc D_{46} in $\mathfrak{F}$ with equations of the form

$$x = x(t), \quad y_i = y_i(t) \qquad (t' \leq t \leq t'')$$

defines a one-parameter family of extremals intersecting it, and a corresponding arc C_{35} in the n-space (127), by means of the functions $a_i(t) = a_i[x(t), y(t)]$. According to the auxiliary theorem cited it is then true that

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

The three terms on the right are completely determined when the end-points of D_{46} are given, since by hypothesis the value $I^*(C_{35})$ is the same for all arcs C_{35} with the same end-points in the n-space (127). Hence the integral I^* is independent of the path in the whole of the region $\mathfrak{F}$, as required by the definition of a field.

The preceding theorem suggests at once a number of methods of constructing fields by means of n-parameter families of extremals. One may take the n-parameter family through a fixed point O and regard the point O as a degenerate n-space (127). Certainly on this degenerate n-space the integral I^* is independent of the path. Every region in xy-space simply covered by the extremals will then be a field with the slope-functions and multipliers (128).

If an n-space

$$x = X(a_1, \cdots, a_n), \quad y_i = Y_i(a_1, \cdots, a_n) \tag{129}$$

and a function $W(a_1, \cdots, a_n)$ are chosen arbitrarily in advance the $n + m$ equations

$$FX_{a_i} + (Y_{ka_i} - y_k' X_{a_i}) F_{y_k'} = W_{a_i}, \quad \phi_a = 0, \tag{130}$$

where the arguments of F, ϕ_a are X, Y_i, y_i', λ_a, may under certain conditions be solved for the $n + m$ variables y_i', λ_a as functions of $a_1, \cdots, a_n$. At each point of the n-space an initial element x, y_i, y_i', λ_a of an extremal is thus determined, and the extremals which have these initial elements form an n-parameter family. The integrand of the integral I^* for this family has the value dW on every arc in the n-space (129), on account of the equations (130), since along such an arc the differentials dx, dy_k have the values

$$dx = X_{a_i} da_i, \quad dy_k = Y_{ka_i} da_i.$$

Hence the integral I^* will be independent of the path on the space (129) and every region of xy-space simply covered by the family of extremals will form a field. If the derivatives W_{a_i} all vanish then an n-space (129) which satisfies the equations (130) with the extremals of the family it is said to cut the family transversally.

A similar discussion can be made for initial spaces (129) of lower dimensions.

30. *Sufficient conditions for a strong relative minimum.* In the following paragraphs the necessary conditions deduced in the preceding chapters will be designated by the numerals I, II, III, IV. These are, respectively, the necessary condition of page 683, the analogue of Weierstrass' condition on page 718, the condition of Clebsch on page 719, and the condition of Mayer on page 722. The notations II′, III′ will be used to designate the conditions II and III when strengthened to exclude the equality sign which occurs in their statements. Similarly IV′ is the stronger condition of Mayer which excludes the conjuate point 3 from the end-point 2 of E_{12}, as well as from the interior of that arc. An arc E_{12} with multipliers $\lambda_0 = 1$, $\lambda_a(x)$ will be said to satisfy the condition II_b' if the inequality

$$E(x, y, y', Y', \lambda) > 0$$

holds for every set of elements (x, y, y', Y', λ) for which the set (x, y, y', λ)

is in a neighborhood of similar sets belonging to E_{12}, and $(x, y, Y') \neq (x, y, y')$ is admissible.

Every extremal arc E_{12} defined on an interval $x_1 x_2$, and on which the determinant R is different from zero, defines an extended extremal on an interval $x_1 - d \leqq x \leqq x_2 + d$ which contains E_{12} as part of it. We may call this longer extremal an extension of E_{12}.

With these agreements we can state the following theorem:

SUFFICIENT CONDITIONS FOR A STRONG RELATIVE MINIMUM. *If an admissible arc E_{12}, without corners and with an extension normal on every subinterval, satisfies the conditions I, II_b', III′, IV′, then there is a neighborhood $\mathfrak{F}$ of the points (x, y) on E_{12} such that the inequality $I(C_{12}) > I(E_{12})$ holds for every admissible arc C_{12} which is in $\mathfrak{F}$ and not identical with E_{12}.*

The minimum furnished by E_{12} is called a relative minimum because it is in a class of arcs restricted to lie in a neighborhood $\mathfrak{F}$ of E_{12}; and it is a strong relative minimum because the neighborhood $\mathfrak{F}$ lays no restriction on the slopes y_i' of comparison arcs which lie in it.

In order to prove the theorem we should note in the first place that the condition I and the normality of E_{12} imply a unique set of multipliers $\lambda_0 = 1$, $\lambda_\alpha(x)$ and constants c_i with which E_{12} satisfies the equations (24) of page 683.

The condition III' now implies that the determinant R of page 684 is different from zero at every element (x, y, y', λ) of E_{12}. For at an element where R vanished the linear equations

$$(131) \qquad F_{y_i' y_k'} \Pi_k + \phi_{\alpha y_i'} \mu_\alpha = 0, \quad \phi_{\alpha y_k'} \Pi_k = 0$$

would have solutions Π_k, μ_α not all zero, with the numbers Π_k also not all zero since the matrix $\| \phi_{\alpha y_i'} \|$ has rank m. But when the first equations (131) are multiplied by $\Pi_1, \cdots, \Pi_n$ and added it is found that

$$F_{y_i' y_k'} \Pi_i \Pi_k = 0,$$

as a result of the second set of equations (131), which would contradict the condition III′.

Since the determinant R is different from zero along E_{12} it follows from the differentiability condition of page 684 that E_{12} must be an extremal. According to the developments of Section 6, page 687, there exists a $2n$-parameter family of extremals

$$y_i = y_i(x, a, b), \quad \lambda_\alpha = \lambda_\alpha(x, a, b)$$

containing E_{12} for special parameter values a_{i0}, b_{i0}. The functions y_i, y_{ix}, λ_α have continuous partial derivatives of the first three orders near the values

(x, a_i, b_i) belonging to E_{12}, and the determinant (36) of page 687 is different from zero at the point 1 on E_{12}.

It will be shown in Section 32 that for an arc E_{12} with an extension normal on every sub-interval there is always an interval $x_1 - h \leq x \leq x_1 + h$ containing no pair of conjugate points, or in other words, containing no two values x, x_0 which satisfy the equation $\Delta(x, x_0, a_0, b_0) = 0$, where Δ is the determinant (104) of page 719. Hence if $x_0 < x_1$ be chosen sufficiently near to x_1 the function $\Delta(x, x_0, a_0, b_0)$ will be different from zero on the interval $x_1 \leq x \leq x_1 + h$, and different from zero also in the interval $x_1 + h \leq x \leq x_2$ on account of the continuity of Δ and condition IV′. The equations

$$(132) \qquad y_i = y_i(x, a, b), \qquad y_{i0} = y_i(x_0, a, b)$$

have now as initial solutions the totality of values (x, y, a, b) belonging to E_{12}, and their functional determinant $\Delta(x, x_0, a, b)$ with respect to the parameters a_i, b_i is different from zero at these initial solutions on account of the choice of x_0 which has just been made. Well-known implicit function theorems then justify the statement that there is a neighborhood $\mathfrak{F}$ of the points (x, y) on E_{12} in which the equations (132) have solutions $a_i(x, y)$, $b_i(x, y)$ with continuous partial derivatives of the first three orders since the functions (132) have such derivatives. This neighborhood $\mathfrak{F}$ is a field with the slope functions and mutlipliers

$$p_i(x, y) = y_{ix}[x, a(x, y), b(x, y)], \quad \lambda_a(x, y) = \lambda_a[x, a(x, y), b(x, y)]$$

since the extremals which simply cover it all pass through the fixed point 0 corresponding, on E_{12} extended, to the value x_0. If the field $\mathfrak{F}$ is taken sufficiently small the values $x, y, p_i(x, y), \lambda_a(x, y)$ belonging to it will remain in so small a neighborhood of the sets (x, y, y', λ) belonging to E_{12} that according to the condition II_b' the inequality

$$(133) \qquad E[x, y, p(x, y), y, \lambda(x, y)] > 0$$

will hold for every admissible element $(x, y, y') \neq (x, y, p)$ in $\mathfrak{F}$. The fundamental sufficiency theorem then justifies the theorem which was to be proved.

31. *Sufficient conditions for a weak relative minimum.* The conditions I, III′, IV′ were the only ones used in the last section up to the very last paragraph. If they only are assumed it is not possible to establish the condition (133). The E-function for admissible elements (x, y, y') in the field $\mathfrak{F}$ is expressible, however, with the help of Taylor's formula with integral remainder term, in the form

$$(134) \qquad E = (y_i' - p_i)(y_k' - p_k)\int_0^1 (1-\theta) F_{y_i' y_k'}[x, y, p + \theta(y' - p), \lambda]\, d\theta$$

where $p_i = p_i(x, y)$, $\lambda_a = \lambda_a(x, y)$ are the slope-functions and multipliers of the field, and the differences $y_i' - p_i$ satisfy the equation

$$\phi_a(x, y, y') - \phi_a(x, y, p) = (y_i' - p_i) \int_0^1 \phi_{a y_i'} [x, y, p + \theta(y' - p)] d\theta = 0.$$

On account of the condition III′ the quadratic form

$$\Pi_i \Pi_k \int_0^1 (1 - \theta) F_{y_i' y_k'} [x, y, p + \theta(y' - p), \lambda] d\theta$$

is positive for all sets (x, y, y', Π) for which (x, y, y') is on the arc E_{12} where $y_i' = p_i(x, y)$, and for which the numbers Π_i satisfy the equations

$$\Pi_i \Pi_i = 1, \quad \Pi_i \int_0^1 \phi_{a y_i'} [x, y, p + \theta(y' - p)] d\theta = 0.$$

Hence it stays positive for sets of values (x, y, y', Π) for which the numbers Π_i satisfy these equations and the set (x, y, y') lies in a sufficiently small neighborhood N of similar sets on E_{12}. It follows readily that the E-function (134) of the field $\mathfrak{F}$ is positive at least for all sets $(x, y, y') \neq (x, y, p)$ in the neighborhood N, and the following theorem is therefore justified:

SUFFICIENT CONDITIONS FOR A WEAK RELATIVE MINIMUM. *If an admissible arc E_{12} without corners and with an extension normal on every sub-interval, satisfies the conditions I, III′, IV′ then there is a neighborhood N of the sets of values (x, y, y') on E_{12} such that the inequality $I(C_{12}) > I(E_{12})$ holds for every admissible arc C_{12} whose elements (x, y, y') are all in N but which is not identical with E_{12}.*

The minimum described in this theorem is called a weak relative minimum because the neighborhood N in which it exists requires the slopes y_i' of the comparison arcs C_{12}, as well as their points (x, y), to be near those on E_{12}.

32. *The justification of a preceding statement.* It was stated on page 736 that there is always an interval $x_1 - h \leq x \leq x_1 + h$ on which no two values x, x_0 can satisfy the equation $\Delta(x, x_0, a_0, b_0) = 0$. The proof of this statement is not simple, but it can be made with the help of properties of solutions of the accessory differential equations

$$(135) \qquad (d/dx)\Omega_{\eta_i'} - \Omega_{\eta_i} = 0, \quad \Omega_{\mu_a} = \Phi_a = 0$$

for the arc E_{12} described on page 724. It is understood that the arc E_{12} is an extremal with an extension normal on every sub-interval and satisfying the condition III′. As a consequence of these properties the determinant R is different from zero at every point of E_{12}.

The equation (114) on page 725

$$u_i\Omega_{v_i} + u_i'\Omega_{v_i'} + \rho_a\Omega_{\sigma_a} = v_i\Omega_{u_i} + v_i'\Omega_{u_i'} + \sigma_a\Omega_{\rho_a},$$

justifies readily the further relation

$$u_i[\Omega_{v_i} - (d/dx)\Omega_{v_i'}] + \rho_a\Omega_{\sigma_a} - v_i[\Omega_{u_i} - (d/dx)\Omega_{u_i'}] - \sigma_\sigma\Omega_{\rho_a}$$
$$= (d/dx)(v_i\Omega_{u_i'} - u_i\Omega_{v_i'}).$$

Hence for every pair of solutions u_i, ρ_a and v_i, σ_a of the accessory equations the expression

$$\psi(u, \rho, v, \sigma) = u_i\Omega_{v_i'} - v_i\Omega_{u_i'}$$

is a constant. If this constant is zero the two solutions are said to be *conjugate solutions.*

There is one and but one set of solutions η_i, μ_a of the accessory equations (135) for which $\eta_i, \zeta_i = \Omega_{\eta_i'}$ take assigned values at the value x_1, as shown for the original xy-problem on pages 685 and 686. A matrix of n solutions u_{ik}, ρ_{ak} $(k = 1, \cdots, n)$ therefore exists for which at the value x_1 the matrix $\| u_{ik} \|$ is the identity matrix and the corresponding matrix of the functions $\zeta_i = \Omega_{\eta_i'}$ has all its elements zero. The solutions u_{ik}, ρ_{ak} $(k = 1, \cdots, n)$ are conjugate in pairs, as one readily verifies, since their functions ζ_i all vanish at x_1. The notations u_i, ρ_a and v_i, σ_a will be used for the linear expressions

$$u_i = a_k u_{ik}, \qquad \rho_a = a_k \rho_{ak},$$
$$v_i = a_k' u_{ik}, \qquad \sigma_a = a_k' \rho_{ak},$$

where the coefficients a_k are functions of x to be determined and the variables a_k' are derivatives of the coefficients a_k with respect to x. Primes attached to expressions involving u_i, ρ_a or v_i, σ_a will always indicate derivatives of those expressions with respect to x calculated as if the coefficients a_k, a_k' were independent of x. One readily verifies, then, the relations

$$(136) \qquad (\Omega_{u_i'})' = \Omega_{u_i}, \quad (\Omega_{v_i'})' = \Omega_{v_i}, \quad u_i\Omega_{v_i'} - v_i\Omega_{u_i'} = 0,$$
$$(d/dx)\Omega_{u_i'} = (\Omega_{u_i'})' + \Omega_{v_i'} = \Omega_{u_i} + \Omega_{v_i'}$$

in which it is understood that the differentiation indicated by d/dx takes account of the fact that the coefficients a_i are functions of x.

Let the functions $\eta_i(x)$ be a set of admissible variations along the arc E_{12}, satisfying therefore the equations $\Phi_a = 0$. The equations

$$\eta_i = u_i = a_k u_{ik}, \quad \mu_a = \rho_a = a_k \rho_{ak}$$

determine uniquely the coefficients a_k and the multipliers μ_a as functions of x on an interval $x_1 - h \leq x \leq x_1 + h$ chosen so small that on it the determinant $| u_{ik} |$ is everywhere different from zero. The derivatives η_i' have the values

$$\eta_i' = a_k u_{ik}' + a_k' u_{ik} = u_i' + v_i. \tag{137}$$

With the help of Taylor's formula, equation (113) of page 725, the equations (136) and (137) above, and the relations $\Omega_{\rho_a} = \Phi_a = 0$, one verifies the further relations

$$\begin{aligned} 2\omega(x, \eta, \eta') &= 2\Omega(x, \eta, \eta', \rho) = 2\Omega(x, u, u' + v, \rho) \\ &= 2\Omega(x, u, u', \rho) + 2v_i\Omega_{u_i'} + F_{y_i' y_k'} v_i v_k \\ &= u_i\Omega_{u_i} + u_i'\Omega_{u_i'} + \rho_a\Omega_{\rho_a} + 2v_i\Omega_{u_i'} + F_{y_i' y_k'} v_i v_k \\ &= u_i[\Omega_{u_i} + \Omega_{v_i'}] + (u_i' + v_i)\Omega_{u_i'} + F_{y_i' y_k'} v_i v_k \\ &= (d/dx)(\eta_i\Omega_{u_i'}) + F_{y_i' y_k'} (\eta_i' - u_i')(\eta_k' - u_k'). \end{aligned}$$

For arbitrary multipliers $\mu_a(x)$ taken with the functions $\eta_i(x)$ it follows therefore that

$$2\Omega(x, \eta, \eta', \mu) = (d/dx)\eta_i\Omega_{u_i'} + F_{y_i' y_k'} (\eta_i' - u_i')(\eta_k' - u_k')$$

and hence with the help of equation (113) on page 725 that

$$\eta_i[\Omega_{\eta_i} - (d/dx)\Omega_{\eta_i'}] + (d/dx)\eta_i(\Omega_{\eta_i'} - \Omega_{u_i'}) = F_{y_i' y_k'} (\eta_i' - u_i')(\eta_k' - u_k').$$

The last equation justifies the following lemma:

LEMMA. *There is an interval $x_1 - h \leq x \leq x_1 + h$ on which there exists no solution $\eta_i(x)$, $\mu_a(x)$ of the accessory equations, except the solution $\eta_i \equiv \mu_a \equiv 0$, whose elements $\eta_i(x)$ all vanish at two points x' and x'' of the interval; or, in other words, there is an interval on which no pair of values x', x'' can define conjugate points on E_{12}.*

This is clear since the last equation shows that for a system of solutions $\eta_i(x)$, $\mu_a(x)$ of the accessory equations the sum $\eta_i(\Omega_{\eta_i'} - \Omega_{u_i'})$ has a non-negative derivative on $x_1 - h \leq x \leq x_1 + h$, on account of the property III′ of E_{12}. If the functions $\eta_i(x)$ all vanish at two points x' and x'' the differences $\eta_i' - u_i' = v_i$ are identically zero on $x'x''$, and this implies that the derivatives a_k' are all zero and the coefficients a_k constants. But since the $\eta_i(x)$ vanish at x' and $| u_{ik} |$ is different from zero these coefficients are then all zero, and the functions $\eta_i(x)$ vanish identically on $x'x''$. The multipliers $\mu_a(x)$ are also zero on $x'x''$. Otherwise they would form with $\lambda_0 = 0$ a set of multipliers for E_{12}, as one readily sees by examining the accessory equations, and this is impossible since the extension of E_{12} is normal on $x'x''$ if the interval $x_1 - h \leq x \leq x_1 + h$ is taken sufficiently small.

As an immediate consequence of this lemma we have the following corollary:

COROLLARY. *There is an interval $x_1 - h \leq x \leq x_1 + h$ on which the determinant*

$$\Delta(x, x_0, a, b) = \begin{vmatrix} y_{ia_k}(x) & y_{ib_k}(x) \\ y_{ia_k}(x_0) & y_{ib_k}(x_0) \end{vmatrix},$$

formed for a family of extremals $y_i = y_i(x, a, b)$, $\lambda_a(x, a, b)$ as described in the theorem of page 687, can not vanish for any pair of points $(x_0, x) = (x', x'')$.

The solutions η_i, μ_a of the accessory equations are all expressible in the form

$$(138) \qquad \eta_i = c_k y_{ia_k} + d_k y_{ib_k}, \qquad \mu_a = c_k \lambda_{aa_k} + d_k \lambda_{ab_k},$$

as was indicated on page 727. If $\Delta(x'', x', a, b) = 0$ for points x', x'' on the interval $x_1 - h \leqq x \leqq x_1 + h$ then there would be constants c_k, d_k not all zero such that the solution (138) has $\eta_i(x') = \eta_i(x'') = 0$, and by the lemma it would follow that $\eta_i \equiv \mu_a \equiv 0$. In that case the corresponding functions

$$\zeta_i = c_k v_{ia_k} + d_k v_{ib_k} = \Omega_{\eta_i'}$$

would also vanish identically, which is impossible since the determinant

$$\begin{vmatrix} y_{ia_k} & y_{ib_k} \\ v_{ia_k} & v_{ib_k} \end{vmatrix}$$

of page 687 is by hypothesis different from zero.

CHAPTER V.

HISTORICAL REMARKS.

A complete history of the problem of Lagrange would require an extensive presentation. The remarks in the following paragraphs are a sketch only of the development of the theory, in which an effort will be made to point out the memoirs which have been especially significant in the preparation of this paper. For more detailed references one should consult the articles on the calculus of variations in the *Encyclopädie der Mathematischen Wissenschaften* by Kneser [1, II A 8] * and Zermelo and Hahn [1, II A 8 a], the translations and extensions of them by Lecat in the *Encyclopédie des Sciences Mathématiques* [2], and the treatise by Bolza [3].

* The numbers in square brackets refer to the following bibliography.

Euler [2, p. 119; 7, p. 114] and Lagrange [8, I, p. 347] both studied special cases of the Lagrange problem which led up to the formulation of the more general problem and its multiplier rule by Lagrange [8, X, p. 420]. The proof of the multiplier rule which Lagrange gave was incomplete. The missing details were provided by A. Mayer [9], Hilbert [10], and Kneser [11, Sections 57-8]. Hahn [12] extended to the multiplier rule for the problem of Mayer, which includes that of Lagrange as a special case, the methods which Du Bois Reymond had applied to simpler problems of the calculus of variations. The argument in the text above is new but was suggested by papers by Hahn [13, p. 271] and Bliss [16].

The distinction between normal and abnormal minimizing arcs seems to have been first mentioned by A. Mayer [9, p. 79] but was emphasized by von Escherich [17] in connection with his theory of the second variation where it played an important role. Hahn [18, p. 152] adopts the definition of von Escherich. The definitions in Sections 7 and 8 above are modeled after that of Bolza [19, p. 440] and are applied to simplify the proof of the multiplier rule in Section 15 for the case when the functions ϕ_α contain no derivatives.

The necessary condition analogous to that of Legendre for simpler problems was first proved for the problem of Lagrange by Clebsch [20] as one of the consequences of his rather elaborate theory of the second variation. The necessary condition analogous to that of Weierstrass seems to have been first proved by Hahn [21] who deduced therefrom the necessary condition of Clebsch without appeal to the theory of the second variation. The method in the text above is that of Bolza [22], who supplied a step missing in the proof of Hahn, but the method is here further simplified by the use of the auxiliary formulas of Section 21 which are generalizations of formulas emphasized by Goursat [23, p. 566].

For the Lagrange problem the necessary condition for a minimum analogous to that of Jacobi for simpler problems is due to A. Mayer [24]. The envelope theorem and the associated geometric proof of the Mayer condition are the work of Kneser [25]. The method of the preceding pages for the development of Kneser's theory is modeled after Bolza [26], but with simplifications due again to the use of the auxiliary formulas of Section 21. The analytic proof of the Mayer condition by means of the theory of the minimum problem of the second variation was suggested by Bliss [27] and applied to the Lagrange problem by D. M. Smith [28]. By this method the advantages of the analytic proof are preserved without the necessity of using any complicated theory of the transformation of the second variation.

The theory of the second variation has been elaborately developed by many writers. The most important of the early papers is that of Clebsch [29] in which he transformed the second variation into its so-called reduced form and derived therefrom his necessary condition analogous to that of Legendre for simpler problems. The methods of Clebsch were modified by A. Mayer [30] who proved the necessity of a condition analogous to that of Jacobi for simpler problems, the so-called condition of Mayer described in the preceding pages. In a series of papers von Escherich [31] discussed in great detail the theory of the second variation and the various consequences which can be deduced from it. A condensed treatment of his theory is given by Bolza [32]. Hahn [33] showed the relationship between the theory of the second variation and certain aspects of the theories of Weierstrass as extended to the problem of Lagrange. The theory of the second variation takes a relatively simple form when it is viewed from the stand-point of the theory of the minimum problem of the second variation, as has been shown by Bliss [27, 34, 35].

The best reference for the sufficiency theorems in Chapter IV above is Bolza [36] to whom the precise formulation of the theorems and many details of the proofs are due. The properties of fields and their relation to the invariant integral analogous to that of Hilbert for simpler cases were first discussed by A. Mayer [37], and further material pertinent to the sufficiency proofs was discussed by Bolza [38] and Carathéodory [39]. The reader may refer to Kneser [11, 2d ed., pp. 290 ff.] for sufficiency proofs for the Mayer problem, and to Bliss [35] for a proof of the integral formula of Weierstrass and other properties of fields for the Lagrange problem.

BIBLIOGRAPHY.

GENERAL REFERENCES.

1. Kneser, "Variationsrechnung," *Encyclopädie der Mathematischen Wissenschaften*, II A 8; Zermelo und Hahn, "Weiterentwicklungen der Variationsrechnung in den letzten Jahren," *ibid.*, II A 8a.

2. Lecat, "Calcul des Variations," *Encyclopédie des Sciences Mathématiques*, II 31 (1913 and 1916).

3. Bolza, *Vorlesungen über Variationsrechnung* (1909).

4. Hadamard, *Leçons sur le Calcul des Variations* (1910).

5. Bliss, *The Calculus of Variations* (1924).

6. Moigno-Lindelöf, *Calcul des Variations* (1861).

THE EULER-LAGRANGE MULTIPLIER RULE.

7. Euler, *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes* (1744).

8. Lagrange, *Oeuvres* I, X. See also the translation in Ostwald's *Klassiker der Exakten Wissenschaften*, Nr. 47, pp. 31 ff.

9. A. Mayer, "Begründung der Lagrangeschen Multiplicatorenmethode der Variationsrechnung," *Mathematische Annalen*, Vol. 26 (1886), pp. 74-82.

10. Hilbert, "Zur Variationsrechnung," *Göttinger Nachrichten* (1905), pp. 159-180; *Mathematische Annalen*, Vol. 62 (1906), pp. 351-368.

11. Kneser, *Lehrbuch der Variationsrechnung* (1900); 2d ed. (1925).

12. Hahn, "Über die Lagrangeschen Multiplicatorenmethode in der Variationsrechnung," *Monatshefte für Mathematik und Physik*, Vol. 14 (1908), pp. 325-342.

13. Hahn, "Über die Herleitung der Differentialgleichungen der Variationsrechnung," *Mathematische Annalen*, Vol. 63 (1907), pp. 253-272.

14. Bliss, "The Solutions of Differential Equations of the First Order as Functions of Their Initial Values," *Annals of Mathematics*, 2d Series, Vol. 6 (1905), pp. 49-68.

15. Bliss, "Solutions of Differential Equations as Functions of the Constants of Integration," *Bulletin of the American Mathematical Society*, Vol. 24 (1918), pp. 15-26.

16. Bliss, "The Problem of Mayer with Variable End-Points," *Transactions of the American Mathematical Society*, Vol. 19 (1918), pp. 305-314.

NORMAL ARCS.

17. von Escherich, see reference 31 below, Vol. 108 (1899), p. 1290.

18. Hahn, "Bemerkungen zur Variationsrechnung," *Mathematische Annalen*, Vol. 58 (1904), pp. 148-168.

19. Bolza, "Über den 'Anormalen Fall' beim Lagrangeschen und Mayerschen Problem mit gemischter Bedingungen und variabeln Endpunkten," *Mathematische Annalen*, Vol. 74 (1913), pp. 430-436.

THE NECESSARY CONDITIONS OF WEIERSTRASS AND CLEBSCH.

20. Clebsch, "Über die Reduction der zweiten Variation auf ihre einfachste Form," *Journal für die reine und angewandte Mathematik*, Vol. 55 (1858), pp. 254-273.

21. Hahn, "Über das allgemeine Problem der Variationsrechnung," *Monatshefte für Mathematik und Physik*, Vol. 17 (1906), pp. 295-304.

22. Bolza, see reference 3 above, pp. 603 ff.

23. Goursat, *Cours d'analyse*, Vol. 3, 4th ed. (1927), pp. 545-660.

THE NECESSARY CONDITION OF MAYER.

24. A. Mayer, "Über die Kriterien des Maximums und Minimums der einfachen Integrale," *Journal für die reine und angewandte Mathematik*, Vol. 69 (1868), pp. 238-263.

25. Kneser, "Die Jacobische Bedingung des Extremums bei einem allgemeinen Typus von Aufgaben der Variationsrechnung," *Mittheilungen der mathematischen Gesselschaft zu Kharkov*, Ser. 2, Vol. 7 (1902), pp. 253-267.

26. Bolza, see reference 3 above, pp. 613 ff.

27. Bliss, "Jacobi's Condition for Problems of the Calculus of Variations in Parametric Form," *Transactions of the American Mathematical Society*, Vol. 17 (1916), pp. 195-206.

28. D. M. Smith, "Jacobi's Condition for the Problem of Lagrange in the Calculus of Variations," *Transactions of the American Mathematical Society*, Vol. 17 (1916), pp. 459-475.

THE THEORY OF THE SECOND VARIATION.

29. Clebsch, see reference 20 above.

30. A. Mayer, see reference 24 above.

31. von Escherich, "Die zweite Variation der einfachen Integrale," *Sitzungsberichten der kaiserlichen Akademie der Wissenschaften in Wien*, Vol. 107 (1898), pp. 1191-1250, 1267-1326, 1383-1430; Vol. 108 (1899), pp. 1269-1340; Vol. 110 (1901), pp. 1355-1421.

32. Bolza, see reference 3 above, pp. 619-634.

33. Hahn, "Über den Zusammenhang zwischen den Theorien der zweiten Variation und der Weierstrass'schen Theorie die Variationsrechnung," *Rendiconti del Circolo Matematico di Palermo*, Vol. 29 (1910), pp. 49-78.

34. Bliss, "A Note on the Problem of Lagrange in the Calculus of Variations," *Bulletin of the American Mathematical Society*, Vol. 22 (1916), pp. 220-225.

35. Bliss, "The Transformation of Clebsch in the Calculus of Variations," *Proceedings of the International Mathematical Congress Held in Toronto*, Vol. 1 (1924), pp. 589-603.

SUFFICIENT CONDITIONS FOR A MINIMUM.

36. Bolza, see reference 3, pp. 635 ff.

37. A. Mayer, "Über den Hilbertschen Unabhangigkeitssatz in der Theorie des Maximums und Minimums der einfacher Integrale," *Leipziger Berichte*, Vol. 55 (1903), pp. 131-145.

38. Bolza, "Weierstrass' Theorem and Kneser's Theorem on Transversals for the Most General Case of an Extremum of simple Definite Integrals," *Transactions of the American Mathematical Society*, Vol. 7 (1906), pp. 459-488.

39. Carathéodory, "Die Methode der Geodätische Equidistanten und das Problem von Lagrange," *Acta Mathematica*, Vol. 47 (1926), pp. 199-236.

40. Kneser, see the reference 11, 2d ed., pp. 290 ff.

41. Bliss, see reference 35, p. 593.