RECENT DEVELOPMENTS IN THE CALCULUS OF VARIATIONS

BY

E. J. McSHANE

1. Since the somewhat indeterminate title of this lecture may have led someone to expect a complete history of the progress of the calculus of variations in the past half-century, I hasten to disclaim any such ambitious intention. My more modest aim is to present, in a manner which I hope will be intelligible to those of us whose chief interest does not lie in this field, a sketch of some of the most significant developments in the past twenty-five years. Moreover, I shall speak only of single integral problems, thereby passing by the recent solutions of the problem of Plateau [1–5];* and I shall not consider applications of the theory. I would feel quite uneasy about ignoring the applications to boundary value problems if the subject had not been so well reported on by W. T. Reid [6].

Let us then state the problems with which we shall be concerned. The simplest problem, often called the *free problem in non-parametric form*, is that of minimizing an integral

$$J[y] \equiv \int_{x_1}^{x_2} f(x, y, y')dx$$

in the class of all functions $y = y(x)$, continuous and having derivatives which are continuous except at a finite number of corners, for which $y(x_1)$ and $y(x_2)$ have assigned fixed values.† The problem as just stated is a plane problem. It becomes a problem in $(n+1)$-space if we choose to regard y as an n-tuple $(y^1, \cdots, y^n)$ and y' likewise.

The *free problem in parametric form* in n-space is that of minimizing an integral

$$J(C) \equiv \int_{t_1}^{t_2} F(y, y')dt,$$

where y denotes an n-tuple $(y^1, \cdots, y^n)$ and the curve

$$C: \quad y^i = y^i(t), \qquad t_1 \leqq t \leqq t_2, \; i = 1, \cdots, n,$$

has a tangent which turns continuously except at a finite number of corners. Since this integral is intended to be a function of the curve C alone

* Numbers in brackets refer to references at end of paper.

† Concerning differentability requirements we shall agree that f and all other functions involved have as many derivatives as needed, unless there is some specific question as to such differentiability requirements.

and not of the particular parametric representation of it, we demand that $J(C)$ remain invariant under the change of parameter

$$(1.1) \qquad t = t(\tau), \quad \tau_1 \leqq \tau \leqq \tau_2; \qquad t'(\tau) > 0, \quad t(\tau_1) = t_1, \quad t(\tau_2) = t_2.$$

For this it is necessary and sufficient that the integrand $F(y, r)$ in $J(C)$ be positively homogeneous of degree 1 in r,

$$F(y, kr) = kF(y, r), \qquad k > 0.$$

Immediate consequences of this homogeneity are

$$(1.2) \qquad r^iF_{,i}(y, r) = F(y, r), \qquad r^iF_{r^ir^j}(y, r)0,$$

where the subscripts denote partial derivatives and the repetition of the affix i indicates as usual summation over the values $1, \cdots, n$ of i.

If instead of seeking to minimize an integral $J[y] = \int f(x, y, y')dx$ in the class of all functions with fixed end values we seek to minimize it in the subclass of such functions for which one or more integrals

$$G^i[y] = \int_{x_1}^{x_2} g^i(x, y, y')dx$$

have assigned constant values γ^i, the problem is called an *isoperimetric problem*. The analogous problem in parametric form is self-suggesting. The problem derives its name from the earliest problem of this type, that of finding the curve having a given length L and enclosing the greatest area. In our notation, we are to minimize

$$\int \tfrac{1}{2}(yx' - xy')dt$$

(the negative of the area integral) in the class of closed curves for which

$$\int [(x')^2 + (y')^2]^{1/2}dt$$

has the assigned value L.

Proceeding at once to the most inclusive variations problem involving single integrals, we state the *problem of Bolza*. In (x, y, r)-space (the y and r being n-tuples) a region R and functions $f(x, y, r)$, $\phi_\beta(x, y, r)$, $(\beta = 1, \cdots, m < n)$ are given and in $(2n+2)$-dimensional space a function $g(x_1, y_1, x_2, y_2)$. We seek to minimize the sum

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

in the class of functions

$$y^i = y^i(x), \qquad x_1 \leqq x \leqq x_2,$$

which satisfy the differential equations

$$\phi_\beta(x,\, y(x),\, y'(x)) \equiv 0,$$

and whose end points determine a point $(x_1,\, y(x_1),\, x_2,\, y(x_2))$ on a point set S in $(2n+2)$-space. This set S may be defined by equations of the form

$$\psi_\mu(x_1,\, y_1,\, x_2,\, y_2) = 0, \qquad \mu = 1, \cdots, p \leqq 2n + 2,$$

or it may be represented parametrically,

$$x_s = x_s(\alpha_1, \cdots, \alpha_r), \qquad y_s^i = y_s^i(\alpha_1, \cdots, \alpha_r), \qquad s = 1, 2, r \leqq 2n + 1.$$

This problem obviously reduces to a fixed end-point problem if the point set S consists of a single point (x_1, y_1, x_2, y_2). It becomes the simplest free problem if $g=0$ and the equations $\phi_\beta=0$ are omitted. The isoperimetric problem is brought into this form by introducing new variables z^i subject to the differential equations

$$x^{i'} - g^i(x,\, y,\, y') = 0$$

and end conditions

$$z^i(x_1) = 0, \qquad z^i(x_2) - \gamma^i = 0.$$

If $g\equiv 0$ we have the *Lagrange problem*, with variable or fixed end points according to the choice of the ψ_μ. If $f\equiv 0$, we have the *problem of Mayer*. If some of the functions ϕ_β are independent of y', the problem is one in which the curves $y=y(x)$ are required to lie on assigned surfaces.

While the problem of Bolza is thus easily seen to include all the other standard single integral problems of the calculus of variations, it obviously does not follow that an important theorem concerning this problem remains important in all of its specialized forms. As a rather crude example, a theorem in which one of the hypotheses is that the integrand f is not zero might be important for Lagrange problems, but it would have no immediate application to Mayer problems.

Thus far the word "minimum" has meant "absolute minimum." The classical calculus of variations has been much more concerned with relative minima. These are of two kinds. A function $y=\gamma(x)$ gives a *strong relative minimum* to an integral $J[y]$ if $J[\gamma]\leqq J[y]$ for all functions $y(x)$ which differ from $\gamma(x)$ by less than some positive ϵ; it gives a *weak relative minimum* if $J[\gamma]\leqq J(y)$ when $y(x)$ and $y'(x)$ differ from $\gamma(x)$ and $\gamma'(x)$ respectively by less than ϵ. Consider now the simplest problem. If $y=\gamma(x)$ gives a weak relative minimum to $J[y]$, and $y=y(x,t)$ is a family of curves passing through the given end points and containing the given curve for $t=0$, the function

$$J_t \equiv \int_{x_1}^{x_2} f(x,\, y(x,t),\, y_x(x,t))dx$$

will have a minimum at $t=0$. It follows that

$$\frac{d}{dt} J_t = \frac{d}{dt} \int_{x_1}^{x_2} f(x, y(x, t), y_x(x, t))dx = 0, \tag{1.3}$$

$$\frac{d^2}{dt^2} J_t \geqq 0, \tag{1.4}$$

for $t=0$. Writing $\eta(x)$ for the "variation" $y_t(x, 0)$, the equation (1.3) becomes

$$\int_{x_1}^{x_2} (f_y \eta + f_{y'} \eta')dx = 0.$$

This must hold for all η vanishing at x_1 and x_2, for the family $y(x,t) = \gamma(x) + t\eta(x)$ has $\eta(x)$ for its variation. By methods older than our Society we deduce two conclusions.

(i) *The Euler equation*

$$\frac{d}{dx} f_r(x, \gamma, \gamma') - f_y(x, \gamma, \gamma') = 0$$

holds between corners of the curve.

(ii) *At a value of x which defines a corner of the curve the Weierstrass-Erdmann corner condition*

$$f_r(x, \gamma(x), \gamma'(x - 0)) = f_r(x, \gamma(x), \gamma'(x + 0))$$

holds.

By definition, an *extremal* is a curve $y = \gamma(x)$ such that γ has continuous first and second derivatives and satisfies the Euler differential equation.

Somewhat more generally, the beginning point may vary on a curve $x = x_1(\alpha)$, $y = y_1(\alpha)$. If the integral represents arc length, it is obvious that the minimizing curve $y = \gamma(x)$ must meet the end-locus orthogonally. The easy generalization of this to general integrands takes the form that the minimizing curve must meet the end-locus *transversally*, by which we mean that the equation

$$(f - \gamma' f_r)x_1' + f_r y_1' = 0$$

must hold. Here the arguments of f and f_r are $(x_1, \gamma(x_1), \gamma'(x_1))$, while (x_1', y_1') is the vector tangent to the end-locus at the point of intersection.

Obvious analogues of these statements hold for parametric problems. For the Bolza problem, analogues hold, but they are no longer obvious or easily established. In place of the Euler equation, we have the *Lagrange multiplier rule*, which is as follows [7–9]: There is a number λ_0 and a set of functions $\lambda_1(x), \cdots, \lambda_m(x)$ (not all zero), continuous except perhaps at corners of the minimizing curve $y = \gamma(x)$, such that for the combination

$$F(x, y, r; \lambda) \equiv \lambda_0 f(x, y, r) + \lambda_\beta(x)\phi_\beta(x, y, r)$$

the Euler equations

$$\frac{d}{dx} F_{r^i} - F_{y^i} = 0, \qquad i = 1, \cdots, n,$$

hold between corners, while the Weierstrass-Erdmann condition holds at corners. Moreover a transversality condition holds which is a fairly direct generalization of that for the simplest case, but we shall not write it explicitly.

2. Returning to the simplest problem we will utilize the inequality (1.4). Straightforward differentiation shows that this can be written in the form

$$I[\eta] = \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx \geqq 0, \tag{2.1}$$

where

$$2\omega(x, \eta, \rho) = f_{rr}(x, \gamma(x), \gamma'(x))\rho^2 + 2f_{yr}(x, \gamma(x), \gamma'(x))\eta\rho + f_{yy}(x, \gamma(x), \gamma'(x))\eta^2.$$

It is not easy to verify whether or not inequality (2.1), as it stands, is satisfied along a given curve. Theoretically at least it would require the investigation of every function $\eta(x)$ vanishing at x_1 and at x_2. Therefore, just as in studying the first variation (1.3), we seek to bring the condition (2.1) into more usable form. The first method proposed for this is more than a century old. It involves a transformation of the second variation $I[\eta]$ into other forms [10, pp. 55, 61, 226, 619–634; 11]. This transformation is rather simple for the simplest problem. For the parametric problem, it already calls for ingenuity, and the complexity of the transformation increases rapidly with that of the original problem, reaching formidable proportions for the general Bolza problem. To avoid these analytic difficulties, Kneser developed a geometric method [12; 10, chap. 7; 13] which, at least in the simpler cases, was quite elegant, but which left some special situations undiscussed. The study of these special situations called for at least as much effort as that of the original case.

In 1916 Bliss [14–16] circumvented these troubles by remarking that the inequality $I[\eta]\geqq 0$ is equivalent to the statement that the function $\eta\equiv 0$ minimizes the integral $I[\eta]$ in the class of all functions $\eta(x)$ vanishing at x_1 and at x_2. Thus the discussion of the second variation becomes the study of a new variations problem, that of minimizing $I[\eta]$. The Euler equation for this "accessory" problem is called the Jacobi equation for the original problem; it is $d\omega_\rho/dx - \omega_y = 0$.

Two points x_3, x_4 are said to determine *conjugate points* on the extremal E: $y=\gamma(x)$ if there is a solution of the Jacobi equations which vanishes at x_3 and x_4 without vanishing identically. Jacobi's necessary condition is that, if the extremal E gives a weak relative minimum to $J[y]$ and along it $f_{rr}\neq 0$, there is no point $(x_3, \gamma(x_3))$ conjugate to the beginning point and having $x_1<x_3<x_2$. Suppose there were such a point x_3; we must show that E would not minimize the integral. Let $\eta(x)$ satisfy the Jacobi equations and vanish at x_1 and x_3 without vanishing identically. Let $u(x)$ equal

$\eta(x)$ for $x_1 \leqq x \leqq x_3$, while it vanishes identically for $x > x_3$. A simple calculation shows that $I[u] = 0$. But since $u[x]$ has a corner at x_3 the relation

$$\omega_\rho(x_3, a(x_3), u'(x_3 + 0)) - \omega_\rho(x_3, u(x_3), u'(x_3 - 0))$$
$$= f_{rr}(x_3, y(x_3), y'(x_3))(u'(x_3 + 0) - u'(x_3 - 0)) = 0$$

fails to hold. That is, the Weierstrass-Erdmann condition is *not* satisfied, and u does *not* minimize $I[\eta]$. So for some η_0 we must have $I[\eta_0] < 0$, violating the condition $I[\eta] \geqq 0$ which is necessary for a minimum.

Bliss first set forth these ideas in connection with the parametric problem [14]. Here the situation is complicated by the fact that the Jacobi equations are not independent, so that there is a superfluity of solutions. In fact, if $y^i = \gamma^i(t)$ is an extremal, then for every continuously differentiable function $\rho(t)$ the product $\rho\gamma^{i'}$ satisfies the Jacobi equations. This lack of independence had been a source of annoyance in transformations of the second variation. Bliss disposed of it by considering only "normal" solutions $\eta(t)$ of the Jacobi equations, that is, solutions for which $\gamma^{i'}\eta^i$ vanishes identically. More recently, other devices for eliminating superfluous solutions have been proposed [17, 18, 19], designed so that the selected solutions $\eta(t)$ will be of a form suitable for some special purpose. Nevertheless, Bliss' original definition of normal solutions can, by a slight modification, be made to answer all special needs so far encountered.*

Since 1916, this method has been applied to variations problems of all degrees of generality with such consistency and success that now it is referred to as the "usual" or "classical" method. It is worthy of remark that a simple by-product of this line of reasoning is a derivation of a transformation of the second variation [15] without any of the massive analytical machinery of von Escherich and his predecessors.

3. In discussing the first and second variations for the simplest problem, it was most convenient that, for every function $\eta(x)$ vanishing at x_1 and x_2, there was a family $y(x, t)$ of functions, all having the assigned end values $y(x_1, t) = \gamma(x_1)$ and $y(x_2, t) = \gamma(x_2)$, containing the extremal E: $y = \gamma(x)$ for $t = 0$, and having $\eta(t)$ for its variation function $y_t(x, 0)$. All we had to do was to define

$$y(x, t) = \gamma(x) + t\eta(x). \tag{3.1}$$

When side conditions are present, matters are not so simple. Consider for example, the Bolza problem. The simple definition (3.1) now fails for the functions $\gamma(x) + t\eta(x)$ will not in general satisfy the differential equations $\phi_\beta = 0$. In fact, if

$$y = y(x, t), \qquad x_1(t) \leqq x \leqq x_2(t), \tag{3.2}$$

* Such as invariance of selected solutions under change of parameter and change of coördinates, and applicability to problems with end points variable on manifolds.

is a family of curves satisfying the differential equations

$$\phi_\beta(x,\ y(x,\ t),\ y_x(x,\ t)) = 0 \tag{3.3}$$

and the end conditions

$$\psi_\mu(x_1(t),\ y(x_1(t),\ t),\ x_2(t),\ y(x_2(t),\ t)) = 0 \tag{3.4}$$

for all t near 0, and if the given extremal E is defined by equations (3.2) with $t=0$, then by differentiating in (3.3) and setting $t=0$ we obtain

$$\Phi_\beta(x,\ \eta,\ \eta') \equiv \phi_{\beta y^i}\eta^i + \phi_{\beta r^i}\eta^{i\prime} = 0, \tag{3.5}$$

where

$$\eta^i(x) = y_t^i(x,\ 0). \tag{3.6}$$

Likewise, by differentiating in (3.4) and setting $t=0$, we obtain

$$\Psi_\mu(\xi_1,\ \eta(x_1),\ \xi_2,\ \eta(x_2)) = 0, \tag{3.7}$$

where

$$\xi_s = x_s'(0), \qquad s = 1,\ 2. \tag{3.8}$$

(We do not need the specific form of the functions Ψ_μ.)

Therefore in order that a set $(\xi_1,\ \xi_2,\ \eta(x))$ be the set of variations derived (by (3.6) and (3.8)) from a family of curves satisfying the differential equations and end conditions, it is necessary that equations (3.5) and (3.7) hold. Annoyingly, these conditions are still not sufficient. If we disregard the end conditions, we can show by theorems on differential equations that every set $\eta(x)$ satisfying the equations of variation (3.5) is derived (by (3.6)) from a family (3.2) of curves satisfying the equations $\phi_\beta=0$. But this family may fail to satisfy the end conditions $\psi_\mu=0$. If we try to determine the end points $x_1(t)$, $x_2(t)$ in such a manner that the end conditions $\psi_\mu=0$ hold, we are essentially trying to solve a set of equations for an unknown function. Recalling the implicit function theorem, we will not be astonished to find that in order to carry through the analysis it is necessary that a certain determinant be not zero. In this case, we need to have a system of sets

$$(\xi_{\nu,1},\ \xi_{\nu,2},\ \eta_\nu^i(x)), \qquad \nu = 1,\ \cdots,\ p,$$

such that the functions $\eta_\nu(x)$ satisfy equations (3.5) and the determinant

$$\left|\ \Psi_\mu(\xi_{\nu,1},\ \eta_\nu(x_1),\ \xi_{\nu,2},\ \eta_\nu(x_2))\ \right|, \qquad \mu,\ \nu = 1,\ \cdots,\ p, \tag{3.9}$$

is different from zero.

The case in which the determinant (3.9) can be made different from zero is thus obviously more tractable than that in which (3.9) vanishes identically. Moreover, we may reasonably expect it to be the usual case.

Accordingly, it has been named the "normal" case,* and in it the extremal E is said to be "normal with respect to the end conditions $\psi_\mu=0$." It is evident that if E is normal with respect to the end conditions $\psi_\mu=0$, it remains normal with respect to any subset of these conditions. Consequently, the most drastic type of normality requirement is that of normality with respect to fixed end points.

The distinction between normal and abnormal arcs was of course ignored in the days of Lagrange. After it was recognized, abnormal cases were systematically avoided, and rigid normality requirements imposed on all arcs considered. The necessary condition of Weierstrass was established, not for abnormal arcs, nor even for all normal arcs, but only for arcs such that every subarc is normal with respect to fixed end points [10, p. 603; 7]. Sets of sufficient conditions [10, §77; 7] included the even stronger hypothesis that the extremal E: $y=\gamma(x)$, $x_1 \leqq x \leqq x_2$, could be extended so as to be defined on a larger interval $x_1-\epsilon \leqq x \leqq x_2+\epsilon$, remaining normal with respect to fixed end points on every subarc of this extended arc.

Normality conditions of this type are undesirable for three reasons. First, we do not wish to restrict results by imposing somewhat artificial conditions designed to exclude the more refractory problems, if it is possible to avoid the use of such conditions. From this point of view it would be desirable to proceed without any use of normality. While this is beyond our reach at present, still we can hope to minimize normality requirements, and, in particular, to avoid the exceedingly drastic requirement of normality on every subarc. Second, even though the normal case is in a sense the usual case, still, in each separate instance, we must verify that the specific arc under consideration does not happen to be abnormal. This verification may be far from easy, especially when it must be made for every subarc. Third, the Mayer problem is a Bolza problem with $f(x, y, r)\equiv 0$; but, as remarked by Carathéodory [21, 22] and others, if $y=\gamma(x)$ is a minimizing curve for this Bolza problem, it cannot possibly be normal (with respect to fixed end points) on any subarc. Consequently, if the theory of the Bolza problem is to be developed to cover that of the more special problems, we must not restrict our attention to problems normal on all subarcs.

Graves [23] showed that the requirement of normality on subarcs is not needed in establishing the necessary condition of Weierstrass. He established the necessity of the Weierstrass condition under the hypothesis that the arc itself is normal with respect to the end conditions $\psi_\mu=0$. Also, for a class of abnormal arcs, he gave a necessary condition analogous to that of Weierstrass, but more complex in its statement.

---

* The distinction between normal and abnormal cases appears already in the problem of minimizing a function $f(y)$ of the variable $y_1, \cdots, y^n$ subject to conditions $g_\mu(y)=0$, $\mu=1, \cdots, p<n$. Bliss [20] has made use of this in presenting a highly readable and instructive introduction to the concept and uses of normality.

The first set of conditions not requiring normality and still sufficient to ensure a minimum was given by Hestenes [24]. His conditions were the following: (I′) The functions $y^i = \gamma^i(x)$ are continuous with their first and second derivatives, and satisfy the Euler equations with continuously differentiable multipliers $\lambda_0 \neq 0$, $\lambda_\beta(x)$. (We can then take $\lambda_0 = 1$). (II_N') The strengthened Weierstrass condition holds along the curve $y = y_0(x)$. (III′) The strengthened Clebsch condition holds; that is, the form

$$u^i F_{r^i r^j}(x, \gamma, \gamma') u^j$$

is positive definite. (IV′) The second variation is positive. This last condition Hestenes showed to be a consequence of any one of three other hypotheses. Marston Morse [25] showed that (IV′) is also a consequence of a hypothesis which is a direct generalization of the Jacobi condition, stated in terms of the zeros of solutions of the Jacobi equations.

A very different method of reaching the same results was used by W. T. Reid [26]. He made no use of fields or of any part of the Weierstrassian theory, returning instead to the venerable expansion proof. This style of proof is based on a Taylor expansion with an estimate of the remainder term. It had been attempted frequently, had yielded a sufficiency theorem for weak relative minima, and finally had been successfully applied by E. E. Levi [27] to establishing sufficient conditions for a strong relative minimum for problems involving no side conditions. Reid's proof uses a rather simple but important lemma on the existence of a system of solutions $\eta_1, \cdots, \eta_n$ of the Jacobi equations and equations of variation

$$\frac{d}{dx}\Omega_{\rho^i} - \Omega_{\eta^i} = 0, \qquad \phi_{\beta y^i}\eta^i + \phi_{\beta r^i}\eta^{i\prime} = 0,$$

which are conjugate in the sense that

$$\eta_k^i \Omega_{\rho^i}(x, \eta_j, \eta_j') - \eta_j^i \Psi_{\rho^i}(x, \eta_k, \eta_k') = 0, \qquad i, j = 1, \cdots, n,$$

and whose determinant $|\eta_j^i(x)|$ is not zero. This lemma was proved independently by Reid, Morse, and Hestenes. Except in this lemma, Reid's proof proceeds without any need of the concept of normality.

The proofs just mentioned apply to the case of separated end conditions, in which none of the functions $\psi_\mu(x_1, y_1, x_2, y_2)$ occurring in the end conditions $\psi_\mu = 0$ actually depend, both on (x_1, y_1) and on (x_2, y_2). The general problem is reducible to this by a device. More recently, Hestenes [28] and Reid [29] have given proofs which apply directly to general (not necessarily separated) end conditions.

It is to be remarked that if the problem is normal, every set of multipliers λ_0, $\lambda_\beta(x)$ has $\lambda_0 \neq 0$. The theorems under discussion require less than this. It is asked only that there exist at least one set of multipliers with $\lambda_0 \neq 0$. Thus the only case uncovered is what might be called the "pure abnormal" case, in which *every* system of multipliers λ_0, $\lambda_\beta(x)$ with which the

Euler equations hold is such that $\lambda_0=0$. This pure abnormal case seems to be hopelessly unmanageable at present. Recently, Bliss has shown [20] that in so far as sufficiency theorems are concerned, the only important distinction is between the pure abnormal and the normal cases. Problems which are abnormal, but still have multipliers λ_0, $\lambda_\beta(x)$ with $\lambda_0\neq 0$, can be quite readily replaced by other problems, equivalent with respect to sufficiency conditions, which are normal with respect to the end conditions. Therefore, it is possible to simplify the proofs of Hestenes, Morse, and Reid by considering only normal extremals, and finally to recover the full generality of these theorems by using the device of Bliss.

Thus the treatment of the problem of Bolza from the classical point of view has now reached a state of completeness, at least in so far as normal arcs are concerned. For these, the necessary and the sufficient conditions are separated only by that gap which is already present in the simplest problem. For abnormal arcs not purely abnormal, the sufficiency theorems are essentially complete, but the same cannot be said of the necessary conditions. It does not seem likely that any simple and general results on such problems will soon be reached.

4. The problems which we have been discussing have required the finding of that one of a given class of functions or curves for which a certain integral assumes its least value. Heretofore we have considered functions or curves which give a *relative* minimum to the integral. It is quite another thing to ask if there actually exists a function or curve in the class for which the integral takes on its least value. Seventy-five years ago it was accepted that if the integrand were positive the existence of a minimum was evident. Weierstrass removed this misapprehension by a simple example. If we are to minimize the integral

$$J[y] = \int_0^1 | y' |^{1/2} dx$$

in the class of all functions having $y(0)=0$ and $y(1)=1$, we see that the lower bound of the values of the integral is zero; for it is never negative, and if $y=x^{2n+1}$, then $J[y]=(2n+1)^{1/2}(n+1)^{-1}$, which is arbitrarily close to zero. But $J[y]$ can never assume its lower bound zero; for if $J[y]=0$ then $y'\equiv 0$, which is incompatible with the end conditions $y(0)=0$, $y(1)=1$.

A similar example in parametric form is not difficult to construct. We first observe that $2x(2^{1/2}-x)\leqq 1$, so that $1/(2^{1/2}-x)\geqq 2x$ if $x<2^{1/2}$. Let us seek to minimize the integral $J(C)=\int F dt$ in the class of rectifiable curves C joining $(0, 0)$ and $(1, 0)$, where

$$(4.1)\qquad F(x, y, \dot{x}, \dot{y}) = (1 + y^2)(\dot{x}^2 + \dot{y}^2)/[(2\dot{x}^2 + 2\dot{y}^2)^{1/2} - \dot{x}].$$

(The integral is a Lebesgue integral, and the dot denotes differentiation with respect to t.) It is convenient to choose the parameter t to be arc length, so that $\dot{x}^2+\dot{y}^2=1$ for almost all t. Then if $L(C)$ is the length of C we have

$$(4.2)\qquad \begin{aligned} J(C) &= \int_0^{L(C)} (1+y^2)/(2^{1/2}-\dot{x})dt \\ &\geqq \int_0^{L(C)} 1/(2^{1/2}-\dot{x})dt \\ &\geqq \int_0^{L(C)} 2\dot{x}\,dt = 2x(L(C)) - 2x(0) = 2. \end{aligned}$$

The value $J(C)$ is arbitrarily near 2 if evaluated along a polygon whose sides all have slopes ± 1 and which lies in a sufficiently narrow strip $|y|<\epsilon$. But there is no curve for which $J(C)=2$. If there were such a curve C: $x=x(t)$, $y=y(t)$, then for it all members of (4.2) would be equal. This would imply that for almost all t, $1+y^2=1$ and $1/(2^{1/2}-\dot{x})=2\dot{x}$. The first of these can hold only if $y\equiv 0$, so that $\dot{y}=0$ and $\dot{x}=\pm 1$. But this contradicts the second.

In 1900 Hilbert [30] published a short but suggestive paper expounding a method by which theorems on the existence of minimizing curves can be established for a very important class of problems. Consider the problem of minimizing an integral $J(C)=\int F\,dt$ (in parametric form) in the class of all curves lying in a bounded closed set A and joining two given points. We first assume that

$$(4.3)\qquad F(y, r) \geqq \epsilon |r|,$$

where ϵ is positive and $|r|$ denotes the length of the vector r. Then $J(C)\geqq 0$; let μ be its lower bound. Let $\{C_n\}$ be a sequence of curves $y=y_n(t)$ for which $J(C_n)$ tends to μ. If we denote the length of C by $L(C)$, we have by (4.3)

$$L(C_n) = \int |\dot{y}_n|\,dt \leqq \int (F/\epsilon)dt = \epsilon^{-1}J(C_n) \to \mu/\epsilon.$$

Hence the curves C_n have a bound on their lengths. We can therefore apply Hilbert's convergence theorem, which states that if C_n is a sequence of curves of uniformly bounded lengths lying in a bounded point set, it is possible to select a subsequence C_m and represent these curves C_m by equations $y^i=y_m^i(t)$, $(0\leqq t\leqq 1)$, in such a manner that the functions $y_m^i(t)$ converge uniformly to limit functions $y_0^i(t)$. Now we have a curve C_0: $y=y_0(t)$, $(0\leqq t\leqq 1)$, which is the limit of our minimizing sequence. It does not follow from our hypotheses that C_0 is a minimizing curve; our example (4.1) shows this. Let us now add the hypothesis that the integral $J(C)$ is *regular*, which by definition means that the quadratic form $F_{r^ir^j}(y, r)u^iu^j$ is positive for all y in A, all $r\neq 0$, and all u which do not satisfy the equations

$u^i = kr^i$. (If $u^i = kr^i$ the quadratic form necessarily vanishes, as follows from equation (1.2)). It is then possible to utilize the classical field construction on successive small arcs of C_0 in such a way as to show that C_0 actually minimizes $J(C)$. Hilbert gave this proof for plane problems. A similar discussion was applied by Carathéodory [31] to certain classes of non-regular integrals, and Bill [32] extended Hilbert's proof to space problems.

The situation can be described in topological terms by observing that the aggregate of all continuous curves $y = y(t)$, $(t_1 \leqq t \leqq t_2)$, lying in a euclidean space (or more generally, in a metric space) can be made into a metric space by a suitable definition of the distance $\|C_1, C_2\|$ between all pairs of curves C_1, C_2. Such a definition was given by Fréchet. For our purposes it is enough to observe that C_n tends to C_0, that is, $\lim \|C_n, C_0\| = 0$, if and only if each C_j, $(j = 0, 1, \cdots)$, can be parametrically represented by functions $y = y_j(t)$, $(0 \leqq t \leqq 1)$, in such a manner that $y_n(t)$ tends to $y_0(t)$ uniformly on the interval $0 \leqq t \leqq 1$.

Let us confine our attention to the (metric) subspace consisting of all curves C which are rectifiable (that is, have finite length) and whose points $y(t)$ lie in a given point set A. To each of these curves the integral $\int F\, dt$ assigns a functional value $J(C)$. We can now state our problem in the following form. Given a class K of rectifiable curves whose points lie in A, and a function $J(C)$ defined on K, we seek a curve C_0 in K such that $J(C) \geqq J(C_0)$ for all curves C in the class K.

It is natural to approach this problem as before; we select first a sequence $\{C_n\}$ of curves of K such that $J(C_n)$ tends to the greatest lower bound μ of $J(C)$ on K. If the set A is bounded and closed, and $J(C)$ has some property which ensures that the C_n have uniformly bounded lengths (for instance, if $F(y, r) \geqq \epsilon |r|$), then by Hilbert's convergence theorem there is a curve C_0 which is the limit of a subsequence of $\{C_n\}$. There is no loss of generality in supposing that this subsequence is the whole sequence $\{C_n\}$. The curve C_0 is a rectifiable curve and its points lie in A, but this alone does not imply that C_0 belongs to the class K. We therefore impose the hypothesis that K is a *complete* class, that is, that it contains all its rectifiable limit curves. For example, K may consist of all rectifiable curves in A joining two fixed points, or joining two fixed closed point sets; there are many other types of examples.

If now the function $J(C)$ were continuous, from the relations $C_n \to C_0$ and $J(C_n) \to \mu$ we would at once have $J(C_0) = \mu$, so that C_0 would be the minimizing curve sought. But it is easy to see that few of the interesting integrals of the calculus of variations yield continuous functions $J(C)$. Even the length integral is not continuous, for arbitrarily near any given curve C we can construct other curves so crinkly as to have lengths greatly in excess of the length of C. However, continuity is much more than we need. In fact, if $J(C)$ is lower semicontinuous, which by definition means that it satisfies the condition

$$\liminf_{n\to\infty} J(C_n) \geqq J(C) \tag{4.4}$$

whenever C_n tends to C, then the curve C_0 is the minimizing curve sought. For then

$$\mu = \lim J(C_n) \geqq J(C_0)$$

by (4.4), while $J(C_0) \geqq \mu$ by the definition of μ. So $J(C_0) = \mu$, as was to be proved.

It was Tonelli [33] who first stressed the importance of this concept of lower semicontinuity in the calculus of variations, and used it to obtain powerful existence theorems. The integrals $J(C)$ which possess this property include all the regular integrals. In fact, if we consider only classes of curves of uniformly bounded lengths (and this restriction on lengths is necessary if we are to use Hilbert's theorem on convergence of curves), semicontinuity follows from the weaker condition that the inequality

$$F_{r^i r^j}(y, r)u^i u^j \geqq 0 \tag{4.5}$$

holds for all y in A, all $r \neq 0$ and all u. More than this, it even follows from the condition that for each fixed y in A the surface (called the *figurative*) defined by the equation

$$z = F(y, r) \tag{4.6}$$

is convex. By the homogeneity of F, this surface is a cone with vertex at the origin; and for functions F which are twice differentiable, the figurative is convex if and only if (4.5) holds.

It is not very difficult to establish the theorem that the convexity of the figurative implies the lower semicontinuity of $J(C)$ on any class of curves of uniformly bounded lengths. Suppose that the curves C_n: $y = y_n(t)$, $(0 \leqq t \leqq 1)$, converge to C_0: $y = y_0(t)$, $(0 \leqq t \leqq 1)$, the functions y_n converging uniformly to y_0 and the derivatives y_n' being all less in absolute value than a constant M. If F is independent of y and $y_0'(t)$ is constant, we can apply Jensen's inequality to the convex function $F(r)$, obtaining

$$\begin{aligned}\int_0^1 F(y_n')dt \geqq F\left(\int_0^1 y_n' dt\right) &= F(y_n(1) - y_n(0)) \\ \to F(y_0(1) - y_0(0)) &= F(y_0'(t)) = \int_0^1 F(y_0')dt.\end{aligned}$$

That is,

$$\liminf J(C_n) \geqq J(C_0). \tag{4.7}$$

(Observe that here we need the convergence of $y_n(t)$ to $y_0(t)$ only at $t = 0$ and $t = 1$.) If C_0 is not a line segment we can inscribe in it a polygon Π: $y = \pi(t)$, $(0 \leqq t \leqq 1)$, with sides short enough so that $J(\Pi)$ is arbitrarily near

$J(C_0)$. Although y_n does not tend to π for all t, we still have $y_n(t)\to\pi(t)$ whenever t defines a vertex of Π. So we can apply the previous proof to each side of Π and obtain

$$\liminf J(C_n) \geqq J(\Pi);$$

and since $J(\Pi)$ is arbitrarily near $J(C_0)$, inequality (4.7) still holds. We can extend these results to the general case, in which F depends on y, by the traditional device of subdividing the interval $0\leqq t\leqq 1$ into short subintervals and using only large n, so that on each subinterval $F(y_j, y_j')$ is independent of y to within an arbitrarily small error ($j=0, 1, \cdots$). This type of proof of semicontinuity does not even require F to have any partial derivatives; it is enough that it be continuous in (y, r) and convex in r for each y.

Tonelli's many contributions to this direct method in the calculus of variations up to 1922 are summed up in his book [33]. In it four types of problems (all in the plane) are considered; the free problem (without side conditions) both in parametric and non-parametric form, and the isoperimetric problem, also in both forms. Since then the development has been considerable in volume and varied in direction.

An early improvement was a weakening of the conditions needed to ensure the existence of a convergent minimizing sequence. The condition

$$F(y, r) \geqq \epsilon \,|\, r \,|, \qquad \epsilon > 0, \tag{4.8}$$

used to keep a bound on the lengths of the curves C_n in a minimizing sequence, was replaced by a weaker condition by Hahn [34]. Hahn's theorem was improved by Carathéodory [35] and Tonelli [36], and extended to n-space by Graves [37] and McShane [38]. The result is that (4.8) can be replaced by the much weaker condition that the figurative (4.6) is convex and non-planar for each y, and there is a finite upper bound for the lengths of all curves C such that $J(C)\leqq 0$. This, with the semicontinuity theorem just discussed, yields a very strong existence theorem for problems in parametric form.

The problem in nonparametric form had been less exhaustively treated until five years ago. For such problems, the most straightforward approach is to show that the minimizing sequence $y=y_n(x)$ has an absolutely continuous limit function. This was the method used by Tonelli in his book. In order to control the behavior of the minimizing sequence, he assumed that there were positive constants a, b, ϵ such that

$$f(x, y, y') \geqq a \,|\, y' \,|^{1+\epsilon} - b$$

for all (x, y) in A and all y'. Nagumo [39] showed that it was sufficient to make the weaker hypothesis that

$$f(x, y, y')/\,|\, y' \,| \to \infty \qquad \text{as} \qquad |\, y' \,| \to \infty$$

uniformly in (x, y). The author [40] approached the problem indirectly. He restated the problem in parametric notation by defining

$$(4.9) \qquad \begin{aligned} F(x, y, x', y') &= x'f(x, y, y'/x'), \qquad x' > 0, \\ F(x, y, 0, y') &= \lim_{x_r \to +0} F(x, y, x', y'). \end{aligned}$$

Then by methods appropriate to the more highly developed parametric problem he showed the existence of a minimizing curve $x = x(t)$, $y = y(t)$ for $\int F\, dt$ in the class of curves with $x'(t) \geqq 0$. Adding any of several hypotheses on $f(x, y, r)$ ensured $x'(t) > 0$ almost everywhere, permitting a return to the nonparametric form. A number of new theorems resulted. Tonelli [41] showed that all these theorems could also be obtained by his method, and added several new ones. The author [42] established a semicontinuity theorem for integrals $\int F\, dt$ where for each y the vectors r may be restricted in direction, the function $F(y, r)$ being merely assumed to be a lower semicontinuous function of its arguments. The analytical assumptions are weak enough so that the theorem can be applied to all the standard single integral problems, both in parametric and in nonparametric form. By use of this semicontinuity theorem he established rather general existence theorems [43] for problems in non-parametric form.

It would be natural to suppose that this relaxing of the continuity requirements on the integrand, so as to permit us to study problems in which the integrand is merely lower semicontinuous, is nothing more than analytic gymnastics. While it is true that most of the problems suggested by physics and chemistry have analytic or at least many times differentiable integrands, the extension is not really trifling. For one thing, when even an analytic and regular nonparametric problem is transformed by equations (4.9) into parametric form it will usually turn out that the transformed integrand is merely lower semicontinuous at points having $x' = 0$. For another, the principal existence theorems for nonparametric problems have hypotheses which do not permit the integrand $f(x, y, r)$ to vanish identically in r for any (x, y). Tonelli has established special theorems to cover certain classes of integrals in which this identical vanishing occurs. These theorems, even in generalized form, can be brought under the principal theorems by a suitable transformation of the independent variable x. But the problem when thus transformed has an integrand which is only lower semicontinuous, and in fact is $+\infty$ identically in r and y for some values of x. Thus the extension of the theory to discontinuous integrands has a rather noticeable unifying effect.

Another existence theorem which does not require continuity and to some extent even dispenses with lower semicontinuity of the integrand has been established by Menger [44–46]. The setting for the theorem is an abstract metric space. The integrand $F(y, r)$, which may be thought of as determined by the point y and the vector from the point $q = 0$ to the point

r, is replaced by a function $\phi(p; q, r)$. This generates a "ϕ-distance" between points, and thus defines the ϕ-lengths of polygons. The functional $J(C)$, now the ϕ-length of C, is defined by inscribing polygons; here the procedure is somewhat like that used by Weierstrass in defining the integral $J(C)$ on rectifiable curves in days when only the Riemann integral was available. Under a certain convexity assumption, the ϕ-length can be proved lower semicontinuous and existence theorems can be established.

Another extension of the calculus of variations to metric spaces was made by Frink [47], who showed that the Carathéodory linear measure was lower semicontinuous on the class of continua and thus established the existence of continua of least linear measure.

Problems with side conditions have as yet not been so completely treated. If the side conditions are in the form of differential equations linear in y' the theorems for free problems apply, as Graves [42, 48, 49] pointed out. The reason that such side conditions cause us no trouble is that the only property required of the class K of curves is completeness, or rather less than that; the class must be such that if C_n is a sequence of curves K having uniformly bounded lengths and converging to C_0, then C_0 must also belong to K. It is quite easy to prove that the class of all curves satisfying a set of linear differential equations has this property. This remark is not trivial, for such side conditions occur (1) when the curves are restricted to lie on a given surface (although here it is easy to see directly that such a class K of curves is complete); (2) in isoperimetric problems for which the integrands in the side integrals do not involve the derivatives; (3) when the problem under consideration is itself the second variation of another Lagrange problem; and (4) when the problem of minimizing an integral

$$\int f(x, y, y', \cdots, y^{(n)})dx$$

involving derivatives of y of higher order is set into Lagrange form. More recently, the problem with higher derivatives in the integrand has been studied in greater detail by Cinquini [50, 51] and the author [43], and now has been as completely solved as the nonparametric problem with only first derivatives in the integrand.

Since, ordinarily, the curves $y = y(x)$ satisfying a system of differential equations $\phi_\beta = 0$ do not form a complete class, the general Mayer, Lagrange, and Bolza problems require special treatment. Manià studied in detail two well known problems, the navigation problem of Zermelo [52] and the problem of finding that curve of descent of a particle through a resisting medium which shall maximize the final velocity [53, 54], and was led to an existence theorem for a class of Mayer and Lagrange problems [55, 56]. The most general theorem of this type, applicable to a class of Lagrange, Mayer, and Bolza problems, is due to Graves [57].

A restriction common to all the theorems mentioned in the preceding paragraph is that certain of the functions $y^i(x)$ must have their final values entirely unrestricted. This does not hinder their application to a number of interesting special problems, but fixed end-point problems are excluded. Among these excluded problems are all isoperimetric problems, so that these problems require special treatment. Some theorems for such problems were established by Tonelli [33] and slightly generalized by McShane [58]. These theorems apply to the problem

$$J(C) = \int F(y, y')dt = \min, \qquad G(C) = \int G(y, y')dt = \gamma,$$

and require a rather close relationship between the integrands F and G. It would seem that methods of proof based on lower semicontinuity are not well fitted to the discussion of these problems. Granted that we have a convergent sequence $\{C_n\}$ for which $J\{C_n\}$ tends to the lower bound μ while $G(C_n)$ tends to γ, it is not very helpful to find that at the limit curve C_0 we know only that $G(C_0) \leqq \gamma$. In some papers as yet unpublished McShane [59] studies isoperimetric problems both in parametric and in nonparametric form without using semicontinuity. His method is an extension of one used by Lewy [60] for regular problems in the plane and later used by Manià [56] for a class of Mayer problems. First a particular type of minimizing sequence is found consisting of a sequence of polygons Π_n such that $J(\Pi_n) \to \mu$ and $G(\Pi_n) \to \gamma$. Each of these polygons Π_n has a minimizing property of its own, in that it minimizes $J(C)$ in the class of polygons having no more vertices than Π_n has and making $G(C)$ equal to $G(\Pi_n)$. From this minimizing property of the Π_n, it can be shown that these polygons satisfy certain relations which are approximations to the usual necessary conditions for a minimum. Two of these conditions are particularly useful. One is the Lagrange multiplier rule; there are numbers λ_0, λ_1 such that for the function $H(y, r) = \lambda_0 F(y, r) + \lambda_1 G(y, r)$ the Euler equations (in the integrated, or DuBois-Reymond, form)

$$H_{r^i}(y, y') = \int_a^t H_{y^i}(y, y')dt + c_i$$

are satisfied. The other is a corner condition established by Dresden [61]. The Dresden corner condition states that if t defines a corner of a minimizing curve $y = y(t)$, then

$$\Omega_H(y(t), y'(t-0), y'(t+0)) \leqq 0,$$

where $\Omega_H(y, p, r) \equiv p_i H_{y^i}(y, q) - q^i H_{y^i}(y, p)$. From the information which these relations yield concerning the derivatives of the functions $y_n(t)$ defining the polygons Π_n, it can be shown that under suitable hypotheses a minimizing sequence exists such that $y_n(t)$ converges uniformly to a limit function $y_0(t)$, while $y_n'(t)$ converges almost everywhere to $y_0'(t)$. It follows at once that for the curve C_0: $y=y_0(t)$ the equations

$$J(C_0) = \lim J(\Pi_n) = \mu, \qquad G(C_0) = \lim G(\Pi_n) = \gamma$$

hold.

It is evident that this style of proof does not require semicontinuity of the integrals involved. More than that, the hypotheses are not strong enough to imply semicontinuity. Thus even for free problems existence theorems are established which are new in that they apply to problems in which the figurative is not convex. (Some such theorems had already been established by Carathéodory.) The existence theorems thus established for isoperimetric problems in non-parametric form are the first of any generality.

5. Although many elementary problems in the differential calculus require the minimizing of a function, there are many other problems in which it is required to find a stationary point of a function. Consider, for example, a rigid convex body. At each point P of its surface we define $f(P)$ to be the distance from the center of mass C of the body to the plane tangent to the surface at P. In other words, if the body is placed on a horizontal table with the point P resting on the table, $f(P)$ is the height of the center of gravity above the table. In order that the body be in equilibrium when placed with the point P_0 touching the table, the function $f(P)$ must be stationary at P_0. If P_0 minimizes $f(P)$, the equilibrium is stable.

Likewise, in the calculus of variations, many important applications require the determination of a curve along which the first variation of a certain integral is zero. Such a curve is an extremal, at least if it is twice continuously differentiable. For instance, the principle of "least" action is of this type; what is sought is a curve of *stationary* action, that is, an extremal curve. In optics, Fermat's principle of "least" time is again of this type. This suggests the importance of studying the stationary points of functions and the extremal curves of integrals quite apart from their minimizing properties.

That such a study is really profitable was definitely established in 1917, when Birkhoff [62] enunciated his minimax principle and utilized it in establishing the existence of new classes of periodic orbits in dynamical problems. Birkhoff defined the minimax points of a function J as follows. *If J_0 is the value of J at a point P_0 and if the inequality $J<J_0-\epsilon$ where ϵ is small and positive defines more than one region near the point P_0, then P_0 will be called a point of minimax. If the inequality defines k regions in the neighborhood of P_0, that point will be said to be of multiplicity $k-1$.* It is easily seen that such a point P_0 is a *critical point* of J, by which we mean that all first partial derivatives of the function vanish at P_0; but P_0 is not a minimum point of the function. In fact, if we introduce a coordinate system $(x^1, \cdots, x^n)$ with origin at P_0 and represent the function J as a

function $f(x)$, then (assuming that f has continuous partial derivatives of order up to 3)

$$f(x) = f(0) + \tfrac{1}{2} f_{x^i x^j}(0) x^i x^j + R_3, \tag{5.1}$$

where R_3 is an infinitesimal of the third order with respect to $|x|$. The quadratic form in (5.1) governs the behavior of f near 0. If it is nonsingular, the critical point is called *nondegenerate*. In this case, by means of a linear transformation to new variables y, the quadratic form in (5.1) can be brought into the form

$$-(y^1)^2 - \cdots - (y^k)^2 + (y^{k+1})^2 + \cdots + (y^n)^2. \tag{5.2}$$

Birkhoff showed [62] that the critical point is a minimax point if and only if $k=1$ in the quadratic form (5.2). (Recall that this is under the assumption that the critical point is nondegenerate.)

Morse observed [63] that among nondegenerate critical points the minimum and minimax points represent two classes out of a total of $n+1$. For if at a nondegenerate critical point the second-order terms are brought into the form (5.2), then k (the *index* of the quadratic form $f_{x^i x^j} x^i x^j$) can have any one of the values $0, 1, \cdots, n$. If we say that k is the *type* of the critical point, then a critical point of type 0 is a minimum point, a critical point of type 1 is a minimax point, and there still remain critical points of type $2, 3, \cdots, n$, a critical point of type n being a maximum point. Morse investigated the interrelations between the numbers of critical points of different types and the Betti numbers of the point set R on which the function is defined. Suppose first that R is compact and without boundary, each point of R having neighborhoods which can be mapped topologically on the interior of a sphere in n-space. For example, if $n=2$ the set R can be the surface of a sphere or of a torus. If $p_0, p_1, \cdots, p_n$ are the Betti numbers of R, and f has only non-degenerate critical points (necessarily finite in number), and M_k is the number of critical points of type k, the following relationships must hold, as Morse proved:

$$\begin{aligned}
M_0 &\geqq p_0, \\
M_1 - M_0 &\geqq p_1 - p_0, \\
M_2 - M_1 + M_0 &\geqq p_2 - p_1 + p_0, \\
&\cdots\cdots\cdots\cdots, \\
M_{n-1} - M_{n-2} + \cdots + (-1)^{n-1} M_0 &\geqq p_{n-1} - p_{n-2} + \cdots + (-1)^{n-1} p_0, \\
M_n - M_{n-1} + \cdots + (-1)^n M_0 &= p_n - p_{n-1} + \cdots + (-1)^n p_0.
\end{aligned} \tag{5.3}$$

The first of these is obvious; in each separate portion of R the function has at least one minimum. The second expresses Birkhoff's minimax principle; it was established by him for analytic functions without assuming that the minimax points are nondegenerate. The final equation was known to

Poincaré [64], at least for $n=2$; a closely related theorem on the sum of the indexes of fixed points of transformations was established by Lefschetz [65] and by Hopf [66] at about the same time that Morse established the relations (5.3).

It might be interesting to look at a rather crude and heuristic sketch of a proof of these inequalities for the case $m=2$. Assume that all the critical points $P_1, \cdots, P_n$ are non-degenerate and therefore finite in number, and that the critical values $f(P_j)$ are distinct. For purposes of visualization think of $f(P)$ as represented by a sort of relief map on K, and for a real number c consider the part of K on which $f(P)<c$ as inundated. This flooded portion has Betti numbers $p_0(c)$, $p_1(c)$, $p_2(c)$; and for $k=0, 1, 2$ $M_k(c)$ is defined as the number of critical points of type k under water, that is, having $f(P_j)<c$. We now proceed to establish

$$\begin{aligned} M_0(c) &\geqq p_0(c), \\ M_1(c) - M_0(c) &\geqq p_1(c) - p_0(c), \qquad (5.4)\\ M_2(c) - M_1(c) + M_0(c) &= p_2(c) - p_1(c) + p_0(c), \end{aligned}$$

for all c.

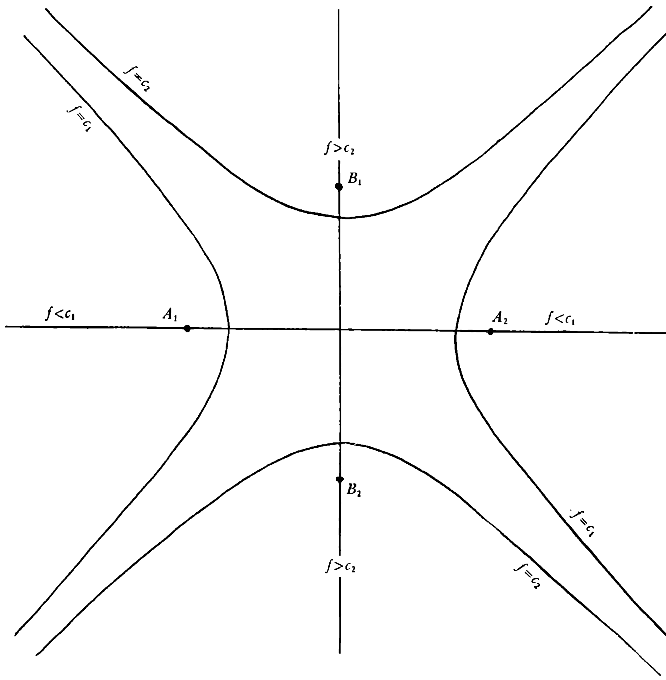
If c is just a little greater than the absolute minimum of f, then $M_0(c)=1$, $M_1(c)=M_2(c)=0$. The inundated portion consists of one small nearly elliptical lake, with Betti numbers 1, 0, 0. Hence (5.4) holds. It is intuitively plausible (and also true) that as c increases from one value c_1 to another c_2 such that no critical value lies in the interval $c_1 \leqq c \leqq c_2$ the Betti numbers $p_j(c)$ do not change, and by definition the numbers $M_k(c)$ remain constant. Therefore it remains only to investigate the behavior of these functions as c increases from a value c_1 slightly less than a critical value $f(P_r)$ to a value c_2 slightly greater than that critical value.

If the critical point P_r happens to be of type 0 the result is easily seen. Except near P_r, there is essentially no change; around P_r a new small lake appears. This introduces no new nonbounding one-cycle or two-cycles, but the Betti number $p_0(c)$ increases by unity. So does $M_0(c)$; and we verify at once that the relationships (5.4), if valid for $c=c_1$, are still valid for $c=c_2$. If P_r is a critical point of type 2, then for $c=c_1$ the point P_r is surrounded by a small island which is flooded when $c=c_2$. Consider a small circle Γ drawn in the flooded region about the island. There are two possibilities. Either Γ bounds a flooded region, which is only possible if the island is the last unflooded land; or it does not. In the first case the whole of K is flooded at $c=c_2$, so that $p_2(c_2)$ is the two-dimensional Betti number of K, which is 1, while $p_2(c_1)$ is zero. Then the right-hand member of the equation in (5.4) increases by 1; so does its left-hand member, since one new critical point of type 2 is flooded. So the equality remains valid; the inequalities are unaffected. In the second case the circle Γ was at $c=c_1$ a nonbounding cycle while at $c=c_2$ it is bounding, since its whole interior is flooded. Then $p_1(c)$ diminishes by unity, since one nonbounding one-cycle is lost. Here, too, we verify that relations (5.4) retain their validity.

Finally, let P_r be a critical point of type 1. By suitable choice of coordinates near P_r, the function $f(P)$ is approximately represented in the neighborhood P_r by the equation

$$f(P) = f(P_r) - (y^1)^2 + (y^2)^2.$$

This yields the diagram



As c goes from c_1 to c_2, the number $M_1(c)$ increases by 1, while $M_0(c)$ and $M_2(c)$ are unchanged. Consider the points A_1, A_2, of the region $f<c_1$: Either they can be joined by an arc lying in $f<c_1$, or they cannot. In the first case this arc, with the segment A_1A_2, constitutes a closed curve lying in $f<c_2$; it is nonbounding in $f<c_2$, for a portion of K bounded by it would have to contain either B_1 or B_2, and neither is in $f<c_2$. Hence the number $p_1(c)$ of nonbounding one-cycles has increased by 1. We verify that relations (5.4) retain validity. In the second case, the pair of points A_1, A_2 constitute in $f<c_1$ a nonbounding zero-cycle. In the set $f<c_2$ they bound the segment A_1A_2. So one nonbounding zero-cycle is lost, and $p_0(c)$ has decreased by 1. Otherwise stated, for $c=c_1$ the pieces containing A_1 and A_2 are separate, while for $c=c_2$ they are united, and the number $p_0(c)$ of

pieces has diminished by 1. In this case, too, we see that (5.4) retains validity.

Since it is now true that (5.4) is valid for all c greater than the absolute minimum of f, we assign c a value greater than the absolute maximum of f. Then relations (5.4) reduce to Morse's relationships (5.3).

The Betti numbers of a sphere are $p_0=1$, $p_1=0$, $p_2=1$. So on a globe which shows elevations in relief, the number M_2 of peaks plus the number M_0 of valley-bottoms equals 2 plus the number M_1 of passes. (This works better on a globe than on the earth itself, since such phenomena as natural bridges alter the Betti numbers of the earth's surface and complicate the notions of pass, etc.)

In stating the Morse relationships, we assumed for simplicity that the set R on which f is defined has no boundary. If it has a boundary, and at each boundary point the derivative of f in the direction of the outwardly-drawn normal is positive, the relations still hold [63]. Without the assumption about the directional derivative, the relations still hold, if the numbers M_k are properly reinterpreted [67].

If we try to apply the theory of critical points to the calculus of variations we encounter an obvious difficulty. The class of rectifiable curves C lying in a manifold constitutes a metric space, and on it an integral of the type considered in the calculus of variations defines a function $J(C)$. But the space is not finite-dimensional, and the function $J(C)$ is not in general continuous; even if the integral is regular and positive definite we can only say that $J(C)$ is lower semicontinuous, as we saw in §4. Thus in each application of the critical point theory it is necessary either to simplify the formulation of the variational problem, to generalize the theory, or to do some of both.

There are some questions in the calculus of variations to which the critical point theory can be readily applied by means of a simple device [68]. Suppose that we are studying the values of an integral taken along a class of curves lying in a manifold R and having the initial and terminal points in submanifolds Z_1, Z_2, respectively. If E is an extremal $y=\gamma(t)$, $(t_1 \leqq t \leqq t_2)$, belonging to this class of curves and meeting Z_1 and Z_2 transversally, cut it by $(n-1)$-dimensional manifolds $T_1, \cdots, T_p$. In Z_1, Z_2 and each T_j we choose coordinate systems. Then if $(A, P_1, \cdots, P_p, B)$ is a sequence of points with A in Z_1, P_j in T_j and B in Z_2, we can represent the whole sequence by a single μ-tuple (z), by writing first the coördinates of A in Z_1, then those of P_1 in T_1, etc. Under everyday hypotheses, if the manifolds T_j divide E into small enough arcs and the points $A, P_1, \cdots, P_p, B$ are near enough to E, they can be joined in that order by uniquely determined short extremal arcs. Thus (z) determines an extremaloid, or broken extremal, with vertices as indicated. We may suppose that (0) determines E itself. The integral along the extremaloid determined by (z) we denote by $J(z)$.

A simple calculation shows that a point (z) is a critical point of $J(z)$ if and only if the extremaloid determined by (z) is free of corners (and is therefore an extremal) and cuts Z_1 and Z_2 transversally, this last condition being considered vacuously satisfied at an end locus Z_i which consists of a single point. In particular, (0) is a critical point. If Z_1 and Z_2 both reduce to single points, the nature of the critical point is expressed by the *index theorem* [71]: *The critical point* $(z)=(0)$ *is degenerate if and only if the beginning and end points of E are conjugate. If they are not conjugate, the type of the critical point is equal to the number of points on E conjugate to the beginning, each counted according to its multiplicity.* A like result holds if only one end manifold reduces to a point; the type of the critical point is the number of points of E which are focal points of the other end manifolds.

From this we deduce several results of interest in the calculus of variations. For example, there is a well known theorem, due to Bliss, concerning the order of focal points along a minimizing curve in the plane problem with two variable end points. By use of broken extremals, Currier [72] extended this theorem to the problem of minimizing an integral on a family of curves in $(n+1)$-space joining two n-dimensional manifolds. He obtained both necessary conditions and sufficient conditions for a minimum. Morse [73] extended this by permitting the end manifolds to be of arbitrary dimensionality. Again, by use of similar concepts, Morse [69] established the following result. Let the terminal manifold Z_2 be a single point not focal to any point of Z_1. Suppose that each point P of Z_1 can be joined to Z_2 by a unique continuously varying extremal. The numbers M_k of such extremals cutting Z_1 transversally and containing k focal points are connected with the Betti numbers of Z_1 by the relations (5.3).

Consider next the aggregate of all extremals joining two fixed points A and B on a manifold R. We can proceed nearly as before. For any large number p, we can choose points $A, P_1, P_2, \cdots, P_p, B$, subject only to the restriction that the distances between successive points be small enough, and then join these by extremal arcs. Again, we can determine the sequence of points by a single np-tuple (z) consisting of the coordinates of P_1, followed by those of P_2, etc. Denoting by $J(z)$ the value of the integral along the extremaloid determined by (z), we find that (z) is a critical point of $J(z)$ if and only if the extremaloid has no corners. But this precludes the possibility of isolated critical points. For if we cut any extremal in any manner into short arcs, the points of division yield a critical point (z), so that each critical point belongs to an n-dimensional spread of such points.

This is harmless if we are using Birkhoff's minimax principle, since that principle is valid without assumptions of nondegeneracy of critical points. In this manner, Birkhoff was able to prove, for example, that on every surface of the topological type of the sphere there exists at least one closed geodesic [62]. However, if we wish to apply the entire set of Morse relations to such problems, the theory of critical points must be extended.

Such an extension was made by Morse [70], who thereby arrived at estimates of the numbers of extremals of different types joining two fixed points.

Thus we see that even in a rather simple variational problem, there is a need for generalization of the critical point theory. This need increases with the complexity of the problem discussed. Particularly, in discussing the problems (first investigated by Poincaré) concerned with the existence and behavior of closed geodesics on a surface we encounter the space of closed curves; and this space presents such topological complications that it is most undesirable to have also to keep guard against the analytical complication arising from the degeneration of critical points.

But if the general theory is to be extended to functions with degenerate critical points, the definition of the type of a critical point must be changed, and some reasonable method of assigning multiplicities must be devised. Our crude sketch of proof of the Morse relations contains a hint as to how this might be done. In the course of the proof it was shown that as c increased through a critical value $f(P_r)$, where P_r is a critical point of type k, ($k=0$, 1 or 2), then either $p_k(c)$ increased by 1 or $p_{k-1}(c)$ decreased by 1. So we can say that the "multiplicity" of P_r as a critical point of type j (which for this simple case should obviously be 0 if $j \neq k$ and 1 if $k=j$) is the sum of the number of new nonbounding j-cycles and the number of $(j-1)$-cycles which become bounding as c increases through $f(P_r)$. Such a topological definition clearly can be applied to degenerate critical points, and even to functions without derivatives. It conforms with Birkhoff's definition [62] of the multiplicity of a minimax point. As a simple example, consider a flat-topped mountain (not the highest in the world). At every point of the top all first partial derivatives of the height $f(x)$ are zero. But as c increases through the height of the mountain top, the number of nonbounding one-cycles in a maximal independent set will diminish by just one, so the whole mountain top is to be regarded as the equivalent of a single critical point of type 2 (maximum type). The "type number sums" for the critical set formed by the mountain top are: 0 for type 0, 0 for type 1, 1 for type 2.

In 1929 W. M. Whyburn [74] established some properties of functions with nonisolated critical points, which in essence were a generalization of the minimax principle. A year later A. B. Brown [75, 76]* extended the critical point theory to analytic functions with isolated critical points. Morse [71] further extended the theory so as to permit its use in treating the Poincaré problem on closed geodesics. He obtained theorems of which the following is a sample. Let R be a Riemannian manifold which is the nonsingular, analytic homeomorph of the unit m-sphere. Then there exists a set G of closed geodesics on R such that the kth type number sum of

* The theory is extended in [76] to critical sets which are complexes. There is a lacuna in one of the proofs in [75] (cf. footnote [76, p. 512]).

the geodesics of G is at least the number of principal ellipses of index k on any ellipsoid $E_m(a)$ for which the principal semi-axes are distinct and sufficiently near unity.*

But the movement toward generalizing the critical point theory did not stop with the papers cited. As we have already mentioned, it is possible to define critical points of various types by means of purely topological concepts without mention of derivatives. Recently the theory has been reinvestigated from the topological aspect, and as a result of the researches of Morse, Birkhoff, and Hestenes it has been extended so as to apply to functions defined and lower semicontinuous on metric spaces [71, 78–80].

While the theory of critical sets is thus extended to a wide class of spaces and functions, we cannot expect that the work in this direction is completed nor that it has assumed its final form. And if it were completed, the theory of the calculus of variations in the large would still be far from being a closed chapter. For we may reasonably hope to see it extended to larger classes of problems, including double integral problems, and we may expect new and significant applications of the general theory, for example, to problems in dynamics.

6. The idea, presented in the preceding paragraphs, of embedding the problems of the calculus of variations in a more general class of problems, is not a new one. Volterra considered the integrals of the calculus of variations as a particular class of functions of lines, and more recently there have probably been many who have regarded these integrals as functions on metric spaces [81]. We have already seen that the existence theorems gain in clarity when regarded from this point of view, and that Menger has studied the problems of the calculus of variations entirely within the framework of semimetric spaces. Moreover we have seen that the critical point theory can be profitably extended to functions defined on metric spaces. However, the same concept can also be used in deriving necessary conditions for a minimum. LeStourgeon [82] studied the problem of minimizing functionals of a certain type, and by use of the differential, established necessary conditions which specialized to the Euler equation and Jacobi condition for free problems. More recently, Lusternik [83] considdered the problem of minimizing a functional f on that part of a linear space A which is carried by a differentiable transformation into the origin of another linear space B. He obtained a necessary condition which generalizes the Lagrange multiplier rule for Lagrange problems. Graves [84] applied theorems developed in general analysis to establish the Euler equations for Lagrange problems in a degree of generality which, at least for parametric problems, seems to leave nothing to be desired. Goldstine [85, 86] set up a problem of which the Bolza problem is a special case, and

* By different methods, depending on the introduction of a new topological invariant could the "category," Lusternik and Schnirelmann earlier established the existence of at least three non-self-intersecting closed geodesics on surfaces of genus zero [77].

for it deduced the analogues of the Euler equations and the Legendre and Weierstrass conditions. The problem of developing the methods of abstract spaces far enough to include with profit a large central portion of the calculus of variations is very far from completion. But in view of the widespread trend in analysis toward the general and abstract, we may well expect to see the program pushed forward in the next few years.

UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VA.

BIBLIOGRAPHY

1. RADÓ, T. *On the Problem of Plateau.* Ergebnisse der Mathematik und ihre Grenzgebiete. 1933.

2. DOUGLAS, J. *Some new results in the problem of Plateau.* Journal of Mathematics and Physics, vol. 15 (1936), pp. 55–64.

3. DOUGLAS, J. *Minimal surfaces of general topological structure with any finite number of assigned boundaries.* Journal of Mathematics and Physics, vol. 15 (1936), pp. 105–123.

4. COURANT, R. *Plateau's problem and Dirichlet's principle.* Annals of Mathematics, (2), vol. 38 (1937), pp. 679–724.

5. COURANT, R. *The existence of a minimal surface of least area bounded by prescribed Jordan arcs and prescribed surfaces.* Proceedings of the National Academy of Sciences, vol. 24 (1938), pp. 97–101.

6. REID, W. T. *Boundary value problems of the calculus of variations.* Bulletin of the American Mathematical Society, vol. 43 (1937), pp. 633–666.

7. BLISS, G. A. *The problem of Lagrange in the calculus of variations.* American Journal of Mathematics, vol. 52 (1930), pp. 674–744.

8. MORSE, M., and MYERS, S. B. *The problems of Lagrange and Mayer with variable end points.* Proceedings of the American Academy of Arts and Sciences, vol. 66 (1931), pp. 235–253.

9. BLISS, G. A. *The Problem of Bolza in the Calculus of Variations.* Mimeographed notes of lectures delivered at the University of Chicago, 1935.

10. BOLZA, O. *Vorlesungen über Variationsrechnung.* Leipzig, Teubner, 1909.

11. VON ESCHERICH, G. *Die zweite Variation der einfachen Integrale.* Sitzungsberichte der Akademie der Wissenschaften zu Wien, (2A),vol. 107 (1898), pp. 1191–1250, 1267–1326, 1381–1428; vol. 108 (1899), pp. 1269–1340.

12. KNESER, A. *Lehrbuch der Variationsrechnung.* Braunschweig, Vieweg, 1900.

13. MASON, M., and BLISS, G. A. *The properties of curves in space which minimize a definite integral.* Transactions of the American Mathematical Society, vol. 9 (1908), pp. 440–466.

14. BLISS, G. A. *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.

15. BLISS, G. A. *The transformation of Clebsch in the calculus of variations.* Proceedings of the International Congress at Toronto, 1924, vol. 1, pp. 589–603.

16. BLISS, G. A. *Some recent developments in the calculus of variations.* Bulletin of the American Mathematical Society, vol. 26 (1919), pp. 343–361.

17. HESTENES, M. R. *A note on the Jacobi condition for parametric problems in the calculus of variations.* Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 297–302.

18. BIRKHOFF, G. D., and HESTENES, M. R. *Natural isoperimetric conditions in the calculus of variations.* Duke Mathematical Journal, vol. 1 (1935), pp. 198–286. (Cf. in particular, p. 264.)

19. MORSE, M. *The index theorem in the calculus of variations.* Duke Mathematical Journal, vol. 4 (1938), pp. 231–246.

20. BLISS, G. A. *Normality and abnormality in the calculus of variations.* Transactions of the American Mathematical Society, vol. 43 (1938), pp. 365–376.

21. CARATHÉODORY, C. *Die Theorie der zweiten Variation beim Problem von Lagrange.* Sitzungsberichte der Bayerischen Akademie der Wissenschaften, 1923, pp. 99–114.

22. CARATHÉODORY, C. *Über die Einteilung der Variationsprobleme von Lagrange nach Klassen.* Commentarii Mathematici Helvetici, vol. 5 (1933), pp. 1–19.

23. GRAVES, L. M. *On the Weierstrass condition for the problem of Bolza in the calculus of variations.* Annals of Mathematics, (2), vol. 33 (1932), pp. 747–752.

24. HESTENES, M. R. *Sufficient conditions for the problem of Bolza in the calculus of variations.* Transactions of the American Mathematical Society, vol. 36 (1934), pp. 793–818.

25. MORSE, M. *Sufficient conditions in the problem of Lagrange without assumptions of normalcy.* Transactions of the American Mathematical Society, vol. 37 (1935), pp. 147–160.

26. REID, W. T. *Sufficient conditions by expansion methods for the problem of Bolza in the calculus of variations.* Annals of Mathematics, (2), vol. 38 (1937), pp. 662–678.

27. LEVI, E. E. *Sui criterii sufficienti per il massimo e per il minimo nel calcolo delle variazioni.* Annali di Matematica Pura ed Applicata, (4), vol. 21 (1913), pp. 173–218.

28. HESTENES, M. R. *A direct sufficiency proof for the problem of Bolza in the calculus of variations.* Transactions of the American Mathematical Society, vol. 42 (1937), pp. 141–154.

29. REID, W. T. *A direct expansion proof of sufficient conditions for the non-parametric problem of Bolza.* Transactions of the American Mathematical Society, vol. 42 (1937), pp. 183–190.

30. HILBERT, D. *Über das Dirichletsche Prinzip.* Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 8 (1900), pp. 184–188.

31. CARATHÉODORY, C. *Über die starken Maxima und Minima bei einfachen Integralen.* Mathematische Annalen, vol. 62 (1906), pp. 449–503.

32. BILL, E. S. *An existence theorem for a problem of the calculus of variations in space.* Transactions of the American Mathematical Society vol. 13 (1912), pp. 50–58.

33. TONELLI, L. *Fondamenti di Calcolo delle Variazioni.* Zanichelli, Bologna, vol. 1, 1921, and vol. 2, 1923.

34. HAHN, H. *Über ein Existenztheorem der Variationsrechnung.* Sitzungsberichte der Akademie der Wissenschaften zu Wien, (2A), vol. 134 (1925), pp. 437–447.

35. CARATHÉODORY, C. *Über die Existenz der absoluten Minima bei regulären Variationsproblemen auf der Kugel.* Annali della Scuola Normale Superiore di Pisa, (2), vol. 1 (1932), pp 79–88.

36. TONELLI, L. *Sull'esistenza del minimo in problemi di calcolo delle variazioni.* Annali della Scuola Normale Superiore di Pisa, vol. 1 (1932), pp. 89–99.

37. GRAVES, L. M. *On an existence theorem of the calculus of variations.* Monatshefte für Mathematik und Physik, vol. 39 (1932), pp. 101–104.

38. MCSHANE, E. J. *Remark concerning Mr. Graves' paper "On an existence theorem of the calculus of variations."* Monatshefte für Mathematik und Physik, vol. 39 (1932), pp. 105–106.

39. NAGUMO, M. *Über die gleichmässige Summierbarkeit und ihre Anwendung auf ein Variationsproblem.* Japanese Journal of Mathematics, vol. 6 (1929), pp. 173–182.

40. MCSHANE, E. J. *Existence theorems for ordinary problems of the calculus of variations.* Annali della Scuola Normale Superiore di Pisa, (2), vol. 3 (1934), pp. 183–211, 287–315.

41. TONELLI, L. *Su gli integrali del calcolo delle variazioni in forma ordinaria.* Annali della Scuola Normale Superiore di Pisa, (2), vol. 3 (1934), pp. 401–450.

42. MCSHANE, E. J. *Semi-continuity of integrals in the calculus of variations.* Duke mathematical Journal, vol. 2 (1936), pp. 597–616.

43. MCSHANE, E. J. *Some existence theorems for problems in the calculus of variations.* Duke Mathematical Journal, vol. 4 (1938), pp. 132–156.

44. MENGER, K. *Metrische Geometrie und Variationsrechnung.* Fundamenta Mathematicae, vol. 25 (1935), pp. 441–458.

45. MENGER, K. *Die metrische Methode in der Variationsrechnung.* Ergebnisse eines Mathematischen Kolloquiums, vol. 8. Vienna, 1937.

46. MENGER, K. *Metric methods in the calculus of variations.* Proceedings of the National Academy of Sciences, vol. 23 (1937), pp. 244–250.

47. FRINK, O. *Geodesic continua in abstract metric space.* American Journal of Mathematics, vol. 58 (1936), pp. 514–520.

48. GRAVES, L. M. *On the existence of the absolute minimum in space problems of the calculus of variations.* Annals of Mathematics, vol. 28 (1927), pp. 153–170.

49. GRAVES, L. M. *On the existence of the absolute minimum in problems of Lagrange.* Bulletin of the American Mathematical Society, vol. 39 (1933), pp. 101–104.

50. CINQUINI, S. *Sopra l'esistenza della soluzione nei problemi di calcolo delle variazioni di ordine n.* Annali della Scuola Normale Superiore di Pisa, (2), vol. 5 (1936), pp. 169–190.

51. CINQUINI, S. *Nuovi teoremi di esistenza dell'estremo in campi illimitati per i problemi di calcolo delle variazioni di ordine n.* Annali della Scuola Normale Superiore di Pisa, (2), vol. 6 (1937), pp. 191–209.

52. MANIÀ, B. *Sopra un problema di navigazione di Zermelo.* Mathematische Annalen, vol. 113 (1936), pp. 584–599.

53. MANIÀ, B. *Esistenza dell'estremo assoluto in un classico problema di Mayer.* Annali della Scuola Normale Superiore di Pisa, (2), vol. 2 (1933), pp. 343–354.

54. MANIÀ, B. *Sulla curva di massima velocità finale.* Annali della Scuola Normale Superiore di Pisa, (2), vol. 3 (1934), pp. 317–336.

55. MANIÀ, B. *Sui problemi di Lagrange e di Mayer.* Rendiconti del Circolo Matematico di Palermo, vol. 58 (1934), pp. 285–310.

56. MANIÀ, B. *Sopra una classe di problemi di Mayer considerati come limiti di ordinari problemi di minimo.* Rendiconti del Seminario Matematico della Università di Padova, 1934, pp. 99–121.

57. GRAVES, L. M. *The existence of an extremum in problems of Mayer.* Transactions of the American Mathematical Society, vol. 39 (1936), pp. 456–471.

58. MCSHANE, E. J. *Semi-continuity in the calculus of variations, and absolute minima for isoperimetric problems.* Contributions to the Calculus of Variations, 1930. University of Chicago, 1931.

59. MCSHANE, E. J. *Some existence theorems in the calculus of variations.* A sequence of five notes to appear in the Transactions of the American Mathematical Society.

60. LEWY, H. *Über die Methode der Differenzengleichungen zur Lösung von Variations- und Randwertproblemen.* Mathematische Annalen, vol. 98 (1928), pp. 107–124.

61. DRESDEN, A. *The second derivatives of the extremal integral.* Transactions of the American Mathematical Society, vol. 9 (1908), pp. 467–486. (Cf. in particular, p. 486.)

62. BIRKHOFF, G. D. *Dynamical systems with two degrees of freedom.* Transactions of the American Mathematical Society, vol. 18 (1917), pp. 199–300.

63. MORSE, M. *Relations between the critical points of a real function of n independent variables.* Transactions of the American Mathematical Society, vol. 27 (1925), pp. 345–396.

64. POINCARÉ, H. *Sur les courbes définies par les équations différentielles.* Journal de Mathématiques Pures et Appliqués, (4), vol. 1 (1885), pp. 167–244.

65. LEFSCHETZ, S. *Continuous transformations of manifolds.* Proceedings of the National Academy of Sciences, vol. 11 (1925), pp. 290–292.

66. HOPF, H. *Vektorfelder in n-dimensionalen Mannigfaltigkeiten.* Mathematische Annalen, vol. 96 (1927), pp. 225–250.

67. MORSE, M., and VAN SCHAACK, G. B. *The critical point theory under general boundary conditions.* Annals of Mathematics, (2), vol. 35 (1934), pp. 545–571; Duke Mathematical Journal, vol. 2 (1936), pp. 220–242.

68. MORSE, M. *The foundations of a theory in the calculus of variations in the large.* Transactions of the American Mathematical Society, vol. 30 (1928), p. 213–274.

69. MORSE, M. *The foundations of the calculus of variations in the large in m-space.* I. Transactions of the American Mathematical Society, vol. 31 (1929), pp. 379–404.

70. MORSE, M. *The foundations of a theory of the calculus of variations in the large in m-space.* II. Transactions of the American Mathematical Society, vol. 32 (1930), pp. 599–631.

71. MORSE, M. *Functional topology and abstract variational theory*, Annals of Mathematics (2), vol. 38 (1937), pp. 386–449.

72. CURRIER, A. E. *The variable end point problem of the calculus of variations including a generalization of the classical Jacobi conditions*. Transactions of the American Mathematical Society, vol. 34 (1932), pp. 689–704.

73. MORSE, M. *The Calculus of Variations in the Large*. American Mathematical Society Colloquium Publications, vol. 18, pp. 64–70 New York, 1934.

74. WHYBURN, W. M. *Non-isolated critical points of functions*. Bulletin of the American Mathematical Society, vol. 35 (1929), pp. 701–708.

75. BROWN, A. B. *Relations between the critical points of a real analytical function of N independent variables*. American Journal of Mathematics, vol. 52 (1930), pp. 251–270.

76. BROWN, A. B. *Critical sets of an arbitrary real analytical function of n variables*. Annals of Mathematics, (2), vol. 32 (1931), pp. 512–520.

77. LUSTERNIK, L., and SCHNIRELMANN, L. *Topological Methods in the Calculus of Variations*. Gosudarstvennoe Izdatelstvo, Moscow, 1930. (In Russian.) [*Méthodes topologiques dans les problèmes variationnels*. Actualités Scientifiques et Industrielles, no. 188. 1934. (French translation of first part of book.)]

78. MORSE, M. *Analysis in the Large*. Lithoprinted notes of lectures delivered at the Institute for Advanced Study, 1935–1936.

79. BIRKHOFF, G. D., and HESTENES, M. R. *Generalized minimax principle in the calculus of variations*. Duke Mathematical Journal, vol. 1 (1935), pp. 413–432.

80. HESTENES, M. R. *Minimax principle for functions*. Proceedings of the National Academy of Sciences, vol. 22 (1936), pp. 115–119.

81. SANGER, R. G. *Functions of lines and the calculus of variations*. Contributions to the Calculus of Variations, 1931–1932. University of Chicago, 1933.

82. LESTOURGEON, F. E. *Minima of functions of lines*. Transactions of the American Mathematical Society, vol. 21 (1920), pp. 357–383.

83. LUSTERNIK, L. A. *Sur les extrêmes relatifs des fonctionelles*. Recueil Mathématique de Moscow, vol. 41 (1934), pp. 390–400. (In Russian.)

84. GRAVES, L. M. *On the problem of Lagrange*. American Journal of Mathematics, vol. 53 (1931), pp. 547–554.

85. GOLDSTINE, H. H. *Conditions for a minimum of a functional*. Contributions to the Calculus of Variations, 1933–1937. University of Chicago, 1937.

86. GOLDSTINE, H. H. *The minima of functionals with associated side conditions*. Duke Mathematical Journal, vol. 3 (1937), pp. 418–425.