educational policies, insist that the next generation has to operate with the same prejudices.

Braun's book has aspects that can please both styles of applied mathematicians. The book could perhaps play a role in giving both pure and applied mathematics students and other science students an appreciation of both the classical and the modern styles of applied mathematics, and so far as this is so, the book may make a healthy contribution to the future direction of applied mathematic education.

JERRY BONA

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 84, Number 1, January 1978

Optimization, a theory of necessary conditions, by Lucien W. Neustadt, Princeton Univ. Press, Princeton, New Jersey, 1977, xii + 424 pp., $22.50.

Optimale Steuerung diskreter Systeme, by W. G. Boltjanski, Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig, 1976, 326 pp.

The qualitative theory of optimal processes, by R. Gabasov and F. Kirillova, Marcel Dekker, Inc., New York, New York, 1976, xlvi + 640 pp., $55.00.

1. Horreur. "Je me détourne avec éffroi et horreur de cette plaie lamentable des fonctions qui n'ont pas de dérivée"; so said Hermite in a letter to Stieltjes. The reader who shares this aversion to nondifferentiable functions will undoubtedly be affronted by the three books in question. But mathematicians have become much more tolerant about the functions they will talk to. This has been most evident in optimization, where the need to consider differential properties of other than smooth functions arises frequently and fundamentally. In fact, these ill-bred functions are now often brought into the discussion from the start and used systematically, rather than being shunned whenever possible. The extent to which this is true is a striking feature of these three books, all of which were written by well-known researchers in the field of optimal control.

The wedge in this breakthrough was the gradual recognition of the central role in optimization of convexity. This first took place in mathematical programming, and now the methods of convex analysis are being systematically applied in other areas as well; their use in optimal control is currently an active subject for research (see [4]). And convexity *implies* nondifferentiability–not just because differentiability is unnecessary, but because clinging to it is simply not feasible. For example, one of the great successes of convex analysis is duality (see [9]) the pairing with an original minimization problem of a certain closely related maximization problem. Besides being rich in interpretation (e.g. stress vs. reaction, utility vs. price) this concept is at the heart of the most successful computational algorithms in mathematical programming. Yet even if the original problem of interest is smooth, its dual may very well fail to be.

We shall encounter presently some further examples of fundamental nondifferentiability. But before we arrive at what Hermite would think of as this sorry pass, let us look back.

2. Remembrance of things past. The basic problem in the calculus of variations will soon be celebrating its three hundredth birthday, but it remains remarkably spry withal. It consists of minimizing a functional of the form

$$\int_a^b L(t, x(t), \dot{x}(t))dt \tag{1}$$

where x belongs to some class of functions satisfying given conditions at a and b. The three great themes of optimization come quickly to mind:

(a) existence–does a minimizing x exist?

(b) necessary conditions–how may we characterize a minimal x, or at least limit the number of suspects?

(c) sufficiency–can we proceed to prove our conviction that a certain suspect is indeed optimal?

The body of classical theory which responds to these points is one of the great success stories of mathematics. These questions, in being answered, spawned so many concepts that the calculus of variations is, in the words of L. C. Young, "a record of the history of mathematical concepts . . . that no other branch of mathematics possesses to an equal extent." (By the way, Young's book **[10]** is so quotable and so amusing that one must feel sorry for anyone who hasn't read it.) Yet its achievements are far from being a collection of dusty relics; even today the calculus of variations plays a central role in the mathematics underlying modern theories of the structure of matter. Why is it, then, that this subject, so important, so "relevant", is so rarely present in the undergraduate curriculum, for which it is ideally suited?

It was Hilbert who led the attack on the first question ("every problem of the calculus of variations has a solution . . . "), followed by Tonelli. Weierstrass and Jacobi showed that the third question is linked to imbedded families of extremals and the theory of quadratic functionals. But the question of necessary conditions has dominated from the very beginning; a selection follows.

For the basic problem, the first (and foremost) of many necessary conditions says that any solution x to the problem must satisfy the Euler-Lagrange equation

$$\frac{d}{dt} D_{\dot{x}} L(x, \dot{x}) = D_x L(x, \dot{x}).$$

With malice aforethought, let us agree to express this in the following way: there is a function $p(t)$ such that

$$(\dot{p}(t), p(t)) = DL(x, \dot{x}). \tag{2}$$

The second most important necessary condition bears the name of Weierstrass, and it says that for the above p, for each t in $[a, b]$,

$$L(x(t), \dot{x}(t) + v) - L(x(t), \dot{x}(t)) \geqslant v \cdot p(t) \quad \forall v. \tag{3}$$

Now suppose that $L(x, \cdot\,)$ is convex for each x, and that the function H defined as follows is differentiable:

$$H(x, p) = \sup_v \{p \cdot v - L(x, v)\}. \tag{4}$$

Then it so happens that (2) is equivalent to

(5) $$(-\dot{p}, \dot{x}) = DH(x, p).$$

This latter type of differential equation is of fundamental importance in physics; it is called a Hamiltonian system, where H is the Hamiltonian. (Classically, H is defined by a local Legendre transform rather than by (4).) These functions play a central role in the theories of invariance and conservation laws, the Hamilton-Jacobi equation, and in variational principles. L. C. Young has said "conceptually, the importance of Hamiltonians compares with that of complex numbers"; we shall return to them later. Let us first turn the century.

3. Optimal control. The basic problem becomes considerably more complex if pointwise constraints, say

(6) $$f(x(t), \dot{x}(t)) = 0 \qquad \forall t \in [a, b],$$

are imposed (this would be called a problem of Lagrange). Much of the activity in the calculus of variations in this century has been concerned with problems such as these, and in particular with finding a rigorous proof of the "multiplier rule". The latter is a theorem saying that any x minimizing (1) subject to (6) is a stationary point for the integrand

(7) $$L(x, \dot{x}) + \lambda(t) \cdot f(x, \dot{x})$$

if $\lambda(t)$ is suitably chosen; the analogy with Lagrange multipliers is clear. Special cases were analysed and partial results obtained, but a proof that such a $\lambda(t)$ exists so that *all* the necessary conditions for (7) are satisfied (and not merely the Euler-Lagrange equation) remained elusive until McShane's 1939 paper [6]. This remarkable achievement was destined to be overshadowed, however.

In the 1950's, L. S. Pontryagin analysed the following basic optimal control problem. Consider any measurable function $u(t)$ taking values in U ("control"), and the resulting solution $x(t)$ ("trajectory") to the differential equation

(8) $$\dot{x}(t) = f(x(t), u(t)).$$

Find the control u and corresponding trajectory x which minimize the "cost functional"

$$\int_a^b g(x(t), u(t))dt$$

subject to certain constraints on $x(a)$, $x(b)$.

In order to state Pontryagin's necessary conditions for this problem, let us define a function H as follows:

(9) $$H(x, p, u) = p \cdot f(x, u) - g(x, u).$$

The celebrated "maximum principle" [7] then says that if (x, u) are optimal, there exists p such that

(10) $$(-\dot{p}(t), \dot{x}(t)) = D_{x,p}H(x(t), p(t), u(t)),$$

(11) $$H(x(t), p(t), u(t)) = \max_{u \in U}\{p(t) \cdot f(x(t), u) - g(x(t), u)\}.$$

In proving this, Pontryagin et al. used McShane's seminal approach via cones of displacements, an idea reused and elaborated upon numerous times since (see [**5**]).

Is optimal control simply the calculus of variations under another guise? Although few would subscribe to such a bald statement, the point has been controversial. The reason for this is that technically, the basic optimal control problem is "often" equivalent to a calculus of variations problem with constraints, and the multiplier rule for the latter problem yields the conclusions of the maximum principle. (To see this in the simplest case, take U to be the whole space, $f(x, u) = u$, and compare (10), (11) to (2), (3): they are the same; a complete discussion is given in [**1**].)

The difference then is largely one of packaging. Yet it is fundamental ("the medium is the message"?). The very statement of the optimal control problem brings to the fore new and interesting considerations (example: given two points, is there a control u such that the solution to (8) joins them?). Its development has been freer of the "artificial assumptions of smoothness" [**10**, p. 214] of the calculus of variations (apologies to Hermite). But most important is the format's emphasis on available choice (control). This represents a major philosophical shift from the calculus of variations, where the underlying "variational principles" are god-given and nature does the optimizing, and it accounts for optimal control's widespread and immediate use in engineering (and later, in economic modeling). It also leads naturally to the consideration of stochastic elements. (Some credit its appeal to its catchy name as well, which is something the calculus of variations certainly lacks!)

Let us forget the maximum principle for the moment and try applying Hamiltonian theory to the basic optimal control problem. A good way to calculate the Hamiltonian is to put the problem into the same form as the basic problem in the calculus of variations and use formula (4) for H. But in order to take account of the constraints that now exist, we must permit L to attain the value $+\infty$ (this bookkeeping device is common in mathematical programming). We define

$$L(x, \dot{x}) = \inf\{g(x, u) : u \in U, \dot{x} = f(x, u)\},$$

where by the usual convention the infimum over the empty set is $+\infty$. Then the optimal control problem becomes that of minimizing the integral of this L. We calculate H via (4) and find

$$H(x, p) = \sup_{u \in U} \{p \cdot f(x, u) - g(x, u)\}. \tag{12}$$

This, rather than the H in the maximum principle, is the (long-lost) "true Hamiltonian" for the problem [**10**, p. 230], although few realize it. Why then is it not used? Because, in contrast to the classical setting, there is now no hope that the function defined by (12) will be differentiable! But having agreed to cease discriminating on such grounds, we might ask if it is possible to obtain an analogue of the Hamiltonian equation (5) for this problem. In fact, this is possible if we use the "generalized gradient" ∂f of a locally Lipschitz function f (see [**2**]). A necessary condition for optimality [**3**] is the existence of a function p such that $(-\dot{p}, \dot{x}) \in \partial H(x, p)$. This classically-rooted approach via "Hamiltonian inclusions" turns out to have several advantages over the maximum principle (which can be obtained from it). One of these is the possibility of treating other situations, such as when U depends on x, or when (8) is replaced by the more general differential inclusion (see [3]) $\dot{x}(t) \in F(x(t))$.

Although necessary conditions have attracted the most attention, optimal control has of course been concerned with the other two main themes as well. In existence theory, (and elsewhere), the relaxed controls of J. Warga (closely related to the generalized curves of Young) have proved central, and the entire question seems well understood. (Interestingly, criteria guaranteeing existence can also be given in Hamiltonian terms [**8**].) However, sufficiency theory in optimal control has not yet attained the satisfactory state found in the calculus of variations.

All these topics are traced in detail in the book of Gabasov and Kirillova, which is more or less a survey of optimal control. Also awarded chapters are the topics of *controllability* and *observability*. The former, which is currently an active area of research, deals with the feasibility of joining given values of the state x via solutions of the controlled system (8). The latter is concerned with deducing properties of the state x of the system when only functions of x are observable. The authors are best known for their work on computational methods in optimal control, and in consequence that subject also occupies an important place. The final chapter in particular is concerned with the theory of discrete-time systems; that is, the problem in which (8) is replaced by

$$x(t) = f(x(t-1), u(t)), \qquad t = 0, 1, \ldots, n.$$

This happens to be the subject of Boltjanski's monograph, which is concerned primarily with establishing a discrete analogue of the Pontryagin maximum principle.

Although the content of Gabasov and Kirillova's book is familiar in the large, many of the particulars will be of interest to specialists; it also serves as a guide to the huge Russian literature. Despite its wide scope, the book does not recommend itself as an introduction to the subject; its penchant for cumbersome technical detail and its stilted language could well smother a nascent enthusiasm.

In contrast, Boltjanski treats a limited subject (but very thoroughly). He obtains the necessary conditions for the discrete-time optimal control problem as a consequence of convex analysis and mathematical programming (this is similar in spirit to the approach of Neustadt described later), both of which are developed from first principles. Readers will be grateful for the opening chapter, which motivates, illustrates and summarizes the rest of the book. This reminds one of a similar chapter in the seminal (and still interesting) book [7] by Pontryagin et al., which is perhaps not surprising, since the author is among the alii. The book is well written and intelligently organized, and can be recommended for the reader who would like to learn simultaneously something about mathematical programming, convex analysis, and optimal control theory.

4. General theories of necessary conditions. Suppose that X is a Banach space, and that f: $X \to R$ is a given function. There are many (very many!) definitions of objects that behave like derivatives of f; let us call one such the

"α-derivative" and denote it $D_\alpha f$ (any resemblance to a term living or dead is purely coincidental). One defines an "α-stationary" point x as a point such that $D_\alpha f(x)$ behaves somehow like zero. Undoubtedly, a theorem exists somewhere which says that if f and g are α-differentiable, and if x minimizes f subject to $g = 0$, then there exist λ_0 and λ in an appropriate space such that $\lambda_0 f + \lambda g$ has an α-stationary point at x. Any such theorem is called a (Lagrange) multiplier rule, and the proliferation of these (for α in a large index set) has given necessary conditions a bad name in some quarters. The point is, of course, that too often the α-multiplier rule is a sterile exercise; it "extends" well-known results by ε, and adds no new possibilities or insights. And the α-derivative is never heard from again.

When then is a multiplier rule worthwhile? There seem to be two raisons d'etre: *if*

(a) the theory treats a special class of functions, but with widely significant and practical results, or

(b) the theory is so general that it is useful in a great variety of truly different problems.

Examples of (a) are the classic continuously differentiable case, and the theory of convex mathematical programming; an example of (b) is the subject of the book by the late L. W. Neustadt (the finishing touches were applied by H. T. Banks).

The book begins by developing a general multiplier rule; we shall not describe in detail the approach, which was developed largely in conjunction with H. Halkin, and which generalizes even the definition of optimality. The rest of the book consists in putting this "abstract maximum principle" through its paces. One of the features of new fields like optimal control is that the important basic problems remain to some extent unidentified, so that new variations continue to appear. The great advantage of a unified general multiplier rule, is that in obtaining conditions for such problems we are saved from starting from scratch each time. Undoubtedly, an approach to a given problem handcrafted from first principles can often yield more than the wheeling in of general machinery, but the latter can be very useful, as has already been demonstrated in the case of Neustadt's theory. The applications presented here are to state-constrained optimal control problems (an early success of the method), systems governed by functional relations, minimax optimization, vector criteria, and discrete problems. The book, which should prove interesting to a variety of workers in the field, is further enhanced by an excellent bibliography and commentary on the literature of necessary conditions in optimal control to about 1973.

References

1. L. D. Berkovitz, *Variational methods in problems of control and programming*, J. Math. Anal. Appl. 3 (1961), 145–169.

2. F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.

3. ______, *Necessary conditions for a general control problem*, in Calculus of Variations and Control Theory (ed. D. L. Russell), Mathematics Research Center Publ. 36, Academic Press, New York, 1976.

4. I. Ekeland and R. Temam, *Convex analysis and variational problems*, North-Holland, Amsterdam, 1976.

5. M. R. Hestenes, *Calculus of variations and optimal control theory*, Wiley, New York, 1966.

6. E. J. McShane, *On multipliers for Lagrange problems*, Amer. J. Math. **61** (1939), 809–819.

7. L. S. Pontryagin et al., *The mathematical theory of optimal processes*, K. N. Trirogoff (Tran.), L. W. Neustadt (Ed.), Interscience Publishers, New York, 1962.

8. R. T. Rockafellar, *Existence theorems for general control problems of Bolza and Lagrange*, Advances in Math. **15** (1975), 312–333.

9. ______, *Conjugate duality and optimization*, S.I.A.M., Philadelphia, 1974.

10. L. C. Young, *Lectures on the calculus of variations and optimal control theory*, Saunders, Philadelphia, 1969.

FRANK H. CLARKE

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 84, Number 1, January 1978

Problems and theorems in analysis, by G. Pólya and G. Szegö, Die Grundlehren der math. Wissenschaften, Springer-Verlag, Berlin and New York; Vol. I, 1972, xix + 389 pp., Vol. II, 1976, xi + 391 pp., $45.10.

Pólya and Szegö, *Aufgaben und Lehrsätze aus der Analysis* was published first in 1925 as volumes 19 and 20 of the "yellow-peril" series. See Tamarkin [**1**] for a review. The inexpensive reprint in 1945 (Dover Publications) by authority of the U. S. Alien Property Custodian made the work widely known in N. America. The four Springer (German) editions through the latest (1970, 1971) are unchanged from the original except for the correction of minor errors.

The present volumes are a revised and enlarged translation of the 4th edition, vol. I translated by Dorothee Aeppli and vol. II by Claude E. Billigheimer.

The work is one of the real classics of this century; it has had much influence on teaching, on research in several branches of hard analysis, particularly complex function theory, and it has been an essential indispensable source book for those seriously interested in mathematical problems. One can think of few books written more than a half century ago that would really be worth translating today. This one certainly was; of course some parts are a bit faded and dated, but much is fresh and exciting and will be consulted for years to come. The translators (whose work is first-rate), authors, and publisher deserve our praise for making Pólya-Szegö available in English to the ever widening set of mathematicians and students who no longer read German.

These volumes contain many extraordinary problems and sequences of problems, mostly from some time past, well worth attention today and tomorrow. Before embarking on my reviewer's responsibility of evaluation and criticism, I want to emphasize, regardless of anything I say below, my personal enormous respect for the mathematics of Pólya-Szegö. This work was written in the early twenties by two young mathematicians of outstanding talent, taste, breadth, perception, perseverence, and pedagogical skill. It broke new ground in the teaching of mathematics and how to do mathematical research.