VARIATIONAL ANALYSIS:

Critical Extremals and Sturmian Extensions

MARSTON MORSE

Institute for Advanced Study

A WILEY-INTERSCIENCE PUBLICATION

JOHN WILEY & SONS New York . London . Sydney . Toronto

To

LOUISE

Library of Congress Cataloging in Publication Data

Morse, Marston, 1892
Variational analysis

(Pure and applied mathematics)
Bibliography
1 Calculus of variations 2. Maxima and minima. I. Title

QA316 M58 515′64 72-8368
ISBN 0-471-61700-8

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

PREFACE

The first concern of Euler, Legendre, Jacobi, Weierstrass, and other mathematicians who pioneered in the study of the calculus of variations was to characterize differentiable mappings $x \to g(x)$ of a finite interval $[a,b]$ into R^m ($m = 1$, or 2) which afforded a minimum to an integral of the form

$$\int_a^b f(x, y_1, \ldots, y_m; y_1', \ldots, y_m')\, dx \tag{1}$$

relative to values of this integral along mappings $x \to y(x)$ neighboring g which satisfied simple boundary conditions. Under conditions on f, g and the mappings $x \to y(x)$ neighboring g, which will be made explicit, the mapping $x \to g(x)$ satisfies the Euler equations and has a graph that is called a *minimizing extremal.*

However, it has become increasingly evident that the restriction of the calculus of variations to a study of *minimizing extremals* is not only unnatural but fails to respond to present needs of mathematics in analysis, differential geometry, mathematical physics, and engineering. This is transparent in global and differential topology. It is hoped that the theorems and methods of this book will make this equally clear in analysis, in particular in that branch of analysis which is concerned with extensions of Sturm separations, comparison, and oscillation theorems.

Apart from the development of Sturm-like theorems in my Colloquium Lectures, most applications of variational methods to the Sturm theory have been restricted to the case $m = 1$, where m is the number of variables $y_1, \ldots, y_m$ in (1). This is the case of one linear second-order differential equation. In the case $m = 1$, A. Kneser, Leighton, Nehari, Reid, Howard, and others have effectively used variational methods. In the case $m = 1$, much attention has also been given to linear, homogeneous DE (differential equations) of the third or fourth order. A book by Swanson summarizes these results and contains an extensive bibliography. It is my belief that when the systematic carrying-over of the methods of this book to the general Bolza problem is completed (as it can be), much light will be thrown on DE of orders higher than the second.

Among the exceptions* to the special emphasis on the case $m = 1$ was the comprehensive paper by Birkhoff and Hestenes in 1935. These authors confirmed some of the principal new theorems that were presented in my Colloquium Lectures of 1932. An exception is a well-known theorem that identifies the least characteristic root λ_1 of quadratic functionals $\eta \to I_r^{\lambda}(\eta)$, associated with the second variation, with the minimum value of $I_r^0(\eta)$, under the condition that the norm $\|\eta\| = 1$. See Hilbert-Courant [1]. We shall replace† this characterization of λ_1 by our Index Theorem 15.2. This theorem effectively characterizes not only λ_1 but each characteristic root λ_n. In particular the classical characterization of λ_1 is confirmed in Corollary 16.2 as a consequence of the Index Theorem and *extended* to the case in which the coefficients of the underlying quadratic form

$$2\omega(x,\eta,\zeta) = R_{ij}(x)\eta_i'\eta_j' + 2Q_{ij}(x)\eta_i'\eta_j + P_{ij}(x)\eta_i\eta_j \tag{2}$$

are required to be no more than continuous.

From my point of view, the principal object of study in variational analysis should be the study of the origin and nature of *critical extremals*, including the study of minimizing extremals as a special case. *A critical extremal is defined as a finite extremal arc that satisfies the transversality conditions associated with the prescribed boundary conditions.* See Definition 9.1.

In addition to the study of minimizing extremals we shall emphasize two principal applications of variational analysis:

1. *Variational topology*
2. *Quadratic functionals*

Variational topology includes and extends differential topology. In variational topology extremals replace the geodesics of differential topology and the topology of the underlying space enters strongly.

The quadratic functionals studied extend the functionals given by the classical second variation. In the "free" form that we give them in Section 15 they are more general than the classical or "derived" second variation. It is a main thesis of this book that Sturm-like separation, comparison, and oscillation theorems are best understood and extended as by-products of a variational study of the relevant quadratic functionals. The study of quadratic

* Also exceptional are paper [2] and book [3] by W. T. Reid.

† In Part V a general structure underlying the theory of characteristic roots is introduced. We follow Sturm and Liouville (see Bôcher, *Méthodes de Sturm*) in allowing the characteristic parameter λ to enter very generally, both in the integrand of the quadratic functional and in the boundary conditions. An Index Theorem 33.5 still holds, even when λ does not enter linearly, and serves to characterize roots λ_n, including the least characteristic root λ_1. A characterization of λ_1 with the aid of a simple isoperimetric condition does not seem possible in the general case.

functionals, as we shall restrict them, will turn out to be the study of a general system of m nonsingular, selfadjoint, linear, homogeneous, second-order DE under the most general selfadjoint boundary conditions. See Sections 30 and 31.

It would not be possible to present an extended variational theory had not the classical theory been so well built by men such as Bolza, Bliss, Hadamard, Kneser, Mayer, von Escherich, Carathéodory, Tonelli, and their pupils. To these names we add the names of present-day expositors such as Hestenes and Gelfand. The contributions to the theory "in the large" by mathematicians such as Deheuvels, Milnor, Serre, Pitcher, Herman, Palais, Bott, Ljusternik, Snirelman, and Smale should be recalled. The lectures by L. C. Young on the "Calculus of Variations and Optimal Control Theory" cover important novel aspects of the theory which, for the sake of brevity and clarity, are not included. In my second volume on *Variational Topology* other names will be cited.

After summarizing and refining the relevant classical theorems on the minimizing extremal this book extends and organizes in a new theory the theorems on "critical" primary and secondary extremals. Theorems on critical secondary extremals take the form of a variational study of quadratic functionals that includes extensions of the Sturm theory. An indication of the methods and technical innovations is given in the Introduction.

I am indebted to Professor George Kozlowski, Professor Ping-Fun Lam, and Professor Everett Pitcher for their critical reading of the text and helpful suggestions, and to Miss Caroline Underwood for her very skillful preparation of the manuscript.

MARSTON MORSE

Princeton, New Jersey
May 1972

CONTENTS

PART III FOCAL CONDITIONS AND STURM-LIKE THEOREMS

PART IV GENERAL BOUNDARY CONDITIONS

PART V PRESTRUCTURES FOR CHARACTERISTIC ROOT THEORY

Introduction

We shall single out from each chapter special theorems or concepts which are particularly significant and indicate in what respect they are significant. These comments should be read after reading the chapter in question but before reading the following chapter.

CHAPTER 1. Under the conditions of Theorem 7.4

$$\int_{a^1}^{a^2} \omega(x,\eta(x),\eta'(x))\,dx > 0. \tag{1}$$

Classical proofs of this theorem require that the mappings,

$$x \to R_{ij}(x), \qquad x \to Q_{ij}(x), \qquad x \to P_{ij}(x) \tag{2}$$

of $[a^1,a^2]$ into R be of class C^1. It suffices that the mappings (2) be continuous. Similar remarks should be made concerning the remaining theorems of Section 7. These Theorems affect the whole body of Sturm-like theorems.

CHAPTER 2. Let C_r be general end point conditions of Section 8 with $r > 0$. The functional (J,C_r), whose values $J(\chi.\alpha)$ are given by (8.19), is the sum of an integral and an "external" function Θ whose values $\Theta(\alpha)$ are determined by the end points $(X^1(\alpha), Y^1(\alpha))$ and $(X^2(\alpha), Y^2(\alpha))$ of the graph of χ.

Theorem 10.1 shows that the second variation of (J,C_r) whose values are given by the right member of (10.13) has a similar structure. The second variation is the sum of an integral and of an *external* quadratic form $b_{hk}u_h u_k$ where the r-tuple $u = (u_1, \ldots, u_r)$ is uniquely determined, as (10.14) shows, by the end points of the graph of the "variation" η.

This similarity of structure of (J,C_r) and its second variation motivated the choice of the representation (8.19) of the functional (J,C_r) and our definition (10.23) of the basic quadratic functional $\eta \to \mathscr{I}_r(\eta)$.

CHAPTER 3. With each critical extremal g of (J^λ, C_r), Definition 14.2 associates a quadratic *index form* Q^λ whose index (Theorem 14.1) equals the "count" of negative characteristic roots of Conditions (11.8), when $r > 0$, and of conditions (12.2) when $r = 0$. This index form is termed *derived* because its definition depends on an antecedent critical extremal g. It is replaced in Section 15 by a similarly structured *free index form* whose definition is independent of any antecedent critical extremal g.

In the problem of relating the geodesics g joining two fixed points A and B, on a differentiable manifold M_n, to the homological characteristics of M_n, a derived index form is associated with each geodesic g joining A to B, and enables one to assign local homological characteristics to each g. This will be elaborated in our second volume on "variational topology." See the analogous treatment of critical points of a nondegenerate function f on a differentiable manifold. Morse and Cairns [1].

CHAPTER 4. Conditions $W_r(\lambda)$ combine the Jacobi differential equations with boundary conditions $0 < r \leq 2m$. For each $\lambda \in R$, conditions $W_r(\lambda)$: (15.0) are uniquely determined by giving $\omega^\lambda(x,\eta,\zeta)$ as in Appendix I, by prescribing a $2m \times r$ matrix $\|c_{ih}^s\|$ of rank r and an r-square symmetric "*comparison*" *matrix** $\|b_{hk}\|$. If λ ranges over R, a system of conditions $W_r(\lambda)$ in which the matrices $\|c_{ih}^s\|$ and $\|b_{hk}\|$ remain invariable, is called a *canonical system* W_r of dimension r.

With each system W_r and value $\sigma \in R$ there is associated, by Definitions 15.1 and 15.2, respectively, a quadratic functional I_r^σ : (15.2) and a *free* index form Q^σ, whose index and nullity are equal (Theorem 15.3) and are given by Index Theorem 15.2. A nonnull solution of conditions $W_r(\sigma)$ has a graph which is a "critical extremal" of I_r^σ. Sturm-like theorems result from a study of critical extremals of I_r^σ. The index form Q^σ is a technical aid in this study. The exceptional case $r = 0$ is also considered.

CHAPTER 5. The focal points of free focal conditions V_r : (17.6) and (17.7) are introduced in Definition 17.4. The Focal Point Theorem† 17.3 gives the "count" of focal points of V_r, $r \geq 0$, in terms of the characteristic roots of a special canonical system $\mathscr{W}_r$ determined by V_r in Definition 17.6. When $r = 0$, Theorem 17.3 (restated as Theorem 17.4) gives the count of conjugate points of $x = a^1$, at $x = a^2$ or in (a^1,a^2) in terms of characteristic roots of the canonical system W_0 of Conditions (15.1).

In Section 18 the theory of focal points is identified with the theory of von

* The matrix $\|b_{hk}\|$ is called a *comparison matrix* because of the basic role it plays in the extended Sturmian Comparison Theorems.

† The Focal Point Theorem was first given on page 61 in Morse [1].

Escherich families of "mutually conjugate" solutions of the system of DE

$$\frac{d}{dx}\omega_{\zeta_i}(x,\eta,\eta') = \omega_{\eta_i}(x,\eta,\eta') \qquad (i = 1,\ldots,m) \tag{3}$$

and leads to our Separation Theorem.

CHAPTER 6. The "Extended Separation* Theorem" 20.1 implies the classical Sturm Separation Theorem when $m = 1$ and replaces it when $m > 1$. It affirms the following. If two von Escherich families F and $\hat{F}$ of the DE (3) have exactly ρ linearly independent solutions in common, then the "count" of focal points of F in any relatively compact subinterval τ of $\aleph$ differs from the corresponding count for $\hat{F}$ by at most $m - \rho$.

In Section 21 we compare the focal points in an interval $(a^1,d]$ of two sets of focal conditions V_r and $\hat{V}_\rho$, as defined in (17.6) and (17.7). When $r = \rho = 0$ the comparison reduces to a comparison of conjugate points. In our first comparison theorem, Theorem 21.1, the hypotheses, when $m = 1$, are similar to classical hypotheses. When $m = 1$ certain classical theorems such as Reid's Theorem (Theorem 1.29 of Swanson) are extended in Section 22. Our second comparison theorem is the Nuclear Comparison Theorem 21.4. It has hypotheses which are definitely weaker than those of Comparison Theorem 21.1 but imply the same conclusions. When $m = 1$ Leighton's Theorem (Theorem 1.4, page 4 of Swanson) is of the same nature as our Nuclear Comparison Theorem. Leighton's Example 1, on page 6 of Swanson, is used by us for the same end. See Section 21.

CHAPTER 7. The Oscillation Theorem 24.1 differs in character from the classical oscillation theorems to which we shall turn in Section 37. Theorem 24.1 gives the exact value of the difference

$$\text{index } W_r - \text{index } W_0 \qquad (r > 0) \tag{4}$$

where index W_r denotes the count of negative characteristic roots of a canonical system W_r for which $\lambda = 0$ is not a characteristic root, and index W_0 denotes the count of conjugate points of $x = a^1$ in (a^1,a^2) of the underlying DE (3). The proof of Theorem 24.1 depends upon an auxiliary theorem, Theorem 25.1, on quadratic forms, proved in Appendix II. Many other applications of Theorem 25.1 exist. We apply Theorem 24.1 to the "periodic case," as defined in Section 27.

* The Extended Separation Theorem was first given as Theorem 7 in Morse [2] and later as Theorem 8.3 on page 104 of Morse [1].

CHAPTER 8. In Section 29, general selfadjoint BC (boundary conditions) associated with prescribed selfadjoint DE are defined, as well as the *equivalence* (Definition 29.1) of any two such sets of BC. Theorem 31.1 then implies that an arbitrary set of selfadjoint BC with an accessory r-plane χ_r, (Definition 31.1), is equivalent, when $r > 0$, to some set of BC of form (29.1). When $r = 0$ it is trivial that selfadjoint BC are equivalent to the conditions, $\eta(a^1) = \eta(a^2) = 0$. The Conditions (15.0) and (15.1) are selfadjoint, and (up to an equivalence of BC) are general in form, as we show in Section 31.

CHAPTER 9. The general systems $\dot{W}_r$ of canonical conditions $\dot{W}_r(\lambda)$ introduced in Section 32, include the systems W_r of canonical conditions $W_r(\lambda)$ of Section 15 as special cases. Theorem 34.2 gives necessary and sufficient conditions that a system $\dot{W}_r$ have infinitely many characteristic roots. It is a corollary that the canonical systems W_r of Section 15 have infinitely many such roots.

In Section 37 Oscillation Criteria are given in a set of theorems whose proofs will be presented in a separate paper. A set of DE of the form (3) is given for $x \in (0, \infty)$ and conditioned as in Appendix I. The DE are termed *oscillatory* if the point $x = 1$ has infinitely many conjugate points following $x = 1$. A D^1-mapping $x \to \gamma(x) : [1,\infty) \to R^m$ is called a *thread* and the condition that

(5) $$\liminf_{x \uparrow \infty} \int_1^x \omega(x,\gamma(x),\gamma'(x))\, dx = -\infty$$

a *thread condition.*

One of the corollaries of the general theorems presented in Section 37 for $m \geq 1$, concerns the DE

(6) $$\eta'' + a(x)\eta = 0 \qquad (0 < x < \infty;\ a(x) \geq 0)$$

where the mapping $x \to a(x)$ is continuous. See Swanson, Chapter 2. The corresponding Thread Condition has the form

(7) $$\liminf_{x \uparrow \infty} \int_1^x (\gamma'^2(x) - a(x)\gamma^2(x))\, dx = -\infty.$$

The corollary takes the form:

COROLLARY. *A necessary and sufficient condition that the DE (6) be oscillatory is that the Thread Condition (7) be satisfied by some thread γ.*

The Leighton-Wintner Theorem that the DE (6) is oscillatory if

$$\int_1^\infty a(x)\, dx = \infty,$$

is implied by the corollary, on taking the thread as the mapping $x \to \gamma(x) \equiv 1$.

Note. The major class of admissible curves is that of x-parameterized curves of class D^1, including broken extremals. This class of curves is adequate for a first simple presentation of the new theorems and structure. It is preferred to a choice of Hilbert Space from the different Banach Spaces which could be profitably used in variational analysis, because it leaves that choice fully open, while preserving a close connection with the classical theory. "Index forms" are definable in terms of broken extremals and permit an easy transition from the Sturmian properties of a critical extremal to properties which are purely topological. In our presentation of *Variational Topology* which is to follow, admissible curves will range from the merely continuous to the real analytic and include curves definable in the terminology of Lebesgue.

Part I

CRITICAL EXTREMALS

CHAPTER 1

Minimizing Extremals g: Fixed Endpoints

This chapter gives a review of classical results found in varying forms in the works of Bliss, Carathéodory, Bolza, Hadamard, Tonelli, and other writers on variational theory. Present day notation is introduced and classical proofs are freely modified.

1. The Euler Equations

To properly condition admissible curves several definitions are needed.

Mappings of Class D⁰. Let $[a^1,a^2]$ be a closed interval of the axis R of real numbers. Two intervals of the R-axis are termed *nonoverlapping* if their intersection is empty or a point. A mapping $x \to h(x) : [a^1,a^2] \to R$ will be said to be of *class* D^0 if $[a^1,a^2]$ is the union of a finite set of closed, nonoverlapping subintervals I, on the interior $\mathring{I}$ of each of which h is continuous and at each endpoint x_0 of which h has a finite limit as x_0 is approached from $\mathring{I}$.

Mappings of Class D¹. A mapping $x \to h(x) : [a^1,a^2] \to R$ will be said to be of *class* D^1 if h is continuous and if $[a^1,a^2]$ is the union of a finite set of nonoverlapping closed subintervals I on the interior $\mathring{I}$ of each of which h' exists and is continuous, and at each endpoint x_0 of which h' has a finite limit when x_0 is approached from $\mathring{I}$.

The Preintegrand f. Let $(x, y_1, \ldots, y_m)$ be written as (x,y) and be the set of rectangular coordinates of a point (x,y) in a euclidean space E_{m+1} of dimension $m + 1$. Let X be an open connected subset of E_{m+1}. Let $p = (p_1, \ldots, p_m)$ be an arbitrary point in the m-fold Cartesian product R^m of R. For brevity we write

$$(x, y_1, \ldots, y_m, p_1, \ldots, p_m) = (x,y,p) \tag{1.1}$$