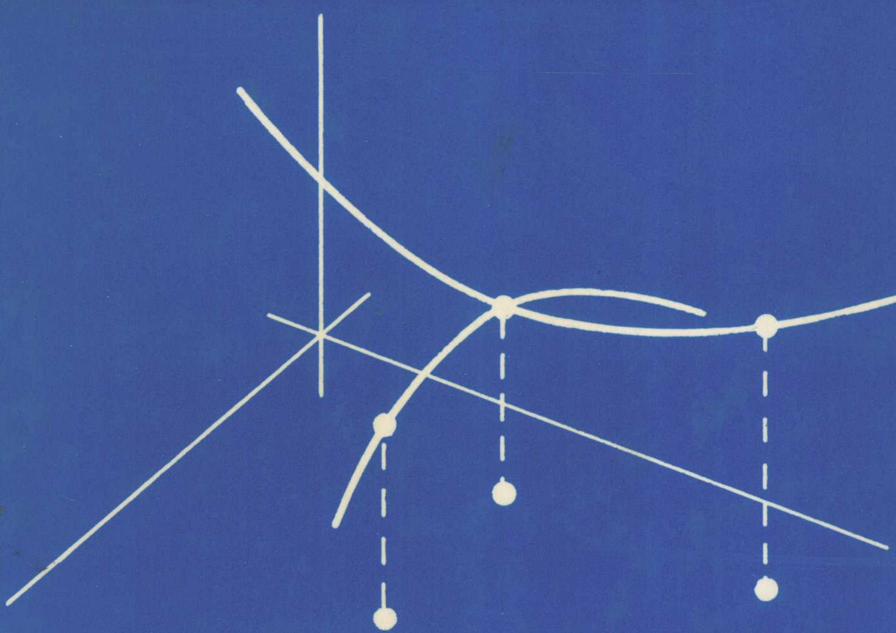
Convexity Methods in Variational Calculus

Peter Smith



RESEARCH STUDIES PRESS

Convexity Methods in Variational Calculus

Peter Smith
University of Keele, England

RESEARCH STUDIES PRESS LTD.
Letchworth, Hertfordshire, England
JOHN WILEY & SONS INC.
New York · Chichester · Toronto · Brisbane · Singapore

RESEARCH STUDIES PRESS LTD.
58B Station Road, Letchworth, Herts. SG6 3BE, England

Marketing and Distribution:

Australia, New Zealand, South-east Asia:
Jacaranda-Wiley Ltd., Jacaranda Press
JOHN WILEY & SONS INC.
GPO Box 859, Brisbane, Queensland 4001, Australia

Canada:
JOHN WILEY & SONS CANADA LIMITED
22 Worcester Road, Rexdale, Ontario, Canada

Europe, Africa:
JOHN WILEY & SONS LIMITED
Baffins Lane, Chichester, West Sussex, England

North and South America and the rest of the world:
JOHN WILEY & SONS INC.
605 Third Avenue, New York, NY 10158, USA

Library of Congress Cataloging in Publication Data:

Smith, Peter, 1935–
Convexity methods in variational calculus.
(Electronic & electrical engineering research studies. Applied and engineering mathematics series; 1)
Bibliography: p.
Includes index.
1. Calculus of variations. 2. Convex domains.
I. Title. II. Series.
QA316.S54 1985 515'.64 84-24950
ISBN 0 86380 022 X (Research Studies Press)
ISBN 0 471 90679 4 (J. Wiley)

British Library Cataloguing in Publication Data:

Smith, P. (Peter) 1935–
Convexity methods in variational calculus.—
(Electronic & electrical engineering research studies. Applied and engineering mathematics series; 1)
1. Calculus of variations. 2. Convex functions
I. Title II. Series
515'.64 QA315

ISBN 0 86380 022 X

ISBN 0 86380 022 X (Research Studies Press Ltd.)
ISBN 0 471 90679 4 (John Wiley & Sons Inc.)

Printed in Great Britain

Preface

The calculus of variations is a subject with a long history stretching back to classical times, with its origins lying in natural questions which arose from problems of maximum areas, minimum times and shortest paths in geometry and mechanics. Most of these applications lead to integrals and, as a consequence, the calculus of variations has been concerned to a large extent with the maximum and minimum properties of integrals. The realization that convexity was an important concept in variational calculus was a more recent development. If the convexity of an integral can be established then it usually also implies a minimum principle. Of course not all variational problems are convex, but it does happen to be so in a large number of important cases. Whilst many standard and classical applications lead to integrals of real functions, there is no reason why the space of real functions cannot be generalized to vector spaces and the mapping developed into a functional instead of an integral. In these terms a variational principle follows when at least one element in the vector space can be found which makes the functional stationary, and an extremum principle is obtained when some assertion can be made about the minimum, maximum or saddle behaviour of the functional at its stationary value.

Any approach to this subject along the lines indicated in the previous paragraph, inevitably requires a background in elementary applied functional analysis. The first chapter attempts to supply this need within the requirements dictated by the other two chapters. This chapter contains sections on vector spaces, norms, inner products, various properties of linear and nonlinear operators and functionals and convexity, together with a brief description of the Lebesgue integral. In the second chapter the calculus of operators is introduced, including the Gâteaux, Fréchet and gradient operators. For differentiable functionals, convexity can be defined in terms of these generalized derivatives, and minimum principles follow from the resulting inequality. In chapter 3 the functionals are defined over the product space of two inner product spaces. If it can be shown that such functionals have a global saddle structure then dual extremum principles follow, and these provide upper and lower bounds for the stationary values of the functionals.

The material presented in this book arose from the author's interests in extremum principles and nonlinear differential equations, and from the teaching, over a number of years, of final year courses in these subjects at the University of Keele. The book is intended as an introduction to various techniques which can be applied to those differentiable convex functionals which arise from operator equations in applied mathematics. The approach is intended to be fairly informal and deliberately limited in length and depth; it is mainly directed to those whose interests lie in the applications of these methods. The treatment is kept at a reasonably intuitive level so that the reader will have results and techniques readily available. To this end there are 65 worked examples in the text and about 90 problems and applications at the ends of the chapters together with brief answers at the end of the book. Most of the specific applications are taken from differential equations which are usually presented in reduced or simplified form; the actual derivation of equations from their original physical or engineering context is not attempted here.

My thanks are due to Peter Kendall for inviting me to contribute this book to the series, and also for his valuable comments on an earlier draft of the book. I am particularly grateful to

Olwen Brindley for her skill and accuracy in the difficult task of setting out the final typescript.

Keele, Staffordshire

1984

Peter Smith

Contents

CHAPTER 1
Spaces, Operators and Convexity

1.1 Introduction

Many of the ideas presented in this book have their origins in concepts in real function theory. It will be instructive to briefly formulate such problems in this context. Let $f(x)$ be a real function of the real variable x defined in the open interval $x \varepsilon (a, b)$ on the real line R. If the interval is denoted by I, then the function may be represented by the mapping notation $f : I \to R$. Here I denotes the domain of the function f, that is, the set of values of x over which f is defined, and the range of f, which is the set of values taken by f, lies in R. This notation is used extensively in this book and will be discussed in more detail in the next section. In addition suppose that f is differentiable on I, so that $f' : I \to R$ where $f'(x)$ is the usual ordinary derivative of $f(x)$ at x. At any particular value of x for which $f'(x)$ vanishes, the function $f(x)$ is said to be *stationary*. Thus if $x_0 \varepsilon I$ and $f'(x_0) = 0$ then $f(x)$ has the stationary value $f(x_0)$. Such a statement expresses a *variational principle*.

It might be possible to obtain more information about the nature of the stationary value: is it a maximum or minimum value or a point of inflection? For real functions, for example, $f''(x_0) > 0$ is a sufficient condition for a minimum value of $f(x)$ at $x = x_0$. A more general condition for a minimum can be expressed in terms of the sign of $f'(x)$ in a neighbourhood of $x = x_0$: there should exist an interval $|x - x_0| \leq \delta$ such that $f'(x) < 0$ for $x \varepsilon [x_0 - \delta, x_0)$ and

$f'(x) > 0$ for $x \in (x_0, x_0 + \delta]$. This minimum value of the function gives a local extremum principle. At least for $x \in [x_0 - \delta, x_0 + \delta]$ and $x \neq x_0$, we can say that

$$f(x) > f(x_0) \tag{1.1}$$

that is, the function value is greater than its stationary value. In this instance we have a local minimum principle. There might also exist further local minimum or maximum values on the interval I. If, however, (1.1) holds for all $x \in I$ then the statement illustrates what is known as a *global extremum principle*, which, of course, does not preclude the existence of other local extremum principles.

There is a particular set of functions for which any minimum value is automatically a global minimum. These are functions which are *convex*. Twice differentiable convex real functions can be characterised by the inequality $f''(x) \geq 0$ for $x \in I$. If the equality is missing from the previous inequality then the function is said to be *strictly convex* and any stationary minimum value is then unique.

This real function analysis can be extended to include functions of more than one variable in which case the stationary value can include the saddle point in addition to the maximum and the minimum. The location of stationary values for real functions is probably the simplest example of a variational principle.

The calculus of variations in its more general context is concerned with the problem of finding the stationary or extreme values of real quantities which are not necessarily real functions of real variables. These quantities or functionals as they are known may be integrals, scalar products, series, etc, which themselves contain unknown functions. The classical calculus of variations problem typically involves the construction of the stationary value of integrals of the form

$$J(y) = \int_{a_1}^{a_2} F(x, y, y')\, dx, \qquad y' \equiv dy/dx, \tag{1.2}$$

subject to boundary conditions on $y(x)$ at $x = a_1$ and $x = a_2$. If the integral $J(y)$ is stationary for the extremal $y = z(x)$, then variations of $J(y)$ are constructed by letting $y = z(x) + \varepsilon\eta(x)$ where ε is a real parameter and $\eta(x)$ is a real twice-differentiable function which vanishes at $x = a_1$ and $x = a_2$. After the expansion of $F(x, z(x) + \varepsilon\eta(x), z'(x) + \varepsilon\eta'(x))$ in powers of ε using a Taylor expansion, the integral becomes

$$J(z + \varepsilon\eta) = J(z) + \delta J + O(\varepsilon^2) \tag{1.3}$$

where δJ is the so-called *first variation* of J given by

$$\delta J = \varepsilon \int_{a_1}^{a_2} \left\{\frac{\partial F}{\partial z} - \frac{d}{dx}\left(\frac{\partial F}{\partial z'}\right)\right\} dx. \tag{1.4}$$

Stationary values occur, by definition, where the first variation vanishes. The argument proceeds by a proof that $\delta J = 0$ for all η if z satisfies the Euler-Lagrange equation

$$\frac{\partial F}{\partial z} - \frac{d}{dx}\left(\frac{\partial F}{\partial z'}\right) = 0. \tag{1.5}$$

Full descriptions of the Euler-Lagrange theory and its many generalisations can be found in standard texts on the calculus of variations (see, for example, Akhiezer (1962), Cesari (1983), Clegg (1968), Pars (1962), Weinstock (1952) and Smith (1974)). Investigations concerning the status of the stationary value usually depend on the behaviour of the second variation in the expansion of (1.3).

The right side of (1.3) is really a generalised Taylor expansion in which the first variation can be considered as the generalised 'first derivative' of the functional which defines J. It is the vanishing of this derivative which leads to the Euler-Lagrange equations. The aim of this text is to emphasize, as far as possible, this particular approach to variational problems. The application of convexity to the calculus of variations has received considerable attention during the past twenty years. Much of the theory is now well-established, and it seems an appropriate time to record a broad view of the current state of the subject. Also there does exist in the literature a large body of applications, illustrations and examples which can be drawn on to develop the theoretical ideas. However many of the examples will be given in reduced or simplified form; it is not possible in a book of this length to fill in the necessary background theory or the physical meaning or interpretation of many of the extremum principles. In many areas the physical principles lead directly to the appropriate variational problem. But this is not always the case. The link between nonlinear equations and an underlying variational principle can be often not intuitively obvious, nor, for that matter, need it be a unique principle. This text is specifically directed towards a study of the stationary values of those functionals which possess some additional property such as convexity or saddle structure.

A broad account of variational methods with applications in classical physics can be found in the book by Mikhlin (1964). The works of Arthurs (1970, second edition 1981) and Robinson (1971) are concerned with dual or complementary variational principles, functional differential and the canonical formulation of physical problems and are relevant to the account presented here. A modern treatment of the variational method based on functional analytic techniques has been given by Oden and Reddy (1976), and an account using the ideas of Gâteaux differentiation and monotone operators can be found in the work of Vainberg (1973). Early work on dual extremum principles was initiated by Noble (1964) and Sewell (1969). A complete study of dual extremum principles and associated Legendre transformations up to 1972 can be found in the survey article by Noble and Sewell (1972). This survey also shows the particular links between extremum principles and Lagrangian and Hamiltonian methods in applied mathematics. This work has been updated in more recent surveys by Sewell (1979, 1982), which also develop new connections with bifurcation theory.

The applications of dual extremum methods are now varied, ranging from the traditional variational problems in solid and fluid mechanics to network theory, optimization and nonlinear programming. Recent research interest has concentrated on the areas of dissipative systems, initial-value problems for ordinary and partial differential equations and nonlinear applications in general. As we remarked previously, nonlinear problems (and some linear ones also) may have no 'natural' variational context. The inverse variational question becomes the following: given a set of equations which arise, say, in some physical setting, can we design a functional such that the stationary value of the functional occurs where the equations are satisfied? Obviously there is no question of this functional being unique. However further constraints may be necessary or desirable. For extremum principles we may require convexity and we may demand that the stationary value takes a prescribed form which is of some physical or practical interest.

Any modern treatment of this subject must draw heavily on the ideas of applied functional analysis. The development of functional analysis in the early part of this century created a natural setting for the abstraction and generalisation of classical calculus of variations. Real variables are replaced by elements of vector spaces, integrals become functionals and the Euler-Lagrange form becomes the gradient of the functional. This chapter continues with an introduction to these parts of applied functional analysis which are necessary for variational principles. Inevitably it has to be of limited scope. More detailed treatments can be found in the books of Griffel (1981), Hutson and Pym (1980), Milne (1980), Wouk (1979), Stakgold (1979) and Hille (1972).

1.2 Vector spaces

We are interested in mappings between sets of *elements*. Suppose that U and V are two sets of elements. We say that a function f *maps* U into V if to each element $u \in U$ there corresponds exactly one element v in V. We write $v = f(u)$ or $f : U \to V$ to emphasise the mapping. In the latter notation U is called the *domain* of f and is sometimes denoted by D_f instead of U to emphasise the function dependence. Frequently we shall assume that $u \in D_f \subseteq U$ that is D_f is a subset of U which can occur, for example, if additional conditions are required of the elements in U. The set of elements given by $f(u)$ is called the *range* of f and denoted by R_f. Since the range is often not precisely or conveniently defined it is often assumed to be part of a larger set V. For example we may know that $f(x)$ is a real function but not precisely the range of its values. If $R_f = V$ we say that the mapping is onto V, and if $R_f \subset V$ we say that the mapping is into V. Strictly the mapping should be written

$$f : D_f \subseteq U \rightarrow R_f \subseteq V$$

as illustrated in figure 1.1. The mapping is said to be *injective* if $u_1 \neq u_2$ implies $f(u_1) \neq f(u_2)$ for every u_1, u_2. The mapping is said to be *surjective* (or a mapping onto V) if $R_f = V$. Finally the mapping is said to be *bijective* if it is both injective and surjective. We can associate inverse mappings with bijective ones as

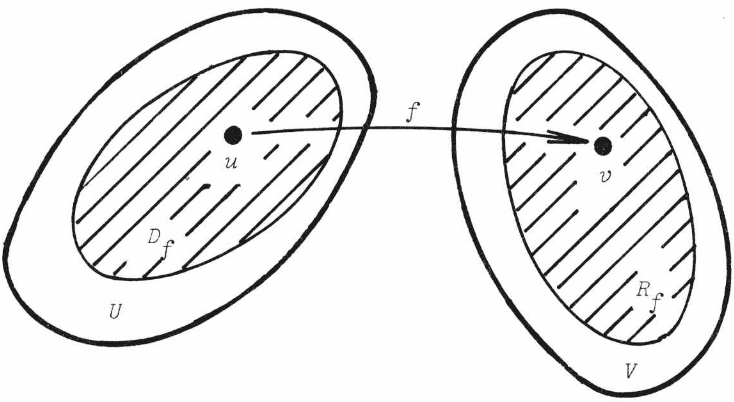


FIG. 1.1

we shall see later. For any given u, v is called its *image*.

Example 1.1 If $U = \{u | (-\infty, \infty)\}$, then U is the real line or the set of real numbers which we shall denote by R. In particular, if $f(u) = \cos u$, then we can put $V = R$ with the range given by $R_f = \{v | [-1, 1]\}$. Hence $f : R \rightarrow R_f \subset R$.

The *Cartesian product* of two non-empty sets U_1 and U_2 is denoted by $U_1 \times U_2$ and represents the set of all ordered pairs (u_1, u_2) with $u_1 \in U_1$ and $u_2 \in U_2$.

Example 1.2 Let U be the set of all n-tuples of real numbers. We write $u = (u_1, u_2, \ldots, u_n)$ where $u_i \in R$ $(i = 1, 2, \ldots, n)$. Each component can be chosen independently from the real numbers. The set U is the product space $R \times R \times \ldots \times R$ (n times) which we write as R^n. It is known as Euclidean n-space.

Example 1.3 We denote the set of real functions which are continuous on the closed interval $[a, b]$ by $C[a, b]$. If they are n times continuously differentiable the set will be represented by $C^n[a, b]$. Let $u \in C^1[a, b]$ be a particular element. Then the mapping $v = du/dt$ maps elements from $C^1[a, b]$ into the space of continuous functions $C[a, b]$. If the mapping is represented symbolically by $v = Du$ then the operation is given by $D : C^1[a, b] \to C[a, b]$. In this example the elements are themselves functions so that the function is an element in some larger set. For this reason $C[a, b]$ is an example of a function set or space.

The terms *operator* and *transformation* are also used as alternatives to function, although operator is sometimes restricted to mappings between function spaces (rather than real numbers) but the usage is by no means firm. However one particular mapping has a special name. If the range of the mapping is a subset of the real numbers then the mapping is known as a *functional* (the term is also used if the range is a subset of the complex numbers but we are almost exclusively interested in real functionals in this context).

Example 1.4 If $u(t) \in C^1[a, b]$, then

$$N(u) = \int_a^b \left[u^2 + \left(\frac{du}{dt}\right)^2\right] dt$$

is a functional. In this case the mapping can be represented by $N : C^1[a, b] \to R$, that is the operator N maps real functions which have continuous first derivatives onto the real line.

We require an algebraic structure for the elements of the set, and this is provided by assuming that the elements form a vector or linear space. To some extent this allows elements in the space to be added and multiplied by constants in a conventional way. For this reason the elements are often referred to as vectors.

Let u, v, w, ... be elements of a set U. Then U is called a *vector space* if vector addition $u + v \in U$ of any two elements u and v can be defined which satisfies the axioms

(i) $u + v = v + u$,

(ii) $u + (v + w) = (u + v) + w$,

(iii) there exists a unique zero vector 0 such that $u + 0 = u$ for every $u \in U$,

(iv) for every vector u there exists a unique vector $-u$ such that $u + (-u) = 0$,

and scalar multiplication $\alpha u \in U$ for any vector u can be defined which satisfies the axioms

(v) $\alpha(\beta u) = (\alpha\beta)u$,

(vi) $1.u = u$,

(vii) $\alpha(u + v) = \alpha u + \alpha v$,

(viii) $(\alpha + \beta)u = \alpha u + \beta u$.

If the scalars are real (complex) numbers then U is called a real (complex) vector space.

As we stated previously the Cartesian product $U_1 \times U_2$ of two non-empty vector spaces U_1 and U_2 is the set of all ordered pairs (u_1, u_2) where $u_1 \in U_1$ and $u_2 \in U_2$. In the axioms of the vector space, vector addition is a mapping $U \times U \to U$ whilst for a real vector space scalar multiplication is a mapping $R \times U \to U$.

Example 1.5 The space R^n with vector addition and scalar multiplication defined by the operations

$$u + v = (u_1 + v_1, \ldots, u_n + v_n),$$

$$\alpha u = (\alpha u_1, \ldots, \alpha u_n),$$

for any vectors $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$ belonging to R^n is a vector space. It is easy to verify that axioms (i) - (viii) are satisfied by these vectors.

Example 1.6 the space $C^n[a, b]$ is a vector space. If $u(x)$ and $v(x)$

are any two functions belonging to $C^n[a, b]$ then $u(x) + v(x)$ and $\alpha u(x)$ are both n times continuously differentiable on the interval. The remaining axioms are easily confirmed.

Example 1.7 Let ℓ_p be the set of all of all real sequences $u = (u_i) = (u_1, u_2, \ldots)$ such that for $p \geq 1$ the series $\Sigma_{i=1}^{\infty} |u_i|^p$ converges. Define the algebraic operations of vector addition and scalar multiplication by

$$u + v = (u_1, u_2, \ldots) + (v_1, v_2, \ldots) = (u_1 + v_1, u_2 + v_2, \ldots),$$

$$\alpha u = \alpha(u_1, u_2, \ldots) = (\alpha u_1, \alpha u_2, \ldots).$$

Then this set forms a vector space.

The proof requires the Minkowski inequality

$$\left[\sum_{i=1}^{\infty} |u_i + v_i|^p\right]^{\frac{1}{p}} \leq \left[\sum_{i=1}^{\infty} |u_i|^p\right]^{\frac{1}{p}} + \left[\sum_{i=1}^{\infty} |v_i|^p\right]^{\frac{1}{p}},$$

to establish the convergence of

$$\sum_{i=1}^{\infty} |u_i + v_i|^p,$$

(see Kreysig (1978) for a proof of this result; some hints are also given in Problem 1.7).

A subspace U' of a vector space U is a nonempty subset $U' \subseteq U$ such that for all $u_1, u_2 \in U'$ and all scalars α and β, $\alpha u_1 + \beta u_2 \in U'$ and U' is itself a vector space. On the other hand a subset of U is simply a collection of elements drawn from U which may not necessarily satisfy the axioms of a vector space.

Example 1.8 Consider the set of real functions which are continuous for $x \in [a, b]$ but which vanish at $x = a$ and $x = b$. Clearly any linear combination of any two elements of this set must also belong to the set, say U', and U' also satisfies the axioms of a vector space. It is a subspace of $C[a, b]$ and we can write $U' \subset C[a, b]$. On the other hand, the subset of functions which take values of 1 and