COMPLEMENTARY VARIATIONAL PRINCIPLES

BY

A. M ARTHURS

Second Edition

COMPLEMENTARY VARIATIONAL PRINCIPLES

BY

A. M ARTHURS

Second Edition

CLARENDON PRESS · OXFORD
1980

Oxford University Press, Walton Street, Oxford OX2 6DP

OXFORD LONDON GLASGOW
NEW YORK TORONTO MELBOURNE WELLINGTON
KUALA LUMPUR SINGAPORE JAKARTA HONG KONG TOKYO
DELHI BOMBAY CALCUTTA MADRAS KARACHI
NAIROBI DAR ES SALAAM CAPE TOWN

Published in the United States by Oxford University Press, New York

First edition 1970

British Library Cataloguing in Publication Data

Arthurs, Arnold Magowan
Complementary variational principles. –
2nd ed. – (Oxford mathematical monographs).
1. Boundary value problems
I. Title II. Series
515′.62 QA379 80–40613

ISBN 0-19-853532-5

Filmset by Universities Press, Belfast
Printed in Great Britain by
Lowe & Brydone Printers Ltd, Thetford, Norfolk

PREFACE

THE book has been mostly rewritten to bring in various improvements and additions. In particular, I have replaced the local theory with a global treatment based on simple ideas of convexity and monotone operators. Another major change is that the class of problems treated is much wider than the Dirichlet type originally discussed. In addition, the variational results are given a geometrical formulation that includes the hypercircle, and error estimates for variational solutions are also described.

The number of applications to linear and nonlinear boundary value problems has been doubled, covering some thirty cases which arise in mathematical physics, chemistry, engineering, and biology. As well as containing new derivations of well-known results such as the Rayleigh and Temple bounds for eigenvalues, the examples contain many results on upper and lower bounds that have only recently been obtained.

The book is written at a fairly elementary level and should be accessible to any student with a little knowledge of the calculus of variations and differential equations.

I wish to thank Professor J L Synge of Dublin and Dr N Anderson of York for many helpful discussions and suggestions on the material presented here. I am also grateful to the editors and the Clarendon Press for including this monograph in their series.

University of York
September, 1979

A. M. A.

CONTENTS

1

VARIATIONAL PRINCIPLES: INTRODUCTION

1.1. Introduction

VARIATIONAL principles play an important part in mathematics and the physical sciences for three main reasons: they (i) unify many diverse fields, (ii) lead to new theoretical results, and (iii) provide powerful methods of calculation. Thus, the well-known Euler–Lagrange principle can be used to derive field equations of many kinds, extremum principles lead to new estimates for important physical quantities, and direct methods form the basis of very accurate computations (cf. Gould 1966, Mikhlin 1964, Mitchell and Wait 1977). Many problems, however, are usually first posed in the form of differential equations, or more generally as operator equations, and there is no guarantee that an equivalent variational problem exists. Even if we know that an equivalent problem does exist, it may not be easy to find an explicit form for the variational expression. Stated in mathematical terms, the problem is to find the potential (or action) corresponding to a given field equation (cf. Vainberg 1963, 1973). Of course in some branches of mathematical physics, such as classical dynamics, the variational problem is known once the Lagrangian is specified. As it turns out, all the results obtained in this book are examples of this latter kind, for which the basic action functional is readily found. Our particular interest centres on principles which lead to variational bounds, and especially in those cases for which both maximum and minimum (complementary) principles can be obtained. In many applications these complementary extremum principles provide upper and lower bounds for quantities of interest, and they are also important because of their utility for establishing bounds on approximate solutions of a wide class of boundary value problems.

One of the earliest examples of complementary principles is provided by the energy principle in the theory of structures, together with the principle of complementary energy (Trefftz 1928). Another example concerns the Dirichlet and Thomson bounds in electrostatics, while yet another is Rayleigh's bounds in acoustics (Rayleigh 1899). There are several methods, closely related, by which complementary principles can be derived. The first such method, due to Friedrichs (1929) and Courant and Hilbert (1953), employs transformations to canonical and involutory form, while another, which applies to certain linear problems, employs the hypercircle approach of Prager and Synge (cf. Synge 1957). This latter method is a geometrical version of the canonical variational method (Arthurs 1977b), though it was originally developed quite separately.

More recently, starting with the work of Noble (1964), these ideas have been expanded and generalized to form a coherent theory of complementary variational principles for boundary-value problems. This theory, which forms the subject of the present monograph, provides a systematic approach to many linear and nonlinear problems involving differential, integral, and matrix equations.

As we shall see, the key results concern differentiable functionals $I(u, \phi)$ of two independent functions. Such functionals are stationary at a solution (u_0, ϕ_0) of the Euler equations

$$I_u = 0, \qquad I_\phi = 0, \tag{1.1.1}$$

where subscripts denote differentiation. If $I(u, \phi)$ is concave in u and convex in ϕ, then the complementary extremum principles

$$I(u_2, \phi_2) \leqslant I(u_0, \phi_0) \leqslant I(u_1, \phi_1) \tag{1.1.2}$$

hold, where the functions (u_1, ϕ_1) and (u_2, ϕ_2) satisfy

$$I_u(u_1, \phi_1) = 0 \quad \text{and} \quad I_\phi(u_2, \phi_2) = 0. \tag{1.1.3}$$

The pair of equations in (1.1.1) represents an abstract form of the canonical Euler–Hamilton equations in the calculus of variations.

1.2. Euler–Lagrange theory

The variational principles described in this book have their origins in the simplest kind of variational problem that can be treated by the Euler–Lagrange theory. Thus they are basically concerned with differentiable functionals of the form

$$E(\Phi) = \int_a^b L(x, \Phi, \Phi')\, dx, \qquad \Phi' = d\Phi/dx, \tag{1.2.1}$$

with fixed end-points

$$\Phi(a) = \alpha, \qquad \Phi(b) = \beta. \tag{1.2.2}$$

Here Φ belongs to the class C_2 of functions which have continuous derivatives up to second order for $a \leqslant x \leqslant b$, and L is assumed to possess continuous second-order derivatives with respect to all its arguments. Of course, the assumptions just made can be relaxed to some considerable extent (cf. Gelfand and Fomin 1963, Pars 1962), but we shall not deal with that aspect of the theory here.

Suppose now that the functional $E(\Phi)$ has an extremum at ϕ. Then we consider variations round ϕ

$$\Phi = \phi + \varepsilon\xi. \tag{1.2.3}$$

If Φ and ϕ both satisfy the end-point conditions (1.2.2), it follows that

$$\xi(a) = \xi(b) = 0. \tag{1.2.4}$$

Since $E(\Phi)$ is differentiable, we can write

$$E(\phi + \varepsilon\xi) = E(\phi) + \delta E(\phi, \varepsilon\xi) + \delta^2 E(\phi, \varepsilon\xi) + \ldots, \tag{1.2.5}$$

where the first variation is

$$\delta E = \varepsilon \int_a^b \left\{ \xi \frac{\partial L}{\partial \phi} + \xi' \frac{\partial L}{\partial \phi'} \right\} \mathrm{d}x, \tag{1.2.6}$$

and the second variation is

$$\delta^2 E = \frac{1}{2} \varepsilon^2 \int_a^b \left\{ \xi^2 \frac{\partial^2 L}{\partial \phi^2} + 2\xi\xi' \frac{\partial^2 L}{\partial \phi\, \partial \phi'} + \xi'^2 \frac{\partial^2 L}{\partial \phi'^2} \right\} \mathrm{d}x. \tag{1.2.7}$$

In (1.2.6)

$$\partial L/\partial \phi = \partial L(x, \Phi, \Phi')/\partial \Phi \quad \text{at} \quad \Phi = \phi,$$

and

$$\partial L/\partial \phi' = \partial L(x, \Phi, \Phi')/\partial \Phi' \quad \text{at} \quad \Phi = \phi,$$

with the second derivatives in (1.2.7) defined similarly. Integrating by parts in (1.2.6), we obtain an alternative expression for the first variation

$$\delta E = \varepsilon \int_a^b \xi \left\{ \frac{\partial L}{\partial \phi} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial L}{\partial \phi'} \right\} \mathrm{d}x + \varepsilon \left[\xi \frac{\partial L}{\partial \phi'} \right]_a^b. \tag{1.2.8}$$

Since the variations are such that ξ vanishes at the end-points, this reduces to

$$\delta E = \varepsilon \int_a^b \xi \left\{ \frac{\partial L}{\partial \phi} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial L}{\partial \phi'} \right\} \mathrm{d}x. \tag{1.2.9}$$

For the functional $E(\Phi)$ to have an extremum at $\Phi = \phi$ it is necessary that the first variation vanish. From (1.2.9) this means that

$$\int_a^b \xi \left\{ \frac{\partial L}{\partial \phi} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial L}{\partial \phi'} \right\} \mathrm{d}x = 0. \tag{1.2.10}$$

Since ξ is arbitrary in the interval (a, b), it follows from (1.2.10) that

$$\frac{\partial L}{\partial \phi} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial L}{\partial \phi'} = 0, \qquad a < x < b, \tag{1.2.11}$$

which is the Euler–Lagrange equation. We therefore can state

THEOREM 1.2.1. *The integral $E(\Phi)$ in (1.2.1) is stationary at ϕ where ϕ is a solution of*

$$\frac{\partial L}{\partial \phi}-\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial \phi'}=0, \qquad a<x<b \tag{1.2.12}$$

with

$$\phi(a)=\alpha, \qquad \phi(b)=\beta. \tag{1.2.13}$$

This is the Euler–Lagrange variational principle.

As a slight extension of this we note that the constraint (1.2.2) on the admissible functions Φ can be removed if instead of $E(\Phi)$ we consider the functional

$$J(\Phi)=\int_a^b L(x,\Phi,\Phi')\,\mathrm{d}x-\left[\frac{\partial L}{\partial \Phi'}(\Phi-\phi_B)\right]_a^b, \tag{1.2.14}$$

where

$$\begin{aligned}\phi_B&=\alpha \quad \text{at} \quad x=a,\\ &=\beta \quad \text{at} \quad x=b.\end{aligned} \tag{1.2.15}$$

For varied curves $\Phi=\phi+\varepsilon\xi$, the first variation of J is

$$\begin{aligned}\delta J&=\varepsilon\int_a^b\left\{\xi\frac{\partial L}{\partial \phi}+\xi'\frac{\partial L}{\partial \phi'}\right\}\mathrm{d}x-\varepsilon\left[\frac{\partial L}{\partial \phi'}\xi+(\phi-\phi_B)\left(\xi\frac{\partial^2 L}{\partial \phi\,\partial \phi'}+\xi'\frac{\partial^2 L}{\partial \phi'^2}\right)\right]_a^b\\ &=\varepsilon\int_a^b\xi\left\{\frac{\partial L}{\partial \phi}-\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial \phi'}\right\}\mathrm{d}x-\varepsilon\left[\left(\xi\frac{\partial^2 L}{\partial \phi\,\partial \phi'}+\xi'\frac{\partial^2 L}{\partial \phi'^2}\right)(\phi-\phi_B)\right]_a^b.\end{aligned} \tag{1.2.16}$$

For any variation ξ, so that ξ is arbitrary in (a, b) and at $x=a, b$, the stationary condition $\delta J=0$ implies that ϕ is a solution of

$$\frac{\partial L}{\partial \phi}-\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial \phi'}=0, \qquad a<x<b$$

with

$$\phi=\phi_B \quad \text{at} \quad x=a, b,$$

that is, ϕ satisfies (1.2.12) and (1.2.13). In this way the boundary conditions on the critical function ϕ come out naturally from the stationary principle for $J(\Phi)$.

Equation (1.2.10) gives a necessary condition for an extremum, but, in general, one which is not sufficient. In many cases, however, the Euler–Lagrange equation (1.2.12) is itself enough to give a complete solution of the problem, and in fact the existence of an extremum is often clear from the physical meaning of the problem. If in such a case there exists only one extremal (critical curve) ϕ satisfying the boundary conditions of the problem, this extremal must be the function for which the extremum is attained.

Assuming that we have found a function ϕ which makes $E(\Phi)$ or $J(\Phi)$ stationary, we now wish to consider the nature of the extremum, that is, its maximum or minimum properties. If the inequality

$$J(\phi) \leqslant J(\Phi) \tag{1.2.17}$$

holds for all admissible functions Φ, we have a *minimum* principle, since $J(\Phi)$ takes its smallest value for $\Phi = \phi$. Similarly, if the inequality

$$J(\Phi) \leqslant J(\phi) \tag{1.2.18}$$

holds for all admissible functions Φ, we have a *maximum* principle. General conditions which lead to these principles are given in section 2.5. Here we shall confine our attention to the integral (1.2.14) for which

$$\begin{aligned}\Delta J &= J(\phi + \varepsilon\xi) - J(\phi)\\ &= \frac{1}{2}\varepsilon^2 \int_a^b \left\{ \xi^2 \overline{\frac{\partial^2 L}{\partial \phi^2}} + 2\xi\xi' \overline{\frac{\partial^2 L}{\partial \phi\, \partial \phi'}} + \xi'^2 \overline{\frac{\partial^2 L}{\partial \phi'^2}} \right\} \mathrm{d}x\\ &\quad + \varepsilon^2 \left[\xi^2 \overline{\frac{\partial^2 L}{\partial \phi\, \partial \phi'}} + \xi\xi' \overline{\frac{\partial^2 L}{\partial \phi'^2}} \right]_a^b,\end{aligned} \tag{1.2.19}$$

where the overbar indicates that the factor is evaluated for some function $\bar{\phi} = \phi + \eta\varepsilon\xi$, $0 < \eta < 1$. This expression may in various cases lead to a definite result for the sign of ΔJ and hence to a maximum or a minimum principle.

Example. To illustrate these results we consider a quadratic L given by

$$L(x, \Phi, \Phi') = \tfrac{1}{2}p(\Phi')^2 + \tfrac{1}{2}w\Phi^2 - q\Phi, \tag{1.2.20}$$

where p, w, and q may be functions of x. By theorem 1.2.1 the associated integral (1.2.14) is stationary at a solution ϕ of the Sturm–Liouville equation

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}\phi}{\mathrm{d}x}\right) + w\phi = q, \qquad a < x < b, \tag{1.2.21}$$

subject to

$$\phi(a) = \alpha, \qquad \phi(b) = \beta. \tag{1.2.22}$$

If we now expand

$$J(\Phi)=\int_a^b \{\tfrac{1}{2}p(\Phi')^2+\tfrac{1}{2}w\Phi^2-q\Phi\}\,dx-[p\Phi'(\Phi-\phi_B)]_a^b \qquad (1.2.23)$$

about the critical curve ϕ, we obtain by (1.2.19)

$$J(\Phi)-J(\phi)=\tfrac{1}{2}\varepsilon^2\int_a^b (p\xi'^2+w\xi^2)\,dx-\varepsilon^2[p\xi'\xi]_a^b, \qquad (1.2.24)$$

there being no terms higher than the second since L is quadratic here. Now, if $p>0$ and $w\geq 0$, the expression in (1.2.24) is non-negative provided we make $\xi=0$, that is $\Phi=\phi_B$ on $\partial[a,b]$. Hence if $p>0$ and $w\geq 0$, we have the minimum principle

$$J(\phi)\leq J(\Phi) \qquad (1.2.25)$$

for all admissible functions Φ satisfying (1.2.2). If $p=1$ and $w=q=0$, this result corresponds to a one-dimensional version of the well-known Dirichlet principle (cf. Pars 1962).

These results show that variational problems which are formulated in terms of finding relative minima (or maxima) of Euler functionals lead in a natural way only to upper (or lower) bounds for the stationary value of the functional. However, certain minimization problems in the calculus of variations can be transformed into maximization problems. From a combination of these two problems it is then possible to obtain upper *and* lower bounds on the stationary value. One approach to these complementary bounds, due to Friedrichs (1929), is based on involutory (Legendre) transformations. A related approach, due to Courant and Hilbert (1953), Synge (1957), and Noble (1964), is based on the canonical form of Euler theory, and it is this method which we now consider.

1.3. Canonical formalism

In the canonical approach functionals are expressed in terms of Φ and of the conjugate variable U defined by

$$U=\partial L/\partial\Phi'. \qquad (1.3.1)$$

The Hamiltonian is defined by

$$H(x,U,\Phi)=U\Phi'-L(x,\Phi,\Phi'), \qquad (1.3.2)$$

where it is assumed that (1.3.1) gives Φ' as a function of x, Φ, and U. Then the action functional (1.2.1) is written as

$$I(U,\Phi)=\int_a^b \{U\Phi'-H(x,U,\Phi)\}\,dx, \qquad (1.3.3)$$

the notation indicating that the integral I is to be treated as a functional of the independent functions U and Φ. As we shall see in theorem 1.3.1, the associated Euler–Lagrange equations are

$$\frac{d\phi}{dx}=\frac{\partial H}{\partial u}, \qquad -\frac{du}{dx}=\frac{\partial H}{\partial \phi}. \tag{1.3.4}$$

These are the so-called canonical Euler (or Hamilton) equations, which are well known for the part they play in analytical dynamics (Lanczos 1966).

To discuss the variational theory associated with (1.3.4) in detail, we introduce the differentiable functional

$$I(U,\Phi)=\int_a^b \{U\Phi'-H(x,U,\Phi)\}\,dx-[U(\Phi-\phi_B)]_a^b, \tag{1.3.5}$$

where the boundary term corresponds to that in (1.2.14) and is included to make the resulting boundary conditions natural. Now let u, ϕ be critical curves of $I(U, \Phi)$ and consider variations round u and ϕ by setting

$$U=u+\varepsilon v, \qquad \Phi=\phi+\varepsilon\xi, \tag{1.3.6}$$

v and ξ being arbitrary admissible functions. Then we obtain

$$I(U,\Phi)=I(u,\phi)+\delta I+O(\varepsilon^2), \tag{1.3.7}$$

where the first-order terms are

$$\begin{aligned}\delta I&=\varepsilon\int_a^b \{u\xi'+v\phi'-\xi H_\phi-vH_u\}\,dx-\varepsilon[v(\phi-\phi_B)+u\xi]_a^b\\ &=\varepsilon\int_a^b \{v(\phi'-H_u)-\xi(u'+H_\phi)\}\,dx-\varepsilon[v(\phi-\phi_B)]_a^b.\end{aligned} \tag{1.3.8}$$

Here subscripts on H denote partial derivatives evaluated at (u, ϕ). For the functional (1.3.5) to be stationary at $U=u$, $\Phi=\phi$, it is necessary that $\delta I=0$. From (1.3.8) this means that

$$\int_a^b \{v(\phi'-H_u)-\xi(u'+H_\phi)\}\,dx-[v(\phi-\phi_B)]_a^b=0, \tag{1.3.9}$$

in which v and ξ are any independent admissible functions. This implies the following principle for the functional $I(U, \Phi)$ defined by (1.3.5)

THEOREM 1.3.1. *The functional $I(U, \Phi)$ in (1.3.5) is stationary at u, ϕ where u, ϕ are solutions of the boundary-value problem*

$$\frac{d\phi}{dx}=\frac{\partial H}{\partial u}, \qquad a<x<b, \tag{1.3.10}$$

$$-\frac{du}{dx}=\frac{\partial H}{\partial \phi}, \qquad a<x<b, \tag{1.3.11}$$

$$\phi=\phi_B \quad on \quad \partial[a, b]. \tag{1.3.12}$$

This result tells us the precise form of the boundary conditions associated with the functional (1.3.5). They arise naturally and in this case are of Dirichlet type. Other boundary conditions can also be included in the theory and we shall deal with them later on.

As we saw in section 1.2, in certain cases a bound $J(\Phi)$ can be obtained for $J(\phi)$. In the present section the admission of variations in U which are independent of those in Φ gives an extra degree of freedom to the action functional and suggests the possibility of a bound on $I(u, \phi)$ different from that given by $J(\Phi)=I(\partial L/\partial \Phi', \Phi)$. Before developing this idea for the functional in (1.3.5), we shall examine a special case corresponding to the example of section 1.2.

Example 1. Consider

$$L=\tfrac{1}{2}p(\Phi')^2+\tfrac{1}{2}w\Phi^2-q\Phi, \tag{1.3.13}$$

where as before p, w, and q may be functions of x. By (1.3.1) the conjugate variable U is given by

$$U=\partial L/\partial \Phi'=p\Phi', \tag{1.3.14}$$

and the associated Hamiltonian H in (1.3.2) is

$$H(x, U, \Phi)=\frac{1}{2p}U^2-\tfrac{1}{2}w\Phi^2+q\Phi. \tag{1.3.15}$$

The functional $I(U, \Phi)$ for this example is then

$$I(U, \Phi)=\int_a^b\left\{U\Phi'-\frac{1}{2p}U^2+\tfrac{1}{2}w\Phi^2-q\Phi\right\}dx-[U(\Phi-\phi_B)]_a^b, \tag{1.3.16a}$$

or

$$I(U, \Phi)=\int_a^b\left\{-U'\Phi-\frac{1}{2p}U^2+\tfrac{1}{2}w\Phi^2-q\Phi\right\}dx+[U\phi_B]_a^b, \tag{1.3.16b}$$

on integrating by parts. By theorem 1.3.1, the action $I(U, \Phi)$ is stationary at u, ϕ where these are solutions of the canonical equations

$$\frac{d\Phi}{dx}=\frac{\partial H}{\partial U}=\frac{1}{p}U, \tag{1.3.17}$$

$$-\frac{dU}{dx}=\frac{\partial H}{\partial \Phi}=-w\Phi+q, \qquad a<x<b, \tag{1.3.18}$$

$$\Phi=\phi_B \quad \text{on} \quad \partial[a, b]. \tag{1.3.19}$$

Elimination of U here brings us back to the Euler–Lagrange equation (1.2.21). Further, we notice that if the first canonical equation (1.3.17) holds for arbitrary U and Φ, then expression (1.3.16a) reduces to the Euler–Lagrange form $J(\Phi)$ of the action integral, that is

$$I(p\Phi', \Phi)=J(\Phi). \tag{1.3.20}$$

This functional is then stationary at $\Phi=\phi$ provided that the second canonical equation (1.3.18) also holds at (u, ϕ). We have seen in section 1.2 that $J(\Phi)=I(p\Phi', \Phi)$ can provide a bound on $J(\phi)=I(p\phi', \phi)=I(u, \phi)$, and so it is natural also to investigate the behaviour of $I(U, \Phi)$ if Φ is determined in terms of U from the second canonical equation (1.3.18).

Suppose then that we choose any admissible function U and determine $\Phi=\Phi(U)$ by making the second canonical equation (1.3.18) hold identically. This gives

$$\Phi=\Phi(U)=\frac{1}{w}(U'+q), \tag{1.3.21}$$

w being non-zero here. Now (1.3.21) and (1.3.16b) specify a form of the action functional dependent on U only, which we shall call $G(U)$. Thus

$$\begin{aligned} G(U)&=I(U, \Phi(U)) \\ &=-\frac{1}{2}\int_a^b\left\{\frac{1}{p}U^2+\frac{1}{w}(q+U')^2\right\}dx+[U\phi_B]_a^b. \end{aligned} \tag{1.3.22}$$

This functional is *complementary* to $J(\Phi)$ above, in the sense that each can be obtained from $I(U, \Phi)$ by eliminating U or Φ with the help of one or other of the canonical equations assumed to hold identically.

From its construction $G(U)$ is stationary at $U=u$ and $G(u)=I(u, \phi)$. To investigate possible bounds let

$$U=u+\varepsilon v.$$

Then we find that

$$G(u)-G(U)=\frac{1}{2}\varepsilon^2\int_a^b\left\{\frac{1}{p}v^2+\frac{1}{w}v'^2\right\}\mathrm{d}x, \tag{1.3.23}$$

there being no terms higher than the second order since H is quadratic here. It follows that, if $p>0$ and $w>0$, the expression in (1.3.23) is non-negative and hence we have the maximum principle

$$G(U)\leqslant G(u)=I(u,\phi) \tag{1.3.24}$$

for all admissible functions U.

We saw earlier in (1.2.25) that if $p>0$ and $w\geqslant 0$, the Euler–Lagrange functional $J(\Phi)$ in (1.2.23), or equivalently in (1.3.20), satisfies the minimum principle

$$I(u,\phi)=J(\phi)\leqslant J(\Phi) \tag{1.3.25}$$

for all admissible functions Φ such that $\Phi=\phi_B$ on $\partial[a,b]$.

Combining the results (1.3.24) and (1.3.25), we see that for this example we have established complementary upper and lower bounds

$$G(U)\leqslant I(u,\phi)\leqslant J(\Phi), \tag{1.3.26}$$

given that $p>0$ and $w>0$.

The stationary properties of $J(\Phi)$ and $G(U)$ are called complementary variational principles, and in (1.3.26) we have an illustration of the case when these can be strengthened into complementary extremum principles. Thus, in this example, the variational principles provide upper and lower bounds for the solution $I(u,\phi)$ of the variational problem. We note that

$$I(u,\phi)=-\frac{1}{2}\int_a^b q\phi\,\mathrm{d}x+\tfrac{1}{2}[p\phi\phi']_a^b, \tag{1.3.27}$$

which follows from (1.3.16). Apart from the boundary term in (1.3.27), we see that upper and lower bounds have been obtained for a certain weighted average $\int q\phi\,\mathrm{d}x$ of the solution ϕ. In many applications the quantity $I(u,\phi)$ is related to the energy of the physical system under investigation, and bounds for it may be of considerable interest.

Example 2. In example 1 we assumed that w was non-zero in the derivation of the $G(U)$ bound. Let us now look at the case when $w=0$. It is readily seen that, apart from equations (1.3.21) and (1.3.22), all the results derived in example 1 hold with $w=0$. Consider then the modifications that are required to (1.3.21) and (1.3.22). These equations are concerned with the derivation of $G(U)$, which is obtained from $I(U,\Phi)$

by seeking $\Phi = \Phi(U)$ so that

$$-\frac{dU}{dx} = \frac{\partial H}{\partial \Phi} \tag{1.3.28}$$

holds identically. But (1.3.28) in this case reads

$$-\frac{dU}{dx} = q \quad \text{in} \quad a < x < b, \tag{1.3.29}$$

and this cannot be solved for $\Phi(U)$. Instead it represents a *constraint* on the trial function U. Setting such a trial U in (1.3.16b) with $w = 0$ gives

$$G(U) = -\frac{1}{2}\int_a^b \frac{1}{p} U^2 \, dx + [U\phi_B]_a^b. \tag{1.3.30}$$

Equations (1.3.29) and (1.3.30) are thus the modified forms of (1.3.21) and (1.3.22) when $w = 0$. We have included this example at this stage as it shows in a simple way how the canonical approach leads to constraints in an automatic fashion.

We now wish to extend the results of these examples so as to include the more general action functional (1.3.5). Such an extension involves the idea of convex functions which we shall briefly review first.

1.4 Convex functions

The definition of a convex function is based on the idea of 'chord above arc'. Thus we have

Definition 1.4.1. A function $f(x) \in \mathbb{R}$ is *convex* on (a, b) if

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \tag{1.4.1}$$

for $0 < \lambda < 1$ and any x_1, x_2 in (a, b). A function f is strictly convex if the inequality here is strict for distinct x_1 and x_2. A function f is *concave* if $-f$ is convex.

When f is differentiable there is an equivalent 'arc above tangent' statement which is also useful. We give this in

LEMMA 1.4.1. *If f is differentiable in (a, b), the following are equivalent statements:*

(i) *f is convex on (a, b);*
(ii) $f(x_1) - f(x_2) - (x_1 - x_2) f'(x_2) \geq 0, \qquad x_1, x_2 \in (a, b).$ (1.4.2)

Proof
(a) We first prove that (1.4.1) implies (1.4.2). From (1.4.1) we have

$$f(x_1) - f(x_2) - \lambda^{-1}\{f(x_2 + \lambda(x_1 - x_2)) - f(x_2)\} \geq 0.$$