A Duality Principle for Non-Convex Optimisation and the Calculus of Variations

J. F. TOLAND

Communicated by C. DAFERMOS

Contents

Abstract

This paper analyses a notion of duality which is defined for a class of non-convex problems which arise in the calculus of variations. The duality principle which is introduced in Section 2.1 is motivated by a specific problem, namely the mathematical description of the rotating heavy chain. Nevertheless the theory applies in many other situations as well.

In the course of the analysis a notion of critical point is introduced, generalising the usual definition of the critical point of a nonlinear functional, and it is found that there is a duality principle for the critical points as well as for the extremals of the functional question.

The duality theory is then applied to explain why there are two distinct variational formulations of the steadily rotating heavy chain problem.

1.1. Introduction

The results of this paper were motivated by some nonlinear problems in the calculus of variations which arise in mechanics. It is a familiar situation to find that the equilibrium states of a mechanical system can be described by the elements of a function space which minimise a given nonlinear functional among a class of admissible functions. There are however some systems, such as the heavy nonlinear rotating chain which is treated in section 4, which can equally well be described by two different variational principles. It might appear that the existence of two variational descriptions of the spinning chain is characteristic of a small class of variational problems. This turns out not to be the case. There is a clear sense in which the two formulations are in duality.

Since neither of the functionals which arise in the theory of the spinning chain is convex, they cannot be in duality in the sense of ROCKAFELLAR ([5, 6, 7, 8] and the references cited therein) and FENCHEL [3]. An exposition of these theories is given in the book of EKELAND & TEMAM [2].

The ideas of these authors are exploited to the full in what follows but the duality results are more obvious than those cited above.

In a recent paper COFFMAN [1] observed that the equations which are usually used to describe the motion of a steadily rotating, heavy chain are the Euler-Lagrange equations corresponding to a functional E_2 defined in Section 4. It is remarkable that this functional is not the energy functional of the string; the energy functional gives rise to a different equation which also describes the motion of the chain. The duality theory of section 2 helps to explain, from purely variational considerations, the connection between these two distinct mathematical formulations of the problem.

The two variational formulations of the chain problem have in common the feature that each involves a functional which can be written as the difference of two convex functionals. The clue to the inter-relationship between the two formulations is contained in the following observation: If F and G are lower semi-continuous convex functionals on a topological vector space V then

$$\inf_{u \in V} \{G(u) - F(u)\} = \inf_{u^* \in V^*} \{F^*(u^*) - G^*(u^*)\}$$

where V^* is a space in duality with V and F^* and G^* are the polar functions of F and G, respectively.

Section 2 begins with a proof of a general form of this result which does not require the convexity of G (Theorem 2.2). The result holds under special circumstances even when neither F nor G is convex (Theorem 2.3).

An important question in any consideration of duality is the following: "Does existence of a solution of the dual problem imply the existence of a solution of the original problem?". This question is treated in Theorem 2.4 and its corollaries, where some positive results in this direction are established. Section 2 concludes with some results on minimising sequences for $G-F$ and $F^* - G^*$, and the possible relationship between them.

In section 3.1 a special case of the duality theory of section 2 is treated. This special case is modelled on a class of non-convex problems typical of those arising in the calculus of variations. In this case the dual principle is formulated, not on V^* as might be expected, but on an associated space Y^* as follows:

Suppose that V, V^*, Y, Y^* are two pairs of dual spaces, and that $\Lambda: V \to Y$ is a bijection with adjoint $\Lambda^*: Y^* \to V^*$. If G and F are convex functionals on Y and V, respectively, and F is lower semi-continuous then, as shown in Section 3.1,

$$\inf_{u \in V} \{G \circ \Lambda(u) - F(u)\} = \inf_{v^* \in Y^*} \{F^* \circ \Lambda^*(v^*) - G^*(v^*)\}.$$

It is this result that establishes the connection between the two variational formulations of the rotating chain problem.

In the course of section 3.1 it is found that if $\underline{u} \in V$ is a minimiser of $G \circ \Lambda - F$ and $\Lambda^* \underline{v}^* \in \partial F(\underline{u})$ (the subdifferential of F at $\underline{u}$), then $\underline{v}^*$ minimises $F^* \circ \Lambda^* - G^*$ and

$$F(\underline{u}) + F^*(\Lambda^* \underline{v}^*) = \langle \Lambda^* \underline{v}^*, \underline{u} \rangle$$
$$G^*(\underline{v}^*) + G \circ \Lambda(\underline{u}) = \langle \Lambda^* \underline{v}^*, \underline{u} \rangle.$$

These two equalities are sufficient to imply that $\underline{u}$ and $\underline{v}^*$ satisfy a weak form of the Euler-Lagrange equations for $G \circ \Lambda - F$ and $F^* \circ \Lambda^* - G^*$, respectively, and that the usual transversality conditions hold. Thus $\underline{u}$ and $\underline{v}^*$ are critical points of the respective functionals. However, unlike the extremal conditions which arise in the theory of convex optimisation ((4.22) and (4.23) of chapter III of [2]), the above equations do not guarantee that $\underline{u}$ and $\underline{v}^*$ are minimisers of $G \circ \Lambda - F$ and $F^* \circ \Lambda^* - G^*$.

In section 3.2 a notion of critical point of $G \circ \Lambda - F$ is introduced which generalises the classical definition in the case where $G \circ \Lambda$ and F are not necessarily differentiable functionals. A one-to-one correspondence between the critical points of $G \circ \Lambda - F$ and $F^* \circ \Lambda^* - G^*$ is established, and the Euler-Lagrange equations and the classical transversality conditions are seen to be necessary conditions for a function to be a critical point of $G \circ \Lambda - F$.

Finally, in section 4, the analysis of the steadily rotating heavy chain is carried out. The two variational formulations of the problem are seen to be dual. The existence of a minimiser for the functional E_2 is used to prove the existence of a minimiser for the energy functional (which is the dual of E_2), and a one-to-one correspondence between the critical points of the two functionals is established. From duality considerations alone we can infer a point-wise relationship between critical points of the respective functionals as well as that the Euler-Lagrange equations and natural transversality conditions hold.

Thus either of these two variational principles provides a complete description, not only of the stable equilibrium configurations of the chain (the minimisation problem), but also of the unstable configurations (the critical points).

1.2. Preliminaries

For the convenience of the reader we introduce in this section the terminology used consistently in the paper. We also record statements of important definitions and results (which will be used without further elaboration in later sections) found in the book of EKELAND & TEMAM [2].

We shall use $\mathbb{R}$ to denote the set of real numbers and $\overline{\mathbb{R}}$ for the extended set of real numbers, with the convention that $\infty - \infty = \infty + \infty = -\infty + \infty = \alpha + \infty = \infty$ for all $\alpha \in \mathbb{R}$.

Let V and V^* be two real vector spaces and let $\langle \ , \ \rangle$ be a bilinear form on the Cartesian product $V \times V^*$. We say that the bilinear form puts the spaces *in duality*. The duality defined by $\langle \ , \ \rangle$ is said to be *separating* if

(i) when $u(\neq 0) \in V$ there exists an element $u^* \in V^*$ such that $\langle u, u^* \rangle \neq 0$,

(ii) when $u^*(\neq 0) \in V^*$ there exists an element $u \in V$ such that $\langle u, u^* \rangle \neq 0$.

If V and V^* are topological vector spaces put in duality by the bilinear form $\langle \ , \ \rangle$ then we can define a linear functional f_{u^*} for each $u^* \in V^*$ by

$$f_{u^*}(u) = \langle u, u^* \rangle \qquad \text{for each } u \in V.$$

Then the *weak topology on V induced by $\langle \ , \ \rangle$* is the coarsest one for which all the linear functionals f_{u^*} are continuous. This topology will be denoted by $\sigma(V, V^*)$. The weak topology on V^* induced by $\langle \ , \ \rangle$ is defined analogously and will be denoted by $\sigma(V^*, V)$.

In order that $\sigma(V, V^*)$ and $\sigma(V^*, V)$ be Hausdorff topologies it is both necessary and sufficient that the duality between V and V^* be separating.

Throughout this paper we shall assume the spaces V and V^* to be in separating duality and to be endowed with the topologies $\sigma(V, V^*)$ and $\sigma(V^*, V)$. *All statements concerning continuity, lower semi-continuity, convergence, closure etc. will be understood with respect to these topologies.*

A functional $F: V \to \overline{\mathbb{R}}$ is said to be *lower semi-continuous* (l.s.c.) if

$$F(u) \leqq \liminf_{n \to \infty} F(u_n)$$

for each $u \in V$ and any sequence $u_n \to u$. A functional $F: V \to \overline{\mathbb{R}}$ is called *convex* if

$$F(\lambda u + (1-\lambda) w) \leqq \lambda F(u) + (1-\lambda) F(w)$$

for all $\lambda \in (0,1)$ and all $u, w \in V$. A functional is called *strictly convex* if it is convex and if the above inequality is strict for $u \neq w$ and $\lambda \in (0,1)$.

Proposition. *If a functional $V \to \overline{\mathbb{R}}$ is convex, lower semi-continuous, and takes the value $-\infty$, then it cannot take any finite value.*

A function $f: V \to \mathbb{R}$ is called *affine continuous on V* if $f(u) = l(u) + \alpha$, where l is a continuous linear functional on V and $\alpha \in \mathbb{R}$.

Proposition. If V and V^* are in separating duality, then all affine continuous functions on V are of the form $f(u) = \langle u, u^* \rangle + \alpha$ for some $u^* \in V^*$.

We shall denote by $\Gamma(V)$ the set of all functions $F: V \to \overline{\mathbb{R}}$ which are pointwise suprema of a family of affine continuous functions on V.

Proposition. *A functional F is in $\Gamma(V)$ if and only if it is convex and lower semi-continuous on V. Moreover if F takes the value $-\infty$, then it must be identically $-\infty$.*

If V and V^* are in separating duality and if $F: V \to \overline{\mathbb{R}}$ is arbitrary, then the *polar $F^*: V^* \to \overline{\mathbb{R}}$ of F* is defined by

$$F^*(u^*) = \sup_{u \in V} \{\langle u, u^* \rangle - F(u)\}.$$

Clearly $F^* \in \Gamma(V^*)$ for arbitrary $F: V \to \overline{\mathbb{R}}$. It is easy to see that if $F, G: V \to \overline{\mathbb{R}}$ and $F \leqq G$, then $G^* \leqq F^*$ and $(F + \alpha)^* = F^* - \alpha$ for all $\alpha \in \mathbb{R}$.

If $F: V \to \overline{\mathbb{R}}$ then $F^*: V^* \to \overline{\mathbb{R}}$. Consequently F^{**} maps V into $\overline{\mathbb{R}}$ and $F^{**} \in \Gamma(V)$. In fact F^{**} is the largest convex lower semi-continuous function on V which does not exceed F. Thus $F^{**}(u) \leqq F(u)$ for all $u \in V$ and $F(u) = F^{**}(u)$ for all $u \in V$, if and only if $F \in \Gamma(V)$.

A function $F: V \to \overline{\mathbb{R}}$ is said to be *subdifferentiable* at $\underline{u} \in V$ if $F(\underline{u})$ is finite and there exists an element $u^* \in V^*$ such that

$$F(\underline{u}) + \langle u - \underline{u}, u^* \rangle \leqq F(u)$$

for all $u \in V$. The set of all u^* with this property is denoted by $\partial F(\underline{u})$. The following observation is crucial in the subsequent analysis.

Theorem 1.1. *Let $F: V \to \overline{\mathbb{R}}$ and let F^* be its polar. Then $u^* \in \partial F(u)$ if and only if*

$$F(u) + F^*(u^*) = \langle u, u^* \rangle.$$

Furthermore $u^* \in \partial F(u)$ implies $u \in \partial F^*(u^*)$.

If in addition $F \in \Gamma(V)$ then $u^* \in \partial F(u)$ if and only if $u \in \partial F^*(u^*)$.

If $F: V \to \overline{\mathbb{R}}$ and $\lim_{\lambda \to 0} \{[F(u + \lambda v) - F(u)]/\lambda\} = \langle v, u^* \rangle$ for some $u^* \in V^*$, then F is said to be *Gateaux differentiable* at u and we write $F'(u) = u^* \in V^*$.

Proposition. *If $F: V \to \overline{\mathbb{R}}$ is convex and Gateaux differentiable at $u \in V$, then F is subdifferentiable at $u \in V$, and $\partial F(u) = \{F'(u)\}$.*

In the application of Theorem 1.1 it is often necessary to calculate the dual of a functional which takes the form of a nonlinear integral on a function space. The next result is central in the application to be carried out in section 4.

Let Ω be an open subset of $\mathbb{R}^n$ and let g be a mapping from $\Omega \times \mathbb{R}^m$ to $\mathbb{R}$ which satisfies the Carathéodory condition (i.e. for all $y \in \mathbb{R}^m$, $x \to g(x, y)$ is measurable, and for almost all $x \in \Omega$, $y \to g(x, y)$ is continuous).

Proposition. *Let $V = L^{\alpha_1}(\Omega) \times \ldots \times L^{\alpha_m}(\Omega)$ where $1 \leqq \alpha_i < \infty$. Suppose that for all $u \in V$, the function $x \to g(x, u(x))$ is integrable. Then the function $u \to g(\cdot, u(\cdot))$ is continuous from V to L^1.*

This allows us to define $G: V \to \mathbb{R}$ by

$$G(u) = \int_\Omega g(x, u(x))\, dx.$$

Now let

$$V^* = L^{\alpha'_1} \times \ldots \times L^{\alpha'_m}, \quad \frac{1}{\alpha_i} + \frac{1}{\alpha'_i} = 1,$$

and let $\langle\ ,\ \rangle$ denote the usual duality between V and V^*. Then we get

Theorem 1.2. *Under the hypotheses above, for any $u^* \in V^*$ we have*

$$\begin{aligned} G^*(u^*) &= \sup_{u \in V} \{\langle u, u^* \rangle - G(u)\} \\ &= \sup_{u \in V} \{\int_\Omega u(x)\, u^*(x) - g(x, u(x))\, dx\} \\ &= \int_\Omega g^*(x, u^*(x))\, dx, \end{aligned}$$

where

$$g^*(x, y) = \sup_{\eta \in \mathbb{R}^m} \{y\eta - g(x, \eta)\}$$

for almost all $x \in \Omega$.

2.1. Duality Theory

Let V and V^* be linear spaces in duality, and let $\langle\ ,\ \rangle$: $V \times V^* \to \mathbb{R}$ denote the corresponding bilinear form (which is compatible with the topologies on V and V^*). If F: $V \to \overline{\mathbb{R}}$ and G: $V \to \overline{\mathbb{R}}$ we may define a functional J: $V \to \overline{\mathbb{R}}$ by

$$J(u) = G(u) - F(u)$$

for all $u \in V$. We shall denote by $\mathscr{P}$ the problem of evaluating

$$\inf_{u \in V} J(u),$$

and we shall call an element $\underline{u} \in V$ a solution of $\mathscr{P}$ if $J(\underline{u})$ is finite and

$$\inf_{u \in V} J(u) = J(\underline{u}).$$

As in the preceding section we let F^* and G^* denote the polar functionals of F and G.

Theorem 2.1. *Suppose that* $J(u) \geqq \alpha$ *for all* $u \in V$. *Then* $F^*(u^*) - G^*(u^*) \geqq \alpha$ *for all* $u^* \in V^*$.

Proof. Since $G(u) - F(u) \geqq \alpha$ for all $u \in V$, it follows that $G(u) \geqq \alpha + F(u)$ for all $u \in V$. Consequently $G^*(u^*) \leqq F^*(u^*) - \alpha$ for all $u^* \in V^*$, which in turn implies that

$$F^*(u^*) - G^*(u^*) \geqq \alpha$$

for all $u^* \in V^*$. This completes the proof of the theorem.

We define a functional $\hat{J}$: $V^* \to \overline{\mathbb{R}}$ by

$$\hat{J}(u^*) = F^*(u^*) - G^*(u^*).$$

Then Theorem 2.1 asserts that

$$\inf_{u \in V} J(u) \leqq \inf_{u^* \in V^*} \hat{J}(u^*). \tag{2.1}$$

We denote by $\mathscr{P}^*$ the problem of evaluating

$$\inf_{u^* \in V^*} \hat{J}(u^*).$$

An element $\underline{u}^* \in V^*$ will be called a solution of $\mathscr{P}^*$ if $\hat{J}(\underline{u}^*)$ is finite and

$$\inf_{u^* \in V^*} \hat{J}(u^*) = \hat{J}(\underline{u}^*).$$

We begin by showing that equality holds in (2.1) provided that F is convex and lower semi-continuous.

Theorem 2.2. *If $F \in \Gamma(V)$ then*

$$\inf_{u \in V} J(u) = \inf_{u^* \in V^*} \hat{J}(u^*). \tag{2.2}$$

Proof. Suppose that $F^*(u^*) - G^*(u^*) \geqq \alpha$ for all $u^* \in V^*$. Then $F^*(u^*) \geqq G^*(u^*) + \alpha$ for all $u^* \in V^*$, and so

$$F(u) = F^{**}(u) \leqq G^{**}(u) - \alpha \leqq G(u) - \alpha$$

for all $u \in V$. Hence $G(u) - F(u) \geqq \alpha$. Thus $\inf_{u \in V} J(u) = +\infty$ if and only if $\inf_{u^* \in V^*} \hat{J}(u^*) = +\infty$, and $\inf_{u \in V} J(u) > -\infty$ if and only if $\inf_{u^* \in V^*} \hat{J}(u^*) > -\infty$. In the latter case we have established equality, and so (2.2) holds.

The next result is similar in spirit to Theorem 2.2, but does not require convexity of F or G.

Theorem 2.3. *Suppose that $\underline{u} \in V$ is a solution of $\mathscr{P}$ and that $\partial F(\underline{u}) \neq \phi$. Then* (2.2) *holds and u^* solves $\mathscr{P}^*$ provided $u^* \in \partial F(\underline{u})$.*

Proof. Since $\underline{u}$ solves $\mathscr{P}$, the quantity $G(\underline{u}) - F(\underline{u})$ is finite and

$$G(\underline{u}) - F(\underline{u}) \leqq G(u) - F(u) \qquad \text{for all } u \in V.$$

Then $F(u) - F(\underline{u}) \leqq G(u) - G(\underline{u})$. Since $u^* \in \partial F(\underline{u})$ we have

$$\langle u - \underline{u}, u^* \rangle + F(\underline{u}) \leqq F(u),$$

which now implies that $u^* \in \partial G(\underline{u})$. Hence

$$G(\underline{u}) + G^*(u^*) = \langle \underline{u}, u^* \rangle$$
$$F(\underline{u}) + F^*(u^*) = \langle \underline{u}, u^* \rangle.$$

By subtraction it is clear that (2.2) holds and that u^* solves $\mathscr{P}^*$.

It is worth noting that if F is everywhere subdifferentiable then $F \in \Gamma(V)$. Hence Theorem 2.2 may be used to infer (2.2) in this case. Theorem 2.3, however, is a stronger statement since it implies the following

Corollary. *If F is everywhere subdifferentiable and if $\min_{u \in V} J(u)$ exists and is finite, then $\min_{u^* \in V^*} \hat{J}(u^*)$ exists and*

$$\min_{u \in V} J(u) = \min_{u^* \in V^*} \hat{J}(u^*).$$

The next theorem gives conditions under which

$$\inf_{u \in V} J(u) = \min_{u^* \in V^*} \hat{J}(u^*).$$

Theorem 2.4. *Suppose $F \in \Gamma(V)$, and let $\{u_n\}$ be a minimising sequence for problem $\mathscr{P}$. Then if $\inf_{u \in V} J(u)$ is finite and*

$$\lim_{n\to\infty}\{\langle u_n, \underline{u}^*\rangle - F(u_n)\} = F^*(\underline{u}^*) < \infty$$

for some $\underline{u}^* \in V^*$, it follows that $\underline{u}^*$ is a solution of problem $\mathscr{P}^*$ and

$$\lim_{n\to\infty}\{\langle u_n, \underline{u}^*\rangle - G(u_n)\} = G^*(\underline{u}^*).$$

Conversely, if $\underline{u}^*$ is a solution of $\mathscr{P}^*$ and there is a sequence $\{u_n\}$ such that

$$\lim_{n\to\infty}\{\langle u_n, \underline{u}^*\rangle - G(u_n)\} = G^*(\underline{u}^*),$$

then

$$\inf_{u\in V} J(u) \quad \text{is finite},$$

$$\lim_{n\to\infty}\{\langle u_n, \underline{u}^*\rangle - F(u_n)\} = F^*(\underline{u}^*) < \infty,$$

and $\{u_n\}$ is a minimising sequence for $\mathscr{P}$.

Proof. Suppose first that $\underline{u}^*$ solves $\mathscr{P}^*$ and that $\lim\{\langle u_n, \underline{u}^*\rangle - G(u_n)\} = G^*(\underline{u}^*)$. Since $F \in \Gamma(V)$ it follows that

$$\inf_{u\in V} J(u) = \inf_{u^*\in V^*} \hat{J}(u^*) = \alpha,$$

where α is real. Since $F^*(\underline{u}^*) - \alpha = G^*(\underline{u}^*)$, the assumptions imply that

$$\lim\{\langle u_n, \underline{u}^*\rangle - G(u_n)\} + \alpha = F^*(\underline{u}^*).$$

By definition, $G(u_n) - F(u_n) \geqq \alpha$ for all n, and so

$$\liminf\{\langle u_n, \underline{u}^*\rangle - F(u_n)\} \geqq F^*(\underline{u}^*).$$

It follows from the definition of $F^*(\underline{u}^*)$ that

$$\lim_{n\to\infty}\{\langle u_n, \underline{u}^*\rangle - F(u_n)\} = F^*(\underline{u}^*).$$

$F^*(\underline{u}^*)$ is clearly finite (since $F^*(\underline{u}^*) - G^*(\underline{u}^*) = \alpha \in \mathbb{R}$), whence

$$G(u_n) - F(u_n) \to \alpha \quad \text{as } n \to \infty.$$

Conversely, suppose that $\{u_n\}$ is a minimising sequence for $\mathscr{P}$ and that

$$\lim\{\langle u_n, \underline{u}^*\rangle - F(u_n)\} = F^*(\underline{u}^*) < \infty.$$

Since $\lim\{G(u_n) - F(u_n)\} = \alpha \in \mathbb{R}$ it follows that

$$\lim\{\langle u_n, \underline{u}^*\rangle - G(u_n)\} = F^*(\underline{u}^*) - \alpha \geqq G^*(\underline{u}^*).$$

Hence $\lim\{\langle u_n, \underline{u}^*\rangle - G(u_n)\} = G^*(\underline{u}^*)$ and $\underline{u}^*$ is a solution of $\mathscr{P}^*$.

Corollary 2.5. *Suppose that $F \in \Gamma(V)$, that G is lower semi-continuous and that $\underline{u}^*$ solves $\mathscr{P}^*$. If there exists a sequence $\{u_n\}$, with $u_n \to \underline{u}$ in the topology $\sigma(V, V^*)$,* *and if*

$$\lim_{n\to\infty}\{\langle u_n, \underline{u}^*\rangle - G(u_n)\} = G^*(\underline{u}^*),$$

then $\underline{u}$ *is a solution of* $\mathscr{P}$.

Proof. Since $u_n \to \underline{u}$, we have $\lim\langle u_n, \underline{u}^*\rangle = \langle \underline{u}, \underline{u}^*\rangle$; thus the lower semi-continuity of G implies that

$$\lim G(u_n) = -G^*(\underline{u}^*) + \langle \underline{u}, \underline{u}^*\rangle \leqq G(\underline{u}) \leqq \liminf G(u_n).$$

Hence $G(u_n)\to G(\underline{u})$. Similarly, $F(u_n)\to F(\underline{u})$ and $\underline{u}$ is a solution of $\mathscr{P}$ since, by Theorem 2.4, $\{u_n\}$ is a minimising sequence for $\mathscr{P}$.

Corollary 2.6. *Under the hypotheses of Corollary* 2.5 *we have*

$$F(\underline{u}) + F^*(\underline{u}^*) = \langle \underline{u}, \underline{u}^*\rangle$$

and

$$G(\underline{u}) + G^*(\underline{u}^*) = \langle \underline{u}, \underline{u}^*\rangle.$$

Proof. Since $F(u_n)\to F(\underline{u})$, $G(u_n)\to G(\underline{u})$, the result follows immediately from the proof of Corollary 2.5. Thus we have established a criterion which ensures that if $\mathscr{P}$ is soluble and finite then so is $\mathscr{P}^*$.

Theorem 2.7. *Suppose that* V *is a reflexive Banach space, that* $F\in\Gamma(V)$, *and that* G *is lower semi-continuous. If* $G(u)/\|u\|\to\infty$ *as* $\|u\|\to\infty$ *and* $\underline{u}^*$ *solves* $\mathscr{P}^*$, *then there exists a solution* $\underline{u}$ *of* $\mathscr{P}$ *and*

$$F(\underline{u}) + F^*(\underline{u}^*) = \langle \underline{u}, \underline{u}^*\rangle$$
$$G(\underline{u}) + G^*(\underline{u}^*) = \langle \underline{u}, \underline{u}^*\rangle.$$

Proof. Since $F\in\Gamma(V)$, we have

$$\inf_{u\in V} J(u) = \inf_{u^*\in V^*}\{F^*(u^*) - G^*(u^*)\}\in\mathbb{R}.$$

Let $\{u_n\}$ be any sequence such that

$$\langle u_n, \underline{u}^*\rangle - G(u_n)\to G^*(\underline{u}^*) \quad \text{as } n\to\infty.$$

Then Theorem 2.4 ensures that $\{u_n\}$ is a minimising sequence for $\mathscr{P}$. By hypothesis, $\{u_n\}$ is bounded and so the reflexivity of V implies that $\{u_n\}$ has a weakly convergent subsequence. Then Corollaries 2.5 and 2.6 are applicable, and the proof is complete.

If G is not coercive $(G(u)/\|u\|\nrightarrow\infty$ as $\|u\|\to\infty)$, then there exists at least one $w^*\in V^*$ such that $w^*\notin\partial G(u)$ for any $u\in V$. If $F(u)=\langle u, w^*\rangle$ for all $u\in V$, it is clear that $F^*(u^*)=\infty$ if $u^*\neq w^*$, and $F^*(w^*)=0$.

Hence $w^*\in V^*$ is a solution of $\mathscr{P}^*$ if $G^*(w^*)$ is finite, yet there may be no solution of $\mathscr{P}$. In some cases, however, the behaviour of F may compensate for a lack of coerciveness of G.

Theorem 2.8. *If the coerciveness assumption on G in Theorem 2.7 is replaced by*

$$G(u)-F(u)\to\infty \qquad as\ \|u\|\to\infty,$$

then the same conclusion holds.

Proof. Since $\inf_{u\in V} J(u)$ is finite and $\{u_n\}$ is a minimising sequence for $\mathscr{P}$, it follows that $\{u_n\}$ is bounded, and in turn that it has a weakly convergent subsequence. The proof then proceeds as before.

The previous two results guarantee the existence of solutions of $\mathscr{P}$ given the existence of solutions $\mathscr{P}^*$. Note that in Theorems 2.7 and 2.8 we do not require the lower semi-continuity of $G-F$, as we would if we were attempting to show directly that $\mathscr{P}$ is soluble. Thus we are led to the problem of proving the existence of solutions of $\mathscr{P}^*$.

Theorem 2.9. *Let V be a reflexive Banach space and V^* its dual. Suppose that G^* has the property that if $u_n^*\to u^*$ in $\sigma(V^*, V)$ then there exists a subsequence $\{u_{n_j}^*\}$ with $G^*(u_{n_j}^*)\to G^*(u^*)$ as $j\to\infty$. If $\hat{J}(u^*)\to\infty$ as $\|u^*\|\to\infty$ and $\hat{J}$ is bounded below, then $\mathscr{P}^*$ has a solution.*

Proof. Let $\{u_n^*\}$ be a minimising sequence for $\mathscr{P}^*$. Then $\{u_n^*\}$ is bounded (except in the trivial case when $\hat{J}\equiv\infty$). Since V is reflexive we can assume that $u_n^*\to u^*$ in $\sigma(V^*, V)$ for some $u^*\in V^*$. Thus there exists a subsequence $\{u_{n_j}^*\}$ with $G^*(u_{n_j}^*)\to G^*(u^*)$. The lower semicontinuity of F^* implies that

$$F^*(u^*)-G^*(u^*)\leqq \liminf\{F^*(u_{n_j}^*)-G^*(u_{n_j}^*)\}.$$

Since $u_{n_j}^*$ is a minimising sequence for $\mathscr{P}^*$, the proof of the theorem is complete.

2.2. On Minimising Sequences

Theorem 2.10. *Suppose that F is everywhere subdifferentiable and that $\{u_n\}$ is a minimising sequence for $\mathscr{P}$. If $u_n^*\in\partial F(u_n)$, then $\{u_n^*\}$ is a minimising sequence for $\mathscr{P}^*$.*

Proof. If $\inf_{u\in V} J(u)=\infty$ then the result is immediate. If $\inf_{u\in V} J(u)=\alpha<\infty$, let $\{u_n\}$ be a minimising sequence. Since $u_n^*\in\partial F(u_n)$, we have

$$F(u_n)+F^*(u_n^*)=\langle u_n, u_n^*\rangle$$

and

$$\lim_{n\to\infty}\{G(u_n)-F(u_n)\}=\alpha.$$

Hence

$$G(u_n)-F(u_n)=G(u_n)-\langle u_n, u_n^*\rangle+F^*(u_n^*)$$
$$\geqq F^*(u_n^*)-G^*(u_n^*).$$

Consequently $\{u_n^*\}$ is a minimising sequence for $\mathscr{P}^*$. If $\lim_{n\to\infty} J(u_n)=-\infty$, then the above argument goes through for any $\alpha\in\mathbb{R}$, and so

$$\lim_{n\to\infty}\{F^*(u_n^*)-G^*(u_n^*)\}=-\infty.$$

The proof of the theorem is complete.

Theorem 2.11. *Let V be a Banach space, and suppose that $F(\not\equiv\pm\infty)$ is convex and lower semi-continuous on V. Let $\{u_n\}$ be a minimising sequence for $\mathscr{P}$ with $\inf_{u\in V} J(u)\in\mathbb{R}$. Then there exists a sequence $\{v_n\}$ such that*

$$\partial F(v_n)\neq\phi \quad \textit{and} \quad \|v_n-u_n\|\leqq\frac{1}{2^n}.$$

Furthermore the sequence $\{v_n^*\}$, $v_n^*\in\partial F(v_n)$, is a minimising sequence for $\mathscr{P}^*$.

Proof. For each n, $F(u_n)=F^{**}(u_n)$; hence

$$F(u_n)=\sup_{u^*\in V^*}\{\langle u_n, u^*\rangle-F^*(u^*)\}.$$

Thus for each n there exists an element $u_n^*\in V^*$ with

$$\langle u_n, u_n^*\rangle-F^*(u_n^*)\geqq F(u_n)-\frac{1}{4^n}.$$

Then, by Theorem 6.2 of Chapter I of [2], there exist elements v_k, v_k^* such that

$$\|u_n-v_k\|\leqq\frac{1}{2^k(1+\|u_n^*\|)},\quad \|u_n^*-v_k^*\|\leqq\frac{(1+\|u_n^*\|)\,2^k}{4^n},\quad v_{k_n}^*\in\partial F(v_{k_n}).$$

Choose k_n so large that $F(v_{k_n})\geqq F(u_n)-1/2^n$, $k_n\geqq n$ (this is possible since F is lower semi-continuous).

Put

$$G(u_n)-F(u_n)=\alpha+\varepsilon_n,\qquad \text{where } \inf_{u\in V} J(u)=\alpha.$$

Since $F(v_{k_n})+F^*(v_{k_n}^*)=\langle v_{k_n}, v_{k_n}^*\rangle$ and $F(v_{k_n})\geqq F(u_n)-1/2^n$, we must have

$$F(u_n)+F^*(v_{k_n}^*)\leqq\langle v_{k_n}, v_{k_n}^*\rangle+\frac{1}{2^n}.$$

Therefore

$$G(u_n)+F^*(v_{k_n}^*)\leqq\alpha+\langle v_{k_n}, v_{k_n}^*\rangle+\frac{1}{2^n}+\varepsilon_n,$$

and so

$$F^*(v_{k_n}^*)+G(u_n)-\langle v_{k_n}, v_{k_n}^*\rangle \leqq \alpha+\frac{1}{2^n}+\varepsilon_n.$$

Hence

$$F^*(v_{k_n}^*)-G^*(v_{k_n}^*)\leqq \alpha+\frac{1}{2^n}+\varepsilon_n+\langle v_{k_n}-u_n, v_{k_n}^*\rangle \leqq \alpha+\frac{1}{2^{n-1}}+\varepsilon_n.$$

Putting $v_n=v_{k_n}$ and $v_n^*=v_{k_n}^*$, we arrive at the assertion of the theorem.

Corollary. *If G is continuous in the norm topology on V then $\{v_n\}$ in Theorem* 2.10 *can be chosen to be a minimising sequence.*

Proof. If G is continuous in the norm topology, then v_n may be chosen so that

$$|G(v_n)-G(u_n)|\leqq \frac{1}{2^n},$$

and at the same time so that

$$F(v_n)\geqq F(u_n)-\frac{1}{2^n}.$$

Hence

$$\begin{aligned} G(v_n)-F(v_n) &\leqq |G(v_n)-G(u_n)|+G(u_n)-F(v_n) \\ &\leqq \frac{1}{2^n}+G(u_n)-F(u_n)+\frac{1}{2^n}. \end{aligned}$$

The proof of the corollary is complete.

3.1. An Important Special Case: the Calculus of Variations

In this section we consider a special case which is of considerable importance in applications.

Let Y and V be locally convex spaces, and let Y^* and V^* be in duality with Y and V, respectively. Without ambiguity we shall use $\langle\ ,\ \rangle$ to indicate the duality between Y and Y^* and between V and V^*. Suppose that Λ is a linear homeomorphism from V to Y, with adjoint Λ^*, and that $F\colon V\to\overline{\mathbb{R}}$ and $G\colon Y\to\overline{\mathbb{R}}$ are convex functionals.

Put $J(u)=G\circ\Lambda(u)-F(u)$ for all $u\in V$. As before; we denote by $\mathscr{P}$ the problem of evaluating

$$\inf_{u\in V} J(u),$$

and an element $\underline{u}$ of V is called a solution of $\mathscr{P}$ if

$$G\circ\Lambda(\underline{u})-F(\underline{u})=\inf_{u\in V} J(u)\in\mathbb{R}.$$

If $F\in\Gamma(V)$, we know from Theorem 2.2 that

$$\inf_{u\in V} J(u)=\inf_{u^*\in V^*}\{F^*(u^*)-(G\circ\Lambda)^*(u^*)\}.$$

In this case, however, we shall introduce a dual problem different from the usual one considered in section 2.

Since $\Lambda: V \to Y$ is a linear homeomorphism, so is $\Lambda^*: Y^* \to V^*$. Now for each $u^* \in V^*$ there exists an element $v^* \in Y^*$ with $u^* = \Lambda^* v^*$. Hence

$$\begin{aligned}(G \circ \Lambda)^*(u^*) &= (G \circ \Lambda)^*(\Lambda^* v^*)\\ &= \sup_{u \in V} \{\langle u, \Lambda^* v^*\rangle - (G \circ \Lambda)(u)\}\\ &= \sup_{u \in V} \{\langle \Lambda u, v^*\rangle - G(\Lambda u)\}\\ &= \sup_{v \in Y} \{\langle v, v^*\rangle - G(v)\} = G^*(v^*).\end{aligned}$$

Hence, by Theorem 2.2 and the above remarks, if $F \in \Gamma(V)$ then

$$\inf_{v^* \in Y^*} \{F^* \circ \Lambda^*(v^*) - G^*(v^*)\} = \inf_{u \in V} \{G \circ \Lambda(u) - F(u)\}.$$

In this section problem $\mathscr{P}^*$ is the problem of evaluating

$$\inf_{v^* \in Y^*} \hat{J}(v^*) = \inf_{v^* \in Y^*} \{F^* \circ \Lambda^*(v^*) - G^*(v^*)\},$$

and a point $\underline{v}^* \in Y^*$ will be called a solution of $\mathscr{P}^*$ if

$$\hat{J}(\underline{v}^*) = \inf_{v^* \in Y^*} \hat{J}(v^*) \in \mathbb{R}.$$

Now if $\underline{u}$ is a solution of $\mathscr{P}$ and $\Lambda^* \underline{v}^* \in \partial F(\underline{u})$, then $\underline{v}^*$ is a solution of $\mathscr{P}^*$. Furthermore

$$F(\underline{u}) + F^* \circ \Lambda^*(\underline{v}^*) = \langle \underline{u}, \Lambda^* \underline{v}^*\rangle$$

and

$$G^*(\underline{v}^*) + G \circ \Lambda(\underline{u}) = \langle \underline{u}, \Lambda^* \underline{v}^*\rangle.$$

These two conditions should be compared with the so called extremality conditions (4.22), (4.23) of Chapter III of [2]. They play a crucial role in the analysis of the spinning chain, which follows in the last section. However, as we shall see in the next section, they are not sufficient to guarantee that $\underline{u}$ and $\underline{v}^*$ solve $\mathscr{P}$ and $\mathscr{P}^*$.

Theorem 3.1. *Suppose that $F \in \Gamma(V)$ and $G \in \Gamma(Y)$. If F is Gateaux differentiable and G^* is subdifferentiable and strictly convex, and if the solution $\underline{u}$ of $\mathscr{P}$ is unique then there exists precisely one solution $\underline{v}^*$ of $\mathscr{P}^*$, and $\Lambda^* \underline{v}^* = F'(\underline{u})$.*

Proof. If $\underline{u}$ is the unique solution of $\mathscr{P}$, then $\Lambda^* \underline{v}^* \in \partial F(\underline{u})$ for some $\underline{v}^* \in V^*$, and $\underline{v}^*$ is a solution of $\mathscr{P}^*$. If there exists another solution of $\mathscr{P}^*$, say v^*, then by assumption there exists an element $\Lambda u \in Y$ with $\Lambda u \in \partial G^*(v^*)$. Furthermore, u is a solution of $\mathscr{P}$, and so $u = \underline{u}$. The fact that $\Lambda \underline{u}$ is an element of both $\partial G^*(\underline{v}^*)$ and $\partial G^*(v^*)$ may be expressed as follows:

$$\begin{aligned}G(\Lambda \underline{u}) + G^*(\underline{v}^*) &= \langle \Lambda \underline{u}, \underline{v}^*\rangle,\\ G(\Lambda \underline{u}) + G^*(v^*) &= \langle \Lambda \underline{u}, v^*\rangle.\end{aligned}$$

The strict convexity of G^* implies that this is impossible unless $v^* = \underline{v}^*$. Note that in the course of the proof we have shown that $\Lambda^* \underline{v}^* = F'(\underline{u})$.

We now discuss the correspondence between the weak forms of the Euler-Lagrange equations which are satisfied by solutions of $\mathscr{P}$ and $\mathscr{P}^*$.

Theorem 3.2. *Suppose that $F \in \Gamma(V)$ and $G \in \Gamma(Y)$ are both Gateaux differentiable and strictly convex. If $\underline{u}$ is a solution of $\mathscr{P}$, then*

$$F'(\underline{u}) = (G \circ \Lambda)'(\underline{u}) = \Lambda^* v^*,$$

where

$$\partial G^*(v^*) = \partial(F^* \circ \Lambda^*)(v^*) = \{\Lambda \underline{u}\},$$

and v^* is a solution of $\mathscr{P}^*$.

Proof. Since F and $G \circ \Lambda$ are Gateaux differentiable, the proof is identical to that of Theorem 3.1 once one notices that

$$F'(\underline{u}) = (G \circ \Lambda)'(\underline{u}),$$

this equality following from the fact that $\underline{u}$ is a solution of $\mathscr{P}$.

3.2. Critical Points of J and $\hat{J}$

Let us consider briefly a problem related to that of finding extremals for the functional J. Let $F \in \Gamma(V)$ and $G \in \Gamma(Y)$. A point $u \in V$ will be called a *critical point* for J if

$$\partial(G \circ \Lambda)(u) \cap \partial F(u) \neq \phi.$$

Theorem 3.3. *Suppose $u \in V$ is a critical point for J. If*

$$\Lambda^* v^* \in \partial(G \circ \Lambda)(u) \cap \partial F(u),$$

then v^* is a critical point for $\hat{J}$, i.e.

$$\partial G^*(v^*) \cap \partial(F^* \circ \Lambda^*) v^* \neq \phi.$$

Furthermore,

$$G \circ \Lambda(u) - F(u) = F^* \circ \Lambda^*(v^*) - G^*(v^*).$$

Proof. Since $\Lambda^* v^* \in \partial F(u)$, we have

$$F(u) + F^* \circ \Lambda^* v^* = \langle u, \Lambda^* v^* \rangle$$

and since $\Lambda^* v^* \in \partial(G \circ \Lambda)(u)$ it follows that

$$G \circ \Lambda(u) + G^*(v^*) = \langle u, \Lambda^* v^* \rangle.$$

Hence

$$\Lambda u \in \partial G^*(v^*) \cap \partial(F^* \circ \Lambda^*)(v^*).$$

Subtracting the two expressions for $\langle u, \Lambda^* v^* \rangle$, we see that

$$G \circ \Lambda(u) - F(u) = F^* \circ \Lambda^*(v^*) - G^*(v^*).$$

Theorem 3.4. *If F and G are both strictly convex and Gateaux differentiable, then*

$$(G \circ \Lambda)'(u) - F'(u) = 0$$

if and only if

$$\partial(G \circ \Lambda)(u) \cap \partial F(u) \neq \phi.$$

Furthermore

$$(G \circ \Lambda)'(u) = F'(u) = \Lambda^* v^*$$

if and only if

$$\{\Lambda u\} = \partial G^*(v^*) = \partial(F^* \circ \Lambda^*)(v^*).$$

On the other hand

$$\Lambda u \in \partial G^*(v^*) \cap \partial(F^* \circ \Lambda^*) v^*$$

if and only if

$$\Lambda^* v^* = F'(u) = (G \circ \Lambda)'(u).$$

Proof. The first part or the theorem follows from the fact that

$$\partial(G \circ \Lambda)(u) = \{(G \circ \Lambda)'(u)\}$$

and

$$\partial F(u) = \{F'(u)\} \qquad \text{for all } u \in V.$$

Now suppose that $(G \circ \Lambda)'(u) - F'(u) = 0$ and $\Lambda^* v^* = F'(u)$. Then

$$F(u) + F^* \circ \Lambda^* v^* = \langle u, \Lambda^* v^* \rangle.$$

Since F is strictly convex, Λu is the unique element in $\partial(F^* \circ \Lambda^*)(v^*)$. Similarly Λu is the unique element in $\partial G^*(v^*)$. Conversely, suppose that

$$\Lambda u \in \partial G^*(v^*) \cap \partial(F^* \circ \Lambda^*)(v^*).$$

Then

$$G^*(v^*) + G \circ \Lambda(u) = \langle \Lambda u, v^* \rangle$$

and

$$F^* \circ \Lambda^*(v^*) + F(u) = \langle \Lambda u, v^* \rangle$$

which imply that

$$\Lambda^* v^* = F'(u) = (G \circ \Lambda)'(u).$$

The proof of the theorem is complete.

4.1. The Nonlinear Heavy Rotating Chain

It is the purpose of this section to explain, in terms of the results of Section 3, the connection between the two different differential equations which describe the motion of a heavy rotating chain. In the original paper on this problem [4], Kolodner considered the chain to be suspended with one end-point fixed, and acted on solely by the forces of gravity and tension T. He proceeded to seek solutions which describe a chain lying on a vertical plane and rotating with constant angular speed $\sqrt{\lambda}$. Taking the independent variable s to be arc-length measured from the free end of the chain, and the dependent variables $v(s)$ and $T(s)$ to be the horizontal displacement and tension, respectively, of the chain at the point at a distance s from the free end, he showed that the position of the chain could be described in terms of a new variable $u(s)=T(s)\,v(s)$ by the differential equation

$$\begin{aligned} u''(s)+\lambda u(s)\big(u(s)^2+s^2\big)^{-1/2}&=0\\ u(0)=u'(1)&=0. \end{aligned} \tag{4.1}$$

(Here all physical constants have been appropriately normalised.)

If the original variable $v(s)$ is retained, then after some manipulation it is found that the displacement v satisfies the equation

$$\begin{aligned} \left\{\frac{s\,v'(s)}{\big(1-v'(s)^2\big)^{1/2}}\right\}'+\lambda v(s)&=0\\ v(1)&=0. \end{aligned} \tag{4.2}$$

On examination it turns out that (4.2) is in fact the Euler-Lagrange equation for the functional

$$E_1(v)=-\int_0^1\left\{\frac{\lambda v(s)^2}{2}+s\big(1-v'(s)^2\big)^{1/2}\right\}ds$$

where $v(s)$ is subject to the conditions

$$|v'(s)|\leqq 1,\qquad v(1)=0.$$

This functional is the total energy of a chain which is lying on a plane in a configuration described by $v(s)$, $s\in[0,1]$, when the plane rotates with angular speed $\sqrt{\lambda}$ (see [9] for further details). On the other hand, (4.1) is the Euler-Lagrange equation for the functional

$$E_2(u)=\int_0^1\left\{\frac{u'(s)^2}{2\lambda}-\big(u(s)^2+s^2\big)^{1/2}\right\}ds,$$

where $u(0)=0$.

It is easy to see (c.f. [1]) that a minimiser for E_2 exists and that it is a classical solution of the Euler-Lagrange equation (4.1).

It is not so easy to show that a minimiser for E_2 exists. If such a minimiser were known to exist, it could lie on the boundary of the domain of definition of the functional, and so satisfy a differential inequality rather than the Euler-Lagrange equations, (4.2).

It is out intention to settle these questions, by investigating the duality of E_1 and E_2.

4.2. A Duality Result for the Heavy Rotating Chain

We shall begin by appropriately specifying the spaces V, V^*, Y and Y^* in section 3. Let V be the set of functions in $W^{1,2}[0,1]$ which vanish at zero. For any C^1-function $v(s)$ which vanishes at $s=0$, we have

$$v^2(s)=2\int_0^s v(t)\,v'(t)\,dt,$$

whence $\|v\|_{L^2}\leqq 2\,\|v'\|_{L^2}$ by Schwarz's inequality. It is possible, then, to define a norm on V by $|v|=\|v'\|_{L^2}$; this definition is equivalent to the usual $W^{1,2}[0,1]$ norm, and makes V a Hilbert space. Let Y be $L^2[0,1]$, and let Λ: $V\to Y$ be given by $\Lambda u=u'\in Y$ for $u\in V$. Here $'$ denotes weak differentiation. For each $v\in Y$ the function

$$u(s)=\int_0^s v(t)\,dt,\qquad s\in[0,1],$$

is an element of V, and $u'=v$. Hence Λ is surjective. If $\Lambda u=0$ for some $u\in V$, then $u=c$, a constant almost everywhere. But u is continuous and $u(0)=0$, and so $u\equiv 0$. Furthermore $\|\Lambda u\|_{L^2}=\|u'\|_{L^2}=|u|$, which implies that Λ is a linear homeomorphism between V and Y. We shall also put $Y^*=L^2\,[0,1]$ and

$$\langle v,v^*\rangle=\int_0^1 v(s)\,v^*(s)\,ds\qquad\text{for}\quad v^*\in Y^*,\ v\in Y.$$

Let us define the functionals F: $V\to\mathbb{R}$ by

$$F(u)=\int_0^1 (u(s)^2+s^2)^{1/2}\,ds,\qquad u\in V$$

and G: $Y\to\mathbb{R}$ by

$$G(v)=\int_0^1 \frac{v^2(s)}{2\lambda}\,ds,\qquad v\in Y.$$

Then, in the context of section 3, we are interested in the problem $\mathscr{P}$ given by

$$\inf_{u\in V} J(u)=\inf_{u\in V}\{G\circ\Lambda(u)-F(u)\}.$$

We wish to establish the existence of a minimiser in V of the functional

$$\int_0^1 \frac{u'(s)^2}{2\lambda}-(u(s)^2+s^2)^{1/2}\,ds.$$

Theorem 4.1. *For all $v^*\in Y^*$ we have*

$$F^*(\Lambda^* v^*)=-\int_0^1 s(1-v^{*\prime}(s)^2)^{1/2}\,ds$$

if $v^*\in W^{1,2}[0,1]$, $v^*(1)=0$ and $|v^{*\prime}(s)|\leqq 1$ a.e.

Proof. We begin by observing that, for all $v^*\in Y^*$ and $u\in V$,

$$\langle \Lambda^* v^*,u\rangle=\int_0^1 v^*(s)\,u'(s)\,ds.$$

For each $v^* \in Y^*$,

$$F^*(\Lambda^* v^*) = \sup_{u\in V}\{\langle u, \Lambda^* v^*\rangle - F(u)\}$$
$$= \sup_{u\in V}\left\{\int_0^1 [v^*(s)\,u'(s) - (u(s)^2 + s^2)^{1/2}]\, ds\right\}.$$

Now we must consider two cases. Suppose first that $v^* \notin W^{1,2}[0,1]$. Then

$$\left|\int_0^1 v^*(s)\, u'(s)\, ds\right|$$

is unbounded with respect to the L^2 norm of u in $W^{1,2}[0,1]$ (i.e. there exists a sequence $\{u_n\}$ in $W^{1,2}[0,1]$ with $\|u_n\|_{L^2} \leqq 1$, but $\int_0^1 v^*(s)\, u_n'(s)\, ds \to \infty$ as $n\to\infty$). Thus if $v^* \notin W^{1,2}[0,1]$ then $F^* \circ \Lambda^* v^* = \infty$.

Suppose on the other hand that $v^* \in W^{1,2}[0,1]$. If ϕ is a C^∞ function on $[0,1]$ with $\phi(0)=0$, then $\phi\in V$ and

$$F^*(\Lambda^* v^*) \geqq \int_0^1 \{v^*(s)\,\phi'(s) - (\phi(s)^2 + s^2)^{1/2}\}\, ds$$
$$= -\int_0^1 \{v^{*\prime}(s)\,\phi(s) + (\phi(s)^2 + s^2)^{1/2}\}\, ds + \phi(1)\, v^*(1).$$

Now we can choose a sequence $\{\phi_n\}$ of C^∞-functions in V with $\|\phi_n\|_{L^2} \leqq 1$ and $\phi_n(1) \to \infty$ as $n\to\infty$. Thus, unless $v^*(1)=0$, we have $F^* \circ \Lambda^* v^* = \infty$.

Finally suppose $v^* \in W^{1,2}[0,1]$, $v^*(1)=0$. Then

$$F^*(\Lambda^* v^*) = \sup_{u\in V} -\int_0^1 \{v^{*\prime}(s)\, u(s) + (u(s)^2 + s^2)^{1/2}\}\, ds$$
$$= \sup_{u\in L^2} -\int_0^1 \{v^{*\prime}(s)\, u(s) + (u(s)^2 + s^2)^{1/2}\}\, ds$$
$$= \begin{cases} -\int_0^1 s(1 - v^{*\prime}(s)^2)^{1/2}\, ds & \text{if } |v^{*\prime}(s)| < 1, \text{ a.e.} \\ \infty & \text{otherwise,} \end{cases}$$

the last equality following from Theorem 1.2 of section 1.

Theorem 4.2. *For* $v^* \in Y^* = L^2$,

$$G^*(v^*) = \int_0^1 \frac{\lambda\, v^*(s)^2}{2}\, ds.$$

Proof. By Theorem 1.2

$$G^*(v^*) = \sup_{v\in L^2[0,1]} \int_0^1 v^*(s)\, v(s) - \frac{v^2(s)}{2\lambda}\, ds$$
$$= \int_0^1 \frac{\lambda\, v^*(s)^2}{2}\, ds.$$

We have now established the duality between the problem of finding

$$\inf\int_0^1 \frac{u'(s)^2}{2\lambda} - \left(u(s)^2 + s^2\right)^{1/2} ds,$$

where $u \in W^{1,2}[0,1]$ and $u(0)=0$, and the problem of finding

$$\inf - \int_0^1 \frac{\lambda v^2(s)}{2} + s\left(1 - v'(s)^2\right)^{1/2} ds,$$

where $v \in W^{1,\infty}[0,1]$, $v(1)=0$, and $|v'|<1$ a.e.

It is clear that both F and G are strictly convex, Gateaux differentiable, and lower semi-continuous. It is known that there exists a unique minimiser $\underline{u}$ for J, and, since G^* is strictly convex and subdifferentiable, Theorem 3.1 implies that there exists a unique minimiser $\underline{v}^*$ for $\hat{J}=E_1$. Theorem 3.2 is also applicable here, so that

$$F'(\underline{u}) = (G \circ \Lambda)'(\underline{u}) = \Lambda^* \underline{v}^*$$

and

$$\partial G^*(\underline{v}^*) = \partial(F^* \circ \Lambda^*)(\underline{v}^*) = \{\Lambda \underline{u}\}.$$

These last two statements imply the equalities

$$-F(\underline{u}) + \langle \Lambda^* \underline{v}^*, \underline{u}\rangle = F^* \circ \Lambda^*(\underline{v}^*) \tag{4.3}$$

and

$$-G(\Lambda \underline{u}) + \langle \underline{v}^*, \Lambda \underline{u}\rangle = G^*(\underline{v}^*). \tag{4.4}$$

We remark that (4.3) implies $|\underline{v}^{*\prime}(s)|<1$ almost everywhere, and $\underline{v}^*(1)=0$. By Theorem 1.2

$$G^*(\underline{v}^*) = \int_0^1 g^*(s, \underline{v}^*(s))\, ds$$

where

$$g^*(s,t) = \sup_{x \in \mathbb{R}} \left\{ t x - \frac{x^2}{2\lambda} \right\} = \frac{\lambda t^2}{2}.$$

Hence

$$\frac{\lambda \underline{v}^*(s)^2}{2} - \underline{v}^*(s)\underline{u}'(s) + \frac{\underline{u}'(s)^2}{2\lambda} \geqq 0$$

almost everywhere, but

$$\int_0^1 \left\{ \frac{\lambda \underline{v}^*(s)^2}{2} - \underline{v}^*(s)\underline{u}'(s) + \frac{\underline{u}'(s)^2}{2\lambda} \right\} ds = 0$$

by (4.4). Therefore the integrand is zero almost everywhere, and so $\underline{v}^*(s) = \dfrac{\underline{u}'(s)}{\lambda}$ almost everywhere.

Now by (4.3), we have $|\underline{v}^{*\prime}(\mathfrak{s})|<1$ almost everywhere, $\underline{v}^*(1)=0$, and

$$\int_0^1 \{-(\underline{u}(\mathfrak{s})^2+\mathfrak{s}^2)^{1/2}-\underline{v}^{*\prime}(\mathfrak{s})\,\underline{u}(\mathfrak{s})+\mathfrak{s}(1-\underline{v}^{*\prime}(\mathfrak{s})^2)^{1/2}\}\,d\mathfrak{s}=0.$$

Again by Theorem 1.2, for $\underline{v}^*\in W^{1,\infty}$ with $\underline{v}^*(1)=0$ and $|\underline{v}^*(\mathfrak{s})|<1$ almost everywhere,

$$F^*\circ\Lambda^*(\underline{v}^*)=\int_0^1 f^*(\mathfrak{s},\underline{v}^{*\prime}(\mathfrak{s}))\,d\mathfrak{s}$$

where

$$f^*(\mathfrak{s},t)=\sup_{x\in\mathbb{R}}\{-tx-(x^2+\mathfrak{s}^2)^{1/2}\}=-\mathfrak{s}\sqrt{1-t^2}.$$

Once again the integrand is positive almost everywhere, and, since the integral is zero, the integrand is zero almost everywhere. Consequently

$$\underline{u}(\mathfrak{s})=\frac{-\mathfrak{s}\,\underline{v}^{*\prime}(\mathfrak{s})}{(1-\underline{v}^{*\prime}(\mathfrak{s})^2)^{1/2}}$$

almost everywhere. But $\underline{v}^*(\mathfrak{s})=\dfrac{\underline{u}'(\mathfrak{s})}{\lambda}$, and so

$$\left\{\frac{\mathfrak{s}\,\underline{v}^{*\prime}(\mathfrak{s})}{(1-\underline{v}^{*\prime}(\mathfrak{s})^2)^{1/2}}\right\}'+\lambda\,\underline{v}^*(\mathfrak{s})=0 \tag{4.5}$$

almost everywhere.

Thus by the duality theory of section 3 it is possible to establish the existence and uniqueness of a minimiser of the functional $\hat{J}=E_2$. Furthermore, we have shown that such a minimiser satisfies a weak form of the Euler-Lagrange equations. Finally we have established a pointwise relationship between the minimisers of J and $\hat{J}$ (which can be verified directly from the Euler-Lagrange equations, though it was established independently of them). Since J has an infinite number of critical points, so has $\hat{J}$ (there is a one-to-one correspondence between the critical points of J and those of $\hat{J}$, according to Theorem 3.3) and each critical point of $\hat{J}$ satisfies the Euler-Lagrange equations (4.5) in its weak form. The same pointwise relationship between critical points of J and $\hat{J}$ holds as in the case of their minimisers.

Note added in proof. The notion of duality considered in sections 2 and 3 has been generalised by the author in "Duality in Nonconvex Optimisation", J. Math. Anal. Appl. (in press), and the definition of critical point introduced in section 3.2 is further elaborated in "A variational method for boundary value problems with discontinuous nonlinearities" by C. A. STUART and J. F. TOLAND (to appear).

Recently the author learnt of a quite independent, but related, approach to the question of duality for non-convex problems due to IVAR EKELAND "Duality in Nonconvex Optimisation and the Calculus of Variations" (to appear in S.I.A.M. Journal of Control).

References

1. COFFMAN, C. V., "On variational principles for sublinear boundary value problems", Jour. Diff. Eqns., **17** 46–60 (1975).
2. EKELAND, I., & TEMAM, R., "Analyse convexe et problemes variationels", Dunod, Gauthiers. Paris 1974. (English translation "Convex analysis and variational problems", North-Holland, Amsterdam 1976.)
3. FENCHEL, W., "Convex cones, sets and functions", Notes de Cours polycopiees, Princeton University, 1951.
4. KOLODNER, I. I., "Heavy rotating string – a non-linear eigenvalue problem", Comm. Pure Appl. Math., **VIII** 395–408 (1955).
5. ROCKAFELLAR, T. R., "Duality and stability in extremal problems involving convex functions", Pac. Jour. Math., **21**, 167–187 (1967).
6. ROCKAFELLAR, T. R., "Convex Analysis", Princeton University Press, 1970.
7. ROCKAFELLAR, T. R., "Convex integral functions and duality", in Contributions to Nonlinear Functional Analysis (E. Zarantonello, ed.), Academic Press, 1971.
8. ROCKAFELLAR, T. R., "Conjugate convex functions in optimal control and the calculus of variations", Jour. Math. Anal. Appl., **32**, 174–222 (1970).
9. TOLAND, J. F., "On the stability of rotating heavy chains", to appear, Jour. Diff. Eqns.

Fluid Mechanics Research Institute
University of Essex
Wivenhoe Park
Colchester

(Received September 30, 1977)