Duality for Non-Convex Variational Principles

Giles Auchmuty*

Department of Mathematics, Indiana University, Bloomington, Indiana 47405

Received November 13, 1981; revised April 2, 1982

This paper treats the construction of dual variational principles for non-convex problems using Lagrangian methods. Besides the well-known saddle-point theory, there are dualities involving minimizing the Lagrangian in both variables and dualities where one first maximizes and then minimizes the Lagrangian in each variable. When the primal problem has certain structure, an associated canonical Lagrangian is defined which leads to the appropriate dual variational principle and to a Hamiltonian description. The properties of, and correspondence between, the critical points are analyzed. The last five sections are devoted to examples of these dual principles.

1. INTRODUCTION

There are many problems in mechanics where a number of different variational principles are known to describe the same system. Many of these involve convex functionals where the different variational principles may be considered as dual problems in the sense of convex analysis. There also are, however, some non-convex problems where this occurs. Toland studied one of these in [7] and subsequently developed a general theory for some of these problems in [8] and [9].

In trying to apply this theory to some other physical problems, it transpired that a Lagrangian version of the duality described in [8] was particularly informative. It could, moreover, be generalized to include some other classes of dual variational principles.

The results described here have largely been motivated by the examples described in Sections 7–12. The study of these examples inspired some of the general results, so it might prove worthwhile to look at the examples first.

The primal problem is always assumed to be a minimization problem. In Section 2, two types of Lagrangian are defined which could be associated with the primal problem. Depending on the type, the standard dual problem could be either a maximization problem (as in the classical theory) or a minimization problem. In certain examples, it was found that there could

* This work was supported in part by NSF Grant MCS 8102461. Present address: Department of Mathematics, University of Houston, Houston, Texas 77004.
also be a non-trivial "anomalous" dual problem. These basic definitions and constructions are given in Section 2, together with descriptions for canonical forms of Lagrangians under various assumptions on the primal functional.

Section 3 describes the possible types of critical points of the Lagrangians and their implications for the dual problems, while Section 4 considers the Hamiltonian formulation. In Section 5, we make a detailed study of the finite-dimensional case under smoothness and convexity assumptions on the Lagrangian. Then in Section 6, we generalize many of these results to general Banach spaces and non-differentiable functionals. The last six sections are devoted to specific examples of non-convex variational problems where this theory applies.

Whenever a term is used in this paper without definition, it should be taken in the sense used by Ekeland and Temam in [4].

2. LAGRANGIANS AND DUAL PRINCIPLES

Let $X, Y$ be two real, locally convex, topological vector spaces. Let $X^*, Y^*$ be their dual spaces and let $\langle \cdot, \cdot \rangle$ represent the pairing between a space and its dual. Suppose $f: X \to \mathbb{R}$ is a function.

The primal problem ($\mathcal{P}$) is to find

$$\alpha = \inf_{x \in X} f(x)$$

and

$$S = \{ x \in X : f(x) = \alpha \}. $$

$\alpha$ is called the value of ($\mathcal{P}$) and an element of $S$ is called a solution of ($\mathcal{P}$).

A function $L: X \times Y^* \to \mathbb{R}$ is said to be a Lagrangian of type I if, for each $x$ in $X$,

$$f(x) = \sup_{y^* \in Y^*} L(x, y^*). \quad (2.1)$$

$L$ is said to be a Lagrangian of type II if, for each $x$ in $X$,

$$f(x) = \inf_{y^* \in Y^*} L(x, y^*). \quad (2.2)$$

When $L$ is a Lagrangian of type I, the dual variational principle ($\mathcal{P}^*$) is to find

$$\alpha^* = \sup_{y^* \in Y^*} g(y^*)$$
and

\[ S^* = \{ y^* \in Y^*: g(y^*) = \alpha^* \}, \]

where \( g: Y^* \to \mathbb{R} \) is defined by

\[ g(y^*) = \inf_{x \in X} L(x, y^*). \tag{2.3} \]

There is also an anomalous dual variational principle \((\mathcal{D}^{\alpha})\) of finding

\[ \alpha^{\alpha} = \inf_{y^{\alpha} \in Y^{\alpha}} h(y^{\alpha}) \]

and

\[ S^{\alpha} = \{ y^{\alpha} \in Y^{\alpha}: h(y^{\alpha}) = \alpha^{\alpha} \}, \]

where

\[ h(y^{\alpha}) = \sup_{x \in X} L(x, y^{\alpha}). \tag{2.4} \]

When \( L \) is a Lagrangian of type I, the dual variational principle \((\mathcal{D}^{\alpha})\) is to find

\[ \alpha^* = \inf_{y^* \in Y^*} g(y^*) \]

and

\[ S^* = \{ y^* \in Y^*: g(y^*) = \alpha^* \}, \]

where \( g \) is again defined by (2.3).

The dual problems \((\mathcal{D}^{\alpha})\) or \((\mathcal{D}^{\alpha^*})\) are said to be non-trivial if there are \( y^* \)'s such that \( |g(y^*)| \neq \infty \), or \( |h(y^*)| \neq \infty \), respectively. Note that, so far we have not used any topological, or vector space, properties of \( X \) or \( Y^* \). They could be arbitrary sets.

From these definitions one sees that if \( L \) is a Lagrangian of type I, the dual problem \((\mathcal{D}^{\alpha^*})\) is a maximization problem and one has the familiar situation of seeking saddle-points of the Lagrangian.

When \( L \) is a Lagrangian of type II, the dual problem \((\mathcal{D}^{\alpha})\) is another minimization problem. If \( \hat{x} \) is a solution of \((\mathcal{D})\) and \( \hat{y}^* \) is a solution of \((\mathcal{D}^{\alpha^*})\), then \( (\hat{x}, \hat{y}^*) \) will minimize \( L \) on \( X \times Y^* \). A particular problem of this type was studied by Toland in [7] and a general theory was described in [8] and [9]. The anomalous variational problem \((\mathcal{D}^{\alpha^*})\) is also a minimization
problem. The reason for its name is the fact that if \( \hat{x} \) is a solution of \((\mathcal{L}^d)\), \( \hat{y}^* \) is a solution of \((\mathcal{G}^d)\), then

\[
f(\hat{x}) = \inf_{x \in X} \sup_{y^* \in Y^*} L(x, y^*)
\]

and

\[
h(\hat{y}^*) = \inf_{y^* \in Y^*} \sup_{x \in X} L(x, y^*).
\]

The surprising fact is that these are equal (see the corollary to Lemma 4.1) under some mild assumptions on \( L \).

To illustrate these definitions, it is worthwhile considering some general classes of problems which yield non-trivial, dual variational principles of the various types. More specific examples are given in the later chapters. Consider \( f: X \to \mathbb{R} \) defined by

\[
f(x) = f_1(Ax),
\]

where \( F: X \times Y \to \mathbb{R} \) is a given function and \( A: X \to Y \) is a continuous linear operator.

Very often

\[
f(x) = f_1(Ax) + f_2(x),
\]

where \( f_1 \) and \( f_2 \) are given functionals.

One says that \( F \) is convex on \( X \) (or \( Y \)) if, for each \( y \) in \( Y \) (respectively, \( x \) in \( X \)), \( F(\cdot, y) (F(x, \cdot)) \) is convex on \( X \) (or \( Y \)). Similarly, \( F \) is concave on \( X \) (or \( Y \)).

Also \( F \) is lower semi-continuous (l.s.c.) on \( X \) if, for each \( y \) in \( Y \), \( F(\cdot, y) \) is l.s.c. on \( X \). Similarly, we define upper semi-continuity (u.s.c.) on \( X \) and upper and lower semi-continuity on \( Y \). If \( X \) is a Banach space then \( f \) is said to be weakly l.s.c. (or u.s.c.) on \( X \) if it is l.s.c. (u.s.c.) on \( X \) when \( X \) has the weak topology.

When \( f: X \to \mathbb{R} \) is a given functional, its polar is the functional \( f^*: X^* \to \mathbb{R} \) defined by

\[
f^*(x^*) = \sup_{x \in X} |\langle x, x^* \rangle - f(x)|.
\]

In this paper we shall also need to define partial polars. Let \( F: X \times Y \to \mathbb{R} \) be given. Then \( F^*: X \times Y^* \to \mathbb{R} \) and \( F^*: X^* \times Y \to \mathbb{R} \) are defined by

\[
F^*(x, y^*) = \sup_{y \in Y} |\langle y, y^* \rangle - F(x, y)|
\]

and

\[
F^*(x^*, y) = \sup_{x \in X} |\langle x, x^* \rangle - F(x, y)|.
\]
These operations may be iterated and we shall write $F^\circ \#$ for $(F^\circ)^\#$ and $F^\# \circ$ for $(F^\#)\circ$. They are proper and non-trivial under conditions similar to those which imply $f^*$ is proper and non-trivial.

When $F$ is given and obeys various convexity (or concavity) and continuity conditions there are natural Lagrangians associated with $(\mathcal{C})$. Here, and henceforth, we shall always assume that $F$, $L$ and their various polars are proper.

**Case I.** Assume $F$ is convex and l.s.c. on $Y$ for each $x$ in $X$.
Define $L: X \times Y^* \to \mathbb{R}$ by

$$L(x, y^*) = \langle Ax, y^* \rangle - F^\circ(x, y^*).$$

(2.8)

This is a Lagrangian of type I because

$$\sup_{y^* \in Y^*} L(x, y^*) = F^\circ(x, Ax)$$

and the right-hand side is $F(x, Ax)$ from Proposition 1.4.1 of [4].

The dual variational principle $(\mathcal{V})$ is to find

$$\alpha^* = \sup_{y^* \in Y^*} g(y^*),$$

where

$$g(y^*) = \inf_{x \in X} L(x, y^*) = -F^*(-A^*y^*, y^*)$$

from an elementary computation. This is the classical dual variational principle for convex functionals and has been extensively studied. Note that in this construction we haven't required that $F$ be convex on $X \times Y$.

The anomalous dual variational principle $(\mathcal{V}^\otimes)$ is to find

$$\alpha^\otimes = \inf_{y^* \in Y^*} h(y^*),$$

where

$$h(y^*) = \sup_{x \in X} L(x, y^*) = F^\circ(A^*y^*, y^*).$$

(2.10)

When the Lagrangian is coercive on $X$ (as is often the case for convex functionals $F$), this functional $h$ will be trivial. In Section 10 we shall describe some problems where $h$ is non-trivial.

**Case II.** Assume $F$ is convex and l.s.c. on $X$ for each $y$ in $Y$.
Define $L: X \times X^* \to \mathbb{R}$ by

$$L(x, x^*) = \langle x, x^* \rangle - F^\#(x^*, Ax).$$

(2.11)
This is a Lagrangian of type I as
\[
\sup_{x^* \in X^*} L(x, x^*) = \sup_{x^* \in X^*} |\langle x, x^* \rangle - F^\#(x^*, Ax)|
\]
\[= F^\#(x, Ax),\]
which equals \( F(x, Ax) \) as before.

In this case the dual functionals \( g: X^* \to \mathbb{R} \) and \( h: X^* \to \mathbb{R} \) given by
\[
g(x^*) = \inf_{x \in X} |\langle x, x^* \rangle - F^\#(x^*, Ax)|
\]
and
\[
h(x^*) = \sup_{x \in X} |\langle x, x^* \rangle - F^\#(x^*, Ax)|
\]
may be defined but, unfortunately, do not have a simple expression in terms of \( F \).

Suppose now that \( A \) and \( A^* \) are surjective and that \( L: X \times Y^* \to \mathbb{R} \) is defined by
\[
L(x, y^*) = \langle Ax, y^* \rangle - F^\#(A^*y^*, Ax). \tag{2.12}
\]
In this expression \( A^*y^* \) replaces \( x^* \) in the right-hand side of (2.11). This Lagrangian is again of type I, since \( A^* \) is surjective and \( F \) is convex and l.s.c. on \( X \).

The dual functional \( g: Y^* \to \mathbb{R} \) defined by (2.3) is
\[
g(y^*) = \inf_{y \in Y} |\langle y, y^* \rangle - F^\#(A^*y^*, y)|
\]
\[= -(-F^\#)^\circ (A^*y^*, -y^*), \tag{2.13}
\]
where the first equality holds because \( A \) is surjective and the second follows by converting the inf to a sup.

The anomalous dual functional \( h: Y^* \to \mathbb{R} \) defined by (2.4) is
\[
h(y^*) = \sup_{y \in Y} |\langle y, y^* \rangle - F^\#(A^*y^*, y)|
\]
\[= F^\#(A^*y^*, y^*). \tag{2.14}
\]

**Case III.** Assume \( F \) is concave and u.s.c. on \( Y \) for each \( x \) in \( X \). Define \( L: X \times Y^* \to \mathbb{R} \) by
\[
L(x, y^*) = \langle Ax, y^* \rangle + (-F)^\circ (x, -y^*), \tag{2.15}
\]
This is a Lagrangian of type II as

\[
\inf_{y^* \in Y^*} L(x, y^*) = - \sup_{y^* \in Y^*} \left[ -\langle Ax, y^* \rangle - (-F)^\circ (x, -y^*) \right]
\]

\[
= -(F)^\circ (x, Ax)
\]

\[
= F(x, Ax)
\]

as \((-F)(x, \cdot)\) is convex and l.s.c. on \(Y\) for each \(x\) in \(X\).

This time, the dual functional \(g: Y^* \rightarrow \mathbb{R}\) defined by (2.3) is

\[
g(y^*) = \inf_{x \in X} [\langle Ax, y^* \rangle + (-F)^\circ (x, -y^*)]
\]

\[
= \sup_{x \in X} [\langle x, A^*y^* \rangle - (-F)^\circ (x, -y^*)]
\]

\[
= -[-(F)^\circ (-A^*y^*, -y^*)]. \quad (2.16)
\]

**Case IV.** Assume \(F\) is concave and u.s.c. on \(X\) for each \(y\) in \(Y\). Define \(L: X \times X^* \rightarrow \mathbb{R}\) by

\[
L(x, x^*) = \langle x, x^* \rangle + (-F)^\# (-x^*, Ax).
\]

This is a Lagrangian of type II because

\[
\inf_{x^* \in X^*} L(x, x^*) = \sup_{x^* \in X^*} \{-\langle x, x^* \rangle - (-F)^\# (-x^*, Ax)\}
\]

\[
= -(F)^\# (x, Ax) = F(x, Ax)
\]

as \((-F)(\cdot, y)\) is convex and l.s.c. on \(X\) for each \(y\) in \(Y\).

The dual functional \(g: X^* \rightarrow \mathbb{R}\) is defined by (2.3) so

\[
g(x^*) = \inf_{x \in X} [\langle x, x^* \rangle + (-F)^\# (-x^*, Ax)]. \quad (2.18)
\]

Once again there doesn't appear to be a simple expression for \(g\) in terms of \(F\) and its partial polars unless \(A\) is \(1 - 1\) and surjective.

When \(A\) and \(A^*\) are surjective, we could define \(L: X \times Y^* \rightarrow \mathbb{R}\) by

\[
L(x, y^*) = \langle Ax, y^* \rangle + (-F)^\# (-A^*y^*, Ax). \quad (2.19)
\]

Here \(A^*y^*\) replaces \(x^*\) in the right-hand side of (2.17) and, since \(A^*\) is surjective, \(L\) is again a Lagrangian of type II.
In this case the dual functional \( g: Y^* \to \mathbb{R} \) defined by (2.3) is

\[
g(y^*) = - \sup_{y \in Y} \{ - \langle y, y^* \rangle - (F^*(-A^*y^*, y)) \} = -(F^*)^* (-A^*y^*, -y^*), \tag{2.20}
\]

where we have used the surjectivity of \( A \) in the first equality.

Thus we have shown that if \( F \) is convex and l.s.c. in either variable, there is a Lagrangian of type I associated with the problem. When \( F \) is concave and u.s.c. in either variable there is a Lagrangian of type II associated with the problem. In particular, when \( F \) is convex and l.s.c. in one variable and concave and u.s.c. in the other, one may construct Lagrangians of both types. The different Lagrangians yield different dual variational principles.

When \( f \) has the special form (2.6), these formulas simplify further and the Lagrangian has one of the forms

\[
L(x, x^*) = \langle x, x^* \rangle + l_1(x) + l_2(x^*)
\]

or

\[
L(x, y^*) = \langle Ax, y^* \rangle + l_1(x) + l_2(y^*),
\]

where \( l_1, l_2 \) are given functionals.

**Case I.** Assume \( f_1 \) is convex and l.s.c. on \( Y \). Then

\[
L(x, y^*) = \langle Ax, y^* \rangle + f_2(x) - f_1^*(y^*), \quad g(y^*) = -f_1^*(y^*) - f_2^*(-A^*y^*), \tag{2.22}
\]

and

\[
h(y^*) = -f_1^*(y^*) + (-f_2^*)^*(A^*y^*).
\]

**Case II.** Assume \( f_2 \) is convex and l.s.c. on \( X \). Then

\[
L(x, x^*) = \langle x, x^* \rangle + f_1(Ax) - f_2^*(x^*) \tag{2.23}
\]

when \( L \) is defined by (2.11).

When \( A \) and \( A^* \) are surjective (2.12) yields

\[
L(x, y^*) = \langle Ax, y^* \rangle + f_1(Ax) - f_2^*(A^*y^*)
\]

so

\[
g(y^*) = -f_1^*(-y^*) - f_2^*(A^*y^*)
\]

and

\[
h(y^*) = (-f_1^*)^*(y^*) - f_2^*(A^*y^*)
\]
Case III. Assume $f_1$ is concave and u.s.c. on $Y$. Then
\[ L(x, y^*) = \langle Ax, y^* \rangle + (-f_1)^*(-y^*) + f_2(x) \]
and
\[ g(y^*) = (-f_1)^*(-y^*) - f_2^*(-A^*y^*). \]

Case IV. Assume $f_2$ is concave and u.s.c. on $X$. Then
\[ L(x, x^*) = \langle x, x^* \rangle + f_1(Ax) + (-f_2)^*(-x^*) \]
when $L$ is defined by (2.17).

When $A$ and $A^*$ are surjective (2.19) becomes
\[ L(x, y^*) = \langle Ax, y^* \rangle + f_1(Ax) + (-f_2)^*(-A^*y^*). \]

This time $g(y^*) = f_1^*(y^*) + (-f_2)^*(A^*y^*)$.

Note that in cases I and II, when $f$ is given by (2.6), the usual dual functional is always the sum of two concave functionals. The dual functionals in cases III and IV and the anomalous functionals in cases I and II are always the difference of two convex functionals.

In cases I and II, the Lagrangian is of type I; in cases III and IV it is of type II.

3. Critical Points of Lagrangians

When one has dual variational principles, the solutions of the dual problems are special critical points of the Lagrangian. In this section we shall describe these relationships.

Let $L: X \times Y^* \rightarrow \mathbb{R}$ be a given function.

A point $(\hat{x}, \hat{y}^*)$ is said to be a saddle-point of $L$ if
\[ I.(\hat{x}, y^*) \leq I.(\hat{x}, \hat{y}^*) \leq I.(x, \hat{y}^*) \]
for all $x$ in $X$, $y^*$ in $Y^*$.

Some well-known results in the theory of convex duality may be summarized as follows (cf. [4, Proposition III.3.1] although we do not assume any convexity in this version).

**Theorem 3.1.** Let $L$ be a Lagrangian of type I. Then $(\hat{x}, \hat{y}^*)$ is a saddle-point of $L$ iff $\hat{x}$ is a solution of ($\mathcal{P}$), $\hat{y}^*$ is a solution of ($\mathcal{P}^*$) and $\alpha = \alpha^*$. 

Proof. Choose \( \tilde{x} \) in \( X \), \( \tilde{y}^* \) in \( Y^* \). Then

\[
f(\tilde{x}) - g(\tilde{y}^*) = \sup_{y \in Y^*} L(\tilde{x}, y) - \inf_{x \in X} L(x, \tilde{y}^*)
\]

\[
= \sup_{x \in X} \sup_{y \in Y^*} |L(\tilde{x}, y) - L(x, \tilde{y}^*)| \geq 0.
\]  \( (3.2) \)

When \((\tilde{x}, \tilde{y}^*)\) is a saddle-point of \( L \) one has, from (3.1) and the definitions of \( f, g \), that

\[
f(\tilde{x}) \leq L(\tilde{x}, \tilde{y}^*) \leq g(\tilde{y}^*).
\]

Combining this with (3.2) one has \( f(\tilde{x}) = L(\tilde{x}, \tilde{y}^*) = g(\tilde{y}^*) \). Thus \( \tilde{x} \) is a solution of \((\tilde{x}^*)\), \( \tilde{y}^* \) is a solution of \((\tilde{y}^*)\) and \( \alpha = \alpha^* \).

Conversely, if \( \tilde{x} \) is a solution of \((\tilde{x}^*)\), \( \tilde{y}^* \) is a solution of \((\tilde{y}^*)\) and \( \alpha = \alpha^* \), then \( f(\tilde{x}) = g(\tilde{y}^*) \).

From the definitions of \( f, g \) one has

\[
\sup_{y \in Y^*} L(\tilde{x}, y) = \inf_{x \in X} L(x, \tilde{y}^*) \leq L(\tilde{x}, \tilde{y}^*).
\]

Thus

\[
L(\tilde{x}, y) \leq L(\tilde{x}, \tilde{y}^*) \quad \text{for all } y \in Y^*.
\]

Similarly,

\[
\inf_{x \in X} L(x, \tilde{y}^*) \geq \sup_{y \in Y^*} L(\tilde{x}, y) \geq L(\tilde{x}, \tilde{y}^*),
\]

so

\[
L(x, \tilde{y}) \geq L(\tilde{x}, \tilde{y}^*) \quad \text{for all } x \text{ in } X.
\]

Thus (3.1) holds or \((\tilde{x}, \tilde{y}^*)\) is a saddle-point of \( L \).  \( \blacksquare \)

When \( L \) is given, we define the partial subdifferentials \( \partial_1 L(\tilde{x}, \tilde{y}^*) \) and \( \partial_2 L(\tilde{x}, \tilde{y}^*) \) as follows:

\[
v \in \partial_1 L(\tilde{x}, \tilde{y}^*) \quad \text{if } v \in X^*
\]

and

\[
L(x, \tilde{y}^*) \geq L(\tilde{x}, \tilde{y}^*) + \langle v, x - \tilde{x} \rangle \quad \text{for any } x \text{ in } X,
\]

\[
w \in \partial_2 L(\tilde{x}, \tilde{y}^*) \quad \text{if } w \in Y^{**},
\]

and

\[
L(\tilde{x}, y^*) \geq L(\tilde{x}, \tilde{y}^*) + \langle w, y^* - \tilde{y}^* \rangle \quad \text{for all } y^* \text{ in } Y^*.
\]
From (3.1) one sees that \((\hat{x}, \hat{y}^*)\) is a saddle-point of \(L\) iff
\[
0 \in \partial_1 L(\hat{x}, \hat{y}^*) \quad \text{and} \quad 0 \in \partial_2 (-L)(\hat{x}, \hat{y}^*). \tag{3.3}
\]

A point \((\hat{x}, \hat{y}^*)\) is said to be a \(\partial\)-critical point of \(L\) if
\[
0 \in \partial_1 L(\hat{x}, \hat{y}^*) \quad \text{and} \quad 0 \in \partial_2 L(\hat{x}, \hat{y}^*). \tag{3.4}
\]

From the definitions of \(\partial_1 L\) and \(\partial_2 L\) one sees that \((\hat{x}, \hat{y}^*)\) is a \(\partial\)-critical point of \(L\) iff
\[
L(x, \hat{y}^*) \geq L(\hat{x}, \hat{y}^*) \quad \text{for all } x \text{ in } X, \tag{3.5}
\]
and
\[
L(\hat{x}, y^*) \geq L(\hat{x}, \hat{y}^*) \quad \text{for all } y^* \text{ in } Y^*. \tag{3.6}
\]

Note that if \(0 \in \partial L(\hat{x}, \hat{y}^*)\) then \((\hat{x}, \hat{y}^*)\) is a \(\partial\)-critical point of \(L\) on \(X \times Y^*\) and, in fact, \(L(\hat{x}, \hat{y}^*)\) is the infimum of \(L\) on \(X \times Y^*\). However, \((\hat{x}, \hat{y}^*)\) may be a \(\partial\)-critical point of \(L\) on \(X \times Y^*\) without minimizing \(L\) on \(X \times Y^*\). When \((\hat{x}, \hat{y}^*)\) is a \(\partial\)-critical point of \(L\), Toland [7] calls \(\hat{x}\) and \(\hat{y}^*\) "critical points in duality."

**Lemma 3.2.** If \(L\) is a Lagrangian of type II and \((\hat{x}, \hat{y}^*)\) is a \(\partial\)-critical point of \(L\), then \(f(\hat{x}) = g(\hat{y}^*) = L(\hat{x}, \hat{y}^*)\).

**Proof.** The fact that \(0 \in \partial_1 L(\hat{x}, \hat{y}^*)\) implies \(L(x, \hat{y}^*) \geq L(\hat{x}, \hat{y}^*)\) for all \(x \in X\). Thus \(g(\hat{y}^*) = \inf_{x \in X} L(x, \hat{y}^*) = L(\hat{x}, \hat{y}^*)\).

Similarly, \(0 \in \partial_2 L(\hat{x}, \hat{y}^*)\) implies \(f(\hat{x}) = \inf_{y \in Y^*} L(\hat{x}, y) = L(\hat{x}, \hat{y}^*)\) so \(f(\hat{x}) = g(\hat{y}^*) = L(\hat{x}, \hat{y}^*)\).

**Theorem 3.3.** Let \(L\) be a Lagrangian of type II, then \(\alpha = \alpha^*\). A point \((\hat{x}, \hat{y}^*)\) minimizes \(L\) on \(X \times Y^*\) iff \(\hat{x}\) is a solution of \((\hat{\varphi})\), \(\hat{y}^*\) is a solution of \((\hat{\psi}^*)\) and \(L(\hat{x}, \hat{y}^*) = \alpha = \alpha^*\).

**Proof.** One has
\[
\alpha = \inf_{x \in X} \left( \inf_{y^* \in Y^*} L(x, y^*) \right),
\]
\[
\alpha^* = \inf_{y^* \in Y^*} \left( \inf_{x \in X} L(x, y^*) \right)
\]
so \(\alpha = \alpha^*\) as one can take the infima in either order.

Suppose \((\hat{x}, \hat{y}^*)\) minimizes \(L\) on \(X \times Y^*\). Then
\[
L(\hat{x}, \hat{y}^*) = \inf_{x \in X} f(x) = \inf_{y^* \in Y^*} g(y^*). \tag{3.6}
\]
Thus \( L(\hat{x}, \hat{y}^*) = a = a^* \). Moreover, from the definition of \( f \),

\[
f'(\hat{x}) = \inf_{y^* \in Y'} L(\hat{x}, y^*) \leq L(\hat{x}, \hat{y}^*).
\]

Combining this with (3.6) one has \( f(\hat{x}) = \inf_{x \in X} f(x) \) or \( \hat{x} \) is a solution of \((\mathcal{P})\). Similarly, \( \hat{y}^* \) is a solution of \((\mathcal{P}^*)\).

Conversely, if \( \hat{x} \) is a solution of \((\mathcal{P})\), then

\[
f(x) = \inf_{y^* \in Y'} L(x, y^*) \geq f(\hat{x}) = a \quad \text{for any } x \in X.
\]

Thus \( L(x, y^*) \geq a \) for any \( x \in X \), \( y^* \in Y^* \). But \( a = L(\hat{x}, \hat{y}^*) \) by assumption so \((\hat{x}, \hat{y}^*)\) minimizes \( L \) on \( X \times Y^* \) as required. \( \blacksquare \)

**Theorem 3.4.** Let \( L \) be a Lagrangian of type II and \((\hat{x}, \hat{y}^*)\) be a \( \partial \)-critical point of \( L \). Then the following statements are equivalent:

(i) \( \hat{x} \) is a solution of \((\mathcal{P})\).

(ii) \( \hat{y}^* \) is a solution of \((\mathcal{P}^*)\), and

(iii) \( L(\hat{x}, \hat{y}^*) = \inf_{X \times Y'} L(x, y^*) \).

**Proof.** Suppose \( \hat{x} \) is a solution of \((\mathcal{P})\). Then

\[
f(\hat{x}) = \inf_{\hat{x}} f(x) = \inf_{X \times Y'} L(x, y^*) = a.
\]

From Lemma 3.2 one has, since \((\hat{x}, \hat{y}^*)\) is a \( \partial \)-critical point, that

\[
g(\hat{y}^*) = L(\hat{x}, \hat{y}^*) = a.
\]

Therefore Theorem 3.3 implies that \( \hat{y}^* \) is a solution of \((\mathcal{P}^*)\) as \((\hat{x}, \hat{y}^*)\) minimizes \( L \) on \( X \times Y^* \).

Statements (ii) and (iii) then follow. Similarly (ii) implies (i) and (iii). The fact that (iii) implies (i) and (ii) comes from Theorem 3.3.

**Theorem 3.5.** Let \( L \) be a Lagrangian of type II which is l.s.c. on \( X \times Y^* \). Suppose that \( E_c = \{(x, y^*) \}: L(x, y^*) \leq c \) is non-empty and sequentially compact for some \( c \). Then \((\mathcal{P})\) and \((\mathcal{P}^*)\) have solutions.

**Proof.** Let \( \{(x_n, y_n^*): n \geq 1\} \) be a minimizing sequence for \( L \) with \( L(x_n, y_n^*) \to a_0 = \inf_{X \times Y'} L(x, y^*) \) as \( n \to \infty \).

Suppose \( c > a_0 \), then there exists \( N \) such that \( \{(x_n, y_n^*): n \geq N\} \subseteq E_c \). Thus there is a convergent subsequence \( \{(x_{n_k}, y_{n_k}^*): n_k \geq N\} \) converging to a point \((\hat{x}, \hat{y}^*)\) in \( E_c \), since \( E_c \) is sequentially compact.
Since $L$ is l.s.c. on $X \times Y^*$, one has

$$L(\bar{x}, \bar{y}^*) \leq \liminf_{n_k \to \infty} L(x_{n_k}, y_{n_k}^*) = \alpha_0.$$ 

Thus $(\bar{x}, \bar{y}^*)$ minimizes $L$ on $X \times Y^*$, and Theorem 3.3 implies that $\bar{x}$ is a solution of $(\mathcal{P}^*)$ and $\bar{y}^*$ is a solution of $(\mathcal{P}^*)$.

When $c = a_0$, any element of $E_c$ provides a pair of solutions of $(\mathcal{P}^*)$ and $(\mathcal{P}^*)$. When $c < a_0$, $E_c$ must be empty.

It is quite easy to construct examples of Lagrangians of type II where the primal problem has a solution but the dual problem does not. An example with $X = Y^* = \mathbb{R}$ is $L(x, y) = \frac{1}{2}x^2 + e^{-y}$. Then $f(x) - \frac{1}{2}x^2$, $g(y) = e^{-y}$ and $\alpha = \alpha^* = 0$. Here $(\mathcal{P}^*)$ has the solution 0 but $g$ does not attain its infimum.

A point $\bar{x}$ in $X$ is said to be critical point of a functional $f: X \to \mathbb{R}$ if $f$ is Gâteaux-differentiable at $\bar{x}$ and

$$Df(\bar{x}) = 0 \quad \text{in } X^*.$$ \hspace{1cm} (3.7)

Here $Df(\bar{x})$ is the Gâteaux-derivative of $f$ at $\bar{x}$.

Similarly, a point $(\bar{x}, \bar{y}^*)$ in $X \times Y^*$ is said to be a critical point of the Lagrangian $L$ if $L$ is Gâteaux-differentiable at $(\bar{x}, \bar{y}^*)$ with respect to both $x$ and $y^*$ separately and

$$D_1 L(\bar{x}, \bar{y}^*) = 0 \quad \text{in } X^*.$$ \hspace{1cm} (3.8)

$$D_2 L(\bar{x}, \bar{y}^*) = 0 \quad \text{in } Y^{**}.$$ 

Here $D_1, D_2$ denote partial Gâteaux-derivatives on $X, Y^*$, respectively.

**Lemma 3.6.** If $f: X \to \mathbb{R}$ is Gâteaux-differentiable at $\bar{x}$ and $\xi \in \partial f(\bar{x})$ then $\xi = Df(\bar{x})$.

**Proof.** When $\xi \in \partial f(\bar{x})$ one has

$$f(\bar{x} + th) - f(\bar{x}) > t\langle \xi, h \rangle$$

for any $t \in \mathbb{R}, h \in X$. Divide by $t$ and let $t \to 0^+$. Then

$$\langle Df(\bar{x}), h \rangle < \langle \xi, h \rangle,$$

or

$$\langle Df(\bar{x}) - \xi, h \rangle \geq 0 \quad \text{for all } h \in X.$$ 

Thus $DF(\bar{x}) = \xi$ as $h$ is arbitrary.
COROLLARY 1. If \((\hat{x}, \hat{y}^*)\) is either a saddle-point of \(L\) or a \(\partial\)-critical point of \(L\) and \(L\) is partially Gâteaux-differentiable at \((\hat{x}, \hat{y}^*)\), then \((\hat{x}, \hat{y}^*)\) is a critical point of \(L\).

**Proof.** When \((\hat{x}, \hat{y}^*)\) is a saddle-point of \(L\), (3.3) and the lemma imply

\[
D_1 L(\hat{x}, \hat{y}^*) = 0 \quad \text{and} \quad -D_2 L(\hat{x}, \hat{y}^*) = 0.
\]

Thus \((\hat{x}, \hat{y}^*)\) is a critical point of \(L\).

When \((\hat{x}, \hat{y}^*)\) is a \(\partial\)-critical point, the result follows from (3.4).

COROLLARY 2. Let \((\hat{x}, \hat{y}^*)\) be either a saddle-point or a \(\partial\)-critical point of \(L\). If \(f'(g)\) is Gâteaux-differentiable at \(\hat{x}(\hat{y}^*)\), then \(Df(\hat{x}) = 0\) \((Dg(\hat{y}^*) = 0)\).

**Proof.** Let \((\hat{x}, \hat{y}^*)\) be a saddle-point of \(L\). From (3.1), \(f(x) \geq f(\hat{x})\) for all \(x\) in \(X\). Thus \(0 \in \partial f(\hat{x})\) and the lemma implies \(\hat{x}\) is a critical point.

Similarly for \(g\) or when \((\hat{x}, \hat{y}^*)\) is a \(\partial\)-critical point.

One would like to have the analog of Corollary 2 when \((\hat{x}, \hat{y}^*)\) is a general critical point of \(L\). To do this one usually has to make more assumptions on \(L, f\) and \(g\).

Consider the following conditions:

1. **(D1)** There exist a neighborhood \(U\) of \(\hat{x}\) in \(X\) and a function \(\xi: U \to Y^*\) such that \(\xi(\hat{x}) = \hat{y}^*\) and

\[
f(x) = L(x, \xi(x)), \quad \text{for} \ x \ \text{in} \ U.
\]

2. **(D2)** There exist a neighborhood \(V\) of \(\hat{y}^*\) in \(y^*\) and a function \(\eta: V \to X\) such that \(\eta(\hat{y}^*) = \hat{x}\) and

\[
g(y^*) = L(\eta(y^*), y^*), \quad \text{for} \ y^* \ \text{in} \ V.
\]

A function \(f: X \to \mathbb{R}\) is said to be \(C\)-differentiable at \(\hat{x}\) in \(X\) if for any curve \(c: (-\delta, \delta) \to X\), which is continuously differentiable and obeys \(c(0) = \hat{x}\).

one has that \(f \circ c: (-\delta, \delta) \to \mathbb{R}\) is differentiable at 0 and the derivative \(d/dt f(c(t))|_{t=0} = \langle \zeta, c'(0) \rangle\), where \(c'(0) \in X\) is the derivative of \(c\) at 0.

Then \(\zeta\) is called the derivative of \(f\) at \(\hat{x}\) and one writes \(Df(\hat{x}) = \zeta\).

When \(f\) is \(C\)-differentiable at \(\hat{x}\) then it is Gâteaux-differentiable as the latter only requires differentiability along straight lines.

**Lemma 3.7.** Suppose \(L\) is \(C\)-differentiable in a neighbourhood of \((\hat{x}, \hat{y}^*)\) in \(X \times Y^*\) and \((\hat{x}, \hat{y}^*)\) is a critical point of \(L\). Assume **(D1)** holds and \(\xi\) is continuously Gâteaux-differentiable near \(\hat{x}\). Then \(\hat{x}\) is a critical point of \(f\).
Consider $\phi: (-\delta, \delta) \to \mathbb{R}$ defined by
\[
\phi(t) = L(\hat{x} + th, \zeta(\hat{x} + th))
\]
for some $h$ in $X$. Consider the curve $c: (-\delta, \delta) \to X \times Y^*$ defined by
\[
c(t) = (\hat{x} + th, \zeta(\hat{x} + th)).
\]
Then since $\zeta$ is continuously Gâteaux-differentiable, $c$ is continuously differentiable and $c'(0) = (h, D\zeta(\hat{x}) h)$, where $D\zeta(\hat{x})$ is the Gâteaux derivative of $\zeta$ at $\hat{x}$.

Since $\phi(t) - L(c(t))$ and $L$ is $C$-differentiable, there is a $\zeta \in X^* \times Y^{**}$ such that $\phi'(0) = \langle \zeta, c'(0) \rangle$.

Write $\zeta = (\zeta_1, \zeta_2)$, then $\zeta_1 = D_1 L(\hat{x}, \hat{y}^*)$ and $\zeta_2 = D_2 L(\hat{x}, \hat{y}^*)$, where the right-hand sides are the partial Gateaux-derivatives. Thus $\zeta = (0, 0)$ as $(\hat{x}, \hat{y}^*)$ is a critical point and consequently $\phi'(0) = 0$. But $\phi(t) = f(\hat{x} + th)$ from (D1) so $f$ is Gateaux-differentiable at $\hat{x}$ and $\langle Df(\hat{x}), h \rangle = 0$ for any $h$ in $X$.

Thus the lemma holds.

**Corollary.** Suppose $L$ as in Lemma 3.6 and (D2) holds. If $\eta$ is continuously Gateaux-differentiable near $\hat{y}^*$ then $\hat{y}^*$ is a critical point of $g$.

**Remark.** When one replaces $g$ by $h$ in (D2) then one can replace $g$ by $h$ in this corollary.

### 4. Hamiltonian Formulation

Many studies of dual variational principles in applied analysis (see, e.g., Arthurs [1] and the references therein) take as their starting point a version of Hamilton's equations. In such formulations there is a special linear operator $A$ and they restrict attention to functionals $f$ of the form (2.5).

When $L$ is a Lagrangian for a variational principle, the associated Hamiltonian is $H: X \times Y^* \to \mathbb{R}$ defined by
\[
H(x, y^*) = \langle Ax, y^* \rangle - L(x, y^*).
\]  

In each of the special cases described in Section 2, the Hamiltonian may be described in terms of the polars of $F$. Namely:

<table>
<thead>
<tr>
<th>Case</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H(x, y^*)$</td>
<td>$f^\circ(x, y^*)$</td>
<td>$F^#(A^<em>y^</em>, Ax)$</td>
<td>$-(-F)^\circ(x, -y^*)$</td>
<td>$(-F)^#(A^<em>y^</em>, Ax)$</td>
</tr>
</tbody>
</table>
In cases II and IV we have assumed that \( A \) and \( A^* \) are surjective. When \( F \)
has the form (2.6), the Hamiltonian is separable with
\[
H(x, y^*) = h_1(x) + h_2(y^*)
\]  
and \( h_1, h_2 \) are given below:

<table>
<thead>
<tr>
<th>Case</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1(x) )</td>
<td>(-f_1(x))</td>
<td>(-f_1(Ax))</td>
<td>(-f_1(x))</td>
<td>(-f_1(Ax))</td>
</tr>
<tr>
<td>( h_1(y^*) )</td>
<td>(f_1^<em>(y^</em>))</td>
<td>(f_2^*(A^<em>y^</em>))</td>
<td>((-f_1^<em>) (-y^</em>))</td>
<td>((-f_2^*) (-A^<em>y^</em>))</td>
</tr>
</tbody>
</table>

Note that \( h_2 \) is convex in cases I and II and concave in cases III and IV
and no special assumptions have been made on \( h_1 \). From (3.3) and (4.1) one
sees that \((\hat{x}, \hat{y}^*)\) is a saddle point of \( I \) iff

\[
-A^*\hat{y}^* \in \partial_1(-H)(\hat{x}, \hat{y}^*)
\]

and

\[
A\hat{x} \in \partial_2 H(\hat{x}, \hat{y}^*).
\]  
Similarly from (3.4) and (4.1), \((\hat{x}, \hat{y}^*)\) is a \( \hat{c} \)-critical point of \( L \) iff

\[
-A^*\hat{y}^* \in \partial_1(-H)(\hat{x}, \hat{y}^*)
\]

and

\[
-A\hat{x} \in \partial_2(-H)(\hat{x}, \hat{y}^*).
\]

Also \((\hat{x}, \hat{y}^*)\) is a critical point of \( L \) iff \( H \) is partially Gâteaux-
differentiable at \((\hat{x}, \hat{y}^*)\) and

\[
A^*\hat{y}^* = D_1 H(\hat{x}, \hat{y}^*),
\]

\[
A\hat{x} = D_2 H(\hat{x}, \hat{y}^*).
\]  

These are the usual abstract forms of Hamilton's equations as used, for
example, in [1].

When a Hamiltonian \( H \) is given, the functionals \( f, g \) and \( h \) may be
computed directly. From (4.1) one has

\[
L(x, y^*) = \langle Ax, y^* \rangle - H(x, y^*).
\]  

When \( L \) is a Lagrangian of type I, then

\[
f(x) = \sup_{y^*} L(x, y^*) = H^0(x, Ax),
\]
\[ g(y^*) = \inf_{x} L(x, y^*) = - \sup_{x} \left| -\langle Ax, y^* \rangle + H(x, y^*) \right| \]
\[ = - (-H)^* (-A^*y^*, y^*) \]  \hspace{1cm} (4.8)

and
\[ h(y^*) = \sup_{x} L(x, y^*) = H^*(A^*y^*, y^*). \]  \hspace{1cm} (4.9)

When \( L \) is a Lagrangian of type II, similar computations yield
\[ f(x) = \inf_{y^*} L(x, y^*) \]
\[ = - \mathcal{H}^0(x, Ax), \]  \hspace{1cm} (4.10)

where \( \mathcal{H}^0(x, y^*) = -H(x, y^*). \) In this case \( g \) is again defined by (4.8) or
\[ g(y^*) = - \mathcal{H}^0(-A^*y^*, y^*). \]

For Lagrangians of type II, we shall use \( \mathcal{H}^* \) instead of \( H. \) Moreover, one sees that if one is given \( L, H \) or \( \mathcal{H}^* \), then one may derive the primal and dual problems quite easily.

The Hamiltonian theory of saddle-points is well known. Here we shall concentrate on the Hamiltonian theory for anomalous duality and when \( L \) is a Lagrangian of type II.

The anomalous dual variational principles for a Lagrangian of type I may be interpreted when they are in Hamiltonian form.

A point \((\hat{x}, \hat{y}^*)\) is said to be an anomalous critical point of \( L \) if
\[ A\hat{x} \in \partial_2 H(\hat{x}, \hat{y}^*) \]
and
\[ A^*\hat{y}^* \in \partial_1 H(\hat{x}, \hat{y}^*). \]  \hspace{1cm} (4.11)

A sufficient, but not necessary, condition that this holds is
\[ (A^*\hat{y}^*, A\hat{x}) \in \partial H(\hat{x}, \hat{y}^*). \]  \hspace{1cm} (4.12)

When \( A = d/dt, \) \( A^* = -d/dt, \) this is precisely the definition of Hamiltonian inclusion as used in Clarke [10].

**Lemma 4.1.** Let \( L \) be a Lagrangian of type 1 and \( H \) be the corresponding Hamiltonian. Then

(i) \( Ax \in \partial_2 H(x, \hat{y}^*) \) implies \( f(x) = L(x, \hat{y}^*) \leq h(\hat{y}^*) \) and

(ii) \( A^*y^* \in \partial_1 H(\hat{x}, y^*) \) implies \( h(y^*) = L(\hat{x}, y^*) \leq f(\hat{x}). \)
Proof. (i) When $Ax \in \partial_2 H(x, \tilde{y}^*)$ one has

$$H(x, y^*) \geq H(x, \tilde{y}^*) + \langle y^* - \tilde{y}^*, Ax \rangle$$

for any $y^*$ in $Y^*$.

Thus $\langle Ax, y^* \rangle - H(x, y^*) \leq \langle Ax, \tilde{y}^* \rangle - H(x, \tilde{y}^*)$.

Taking the supremum over $Y^*$, and using the definition of $h$, one has

$$f(x) = L(x, \tilde{y}^*) \leq h(\tilde{y}^*)$$

as required.

Similarly for (ii).

Corollary. Let $L$ be a Lagrangian of type 1. If $(\hat{x}, \hat{y}^*)$ is an anomalous critical point of $L$, then $f(\hat{x}) = h(\hat{y}^*) = L(\hat{x}, \hat{y}^*)$.

Lemma 4.2. Let $L$ be a Lagrangian of type 1 and $(\hat{x}, \hat{y}^*)$ be an anomalous critical point of $L$. Suppose $L$ is partially Gâteaux-differentiable at $(\hat{x}, \hat{y}^*)$, then $(\hat{x}, \hat{y}^*)$ is a critical point of $L$.

Proof. One has

$$t^{-1} |L(\hat{x} + th, \hat{y}^*) - (L(\hat{x}, \hat{y}^*))| = \langle Ah, y^* \rangle - t^{-1} |H(\hat{x} + th, \hat{y}^*) - H(\hat{x}, \hat{y}^*)|.$$

Since $L$ is partially Gâteaux-differentiable at $(\hat{x}, \hat{y}^*)$, one has that the left-hand side converges to a limit $\langle \xi, h \rangle$ as $t$ does to 0. Thus

$$\lim_{t \to 0} t^{-1} |H(\hat{x} + th, \hat{y}^*) - H(\hat{x}, \hat{y}^*)| = \langle A^* \hat{y}^* - \xi, h \rangle$$

for all $h$ in $X$.

But $A^* \hat{y}^* \in \partial_1 H(\hat{x}, \hat{y}^*)$, so from Lemma 3.6 one has

$$D_1 H(\hat{x}, \hat{y}^*) = A^* \hat{y}^*.$$

Combining this with the previous equation one finds $\xi = 0$ or $D_1 L(\hat{x}, \hat{y}^*) = 0$. Similarly for $D_2 L(\hat{x}, \hat{y}^*)$.

When $(\hat{x}, \hat{y}^*)$ is a critical point of $L$ then, provided $L$ is $C$-differentiable on a neighborhood of $(\hat{x}, \hat{y}^*)$ and (D1) holds, $\hat{x}$ will be a critical point of $f$ from Lemma 3.7. Similarly, when the analog of (D2) holds with $g$ replacing $h$ then $\hat{y}^*$ will be a critical point of $g$.

The following identities are also important.
Proposition 4.3. Let \((\tilde{x}, \tilde{y}^*)\) be an anomalous critical point of \(L\). Then

\[
\begin{align*}
(\text{i}) & \quad H(\tilde{x}, \tilde{y}^*) + H^*(\tilde{x}, A\tilde{x}) = \langle A\tilde{x}, \tilde{y}^* \rangle, \\
(\text{ii}) & \quad H(\tilde{x}, \tilde{y}^*) + H^*(A^*\tilde{y}^*, \tilde{y}^*) = \langle A\tilde{x}, \tilde{y}^* \rangle.
\end{align*}
\]

(4.13)

Proof. From the definition of polar

\[
H(\tilde{x}, \tilde{y}^*) + H^*(\tilde{x}, A\tilde{x}) \geq \langle A\tilde{x}, \tilde{y}^* \rangle
\]

for all \(y^*\) in \(Y^*\) with equality holding if and only if

\[
A\tilde{x} \in \partial H(\tilde{x}, \tilde{y}^*).
\]

Thus (i) holds as \((\tilde{x}, \tilde{y}^*)\) is an anomalous critical point. (ii) is proven similarly.

When \(L\) is a Lagrangian of type II, the individual relationships in (4.4) have important consequences.

Lemma 4.4. Let \(L\) be a Lagrangian of type II, \(H\) be the corresponding Hamiltonian and \(\mathscr{H} = -H\),

(i) if \(-A^*\tilde{y}^* \in \partial_1 \mathscr{H}(\tilde{x}, \tilde{y}^*)\) then \(g(\tilde{y}^*) \geq f(\tilde{x})\), and

(ii) if \(-A\tilde{x} \in \partial_2 \mathscr{H}(\tilde{x}, \tilde{y}^*)\) then \(f(\tilde{x}) \geq g(\tilde{y}^*)\).

Proof. (i) If \(-A^*\tilde{y}^* \in \partial_1 \mathscr{H}(\tilde{x}, \tilde{y}^*)\) and \(x \in X\), then

\[
\mathscr{H}(x, \tilde{y}^*) \geq \mathscr{H}(\tilde{x}, \tilde{y}^*) - \langle x - \tilde{x}, A^*\tilde{y}^* \rangle.
\]

Thus \(L(x, \tilde{y}^*) \geq L(\tilde{x}, \tilde{y}^*)\) for any \(x\) in \(X\). Therefore \(g(\tilde{y}^*) = L(\tilde{x}, \tilde{y}^*) \geq f(\tilde{x})\).

Similarly for (ii).

Corollary. When \(L\) is a Lagrangian of type II and (4.4) holds then \(f(\tilde{x}) = g(\tilde{y}^*) = L(\tilde{x}, \tilde{y}^*)\).

This follows from (i) and (ii) of Lemma 4.4 and is the same result as Lemma 3.2.

Proposition 4.5. Let \((\tilde{x}, \tilde{y}^*)\) be a \(\partial\)-critical point of \(L\). Then

\[
\begin{align*}
(\text{i}) & \quad \mathcal{H}(\tilde{x}, \tilde{y}^*) + \mathcal{H}^*(-A^*\tilde{y}^*, \tilde{y}^*) = -\langle A\tilde{x}, \tilde{y}^* \rangle \\
(\text{ii}) & \quad \mathcal{H}(\tilde{x}, \tilde{y}^*) + \mathcal{H}^*(\tilde{x}, -A\tilde{x}) = -\langle A\tilde{x}, y^* \rangle.
\end{align*}
\]

(4.14)
Proof. From the definition of \( \mathcal{H}(\cdot, \cdot) \) one has
\[
\mathcal{H}(\cdot, \cdot) = \sup_{x \in X} \{ -\langle A^* y^*, x \rangle + H(x, y^*) \} = -\inf_{x \in X} L(x, y^*) \quad \text{from (4.1)},
\]
\[
= -L(\hat{x}, \hat{y}^*)
\]
as \((\hat{x}, \hat{y}^*)\) is a \(\partial\)-critical point of \(L\). Using (4.1) again one has (i).
Similarly for (ii).

Theorem 4.6. Suppose \(L\) is a Lagrangian of type I, \(A : X \to Y\) is a continuous linear operator and \(H\) is the corresponding Hamiltonian. Suppose \(H\) is u.s.c. on \(X\) for each \(y^*\) in \(Y^*\) and on \(Y^*\) for each \(x\) in \(X\); that \(f(x) = H^o(x, Ax)\) and \(h(y^*) = H^o(A^* y^*, y^*)\) are non-trivial functionals and that the sets
\[
E_{c_1} = \{ x \in X : f(x) \leq c_1 \} \quad \text{and} \quad H_{c_2} = \{ y^* \in Y^* : h(y^*) \leq c_2 \}
\]
are sequentially compact and non-empty for some \(c_1\) and \(c_2\). Then \((\mathcal{F}^1)\) and \((\mathcal{F}^o)\) have solutions.
Proof: From the definitions
\[
f(x) = H^o(x, Ax) = \sup_{y^*} \{ \langle x, A^* y^* \rangle - H(x, y^*) \}.
\]
If \(H\) is u.s.c. on \(X\) for each \(y^*\) in \(Y^*\), then \(f\) is l.s.c. on \(X\). Similarly, \(h\) will be l.s.c. on \(Y^*\).
If \(E_{c_1}\) is sequentially compact and non-empty for some \(c_1\), then since \(f\) is l.s.c. on \(X\), one has that \((\mathcal{F}^1)\) has a solution \(\hat{x}\). Similarly for \((\mathcal{F}^o)\).

5. Finite Dimensional Cases

The theory has a particularly nice form when \(X\) and \(Y\) are finite dimensional. We shall focus most attention in this section on the anomalous dual principles for Lagrangians of type I and on duality for Lagrangians of type II, since ordinary saddle-point duality has been so thoroughly studied. The results described here may be compared with Ekeland [11], where a different theory is developed.
Assume first that one has a Lagrangian \(L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}\) of type I and define \(f\) and \(h\) by (2.1) and (2.4) so that
\[
f(x) = \sup_{y^*} L(x, y), \quad \text{and} \quad h(y) = \sup_{x \in X} L(x, y).
\]
Let $D_f(D_h)$ be the sets in $\mathbb{R}^n(\mathbb{R}^m)$ on which $f(h)$ is finite and assume:

(L1) For each $x$ in $D_f$, there is a unique $\xi(x)$ in $\mathbb{R}^m$ such that
$$f(x) = L(x, \xi(x)). \quad (5.3)$$

(L2) For each $y$ in $D_h$, there is a unique $\eta(y)$ in $\mathbb{R}^n$ such that
$$h(y) = L(\eta(y), y). \quad (5.4)$$

**Lemma 5.1.** Assume that $L$ as above obeys (L1)–(L2). Suppose $(\hat{x}, \hat{y})$ is a critical point of $L$, that $L$ is $C^k$ on an open neighborhood of $(\hat{x}, \hat{y})$, that $\hat{x}$ is an interior point of $D_f$ and $\hat{y}$ is an interior point of $D_h$ and that $\xi, \eta$ are $C^k$-functions on neighborhoods of $\hat{x}, \hat{y}$, respectively, with $k \geq 1$. Then $\hat{x}$ is a critical point of $f$ and $\hat{y}$ is a critical point of $h$.

**Proof.** Applying the chain rule to (5.3) one has
$$Df(\hat{x}) = D_1L(\hat{x}, \xi(\hat{x})) + D_2L(\hat{x}, \xi(\hat{x})) D\xi(\hat{x}) = 0$$
since $\xi(\hat{x}) = \hat{y}$ and $(\hat{x}, \hat{y})$ is a critical point of $L$.

Similarly for $h$.

When $f: \mathbb{R}^n \to \mathbb{R}$ has a critical point $\hat{x}$ and $f$ is $C^2$ on an open neighborhood of $\hat{x}$, the type of the critical point may be described by the eigenvalues of the Hessian matrix $D^2f(\hat{x})$.

The critical point $\hat{x}$ is said to be non-degenerate if
$$\det D^2f(\hat{x}) \neq 0.$$  

When the critical point $\hat{x}$ is non-degenerate, the (Morse) index $i(\hat{x})$ of $\hat{x}$ is defined to be the number of negative eigenvalues of $D^2f(\hat{x})$.

$\hat{x}$ will be a local minimum iff $i(\hat{x}) = 0$ and it will be a local maximum iff $i(\hat{x}) = n$.

The surprising fact is that if $(\hat{x}, \hat{y})$ is a critical point of $L$ and $\hat{x} = \xi(\hat{x})$ and $\hat{y} = \eta(\hat{y})$ are non-degenerate critical points of $f$, $(h)$, respectively, then their indices are related. To prove this one needs the following lemma.

**Lemma 5.2.** Let the assumptions of Lemma 5.1 hold with $k \geq 2$, and that $\hat{y} = \xi(\hat{x})$ and $\hat{x} = \eta(\hat{y})$. Let
$$D^2L(\hat{x}, \hat{y}) = \begin{pmatrix} A & A \\ A^* & B \end{pmatrix} \quad (5.5)$$

with $A$ and $B$ non-singular. Then
$$D^2f(\hat{x}) = A - AB^{-1}A^*$$
and
$$D^2h(\hat{y}) = B - A^*A^{-1}A \quad (5.6)$$
Proof: From the definition of $\xi$, one has that there is a neighborhood $U$ of $\hat{x}$ in $\mathbb{R}^n$ such that

$$D_2L(x, \xi(x)) \equiv 0 \quad \text{for } x \text{ in } U.$$ 

Here $D_2$ represents differentiation with respect to $y$.

Differentiating each component of this with respect to $x_j$ one has

$$\frac{\partial^2 L}{\partial x_j \partial y_k} (x, \xi(x)) + \sum_{l=1}^m \frac{\partial^2 L}{\partial y_k \partial y_l} (x, \xi(x)) \frac{\partial \xi_l}{\partial x_j} (x) = 0 \quad (5.7)$$

for all $x$ in $U$ and $1 \leq j \leq n$, $1 \leq k \leq m$.

Let $E(x) = (E_{ij}(x))$ have components $E_{ij}(x) = \partial \xi_j / \partial x_i$. Then putting $x = \hat{x}$ in (5.7) and using (5.5) one has

$$A^* + BE(\hat{x}) = 0. \quad (5.8)$$

Similarly, using the definition of $q$ and letting $N(y)$ be the matrix whose $(k, l)$th entry is $N_{kl}(y) = \partial q_k(y) / \partial y_l$ one finds that

$$A + AN(\hat{y}) = 0. \quad (5.9')$$

Moreover, for all $x$ in $U$ one has

$$\frac{\partial f}{\partial x_i} (x) = \frac{\partial L}{\partial x_i} (x, \xi(x)),$$

so that

$$\frac{\partial^2 f}{\partial x_i \partial x_j} (x) = \frac{\partial^2 L}{\partial x_i \partial x_j} (x, \xi(x)) + \sum_{k=1}^m \frac{\partial^2 L}{\partial x_i \partial y_k} (x, \xi(x)) \frac{\partial \xi_k}{\partial x_j} (x).$$

Evaluating this at $\hat{x}$ one finds

$$\frac{\partial^2 f}{\partial x_i \partial x_j} (\hat{x}) = a_{ij} + \sum_{k=1}^m A_{ik} E_{kj}(\hat{x}),$$

where $A$ has entries $a_{ij}$ and $A_{ik}$ is the $(i, k)$th entry of $A$. Thus in matrix form, using (5.8), one has

$$D_2 f(\hat{x}) = A - AB^{-1}A^*.$$

Similarly, $D^2 h(\hat{y}) = B + A^* N(\hat{y}) = B - A^* A^{-1} A$.

Theorem 5.3. Let the assumptions of Lemma 5.2 hold with $n - m$ and assume that either $A$ or $B$ is negative-definite and that $A$ is non-singular.
Then \( \hat{x} \) is a non-degenerate critical point of \( f \) iff \( \hat{y} \) is a non-degenerate critical point of \( h \). When this holds, then \( i(\hat{x}) = i(\hat{y}) \).

**Proof.** Assume \( A \) is negative-definite. Then from the theory of diagonalization of quadratic forms, there is a non-singular matrix \( V \) such that

\[
V^*AV = -I
\]

and

\[
V^*AB^{-1}A^*V = D,
\]

where \( D \) is a diagonal matrix and \( I \) is the \( n \times n \) identity.

Thus \( D^2f(\hat{x}) \) is congruent to the matrix \(- (I + D)\) and from Sylvester's law \( \hat{x} \) is non-degenerate iff \((I + D)\) is non-singular and the Morse index of \( \hat{x} \) is the number of negative eigenvalues of \(-(I + D)\).

From (5.10) one has \( A^{-1} = -VV^* \) and \( B = A^*VD^{-1}V^*A \). Thus \( D^2h(\hat{y}) = A^*V(D^{-1} + I)V^*A \).

Since \( V^*A \) is non-singular, this says \( D^2h(\hat{y}) \) is congruent to \((D^{-1} + I)\) and so \( \hat{y} \) is non-degenerate iff \((I + D^{-1})\) is non-singular and the Morse index of \( \hat{y} \) is the number of negative eigenvalues of \( I + D^{-1} \).

Thus \( \hat{x} \) and \( \hat{y} \) are non-degenerate iff \(-1\) is not an eigenvalue of \( D \) and the Morse index of both \( \hat{x} \) and \( \hat{y} \) is the number of eigenvalues of \( D \) which are greater than \(-1\).

When \( B \) is negative-definite, one interchanges the roles of \( A \) and \( B \).

This result may be applied to problems where the anomalous dual variational principles are well defined.

**Corollary 1.** Suppose \((\hat{x}, \hat{y})\) is a critical point of a Lagrangian \( L \) and the assumptions of the theorem hold. Then a non-degenerate critical point \( \hat{x} \) of \( f \) is a local minimum (maximum) of \( f \) iff \( \hat{y} \) is non-degenerate and a local minimum (maximum) of \( h \).

**Proof.** From the theorem \( \hat{x} \) is non-degenerate iff \( \hat{y} \) is non-degenerate and 
\( i(\hat{x}) = i(\hat{y}) \). Thus \( i(\hat{x}) = 0 \) iff \( i(\hat{y}) = 0 \) so local minima correspond. Similarly if \( i(\hat{x}) = i(y) = n \).

It is instructive to see how the usual dual variational principles fit into this framework. Suppose \( g: \mathbb{R}^n \rightarrow \mathbb{R} \) is defined by

\[
g(y) = \inf_{x \in \mathbb{R}^n} L(x, y)
\]

and \( D_g \) is the essential domain of \( g \). Consider (L3): For each \( y \) in \( D_g \), there is a unique \( \mu(y) \) in \( \mathbb{R}^n \) such that

\[
g(y) = L(\mu(y), y).
\]
Then if (L3) replaces (L2) and \( g, \mu \) replace \( h, \eta \) the analogs of Lemmas 5.1 and 5.2 hold.

Usually for saddle-point problems such as this, one assumes \( L \) is strictly convex in \( x \) near \( \hat{x} \) and strictly concave in \( y \) near \( \hat{y} \) when \((\hat{x}, \hat{y})\) is a critical point of \( L \). This implies \( A \) is positive-definite and \( B \) is negative-definite. Thus \( D^2 f(\hat{x}) \) will be positive-definite or \( \hat{x} \) is necessarily a local minimum. Similarly, \( D^2 h(\hat{y}) \) is negative-definite and \( \hat{y} \) must be a local maximum.

When \( L: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) is a Lagrangian of type II, assume

\[
f(x) = \inf_{y \in \mathbb{R}^n} L(x, y)
\]

and let \( g \) be defined by (5.11).

Let \( D_f \) and \( D_g \) be defined as before. Once again if (L3) replaces (L2) and \( g, \mu \) replace \( h, \eta \) the analogs of Lemmas 5.1 and 5.2 hold. The analog of Theorem 5.3 is the following.

**Theorem 5.4.** Assume \( L \) obeys (L1) and (L3) and \( n = m \). Let \((\hat{x}, \hat{y})\) be a critical point of \( L \) and the assumptions of Lemma 5.2 hold with \( g, \mu \) replacing \( h, \eta \). Assume either \( A \) or \( B \) is positive-definite and \( A \) is non-singular. Then \( \hat{x} \) is a non-degenerate critical point of \( f \) iff \( \hat{y} \) is a non-degenerate critical point of \( h \). In this case \( i(\hat{x}) = i(\hat{y}) \).

**Proof.** The only difference from the proof of Theorem 5.3 is that now \( V^* A V = I \) and thus \( D^2 f(\hat{x}) \) is congruent to \( I - D \).

The index of \( \hat{x} \) is the number of eigenvalues of \( D \) less than 1.

Moreover, the corollary to Theorem 5.3 also applies here so local minima and maxima are in correspondence.

### 6. Correspondence of Critical Points

In the preceding sections, we saw that if \((\hat{x}, \hat{y}^*)\) is a critical point of the Lagrangian \( L \), then, subject to certain technical conditions, \( \hat{x} \) and \( \hat{y}^* \) are critical points of a functional and its dual. Moreover, when \( X \) and \( Y^* \) were of the same finite dimension and one considered either anomalous dual principles, or Lagrangians of type II, then the critical points were of the same type.

In this section we shall try to obtain other results along these lines when \( X \) and \( Y^* \) are not necessarily finite dimensional. Theorem 3.4 already embodies one such result and says that for Lagrangians of type II their global minima are in correspondence.

To describe these results we shall need certain conditions on \( L \).
(L1) \( L(x, y^*) > -\infty \) for each \( x \) in \( X \), \( y^* \) in \( Y^* \).

(L2) \( L(\cdot, y^*) \) is l.s.c. on \( X \) for each \( y^* \) in \( Y^* \) and \( L(x, \cdot) \) is l.s.c. on \( Y^* \) for each \( x \) in \( X \).

(L3) \( L(\cdot, y^*) \) is convex on \( X \) for each \( y^* \) in \( Y^* \).

(L4) \( L(x, \cdot) \) is convex on \( Y^* \) for each \( x \) in \( X \).

As in Section 4, let \( \mathcal{H}: X \times Y^* \rightarrow \mathbb{R} \) be defined by

\[
\mathcal{H}(x, y^*) = L(x, y^*) - \langle Ax, y^* \rangle.
\] (6.1)

One observes that if \( \varphi: X \rightarrow Y \) is continuous, then \( \mathcal{H} \) obeys (L1)–(L4) whenever \( L \) does.

Suppose \( G: X \rightarrow \mathbb{R} \) is a given functional and \( \partial G: X \rightarrow 2^{X^*} \) is its subdifferential. \( \partial G \) is said to be set-valued u.s.c. at a point \( \hat{x} \) in \( X \) if for any neighborhood \( V \) of \( \partial G(\hat{x}) \) in \( X^* \) there is a neighborhood \( U \) of \( \hat{x} \) such that

\[
\partial G(U) \subseteq V.
\]

Consider the continuity conditions:

(L5) \( \partial_1 L(\cdot, y^*) \) is set-valued u.s.c. on \( X^{**} \) for each \( y^* \) in \( Y^* \).

(L6) \( \partial_2 L(\cdot, \cdot) \) is set-valued u.s.c. on \( Y^{**} \) for each \( x \) in \( X \).

A point \( \hat{x} \) in \( X \) is said to be a local minimum for \( G: X \rightarrow \mathbb{R} \) if there is an open neighborhood \( U \) of \( \hat{x} \) such that \( x \) in \( U \) implies \( G(x) \geq G(\hat{x}) \).

**Proposition 6.1.** Suppose \( L \) is a Lagrangian of type II which obeys (L1)–(L3) and (L5). Assume \( (\hat{x}, \hat{y}^*) \) is a \( \partial \)-critical point of \( L \) on \( X \times Y^* \) and \( \Lambda \) is a homeomorphism. Assume (L7): there is a neighborhood \( V^* \) of \( \hat{y}^* \) in \( Y^* \) such that \( \partial_1 L^*(-A^*y^*, \hat{y}^*) \neq \emptyset \) for all \( y^* \in V^* \cap \text{dom } g \). If \( \hat{x} \) is a local minimum of \( f \) on \( X \), then \( \hat{y}^* \) is a local minimum of \( g \) on \( Y^* \).

**Proof.** Since \( (\hat{x}, \hat{y}^*) \) is a \( \partial \)-critical point of \( L \), from (4.14) one has

\[
\mathcal{H}(\hat{x}, \hat{y}^*) + \mathcal{H}^*(-A^*\hat{y}^*, \hat{y}^*) = -\langle A\hat{x}, \hat{y}^* \rangle.
\]

Thus from (L3) and Proposition I.5.1 of [4] one has

\[
\{\hat{x}\} \subseteq \partial_1 \mathcal{H}^*(-A^*\hat{y}^*, \hat{y}^*).
\]

Let \( U \) be a neighborhood of \( \partial_1 \mathcal{H}^*(-A^*\hat{y}^*, \hat{y}^*) \) in \( X^{**} \). \( U \) will be a neighborhood of \( \hat{x} \). Then from (L5) there is a neighborhood \( W^* \) of \( -A^*\hat{y}^* \) in \( X^* \) such that

\[
\partial_1 \mathcal{H}^*(W^*, \hat{y}^*) \subseteq U.
\]
Let \( V_1^* = (A^*)^{-1}(-W^*) \). Then \( V_1^* \) is a neighborhood of \( \hat{y}^* \) in \( y^* \) as \( A^* \) is a homeomorphism. Choose \( y^* \) in \( V_1^* \). If \( \partial_1 \mathcal{H}^*(-A^*y^*, \hat{y}^*) = \emptyset \) then \( g(y^*) = +\infty \) by assumption. If \( x \in \partial_1 \mathcal{H}^*(-A^*y^*, \hat{y}^*) \), then \( x \in U \) and since \( \mathcal{H}^*(\cdot, y^*) \) is convex and l.s.c. on \( X \) one has \(-A^*y^* \in \partial_1 \mathcal{H}^*(x, \hat{y}^*)\). Thus from (i) of Lemma 4.4 one has \( g(y^*) \geq J'(x) > f(\hat{x}) = g(\hat{y}^*) \). Hence \( \hat{y}^* \) is a local minimum of \( g \).

**Corollary.** Suppose \( L \) is a Lagrangian of type II which obeys (L1)–(L2), (L4) and (L6). Assume \((\hat{x}, \hat{y}^*)\) is a \( \partial \)-critical point of \( L \) and \( A \) is a homeomorphism. Assume (L8): there is a neighborhood \( U \) of \( x \) in \( X \) such that \( \partial_2 L^2(\hat{x}, Ax) \neq \emptyset \) for all \( x \) in \( U \cap \text{dom} \, f \). If \( \hat{y}^* \) is a local minimum of \( g \) then \( \hat{x} \) is a local minimum of \( f \).

**Proof.** Just interchange the roles of \( x, y^* \) in the proposition.

**Theorem 6.2.** Suppose \( L \) is a Lagrangian of type II which obeys (L1)–(L8), let \((\hat{x}, \hat{y}^*)\) be a \( \partial \)-critical point of \( L \) and \( A \) be a homeomorphism. Then \( \hat{x} \) is a local minimum of \( f \) on \( X \) iff \( \hat{y}^* \) is a local minimum of \( g \) on \( Y^* \).

**Proof.** This follows from Proposition 6.1 and its corollary.

In a particular case, a result similar to this was proven by Toland (see [7, Theorem 4.3]).

The conditions (L1)–(L8) simplify considerably when \( f \) has the form (2.6) or \( L \) has one of the forms (2.21). In particular, each of the conditions (L5)–(L8) involve only one simpler functional.

A similar analysis may be performed for anomalous critical points and is summarized by the following theorem. When we say \( H \) obeys (L1), we mean \( H \) obeys (L1) with \( L \) replaced by \( H \). In (L7) \( g \) is replaced by \( h \).

**Theorem 6.3.** Let \((\hat{x}, \hat{y}^*)\) be an anomalous critical point of \( L \) on \( X \times Y^* \), \( A \) be a homeomorphism and \( H \) be the corresponding Hamiltonian.

If \( H \) obeys (L1)–(L3), (L5) and (L7) and \( \hat{y}^* \) is a local minimum of \( h \) on \( Y^* \) then \( \hat{x} \) is a local minimum of \( f \) on \( X \).

If \( H \) obeys (L1)–(L2), (L4), (L6) and (L8) and \( \hat{x} \) is a local minimum of \( f \) on \( X \), then \( \hat{y}^* \) is a local minimum of \( h \) on \( Y^* \).

When \( H \) obeys (L1)–(L8) then \( \hat{x} \) is a local minimum of \( f \) on \( X \) iff \( \hat{y}^* \) is a local minimum of \( h \) on \( Y^* \).

**Proof.** Since \((\hat{x}, \hat{y}^*)\) is an anomalous critical point of \( L \) one has

\[
\hat{x} \in \partial_1 H^*(A^*\hat{y}^*, \hat{y}^*).
\]

Let \( U \) be a neighborhood of \( \partial_1 H^*(A^*\hat{y}^*, \hat{y}^*) \) in \( X^{**} \). Then, since \( H \) obeys (L5), there is a neighborhood \( W^* \) of \( A^*\hat{y}^* \) in \( X^* \) such that

\[
\partial_1 H^*(W^*, \hat{y}^*) \subset U.
\]
Let $V^* = (A^*)^{-1}(W^*)$, then $V^*$ is a neighborhood of $\tilde{y}^*$ and choose $y^* \in V^*$.

If $\partial_1 H^*(A^* \tilde{y}^*, \tilde{y}^*) = \emptyset$, then $h(y^*) = \infty$ from (L7). If $x \in \partial_1 H^*(A^* \tilde{y}^*, \tilde{y}^*)$ then $x \in U$ and, from (L2), (L3),

$$A^* \tilde{y}^* \in \partial_1 H(x, \tilde{y}^*).$$

Thus, from (ii) of Lemma 4.1,

$$h(\tilde{y}^*) \leq f(x).$$

But $f(\tilde{x}) - h(\tilde{y}^*) \leq h(y^*)$ so $f(\tilde{x}) \leq f(x)$ or $\tilde{x}$ is a local minimum of $f$ on $X$.

Similarly for the second part and then the last statement is a combination of the first two.

7. The Difference of Two Convex Functionals

The theory described in the preceding sections takes a particularly interesting form when $f$ is the difference of two convex functionals.

Consider first the case where $k_1 : Y \to \bar{R}$ and $k_2 : X \to \bar{R}$ are proper, convex and l.s.c. functionals, $A : X \to Y$ is a continuous linear operator and

$$f(x) = k_1(Ax) - k_2(x). \quad (7.1)$$

The Lagrangian of type I associated with this problem is

$$L(x, y^*) = \langle Ax, y^* \rangle - k_1^*(y^*) - k_2(x) \quad (7.2)$$

and thus the Hamiltonian is

$$H(x, y^*) = k_1^*(y^*) + k_2(x) \quad (7.3)$$

and is convex on $X \times Y^*$, being the sum of two convex functionals.

The anomalous dual functional associated with this is

$$h(y^*) = k_1^*(A^* y^*) - k_2^*(y^*). \quad (7.4)$$

This will be non-trivial if $y^* \in \text{dom } k_1^*$ and $A^* y^* \in \text{dom } k_2^*$ and the anomalous dual problem is to minimize $h$ on $Y^*$.

A point $(\tilde{x}, \tilde{y}^*)$ will be an anomalous critical point of the Lagrangian (7.2) provided

$$A \tilde{x} \in \partial k_1^*(\tilde{y}^*)$$

and

$$A^* \tilde{y}^* \in \partial k_2(\tilde{x}). \quad (7.5)$$
For this problem one may also construct Lagrangians of type II. In general, \( L: X \times X^* \rightarrow \mathbb{R} \) is given by

\[
L(x, x^*) = -(x, x^*) + k_1(Ax) + k_2^*(x^*). \tag{7.6}
\]

The dual problem is to minimize \( g \) on \( X^* \) where

\[
g(x^*) = k_2^*(x^*) - (k_1 \circ A)^*(x^*) \tag{7.7}
\]

and a point \((\hat{x}, \hat{x}^*)\) is a \( \partial \)-critical point of \( L \) provided

\[
\hat{x}^* \in \partial(k_1 \circ A)(\hat{x}) = A^* \partial k_1(A\hat{x})
\]

and

\[
\hat{x} \in \partial k_2^*(\hat{x}^*). \tag{7.8}
\]

When \( A^* \) is surjective, then \( L_2: X \times Y^* \rightarrow \mathbb{R} \) given by

\[
L_2(x, y^*) = -(Ax, y^*) + k_1(Ax) + k_2^*(A^*y^*) \tag{7.9}
\]

is also a Lagrangian of type II.

Now the dual problem is to minimize \( g_1 \) on \( Y^* \) where

\[
g_1(y^*) = k_2^*(A^*y^*) - \sup_x [\langle Ax, y^* \rangle - k_1(Ax)]
\]

\[
= k_2^*(A^*y^*) - k_1^*(y^*) \tag{7.10}
\]

if \( A \) is surjective.

Thus, when \( A \) and \( A^* \) are surjective, this dual problem is exactly the same as that found for the anomalous dual problem associated with the Lagrangian (7.2) of type I. In deriving (7.4), however, we did not have to assume that \( A \) or \( A^* \) was surjective.

Moreover, when \( A \) is a homeomorphism, one sees that \((\hat{x}, \hat{x}^*)\) is a \( \partial \)-critical point of the Lagrangian \( L \) defined by (7.6) provided

\[
(A^*)^{-1} \hat{x}^* \in \partial k_1(A\hat{x})
\]

and

\[
\hat{x}^* \in \partial k_2^*(\hat{x})
\]

since \( k_2 \) is proper and convex.

Comparing this with (7.5) one sees that if

\[
A\hat{x} \in \partial k_1^*(\hat{y}^*),
\]

and

\[
\hat{x} \in \partial k_2^*(\hat{x}^*),
\]

and

\[
A^*\hat{y}^* = \hat{x}^*,
\]

DUALITY FOR NON-CONVEX VARIATIONAL PRINCIPLES 107
then \((\hat{x}, \hat{y}^*)\) will be an anomalous critical of the Lagrangian (7.2) and 
\((\hat{x}, \hat{x}^*)\) will be a \(\partial\)-critical point of the Lagrangian (7.6). Thus there is a 
correspondence between the critical points of these two Lagrangians.

**Theorem 7.1.** Suppose \(f\) is given by (7.1) and \(A, A^*\) are surjective. 
Assume \(k_1, k_2\) are proper, convex functionals and \(h\) is given by (7.4). Then 
\[a = \inf_x f(x) = a^\ominus = \inf_y h(y^*).\]

**Proof.** From (7.4) one has that 
\[h(y^*) = k_2^*(A^*y^*) - k_1^*(y^*)\]
\[= k_2^*(A^*y^*) - \sup_x |\langle Ax, y^* \rangle - k_1(Ax)|\quad\text{as } A\text{ is surjective},\]
\[= \inf_x L_2(x, y^*).\]

Now 
\[a^\ominus = \inf_y h(y^*) = \inf_y \inf_x L_2(x, y^*) = \inf_x f(x) = a.\]

One consequence of this theorem is that, under these conditions on \(f\) and 
\(A, (\mathscr{F})\) has a finite value if and only if \((\mathscr{F} a^\ominus)\) has a finite value.

One might also consider the case where 
\[f(x) = k_1(x) - k_2(Ax), \quad \text{(7.11)}\]
with \(k_1, k_2\) and \(A\) as before.

The Lagrangian of type I associated with this functional is 
\[L: X \times X^* \to \mathbb{R}, \text{ where}\]
\[L(x, x^*) = \langle x, x^* \rangle - k_2^*(x^*) - k_1(Ax). \quad \text{(7.12)}\]

The associated Hamiltonian is 
\[H(x, x^*) = k_2^*(x^*) + k_1(Ax) \quad \text{(7.13)}\]
and \(H\) is convex on \(X \times X^*\).

The anomalous dual functional is 
\[h(x^*) = -k_2^*(x^*) + (k_1 \circ A)^* (x^*). \quad \text{(7.14)}\]

The Lagrangian of type II associated with \(f\) is 
\[L_1(x, y^*) = \langle Ax, y^* \rangle + k_2(x) + k_1^*(-y^*). \quad \text{(7.15)}\]

The dual problem is to minimize \(g\) on \(Y^*\), where
DUALITY FOR NON-CONVEX VARIATIONAL PRINCIPLES

\[ g(y^*) = \inf_x L_1(x, y^*) \]
\[ = k_1^*(-y^*) - k_2^*(-A*y^*). \tag{7.16} \]

For this Lagrangian the associated Hamiltonian is
\[ H(x, y^*) = -k_2(x) - k_1^*(-y^*) \]
and \( H \) is concave on \( X \times Y^* \).

At a \( \partial \)-critical point of \( L_1 \) one has that
\[ Ax \in \partial k_1^*(-y^*) \]
and
\[ -A^*y^* \in \partial k_2(x). \tag{7.17} \]

This time the dual problem using the Lagrangian of type II is, in general, somewhat simpler than the anomalous dual problem for \( h \).

When \( A \) is \( 1 \)-1 and onto, one has that
\[ (k_1 \circ A)^*(x^*) = k_1^*((A^{-1})^*x^*), \]
so
\[ h(x^*) = k_1^*((A^{-1})^*x^*) - k_2^*(x^*) \]
\[ = g(-(A^*)^{-1}x^*). \tag{7.18} \]

Thus the two problems \((\partial^\infty)\) and \((\partial^\infty)\) of minimizing \( h \) on \( X^* \) and \( g \) on \( Y^* \) are equivalent in this case. This leads to the following result.

**Theorem 7.2.** Suppose \( f \) is given by (7.11) and \( A \) is \( 1 \)-1 and onto. Assume \( k_1, k_2 \) are proper convex functionals and \( h \) is defined by (7.14). Then
\[ a = \inf_x f(x) = \alpha^\infty = \inf_x h(x^*). \]

**Proof:** One has \( a = \alpha^* = \inf_y g(y^*) \) from Theorem 3.3. From (7.18) one has \( g(y^*) = h(-A^*y^*) \) for all \( y^* \). Thus \( \inf_y g(y^*) = \inf_x h(x^*) \) as \( A^* \) is surjective so \( a = \alpha^* = \alpha^\infty \) as required.

Suppose now that \( X, Y \) are Banach spaces and \( k_1, k_2 \) and \( A \) obey
\[ \begin{align*}
(a) \quad & k_1(y) \geq \xi \| y \|^\gamma + C_1, \\
(b) \quad & k_2(x) \leq \sigma \| x \|^\alpha + C_2, \tag{7.19} \\
(c) \quad & \| Ax \| \geq v \| x \|, \quad \| A^*y^* \| \geq v \| y^* \|
\end{align*} \]
for all \( x \in X, y \in Y \) and \( y^* \) in \( Y^* \).
Here $\xi, \nu, \gamma, \sigma$ and $\mu$ are positive constants, and $C_1$ and $C_2$ are constant. When these inequalities hold, the functional $f$ defined by (7.1) obeys

$$f(x) \geq \xi \nu^\gamma \|x\|^{\gamma} - \sigma \|x\|^\mu + C_3,$$

where $C_3$ is a constant.

**Theorem 7.3.** Let $k_1, k_2$ be proper convex functionals, $f$ be defined by (7.1), $h$ by (7.4) and $X, Y$ be Banach spaces. Assume (7.19) holds with $\gamma > \mu \geq 1$. Then

$$\liminf_{\|x\| \to \infty} f(x) = \liminf_{\|y^*\| \to \infty} h(y^*) = +\infty.$$

**Proof.** From (7.20) with $\gamma > \mu \geq 1$ one sees that

$$\liminf_{\|x\| \to \infty} f(x) = +\infty.$$

Assume first that $\mu > 1$; then $K_1(y) = \xi \|y\|^{\gamma} + C_1$ is a convex functional of $Y$ and $K_2^*(y^*) = D_1 \|y^*\|^{\gamma} - C_1$, where $\gamma' = \gamma/(\gamma - 1)$, $D_1 = (\gamma')^{-1} (\xi \nu)^\gamma$ and $a = (1 - \gamma)^{-1}$. Here $\|y^*\|$ is the usual dual norm on $Y^*$.

Similarly $K_2(x) = \sigma \|x\|^\mu + C_2$ is convex on $X$ and $K_2^*(x^*) = D_2 \|x^*\|^\mu - C_2$, where $\mu' = \mu/(\mu - 1)$, $D_2 = (\mu')^{-1} (\sigma \nu)^\mu$ and $b = (1 - \mu)^{-1}$.

Since $k_1(y) \geq K_1(y)$ for all $y \in Y$ one has $K_1^*(y^*) \geq k_1^*(y^*)$ for all $y^*$ in $Y^*$ and similarly for $K_2, k_2$. Thus

$$h(y^*) = k_2^*(A^* y^*) - k_1^*(y^*)$$

$$\geq K_2^*(A^* y^*) - K_1^*(y^*)$$

$$= D_2 \|A^* y^*\|^{\mu'} - D_1 \|y^*\|^{\gamma'} + C_3. \quad (7.21)$$

Now $\gamma > \mu$ implies $\mu' > \gamma'$ so $\liminf_{\|y^*\| \to \infty} h(y^*) = \infty$ from this equation and (7.19c).

When $\mu = 1$

$$K_2^*(x^*) = -C_2 \quad \text{if } \|x^*\| \leq \sigma,$$

$$= \infty \quad \text{if } \|x^*\| > \sigma. \quad (7.22)$$

and thus $h$ will obey $h(y^*) = +\infty$ if $\|A^* y^*\| > \sigma$.

**Corollary.** Assume $f, h, k_1, k_2$ as above and (7.19) holds with $\gamma = \mu > 1$ and $\xi \nu^\gamma > \sigma$. Then

$$\liminf_{\|x\| \to \infty} f(x) = \liminf_{\|y^*\| \to \infty} h(y^*) = +\infty.$$
DUALITY FOR NON-CONVEX VARIATIONAL PRINCIPLES

Proof. When \( \gamma = \mu \), from (7.20) one sees that \( \liminf_{\|x\| \to \infty} f(x) = +\infty \).

From (7.21) and (7.19) one has, when \( \gamma = \mu > 1 \),

\[
h(\gamma^*) \geq (D_2 \gamma - D_1) \|\gamma^*\|^\gamma + C\gamma.
\]

From the formulae for \( D_1, D_2 \) one finds

\[
D_2 \gamma - D_1 = (\gamma')^{-1} \gamma^a (\sigma^a \gamma - \xi^a),
\]

where \( a = (1 - \gamma)^{1/\gamma} \). This is positive provided \( \xi^a \gamma > 0 \) as assumed.

This theorem and its corollary say, essentially, that when \( k_1, k_2, A \) obey estimates such as (7.19) which guarantee that \( f \) is coercive, then \( h \) will also be coercive. One might, thus, expect that such estimates will also guarantee that \((\varphi, \psi)\) has a solution and the Lagrangian \( L \) has an anomalous critical point. The following theorem indicates one such result.

**Theorem 7.4.** Let \( X, Y \) be reflexive Banach spaces; \( k_1, k_2 \) be proper, convex functionals and suppose (7.19) holds with either \( \gamma > \mu \geq 1 \) or \( \gamma = \mu > 1 \) and \( \xi^a \gamma > 0 \). Let \( f \) defined by (7.1) and \( h \) by (7.4) be weakly lower semi-continuous, then \((\varphi, \psi)\) and \((\varphi^\infty, \psi^\infty)\) have solutions. Let \( \hat{x} \) be a solution of \((\varphi, \psi)\), \( \hat{\psi}^* \) be a solution of \((\varphi^\infty, \psi^\infty)\) then

1. \( \alpha = f(\hat{x}) = h(\hat{\psi}^*) = \alpha^\infty > -\infty \) and
2. \((\hat{x}, \hat{\psi}^*)\) is an anomalous critical point of \( L \) with \( L(\hat{x}, \hat{\psi}^*) = \alpha = \alpha^\infty \).

**Proof:** Theorem 7.3 and its corollary imply that \( f \) and \( h \) are coercive on \( X, Y^* \), respectively. Since \( X, Y^* \) are reflexive, and \( f, h \) are w.l.s.c. then \( f, h \) attain their infima on \( X, Y^* \), and this infimum is finite.

When (7.19c) holds, \( A \) and \( A^* \) are surjective so \( h(\gamma^*) = g_1(\hat{\psi}^*) \), where \( g_1 \) is given by (7.10).

If \( \hat{x} \) is a solution of \((\varphi, \psi)\) and \( \hat{\psi}^* \) is a solution of \((\varphi^\infty, \psi^\infty)\) then \((\hat{x}, \hat{\psi}^*)\) is a global minimum and a \( \partial \)-critical point of \( L \) defined by (7.9). Hence, from Lemma 3.2,

\[
\alpha = f(\hat{x}) = h(\hat{\psi}^*) = \alpha^\infty > -\infty.
\]

This \( \partial \)-critical point obeys

\[
A\hat{x}^* \in \partial(k_2^* \circ A^*)(\hat{\psi}^*) = A^\partial k_2^*(A^*\hat{\psi}^*)
\]

and

\[
A^*\hat{\psi}^* \in \partial(k_1 \circ A)(\hat{x}) = A^\partial k_1(A\hat{x}).
\]
Since $A$ and $A^*$ are 1–1, one has
\[ \hat{x} \in \partial k_2^+(A^*\hat{y}^*) \]
and
\[ \hat{y}^* \in \partial k_1(A\hat{x}). \]
Thus (7.5) holds as $k_1, k_2$ are proper convex functionals so $(\hat{x}, \hat{y}^*)$ is an anomalous critical point of the Lagrangian defined by (7.2).

When $X, Y$ are Banach spaces and $f$ is given by (7.11) one has analogous results. Assume now that $k_1, k_2, A$ obey.

\begin{align*}
(a) \quad & k_2(x) \geq \xi \|x\|^\gamma + C_1, \\
(b) \quad & k_1(y) \leq \sigma \|y\|^\mu + C_2, \\
(c) \quad & \|Ax\| \leq v\|x\|, \quad \|A^*y^*\| \leq v\|y^*\|
\end{align*}

for all $x$ in $X$, $y$ in $Y$ and $y^*$ in $Y^*$. Here $\xi, v, \mu, \sigma$ and $\gamma$ are positive constants and $C_1, C_2$ are constants.

When (7.23) holds and $f$ is defined by (7.11) then
\[ f(x) \geq \xi \|x\|^\gamma - \sigma v^\mu \|x\|^\mu + C_3, \]
where $C_3$ is a constant.

**Theorem 7.5.** Let $k_1, k_2$ be proper convex functionals, $X, Y$ be Banach spaces and $f, g$ be defined by (7.11), (7.16), respectively. Assume (7.23) holds with either $\gamma > \mu \geq 1$ or $\gamma = \mu > 1$ and $\xi > \sigma v^\gamma$, then $\liminf_{\|x\| \to \infty} f(x) = \liminf_{\|y^*\| \to \infty} g(y^*) = +\infty$.

**Proof.** From (7.24), under the assumed conditions, one sees that
\[ \liminf_{\|x\| \to \infty} f(x) = +\infty. \]
Define $K_2(x) = \xi \|x\|^\gamma + C_1$ and $K_1(y) = \sigma \|y\|^\mu + C_2$ then $k_2(x) \geq K_2(x)$ and $k_1(Ax) \leq K_1(Ax)$ for all $x$ in $X$. Just as before,
\[ K_2^x(x^*) = D_1 \|x^*\|^\gamma - C_1 \]
and
\[ K_1^y(y^*) = D_2 \|y^*\|^\mu - C_2 \quad \text{if} \quad \gamma, \mu > 1. \]
Thus
\[ g(-y^*) = k_1(y^*) - k_2(A^*y^*) \]
\[ \geq D_2 \| y^* \|^\gamma - D_1 \| A^*y^* \|^{\gamma'} + C_3 \]
\[ \geq D_2 \| y^* \|^\gamma - D_1 v^{\gamma'} \| y^* \|^{\gamma'} + C_3. \]

When \( \gamma > \mu > 1 \), then \( \mu' > \gamma' > 1 \) so
\[ \liminf_{y^* \to -\infty} g(y^*) = +\infty. \]

When \( \gamma = \mu > 1 \), then \( \gamma' = \mu' \) and
\[ g(y^*) \geq (D_2 - D_1 v^{\gamma'}) \| y^* \|^{\gamma'} + C_3 \]
with \( D_1, D_2 \) defined as before. Thus
\[ D_2 - D_1 v^{\gamma'} = (y')^{\gamma'} | y^{\gamma'} | \nu^{\gamma'} | x_0 - \xi v^{\gamma'} |, \]
where \( a = 1/(1 - \gamma) < 0 \). This is positive whenever \( \xi > \sigma v^{\gamma'} \) and then
\[ \liminf_{\| y^* \| \to \infty} g(y^*) = +\infty. \]

When \( \mu = 1 \), just as in Theorem 7.3, one has
\[ K^*(y^*) = +\infty \quad \text{if} \quad \| y^* \| > \sigma, \]
and so \( g(y^*) = +\infty \) for \( \| y^* \| > \sigma \) if \( \gamma > \mu \).

This result says essentially that if \( k_1, k_2, A \) obey the estimates (7.23) which guarantee that \( f \) is coercive, then the dual functional \( g \) will also be coercive. This leads to the following existence theorem for the dual problem.

**Theorem 7.6.** Let \( Y \) be a reflexive Banach space; \( k_1, k_2 \) be proper, convex functionals and suppose (7.23) holds with either \( \gamma > \mu \geq 1 \) or \( \gamma = \mu > 1 \) and \( \xi v^{\gamma'} > \sigma \). Assume \( g \) defined by (7.16) is weakly lower semicontinuous on \( Y^* \). Then there exists a solution \( \hat{y}^* \) of (\( J^* \)) and \( \alpha^* = g(\hat{y}^*) = \inf_{y^*} g(y^*). \)

**Proof.** When (7.23) holds with \( \gamma > \mu \geq 1 \) or \( \gamma = \mu > 1 \) and \( \xi v^{\gamma'} > \sigma \) then
\[ \liminf_{x \to \infty} f(x) = +\infty. \]
From Theorem 7.5, one has \( \liminf_{\| y^* \| \to \infty} g(y^*) = +\infty. \) Thus (\( J^* \)) has a solution and a finite value as \( g \) is w.l.s.c. and \( Y^* \) is reflexive.

Let \( \hat{y}^* \) be a solution, then \( \alpha^* = g(\hat{y}^*) = \inf_{y^*} g(y^*). \) From Lemma 3.2, since the Lagrangian \( L \) defined by (7.15) is of type II, one has that \( \alpha^* = \alpha = \inf_x f(x) \) as required. \( \square \)

This result enables one to prove the existence of a solution of the dual problem even if there isn't a solution of the primal problem (\( J^* \)). Upon
making more assumptions on \( f \) and \( X \) one obtains the following analog of Theorem 7.4 in which one obtains existence for both \((\mathcal{S})\) and \((\mathcal{S}^*)\).

**Theorem 7.7.** Let \( X, Y \) be reflexive Banach spaces; \( k_1, k_2 \) be proper convex functionals and assume (7.23) holds with either \( y > \mu > 1 \) or \( y = \mu > 1 \) and \( \xi v^2 > \sigma \). Assume \( f \) defined by (7.11) and \( g \) by (7.16) are weakly lower semi-continuous. Then \((\mathcal{S})\) and \((\mathcal{S}^*)\) have solutions. If \( \hat{x} \) is a solution of \((\mathcal{S})\), \( \hat{y}^* \) is a solution of \((\mathcal{S}^*)\), then

1. \( a = f(\hat{x}) = g(\hat{y}^*) = a^* > -\infty \) and
2. \((\hat{x}, \hat{y}^*)\) is a \( \partial \)-critical point of \( L \) with \( L(\hat{x}, \hat{y}^*) = a \).

**Proof.** The existence follows immediately since both \( f, g \) are coercive w.l.s.c. functionals on a reflexive space. The other results follow since \( L \) given by (7.15) is a Lagrangian of type II.

8. **Example 1: Rayleigh's Principle**

Let \( A \) be real symmetric \( n \times n \) matrix. Rayleigh's principle is that the largest eigenvalue \( \lambda_1 \) of \( A \) is given by

\[
\lambda_1 = \sup_{u \neq 0} \frac{\langle Au, u \rangle}{\langle u, u \rangle} = \sup_{u \in S_1} \langle Au, u \rangle,
\]

where \( \langle x, y \rangle = \sum_{i=1}^n x_i y_i \) is the usual inner product on \( \mathbb{R}^n \) and \( S_1 = \{ u \in \mathbb{R}^n : \langle u, u \rangle = 1 \} \) is the unit sphere in \( \mathbb{R}^n \). See Courant and Hilbert [3, Chap. 1] for a thorough discussion of this problem.

Let \( B_1 = \{ u \in \mathbb{R}^n : \langle u, u \rangle \leq 1 \} \) and define \( f : \mathbb{R}^n \to \mathbb{R} \) by

\[
f(u) = -\langle Au, u \rangle + \chi_{B_1}(u)
\]

where

\[
\chi_{B_1}(u) = \begin{cases} 0 & \text{if } u \in B_1, \\ \infty & \text{otherwise}. \end{cases}
\]

The primal problem \((\mathcal{S})\) is to find \( \inf_{u \in \mathbb{R}^n} f(u) \). From (8.1) and some elementary analysis, one observes that the value of \((\mathcal{S})\) is \(-\lambda_1\). In general the functional \( f \) is non-convex. It is convex if and only if \( A \) is negative semi-definite (i.e., all the eigenvalues of \( A \) are less than or equal to 0).
DUALITY FOR NON-CONVEX VARIATIONAL PRINCIPLES

There are a number of different dual principles to $(\mathcal{D})$. Assume, without loss of generality, that $A$ is positive definite, so there exists $c > 0$ such that

$$\langle Au, u \rangle \geq c \|u\|^2 \quad \text{for all } u \in \mathbb{R}^n.$$  

where $\|u\|^2 = \langle u, u \rangle$. (If $A$ is not positive-definite, consider $\mu I + A$, where $\mu$ is sufficiently large.)

Then $f_1(u) = -\langle Au, u \rangle$ is concave on $\mathbb{R}^n$ and $f$ is the difference of two convex functionals.

(a) Use case III of Section 2 with $A = A^{1/2}, f_2(u) = \chi_{B_1}(u)$. Then $f_1(Au) = -\langle Au, Au \rangle$ and

$$(-f_1)^*(v) = \sup_{w \in \mathbb{R}^n} \left( \langle v, w \rangle - \langle w, w \rangle \right) = \frac{1}{4} \langle v, v \rangle.$$

The Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by (2.24) is

$$L(u, v) = \langle A^{1/2} u, v \rangle + \frac{1}{4} \langle v, v \rangle + \chi_{B_1}(u). \quad (8.2)$$

This is a Lagrangian of type II.

The dual functional $g : \mathbb{R}^n \to \mathbb{R}$ is

$$g(v) = \frac{1}{4} \langle v, v \rangle + \inf_{\|u\| \leq 1} \langle A^{1/2} u, v \rangle$$

$$= \frac{1}{4} \langle v, v \rangle - \langle Av, v \rangle^{1/2}. \quad (8.3)$$

One observes $g(v)$ is finite for each $v$ in $\mathbb{R}^n$; it is $C^\infty$ on $\mathbb{R}^n \setminus \{0\}$; it is difference of two convex functions and $g(v) \to \infty$ as $\|v\| \to \infty$.

The dual variational principle $(\mathcal{D}^*)$ is to minimize $g$ on $\mathbb{R}^n$. It is an unconstrained problem.

(b) Use case III of Section 2 with $A = I, f_2(u) = \chi_{B_1}(u)$. Then $f_1(u) = -\langle Au, u \rangle$ and $(-f_1)^*(v) = \frac{1}{4} \langle A^{-1} v, v \rangle$.

The Lagrangian $L_2 : \mathbb{R}^{2n} \to \mathbb{R}$ defined by (2.24) is

$$L_2(u, v) = \langle u, v \rangle + \frac{1}{2} \langle A^{-1} v, v \rangle + \chi_{B_1}(u) \quad (8.4)$$

and is of type II.

The dual functional $g_2 : \mathbb{R}^n \to \mathbb{R}$ is

$$g_2(v) = \inf_{u \in \mathbb{R}^n} \left[ \langle u, v \rangle + \chi_{B_1}(u) \right] + \frac{1}{4} \langle A^{-1} v, v \rangle$$

$$= \frac{1}{4} \langle A^{-1} v, v \rangle - \|v\|. \quad (8.5)$$

Again $g_2(v)$ is finite for each $v$ in $\mathbb{R}^n$. $g_2$ is $C^\infty$ on $\mathbb{R}^n \setminus \{0\}$, is the difference of two convex functions and $g_2(v) \to \infty$ as $\|v\| \to \infty$. 

The dual principle \((\mathcal{P}^*_{2})\) is to minimize \(g_2\) on \(\mathbb{R}^n\) and again it is an unconstrained problem.

From these constructions one sees that \(\mathcal{P}, \mathcal{P}^*\) and \(\mathcal{P}^*_{2}\) all have solutions. From Section 3, one would like to study the \(\partial\)-critical points of \(L\) and especially its global minima.

The analysis of the critical points of \(f\) is well known, at least in a classical manner. See [3, Chap. 1]. The results for \(g\) may be summarized as follows.

**Theorem 8.1.** Suppose \(A\) is positive-definite and \(\hat{v}\) is a critical point of \(g\). Then \(\hat{v}\) is an eigenvector of \(A\) corresponding to the eigenvalue \(\hat{\lambda} = \frac{1}{2}\langle A\hat{v}, \hat{v}\rangle^{1/2}; \|\hat{v}\| = 2\sqrt{\hat{\lambda}}\) and \(g(\hat{v}) = -\hat{\lambda}\).

\(\hat{v}\) is a non-degenerate critical point of \(g\) iff \(\hat{\lambda}\) is a simple eigenvalue of \(A\). In this case, the Morse index of \(\hat{v}\) is \(k\) iff \(\hat{\lambda}\) is the \((k + 1)\)st largest eigenvalue of \(A\).

**Proof:** Differentiating (8.3) one has

\[
\frac{\partial g}{\partial v_i}(v) = \frac{1}{2} v_i - \langle Av, v \rangle^{-1/2} \sum_{j=1}^{n} a_{ij} v_j, \quad 1 \leq i \leq n,
\]

where \(a_{ij}\) are the components of \(A\).

Thus if \(\hat{v}\) is a critical point of \(g\), \(\|\hat{v}\| \neq 0\) and

\[
A\hat{v} = \frac{1}{2}\langle A\hat{v}, \hat{v}\rangle^{1/2} \hat{v}.
\]

That is, \(\hat{v}\) is an eigenvector of \(A\) corresponding to the eigenvalue

\[
\hat{\lambda} = \frac{1}{2}\langle A\hat{v}, \hat{v}\rangle^{1/2}.
\]

Take the inner product of (8.6) with \(\hat{v}\), then

\[
2\hat{\lambda} = \langle A\hat{v}, \hat{v}\rangle^{1/2} = \frac{1}{2} \|\hat{v}\|^2
\]

or

\[
\|\hat{v}\| = 2\sqrt{\hat{\lambda}}.
\]

Substituting into (8.3) one finds that \(g(\hat{v}) = -\hat{\lambda}\).

Taking the second derivative of \(g\) at \(v\) \((\neq 0)\) one finds

\[
\frac{\partial^2 g}{\partial v_i \partial v_j}(v) = \frac{1}{2} \delta_{ij} - \langle Av, v \rangle^{-1/2} a_{ij}
\]

\[
+ \langle Av, v \rangle^{-3/2} \left( \sum_{k=1}^{n} a_{ik} v_k \right) \left( \sum_{l=1}^{n} a_{jl} v_l \right).
\]
Thus
\[
\frac{\partial^2 g}{\partial v_i \partial v_j}(\hat{v}) = \frac{1}{2} \delta_{ij} \frac{a_{ij}}{2\lambda} + \frac{1}{8\lambda} \hat{v}_i \hat{v}_j,
\]
where \( \hat{v} \) has components \( \hat{v}_i \) and we have used the properties of \( \hat{v} \).

Let \( \hat{G} \) be the Hessian matrix of \( g \) at \( \hat{v} \) and let \( \hat{V} \) be the matrix whose entries are \( \hat{v}_i \hat{v}_j \). Then
\[
\hat{G} = \frac{1}{2} I - \frac{1}{2\lambda} A + \frac{1}{8\lambda} \hat{V}.
\]

Let \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \) be the eigenvalues of \( A \) and let \( \{w^1, w^2, \ldots, w^n\} \) be a corresponding orthonormal family of eigenvectors. These may be chosen so that
\[
\hat{v} = \pm 2 \sqrt{\lambda_k} w^k \quad \text{for some} \quad 1 \leq k \leq n.
\]

Now
\[
(Vw^l)_i = \sum_{j=1}^{n} \hat{v}_i \hat{v}_j w^l_j = 4\lambda_k w^k_i \langle w^k, w^l \rangle.
\]
So
\[
Vw^l = 0 \quad \text{if} \quad l \neq k,
\]
\[
= 4\lambda_k w^k \quad \text{if} \quad l = k.
\]

Thus
\[
\hat{G}w^l = \frac{1}{2} \left( 1 - \frac{\lambda_l}{\lambda_k} \right) w^l \quad \text{if} \quad l \neq k,
\]
\[
= \frac{1}{2} w^l \quad \text{if} \quad l = k,
\]
so \( \hat{v} \) is non-degenerate iff \( \lambda_k \) is a simple eigenvalue of \( A \).

Moreover, the Morse index of \( \hat{v} \) is precisely \( k - 1 \) whenever \( \hat{v} \) is proportional to the \( k \)th eigenvector \( w^k \).

One may ask how do the critical points of \( f \) and \( g \) correspond? Unfortunately one cannot use the theory of Section 5 directly since \( L \) does not obey all the necessary differentiability conditions. Nevertheless, one has the following.
THEOREM 8.2. Suppose $A$ is positive-definite with $n$ distinct eigenvalues. The $\partial$-critical points of the Lagrangian (8.2) are precisely the points $\pm (w_k, -2 \sqrt{\lambda_k} \, w_k)$ for $1 \leq k \leq n$.

The Lagrangian is minimized at $\pm (w_1, -2 \sqrt{\lambda_1} \, w_1)$ and

$$ L(w_1, -2 \sqrt{\lambda_1} \, w_1) = f(\pm w_1) = g(\pm 2 \sqrt{\lambda_1} \, w_1) = -\lambda_1. $$

Proof: One observes from (8.2) that the Lagrangian is convex in $u$ for each $v$ and in $v$ for each $u$. Thus

$$ \partial_1 L(u, v) = A^{1/2} v + \partial \chi(u) $$

for $u$ in $B_1$, and

$$ \partial_2 L(u, v) = A^{1/2} u + \frac{1}{2} v. $$

Hence the $\partial$-critical points of $L$ are those points $(\hat{u}, \hat{v})$ obeying

$$ v = -2A^{1/2} u $$

and

$$ -A^{1/2} v \in \partial \chi(u). $$

Thus $\hat{u}$ obeys $2A \hat{u} \in \partial \chi(u)$ and hence $\|\hat{u}\| = 1$ and $A \hat{u} = \mu \hat{u}$ for some $\mu$ in $\mathbb{R}$ from Schwartz's inequality.

Thus $\hat{u}$ is a normalized eigenvector of $A$ and $A^{1/2} \hat{u} = \mu^{1/2} \hat{u}$, so $\hat{v} = -2 \mu^{1/2} \hat{u}$. Hence the first statement holds. The second statement follows from Theorem 3.3.

For the functional $g_2$ and the Lagrangian $L_2$ defined by (8.4) and (8.5) one has similar results.

THEOREM 8.3. Suppose $A$ is positive-definite and $\hat{v}$ is a critical point of $g_2$. Then $\hat{v}$ is an eigenvector of $A$ corresponding to the eigenvalue $\tilde{\lambda} = \frac{1}{2} \|\hat{v}\|$. $\hat{v}$ is a non-degenerate critical point of $g_2$ iff $\tilde{\lambda}$ is a simple eigenvalue of $A$. In this case the Morse index of $\hat{v}$ is $k - 1$ implies $\hat{v} = \pm 2\lambda_k \, w_k$ where $w_k$ is a normalized eigenvector corresponding to the $k$th eigenvalue $\lambda_k$.

When the eigenvalues of $A$ are distinct, the $\partial$-critical points of the Lagrangian $L_2$ defined by (8.4) are precisely $\pm (w_k, -2 \lambda_k \, w_k)$ for $1 \leq k \leq n$ and the Lagrangian takes its infimum value of $-\lambda_1$ at $\pm (w_1, -2 \lambda_1 \, w_1)$.

Proof: Differentiating (8.5) one has

$$ \frac{\partial g_2}{\partial v_j}(v) = \frac{1}{2} \sum_{j=1}^{n} a^{(0)}_{ij} v_j - \frac{v_j}{\|v\|} \quad \text{if} \quad v \neq 0. $$
Here $a_{ij}^{(-1)}$ are the entries in $A^{-1}$. Thus if $\hat{v}$ is a critical point of $g_2$ one has $\|\hat{v}\| \neq 0$ and $\frac{1}{2}A^{-1}\hat{v} = \hat{v}/\|\hat{v}\|$. Equivalently $\hat{v}$ is an eigenvector of $A$ corresponding to the eigenvalue $\hat{\lambda} = \|\hat{v}\|/2$.

Taking second derivatives, one finds

$$\frac{\partial^2 g_2}{\partial v_i \partial v_j} (\hat{v}) = \frac{1}{2} a_{ij}^{(-1)} - \frac{\delta_{ij}}{\|\hat{v}\|} - \frac{\hat{v}_i \hat{v}_j}{\|\hat{v}\|^3},$$

or

$$H = \frac{1}{2} A^{-1} - \frac{1}{2\hat{\lambda}} I - \frac{V}{8\hat{\lambda}^3},$$

where $V$ as before and $H$ is the Hessian matrix of $g_2$ at $\hat{v}$.

Just as before, the eigenvectors of $H$ are $\{w^1, \ldots, w^m\}$ and if $\hat{\lambda} = \lambda_k$ then

$$Hw^l = \frac{1}{2\lambda_l} \left(1 - \frac{\lambda_l}{\lambda_k}\right) w^l \quad \text{if} \quad l \neq k,$$

$$= \frac{1}{2\lambda_k^2} w^k \quad \text{if} \quad l = k.$$ 

Thus the critical point $\hat{v}$ is non-degenerate iff $\hat{\lambda}$ is a simple eigenvalue. The Morse index of $\hat{v}$ is the number of eigenvalues of $A$ greater than $\hat{\lambda}$, so it is $k - 1$ and in this case $\hat{\lambda} = \lambda_k$ and $\hat{v} = \pm 2\lambda_k w^k$.

From (8.4) one sees that the Lagrangian is convex in $\nu$ for each $\nu$ and in $\nu$ for each $\nu$ as $A^{-1}$ is also positive-definite. Thus

$$\partial_1 L(u, v) = v + \partial_{X_{B_1}}(u) \quad \text{for} \quad u \text{ in } B_1,$$

and

$$\partial_2 L(u, v) = u + \frac{1}{2} A^{-1} v.$$

Thus the $\partial$-critical points of $L$ are those points $(\hat{u}, \hat{v})$ obeying

$$-v \in \partial_{X_{B_1}}(u)$$

and

$$u + \frac{1}{2} A^{-1} v = 0 \quad \text{or} \quad A u = -\frac{1}{2} v.$$

Thus $\hat{u}$ obeys $2A\hat{u} \in \partial_{X_{B_1}}(\hat{u})$. Just as in the last theorem this implies $\hat{u}$ is a normalized eigenvector of $A$ but now

$$\hat{v} = -2\hat{\lambda} \hat{u}.$$
Thus the $\partial$-critical points of $L_2$ are precisely $\pm(w^k, -2\lambda_k w^k)$ and one easily verifies that $L_2(w^1, -2\lambda, w^1) = -\lambda_1$ as required.

It is interesting to compare these dual principles with Rayleigh’s principle.

First one notes that the dual problems are unconstrained variational problems for a functional which is always finite and is bounded below.

The critical points occur in pairs at $\pm 2\sqrt{\lambda_k} w^k$ for $g$ and at $\pm 2\lambda_k w^k$ for $g_2$.

Also, both the norm of the critical point and the value of the functional at the critical point are functions of the eigenvalue $\lambda_k$ only.

Finally, the Morse index of a non-degenerate critical point tells one which eigenvalue and eigenvector one has and is analogous to Courant’s mini-max results for Rayleigh’s principle.

One might consider the problem of generalizing this result to self-adjoint compact operators on a real Hilbert-space $H$. Assume $\mathcal{C}: H \to H$ is a compact linear operator and $\langle \cdot, \cdot \rangle$ defines the inner product on $H$. Let $B_1 = \{u \in H: \|u\| < 1\}$ be the closed unit ball in $H$ and define $f$ as above with $A$ replacing $\mathcal{C}$.

If $\mathcal{C}$ obeys $\langle \mathcal{C}u, u \rangle > 0$ for any $u(\not= 0) \in H$ or if $A = \mu I + \mathcal{C}$, where $\mu$ is sufficiently large that $A$ is positive-definite, then one may define the Lagrangian (8.2) and the dual functional $g$ as before with $H \times H$ replacing $\mathbb{R}^n \times \mathbb{R}^n$. In these cases analogues of Theorems 8.1 and 8.2 hold.

To generalize the second method one must introduce $A = \mu I + \mathcal{C}$ with $\mu$ sufficiently large that $A$ is positive-definite and then analogous results to Theorem 8.3 hold.

9. **Example II: Convex Programming**

The preceding methods may be generalized to cover the problems of extremizing convex or concave functions on closed, bounded sets in $\mathbb{R}^n$ or extremizing continuous functions on closed, convex sets in $\mathbb{R}^n$.

(a) Suppose $k: \mathbb{R}^n \to (-\infty, \infty]$ is a non-trivial, convex function and $B$ is a bounded closed set in $\mathbb{R}^n$. The primal problem $(\mathcal{P}_0)$ is to find

$$a_n = \sup_{x \in \mathbb{R}^n} k(x).$$

Define $f: \mathbb{R}^n \to \mathbb{R}$ by $f(x) = -k(x) + \chi_B(x)$, where $\chi_B$ is the indicator function of $B$.

$$\chi_B(x) = 0 \quad \text{if} \quad x \in B,$$
$$= \infty \quad \text{if} \quad x \not\in B.$$
The problem \( (\mathcal{P}_0) \) of maximizing \( k \) on \( B \) is equivalent to the problem \( (\mathcal{P}^*) \) of minimizing \( f \) on \( \mathbb{R}^n \).

Define \( L : \mathbb{R}^{2n} \to \mathbb{R} \) by (2.24) with \( A = I \) and \( X = Y = \mathbb{R}^n \). Then

\[
L(x, y) = \langle x, y \rangle + k^*(-y) + \chi_B(x) \tag{9.1}
\]

and \( L \) is a Lagrangian of type II for \( (\mathcal{P}^*) \).

The dual functional \( g : \mathbb{R}^n \to \mathbb{R} \) associated with this Lagrangian is

\[
g(y) = j_B(y) + k^*(-y),
\]

where

\[
j_B(y) = \inf_{x \in B} \langle x, y \rangle.
\]

The dual problem \( (\mathcal{P}^{**}) \) is to minimize \( g \) on \( \mathbb{R}^n \).

**Lemma 9.1.** Suppose \( k \) is continuous on \( B \). Then \( \alpha_0 \) is finite and there exists \( \hat{x} \in B \) such that \( k(\hat{x}) = \alpha_0 = \sup_{x \in B} k(x) \).

**Proof:** This is just a restatement of the fact that a continuous function attains its maximum on the compact subset \( B \) of \( \mathbb{R}^n \).

**Lemma 9.2.** Suppose that \( B \subseteq \{ x \in \mathbb{R}^n : \|x\| \leq M \} \) and that \( k^* \) obeys

\[
\liminf_{\|y\| \to \infty} \frac{k^*(y)}{\|y\|} = M > M_1 > 0. \tag{9.2}
\]

Then \( (\mathcal{P}^{**}) \) has a finite value \( \alpha^* \) and there exists \( \hat{y} \) in \( \mathbb{R}^n \) such that

\[
g(\hat{y}) = \inf_{y \in \mathbb{R}^n} g(y).
\]

**Proof:** When \( B \) is bounded, \( j_B(y) \) is finite for each \( y \) in \( \mathbb{R}^n \) and \( j_B(y) \geq -M \|y\| \) when \( M \) as above.

Now \( j_B \) is concave and thus continuous on \( \mathbb{R}^n \), \( f^* \) is convex and l.s.c. so \( g \) is l.s.c. on \( \mathbb{R}^n \).

When \( k^* \) obeys (9.2), then

\[
\liminf_{\|y\| \to \infty} \frac{g(y)}{\|y\|} \geq M_1 - M > 0.
\]

Thus \( g \) is coercive, so it attains its minimum on a compact subset of \( \mathbb{R}^n \).
This minimum is finite since neither \( j_B \) nor \( k^* \) can be \(-\infty\) on a compact subset of \( \mathbb{R}^n \).
One may well ask when condition (9.2) holds. A condition that guarantees (9.2) is that $k$ should be bounded on the ball of radius $M_1$, center at the origin.

When this holds, one has

$$f'(x) \leq C + \psi(x),$$

where $\psi(x)$ is zero if $\|x\| \leq M_1$ and is $+\infty$ if $\|x\| > M_1$ and $C$ is the bound on $f$.

Then

$$f^*(y) \geq -C + \psi^*(y) \quad \text{for all } y \in \mathbb{R}^n,$$

as $\psi^*(y) = M_1\|y\|$. Thus (9.2) holds.

**Theorem 9.3.** Let $B$ be a bounded closed set in $\mathbb{R}^n$, $k: \mathbb{R}^n \to (-\infty, \infty]$ be a non-trivial, convex function which is continuous on $B$ and let the Lagrangian $L$ be given by (9.1). Then:

(i) If $(\mathcal{P})$ has a solution $\tilde{x}$, then $\alpha = -\alpha_0 = \alpha^*$ is finite. If $\tilde{y} \in \partial k(\tilde{x})$ then $-\tilde{y}$ is a solution of $(\mathcal{P}^*)$ and $(\tilde{x}, -\tilde{y})$ is a $\partial$-critical point of $L$.

(ii) If $(\mathcal{P}^*)$ has a solution $\tilde{y}$, then $\alpha = \alpha^*$ is finite. If $\tilde{x}$ in $B$ obeys $j_B(\tilde{y}) = \langle \tilde{x}, \tilde{y} \rangle$, then $\tilde{x}$ is a solution of $(\mathcal{P}^*)$ and $(\tilde{x}, \tilde{y})$ is a $\partial$-critical point of $L$.

**Proof.** (i) For any Lagrangian of type II, Theorem 3.3 implies that $\alpha = \alpha^*$. From the construction of $f$, one sees that $\alpha = -\alpha_0$ is finite.

If $\tilde{y} \in \partial k(\tilde{x})$ one has $k(\tilde{x}) + k^*(\tilde{y}) = \langle \tilde{x}, \tilde{y} \rangle$ and from the definition of polars one has

$$k(\tilde{x}) + k^*(y) \geq \langle \tilde{x}, y \rangle \quad \text{for all } y \in \mathbb{R}^n.$$

Thus $f(\tilde{x}) = \alpha$ implies

$$\alpha = -k(\tilde{x}) + \chi_B(\tilde{x})$$

$$=-\langle \tilde{x}, \tilde{y} \rangle + k^*(\tilde{y}) + \chi_B(\tilde{x})$$

or $L(\tilde{x}, -\tilde{y}) = \alpha = \alpha^*$. Thus $(\tilde{x}, -\tilde{y})$ is a global minimizer for $L$. From Theorem 3.3, $-\tilde{y}$ is a solution of $(\mathcal{P}^*)$, and also $0 \in \partial L(\tilde{x}, -\tilde{y})$ so $(\tilde{x}, -\tilde{y})$ is a $\partial$-critical point of $L$.

The proof of (ii) is analogous. One has $\alpha = \alpha^*$ automatically and if $\tilde{y}$ is a solution of $(\mathcal{P}^*)$, then

$$\alpha^* = g(\tilde{y}) = j_B(\tilde{y}) + k^*(-\tilde{y}).$$
Now from the conditions on $B$, one sees that for each $y \in \mathbb{R}^n$, there exists $x = T(y) \in B$ such that $j_B(y) = \langle T(y), y \rangle$.

Let $\tilde{x} = T(\tilde{y})$, then

$$g(\tilde{y}) = \langle \tilde{x}, \tilde{y} \rangle + k^*(-\tilde{y})$$

$$= L(\tilde{x}, \tilde{y}) = \alpha^*.$$

Thus $(\tilde{x}, \tilde{y})$ minimizes $L$ on $\mathbb{R}^{2n}$. Hence from Theorem 3.3 $\tilde{x}$ is a solution of $(\mathcal{P}^*)$ and also $(\tilde{x}, \tilde{y})$ is a $\tilde{\dot{e}}$-critical point of $L$.

Remarks. This theorem says essentially that if one knows that $(\mathcal{P})$ has a solution $\tilde{x}$ and $\partial k(\tilde{x}) \neq \emptyset$, then $(\mathcal{P}^*)$ has a solution. One condition that guarantees that $\partial k(\tilde{x}) \neq \emptyset$ is Proposition 1.5.2 of [4], which essentially requires that $\tilde{x}$ be an interior point of the essential domain of $k$. If, however, one knows that $(\mathcal{P}^*)$ has a solution then $(\mathcal{P})$ automatically has a solution from (ii). Lemmas 9.1 and 9.2 provide direct conditions for $(\mathcal{P})$ and/or $(\mathcal{P}^*)$ to have solutions.

(b) Suppose now that $k$ and $B$ are as in (a) but the primal problem $(\mathcal{P})$ is to minimize $k$ on $B$.

Define $F: \mathbb{R}^n \to \mathbb{R}$ by

$$F(x) = k(x) + \chi_B(x)$$

and $L: \mathbb{R}^{2n} \to \mathbb{R}$ by (2.22) with $A = I$ so that

$$L(x, y) = \langle x, y \rangle + \chi_B(x) - k^*(y). \quad (9.3)$$

This Lagrangian is of type I, and the dual variational problem $(\mathcal{P}^*)$ is to maximize $g$ on $\mathbb{R}^n$ where

$$g(y) = \inf_{x \in B} \langle x, y \rangle - k^*(y) = j_B(y) - k^*(y)$$

and $j_B$ is defined as in (a). This functional is the sum of two concave functions.

**Lemmma 9.4.** Let $B$ be a bounded, closed set in $\mathbb{R}^n$ and $k: \mathbb{R}^n \to (-\infty, \infty]$ be a non-trivial, l.s.c. convex function. Then $\alpha$ and $\alpha^*$ are finite and $(\mathcal{P})$ has a solution. Suppose also that $B$ and $k^*$ obey the hypotheses of Lemma 9.2, then $(\mathcal{P}^*)$ has a solution.

**Proof:** Since $k$ is l.s.c. and $B$ is compact, then $\alpha = \inf_B k(x)$ is attained on $B$. Since $k$ does not take the value $-\infty$, $\alpha$ is finite.

The dual functional $g$ is finite for some $\tilde{y}$ in $\mathbb{R}^n$, so $-\infty < \alpha^* \leq \alpha$ and thus $\alpha^*$ is finite.
When $B$ and $k^*$ obey the hypotheses of Lemma 9.2, one has that $g$ is u.s.c. and $\liminf_{\|y\| \to \infty} g(y)/\|y\| \leq (M - M_1) < 0$.

Thus $g(y) \to -\infty$ as $\|y\| \to \infty$, so $g$ attains its supremum on a compact subset of $\mathbb{R}^n$ and $(\mathcal{J}^*)$ has a solution.

As is well known, if $\hat{x}$ is a solution of $(\mathcal{J})$, $\hat{y}$ is a solution of $(\mathcal{J}^*)$, then $(\hat{x}, \hat{y})$ need not be a saddle-point of the Lagrangian (9.3). When $B$ is convex though, $L$ does have a saddle-point and $\alpha = \alpha^*$.

(c) Suppose now that $C$ is a non-empty closed convex set in $\mathbb{R}^n$ and $w: \mathbb{R}^n \to \mathbb{R}$ is a continuous function.

Consider the primal problem $(\mathcal{J})$ of minimizing $w$ on $C$.

When $C$ is bounded, $a = \inf_{x \in C} w(x)$ is finite and $(\mathcal{J})$ has a solution.

Let $F: \mathbb{R}^n \to \mathbb{R}$ be defined by $F(x) = w(x) + \chi_C(x)$, where $\chi_C$ is the indicator function of $C$.

The polar functional $\chi_C^*(y) = \sup_{x \in C} \langle x, y \rangle$.

The Lagrangian $L: \mathbb{R}^n \to \mathbb{R}$ defined by (2.23) with $A = I$ is

$$L(x, y) = \langle x, y \rangle + w(x) - \chi_C^*(y). \quad (9.4)$$

The dual problem $(\mathcal{J}^*)$ is to maximize $g$ on $\mathbb{R}^n$ where $g: \mathbb{R}^n \to \mathbb{R}$ is defined by

$$g(y) = -\chi_C^*(y) - k^*(-y).$$

One observes that $g$ is the sum of two concave functions and the problem of finding simultaneous solutions of $(\mathcal{J})$ and $(\mathcal{J}^*)$ is equivalent to that of finding saddle-points of the Lagrangian $L$ defined by (9.4).

The anomalous dual problem $(\mathcal{J}^\circ)$ is to minimize $h$ on $\mathbb{R}^n$ where $h: \mathbb{R}^n \to \mathbb{R}$ is defined by

$$h(y) = (-w)^* (y) - \chi_C^*(y).$$

Here the functional $h$ is the difference of two convex functionals while $F$, in general, is not. The following lemma says that $(\mathcal{J})$ has a solution when $-\omega$ obeys the conditions of Lemma 9.2.

**Lemma 9.5.** Suppose $C$ is a bounded, closed convex set in $\mathbb{R}^n$ with $\|x\| \leq M$ for all $x$ in $C$. Suppose $w$ is bounded above on $\mathbb{R}^n$ and that

$$\liminf_{\|y\| \to \infty} \frac{(-w)^*(y)}{\|y\|} \geq M_1 > M. \quad (9.5)$$

Then $(\mathcal{J}^\circ)$ has a solution.

**Proof.** One has that $\chi_C^*(y)$ obeys $|\chi_C^*(y)| \leq M \|y\|$ and $\chi_C^*$ is convex. Thus it is continuous on $\mathbb{R}^n$. 

From the definitions of polars, one sees that

\[ (-w)^* (y) = \sup_{x \in \mathbb{R}^n} |\langle x, y \rangle + w(x)| \]

is convex and l.s.c. on \( \mathbb{R}^n \) and \( (-w)^*(0) = \sup_{x \in \mathbb{R}^n} w(x) \) is finite. Thus \( (-w)^* \) is non-trivial, so \( h \) is non-trivial and l.s.c.

Also

\[
\liminf_{\|y\| \to \infty} \frac{h(y)}{\|y\|} \geq M_1 - M > 0
\]

from (8.5). Thus \( h \) is coercive so \( (\mathscr{F}_w) \) has a solution. \( \qed \)

From (9.4) the Hamiltonian for this system is

\[ H(x, y) = \chi_*(y) - w(x). \] (9.6)

This Hamiltonian is convex in \( y \); it will be convex in \( x \) only if \( w \) is concave. In this case, one can use the results of Section 7 to prove the following.

**Theorem 9.6.** Assume \( w \) is concave on \( \mathbb{R}^n \), then \( \alpha = \alpha^\circ \). If, moreover, \( \alpha \) is finite, \( \hat{x} \) is a solution of \( (\mathscr{F}) \) and \( \hat{y}^* \) is a solution of \( (\mathscr{F}^\circ) \), then \( (\hat{x}, \hat{y}^*) \) is an anomalous critical point of the Lagrangian (9.4).

If, also, \( H \) obeys (L5)–(L8) and \( (\hat{x}, \hat{y}^*) \) is an anomalous critical point of the Lagrangian (9.4), then \( \hat{x} \) is a local minimum of \( f \) on \( C \) iff \( \hat{y}^* \) is a local minimum of \( h \) on \( \mathbb{R}^n \).

**Proof.** When \( w \) is concave on \( \mathbb{R}^n \), then \( F \) is the difference of two convex functionals. Take \( X = Y = \mathbb{R}^n \) and \( A = I \) in Theorem 7.1 and one has \( \alpha = \alpha^\circ \).

The second part follows just as in the proof of Theorem 7.4 and the last part follows from Theorem 6.3 since the Hamiltonian (8.6) already obeys (L1)–(L4). \( \qed \)

When \( w \) is not concave, there appear to be many more possibilities. Some relatively obvious results are the following.

**Lemma 9.7.** Assume that \( (\hat{x}, \hat{y}) \) in \( \mathbb{R}^{2n} \) obeys

\[ \hat{y} \in \partial \chi_C(\hat{x}) \cap \partial (-w)(\hat{x}). \] (9.7)

then \( (\hat{x}, \hat{y}) \) is an anomalous critical point of the Lagrangian (9.4).

**Proof.** If \( \hat{y} \in \partial \chi_C(\hat{x}) \) then \( \hat{x} \in \partial \chi_C^\circ(\hat{y}) \) since \( \chi_C \) is convex. Hence, from (9.6), \( (\hat{x}, \hat{y}) \) is an anomalous critical point.
Corollary. Suppose $\hat{x}$ is a solution of $(\mathcal{P})$ and (9.7) holds then $a^{\ominus} \leq a$.

Proof: Since (9.7) holds, $(\hat{x}, \hat{y})$ is an anomalous critical point of $L$. Thus from the corollary to Lemma 4.1,

$$h(\hat{y}) = f(\hat{x}) = a.$$

Hence

$$a^{\ominus} = \inf_{K^*} h(y) \leq a.$$

10. Example III: Semilinear Operator Equations

Let $X$ be a real Banach space and $X^*$ be its dual with respect to a pairing $\langle \cdot, \cdot \rangle$.

Let $A: X \rightarrow X^*$ be a continuous linear operator obeying

$$\langle Au, u \rangle \geq c \|u\|^2 \tag{10.1}$$

for all $u \in X$ and some $c \geq 0$.

Define $\mathcal{F}: X \rightarrow \mathbb{R}$ by

$$\mathcal{F}(u) = \frac{1}{2}\langle Au, u \rangle + f_2(u), \tag{10.2}$$

where $f_2: X \rightarrow \mathbb{R}$ is a given functional. Let the primal problem $(\mathcal{P})$ be to minimize $\mathcal{F}$ on $X$.

Suppose there is a Hilbert space $H$ and a continuous linear operator $A: X \rightarrow H$ such that $A = A^*A$.

Then

$$\mathcal{F}(u) = f_1(Au) + f_2(u), \tag{10.3}$$

where $f_1: H \rightarrow \mathbb{R}$ is defined by $f_1(v) = \frac{1}{2}\langle v, v \rangle$. $f_1$ is convex on $H$, so one may define a Lagrangian $L: X \times H \rightarrow \mathbb{R}$ by (2.22). Hence

$$L(u, v) = \langle Au, v \rangle + f_2(u) - \frac{1}{2}\langle v, v \rangle \tag{10.4}$$

is a Lagrangian of type I.

If the dual functional $f_2^*$ is non-trivial, the usual dual problem $(\mathcal{P}^*)$ is to maximize $g$ on $H$ where

$$g(v) = -\frac{1}{2}\langle v, v \rangle - f_2^*(-A^*v).$$

It may well be that $f_2^*$ is not non-trivial but that $(-f_2)^*$ is, in which case
one could consider the anomalous dual problem \((\mathcal{P}^\circledcirc)\) of minimizing \(h\) or
\[
\begin{align*}
  h(v) &= -\frac{1}{2}\langle v, v \rangle + (-f_2^*)^* (A^*v).
\end{align*}
\]  

When \(f_2\) is concave and u.s.c. on \(X\), one might try to define a duality \((2.25)\) or \((2.26)\). When \(A\) is surjective, one could take \(L:X \times H \to \mathbb{R}\) with
\[
\begin{align*}
  L(u, v) &= \langle Au, v \rangle + \frac{1}{2}\langle Au, Au \rangle + (-f_2^*)^* (-A^*v).
\end{align*}
\]  

This time the Lagrangian is of type II and the dual problem \((\mathcal{P}^\diamondsuit)\) is minimize \(g\) on \(H\) where
\[
\begin{align*}
  g(v) &= -\frac{1}{2}\langle v, v \rangle + (-f_2^*)^* (-A^*v).
\end{align*}
\]  

This is the same as the anomalous dual problem \((\mathcal{P}^\circledcirc)\) defined by \((1\alpha)\) with \(v\) replaced by \(-v\).

To illustrate the various possibilities we shall consider certain classes variational principles for non-linear integral equations.

Let \(\Omega\) be a bounded open set in \(\mathbb{R}^n\) and \(X=L^p(\Omega)\) with \(2 \leq p < \infty\). Define \(G: L^2(\Omega) \to L^1(\Omega)\) by
\[
G_\lambda u(x) = \lambda u(x) + \int_\Omega G(x, y) u(y) \, dy.
\]  

where \(\lambda \geq 0\) and \(G: \Omega \times \Omega \to \mathbb{R}\) is a Lebesgue measurable function obeyi
\[
\begin{align*}
  & (i) \quad G(x, y) = G(y, x) \text{ for all } x, y \text{ in } \Omega, \\
  & (ii) \quad \int_\Omega \int_\Omega G(x, y) u(x) u(y) \, dx \, dy \geq 0 \text{ for all } u \in L^2(\Omega), \quad \text{and} \\
  & (iii) \quad G_\lambda \text{ is a continuous operator.}
\end{align*}
\]  

Throughout this section we shall use Lebesgue measure on \(\Omega\).

We shall be interested in the variational principle obeyed by the soluti
of the semi-linear integral equation
\[
\lambda u(x) + \int_\Omega G(x, y) u(y) \, dy + k(x, u(x)) = 0.
\]  

Here \(k:\Omega \times \mathbb{R}^1 \to \mathbb{R}^1\) is a Caratheodory function (see [4, Chap. IV.1.] obeying
\[
\begin{align*}
  & (K1) \quad \text{There exists } C > 0, a \in L^2(\Omega) \text{ and } 0 < r < \infty \text{ such that} \\
  & \quad |k(x, t)| \leq C |t|^r + a(x) \quad \text{ a.e. on } \Omega \text{ for all } t \in \mathbb{R}.
\end{align*}
\]
Define \( K(x, t) = \int_0^t k(x, s) \, ds \) and \( p = \max(2, r + 1) \). Define \( f_2 : L^p(\Omega) \to \mathbb{R} \) by

\[
f_2(u) = \int_\Omega K(x, u(x)) \, dx.
\]

and \( \mathcal{F} : L^p(\Omega) \to \mathbb{R} \) by

\[
\mathcal{F}(u) = \frac{1}{2} \langle G, u, u \rangle + \int_\Omega K(x, u(x)) \, dx.
\]

We shall show that the critical points of \( \mathcal{F} \) on \( L^p(\Omega) \) obey (10.8) and that under quite reasonable assumptions on \( K \) the primal problem (\( \mathcal{P} \)) of minimizing \( \mathcal{F} \) on \( L^p(\Omega) \) has a solution. Then we shall study some dual problems associated with this primal problem.

If \( C \) is a convex subset of a vector space, then a function \( f : C \to \mathbb{R} \) is said to be quasi-convex if for each \( c \in \mathbb{R} \), the set \( \{ x \in C : f(x) < c \} \) is convex.

Consider the following condition:

\[(K2) \ f_2 \text{ is quasi-convex on } L^p(\Omega) \text{ and there exist a } C_0 (\neq -\infty) \text{ such that}
\]

\[
\inf_{t \in \mathbb{R}} K(x, t) \geq C_0 \quad \text{a.e. on } \Omega.
\]

**Lemma 10.1.** Assume \( k, K \) obey (K1)–(K2). Then \( f_2 \) is bounded on bounded subsets of \( L^p(\Omega) \), it is continuous in the norm topology, weakly l.s.c., and Gâteaux-differentiable at each point \( u \) in \( L^p(\Omega) \) with

\[
Df_2(u)(x) = k(x, u(x)) \in L^p(\Omega).
\]

**Proof.** From (K1) and the definition of \( K \), one has

\[
|K(x, t)| \leq \frac{C}{r + 1} |t|^{r+1} + |a(x)| t,
\]

\[
|f_2(u)| \leq \int_\Omega |C_1 |u(x)|^{r+1} + |a(x)||u(x)|| \, dx,
\]

where \( C_1 = C/(r + 1) \). Thus \( f_2 \) is bounded on bounded subsets of \( L^p(\Omega) \) from our choice of \( p \) and the fact that \( a \in L^2(\Omega) \).

Suppose \( \{u_n\} \) is a sequence in \( L^p(\Omega) \) which converges strongly to \( u \) in \( L^p(\Omega) \). Then there is a subsequence \( \{u_{n_j}\} \) such that \( \{u_{n_j}\} \) converges a.e. to \( u \) on \( \Omega \). Then \( K(x, u_{n_j}(x)) \) converges a.e. to \( K(x, u(x)) \) on \( \Omega \) as \( K(x, \cdot) \) is continuous.

From the dominated convergence theorem and the estimate on \( K \) one has
\[ \lim_{n \to \infty} \int_{\Omega} K(x, u_n(x)) \, dx = \int_{\Omega} K(x, u(x)) \, dx. \]

Since the limit function \( u \) is unique, one has \( \lim_{n \to \infty} f_2(u_n) = f_2(u) \) or \( f_2 \) is strongly continuous.

To show that \( f_2 \) is weakly l.s.c. it suffices to show that \( U_c = \{ u \in L^p(\Omega) : f_2(u) \leq c \} \) is weakly closed for any real \( c \).

Since \( f_2 \) is strongly continuous this set is strongly closed and \( f_2 \) is quasi-convex implies \( U_c \) is convex. Thus \( U_c \) is weakly closed.

Finally,

\[
\left| f_2(u + tv) - f_2(u) \right| \leq \int_{\Omega} k(x, u(x)) v(x) \, dx
\]

From the definition of \( K \) one has

\[
\left| K(x, (u + tv)(x)) - K(x, u(x)) \right| = k(x, u(x)) \, v(x)
\]
a.e. on \( \Omega \) and from the estimates on \( k, K \) and the Lebesgue dominated convergence theorem one finds

\[
\lim_{t \to 0} t^{-1} \left| f_2(u + tv) - f_2(u) \right| = \int_{\Omega} k(x, u(x)) \, v(x) \, dx.
\]

The fact that \( K(x, u(x)) \in L^p(\Omega) \) follows from (K1).

**Theorem 10.2.** Assume \( k, K \) obey (K1)--(K2) and \( \mathcal{F} \) is defined by (10.10). Then \( \mathcal{F} \) is weakly l.s.c. and bounded below on \( L^p(\Omega) \).

**Proof.** We have that \( \frac{1}{2} \langle G_\lambda u, u \rangle \) is convex on \( L^p(\Omega) \) and non-negative since \( \lambda \geq 0 \) and condition (ii) on \( G \) holds.

Also this functional is strongly continuous as \( \{ u_n \} \) converges to \( \hat{u} \) in \( L^p(\Omega) \) implies \( \{ u_n \} \) converges to \( \hat{u} \) in \( L^2(\Omega) \) and thus \( \{ G_\lambda u_n \} \) converges to \( G_\lambda \hat{u} \) in \( L^2(\Omega) \) and thus in \( L^{p'}(\Omega) \) as \( p \geq 2, |\Omega| < \infty \).

Since it is convex and strongly continuous it is weakly l.s.c. on \( L^p(\Omega) \).

Combining this with the results of Lemma 10.1, one has that \( \mathcal{F} \) is weakly l.s.c. on \( L^p(\Omega) \) and \( \mathcal{F}(u) \geq f_2(u) \geq C_0 |\Omega| \) so \( \mathcal{F} \) is bounded below, from (K2).

**Corollary.** Suppose that for some \( \epsilon > 0 \), there exists \( R > 0 \) such that \( \mathcal{F}(u) \leq a + \epsilon \) implies \( \| u \|_p \leq R \) where \( a = \inf_{u \in L^p(\Omega)} \mathcal{F}(u) \). Then \( (\mathcal{F}) \) has a solution.

**Proof.** The set \( \{ u \in L^p(\Omega) : \| u \|_p \leq R \} \) is weakly compact as \( 2 \leq p < \infty \) so this result follows from the standard existence theorem.
Now we shall be interested in various dual problems associated with \((\mathcal{P})\). To use the abstract results introduced at the beginning of this section take 
\[ H = L^2(\Omega) \] and note that the operator \( G_\lambda \) is self-adjoint and 
\[ \langle G_\lambda u, u \rangle \geq \lambda \|u\|^2. \]

The operator \( G_\lambda \) has a principal square root \( G_\lambda^{1/2} \) which is also a self-adjoint, positive-definite and bounded linear operator when \( \lambda > 0 \).

Let \( A_\lambda \) be the restriction of \( G_\lambda^{1/2} \) to \( L^p(\Omega) \). Then \( A_\lambda : L^p(\Omega) \to L^2(\Omega) \) will be a continuous linear operator and one has that \( G_\lambda : L^p(\Omega) \to L^p(\Omega) \) obeys 
\[ G_\lambda = A_\lambda^* A_\lambda. \]

Then \( \mathcal{F} : L^p(\Omega) \to \mathbb{R} \) is given by 
\[ \mathcal{F}(u) = f'_1(A_\lambda u) + f'_2(u), \]
where \( f'_1(v) = \frac{1}{2} \langle v, v \rangle \) as in (10.3).

A Lagrangian associated with this is \( L : L^p(\Omega) \times L^2(\Omega) \to \mathbb{R} \) defined by (10.4)
\[ L(u, v) = \langle A_\lambda u, v \rangle + f'_2(u) - \frac{1}{2} \langle v, v \rangle. \quad (10.11) \]

The standard dual problem \((\mathcal{P}^*)\) is to maximize \( g \) on \( L^2(\Omega) \) where 
\[ g(v) = -\frac{1}{2} \langle v, v \rangle - f'_2(-A_\lambda^* v). \quad (10.12) \]

To compute \( g \) one needs to compute \( f'_2^* \). From [4, Chap. IV.1.2, Proposition 1.2] one has that \( f'_2^* : L^p(\Omega) \to \mathbb{R} \) is given by 
\[ f'_2^*(v) = \int_\Omega K^*(x, v(x)) \, dx, \quad (10.13) \]
where 
\[ K^*(x, t) = \sup_{y \in \mathbb{R}} (ty - K(x, y)). \]

From this definition and the basic properties of polars one has

**Lemma 10.3.** \( f'_2^* \) is weakly l.s.c., convex and non-negative on \( L^p(\Omega) \).

**Proof.** Polar functionals are automatically weakly l.s.c. and convex. One has \( K^*(x, t) \geq -K(x, 0) - 0 \) for all \( t \), so \( f'_2^*(v) \geq 0 \) as required.

**Theorem 10.4.** Assume \( k, K \) obey \((K1)-(K2)\) and \( g \) is defined by (10.12). Then \( a^* = \sup_{v \in L^2(\Omega)} g(v) \leq 0 \) and there is a unique \( \hat{v} \) in \( L^2(\Omega) \) such that \( g(\hat{v}) = a^* \).

**Proof.** The expression \( \frac{1}{2} \langle v, v \rangle \) is weakly l.s.c. on \( L^2(\Omega) \), so \(-g\) is the sum of two weakly l.s.c. functionals, and thus \( g \) is weakly u.s.c. on \( L^2(\Omega) \).
Also $-g$ is the sum of a strongly convex functional and a convex functional so $-g$ is strongly convex and $-g(v) \geq \frac{1}{2} \langle v, v \rangle \geq 0$. Thus $g$ has a unique maximum $v$ and $\alpha^* \leq 0$.

The dual functional $g$ may be written

$$g(v) = -\int_{\Omega} \left[ \frac{1}{2} v(x)^2 + K^*(x, -A^*_x v(x)) \right] dx. \quad (10.14)$$

If $f^*_x$ is Gâteaux-differentiable at $\hat{w} = -A^*_x \hat{v}$ in $L^p(\Omega)$ and one has

$$Df^*_x(\hat{w}) = k^*(x, \hat{w}(x)), \quad (10.15)$$

where $k^*: \Omega \times \mathbb{R} \to \mathbb{R}$ is some measurable function, then the fact that $Dg(\hat{v}) = 0$ implies

$$-\hat{v}(x) + A^*_x k^*(x, \hat{w}(x)) = 0.$$

Applying $A^*_x$ to this one sees that $\hat{w}$ obeys

$$w(x) + \left[ \hat{\lambda} k^*(x, w(x)) + \int_{\Omega} G(x, y) k^*(y, w(y)) dy \right] = 0 \quad (10.16)$$

a.e. on $\Omega$ upon using $G^*_x = A^*_x A_x$.

Thus the critical point of the dual functional obeys an equation of Hammerstein type.

Some other interesting examples of this duality occur in finite dimensional problems. Let $X = X^* = H = \mathbb{R}^n$ and $A$ be a real-symmetric positive-definite matrix obeying (10.1) with $c > 0$.

Let $f^*_2: \mathbb{R}^n \to \mathbb{R}$ be a continuous function and consider a functional $\mathcal{F}: \mathbb{R}^n \to \mathbb{R}$ defined by (10.2). The primal problem $(\mathcal{P})$ is to minimize $\mathcal{F}$ on $\mathbb{R}^n$.

Let $A = A^{1/2}$ be the principal square root of $A$ and consider $L: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by (10.3) so that

$$L(u, v) = \langle A^{1/2} u, v \rangle - \frac{1}{2} \langle v, v \rangle + f_2(u). \quad (10.17)$$

Let $\| \cdot \|_1$, $\| \cdot \|_2$ and $\| \cdot \|_\infty$ represent the $l^1$, $l^2$ and $l^\infty$ norms on $\mathbb{R}^n$ and consider the following condition:

(F1) There exist constants $C_1$, $C_2$ and $p$ with $C_1 > 0$ and $0 < p \leq 2$ such that

$$f_2(u) \geq -C_1 \| u \|_2^p - C_2 \quad (10.18)$$

for all $u$ in $\mathbb{R}^n$. 
Lemma 10.5. Suppose $A, f_2$ as above and $f_2$ obeys (F1). If $0 < p < 2$, or if $p = 2$ and $2C_1 < c$, then $\mathcal{F}$ attains its infimum on $\mathbb{R}^n$.

Proof. We have assume that $\mathcal{F}$ is continuous, hence l.s.c. on $\mathbb{R}^n$. To prove existence, it suffices to prove that $\mathcal{F}$ is coercive. From (10.18),

$$\mathcal{F}(u) \geq \frac{c}{2} \|u\|^2_2 - C_1 \|u\|^p_2 - C_2.$$ 

Thus if $0 < p < 2$, $\mathcal{F}(u) \to \infty$ as $\|u\|_2 \to \infty$, so $(\mathcal{F})$ has at least one solution. Similarly if $p = 2$ and $C_1 < c/2$.

When $f_2$ is convex, or quasi-convex, the usual dual function may be constructed using (10.4) and one can study the problem of finding saddle-points of the Lagrangian. Since these problems have been treated elsewhere we shall not go through the details here.

It is of some interest, though, to construct some examples of non-trivial anomalous dual variational principles to which the theory of Section 5 applies.

Consider the following conditions:

(F2) $f_2$ is concave and bounded above on $\mathbb{R}^n$.

(F3) There exist positive constants $C_3, C_4$ and $\gamma$ such that $f_2(u) < -C_3 \|u\|^{1+\gamma} + C_4$ for all $u$ in $\mathbb{R}^n$.

When $f_2$ obeys (F2), define $\mu = f_2(0)$ and $\beta = \sup_{u \in \mathbb{R}^n} f_2(u)$. Then $-f_2$ has a non-trivial polar with $(-f_2)^*(0) = \beta$ and $(-f_2)^*(w) \geq \mu$ for all $w$ in $\mathbb{R}^n$.

Let $h$ be the anomalous dual functional on $\mathbb{R}^n$ defined by (10.5) with $A^* = A^{1/2}$. Since $(-f_2)^*$ is non-trivial and convex, the functional $h$ will be non-trivial and is the difference of two convex functionals.

Lemma 10.6. Assume $f_2$ obeys (F2) and (F3). Then $(-f_2)^*$ is l.s.c. and there are constants $C_5, C_6$ such that

$$(-f_2)^*(w) \leq C_5 \|w\|_2^{s} + C_6$$

where $s = \gamma^{-1}(\gamma + 1)$.

Proof. One has $(-f_2)^*(w) = \sup_{u \in \mathbb{R}^n} \{\langle w, u \rangle + f_2(u)\}$. From (F3) $\langle w, u \rangle + f_2(u) \leq \|w\|_2 \|u\|_2 - C_3 \|u\|_2^{1+\gamma} + C_4$. This right-hand side is maximized when $\|u\|_2 = \|w\|_2/(\gamma + 1) C_3^{1/\gamma}$ and at this maximum it takes the value $C_5 \|w\|_2^{s} + C_4$, where $C_5$ depends only on $\gamma$ and $C_3$.

Corollary. Assume $f_2$ obeys (F2) and (F3). Then $h$ defined by (10.5) is l.s.c. on $\mathbb{R}^n$ and $h(w) \leq C_5 \|A^{1/2}w\|_2^{s} + C_6 - \frac{1}{2} \|w\|_2^{s}$ with $s = \gamma^{-1}(\gamma + 1)$. 

GILES AUCHMUTY
Thus when \( f_2 \) obeys (F2) and (F3), the anomalous dual functional is defined. To show that the dual problem (\( \mathcal{P}^\diamond \)) has a solution one generally has to show that \( h \) is coercive or else verify the conditions of Theorem 5.3. For this particular problem such analysis leads to the following results.

Assume

(F4) \( f_2 \) is strictly concave on \( \mathbb{R}^n \) and \( \liminf_{\|u\| \to \infty} -f_2(u) = \infty \).

(F5) \( f_2 \) is \( C^k \) on \( \mathbb{R}^n \) with \( k \geq 2 \).

From (F4) one has that (L1) and (L2) hold.

A critical point of the Lagrangian (10.17) is a point \((\hat{u}, \hat{v})\) in \( \mathbb{R}^{2n} \) obeying the system of equations

\[
A^{1/2} u = v
\]

and

\[
A^{1/2} v = -Df_2(u).
\]

Thus \( \hat{u} \) obeys \( Au + Df_2(u) = 0 \) which is the condition that \( u \) be a critical point of \( \mathcal{F} \).

Similarly, from (10.19) one sees that \( \hat{v} \) obeys

\[
A^{1/2} v = -Df_2(A^{-1/2} v).
\]

When (F4) and (F5) hold, Lemma 5.1 implies that if \((\hat{u}, \hat{v})\) is a solution of (10.19) then \( \hat{v} \) is a critical point of the functional \( h \).

Now

\[
D^2 L(\hat{u}, \hat{v}) = \begin{pmatrix} D^2 f_2(\hat{u}) & A^{1/2} \\ A^{1/2} & -I \end{pmatrix}
\]

in this case, so the assumptions of Theorem 5.3 hold. Thus one has the following result.

**Theorem 10.7.** Assume \( f_2 \) obeys (F4) and (F5). Let \((\hat{u}, \hat{v})\) be a solution of (10.19), then \( \hat{u} \) is a critical point of \( \mathcal{F} \) and \( \hat{v} \) is a critical point of \( h \). \( \hat{u} \) is a non-degenerate critical point of \( \mathcal{F} \) iff \( \hat{v} \) is a non-degenerate critical point of \( h \) and \( i(\hat{u}) = i(\hat{v}) \). In particular, if either \((\mathcal{P})\) or \((\mathcal{P}^\diamond)\) has a solution so do the other and \( \alpha = \alpha^\diamond \).

**Proof.** The first two statements follow immediately from Lemma 5.1. Theorem 5.3. Assume \((\mathcal{P})\) has a solution \( \hat{u} \) and \( \mathcal{F}(\hat{u}) = \alpha = \inf_{u \in \mathcal{U}} \mathcal{F}(u) \). Then \( \hat{u} \) is a critical point of the differentiable functional \( \mathcal{F} \). Define \( \hat{v} = A^{1/2} \hat{u} \).

Then \( A^{1/2} \hat{v} = A\hat{u} = -Df_2(\hat{u}) \) as \( \hat{u} \) is a critical point of \( \mathcal{F} \). Thus \( (\hat{u}, \hat{v}) \) satisfies (10.19) so \( \hat{v} \) is a critical point of \( h \).
From Theorem 5.3, if $\tilde{u}$ is non-degenerate then $\tilde{v}$ is non-degenerate and in this case $i(\tilde{v}) = 0$ as $i(\tilde{u}) = 0$, so $\tilde{v}$ will be a local minimum.

We must show that $\tilde{v}$ is a global minimum of $h$.

For this problem the Hamiltonian $H$ defined by (4.1) is

$$H(u, v) = \frac{1}{2}\langle v, v \rangle - f_2(u),$$

and from (F4), $H$ is strictly convex in $u$ and $v$ and

$$\partial_1 H(u, v) = -Df_2(u),$$

$$\partial_2 H(u, v) = v.$$

as $H$ is $C^k$ on $\mathbb{R}^{2n}$ with $k \geq 2$.

Thus $(\tilde{u}, \tilde{v})$ is an anomalous critical point of $L$ so from the corollary to Lemma 4.1, one has $\tilde{u}$ is a solution of $(\mathcal{L}_2, \mathcal{J}'(\tilde{u})) = \mathcal{L}(\tilde{v})$ and $\alpha = \alpha^{\circ}$ as required.

When $f_2$ is not $C^2$ on $\mathbb{R}^n$, one may still obtain similar results. A simple explicit example of these cases is the following.

Consider the functional $\mathcal{F}: \mathbb{R} \to \mathbb{R}$ defined by

$$\mathcal{F}(x) = \frac{1}{2}ax^2 - c|x|^{1+\gamma},$$

where $0 < \gamma < 1$ and $a$, $c$ are positive numbers.

Here $X = \mathbb{R}$, $A = (a)$ and $f_2'(x) = -c|x|^{1+\gamma}$.

The Lagrangian $L: \mathbb{R}^2 \to \mathbb{R}$ is given by

$$L(x, y) = \sqrt{a}xy - c|x|^{1+\gamma} - \frac{1}{2}y^2.$$

One observes that $f_2$ is strictly concave and obeys (F1), (F4) and $(-f_2^*)^\gamma(y) = K|y|^{\gamma(1+1)/\gamma}$, where $K$ depends on $C$ and $\gamma$ only.

The anomalous dual functional is

$$h(y) = K|a^{1/2}y|^{\gamma(1+1)/\gamma} - \frac{1}{2}y^2$$

and this obeys $\inf_{y \in \mathbb{R}}h(y) = \alpha^{\circ} < 0$ and $h$ is coercive on $\mathbb{R}$ as $1 + 1/\gamma > 2$. A simple check shows that $\alpha = \inf_{x \in \mathbb{R}}\mathcal{F}(x) = \alpha^{\circ}$.

The critical points of $L$ obey $\sqrt{a}x = y$ and $\sqrt{a}y = -c(y+1)|x|^\gamma \text{sgn } x$ when $|x| \neq 0$. Critical points of a similar nature correspond.

This example may be generalized in a number of ways to provide other variational problems for which the anomalous dual variational principles are well posed.
11. Example IV: Classical Dynamical Systems

Many of the best-known variational principles originally arose in mechanics, and their dual principles were derived independently by studying the same system in different coordinate systems.

Perhaps the most familiar example is the motion of a classical dynamical system, with a well-defined action integral.

Suppose \( \mathcal{L}: [0, 1] \times C_1 \times C_2 \rightarrow \mathbb{R} \) is the (mechanical) Lagrangian of a dynamical system with \( n \) degrees of freedom. Then \( C_1 \subseteq \mathbb{R}^n \) and \( C_2 \subseteq \mathbb{R}^n \), and we have normalized the time interval to be \([0, 1]\). The action integral is

\[
f(u) = \int_0^1 \mathcal{L}(t, u(t), \dot{u}(t)) \, dt,
\]

where \( u: [0, 1] \rightarrow C_1 \) is the function defining the position of the system at time \( t \) and \( \dot{u}(t) = du(t)/dt \) represents the (generalized) velocities.

For simplicity of exposition we shall assume \( C_1 = C_2 = \mathbb{R}^n \); that \( u(0) = u(1) = 0 \) (i.e., we normalize the initial and end points of the motion) and that

\[
\mathcal{L}(t, u, \dot{u}) = T(u, \dot{u}) - V(t, u),
\]

where \( T: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \) is a convex function in the second variable and \( V: [0, 1] \times \mathbb{R}^n \rightarrow \mathbb{R} \) is a well-defined function.

Usually

\[
T(u, \dot{u}) = \frac{1}{2} \sum_{i=1}^n a_i(u) \dot{u}_i^2
\]

with \( a_i(u) \geq 0 \) for all \( u \in \mathbb{R}^n \).

Let \( X = \mathcal{W}^{1,2}_{+1}([0, 1]; \mathbb{R}^n) \). This is the set of all \( L^2 \)-functions \( u: [0, 1] \rightarrow \mathbb{R}^n \) whose components \( u_i \) have distributional derivatives \( \dot{u}_i(t) \) which also lie in \( L^2(0, 1) \) and which obey \( u_i(0) = u_i(1) = 0 \) for all \( i \). \( X \) is a real Hilbert space under the inner product

\[
\langle u, v \rangle_{1,2} = \sum_{i=1}^n \int_0^1 \left[ u_i(t) v_i(t) + \dot{u}_i(t) \dot{v}_i(t) \right] \, dt.
\]

Let \( Y_1 = L^2((0, 1); \mathbb{R}^n) \) be the Hilbert space of all measurable functions \( v: [0, 1] \rightarrow \mathbb{R}^n \) which are square integrable on \([0, 1]\). The inner product on \( Y_1 \) is defined by

\[
\langle u, v \rangle = \sum_{i=1}^n \int_0^1 u_i(t) v_i(t) \, dt.
\]
Let $Y$ be the subspace of $Y_1$ consisting of all functions for which

$$\int_0^1 v_i(t) \, dt = 0, \quad 1 \leq i \leq n.$$ 

Let $Au = du/dt$, then $A: W_0^{1,2}((0,1); \mathbb{R}^n) \to Y$ is a continuous linear operator; $A$ is a homeomorphism with

$$(A^{-1}v)_i(t) = \int_0^t v_i(t) \, dt$$

and

$$(A^*v)_i(t) = -\frac{dv_i}{dt}$$

so $A$ is skew-symmetric.

Let the primal problem $(\mathcal{P})$ be to minimize $f$ on $X$ where $f$ is given by (11.1), (11.2) and (11.3).

One can write

$$f(u) = F(u, Au), \quad (11.4)$$

where $F$ is convex on $Y$ for each $u$ in $X$. Thus using case I, the (duality) Lagrangian defined by (2.8) is the functional $L: X \times Y \to \mathbb{R}$ defined by

$$L(u, v) = \langle Au, v \rangle - F(u, v)$$

$$= \sum_{i=1}^n \left[ \int_0^1 \dot{u}_i(t) v_i(t) \, dt - \int_0^1 T^o(u(t), v(t)) \, dt \right]$$

$$- \int_0^1 V(t, u(t)) \, dt.$$ 

Here $T^o(u, v) = \sup_{y \in \mathbb{R}^n} \left| \sum_{i=1}^n v_i y_i - T(u, y) \right|$ is the polar of $T$ with respect to the second variable,

$$= \frac{1}{2} \sum_{i=1}^n a_i(u)^{-1} v_i^2$$

when $T$ is given by (11.3) and $a_i(u) \neq 0$ for $1 \leq i \leq n$, $u$ in $\mathbb{R}^n$.

The dual variables $v$ introduced above are the generalized momenta of classical mechanics and $T^o(u, v)$ is the kinetic energy of the system at a particular point in position-momentum space.

The standard dual problem $(\mathcal{P}^*)$ is to maximize $g$ on $Y$ where $g(v) = \inf_{u \in X} L(u, v)$. 

When $T$ is independent of $u$ so that the coefficients $a_i(u)$ in (10.3) are constant, then
\[ g(v) = -\int_0^1 T^*(v(t)) \, dt - \int_0^1 (-V)^*(t, \dot{v}(t)) \, dt \]
provided $\dot{v}$ is in $Y_1$. In general this dual functional is well defined if $V(t, \cdot)$ is concave on $\mathbb{R}^n$.

When the potential $V(t, \cdot)$ is convex on $\mathbb{R}^n$, one introduces the anomalous dual functional
\[ h(v) = -\int_0^1 T^*(v(t)) \, dt + \sup_{u \in X} \left( \langle Au, v \rangle - \int_0^1 V(t, u(t)) \, dt \right) \]
\[ = -\int_0^1 T^*(v(t)) \, dt + \int_0^1 V^*(t, A^*(t)) \, dt \]
whenever $A^* v = -\dot{v} \in Y_1$.

From (4.1) one sees that the Hamiltonian of this system is
\[ H(u, v) = \int_0^1 T^o(u(t), v(t)) \, dt + \int_0^1 V(t, u(t)) \, dt \]
and this is the integral over $[0, 1]$ of the mechanical Hamiltonian. This justifies the choice of (4.1) as the definition of the Hamiltonian.

It is worth noting that Toland's example [7] of the heavy rotating chain has one degree of freedom and
\[ f(u) = \int_0^1 \left[ \frac{\dot{u}(t)^2}{2\lambda} - (u(t)^2 + t^2)^{1/2} \right] \, dt. \]

In this case the potential $V$ is convex in $u$ for each $t$ and thus the functional $F$ defined by (11.4) is concave in $u$.

Suppose one introduces the Lagrangian $L$ defined by (2.26). Then $L: X \times Y \to \mathbb{R}$ is defined by
\[ L(u, v) = \langle Au, v \rangle + \int_0^1 \frac{\dot{u}(t)^2}{2\lambda} \, dt + \int_0^1 V^*(t, \dot{v}(t)) \, dt, \]
where $V(t, u) = (u^2(t) + t^2)^{1/2}$ and thus
\[ V^*(t, v) = \sup_{u \in \mathbb{R}} (uv - \sqrt{t^2 + u^2(t)}) \]
\[ = \infty \quad \text{if} \quad |v| \geq 1, \]
\[ = t/\sqrt{1 - v^2} \quad \text{if} \quad |v| \leq 1. \]
This is a Lagrangian of type II.

The dual variational principle ($\mathcal{P}^*$) studied by Toland is to minimize $g$ on $Y$ where

$$g(v) = \inf_{u \in X} L(u, v)$$

and his results on the correspondence of critical points of this functional are generalized in Section 6 of this paper.

### 12. Example V: Nonlinear Elliptic Problems

Variational principles for elliptic partial differential equations have been a major topic of study for a long time. For two recent reviews of these problems see Bombieri [2] or Serrin [6].

Throughout this section $\Omega$ is a bounded open set in $\mathbb{R}^n$ with a smooth boundary $\partial \Omega$ and $W_{0}^{1,p}(\Omega)$ is the usual Sobolev space as defined, for example, in [5]. The norm on $W_{0}^{1,p}(\Omega)$ is

$$\|u\|_{1,p} = \left( \int_{\Omega} \left( |u(x)|^p + \sum_{i=1}^{n} |D_i u(x)|^p \right) dx \right)^{\frac{1}{p}},$$

(12.1)

where $D_i u(x) = \frac{\partial u(x)}{\partial x_i}$ and $1 \leq p < \infty$.

Assume $F: \Omega \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is a Caratheodory function obeying

(i) $F(x, u, \cdot)$ is convex on $\mathbb{R}^n$ for each $(x, u)$ in $\Omega \times \mathbb{R}$ and

(ii) there exist constants $C_1$ and $C_2$ such that

$$|F(x, u, \xi)| \leq C_1 \|\xi\|^p_1 + |u|^p + C_2$$

for all $x$ in $\Omega$, $(u, \xi)$ in $\mathbb{R}^{n+1}$.

Let $X = W_{0}^{1,p}(\Omega)$ and consider the functional $f: X \to \mathbb{R}$ with $1 < p < \infty$ defined by

$$f(u) = \int_{\Omega} F(x, u, Du(x)) \, dx.$$

(12.2)

**Lemma 12.1.** Suppose $f$ is defined by (11.2) and $F$ obeys (i)–(ii), then $f$ is strongly continuous on $X$. Suppose also that whenever a sequence $\{u_n: n \geq 1\}$ of functions in $L^p(\Omega)$ converges strongly to a function $u$ in $L^p(\Omega)$ then

$$\lim_{n \to \infty} \int_{\Omega} |F(x, u_n(x), Dv(x)) - F(x, u(x), Dv(x))| \, dx = 0$$

uniformly on bounded subsets of $W_{0}^{1,p}(\Omega)$. Then $f$ is weakly l.s.c. on $X$. 
Proof. The fact that \( f \) is strongly continuous follows directly from (ii) and the fact that \( F(x, \cdot, \cdot) \) is continuous on \( \mathbb{R}^{n+1} \) for almost all \( x \) using an analog of Lemma 10.1 (see also [4, Chap. IX.3, Lemma 3.2]).

Suppose \( \{u_n : n \geq 1\} \) is a sequence of functions in \( X \) which converges weakly to \( u \) in \( X \). Then \( \{u_n\} \) converges strongly to \( u \) in \( L^p(\Omega) \) and \( \|u_n\|_1, \rho \) is bounded.

Let \( \mathcal{F}(u, v) = \int_\Omega F(x, u(x), Dv(x)) \, dx \).

Then \( f(u) = \mathcal{F}(u, u) \) and \( f(u_n) - f(u) = \mathcal{F}(u_n, u_n) - \mathcal{F}(u, u_n) + \mathcal{F}(u_n) - \mathcal{F}(u, u) \). Now \( \liminf_{n \to \infty} |\mathcal{F}(u_n, u_n) - \mathcal{F}(u, u_n)| \geq 0 \) as \( \mathcal{F}(u, \cdot) \) is weakly l.s.c. on \( X \) because it is convex and strongly continuous.

Also \( \lim_{n \to \infty} |\mathcal{F}(u_n, u_n) - \mathcal{F}(u, u_n)| = 0 \) by assumption since \( \{u_n\} \) is bounded on \( X \). Thus \( \liminf_{n \to \infty} (f(u_n) - f(u)) \geq 0 \) or \( f \) is weakly l.s.c. on \( X \).

Let the primal problem \((\mathcal{P})\) be to minimize \( f \) on \( X \).

Let \( Y = L^p(\Omega; \mathbb{R}^n) \) be the real vector space of all functions \( u : \Omega \to \mathbb{R}^n \) having their components \( u_i(x) \in L^p(\Omega) \). \( Y \) is a Banach space under the norm

\[
\|u\|_p^p = \sum_{i=1}^n \int_\Omega |u_i(x)|^p \, dx.
\]

Let \( A : X \to Y \) be the continuous linear operator defined by

\[
Au(x) = (D_1 u(x), D_2 u(x), \ldots, D_n u(x)) = \text{grad } u(x).
\]

Let the pairing between \( X \) and \( X^* \) be

\[
\langle u, v \rangle = \int_\Omega u(x) v(x) \, dx.
\]

Then \( X^* = W^{-1,p'}(\Omega) \) is a space of distributions and \( Y^* = L^{p'}(\Omega; \mathbb{R}^n) \) with \( p' = p/(p - 1) \). The adjoint operator \( A^* : Y^* \to X^* \) is defined by

\[
A^* v = - \sum_{i=1}^n D_i v_i = - \text{div } v,
\]

where the derivatives are defined distributionally.

With these definitions, the Lagrangian given by (2.8) is \( L : X \times Y^* \to \mathbb{R} \) with

\[
L(u, v) = \langle Au, v \rangle - \sup_{w \in Y} \int_\Omega \left[ \sum_{i=1}^n v_i(x) w_i(x) - F(x, u(x), w(x)) \right] \, dx
\]

\[
=: \langle Au, v \rangle - \int_\Omega F^0(x, u(x), v(x)) \, dx,
\]

(12.3)

where \( F^0(x, u, \eta) = \sup_{\xi \in \mathbb{R}^n} [\langle \xi, \eta \rangle - F(x, u, \xi)] \).
This is a Lagrangian of type I. The dual functional \( g: Y^* \to \mathbb{R} \) is defined by
\[
g(v) = \inf_{u \in X} L(u, v) = -\int_{\Omega} F^*(x, \text{div } v(x), v(x)) \, dx \tag{12.4}\]
whenever \( \text{div } v \in L^{p'}(\Omega) \).

If \( L(\cdot, v) \) is bounded above on \( X \) for some \( v \in L^p(\Omega) \) one might also define the anomalous dual functional \( h: Y^* \to \mathbb{R} \) by
\[
h(u) = \sup_{v \in X} L(u, v). \tag{12.5}\]

The standard dual problem \((\mathcal{P}^*)\) is to maximize \( g \) on \( Y^* \) and then the problem of finding solutions of \((\mathcal{P})\) and \((\mathcal{P}^*)\) is equivalent to the problem of finding saddle-points of the Lagrangian (12.3).

The anomalous dual problem \((\mathcal{P}^\otimes)\) is to minimize \( h \) on \( Y^* \).

Assume that \( F \) has the special form
\[
F(x, u, \xi) = \frac{1}{2} \sum_{i=1}^{n} \xi_i^2 - K(x, u), \tag{12.6}\]
where (iii) there exist constants \( \gamma, C_1 \) and \( C_2 \) with \( 0 \leq \gamma < 1 \) such that
\[
|K(x, u)| \leq C_1 |u|^{1+\gamma} + C_2 \quad \text{on } \Omega \times \mathbb{R},
\]
and (iv) \( K(x, \cdot) \) is convex on \( \mathbb{R} \) for all \( x \) in \( \Omega \).

Take \( p = 2 \) in the preceding analysis and let \( f_2: X \to \mathbb{R} \) be defined by
\[
f_2(u) = -\int_{\Omega} K(x, u(x)) \, dx.
\]

Then the anomalous dual functional \( h: L^2(\Omega; \mathbb{R}^n) \to \mathbb{R} \) is given by (see Section 2, Case 1).
\[
h(v) = -\frac{1}{2} \langle v, v \rangle + \sup_{u \in X} \left[ \langle Au, v \rangle - \int_{\Omega} K(x, u(x)) \, dx \right], \tag{12.7}
\]
where \( \langle \cdot, \cdot \rangle \) is the inner product on \( L^2(\Omega; \mathbb{R}^n) \).

Thus \( h \) will be the difference of two convex functionals on \( Y \).

It is somewhat more informative to treat this particular case using a Lagrangian of type II which yields some different dual problems. Before doing this, we should note the basic existence results for \((\mathcal{P}^\otimes)\) when \( F \) is given by (12.6).

**Theorem 12.2.** Suppose \( f: W^{1,2}_0(\Omega) \to \mathbb{R} \) is defined by (12.2) and (12.5)
and $K$ obeys (iii). Then $f$ is strongly continuous and weakly l.s.c. on $W^{1,2}_0(\Omega)$. In (iii), if $0 \leq \gamma < 1$, or if $\gamma = 1$ and $C_1$ is sufficiently small, then $(\mathcal{P})$ has a solution.

**Proof:** From Lemma 12.1, $f$ is strongly continuous. Suppose $\{u_n: n \geq 1\}$ converges strongly to $u$ in $L^2(\Omega)$. Then

\[
\lim_{n \to \infty} \int_{\Omega} |F(x, u_n(x), Du_n(x)) - F(x, u, Du(x))| \, dx
\]

\[
= \lim_{n \to \infty} \int_{\Omega} |K(x, u(x)) - K(x, u_n(x))| \, dx = 0
\]
as $R(u) = \int_{\Omega} K(x, u(x)) \, dx$ is strongly continuous on $L^2(\Omega)$ when $K$ obeys (iii) (see Lemma 10.1). Thus from Lemma 11.1, $f$ is weakly l.s.c. on $X = W^{1,2}_0(\Omega)$. Thus to prove $(\mathcal{P})$ has a solution, since $X$ is reflexive, it suffices to show that $f$ is coercive.

Since $\Omega$ is bounded and $\partial \Omega$ is smooth, Poincaré's inequality states that there is a constant $\lambda_1$ depending only on $\Omega$ such that

\[
\int_{\Omega} |\nabla u(x)|^2 \, dx \geq \lambda_1 \int_{\Omega} u(x)^2 \, dx.
\]

Thus

\[
f(u) \geq \frac{\lambda_1}{2} \int_{\Omega} u(x)^2 \, dx - C_1 \int_{\Omega} |u(x)|^{1+\gamma} \, dx - C_2 |\Omega|
\]

from (iii). When $0 \leq \gamma < 1$ and given $\varepsilon > 0$, there exists $C_3(\varepsilon)$ such that

\[
|u(x)|^{1+\gamma} \leq \varepsilon |u(x)|^2 + C_3(\varepsilon)
\]

for all $x$ in $\Omega$. Choose $\varepsilon = \lambda_1/4C_1$; then (12.8) implies

\[
f(u) \geq \frac{\lambda_1}{4} \int_{\Omega} |u(x)|^2 \, dx - C_4 |\Omega|
\]

where $C_4$ is a constant depending on $\lambda_1$, $C_1$, $C_2$ and $C_3$. Thus $f$ is coercive and so $(\mathcal{P})$ has a solution.

When $\gamma = 1$, $f$ will be coercive provided $0 \leq C_1 < \frac{1}{2} \lambda_1$.

Now consider the functional $\tilde{f}: L^2(\Omega) \to \mathbb{R}$ defined by

\[
\tilde{f}(u) = f(u) \quad \text{if} \quad u \in W^{1,2}_0(\Omega),
\]

\[
= +\infty \quad \text{otherwise},
\]

\[
= f_1(u) - R(u),
\]
where \( f_i \) is defined by
\[
f_i(u) = \begin{cases} \frac{1}{2} \int_{\Omega} |\nabla u(x)|^2 \, dx & \text{if } u \in W_0^{1,2}(\Omega), \\ +\infty & \text{otherwise}, \end{cases}
\]
and \( \mathcal{K} \) is defined by
\[
\mathcal{K}(u) = \int_{\Omega} K(x, u(x)) \, dx.
\]

Let \( (\mathcal{P}) \) be the problem of minimizing \( \mathcal{F} \) on \( L^2(\Omega) \). One sees that \( (\mathcal{P}) \) has a solution iff \( (\mathcal{P}) \) does and that the values of \( (\mathcal{P}) \) and \( (\mathcal{P}) \) are the same.

In the definitions of Section 2, take \( X = Y = L^2(\Omega) \) and let \( A \) be the identity. Define \( L: L^2(\Omega) \times L^2(\Omega) \to \mathbb{R} \) by
\[
L(u, v) = \int_{\Omega} u(x) v(x) \, dx + f_i(u) + \mathcal{K}(v). \tag{12.9}
\]

**Lemma 12.3.** If \( K \) obeys (iii) and (iv), then \( L \) defined by (12.9) is a Lagrangian of type II for \( (\mathcal{P}) \).

**Proof.** One has
\[
\inf_{v \in L^2(\Omega)} L(u, v) = f_i(u) - \sup_{v \in L^2(\Omega)} \left[ -\int_{\Omega} uv \, dx - \mathcal{K}^*(v) \right] = f_i(u) - \mathcal{K}^*(u).
\]

When \( k \) obeys (iii) and (iv), then \( \mathcal{K} \) is strongly continuous and convex, hence weakly l.s.c. on \( L^2(\Omega) \). Thus \( \mathcal{K}(u) = \mathcal{K}^*(u) \) for all \( u \) in \( L^2(\Omega) \) or \( L \) is a Lagrangian of type II for \( \mathcal{F} \).

Just as in Section 10 (see Eq. (10.13)) one has \( \mathcal{K}^*: L^2(\Omega) \to \mathbb{R} \) is given by
\[
\mathcal{K}^*(v) = \int_{\Omega} K^*(x, v(x)) \, dx. \tag{12.10}
\]

The dual functional \( g: L^2(\Omega) \to \mathbb{R} \) arising from (12.9) is
\[
g(v) = \inf_{u \in L^2(\Omega)} \left[ \int_{\Omega} u(x) v(x) \, dx + f_i(u) \right] + \mathcal{K}^*(-v).
\]

The first term on the right-hand side of this is
\[
g_1(v) = \inf_{u \in W_0^{1,2}(\Omega)} \left[ \int_{\Omega} u(x) v(x) \, dx + \frac{1}{2} \int_{\Omega} |\nabla u(x)|^2 \, dx \right].
\]

This is the Dirichlet principle for finding solutions of Poisson's equation \( \Delta u(x) = v(x) \) in \( \Omega \) subject to \( u(x) = 0 \) on \( \partial \Omega \).
DUALITY FOR NON-CONVEX VARIATIONAL PRINCIPLES

From the theory of elliptic boundary value problems (see [5, Chap. 8]) this system has a unique solution \( u = -Gv \in W_0^{1,2}(\Omega) \cap W^{2,2}(\Omega) \). One may prove \( G: L^2(\Omega) \to L^2(\Omega) \) is a compact self-adjoint linear operator with \( \langle Gw, w \rangle > 0 \) for all \( w (\neq 0) \) in \( L^2(\Omega) \).

Thus \( g_1(v) = -\frac{1}{2} \langle Gv, v \rangle \) is concave on \( L^2(\Omega) \) and

\[
g(v) = \mathcal{R}^*(-v) - \frac{1}{2} \langle Gv, v \rangle \tag{12.11}
\]

is the difference of two convex functionals.

The dual problem \((\mathcal{R}^*)^*\) is to minimize \( g \) on \( L^2(\Omega) \).

**Theorem 12.4.** Suppose \( g \) is defined by (12.11), then \( g \) is weakly l.s.c. on \( L^2(\Omega) \). If \( K \) obeys (iii) with \( 0 < \gamma < 1 \), or with \( \gamma = 1 \) and \( 0 \leq C_1 < \frac{1}{2} \lambda_1 \), then \((\mathcal{R}^*)^*\) has a solution.

**Proof.** \( \mathcal{R}^*(-v) \) is weakly l.s.c. on \( L^2(\Omega) \) because it is a polar. Since \( G \) is compact, \( g_1 \) is weakly continuous and thus \( g \) is weakly l.s.c.

From (iii), \( uv - K(x, u) \geq uv - C_1 |u|^{\gamma+1} - C_2 \) for all \( u \in \mathbb{R}^+ \). Thus

\[
\sup_{u \in \mathbb{B}^+} [uv - K(x, u)] \geq \sup_{u \in \mathbb{R}^+} (uv - C_1 |u|^{\gamma+1} - C_2)
\]

\[
= C_3 |v|^{(1+\gamma)/\gamma} - C_2,
\]

where \( C_3 = C_4(\gamma)/C_1^{1/\gamma} > 0 \) and \( C_4(1) = \frac{1}{4} \).

Thus \( K^*(x, v(x)) \geq C_3 |v(x)|^{(\gamma+1)/\gamma} - C_2 \) for all \( x \in \Omega \) and

\[
g(v) \geq C_3 \int_\Omega |v(x)|^{(\gamma+1)/\gamma} dx - C_2 |\Omega| - \frac{1}{2\lambda_1} \int_\Omega |v(x)|^2 dx,
\]

where we have used (12.10) and the fact that \( \langle Gv, v \rangle \leq (1/\lambda_1) \int_\Omega |v(x)|^2 dx \), where \( \lambda_1 \) is defined by Poincaré's inequality in Theorem 12.2.

When \( 0 < \gamma < 1 \), let \( q = (\gamma + 1)/\gamma \). Then \( q = 1 + 1/\gamma > 2 \) and for any \( C > 0 \), there exists a real \( M \) such that

\[
|v|^q \geq C |v|^2 + M \quad \text{for all} \ v \in \mathbb{R}^2.
\]

Thus

\[
g(v) \geq (C_3 C^q - \frac{1}{2\lambda_1}) \int_\Omega |v(x)|^2 dx + (C_3 M - C_2) |\Omega|.
\]

Choose \( C \) sufficiently large that \( C_3 C > 1/2\lambda_1 \); then \( g \) is coercive on \( L^2(\Omega) \) and hence \((\mathcal{R}^*)^*\) has a solution.

When \( \gamma = 1 \), then \( (\gamma + 1)/\gamma = 2 \) and \( g \) will be coercive provided \( C_3 > 1/2\lambda_1 \). Using the expression for \( C_3 \) one finds \( g \) is coercive precisely
when $C_1 < \frac{1}{2} \lambda_1$ which was exactly the condition obtained in Theorem 12.2 for $f$ to be coercive. Hence the result.

**Corollary.** Suppose $K$ obeys (iii) and (iv), and in (iii) either $0 < \gamma < 1$ or $\gamma = 1$ and $0 < C_1 < \frac{1}{2} \lambda_1$. Then the Lagrangian $L$ defined by (12.9) is bounded below on $L^2(\Omega) \times L^2(\Omega)$ and there exists $(\bar{u}, \bar{v})$ in $L^2(\Omega) \times L^2(\Omega)$ which minimizes $L$.

**Proof:** From Lemma 12.3 one knows $L$ is a Lagrangian of type II and from Theorems 12.2 and 12.4 we know that $(\mathcal{P})$ and $(\mathcal{P}^*)$ have finite values and have solutions. Theorem 3.3 then yields this corollary.

This corollary implies that $(\bar{u}, \bar{v})$ is a $\dot{\psi}$-critical point of this Lagrangian. That is, there is a solution of the equations

$$-v \in \partial f_1(u),$$

$$u \in \partial \mathcal{H}^*(-v).$$

Note that, in this case, $f_1$ is not Gâteaux-differentiable at any point in its domain and $\mathcal{H}^*$ need not, in general, be Gâteaux-differentiable but these subdifferentials are always well defined.

It is also worth noting that the dual problem here is similar to the problem of minimizing $\mathcal{F}$ defined by (10.10) with $p = 2$, $\lambda = 0$ but with $-G$ in place of $G$! Also Toland’s example, described at the end of Section 10, fits into this framework with $n = 1$, $\Omega = (0, 1)$.

**Acknowledgments**

Much of this work was done while the author was on leave at the Mathematical Institute of Oxford University and he would like to thank the Institute and Wadham College for their hospitality.

**References**


