SHADOW PRICES, DUALITY AND GREEN'S FORMULA FOR A CLASS OF OPTIMIZATION PROBLEMS

NOV 78 J P AUBIN, F H CLARKE

DAA 6249-75-C-0024
MRC Technical Summary Report #1900

SHADOW PRICES, DUALITY AND GREEN'S FORMULA FOR A CLASS OF OPTIMAL CONTROL PROBLEMS

J. P. Aubin and F. H. Clarke

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

November 1978

Received September 20, 1978

Approved for public release
Distribution unlimited

Sponsored by
U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

National Research Council of Canada
Vancouver, B. C. Canada V6T 1W5
A class of optimal control problems is considered in which the cost functional is locally Lipschitz (not necessarily convex or differentiable) and the dynamics linear and/or convex. By using generalized gradients and duality methods of functional analysis, necessary conditions are obtained in which the dual variables admit interpretation as shadow prices (or rates of change of the value function). Applications are presented in three settings: infinite-horizon optimal control, optimal control of partial differential equations, and a variational problem with unilateral state constraints. A theorem is proved which characterizes the generalized gradients of integral functionals on $L^p$.

AMS(MOS) Subject Classification: 49B30

Key Words: maximum principle, sensitivity, infinite horizon, generalized gradients

Work Unit No. 5 - Mathematical Programming and Operations Research
SIGNIFICANCE AND EXPLANATION

We develop in this paper a theory of optimality conditions for an abstract optimal control problem which encompasses a variety of situations that are met in practice, but not covered by the standard theory. Three such specific applications are made. The first of these deals with a control problem over an infinite interval, while the second and third involve the control of partial differential equations and state constraints, respectively. It is shown in each case that the general theory yields new and stronger necessary conditions, two significant features of which are the capability of treating nondifferentiable functions, and the interpretation of a certain quantity as the sensitivity of the problem with respect to certain perturbations.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.
Introduction

We treat in this article a general class of optimal control problems in which the cost functional need only be locally Lipschitz, and in which the constraints exhibit linearity and convexity. We obtain for these hybrid problems necessary conditions with features usually associated only with the fully convex case. Chief among these is the interpretation of the multipliers as rates of change of the optimal value in the problem as the constraints undergo perturbation. In the context of mathematical programming such interpretations have figured importantly in mathematical economics (see for example [13] for a discussion), where the rates of change are in turn interpreted as "shadow prices". A further feature is the ability to guarantee the strong (Kuhn-Tucker or normal) form of the necessary conditions by means of a Slater-type constraint qualification independent of the solution to the problem. The method used here to treat these problems is novel, and employs the generalized gradient of the value function, a minimax theorem, and an abstract Green formula.

We present three applications of the abstract theory which we believe to be of independent interest; they are in settings characterized by technical difficulties. The first of these ($\S$1) involves a control problem over an infinite interval, and sheds new light on the sensitive relationship between the growth rates of the cost functional and the adjoint variables, and on the role played by the size of the discount factor. In $\S$2 we give an example to illustrate the importance of this consideration, while $\S$3 gives the proof of the theorem in $\S$1. We have placed this application before the abstract theory in order to display the line of reasoning common to both in a more concrete, and hence more easily assimilated setting.

Section 4 recalls a stability result used in the proof, while $\S$5 and 6 develop characterizations of generalized gradients of certain functionals which are useful in interpreting the abstract conditions; these results complement other characterizations given in [9]. The abstract problem and its analysis appear in $\S$7 and 8. In $\S$9 an application is made to a non-differentiable, nonconvex problem involving control of partial differential equations. While this is the first such result that we know of, it should be clear that many other similar problems could be framed within the abstract theory. The final section illustrates the use of the theory in the presence of unilateral state constraints.

We conclude by recalling for the reader's convenience the definition of the generalized gradient in the case of a locally Lipschitz function $f: X \rightarrow \mathbb{R}$, where $X$ is a Banach space (details appear in [7] [9]). Given $v$, the generalized directional derivative $f^*(x;v)$ is defined by

$$ f^*(x;v) = \limsup_{y \to x} \frac{f(y) - f(x)}{\|y-x\|}, $$

where the upper limit is taken as $y$ in $X$ converges to $x$ and $\lambda$ in $(0,\infty)$ converges to 0. The generalized gradient of $f$ at $x$, $\partial f(x)$, consists of all $\xi$ in $X^*$ such that

$$ \langle v, \xi \rangle \leq f^*(x;v) \quad \text{for all } v \quad \text{in } X. $$

It follows that

$$ f^*(x;v) = \max \{ \langle v, \xi \rangle : \xi \in \partial f(x) \}, $$

and that $\partial f(x)$ reduces to the derivative if $f$ is continuously differentiable or to the subdifferential of convex analysis if $f$ is convex.
1. An infinite-horizon optimal control problem.

We are given a locally Lipschitz function \( g: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \), i.e., whenever \((x, u), (y, v)\) are restricted to a suitably chosen neighborhood of any given point in \( \mathbb{R}^n \times \mathbb{R}^n \), there is a constant \( K \) such that

\[
|g(x, u) - g(y, v)| \leq K|x - y|.
\]

Also on hand are matrices \( F \) and \( G \), \( n \times n \) and \( n \times n \) respectively, a point \( x_0 \) in \( \mathbb{R}^n \), a positive number \( \delta \) and a compact convex subset \( U \) of \( \mathbb{R}^n \).

The problem we consider is that of minimizing

\[
\int_0^\infty e^{-\delta t} g(x(t), u(t)) dt
\]

over the (so-called) trajectory-control pairs \((x, u)\) which satisfy

\[
\begin{align*}
\dot{x}(t) &= Fx(t) + Gu(t), \text{ a.e. } t \geq 0 \\
x(0) &= x_0
\end{align*}
\]

where \( u(\cdot) \) need only be measurable and \( x \) almost surely continuous.

We posit the following growth condition on \( g \): there are numbers \( c, \tau \geq 0 \) such that, for every \((x, u)\), for every \( \xi \) in the generalized gradient \( \partial g(x, u) \) of \( g \) at \((x, u)\), we have

\[
|\xi| \leq c|\xi|g(x, u)|^{\tau}.
\]

**Definition.** We denote by \( a(a) \) the infimum in the above problem when the initial condition is \( x(0) = x_0 \) (rather than \( x(0) = x_0^* \)), and by \( \lambda(F) \) the maximum of the real parts of the eigenvalues of \( F \).

We now suppose given a trajectory-control pair \((x, u)\) which solves the above problem (with the original initial condition \( x(0) = x_0 \)).

**Theorem 1.** Let \((x, u)\) solve the problem as described above, and suppose \( \delta > (r+1)\lambda(F) \). Then \( a(\cdot) \) is locally Lipschitz, and there exist an absolutely continuous function \( p(\cdot) \) and a measurable function \((c_1(\cdot), c_2(\cdot))\) such that:

\[
\begin{align*}
\dot{p}(t) &= p(t) - e^{-\delta t} c_1(t) \quad \text{a.e.} \\
(c_1(t), c_2(t)) &\in \partial g(x(t), u(t)) \quad \text{a.e.} \\
\max\{p(t) + c_2(t) : w \in U\} &= \min\{p(t), c_1(t)\} - e^{-\delta t} (c_1(t), c_2(t)) \quad \text{a.e.} \\
\int_0^\tau e^{-(r+1)\delta t}|p(t)|^q dt < \infty, \quad \int_0^\tau e^{-(r-1)\delta t}|p(t)|^q dt < \infty, \quad \text{where } q > 1
\end{align*}
\]

is defined by \( 1/q + 1/(r+1) = 1 \), provided \( r > 0 \). If \( r = 0 \), we have instead: \( e^{\delta t}|p(t)| \) and \( e^{\delta t}|p(t)| \) are bounded.

\[
\begin{align*}
\lim_{t \to \infty} e^{-(r+1)\delta t}|p(t)|^q &= 0 \text{ if } r > 0, \text{ if } r = 0, \text{ then we have instead: } \\
e^{\delta t}p(t) \text{ tends to a finite limit as } t \text{ goes to } \infty.
\end{align*}
\]

1.10 Remarks (a) As shown in the proof, our hypotheses imply that the integral (1.1) is well-defined for all admissible pairs \((x, u)\), so that no ambiguities result from our use of the word "solve".

(b) Except for the infinite interval of integration (called the case of an "infinite horizon" in economics) and the nondifferentiability of the cost functional, the above problem is a standard one in optimal control, and of the sort that arises often in mathematical economics. It is
in this connection especially that the interpretation of \( p(0) \) as a marginal cost ("shadow price") associated with perturbing the initial condition, as afforded by (1.9), is useful. The roles of the infinite horizon and the discount factor are discussed in T.C. Koopmans [13]; see also [16].

(c) It is worth noting that in the above version of Pontryagin's maximum principle, the necessary conditions are stated in a strong form as one could hope; i.e. "normally", without the presence of a possibly vanishing multiplier, and with strong "transversality conditions" at infinity. Similar statements have been derived (incorrectly in many instances) by reasoning by analogy with the finite horizon case. It was H. Malkin [11], to our knowledge, who first pointed out the perhaps unexpected pathology that can arise due to the infinite horizon. Based on the results of this paper, one could say that the necessary conditions may be expected to hold in the strong form provided the "discount rate" \( \delta \) is sufficiently large.

A further moral is that in the infinite-horizon case, growth (or dual) conditions such as (1.7) on the adjoint variable \( \mu \) are more natural than pointwise transversality conditions such as (1.8) (which are simply consequences of (1.7)). This fact was foreshadowed in Theorem A.8 of [4], which can be obtained as a corollary of Theorem 1. Finally, it is well-known (and true) that the conclusions of the theorem are sufficient for \((s,u)\) to be optimal if \( g \) is jointly convex in \((s,u)\).

2. An example

Consider the control system on \( \mathbb{R}^2 \) given by

\[ \begin{align*}
\dot{x}_1 &= -x_2 \\
\dot{x}_2 &= -x_1 + u
\end{align*} \]

\[-1 \leq u \leq 0, \]

\[ x_1(0) = 0, \quad x_2(0) = 0. \]

with the view of minimizing \( \int_0^\infty e^{-\delta t} x_1(t) dt \). This is the case of the preceding section in which

\[ F = \begin{bmatrix} 0 & -1 \\
-1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 \\
0 \end{bmatrix}, \quad U = [-1,0]. \]

It follows that \( x_1 \) satisfies

\[ x_1(0) = \dot{x}_1(0) = 0, \quad x_1(t) \geq x_2(t), \]

so that, by a simple argument, \( x_2(t) \geq 0 \) for all \( t \geq 0 \). Consequently, the control \( u \equiv 0 \) and response \( x \equiv 0 \) are optimal for the problem.

The adjoint equation (1.4) becomes

\[ \dot{\mu}_1 = -\mu_2 - e^{-\delta t}, \quad \dot{\mu}_2 = -\mu_1, \]

and hence

\[ \mu_2 = \mu_2 + e^{-\delta t}. \]

The condition (1.6) yields the fact that \( \dot{x}_2(t) \) must equal \(-1\) if \( p_2(t) \) is negative. Thus we must have \( p_2(t) \geq 0 \) for all \( t \).
The solution of (2.1) is of the form

\[ p_2(t) = ae^t + be^{-t} + e^{-\delta t}/(\delta^2 - 1) \]

It is easy to see that if \( 0 < \delta < 1 \) and \( p_2(t) \) converges to \( 0 \) as \( t \to +\infty \) (as must be the case if (1.8) is to hold for any \( r > 0 \)), then \( p_2(t) \) is negative for large \( t \). This contradiction is not inconsistent with the theorem, since in this case \( \lambda(\gamma) = 1 \). Of course, if \( \delta > 1 \), the conclusions of the theorem hold.

Let us note finally that for \( \delta > 1 \), we may take \( r = 0 \) in applying the theorem, and it follows that \( p_2(t) \) must be \( e^{-\delta t}/(\delta^2 - 1) \). The condition (1.9) now gives

\[ (6/(\delta^2 - 1)) < \lambda(0,0), \]

which makes evident the increasingly unstable dependence on the initial condition of the problem as \( \delta \) nears 1.

### 3. Proof of Theorem 1

In this section, we denote the optimal pair \((\vec{x}, \vec{u})\).

**STEP 1** We denote by \( \mu \) the measure on \([0,\omega]\) having density \( e^{-\delta t} dt \). We set \( p = 1 + r \) and we let \( X \) be the space \( L^p_n \times L^p_{\mathbb{R}^N} \), where \( L^p_n \) means \( L^p([0,\omega], \mathbb{R}^N) \) (with respect to \( \mu \)) and where \( L^p_{\mathbb{R}^N} \) denotes the Sobolev space of absolutely continuous functions \( x \) such that \( x \) and \( \dot{x} \) belong to \( L^p_n \). (Recall that the norm of \( x \) in \( L^p_n \) is given by \( ||x|| + ||\dot{x}||_p \).)

We set \( Y = L^p_n \times L^p_{\mathbb{R}^N} \), and we define the subset \( Z \) of \( Y \) via

\[ Z = \{(0) \times K \times (\vec{x}_0)\}, \]

where \( K \) is the set of functions \( u \) in \( L^p_n \) satisfying

\[ u(t) \in U \text{ a.e.} \]

We define \( f : X \to Y \) by

\[ f(x,u) = \int_0^{\omega} e^{-\delta t} g(x(t), u(t)) dt, \]

and \( A : X \to Y \) by

\[ A(x,u) = \{\dot{x} - Fx - Gu, u, x(0)\}. \]

Theorem 3 of §6 implies that \( f \) is locally Lipschitz on \( L_n^p \times L_{\mathbb{R}^N}^p \), and so all the more on \( W^{1,p}_n \times L_{\mathbb{R}^N}^p \). The optimal control problem of §1 may be phrased as that of minimizing \( f(x,u) \) over \( X \) subject to \( A(x,u) \in Z \).

We shall denote the given solution to this problem \((\vec{x}, \vec{u})\). Proposition 4.1 of §4 will be available to us as soon as (4.1) is verified, and the reader may check that this is equivalent to the following result: (we shorten \( \lambda(F) \) to \( \lambda \)).
Lemma 3.1

If \( \delta > \lambda \) and \( v \) belongs to \( L^p_0 \), then the solution \( x \) to

\[
\dot{x} = Fx + v, \quad x(0) = x_0
\]

belongs to \( U_0^{1,p} \). In fact, \( \|x\| \leq c \|v\| \) for some constant \( c \).

Proof: It suffices to show that the function

\[
x(t) = e^{Ft}x_0 + \int_0^t e^{F(t-s)}v(s)ds
\]

belongs to \( L^p_0 \). Since \( |e^{Ft}| \) is bounded by \( e^{\lambda t} \), it is clear that \( e^{Ft}x_0 \)

belongs to \( L^p_0 \), so we need only study the last term. Now, choosing any \( q \)

between \( \lambda \) and \( \delta/p \), and limiting ourselves to the case \( r > 0 \) (the case \( r = 0 \) calls for straightforward modifications), we have

\[
\begin{align*}
&\int_0^r \int_0^t e^{F(t-s)}v(s)ds|P|dt \\
&\leq \int_0^r \int_0^t e^{F(t-s)}|v(s)||P|ds|dt \\
&\leq \int_0^r \int_0^t e^{(\lambda - \delta/p)t}e^{\lambda (q-\lambda)P|P|}dt \\
&\leq \int_0^r e^{(\lambda - \delta/p)t} \left[ \frac{1}{\gamma} e^{\frac{(q-\lambda)}{p}P|P|} \right] \left[ \frac{1}{\gamma} e^{\gamma \delta P|P|} \right] dt \\
&\leq \varepsilon \int_0^r e^{(\lambda - \delta/p)t}e^{\frac{\delta}{p}P|P|}dt \quad \text{(Hölder)} \\
&\leq \varepsilon \int_0^r e^{\frac{\delta}{p}P|P|}dt \quad \text{(Fubini)}
\end{align*}
\]

Q.E.D.

Proposition 4.1 now assures us that the function \( \alpha \) defined in §1 is Lipschitz in a neighborhood of \( x_0 \) (in fact, it says rather more about a function involving more arguments). Note that if \( (x,u) \) in \( X \) is such that

\[
A(x,u) \in (0) \times K, \quad A(x,u) = (\hat{x} - Fx - Gx,u),
\]

then

\[
f(x,u) = \alpha(x(0))
\]

with equality when \( (x,u) = (\hat{x},\hat{u}) \).

Thus \( (\hat{x},\hat{u}) \) minimizes

\[
f(x,u) = f(x,u) - \alpha(x(0))
\]

over \( X \) subject to \( \hat{A}(x,u) \in (0) \times K \). By redefining \( \hat{f} \) away from \( (\hat{x},\hat{u}) \), we can arrange to preserve this state of affairs while assuring that \( \hat{f} \) be locally Lipschitz on \( X \). We may now apply Proposition 4.1 again to deduce that the function \( h: L^p_0 \times L^p_0 \times K \) is Lipschitz in a neighborhood of \((0,0)\),

where \( K \) is the closed unit ball in \( K_0 \).

\[
h(v,w) = \inf(f(x,u) : A(x,u) \in (0) \times (v,w), x(0) = x_0 + B).
\]

**STEP 2** Let us now choose any \((x,u) \in X \) and \( w \in K \). Upon observing that for any \( \varepsilon > 0 \) sufficiently small

\[
\hat{A}(\hat{x},\hat{u}) + \varepsilon(x-\hat{x},u-\hat{u}) \in (0) \times K + \varepsilon(\hat{x} - Fx - Gx,u-u),
\]

\[
\hat{A}(x,u) \hat{A}(x,u) \in (0) \times B
\]

we deduce

\[
f((\hat{x},\hat{u}) + \varepsilon(x-\hat{x},u-\hat{u}) \hat{v} = h(v,w)) \]

After subtracting \( f((\hat{x},\hat{u}), \hat{v} = h(0,0) \) from either side, dividing by \( \varepsilon \) and
taking upper limits, we arrive at

\[-a^{-}(0,0; \hat{x} - Fx - Gw, u - w) + a^{-}(u_n x(0), x - u_n) \geq 0.\]

Recall [7](9) that \(f^{*}(x,v)\) equals \(\max \{<\zeta, v> \mid \zeta \in \text{dom}(x)\}\). Consequently,

\[
\begin{align*}
\min_{x,w} \max_{y > v} <\gamma_1, \hat{x} - Fx - Gw > + <\gamma_2, u - w > + <\zeta_1, x - \tilde{u} > \\
+ <\zeta_2, u - \tilde{u} > + <v, x(0) - \tilde{u} > = 0,
\end{align*}
\]

where the max is over all \((x,w)\) in \(X \times K\) and the min over all \(y = (\gamma_1, \gamma_2)\), \(\zeta = (\zeta_1, \zeta_2)\), and \(v\) lying in \(\hat{3}(0,0), 3(\tilde{u}, \tilde{u})\), and \(3(\tilde{u}, \tilde{u})\) respectively. Because these sets are \(w\)-compact, the "lop-sided" minimax theorem [1] applies, and we deduce the existence of \(\zeta, \gamma, v\) such that, for all \((x,u)\) in \(X\) and \(w\) in \(K\),

\[
<\gamma_1, \hat{x} - Fx - Gw > + <\gamma_2, u - w > + <\zeta_1, x - \tilde{u} > \\
+ <\zeta_2, u - \tilde{u} > + <v, x(0) - \tilde{u} > \geq 0.
\]

**STEP 3** Let us set \(w = u = \tilde{u}\) in (3.1). We then deduce

\[
<\gamma_1, \hat{x} - Fx - G\tilde{u} > + <\gamma_2, u - \tilde{u} > + <\zeta_1, x - \tilde{u} >
\]

for all \(x\) in \(U_1^{1,p}\). Because \(U_1^{1,p}\) is dense in \(L_1^{p}\) (this follows, for example, from the fact that \(P_n(0,\infty)\) is dense in \(L_1^{p}\)), it follows from Proposition 5.1, 15, that \(\gamma_1\) belongs to \(L_1^{p}\), the dual of \(L_1^{p}\), rather than merely to the dual of \(U_1^{1,p}\) (which is best avoided). Of course, \(\zeta_2, \gamma_2\) belong to \(L_1^{p}\), \(\gamma_1\) to \(L_0^{p}\), and \(v\) to \(E^p\). Thus we have, for all \(x\) in \(U_1^{1,p}\),

\[
\int_0 \varrho e^{-dt} \gamma_1(t) \cdot (\hat{x} - Fx)dt + \int_0 \varrho e^{-dt} \gamma_2(t) + x(t)dt + v \cdot x(0) = 0.
\]

A classical, now familiar argument (DuBois-Reymond lemma) employing integration by parts derives from this the conclusion that \(\gamma_1\) is absolutely continuous, and that

\[
\frac{d}{dt} (\varrho e^{-dt} \gamma_1(t)) = -\varrho e^{-dt} \gamma_1(t) + \varrho e^{-dt} \gamma_2(t),
\]

\(\gamma_1(0) = v\).

If we set \(p(t) = e^{-dt} \gamma_1(t)\), then the above immediately yields (1.4), (1.7), (1.9), while (1.5) is a consequence of Theorem 1, 16. Condition (1.8) is an elementary consequence of (1.7), the line of argument being that found in [4, Lemma A.7]. It remains to prove (1.6).

**STEP 4** Set \(x = \bar{x}\), \(w = \tilde{w}\) in (3.1). Then

\[
-<\gamma_1, G(w(\tilde{u})) > + <\gamma_2, u - \tilde{u} > + <\zeta_2, u - \tilde{u} > = 0
\]

for all \(u\) in \(L_1^p\). This implies

\[
(3.2) \quad e^{-dt} \gamma_2(t) = G^p(t) - e^{-dt} \zeta_2(t) \quad \text{a.e.}
\]

Now set \(x = \bar{x}\), \(w = \tilde{u}\) in (3.1). We deduce, for every \(v\) in \(K\),

\[
(3.3) \quad -<\gamma_2, v - \tilde{u} > \geq 0.
\]

It follows that for almost each \(v\), we have

\[
y_2(t) \cdot v \leq y_2(t) \cdot \tilde{u}(t) \quad \text{for all } w \text{ in } U
\]

(for if this were false we could, from the measurable selection theorem,
find \( w(\cdot) \) in \( K \) contradicting (3.1). Recalling (3.2), we arrive at

\[
(Cp(t) - e^{-\delta t}C_p(t)) \cdot (\nu - \tilde{\nu}(t)) \leq 0,
\]

which is (1.6). Q.E.D.

4. A stability result recalled

Let \( X \) and \( Y \) be Banach spaces and \( A: X \rightarrow Y \) a continuous linear operator. Consider, for \( \epsilon \) in a neighborhood of \( 0 \) in \( Y \), the family of optimization problems consisting of minimizing a given function \( F: X \rightarrow \mathbb{R} \) subject to the constraints

\[
x \in \Omega
\]

\[
Ax \in Z + \epsilon
\]

where \( \Omega, Z \) are given subsets of \( X \) and \( Y \) respectively. We denote by \( C(\epsilon) \) the set of \( x \) satisfying these constraints, and by \( a(\epsilon) \) the infimum in the above problem. Thus

\[
a(\epsilon) = \inf\{F(x) : x \in C(\epsilon)\}.
\]

The following, a special case of [3, Proposition 2], is stated here for the reader’s convenience.

**Proposition 4.1** Let \( Z \) and \( \Omega \) be closed and convex, and \( Y \) locally Lipschitz. Suppose that there is a bounded subset \( K \) such that \( C(\epsilon) \subseteq \Omega \) when \( \epsilon \) is near \( 0 \), and such that \( F \) is Lipschitz on \( K \). Then, if the condition

\[
(\epsilon, 1)
\]

\[
0 \in \text{int}(\epsilon_0 - Z)
\]

is satisfied, and if \( a(0) = \rightarrow \), the function \( a \) is Lipschitz near \( 0 \).
5. A result on generalized gradients

Let \( X \) and \( Y \) be Banach spaces such that \( X \) is continuously
embedded in \( Y \) and such that \( X \) as a subset of \( Y \) is dense. Let \( f : Y \to \mathbb{R} \)
be a function which is locally Lipschitz. It follows that its restriction
to \( X \) is locally Lipschitz in the norm of \( X \). We denote by \( \mathcal{M}_X \) the
generalized gradient (see below) (in \( X^\ast \)) of this restriction, and by \( \mathcal{M}_Y \) the generalised gradient (in \( Y^\ast \)) of the function \( f \) defined on \( Y \).

**Proposition 5.1** Let \( x \) be a point of \( X \). Then

\[
\mathcal{M}_X(x) \subseteq \mathcal{M}_Y(x)
\]

in the sense that every \( \xi \) in \( \mathcal{M}_X(x) \) admits a unique extension to an
element of \( \mathcal{M}_Y(x) \).

**Proof:** Recall [7][9] the generalised directional derivative \( f'_X(x;v) \),
defined by

\[
f'_X(x;v) = \limsup(f(y + \lambda v) - f(y))/\lambda,
\]

where the upper limit is taken as \( y \) converges to \( x \) in \( Y \) and \( \lambda \) decreases
to 0 in \( \mathbb{R} \). The generalised gradient \( \mathcal{M}_X(x) \) consists by definition of
those elements \( \xi \) in \( Y^\ast \) satisfying

\[
f'_X(x;v) \geq \langle \xi, v \rangle \quad \text{for all } v \in Y.
\]

It follows easily from density and from the fact that convergence
in \( X \) implies convergence in \( Y \) that

\[
f'_X(x;v) \leq f'_Y(x;v)
\]

whenever \( x \) and \( v \) lie in \( X \). Now let \( \xi \) belong to \( \mathcal{M}_X(x) \). Then (by
definition)

\[
\langle \xi, v \rangle \leq f'_X(x;v) \quad \text{for all } v \in X,
\]

and by the preceding, along with the fact that the function \( v \to f'_X(x;v) \)
is bounded on bounded subsets of \( Y \), we deduce that the function

\[
v \to \langle \xi, v \rangle
\]

mapping \( X \), with its topology induced by \( Y \), to \( \mathbb{R} \) is bounded on bounded
subsets. It follows that this linear function is continuous on \( X \) (with
the induced topology) and hence, by a standard argument, admits a unique
linear extension to the complete space \( Y \) in which \( X \) is dense. The
extension, which we also label \( \xi \), still satisfies

\[
\langle \xi, v \rangle \leq f'_Y(x;v)
\]

for all \( v \) in \( Y \), and hence belongs to \( \mathcal{M}_Y(x) \) by definition.

Q.E.D.
6. Generalized gradients on $L^p$

We are given a complete measure space $(T, T, u)$ with $u(T) < +\infty$, a separable Banach space $X$, and a function $g : T \times X \rightarrow \mathbb{R}$. We assume that the mapping $t \mapsto g(t, x)$ is measurable for each $x$, that $x \mapsto g(t, x)$ is locally Lipschitz for each $t$, and we posit the existence of scalars $p \geq 1$ and $c > 0$ such that for every $(t, x)$, for every element $\zeta$ of $\partial_d g(t, x)$ (generalized gradient in $x$ for $t$ fixed), the following bound holds:

\[
|\zeta| \leq c(1 + |x|^{p-1}) .
\]

Finally, we suppose that $t \mapsto g(t, 0)$ is (finitely) integrable, and we set, for $x$ in $L^p(T, X)$,

\[
F(x) = \int_T g(t, x(t)) u(dt) .
\]

As usual, $p_\alpha$ denotes the (possibly infinite) quantity satisfying

\[
\frac{1}{p} + \frac{1}{p_\alpha} = 1 .
\]

**Theorem 2.** Under the above hypotheses, the function $F : L^p \rightarrow \mathbb{R}$ given by (6.2) is well-defined (finitely) and locally Lipschitz (in fact, Lipschitz on bounded subsets of $L^p$), and we have

\[
\delta F(x) \subset \partial_d g(t, x(t)) u(dt) .
\]

This means that corresponding to any element $\zeta$ of $\partial F(x)$, there is a function $\zeta(\cdot)$ in $L^{p_\alpha}(T, X)$ satisfying

\[
\zeta(t) \in \partial_d g(t, x(t)) \text{ u.a.e.}
\]

and such that, for every $y$ in $L^p(T, X)$,

\[
<\zeta, y> = \int_T \zeta(t), y(t) u(dt) .
\]

**Proof.** The growth condition (6.1) is easily seen to imply

\[
|g(t, x)| \leq |g(t, 0)| > c(|x| + |x|^p) ,
\]

which, combined with the fact that $t \mapsto g(t, x(t))$ is measurable, yields the fact that $F(x)$ is defined and finite whenever $x$ lies in $L^p$. We now prove that $F$ is locally Lipschitz. Invoking the mean value theorem for generalized gradients (14) we have

\[
|F(x) - F(y)| = \int_T |\zeta(t), x(t) - y(t)| u(dt)
\]

where $x(t)$ belongs to $\partial_d g(t, w(t))$ and $y(t)$ lies between $x(t)$ and $y(t)$. Using (6.1) and Hölder's inequality, we have this last expression bounded above by

\[
\begin{align*}
&c \sqrt{\frac{1}{p} + \frac{1}{p_\alpha}} |x-y| + c \sqrt{\frac{1}{p} + \frac{1}{p_\alpha}} \sqrt{\frac{1}{p_\alpha}} |x-y|_p \\
&\leq c_1 |x-y|_p + c \frac{|(|x| + |y|)|^{p-1}|}{p_\alpha} \sqrt{\frac{1}{p} + \frac{1}{p_\alpha}} \sqrt{\frac{1}{p_\alpha}} |x-y|_p \\
&\leq c_2 + c \frac{|(|x| + |y|)|^{p-1}|}{p_\alpha} \sqrt{\frac{1}{p} + \frac{1}{p_\alpha}} \frac{1}{p_\alpha} |x-y|_p
\end{align*}
\]

as long as $x$ and $y$ remain in a bounded subset of $L^p$.

Now let $\zeta$ belong to $\partial F(x)$ (and hence to the dual of $L^p(T, X)$, $L^{p_\alpha}(T, X)$ (5, 12.6, Prop. 10 and No. 21)). Then for any $v$ in $L^p$,
\( F^*(x;v) \geq \langle \xi, v \rangle \). Using Fatou's Lemma (cf. [8, Lemma 3]) we may show

\[
\int \mathcal{B}^*_x(t, x(t); v(t)) u(dt) \geq F^*(x;v),
\]

from which ensues

\[(6.3) \quad \int \mathcal{B}^*_x(t, x(t); v(t)) u(dt) \geq \langle \xi(t), v(t) \rangle u(dt) .\]

For any \( \varepsilon > 0 \), define a multifunction \( \Gamma \) as follows:

\[
\Gamma(t) = (0) \text{ if } \mathcal{B}_x^*(t, x(t); v) > \langle \xi(t), v \rangle - \varepsilon \text{ for all } v \text{ in } X ,
\]

\[
= (v; \mathcal{B}_x^*(t, x(t); v) \leq \langle \xi(t), v \rangle - \varepsilon) \text{ otherwise .}
\]

The map \((t, v) \rightarrow \mathcal{B}_x^*(t, x(t); v)\) is of course continuous in \( v \), and is seen to be measurable in \( t \) as a consequence of the fact that \( \mathcal{B}_x^*(t, x(t); v) \) can be expressed as the upper limit of a countable family of measurable functions (we use here the fact that \( X \) is separable).

It follows that the multifunction \( \Gamma \) is "measurable" and admits a measurable selection \( v(t) \) [19, Theorem 4.1]. Now (6.3) implies that the set

\[
(t; \Gamma(t) \neq 0)
\]

must have \( \mu \)-measure 0, and since \( \varepsilon \) is arbitrary we deduce that for \( \mu \)-almost all \( t \), for all \( v \) in \( X \),

\[
\mathcal{B}_x^*(t, x(t); v) \geq \langle \xi(t), v \rangle .
\]

Consequently \( \xi(t) \) belongs to \( B_x g(t, x(t)) \) \( \mu \)-a.e.

Q.E.D.

Remark. The case \( p = * \) is treated in [9]. As in that case, equality will hold in the theorem if \( g(t, \cdot) \) is "regular" for \( \mu \)-almost all \( t \).
7. An abstract approach

A. The line of reasoning used to prove Theorem 1 will work in much more
general situations, as we now show by considering the following abstract
optimization problem. We are given Banach spaces \( U, V, W, T, Z \) together with
continuous linear operators \( L : W \to V, \gamma : W \to T, C : U \to Z \), locally Lipschitz
functions \( f : U \to R \) and \( g : T \to R \), and a multifunction \( E \) mapping \( U \) to
\( V \) (i.e. \( E(x) \) is a subset of \( V \) for \( x \in U \)). \( W \), the domain of \( L \),
is assumed to be a subset of \( U \).

We consider the problem of minimizing

\[
(7.1) \quad f(x) + g(\gamma x)
\]

over the elements \( x \) of \( W \) which satisfy

\[
(7.2) \quad Lx \in E(x)
\]
\[
(7.3) \quad \gamma x \in S
\]
\[
(7.4) \quad Cx \in Y,
\]

where \( S \) and \( Y \) are specified subsets of \( T \) and \( Z \) respectively. Here,
\( L \) plays the role of a differential operator and \( \gamma \) that of a trace operator,
so that (7.1) is a type of Bolza functional, (7.2) is a "differential inclusion",
(7.3) a boundary condition, and (7.4) an explicit "state constraint".

An implicit constraint is also incorporated by (7.2), which demands that \( x \)
belong to the domain of \( E \); i.e. the set of points \( x \) for which \( E(x) \neq \emptyset \).

We posit the following conditions:

(\( W_1 \)) The sets \( S, Y, \) and \( \text{Gr}(E) \) are closed and convex (\( \text{Gr}(E) \), the graph of
\( E \), is the set of points \((x,v)\) such that \( v \) belongs to \( E(x) \)).

(\( W_2 \)) ("Trace property") The injection from \( W \) into \( U \) is continuous, \( \gamma \)
has a continuous right inverse, and the kernel \( W_0 \) of \( \gamma \) is dense in \( U \).

Remark 7.5 The type of setting we have described above is most familiar in
partial differential equations. The point as far as we are concerned is,
first, to avoid the dual \( V^* \) of \( V \) (typically nasty) and have intervene
instead the dual \( U^* \) of \( U \), and, second, to avoid the use of the transpose
\( L^* \) of \( L \) and to replace it by the transpose \( L_0^* \in L(V^*, W^*) \) of the restriction
\( L_0 \in L(W, V) \) of \( L \) to \( W_0 \). The motivation stems from the familiar
case in which \( L \) is a differential operator and \( W \) is a space of functions
or distributions; when \( W_0 \) is also the closure of the space of infinitely
differentiable functions with compact support, then \( W_0^* \) is a subspace of
distributions, and \( L_0^* \) is also a differential operator (in the sense of
distributions) and can be computed, whereas the transpose \( L^* \) cannot in
general be expressed in terms of differential operators.

When the trace property \( (W_2) \) holds, we can compare \( L^* \) and \( L_0^* \)
by means of an abstract Green formula in the following way (see Aubin [2]).

First we introduce the domain \( V_0^* \) of \( L_0^* \) defined by

\[
(7.6) \quad V_0^* = \{ p \in V^* : L_0^* p \in U^* \},
\]

where \( U^* \) is identified with a subspace of \( V_0^* \) (indeed, the transpose of
the injection from \( W_0 \) into \( U \) is an injection from \( U^* \) to \( V_0^* \)).

Equipped with the graph norm, \( V_0^* \) is a Banach space, and if \( (W_2) \) holds then there
exists a unique operator \( \tilde{B} \in L(V_0^*, T^*) \) such that

\[
\tilde{B}_0^* p, x > - <p, Lx> + <\tilde{B} p, y^*>
\]

for all \( x \in W \), for all \( p \) in \( V_0^* \).

(7.7)

(This will replace the integration by parts in the proof of Theorem 1.)

We associate to all \((u,v,t,s)\) in \( U \times V \times X \times Z \) the set

-21-
\( \Gamma(u,v,t,s) \) of points \( x \) in \( W \) satisfying

\[
\text{Lx} \in E(x-u) + v, \quad \gamma x \in S + t, \quad Cx \in T + x.
\]

Notice that our optimization problem involves minimizing a functional over \( \Gamma(0,0,0,0) \). We make the following controllability assumption:

1. There is a bounded subset \( \Gamma_0 \) of \( U \) such that, for every \( (u,v,t,s) \)
2. sufficiently small, the set \( \Gamma(u,v,t,s) \) is nonempty and contained in \( \Gamma_0 \). The functions \( f \) and \( g \) are Lipschitz on neighborhoods of \( \Gamma_0 \)

We define the value function \( a \) on \( U \times V \times T \times I \) via

\[
a(u,v,t,s) = \inf \{ f(x) + g(y) : x \in \Gamma(u,v,t,s) \},
\]

and the Hamiltonian function \( H \) on \( U \times V \times T \times I \) via

\[
H(x,p) = \sup \{ \langle p, v \rangle : v \in E(x) \}.
\]

Since the graph of \( E \) is convex, \( H \) is concave with respect to \( x \) and convex with respect to \( p \). We denote by \( \partial H(x,p) \) the subdifferential in the sense of convex analysis [17] of the convex function \( p + H(x,p) \), and by \( \partial^* H(x,p) \) the superdifferential of the concave function \( -H(x,p) \).

**Theorem 3** We posit \((\nabla_1 \Gamma_0) - (\nabla_3 \Gamma_0)\). If \( x \) in \( W \) minimizes (7.1) subject to the constraints (7.2)-(7.4), then there exist \( p \in V^* \), \( \gamma \in C \) satisfying:

\[
\begin{align*}
\text{Lx} & \in \partial H(x,p) \\
\sum \partial^* p & \in \partial^* H(x,p) - \gamma - \gamma \in C \quad \text{yields} \quad \text{satisfying} \quad (7.9) \\
s^p & \in \partial^* H(x,p)
\end{align*}
\]

Furthermore, the value function \( a \) is Lipschitz in a neighborhood of \((0,0,0,0)\), and we have:

\[
\sum \partial^* p \in C^* \gamma, \quad -p + \gamma \in C^* \gamma + H_\gamma(y).
\]

**Remark.** We recognize (7.9)-(7.10) to be Hamiltonian equations and (7.11) a transversality condition. The notation \( H_\gamma(y) \), for example, refers to the normal cone to the convex set \( S \) at the point \( y \gamma \) (i.e., the set of \( p \) in \( T^* \) such that \( \langle p, s - y \gamma \rangle \leq 0 \) for all \( s \) in \( S \)).

**B.** The shadow price information (7.12) exists with respect to four perturbations in the problem. In some situations it may be deemed more natural or more convenient to consider instead a performance function \( a \) depending on fewer variables. For example, in the problem of 11, we chose to mention only the interpretation of \( p(0) \). One can easily imagine situations (e.g., sensitivity analysis) in which the generality of (7.12) would come into play in other ways.

As far as reduced versions of (7.12) is concerned, one cannot simply drop components to obtain them, since it is not generally true that the relation \( (a,b) \in \partial f(x,y) \) implies \( a \in \partial f(x,y) \) (this is discussed in [9, art. 14]). However, one can modify the proof of Theorem 3 to attain the desired result. We shall not attempt to consider the most general situation, but rather one that seems natural in many settings. We consider the case in which \( S \) is a singleton set \( (s_0) \) and \( g \) is identically zero (i.e., boundary conditions reduce to \( y \gamma = s_0 \)). We let \( \delta(s) \) be the infimum in the problem in which the boundary condition is \( y \gamma = s \), and we maintain the hypotheses of the theorem.
Corollary 4. If \( x \) is optimal for the above problem, there exist \( p,q \) and \( t \)
as in Theorem 3 such that (7.9) (7.10) hold. Further, \( \delta \) is Lipschitz near \( x_0 \), and

\[
-\delta_p \in 2\delta(x_0).
\]

(7.13)

We consider now a different version of the above problem, in which there is an explicit dependence on a control parameter, and unilateral constraints on the state variable. We introduce a Banach space \( X \), closed convex subsets \( K \) of \( X \) and \( S \) of \( U \), and linear operators \( F \) and \( G \)
in \( L(U,V) \) and \( L(V,\nu) \) respectively. We now seek to minimise

\[
f(x,\nu) + g(\nu) \]

subject to

\[
x \in S,
Lx = Lx + Gz
\]
\[
q \in K
\]
\[
\gamma \in S
\]
\[
C_\nu \in Y,
\]

where \( f: U = X - K \), and \( S,T,L,S,C,T \) are unchanged from Theorem 3.
We denote by \( \Gamma(u,s,v,c,t) \) the set of \( (x,\nu) \) in \( U \times \nu \) satisfying

\[
x \in S + u
Lx = Lx + Gz + v
q \in K + c
\]
\[
\gamma \in S + \epsilon
C_\nu \in Y + s,
\]

and we define \( \delta \) on \( U = X \times V \times \nu = X \) via

\[
a(u,v,\nu,\zeta) = \inf\{ f(x,\nu) + g(\nu) : (x,\nu) \in \Gamma(u,v,\nu,\zeta) \}.
\]

There is a bounded subset \( \Gamma_0 \) of \( U = X \times V \) such that, for every \( (u,v,\nu,\zeta) \) sufficiently small, \( \Gamma(u,v,\nu,\zeta) \) is non-empty and contained in \( \Gamma_0 \). Furthermore, \( f \) is Lipschitz on a neighborhood of \( \Gamma_0 \).

Corollary 5. Under the stated assumptions, if \( (x,\nu) \) solves the above problem, there exist \( p \in V_0^* \), \( (\theta_1,\theta_2) \) in \( Sf(x,\nu) \), \( \phi \in 2g(\nu) \), and \( \nu \in K_\nu (C_\nu) \) satisfying

\[
-\delta_p + \delta_p - \delta \in Q \subset C \delta + H_0(x)
\]
\[
\delta \in \theta + H_0(\nu)
\]
\[
\delta \in \delta + H_0(\nu)
\]
\[
(\delta + C \delta + H_0(\nu)) \times \times (\delta + C \delta + H_0(\nu)) \times \delta \in \delta(0,0,0,0,0).
\]

Remark 7.12: As before, it is possible to replace (7.18) by alternate relations, such as (7.13) when the boundary conditions are simply \( \gamma_0 = \gamma_0 \). The reader may wish to verify that Theorem 1 is a consequence of this Corollary with the identifications (see 83 for notation):

\[
U = U = U^* = U^*, \quad L = L = L^* = L^*, \quad U = V = L^* = L^*
\]
\[
K = (c \in X : c(t) \in U \text{ a.s.})
\]
\[
f(x,\nu) = \int \delta x \text{ a.e.}
\]
\[
Lx = x, \quad \gamma_0 = x(0)
\]

-26-
\[ Z = W = Y, C = \text{identity} \]
\[ L^*_0 p = -\hat{p}, \quad p^*_0 p = p(0) \]
\[ V^*_0 = \{ p \in L^*_0; \quad p \in L^*_n \} \]

Of course, in applying the theorem to specific problems, there will be a need for appropriate characterizations of generalized gradients (such as Theorem 2 of §6) and normal cones (such as the lemma in §10). A further element will be the verification of the controllability hypothesis \((H_3)\). In the problem of §1, this was a result of taking \( \alpha \) large enough; in the applications of §9 and 10, it follows from postulating "strict feasibility points", a concept akin to the Slater condition of mathematical programming.

8. Proof of the abstract necessary conditions.

We shall merely sketch the proof, since the steps are identical to those in the proof of Theorem 1, albeit in a more general setting. We define, in the notation of §4,

\[ F(x) = f(x) + g(\gamma x) \]
\[ X = W \]
\[ Y = U \times V \times T \times Z \]
\[ Z = \text{Gr}(E) \times S \times Y \]
\[ Ax = (x, LX, Lx, Cx) \]

Then \((H_3)\) and Proposition 4.1 imply that the function \( \alpha \) is Lipschitz in a neighborhood of \( 0 \). As in §2, we denote the optimal solution \( \bar{x} \), and we note that for any \( \theta \) in \((0,1)\), for any \( x \) in \( W \) and \((u, v, t, x)\) in \( \text{Gr}(E) \times S \times Y \), we have

\[ \bar{x} + \theta(x - \bar{x}) \in \mathcal{F}(\theta[x-u, Lx-v, \gamma x-t, Cx-x]) \]

Consequently,

\[ \alpha(\theta[u, v, t, x]) \leq \mathcal{F}(\bar{x} + \theta(x - \bar{x})) \]

and taking generalized directional derivatives leads to

\[ 0 \leq \mathcal{F}(\bar{x}; x - \bar{x}) + \alpha^*(0, u - x, v - Lx, \gamma x - Cx) \]

Applying just as in §2 the left-sided minimax theorem, we deduce the existence of \((q, p, r, t) \in \mathcal{F}(0,0,0,0), \phi \in \mathcal{F}(\bar{x}), \phi \in \mathcal{F}(\gamma \bar{x})\) such that, for all \( x \) in \( W \) and \((u, v, t, x)\) in \( \text{Gr}(E) \times S \times Y \),

-29-
\[(8.1) \quad 0 < c_q(w, z) - c_q(z, z) \leq c_r(z, z) - c_r(x, z) \leq c_r(x, z) + c_r(z, z) + c_q(x, z) + c_q(z, z).
\]

If we set \( (u, v, t, x) = (\hat{u}, \hat{L}, \gamma, \hat{z}, \hat{C}) \) in \( (8.1) \), we obtain that for every \( x \in V \),
\[(8.2) \quad c_q(\hat{u}, x) - c_q(x, x) \leq c_r(x, x) - c_r(u, x) = 0.
\]

As \( x \) ranges over \( \hat{x} + W_0 \), this implies
\[q = L_0^* p + C^* \phi + c q \cdot
\]

Now by definition, \( (q, -p, r, t) \) belongs to \( H(0, 0, 0, 0) \) and hence to
\[U^* \times V^* \times T^* \times Z^*,
\]
while by the result of §5, \( \phi \) belongs to \( U^* \) (rather than merely to \( U^* \)); of course \( C \) belongs to \( U^* \). These facts and the preceding equation imply that \( L_0^* p \) belongs to \( U^* \); i.e., that \( p \) belongs to \( V^* \). Consequently we can use the Green formula \( (7.7) \) in \( (8.2) \) to deduce that for any \( x \) in \( V \),
\[c q(\hat{u}, x) - c q(\hat{u}, x) = 0.
\]

Since \( \gamma \) is surjective, this yields
\[q = \phi - \delta p \cdot
\]

and now \( (7.12) \) follows.

If in \( (8.1) \) we take \( x = \hat{x} \), it follows immediately that
\[(-q, p) \in H_{\frac{1}{2}}(G, L, \hat{z}), \quad -r \in H_{\frac{1}{2}}(\gamma, z), \quad \tau \in H_{\frac{1}{2}}(C, \hat{z}).
\]

It follows from convex analysis that \((-q, p)\) lies in \( H_{\frac{1}{2}}(G, L, \hat{z}) \) iff
\[q \in \frac{1}{2} H(\hat{z}, p), \quad \hat{L} \in \frac{1}{2} H(\hat{z}, p),
\]

and all the conclusions of the theorem ensue.

Q.E.D.

The proof of Corollary 4 uses the device introduced in §3: we observe that \( \hat{x} \) minimizes
\[f(x) = \delta(y_0)
\]
subject to
\[\hat{x} \in K(x), \quad C \in T, \quad y_0 \in \sigma_{y_0} + A
\]

where \( \delta > 0 \) is small, and we apply Theorem 3 to this new problem. The proof of Corollary 3 consists of applying Theorem 3 after the following relabelings:

\[U = U \times \Sigma, \quad W = W \times \Sigma, \quad V = V
\]
\[f(x, y) = f(x), \quad L(x, y) = L_0, \quad C(x, y) = C_0, \quad \gamma(x, y) = y_0
\]
\[E(x, y) = F(x) + G_0 \text{ if } y \in \Omega \text{ and } x \in K
\]
\[= \phi \text{ otherwise.}
\]

Let \( \Omega \) be a bounded open subset of \( \mathbb{R}^n \) whose boundary is a smooth differential manifold. We introduce functionals \( f \) and \( g \) defined by

\[
 f(x,\sigma) = \int_\Omega \phi(w,x(w),\sigma(w))dw \\
 g(\lambda) = \int_\Gamma \phi(w,\xi(w))d\mu(w)
\]

where \( d\mu(w) \) is a measure on \( \Gamma \).

We consider the solutions \( x \in H^2(\Omega) \) to the Dirichlet problem for the Laplacian:

\[
 \begin{align*}
 i) & \quad -\Delta x + x = \sigma \quad (\sigma \text{ ranges over } L^2(\Omega)) \\
 ii) & \quad \sigma|\Gamma = 0
\end{align*}
\]

(9.1)

We denote by \( \frac{\partial}{\partial n} \) the normal derivative. Let \( A \) be a compact set and let \( c \) be a function \( (w,\lambda) \in \Omega \times A \rightarrow c(w,\lambda) \in \mathbb{R} \) satisfying

\[
 i) \quad \lambda \in A, \ w \rightarrow c(w,\lambda) \text{ belongs to } L^2(\Omega) \\
 ii) \quad \text{for almost all } w \in \Omega, \ \lambda \rightarrow c(w,\lambda) \text{ is continuous.}
\]

(9.2)

We consider the following problem:

\[
 \text{Minimize } \int_\Omega \phi(w,x(w),\sigma(w))dw + \int_\Gamma \phi(w,\frac{\partial x}{\partial n}(w))d\mu(w)
\]

subject to the constraints (9.1) and

\[
 \begin{align*}
 i) & \quad \sigma \text{ belongs to the unit ball of } L^2(\Omega) \\
 ii) & \quad \psi \in A, \int_\Omega c(w,\lambda)u(w)dw \geq b(\lambda)
\end{align*}
\]

(9.3)

\[
 \psi \in A, \int_\Omega c(w,\lambda)u(w)dw \geq b(\lambda)
\]

(9.4)

where \( \lambda = 1(\lambda) \) is a continuous function. We posit the following controllability assumption.

\[
 \begin{array}{l}
 \text{There exist } \sigma_0 \in L^2(\Omega) \text{ such that } \|\sigma_0\|_{L^2(\Omega)} \leq 1 \text{ and such that the solution } x_0 \text{ of the Dirichlet problem } -\Delta x_0 + x_0 = \sigma_0 \\
 \quad \text{and } x_0|\Gamma = 0 \text{ satisfies:}
 \end{array}
\]

\[
 \psi \lambda \in A, \int_\Omega c(w,\lambda)x_0(w)dw \geq b(\lambda) .
\]

Finally, we assume that the functions \( \phi \) and \( \theta \) satisfy

\[
 i) \quad \text{the functions } w \rightarrow \phi(w,x,\sigma) \text{ and } \theta(w,\xi) \text{ are measurable for each } x, \sigma \text{ and } \xi \\
 ii) \quad \text{the functions } (w,\sigma) \rightarrow \phi(w,x,\sigma) \text{ and } \xi \rightarrow \theta(w,\xi) \text{ are locally Lipschitz for almost all } w .
\]

(9.5)

\[
 \text{i) there exists a constant } C > 0 \text{ such that}
 \begin{align*}
 \phi(w,x,\sigma) & \leq c(|x|)|\sigma|b(\lambda) \\
 \theta(w,\xi) & \leq c(|\xi|)b
\end{align*}
\]

(9.6)

iv) \( w \rightarrow \phi(w,0,0) \) and \( \theta(0,0) \) are (finitely) integrable.

Theorem 5. We posit assumptions (9.2), (9.5) and (9.6). Let \( (\xi,\delta) \) be an optimal solution. Then there exist \( p \in L^2(\Omega) \) whose Laplacian \( \Delta p \) belongs to \( L^2(\Omega) \) and functions \( \phi_1,\phi_2 \in L^2(\Omega) \) satisfying

\[
 \begin{align*}
 \phi_1(w),\phi_2(w) & \leq b(w,\xi(w),\sigma(w)) \text{ for almost all } w \in \Omega \\
 \text{and a non-negative Radon measure } \delta \text{ on } A \text{ satisfying}
 \end{align*}
\]

\[
 \begin{align*}
 \int_A \int_\Omega \left( \frac{1}{2}c(w,\lambda)x_0(w) \right)dw & \leq b(\lambda)
\end{align*}
\]

such that
\[ \int_{\Omega} \langle p(w) - \phi_2(w), \phi_2(w) \rangle = \| \phi - \phi_2 \|_{L^2(\Omega)}^2. \]

**Remark.** Among the possible shadow price interpretations we can add to the above is the following: if \( a(\xi) \) is the infimum in the above problem when the boundary condition is \( x|_{\Gamma} = \xi \) (instead of \( x|_{\Gamma} = 0 \)), then \( a \) is Lipschitz on a neighborhood of \( 0 \) in \( H^{3/2}(\Gamma) \) and:
\[
- \frac{2p}{\mu} \in \Delta_c(\Omega)
\]

**Proof.** We use Corollary 5 in the following case:

\[
\begin{align*}
U &= \mathcal{V} = \mathcal{L}^2(\Omega), \quad U = H^2(\Omega), \quad G = U, \\
\mathcal{T} &= H^{3/2}(\mathcal{T}) = H^{1/2}(\mathcal{T}), \quad F = 0, \quad G = \text{identity}, \quad K = \text{unit ball in } L, \\
\gamma &= \text{the operator defined by } \gamma x = (x|_\Gamma, \frac{\partial x}{\partial n}|_\Gamma).
\end{align*}
\]

The trace theorem [15] implies that \( \gamma \) is surjective and that

\[ \ker \gamma = H^2_0(\mathcal{T}) \quad \text{the closure of } \mathcal{D}(\gamma) \quad \text{in } H^2(\Omega). \]

We define \( L \) by

\[ L x = -\Delta x + x. \]

Hence \( L^\infty_x \) is defined by \( L^\infty_x p = -\Delta p + p \) (in the sense of distributions) and its domain is the space \( H^2(\Omega, \mathcal{A}) \) of functions \( p \in L^2(\Omega) \) such that \( \Delta p \in L^2(\Omega) \).

The Green formulas can be written

\[
\begin{align*}
(9.7) \quad \delta^* p(w) &= \left( \int_\Omega \frac{1}{\mu} p(w) x(w) dm(w) - \int_\Gamma p(w) \frac{\partial}{\partial n} x(w) dm(w) \right) \\
&\text{for smooth functions. By the abstract Green formula it still holds true when } \\
x \in H^2(\Omega), \quad p \in H^*(\Omega, \mathcal{A}) \quad \text{since the operator } \delta^* \text{ defined by } \delta^* p(w) = \left( \int_\Omega \frac{1}{\mu} p(w) - p(w) \right) \text{ can be extended in a unique way to a continuous linear operator } \delta^* \text{ from } \\
H^*(\Omega, \mathcal{A}) \to H^{-3/2}(\mathcal{T}) \times H^{-1/2}(\mathcal{T}). \]
\]

Finally, we choose \( S = (0) \times H^{1/2}(\mathcal{T}) \), and the (isoperimetric) constraints are defined by the Banach space \( Z = C(\Lambda) \), the subset \( Y = b + C_y(\Lambda) \) and the map \( C \) defined by \( C_y(\Lambda) = \int_\Omega c(x, y) x(w) \, dw \).

We now show that the controllability assumption \( \mathcal{H}_d \) of Corollary 5 is satisfied. Indeed, the operator \( x + (\Delta x, x|_\Gamma) \) is an isomorphism from \( H^2(\Omega) \) onto \( L^2(\Omega) \times H^{1/2}(\mathcal{T}) \). Let \( \mathcal{H}_d \) be the norm of its inverse,

\[ \mathcal{H}_d = \sup_{\lambda \in \Lambda} \| C(\cdot, \lambda) \|_{L^2(\Omega)} \]

We choose \( \gamma > 0 \) such that \( \gamma \geq \mathcal{H}_d/(1 + \mathcal{H}_d \sup_{\lambda \in \Lambda} \| \cdot \|_{L^2(\Omega)}) \). Now, if \( \| x \|_{L^2(\Omega)} \leq 1 \) and \( \| x \|_{L^2(\Omega)} \leq \gamma \), then the solutions

\[ (x, y) \in H^2(\Omega) \]

of

\[
\begin{align*}
(9.7) \quad \delta^* p(w) &= \left( \int_\Omega \frac{1}{\mu} p(w) x(w) dm(w) - \int_\Gamma p(w) \frac{\partial}{\partial n} x(w) dm(w) \right) \\
&\text{for smooth functions. By the abstract Green formula it still holds true when } \\
x \in H^2(\Omega), \quad p \in H^*(\Omega, \mathcal{A}) \quad \text{since the operator } \delta^* \text{ defined by } \delta^* p(w) = \left( \int_\Omega \frac{1}{\mu} p(w) - p(w) \right) \text{ can be extended in a unique way to a continuous linear operator } \delta^* \text{ from } \\
H^*(\Omega, \mathcal{A}) \to H^{-3/2}(\mathcal{T}) \times H^{-1/2}(\mathcal{T}). \]
\]

Finally, we choose \( S = (0) \times H^{1/2}(\mathcal{T}) \), and the (isoperimetric) constraints are defined by the Banach space \( Z = C(\Lambda) \), the subset \( Y = b + C_y(\Lambda) \) and the map \( C \) defined by \( C_y(\Lambda) = \int_\Omega c(x, y) x(w) \, dw \).

We now show that the controllability assumption \( \mathcal{H}_d \) of Corollary 5 is satisfied. Indeed, the operator \( x + (\Delta x, x|_\Gamma) \) is an isomorphism from \( H^2(\Omega) \) onto \( L^2(\Omega) \times H^{1/2}(\mathcal{T}) \). Let \( \mathcal{H}_d \) be the norm of its inverse,

\[ \mathcal{H}_d = \sup_{\lambda \in \Lambda} \| C(\cdot, \lambda) \|_{L^2(\Omega)} \]

We choose \( \gamma > 0 \) such that \( \gamma \geq \mathcal{H}_d/(1 + \mathcal{H}_d \sup_{\lambda \in \Lambda} \| \cdot \|_{L^2(\Omega)}) \). Now, if \( \| x \|_{L^2(\Omega)} \leq 1 \) and \( \| x \|_{L^2(\Omega)} \leq \gamma \), then the solutions

\[ (x, y) \in H^2(\Omega) \]

of

\[
\begin{align*}
(9.7) \quad \delta^* p(w) &= \left( \int_\Omega \frac{1}{\mu} p(w) x(w) dm(w) - \int_\Gamma p(w) \frac{\partial}{\partial n} x(w) dm(w) \right) \\
&\text{for smooth functions. By the abstract Green formula it still holds true when } \\
x \in H^2(\Omega), \quad p \in H^*(\Omega, \mathcal{A}) \quad \text{since the operator } \delta^* \text{ defined by } \delta^* p(w) = \left( \int_\Omega \frac{1}{\mu} p(w) - p(w) \right) \text{ can be extended in a unique way to a continuous linear operator } \delta^* \text{ from } \\
H^*(\Omega, \mathcal{A}) \to H^{-3/2}(\mathcal{T}) \times H^{-1/2}(\mathcal{T}). \]
\]

Finally, we choose \( S = (0) \times H^{1/2}(\mathcal{T}) \), and the (isoperimetric) constraints are defined by the Banach space \( Z = C(\Lambda) \), the subset \( Y = b + C_y(\Lambda) \) and the map \( C \) defined by \( C_y(\Lambda) = \int_\Omega c(x, y) x(w) \, dw \).

We now show that the controllability assumption \( \mathcal{H}_d \) of Corollary 5 is satisfied. Indeed, the operator \( x + (\Delta x, x|_\Gamma) \) is an isomorphism from \( H^2(\Omega) \) onto \( L^2(\Omega) \times H^{1/2}(\mathcal{T}) \). Let \( \mathcal{H}_d \) be the norm of its inverse,

\[ \mathcal{H}_d = \sup_{\lambda \in \Lambda} \| C(\cdot, \lambda) \|_{L^2(\Omega)} \]

We choose \( \gamma > 0 \) such that \( \gamma \geq \mathcal{H}_d/(1 + \mathcal{H}_d \sup_{\lambda \in \Lambda} \| \cdot \|_{L^2(\Omega)}) \). Now, if \( \| x \|_{L^2(\Omega)} \leq 1 \) and \( \| x \|_{L^2(\Omega)} \leq \gamma \), then the solutions

\[ (x, y) \in H^2(\Omega) \]

of

\[
\begin{align*}
(9.7) \quad \delta^* p(w) &= \left( \int_\Omega \frac{1}{\mu} p(w) x(w) dm(w) - \int_\Gamma p(w) \frac{\partial}{\partial n} x(w) dm(w) \right) \\
&\text{for smooth functions. By the abstract Green formula it still holds true when } \\
x \in H^2(\Omega), \quad p \in H^*(\Omega, \mathcal{A}) \quad \text{since the operator } \delta^* \text{ defined by } \delta^* p(w) = \left( \int_\Omega \frac{1}{\mu} p(w) - p(w) \right) \text{ can be extended in a unique way to a continuous linear operator } \delta^* \text{ from } \\
H^*(\Omega, \mathcal{A}) \to H^{-3/2}(\mathcal{T}) \times H^{-1/2}(\mathcal{T}). \]
\]
satisfy $||a|| < 1 + \gamma$ and $||x|| < N_2(1+\gamma)$. Further, such solutions exist; for take $\sigma = C_a + \sigma$. Then the solution $x$ to $-Ax = a + v$ satisfies $||x - x_0|| < 3N_1$, and hence

$$Cz = Cx_0 + b + C(x-a) + b$$
$$= N_3 - N_2||x - x_0|| + b > N_3 - N_2 - 3N_1 + b$$
$$> \gamma + b > b + z.$$  

This yields $(H_3)$ in our present setting.

Now, assumptions (9.6) on the functions $\phi$ and $\theta$ imply that the functionals $f$ and $g$ are Lipschitz on bounded subsets of $L^2(\Omega)$ and $L^2(\gamma)$ respectively (see 16).

The assumptions of Corollary 5 are therefore satisfied. It remains to interpret the conclusions. There exist $p \in \mathbb{R}^n(\Omega, A), \phi_1$ and $\phi_2 \in L^2(\Omega)$ such that, by theorem 2,

$$(\phi_1(v), \phi_2(w)) \in \phi(v, \tilde{u}(v), \tilde{o}(w))$$

for almost all $v \in \Omega$

$\zeta \in L^2(\gamma)$ such that

$$\zeta(v) \in \phi(v, \tilde{u}(w), \tilde{o}(w))$$

for almost all $v \in \gamma$ and a nonnegative Radon measure $\tilde{u} \in C(\Omega)$ (i.e. $\gamma$ of Corollary 5) such that

$$\int \left[ \int \phi(v, w) \tilde{u}(v) d\gamma - \tilde{b}(\gamma) \right] \tilde{u}(v) = 0$$

They satisfy the following equations:

1) $-dp + p + \phi_1 = \int \phi(v, z) \tilde{u}(v) d\gamma$

2) $-p|\gamma = \xi$

3) $\int \phi(v, w) - \phi_2(w), \tilde{o}(w) = ||p - \phi_2||^2_{L^2(\Omega)}$  

Q.E.D.

10. Example - a variational problem with unilateral constraints.

As a final example, we consider the following problem in the calculus of variations: to minimize

$$\int_0^1 g(x(t), x'(t)) dt$$

over the absolutely continuous arcs $x : [0,1] \rightarrow \mathbb{R}$ satisfying

(10.1) $x(t) \in \Omega$

(10.2) $\dot{x}(t) \in K$ a.e.

(10.3) $x(0) = x_0, x(1) = x_1$

where $\Omega$ and $K$ are given closed convex subsets of $\mathbb{R}^n, x_0$ and $x_1$ are given points in $\mathbb{R}^n$, and $g: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is a locally Lipschitz function.

We make the following assumptions: $K$ is compact, and there is an arc $x_0(t)$ joining $x_0$ to $x_1$ and an $\epsilon > 0$ such that

$$\dot{x}_0(t) + \epsilon \in K$$

and for almost every $t$, $x_0(t)$ belongs to the interior of $\Omega$.

For all $t$. We denote by $\sigma(t)$ the infimum in the above problem when the boundary conditions, instead of (10.3), are given by $x(0) = x_0, x(1) = t$.

Theorem 6 If $x$ solves the above problem, then $\sigma$ is Lipschitz near $x_1$

and there exist an absolutely continuous function $p$, an element $\lambda$ of $L^1$, and a nonnegative Radon measure $\mu$ on $[0,1]$ such that $\{t\}$ denotes integration over $[t,1)$:

(10.4) $\lambda(t) \in H_2(\tilde{U}(t))$ a.e.

(10.5) $(\tilde{u}(t), p(t)) \int_{-\lambda(t)}^0 \lambda(s) \sigma ds) \leq g(\tilde{u}(t), \dot{x}(t)) + (0) \in H_2(\tilde{U}(t))$ a.e.

(10.6) $p(t) \in \Omega(\tilde{u}(t))$.
Remark: Certain closely related, but not strictly comparable results appear in the literature. The convex problem is treated by R.T. Rockafellar [18], who obtains necessary conditions couched in terms of vector measures and the Hamiltonian. Related cases are treated by J. Warga [20], H. Markin [12], and F.H. Clarke [9, 16]. In these, no relation of the form (10.6) is obtained.

Proof. We may suppose that $x_0 = 0$. Note that the solutions $x$ of (10.2), (10.3) are bounded, so there is no loss of generality in assuming $S^T$ compact, and in supposing that $z$ satisfies the growth condition (1.3) of §1 with $g = 0$. We denote $\lambda^l_0$ the set of absolutely continuous $x: [0,1] \to \mathbb{R}^n$ such that $x(0) = 0$ and $\dot{x}$ belongs to $L^1$. If the norm of $x$ is taken as $\int_0^1 |x'|^2 dt$, then $\lambda^l_0$ may be identified with $L^1$ with the duality pairing

$$<x, y> = \int_0^1 \dot{x}y dt.$$ 

We shall apply Corollary 4 of §7 with the following identifications:

$$x = (x,y)$$
$$u = L^1 \times L^1$$
$$W = \lambda^l_0 \times L^1$$
$$f(x,y) = \int_0^1 g(x,y) dt$$
$$z = \lambda^l_0$$
$$Y = \{z \in Z : \sup_{t \in [0,1]} x(t) < 0\}$$
$$G(x,y) = \int_0^1 y(x) dt$$

$$y(x,y) = x(1)$$
$$z = T$$
$$S = \{x_1\}$$
$$L(x,y) = \dot{x} - y$$
$$v = L^1$$

$$E(x,y) = \{0\} \text{ if } y \notin K_1, \quad \emptyset \text{ otherwise},$$

where $K_1$ is the set of $y$ in $L^1$ satisfying $y(t) \in K$ a.e. It is routine to check that $(\dot{x},y) = (\ddot{x},\ddot{y})$ then solves the abstract problem of §7, and to verify the hypotheses of the theorem. The controllability hypothesis follows mainly (much as in §9) from the observation that if $(u,v,t,s)$ is sufficiently small in $L^1 \times L^1 \times L^n \times Z$, then the element $(x,y)$ of $W$ defined by

$$y(t) = \dot{x}_0(t) + u(t) + t - \int_0^1 (uv) dt$$
$$\dot{x}(t) = y + v$$

satisfies

$$y \in K_1 + u$$
$$x(1) = x_1 + t$$
$$C(x,y) \in C + s.$$ 

We calculate

$$y^0 = L^n$$
$$\lambda^p_0 = \{p \in L^n : \dot{p} \in L^n\} \text{ (in the sense of distributions)}$$

$$y^0 = \{p \in L^n : \dot{p} \in L^n\}.$$
\[ C^v = \langle 0, v \rangle \]
\[ \tilde{\gamma}^p = -p(1) \]
\[ H(x, y, p) = 0 \text{ if } y \in H_x, \quad \infty \text{ otherwise.} \]

A straightforward interpretation of \((7.9)\) (with the help of Theorem 2, 14) then yields the existence of an arc \(p\) and elements \(v, \phi\) of \(H_x(G(x, y))\) and \(H_{\tilde{\gamma}}\) such that
\[
(\tilde{\gamma}, p) = \tilde{\gamma}(\tilde{\gamma}, x) + (0, \tau)v
\]
\[ p(1) \in \gamma(x, y) \]

Thus, \(v\) is a function in \(L^m\) satisfying
\[
\frac{1}{\sigma} \int v(t) \cdot (h(t) - \tilde{h}(t)) dt \leq 0
\]
whenever \(h\) in \(L^1\) is such that \(h(1) \in E \ a.e.\). It follows readily from the measurable selection theorem [19, Theorem 4.1] that for almost every \(t\), for every \(h\) in \(E\),
\[
\sigma(t) \cdot (h - \tilde{h}(t)) \leq 0,
\]
\[ i.e. \quad \sigma(t) \in H_x(h(t)) \]

It remains now to characterize \(\tau\). We have by definition that for any arc \(x\) satisfying \(x(0) = 0, x(t) \in \Omega\),
\[
\frac{1}{\tau} \int v(t) \cdot (\tilde{h}(t) - \tilde{h}(t)) dt \leq 0,
\]

where, as explained earlier, \(\tau\) is identified with an element of \(L^m\).

The lemma below then yields an expression for \(\tau\) that concludes the proof.

Q.E.D.

We denote \(A^m\) the set of absolutely continuous functions \(x: [0,1] \to H^m\) with derivative \(\dot{x}\) in \(L^m\); \(p\) is an scalar \(\lambda_1\).

**Lemma 10.1** Let \(\Omega\) be a compact convex set in \(H^m\) containing \(0\) in its interior, and let \(\tau\) belong to \(L^2([0,1]; H^m)\). If \(\hat{x}\) in \(A^m\) solves the problem of maximizing
\[
\frac{1}{2} \int \tau(s) \cdot \dot{x}(s)ds
\]
over the elements \(x\) of \(A^m\) satisfying \(x(0) = 0, x(t) \in \Omega\) for all \(t\), then there is a function \(\lambda\) in \(L^m\) and a nonnegative Radon measure \(\mu\) on \([0,1]\) such that:

\[
(10.7) \quad \lambda(t) \in H_x(\hat{x}(t))
\]
\[
(10.8) \quad \tau(t) = \int \lambda(s)\mu(ds) a.e.
\]

**Proof.** Let \(h\) denote the support function of \(\Omega\); i.e.
\[
h(p) = \max \{p \cdot v : v \in \Omega\}.
\]

If \(S\) denotes the unit sphere of \(H^m\), then \(x\) belongs to \(\Omega\) iff
\[
g(x) < 0, \quad \text{where}
\]
\[
g(x) = \max \{p \cdot x - h(p) : p \in S\}.
\]

The function \(g\) is locally Lipschitz, and it follows from [6, Theorem 2.1] that when \(g(x) = 0\), \(\tilde{g}(x) \in H_x(\Omega)\). Further, because \(0 \in \text{int} \Omega\) we have \(h(p) \geq 0\) when \(p \in S\), and this implies \(0 \notin g(x)\) when \(g(x) = 0\).

The statement of the lemma implies then that \(\hat{x}\) minimizes
A class of optimal control problems is considered in which the cost function is locally Lipschitz (not necessarily convex or differentiable) and the dynamics are linear and/or convex. By using generalized gradients and duality methods of functional analysis, necessary conditions are obtained in which the dual variables admit interpretation as shadow prices (or rates of change of the value function). Applications are presented in three settings: infinite-horizon optimal control, optimal control of partial differential equations and a variational problem with unilateral state constraints. A theorem is proved which characterizes the generalized gradients of integral functionals on $L^p$. 
\[
\int_{-1}^{1} x \, dt
\]

subject to \( x(0) = 0 \), \( g(x(t)) \leq 0 \). We may apply [9, Theorem 4] to deduce the existence of an arc \( y \), a function \( \lambda \) in \( L^2 \) and a Radon measure \( m \) supported on the set \( \{ t : g(x(t)) = 0 \} \) such that

\[
(y, y) + \int_{0}^{1} \langle 0, \lambda \rangle \, dm = (0, -\gamma)
\]

Because \( x(1) \) is free, the transversality condition

\[
-\gamma(1) = \int_{0}^{1} \lambda \, dm
\]

also pertains. Thus \( p \) is identically equal to this last quantity, and the result follows.

Q.E.D.

REFERENCES

5. N. Bourbaki, "Integration", Fascicule XXV, Chapitre VI, Hermann, Paris
11. H. Maltin, Necessary conditions for optimal control problems with
12. __________, Optimization without differentiability, in "Proceedings of
the Conference on Optimal Control Theory" (Canberra, August 1977)
13. T.C. Koopmans, Concepts of optimality and their uses, Nobel Lectures
1975; reproduced in Amer. Econ. Rev. 67 (1977) 261-274.
14. G. Lebourg, Valeur moyenne pour gradient généralisé, Comptes Rendus
15. J-L. Lions, E. Magenes, "Problèmes aux limites non homogènes et appli-
16. R. Pallu de la Barrière, On the cost of constraints in dynamical optimis-
ation, in "Mathematical Theory of Control" (Ed. A.V. Balakrishnan,
17. R.T. Rockafellar, "Convex Analysis", Princeton Univ. Press, Princeton,
18. __________, State constraints in convex control problems of
Bolza, SIAM. J. Control and Optimization 10 (1972) 691-711.
20. J. Warga, Controllability and necessary conditions in unilateral
problems without differentiability assumptions, SIAM J. Control