

MAXIMUM PRINCIPLES WITHOUT DIFFERENTIABILITY¹

BY FRANK H. CLARKE

Communicated by Alistair Lachlan, August 1, 1974

Introduction. For reasons both theoretical and practical, there has arisen an interest in variational problems not possessing the customary differentiability hypotheses. The author developed in his dissertation a theory of necessary conditions for certain general problems in optimal control and the calculus of variations. In this article we describe the kind of results that may be obtained when this theory is applied to control problems lacking differentiability or smoothness. In §1 we describe the problem and some terminology, and state a theorem similar to the Pontryagin Maximum Principle. §2 gives a “deparametrized” form of the necessary conditions. In neither case are the results stated in their greatest possible generality. Proofs, details and more general results will appear elsewhere.

1. Let there be given functions $l: R^n \rightarrow R$, $g: [0, 1] \times R^n \times R^m \rightarrow R$ and $f: [0, 1] \times R^n \times R^m \rightarrow R^n$, as well as a *multifunction* $U: [0, 1] \rightarrow R^m$ (i.e. a mapping from $[0, 1]$ to the subsets of R^m).

An admissible *control-response pair* is a pair (u, x) such that u is a (Lebesgue) measurable function from $[0, 1]$ to R^m with $u(t) \in U(t)$ a.e., x is an absolutely continuous function from $[0, 1]$ to R^n , and

$$\dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e.}$$

The illustrative optimal control problem we consider is

(1) minimize $l(x(1)) + \int_0^1 g(t, x(t), u(t))dt$ over the control-response pairs (u, x) satisfying $x(0) \in S$, where S is a given closed subset of R^n .

DEFINITION 1. Let $\phi: R^n \rightarrow R$ be Lipschitz continuous. The *generalized gradient* of ϕ at s , denoted $\partial\phi(s)$, is the set

AMS (MOS) subject classifications (1970). Primary 49B10, 49B35.

Key words and phrases. Optimal control, maximum principle, nondifferentiable functions.

¹Part of this paper is taken from the author's Ph. D. thesis, written under the supervision of R. T. Rockafellar, University of Washington, June 1973.

$$\text{co} \left\{ \lim_{i \rightarrow \infty} \nabla \phi(s_i) \right\},$$

where we consider all sequences s_i converging to s such that $\nabla \phi(s_i)$ exists and converges (co denotes "convex hull"). We remark only that if ϕ is C^1 , $\partial \phi(s) = \{\nabla \phi(s)\}$, and that if ϕ is convex, $\partial \phi(s)$ is the set of subgradients of ϕ at s (see [2]).

DEFINITION 2. We say $p \in R^n$ is *normal* to S at s if for some positive ϵ ,

$$\epsilon p \in \partial d_S(s),$$

where $d_S(s) = d(s, S)$, the (Euclidean) distance function.

It can be shown that when S is a C^1 manifold, the normals in the above sense coincide with the usual normal space at s , and that when S is convex, we get precisely the normals in the sense of convex analysis.

THEOREM 1. Let the control-response pair (v, z) solve problem (1) (locally), where the following hold:

- (a) f and g are measurable in t and continuous in (x, u) .
- (b) l is Lipschitz near $z(1)$.
- (c) U is a closed-valued and measurable multifunction (see [3]).
- (d) For almost every t , $f(t, \cdot, v(t))$ is C^1 near $z(t)$.
- (e) For every positive integer j there is an integrable function k_j such that

$$|f(t, s, u) - f(t, v, u)| \leq k_j(t) |s - v|$$

whenever u lies in $U(t)$, $|u - v(t)| < j$, and s and v are near $z(t)$, and a similar condition holds for g .

Then there exists an absolutely continuous function $p: [0, 1] \rightarrow R^n$ such that

$$(2) \quad \dot{p}(t) + p(t) J_s f(t, z(t), v(t)) \in \partial_s g(t, z(t), v(t)) \text{ a.e.}$$

(3) $p(t) \cdot f(t, z(t), u) - g(t, z(t), u) \leq p(t) \cdot f(t, z(t), v(t)) - g(t, z(t), v(t))$ for all u in $U(t)$, a.e.

$$(4) \quad p(0) \text{ is normal to } S \text{ at } z(0) \text{ (Definition 2) and } -p(1) \in \partial l(z(1)).$$

REMARKS. In contrast to the assumptions usually present, neither v nor U is required to be bounded; note also that the control set U may vary with time. As mentioned earlier, the generalized gradients in (2) and (4), and the normal in (4) reduce to familiar concepts when smoothness or convexity is present.

Note that in hypothesis (d), f need be C^1 in the state variable only at the optimal control ν . In the next section, we shall dispense with smoothness completely. It is easy to see that in the classical case, where in particular f and g have derivatives in s continuous in (t, s, u) , and where ν is essentially bounded, hypotheses (a)–(e) are satisfied. J_s refers to the Jacobian matrix.

2. Let there be given a compact-valued multifunction $E: [0, 1] \times R^n \rightarrow R^n$. The absolutely continuous function $x: [0, 1] \rightarrow R^n$ is a *trajectory* for E if $\dot{x}(t) \in E(t, x(t))$ a.e.

The control problem we consider now is to minimize $l(x(1))$ over the trajectories x for E satisfying $x(0) \in S$, where l and S are as in Theorem 1.

DEFINITION 3. E is *Lipschitz in s near z* , where z is a given function from $[0, 1]$ to R^n , if there exists an integrable function k such that when s_1 and s_2 are near $z(t)$, the Hausdorff distance between $E(t, s_1)$ and $E(t, s_2)$ is no greater than $k(t)|s_1 - s_2|$.

THEOREM 2. Let the trajectory z solve the above problem (locally), where E is measurable in t and Lipschitz in s near z . Then there exists an absolutely continuous function $p: [0, 1] \rightarrow R^n$ such that:

$$(5) \quad (-\dot{p}(t), \dot{z}(t)) \in \partial H(t, z(t), p(t)) \text{ a.e.}$$

$$(6) \quad p(t) \cdot v \leq p(t) \cdot \dot{z}(t) \text{ for all } v \text{ in } E(t, z(t)), \text{ a.e.}$$

(7) $p(0)$ is normal to S at $z(0)$ and $-p(1) \in \partial l(z(1))$ where H , the generalized or maximized Hamiltonian, is the function defined by

$$H(t, x, p) = \max\{p \cdot v: v \in E(t, x)\},$$

and ∂H refers to the generalized gradients of H with respect to the (x, p) variables.

REMARKS. An important special case is when E has the form

$$E(t, s) = \{f(t, s, u): u \in U(t)\}.$$

We then require only that f be Lipschitz in s in order for E to be Lipschitz in s . The correspondence between (5) and (2) may be clarified by showing, as can be done, that (5) is equivalent to:

$$(-\dot{p}(t), \dot{z}(t)) \in \text{co}\{(p(t) \cdot v, f(t, z(t), u)): v \in \partial_s f(t, z(t), u), u \in M(t)\},$$

where

$$M(t) = \{u \in U(t): p(t)f(t, z(t), u) = H(t, z(t), p(t))\},$$

and $\partial_s f$ is defined as in Definition 1, replacing gradients by Jacobian matrices. Conditions similar to these, for a relaxed control problem, have recently been announced by Warga [4]. Hamiltonian conditions for convex problems were discussed by Rockafellar in [3].

REFERENCES

1. F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. (to appear).
2. R. T. Rockafellar, *Convex analysis*, Princeton Math. Ser., no. 28, Princeton Univ. Press, Princeton, N. J., 1970. MR 43 #445.
3. ———, *Generalized Hamiltonian equations for convex problems of Lagrange*, Pacific J. Math. 33 (1970), 411–427. MR 43 #2593.
4. J. Warga, *Necessary conditions in optimal control without differentiability assumptions*, J. Differential Equations. (to appear).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA,
VANCOUVER 8, BRITISH COLUMBIA, CANADA

Current address: Mathématiques de la Décision, Université de Paris IX (Dauphine),
Paris 16, France