A GENERALIZATION OF THE HAMILTON - JACOBI EQUATION

by

David Ian Havelock

B.Sc., Carleton University, Ottawa, 1974

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

in the Department

of

MATHEMATICS

We accept this thesis as conforming to the

required standard

THE UNIVERSITY OF BRITISH COLUMBIA

September, 1977

Department of MATHEMATICS

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date 20 SEPT 1977

ABSTRACT

After a brief review of the relevant classical theory and a presentation of the concept of generalized gradients, it is demonstrated that, in analogy with the classical case, a locally lipschitz value function satisfies a generalized version of the Hamilton-Jacobi equation. A sufficiency condition for optimality is developed and some examples illustrating various aspects of the generalized theory are presented.

ACKNOWLEDGEMENTS

I wish to acknowledge the patient and helpful assistance of my thesis supervisor, Frank Clarke, during the development of this thesis. I would also like to thank my Aunt Pauline for her unlimited hospitality and encouragement during my stay in Vancouver.

David Havelock

Ottawa
September 29, 1977.

TABLE OF CONTENTS

INTRODUCTION

The basic problem in the Calculus of Variations is that of finding a piecewise smooth curve y(x) which minimizes the definite integral

$$\int_{a_o}^{a} F(x,y(x),y'(x))dx$$

and joins two fixed points (a_o,b_o) and (a,b). The integrand F is classically at least once continuously differentiable. The set of curves over which the minimum is sought is called the Set of Admissible Curves.

Caratheodory [2;P205,VOL II] took the approach of considering problems which were equivalent, in some sense, to a nice type of problem. A 'Nice' problem was said to be one for which the integrand F^* satisfies

$$\min_q F^*(x,y,q) = 0$$

for all x,y. A problem with an integrand F is said to be Equivalent to a 'nice' problem with an integrand F^* if there exists a smooth function R(x,y) with

$$F^*(x,y,q) = F(x,y,q) - R'(x,y;1,q)$$

where R'(x,y;1,q) is the directional derivative of R in the direction (1,q). The definite integrals of F and F^* along any admissible curve will differ by the value $R(a,b) - R(a_o,b_o)$ and hence we are assured that a curve will solve the nice problem (ie; will be optimal) if and only if it solves the equivalent one.

It is found that a problem is equivalent to a 'nice' problem exactly where there exists a smooth solution to a certain partial differential equation called the Hamilton-Jacobi Equation (H.-J.Eq.)

$$H(x,y,R_2(x,y)) + R_1(x,y) = 0$$

where R_1 and R_2 are respectively the first partial derivatives of R in the first and second variables. The Hamiltonian Function $H(a,b,p)$ is defined as $H(a,b,p) = pq_p - F(a,b,q_p)$ where q_p and p are related implicitly by the relation

$$\frac{\partial}{\partial q}F(a,b,q)\Big|_{q=q_p} = p \quad .$$

The Value Function, or Hamilton's Characteristic Function S is given as a function of the end point (a,b):

$$S(a,b) = \text{MIN} \int_{a_o}^{a} F(x,y(x),y'(x))dx$$

with the minimum being over admissible curves from (a_o,b_o) to (a,b). Where the value function is defined and smooth, it satisfies the Hamilton - Jacobi Equation. This indicates that many problems are equivalent to 'nice' problems.

There is a greater variety of necessary conditions for optimality than of sufficiency conditions, but for 'nice' problems we have a particularly simple sufficiency condition at hand: " if $F(x,y(x),y'(x)) = 0$ almost everywhere along the curve y, then y is optimal. " For problems equivalent to a 'nice' problem, the corresponding condition would require that

$$F(x,y(x),y'(x)) - R'(x,y(x);1,y'(x)) = 0$$

almost everywhere along y, with R being some smooth solution to the H.-J.Eq. If the value function S is smooth along the optimal curve y then it turns out that y necessarily satisfies the above relation with R = S.

A major difficulty in applying this theory is the requirement of differentiability. In many interesting cases there is no guarantee that the value function will be smooth (and it often is not). The classical soap bubble problem, discussed in example III of chapter IV, is such a case in which the value function actually fails to be smooth. By employing the concept of Generalized Gradients as defined for locally lipschitz functions, the theory may be studied using functions which fail to be differentiable. Although the generalized gradient has been defined for a larger class of functions (see [3]) we will only consider it for locally lipschitz functions here.

For a locally lipschitz function f the generalized gradient at a point x, denoted $\partial f(x)$, is a compact, convex, non-empty set.

If f is a convex function ∂f coincides with the subdifferential of f. Accompanying the generalized gradient is the generalized directional derivative, denoted $f^{o}(x,y;a,b)$. Like the standard directional derivative $f'(x,y;a,b)$, the generalized directional derivative is a single valued mapping.

Classical theory presumes the existence of several minima and maxima which our generalized theory replaces with infima and suprema. We use a <u>Generalized Hamiltonian</u>

$$H(x,y,p) = \operatorname*{SUP}_{q} \left(pq - F(x,y,q) \right)$$

to accomodate a greater variety of integrands. This definition of H, unlike that of the classical Hamiltonian function, is not predicated on the existence of the Legendre transform. It is found that a locally lipschitz solution R to a <u>Generalized Hamilton - Jacobi Equation (G.H.-J.Eq.)</u>

$$\operatorname*{MAX}_{(u,v)\varepsilon\partial R(x,y)} \left(H(x,y,v) + u \right) = 0$$

(where the 'MAX' exists automatically) establishes an equivalence with a generalized type of 'nice' problem: those satisfying

$$\mathrm{INF}\; F(x,y,q) = 0 .$$

If the value function, as defined in classical theory, is locally lipschitz, it is found to satisfy the G.H.-J.Eq. For the generalized theory, the value function is defined for all (a,b) as

$$S(a,b) = \mathrm{INF} \int_{a_o}^{a} F(x,y(x),y'(x))dx .$$

If the infimum of the value function is not attained (no optimal curve exists) then the value function may fail to satisfy the G.H.-J.Eq., but will satisfy the following <u>Generalized Hamilton-Jacobi Inequality</u> (<u>G.H.-J.Ineq.</u>) :

$$H(a,b,v) + u \geq 0$$

for all (u,v) in $\partial S(a,b)$.

A sufficient condition for optimality is provided by the solutions of the G.H.-J.Ineq. as follows: If R is a solution to the G.H.-J.Ineq. on a region and y_o is a curve lying in this region and satisfying

$$F(x,y_o(x),y_o'(x)) - \frac{d}{dx}R(x,y_o(x)) = 0$$

almost everywhere, then y_o is an optimal curve (provides the infimum in the value function).

CHAPTER I Preliminaries

By a curve we will mean a lipschitz function mapping an interval $[a_o,a]$ of the real line $\mathbb{R}$ into the n-dimensional real space $\mathbb{R}^n$. We fix a point (a_o,b_o) in $\mathbb{R} \times \mathbb{R}$ and henceforth consider only curves y which satisfy $y(a_o) = b_o$. Given a set U in $\mathbb{R}^{n+1}$ we say a curve y lies in U almost everywhere if $(x,y(x))$ is in U for almost all x in $[a_o,a]$.

Let $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous function, let U be a set in $\mathbb{R}^{n+1}$, and let (a,b) be a point in $\mathbb{R}^{n+1}$.

Definition

The following problem is referred to as a basic problem: find

$$\underset{y\in A(a,b)}{\mathrm{INF}} \int_{a_o}^{a} F(x,y(x),y'(x))dx$$

where $A(a,b) = \{$curves y lying in U almost everywhere with $y(a) = b\}$.

It is implicit that $y(a_o) = b_o$ and that y is lipschitz (hence y' exists almost everywhere on $[a_o,a]$). Notice that it is assumed that $a > a_o$. For brevity, where no confusion results, $F(x,y(x),y'(x))$ will be written as $F(x,y,y')$.

The basic problem is characterized by three things: the function F; the set U; and the point (a,b). The set U will be referred to as the domain, the point (a,b) as the terminal point, and the function F as simply the integrand. The set $A(a,b)$ will be referred to as the set of admissible curves to (a,b). Notice that the point (a,b) must lie in the closure of the domain U, $\overline{U}$.

For simplicity of presentation we will henceforth consider only one-dimensional problems, that is; $\mathbb{R}^n = \mathbb{R}^1$, wherein curves will map $\mathbb{R}$ into $\mathbb{R}$.

Consider the family of basic problems determined by a fixed integrand F, a fixed domain U, and a set of terminal points Ω. We will require that as a subset of $\mathbb{R}^2$, Ω be an open set and we will refer to Ω as the <u>set of termination</u> for the family of problems.

<u>Definition</u>

<u>For the family of problems described above, we define the value function</u> S <u>on</u> Ω <u>as follows</u>: <u>for</u> (a,b) <u>in</u> Ω

$$S(a,b) \equiv \inf_{y \in A(a,b)} \int_{a_o}^{a} F(x,y,y')dx \;.$$

If the infimum is attained by some y in A(a,b), we say that y <u>is an optimal curve to</u> (a,b) <u>or</u> y <u>is optimal in</u> A(a,b). Furthermore, if y is an admissible curve to (a,b) satisfying

$$\int_{a_o}^{a} F(x,y,y')dx \leq S(a,b) + \delta, \qquad (\text{for } \delta>0)$$

then we say that y <u>is a δ-near optimal curve to</u> (a,b). A sequence of curves $\{y_n\}_{n=1}^{\infty}$ with each y_n a δ_n-near optimal curve to (a,b) with $\delta_n \to 0$ as $n \to \infty$ will be called a <u>minimizing sequence</u>. Notice that as long as S(a,b) is finite, there must be a minimizing sequence in A(a,b).

Equivalent problems

Two integrands F and F^* are said to be equivalent on a set (U say) in $\mathbb{R}^2$ if there exists a function R which is continuously differentiable on U and satisfies

$$\nabla R(a,b) \cdot (1,q) = F(a,b,q) - F^*(a,b,q)$$

for all (a,b) in U, and all q in $\mathbb{R}$. The symbol ∇ represents the usual vector gradient while the symbol $\cdot$ represents the usual inner (scalar) product on $\mathbb{R}^2$.

Consider the two basic problems

$$\underset{y \varepsilon A(a,b)}{\mathrm{INF}} \int_{a_o}^{a} F(x,y,y')dx$$

and

$$\underset{y \varepsilon A(a,b)}{\mathrm{INF}} \int_{a_o}^{a} F^*(x,y,y')dx$$

with F and F^* equivalent on the domain U. Since

$$\int_{a_o}^{a} F(x,y,y')dx - \int_{a_o}^{a} F^*(x,y,y')dx$$

$$= \int_{a_o}^{a} \nabla R(x,y(x)) \cdot (1,y'(x))dx$$

$$= R(a,b) - R(a_o,b_o) ,$$

a curve y in A(a,b) will be optimal or δ-near optimal for one problem if and only if it is optimal or δ-near optimal respectively for the other. Accordingly, the two problems are said to be equivalent problems.

A basic problem or its integrand F is called <u>nice</u> (<u>on the domain</u> U) if

$$\operatorname*{MIN}_{q\in\mathbb{R}} F(x,y,q) = 0$$

for all (x,y) in U. Notice that when a basic problem is nice, it is sufficient (although not necessary) for y to be optimal in A(a,b) to have $F(x,y,y') = 0$ for almost all x in $[a_o,a]$.

The concepts of equivalent problems and nice problems were introduced by Caratheodory ([2;§227 vol II]). In their classical setting it is assumed that both the integrand F and the function R are smooth. In chapter III we will alter the definitions slightly, freeing us from these smoothness assumptions.

<u>The Classical Hamilton-Jacobi Equation</u>

Let $F(x,y,q)$ be a twice continuously differentiable integrand with $\frac{\partial^2}{\partial q^2} F(x,y,q) \geq 0$ for all q in $\mathbb{R}$. The mapping from $\{(x,y)\} \times \mathbb{R}$ into $\{(x,y)\} \times \mathbb{R}$ given by

$$(x,y,q) \rightarrow (x,y,\frac{\partial}{\partial z}F(x,y,z)|_{z=q})$$

is one to one (although the range may be a proper subset of $\{(x,y)\} \times \mathbb{R}$) The classical Hamiltonian function H is defined as

$$H(x,y,p) = p \cdot q_o - F(x,y,q_o)$$

where q_o is determined implicitly by the relationship

$$\frac{\partial}{\partial q} F(x,y,q)|_{q=q_o} = p .$$

The transformation of the system (x,y,q,F) into the system (x,y,p,H) is known as a Legendre Transform [6,§7.1]. Notice that $H(x,y,p)$ may not be defined for all p in $\mathbb{R}$.

Assume that H is defined at the point (x,y,p) and consider the following function in q: $\Psi(q) = pq - F(x,y,q)$. Since this function is concave and the first derivative vanishes at q_o ($p - \frac{\partial}{\partial q}F(x,y,q_o) = 0$), we must have a maximum occuring at q_o. Consequently we see that where $H(x,y,p)$ is defined

$$H(x,y,p) = \max_{q\in\mathbb{R}} (pq - F(x,y,q)) .$$

Furthermore, notice that the Hamiltonian will be defined <u>exactly</u> where the maximum in the above expression exists.

Suppose that R establishes an equivalence between F and a nice integrand F^* on U, then

$$F^*(x,y,q) = F(x,y,q) - \nabla R(x,y)\cdot(1,q)$$

for (x,y) in U and q in $\mathbb{R}$. It follows that

$$\begin{aligned}
&\max_q (-F^*(x,y,q)) \\
&\quad = \max_q \left(\frac{\partial R}{\partial x} + \frac{\partial R}{\partial y}\cdot q - F(x,y,q) \right) \\
&\quad = \frac{\partial R}{\partial x} + H\left(x, y, \frac{\partial R}{\partial y}\right) \\
&\quad = 0,
\end{aligned}$$

and so R satisfies the partial differential equation

$$H\left(x , y , \frac{\partial}{\partial y}R(x,y) \right) + \frac{\partial}{\partial x}R(x,y) = 0$$

on U. This equation is called the Hamilton-Jacobi Equation (H.-J.Eq.). Notice that conversely, if R satisfies the H.-J.Eq. on U then it provides an equivalence between F and the nice integrand

$$F^{*}(x,y,q) = F(x,y,q) - \nabla R(x,y) \cdot (1,q)$$

on U.

Caratheodory's lemma [7,theorem 5.1] asserts that if the infimum of the value function is attained at every point of an open set V and if it is continuously differentiable on V, then the value function establishes an equivalence between F and a nice integrand on V. Consequently, the value function will satisfy the H.-J.Eq. on the open set V.

The Non-Classical Situation

For integrands which may not satisfy the classical assumption of being twice continuously differentiable, we define the more versatile generalized Hamiltonian

$$H(x,y,p) \equiv \sup_{q\in\mathbb{R}} (pq - F(x,y,q)) .$$

As a supremum of affine functions in p the generalized Hamiltonian will be convex in p, and may assume values of plus infinity. Unlike the classical Hamiltonian it may happen that the supremum is not attained by any q in $\mathbb{R}$, or that if the supremum is attained it is attained for more than one q. Whenever the assumptions that F is twice continuously differentiable and strictly convex in q are valid, the classical Hamiltonian, where it is defined, will equal the generalized Hamiltonian. Without confusion, then "Hamiltonian" will henceforth refer to the generalized Hamiltonian.

Consider the H.-J.Eq. in which the generalized Hamiltonian is employed. If R is a continuously differentiable solution to this equation, then setting $F^*(x,y,q) = F(x,y,q) - \nabla R(x,y)\cdot(1,q)$ gives

$$\begin{aligned}
&\inf_q F^*(x,y,q) \\
&\quad = -\sup_q [\nabla R(x,y)\cdot(1,q) - F(x,y,q)] \\
&\quad = -\frac{\partial}{\partial x}R(x,y) - H\left(x, y, \frac{\partial}{\partial y}R(x,y)\right) \\
&\quad = 0 .
\end{aligned}$$

Unlike the situation in the classical case, we are not assured here that the infimum is actually attained. In order to extend the classical theory then, it seems natural to modify the definition of nice problems or integrands as follows:

DEFINITION

A basic problem or its integrand F is said to be nice on U if

$$\operatorname*{INF}_{q} F(x,y,q) = 0$$

for all (x,y) in U.

The classical theory demands not only that the integrand be twice differentiable, but also that

(i) the value function be continuously differentiable and

(ii) optimal curves exist (ie: the infimum defining the value function is always attained)

(see [7,§3]).

These two constraints on the value function are distinct from smoothness and convexity in q imposed on the integrand. As we see in example III chapter IV, the assumptions on the integrand are satisfied, however those on the value function are not. If the integrand need not be convex it is very easy to construct simple examples in which optimal curves do not exist. Consider for example,

$$F(x,y,q) = (1 + q^2)^{-1} .$$

Using "sawtooth shaped" curves we can get

$$\int_{a_o}^{a} F(x,y,y')dx$$

as close to zero as we wish, yet since F is strictly positive, no curve will provide the infimum value of zero.

If the integrand need not be smooth, it is easy to construct examples in which the value function is not smooth. The simple integrand $F(x,y,q) = |q|$ for example, yields the non-smooth value function $S(x,y) = |y|$. Studies have been made of integrands which may fail to be smooth but remain convex, or lipschitz (see [5] and the references provided there).

Although the value function frequently fails to be smooth, the author knows of no case in which the value function is finite but not locally lipschitz as long as the integrand remains continuous. In fact, if F is lipschitz and it happens that we can find a real K such that for all (a,b) in a neighbourhood η, the optimal curve y_{ab} to (a,b) satisfies $|y'| < K$, then the value function S will be assured of being lipschitz in the neighbourhood. The bound K can be guaranteed by growth conditions on the integrand.

LEMMA I   CONDITIONS FOR A LIPSCHITZ VALUE FUNCTION

Let Ω be a set of termination of the form

$$\Omega = \{ (x,y) \in \mathbb{R}^2 \mid x > \varepsilon > 0, \text{ and } |a_o - x| + |b_o - y| < M < \infty \}$$

for fixed ε, M in $\mathbb{R}$, and let the domain U be $\mathbb{R}^2$. The value function

$$S(a,b) = \underset{y \in A(a,b)}{\mathrm{INF}} \int_{a_o}^{a} F(x,y,y')dx$$

will be lipschitz on Ω if there exist positive constants k_1 and k_2 such that for all $(x,y,q) \in \mathbb{R}^3$ and $(\alpha,\beta,\gamma) \in \mathbb{R}^3$, F satisfies

(i) $F(x,y,q) \geq 1$ and

(ii) $F(x+\alpha, y+\beta, q+\gamma) \leq F(x,y,q) \cdot \exp\left[\dfrac{k_1 + ||(\alpha,\beta,\gamma)||}{(k_2 + ||(x,y,q)||)}\right].$

Condition (i) requires only that F be bounded below, since the addition of a constant to the integrand merely adds a linear term to the value function. Condition (ii) is a growth condition which, for differentiable functions, assumes the form

$$f'(x) \leq f(x) \cdot \frac{k_1}{|x| + k_2}$$

Functions of the form $f(x) = |x|^\alpha$ for $\alpha > 1$ satisfy this condition but functions with cusps pointing downward, such as $f(x) = |x|^{\frac{1}{2}}$ do not.

Proof of Lemma 1

We begin by showing that S is bounded on Ω. Assume, without loss of generality, that $(a_o,b_o) = (0,0)$. For (a,b) in Ω we consider the straight line $y \varepsilon A(a,b)$ given by $y(x) = bx/a$.

$$S(a,b) \leq \int_0^a F(x,bx/a,b/a)dx$$

$$\leq \int_0^a F(0,0,0)\cdot \exp[\ k_1 ||(x,bx/a,b/a)||/k_2\ ]dx$$

$$\leq \int_0^a F(0,0,0)\cdot \exp[\ k_1 ||(M,M,M/\varepsilon)||/k_2\ ]dx$$

$$\leq M\cdot F(0,0,0)\cdot \exp[\ k_1 ||(M,M,M/\varepsilon)||/k_2\ ]$$

$$< \infty \quad .$$

Thus S has an upper bound and a lower bound as desired (the lower bound being zero, since F is positive).

To show that S is lipschitz on Ω, it is sufficient to show that it is lipschitz separately in the variables x and y. Let y_1 and y_2 be δ-near optimal curves to $(a_1,b_1) \varepsilon \Omega$ and $(a_1,b_2) \varepsilon \Omega$, respectively. Define the curve $z \varepsilon A(a_1,b_1)$ by

$$z(x) = y_2(x) + (b_1-b_2)x/a_1 \ .$$

Now,

$$S(a_1,b_1) \le \int_0^{a_1} F(x,z(x),z'(x))dx$$

$$\le \int_0^{a_1} F(x,y_2,y_2')\cdot \exp\left[\frac{k_1||(0,\frac{b_1-b_2}{a_1}x,\frac{b_1-b_2}{a_1})||}{||(x,y_2,y'_2)||+k_2}\right] dx$$

$$\le \int_0^{a_1} F(x,y_2,y_2')\cdot \exp\left[\frac{k_1}{k_2}||(0,\frac{b_1-b_2}{a_1}x,\frac{b_1-b_2}{a_1})||\right] dx$$

$$\le \int_0^{a_1} F(x,y_2,y_2')\cdot \exp\left[\frac{k_1}{k_2}(|b_1-b_2| + \frac{|b_1-b_2|}{a_1})\right] dx .$$

From elementary calculus, if x is bounded then $\exists K \in \mathbb{R}$ such that $e^{|x|} \le 1 + |Kx|$. Using this fact we see that

$$S(a_1,b_1) \le \int_0^{a_1} F(x,y_2,y_2')dx\cdot [\ 1 + K|b_1-b_2|\ ]$$

$$\le [\ \delta + S(a_1,b_2)\ ]\cdot[\ 1 + K|b_1-b_2|\ ] .$$

Rearranging and letting $\delta \to 0$,

$$S(a_1,b\ ) - S(a_1,b_2) \le \hat{K}\cdot|b_1-b_2| \quad ,$$

where $\hat{K}$ can be chosen independently of a_1, b_1, b_2.
Interchanging b_1 and b_2 in the above argument we get

$$S(a_1,b_2) - S(a_1,b_1) \le \hat{K}\cdot |b_1-b_2| \quad ,$$

and so S is lipschitz in the second variable.

To see that S is lipschitz in the first variable, we redefine the curves y_1 and y_2 to be δ-near optimal curves to $(a_1,b_1) \in \Omega$ and

$(a_2,b_1) \in \Omega$, respectively. Let $u(x) = a_2 x/a_1$ and define $z \in A(a_1,b_1)$ as $z(x) = y_2(u(x))$. We have,

$$S(a_1,b_1) \leq \int_0^{a_1} F(x,z,z')dx$$

$$= \frac{a_1}{a_2} \cdot \int_0^{a_2} F(a_1 u/a_2 \ , \ y_2(u) \ , \ a_2 y_2'(u)/a_1)du$$

$$\leq \frac{a_1}{a_2} \cdot \int_0^{a_2} F(u,y_2,y_2') \cdot \exp\left[\frac{k_1 || (\frac{a_1-a_2}{a_2}u \ , \ 0 \ , \ \frac{a_1-a_2}{a_1}y' \ ) ||}{k_2 + ||(u,y_2,y_2')||}\right] du$$

$$\leq \frac{a_1}{a_2} \cdot \int_0^{a_2} F(u,y_2,y_2') \cdot \exp\left[k_1 (\ |a_1-a_2|/a_2 + |a_1-a_2|/a_1 \ )\right] du.$$

As before, the exponential term can be bounded by $K|a_1-a_2|$ and we get

$$S(a_1,b_1) \leq \frac{a_1}{a_2} \cdot \int_0^{a_1} F(x,y_2,y_2')dx \cdot [\ 1 + K|a_1-a_2| \ ]$$

$$\leq \quad [\ \delta + S(a_2,b_1) \ ] \cdot [\ 1 + \frac{(a_1 K + 1)}{a_2} \cdot |a_1-a_2| \ ] \ .$$

Rearranging and letting $\delta \to 0$;

$$S(a_1,b_1) - S(a_2,b_1) \leq \hat{K} \cdot |a_1-a_2| \quad ,$$

with $\hat{K}$ independent of a_1, a_2, and b_1. As before, by interchanging a_1 and a_2 we obtain

$$S(a_2,b_1) - S(a_1,b_1) \leq \hat{K} \cdot |a_1-a_2| \quad ,$$

and so S is also lipschitz in the first variable. This completes the proof.

Chapter II GENERALIZED GRADIENTS

Let $f:\mathbb{R}^n \to \mathbb{R}^n$ be a locally lipschitz function. By Rademacher's theorem, f is differentiable almost everywhere although it need not be continuously differentiable anywhere. If V is a subset of $\mathbb{R}^n$, coV will denote the convex hull of V in $\mathbb{R}^n$.

Definition (see [3])

The generalized gradient of a lipschitz function f at a point $x \in \mathbb{R}^n$, $\partial f(x)$, is defined as follows:

$$\partial f(x) \equiv co\{\lambda \in \mathbb{R}^n \mid \lambda = \lim_{n\to\infty} \nabla f(x_n) \ , \ x = \lim_{n\to\infty} x_n \ \}.$$

(This definition has been extended to a larger class of functions - see [3]).

The following examples illustrate the generalized gradient of two simple functions.

Example Let $f(x) = |x|$, then

$$\partial f(x) = \begin{cases} \{1\} & \text{if } x > 0, \\ \{-1\} & \text{if } x < 0, \\ [-1,1] & \text{if } x = 0. \end{cases}$$

Example Let $f(x) = x^2 \sin(1/x)$. For $x \neq 0$ $f'(x) = 2x\sin(1/x) - \cos(1/x)$. For $x = 0$, using basic principles, we find $f'(0) = 0$. The generalized gradient, however, is given by

$$\partial f(x) = \begin{cases} \{f'(x)\} & \text{if } x \neq 0, \\ [-1,1] & \text{if } x = 0. \end{cases}$$

The generalized gradient $\partial f(x)$ is a closed, compact, convex, non-empty subset of $\mathbb{R}^n$ (see [3]). If L is a local lipschitz constant for f about x, then it is easy to see that for all $\lambda \in \partial f(x)$, $||\lambda|| \leq L$.

If f is continuously differentiable at x then clearly $\partial f(x) = \{\nabla f(x)\}$. It may happen, as in the example $f(x)=x^2\sin(1/x)$, that $\nabla f(x)$ exists but $\partial f(x) \neq \{\nabla f(x)\}$. In any case, we are always assured of the following property:

Property 1

If $\nabla f(x)$ *exists, then* $\nabla f(x) \in \partial f(x)$.

The lipschitz property of f provides a very useful form of continuity for the generalized gradient:

Property 2

The generalized gradient is upper semi-continuous (U.S.C.), that is; if $\{x_n\}_{n=1}^{\infty}$ *converges to* x, *and* $\lambda_n \in \partial f(x_n)$ *for* $1 \leq n < \infty$, *then any limit point* λ *to* $\{\lambda_n\}_{n=1}^{\infty}$ *satisfies* $\lambda \in \partial f(x)$.

Notice that $\{\lambda_n\}_{n=1}^{\infty}$ is assured of having at least one limit point since for large n, λ_n lies within the (compact) sphere of radius L where L is a local lipschitz constant for f about x.

Definition (See [3])

The generalized directional derivative of a lipschitz function f at a point $x \in \mathbb{R}^n$ in a (non-zero) direction $v \in \mathbb{R}^n$, $f^o(x;v)$, is given by

$$f^o(x;v) = \limsup_{\substack{h \to 0 \\ \delta \downarrow 0}} [(f(x+h+\delta v) - f(x+h))/\delta].$$

Like the usual one sided directional derivative, denoted $f'(x;v)$ when it exists, the generalized directional derivative is a mapping from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$. Notice that from the definition of f^o we have the following:

Property 3

$$f'(x;v) \leq f^o(x;v) \quad \text{for all } x,v \text{ in } \mathbb{R}^n.$$

Example Let $f(x) = -|x|$, and consider $f^o(0;v)$. Notice that $|\delta v| \geq |h| - |h+\delta v|$ and so $|v| \geq \limsup_{\substack{h \to 0 \\ \delta \downarrow 0}} (-|h+\delta v|+|h|)/\delta$.
If we let $h=-\delta v$ then the reverse inequality is established and we conclude that $f^o(0;v) = |v|$. This is in contrast to the classical situation, in which $f'(0;v) = -|v|$.

An equivalent, and often more convenient, definition of the generalized directional derivative is given by

$$f^o(x;v) = \text{MAX}\{\lambda \cdot v \mid \lambda \in \partial f(x)\}$$

(see [3]). The maximum is attained at an extreme point in $\partial f(x)$, and since all extreme points are of the form $\lim_{n\to\infty} \nabla f(x_n)$ for some sequence $\{x_n\}$ converging to x, the following property holds:

Property 4

$f^o(x;v) = v \cdot \lim_{n\to\infty} \nabla f(x_n)$ for some sequence $\{x_n\}_{n=1}^{\infty}$ converging to x.

Suppose now that $f'(x;v)$ exists. Then

$$\begin{aligned}
&\mathrm{MAX}\ \{\lambda\cdot v \mid \lambda\varepsilon\partial f(x)\} \\
&= f^o(x;v) \\
&\geq f'(x;v) \\
&= -f'(x;-v) \\
&\geq -f^o(x;-v) \\
&= -\mathrm{MAX}\ \{ -\lambda\cdot v \mid \lambda\ \varepsilon\ \partial f(x) \} \\
&= \mathrm{MIN}\ \{ \lambda\cdot v \mid \lambda\ \varepsilon\ \partial f(x) \}
\end{aligned}$$

and we have the following property:

Property 5

$$\begin{aligned}
f'(x;v)\ \varepsilon\ &\{\lambda\cdot v \mid \lambda\ \varepsilon\ \partial f(x)\} \\
&= [-f^o(x;-v)\ ,\ f^o(x;v)]\ .
\end{aligned}$$

Let E be any set of zero measure. Define ∂f_E on $\mathbb{R}^n$ as

$$\partial f_E(x) = \mathrm{co}\{\lambda\varepsilon\mathbb{R}^n \mid \lambda = \lim_{n\to\infty} \nabla f(x_n), x_n \to x \text{ as } n\to\infty, \text{ and } x_n \notin E\},$$

and define the function f^o_E on $\mathbb{R}^n \times \mathbb{R}^n$ as

$$f_E^o(x;v) = \limsup_{\substack{h \to 0 \\ \delta \downarrow 0}} \{ [f(x+h+\delta v)-f(x+h)]/\delta \mid x+h, x+h+\delta v \notin E\}.$$

As is demonstrated in [3] , the following equalities hold:

Property 6

$$\partial f_E(x) = \partial f(x) ,$$

$$f_E^o(x;v) = f^o(x;v) ,$$

and so also

$$f_E^o(x;v) = \text{MAX } \{\lambda \cdot v \mid \lambda \in \partial f_E(x)\} .$$

The generalized gradient and directional derivative extend the concept of the subdifferential of a convex function. If f is a convex function then

(i) the generalized gradient is identical to the subdifferential

(ii) $f^o(x;v) = f'(x;v)$

(iii) $f(x) = \{\alpha\}$ if and only if $\nabla f(x)$ exists and $\alpha = \nabla f(x)$.

Several results employing the subdifferential have been extended by the use of the generalized gradient (for an example and a brief discussion see [4,introduction]).

We will have occasion to consider the generalized gradient of a function at an end point of the interval on which it is defined. In this situation the generalized gradient will be determined by limits of sequences restricted to the interval of definition of the function. If $f:[0,1] \to \mathbb{R}$ is lipschitz,

$$\partial f(1) = \mathrm{co}\{\alpha \mid \alpha = \lim_{n\to\infty} \nabla f(x_n), x_n \to 1, x_n \in [0,1]\} .$$

This corresponds to extending f symmetrically about 1, and so the properties given in this chapter will hold, with the appropriate restrictions. Notice that the "one-sided" generalized gradient described here will be a subset of the "standard" generalized gradient, if the latter exists.

CHAPTER III THE GENERALIZED HAMILTON-JACOBI EQUATION

Recall Caratheodory's definition of equivalent problems; we will generalize the concept as follows:

DEFINITION

One problem, or its integrand F, is said to be equivalent to another (with its integrand F^*) on a set $U \in \mathbb{R}^2$ if there exists a locally lipschitz function R with its generalized gradient defined on U and satisfying

$$F^*(a,b,q) = F(a,b,q) - R^o(a,b;1,q)$$

for each $(a,b) \in U$ and $q \in \mathbb{R}$.

If in the above definition F^* is a nice integrand we have, for (a,b) in U and (u,v) restricted to $\partial R(a,b)$, that

$$\begin{aligned}
&\operatorname*{INF}_{q} [F(a,b,q) - R^o(a,b;1,q)] \\
&\quad = -\operatorname*{SUP}_{q} \Big[\operatorname*{MAX}_{(u,v)} (u,v)\cdot(1,q) - F(a,b,q) \Big] \\
&\quad = -\operatorname*{MAX}_{(u,v)} \Big\{ \operatorname*{SUP}_{q} [vq - F(a,b,q)] + u \Big\} \\
&\quad = -\operatorname*{MAX}_{(u,v)} [H(a,b,v) + u] \\
&\quad = 0 .
\end{aligned}$$

DEFINITION

The relation

$$\max_{(u,v)\in\partial R(a,b)} [\, H(a,b,v) + u \,] = 0 \qquad (a,b) \in U$$

will be referred to as the generalized Hamilton-Jacobi equation (G.H.-J.Eq.) for R on the region $U \in \mathbb{R}^2$.

Because the generalized gradient ∂R is a compact set and the Hamiltonian is convex in the third variable (hence continuous on open sets, where it is finite), the use of a maximum, as opposed to a supremum, is justified in the above definition. Since H may assume values of ∞ however, the G.H.-J.Eq. implicitly requires that H(a,b,v) be finite for all (u,v) in $\partial R(a,b)$.

Notice that in analogy with the classical case, a locally lipschitz function R will satisfy the G.H.-J.Eq. on U if and only if it establishes an equivalence between the problem and a nice problem on U.

If F^*, given by

$$F^*(a,b,q) = F(a,b,q) - R^o(a,b;1,q)$$

is known only to be positive, then we find that R satisfies

$$\max_{(u,v)\in\partial R(a,b)} [\, H(a,b,v) + u \,] \leq 0 ,$$

which will be referred to as the generalized Hamilton-Jacobi inequality for R (G.H.-J.Ineq.) . A locally lipschitz function will establish an equivalence with a positive integrand on U if and only if it satisfies the G.H.-J.Ineq. on U.

The following is the central result.

<u>THEOREM I</u> <u>NECESSARY CONDITIONS</u>

a) <u>Let</u>

$$S(x,y) = \sup_{\hat{y}\varepsilon A(x,y)} \int_{a_o}^{x} F(\hat{x},\hat{y}(\hat{x}),\hat{y}'(\hat{x}))\,d\hat{x}$$

<u>where</u> A(x,y) = { <u>curves to</u> (x,y) <u>which lie in the domain</u> U <u>almost everywhere</u> }. <u>If</u>

(i) S <u>is defined on a neighbourhood</u> $\eta \subset U$ <u>of a point</u> (a,b),

(ii) S <u>is locally lipschitz in</u> η , <u>and</u>

(iii) F(x,y,q) <u>is continuous in all three variables</u>,

<u>then</u>

$$F(a,b,q) - S^{o}(a,b;1,q) \geq 0$$

<u>for all</u> $(u,v) \varepsilon \partial S(a,b)$ <u>and</u> $q \varepsilon \mathbb{R}$.

b) <u>If, in addition,</u> y_o <u>is an optimal curve to</u> (a,b) <u>then</u>

$$F(a,b,q_o) - S^{o}(a,b;1,q_o) = 0$$

<u>for some</u> $q_o \varepsilon \partial y_o(a)$.

Part (a) of the theorem states that interior to where S is defined and locally lipschitz the problem will be equivalent to one with a positive integrand. As we have seen, this is the same as saying that where S is defined and locally lipschitz, it will satisfy the G.H.-J.Ineq. The value function is thus seen to be closely linked to the Hamiltonian function by the G.H.-J.Ineq. Similarly, part (b) of the theorem assures us that under stronger hypotheses S will be a solution to the G.H.-J.Eq. This relationship between the value function and the Hamiltonian function is futher considered, in its classical setting in [7,chapter 9]. Notice that y_o need not be defined outside the interval, $[a_o,a]$, hence $\partial y_o(a)$ requires the interpretation discussed at the end of Chapter II.

Proof of Theorem I

In addition to the hypotheses of the theorem, assume that $\nabla S(a,b)$ exists. For $q \in \mathbb{R}$ let $y_q(x) = b + (x - a)q$ be the line through (a,b) with slope q. Select $\varepsilon > 0$ small enough that $a_o < a - \varepsilon$ and the line segment $\{(x,y_q(x)) \mid x\varepsilon[a-\varepsilon,a+\varepsilon]\}$ lies within η. Let x_1 and x_2 lie in $[a-\varepsilon,a+\varepsilon]$ with $x_1 < x_2$ and let $y_\delta(x)$ be a δ-near optimal curve from (a_o,b_o) to $(x_1,y_q(x_1))$. Define $\hat{y}_\delta(x) \in A(x_2,y_q(x_2))$ as

$$\hat{y}_\delta(x) = \begin{cases} y_\delta(x) & \text{if } a_o \leq x \leq x_1 , \\ y_q(x) & \text{if } x_1 \leq x \leq x_2 . \end{cases}$$

As a function of x, $S(x,y_q(x))$ is locally lipschitz on $[x_1,x_2]$ and so $\frac{d}{dx}S(x,y_q(x))$ exists almost everywhere on $[x_1,x_2]$. Now,

$$\int_{x_1}^{x_2} [\ F(x,y_q,y_q') - \frac{d}{dx}S(x,y_q(x))\ ]dx$$

$$= \int_{a_o}^{x_2} F(x,\hat{y}_\delta,\hat{y}_\delta')dx - \int_{a_o}^{x_1} F(x,y_\delta,y_\delta')dx - \int_{x_1}^{x_2} \frac{d}{dx}S(x,y_q(x))dx$$

$$\geq S(x_2,y_q(x_2)) - [S(x_1,y_q(x_1)) + \delta]$$

$$- S(x_2,y_q(x_2)) - S(x_1,y_q(x_1))$$

$$= -\delta .$$

Since we may independently choose δ arbitrarily small and since the interval $[x_1,x_2] \subset [a-\varepsilon,a+\varepsilon]$ is arbitrary, we see that for almost all x in $[a-\varepsilon,a+\varepsilon]$, $\frac{d}{dx}S(x,y_q(x))$ exists and

$$F(x,y_q(x),q) - \frac{d}{dx}S(x,y_q(x)) \geq 0 .$$

Let $E = \{x\varepsilon[a-\varepsilon,a+\varepsilon] \mid \frac{d}{dx}S(x,y_q(x))$ does not exist, or $F(x,y_q(x),q) - \frac{d}{dx}S(x,y_q(x)) < 0 \}$, then E has zero measure. select a sequence $\{x_n\}_{n=1}^{\infty}$ in $[a-\varepsilon,a+\varepsilon]$ with $x_n \to a$ as $n \to \infty$ and $\{x_n\}_{n=1}^{\infty} \cap E = \phi$, then recalling that F is continuous,

$$\underset{n \to \infty}{\text{LIMSUP}} \left[F(x_n, y_q(x_n), q) - \frac{d}{dx} S(x_n, y_q(x_n)) \right]$$

$$= F(a,b,q) - \underset{n \to \infty}{\text{LIMINF}} \frac{d}{dx} S(x_n, y_q(x_n))$$

$$\geq 0 \ .$$

Denote the generalized gradient of $S(x,y_q(x))$ as a function in x alone, as $\partial_x S(x,y_q(x))$, (not to be confused with $\partial S(x,y(x))$).

Now,

$$\partial_x S_E(a,y_q(a)) \equiv \text{co} \left\{ \alpha \;\middle|\; \alpha = \lim_{n\to\infty} \frac{d}{dx} S(x_n, y_q(x_n)) \ , \text{ with } x_n \notin E \text{ and } x_n \to a \text{ as } n \to \infty \right\} ,$$

and as we have seen, for each α:

$$F(a,b,q) - \alpha \geq 0 \ .$$

Taking the convex hull of the α's preserves this property, that is; $\forall\ \alpha \in \partial_x S_E(a,y_q(a))$, $F(a,b,q) - \alpha \geq 0$. By property 6 of chapter II, $\forall\ \alpha \in \partial_x S(a,y_q(a))$, $F(a,b,q) - \alpha \geq 0$. In particular, $\frac{d}{dx}S(a,y_q(a))$ lies in $\partial_x S(a,y_q(a))$ (property 1 of chapter II) and also, $\frac{d}{dx}S(a,y_q(a)) = \nabla S(a,b) \cdot (1,q)$, so:

$$F(a,b,q) - \nabla S(a,b) \cdot (1,q) \geq 0 \ . \qquad \text{(Equation 1)}$$

We now drop the assuption that $\nabla S(a,b)$ exists. Let $(u_o,v_o) \in \partial S(a,b)$ be such that $S^o(a,b;1,q) = (u_o,v_o) \cdot (1,q)$

and such that there is a sequence $\{a_n, b_n\}_{n=1}^{\infty}$ converging to (a,b) with $(u_o, v_o) = \underset{n\to\infty}{\text{LIM}}\ \nabla S(a,b)$. Existence of (u_o, v_o) is assured by property 4 of chapter II. For n large enough that (a_n, b_n) lies interior to η we have

$$F(a_n, b_n, q) - \nabla S(a_n, b_n) \cdot (1,q) \geq 0 ,$$

and so, taking the limit as $n \to \infty$,

$$F(a,b,q) - S^o(a,b;1,q) \geq 0 ,$$

as stated in (a) of the theorem.

<u>Part II</u> We proceed to show that the minimum value is zero and that it is always attained if an optimal curve exists. Let y_o be an optimal curve to (a,b) and select $x_1 > a$ so that $(x_1, y_o(x_1))$ is interior to the neighbourhood η, and such that;

(i) $y_o'(x_1)$ exists, and

(ii) $\frac{d}{dx}\int_{a_o}^{x} F(x, y_o, y_o')dx = F(x_1, y_o(x_1), y_o'(x_1))$ at $x = x_1$.

Note that (i) and (ii) hold almost everywhere along y_o. Now

$$S^o(x_1, y_o(x_1); 1, y_o'(x_1))$$

$$= \underset{\substack{(h_1,h_2)\to 0 \\ \lambda \downarrow 0}}{\text{LIMSUP}} \Big[S(x_1 + h_1 + \lambda\ ,\ y_o(x_1) + h_2 + \lambda y_o'(x_1))$$

$$- S(x_1 + h_1\ ,\ y_o(x_1) + h_2) \Big] / \lambda$$

$$\geq \limsup_{\lambda \downarrow 0} \left[S(x_1+\lambda, y_o(x_1) + \lambda y_o'(x_1)) - S(x_1, y_o(x_1)) \right] / \lambda$$

$$= \limsup_{\lambda \downarrow 0} \left[S(x_1+\lambda, y_o(x_1) + \lambda y_o'(x_1)) - S(x_1+\lambda, y_o(x_1+\lambda)) \right.$$

$$\left. + \quad S(x_1+\lambda, y_o(x_1+\lambda)) - S(x_1, y_o(x_1)) \right] / \lambda$$

Let M be the lipschitz constant for S in a neighbourhood of (a,b). Since $y_o(x_1+\lambda) = y_o(x_1) + \lambda y_o'(x_1) + o(\lambda)$, for small λ we have:

$$S^o(x_1, y_o(x_1); 1, y_o'(x_1))$$

$$\geq \limsup_{\lambda \downarrow 0} \left[o(\lambda)\cdot M + S(x_1+\lambda, y_o(x_1+\lambda)) - S(x_1, y_o(x_1)) \right] / \lambda$$

$$= \frac{d}{dx} \int_{a_o}^{x} F(x, y_o, y_o'))dx \qquad (\text{at } x = x_1) .$$

Since η is a neighbourhood of $(x_1, y_o(x_1))$, part (a) of the theorem holds here, that is:

$$F(x_1, y_o(x_1), y_o'(x_1)) \quad - \quad S^o(x_1, y_o(x_1); 1, y_o'(x_1))$$

is non-negative. Since

$$F(x_1, y_o(x_1), y_o'(x_1)) \quad = \quad \frac{d}{dx} \int_{a_o}^{x} F(x, y_o, y_o')dx \qquad (\text{at } x = x_1)$$

we combine this with the previous inequality to get:

$$F(x_1,y_o(x_1),y_o'(x_1)) \;-\; S^o(x_1,y_o(x_1);1,y_o'(x_1)) \;=\; 0 \,. \qquad \text{(Equation 2)}$$

Now select a sequence $\{x_n\}_{n=1}^{\infty}$ with the following properties: for each $n \geq 1$;

(i) $y_o'(x_n)$ exists ,

(ii) $\frac{d}{dx}\int_{a_o}^{x} F(\hat{x},y_o,y_o')d\hat{x} = F(x,y_o(x),y_o'(x))$ at $x = x_n$,

(iii) $(x_n,y_o(x_n))$ lies interior to η ,

(iv) $x_n \to a$ as $n \to \infty$, and finally,

(v) $\{y_o'(x_n)\}_{n=1}^{\infty}$ converges (to q_o say) .

Conditions (i) and (ii) are satisfied almost everywhere along y_o , as mentioned earlier, so conditions (i) through (iv) are easily met. Since y_o is lipschitz, y_o' is bounded and any sequence satisfying (i) through (iv) will have a sub-sequence satisfying (v) as well. For such a sequence $\{x_n\}_{n=1}^{\infty}$ let $\{(u_n,v_n)\}_{n=1}^{\infty}$ be a sequence satisfying $(u_n,v_n) \in \partial S(x_n,y_o(x_n))$ and

$$S^o(x_n,y_o(x_n);1,y_o'(x_n)) \;=\; (u_n,v_n)\cdot(1,y_o'(x_n))$$

for each $n \geq 1$.

For each x_n "Equation 2" above will hold and we may rewrite it in the form:

$$F(x_n,y_0(x_n),y_o'(x_n)) - (u_n,v_n)\cdot(1,y_o'(x_n)) = 0 .$$

Since S is lipschitz $\{(u_n,v_n)\}_{n=1}^{\infty}$ has a convergent sub-sequence and its limit (u_o,v_o) lies in $\partial S(a,b)$ (property 2 of chapter II). Similarily q_o , the limit of $\{y_o'(x_n)\}_{n=1}^{\infty}$ lies in $\partial y_o(a)$. Taking the limit as $n \to \infty$ in the above equation we get

$$F(a,b,q_o) - (u_o,v_o)\cdot(1,q_o) = 0 .$$

Now

$$\begin{aligned} F(a,b,q_o) &- \max_{(u,v)\varepsilon\partial S(a,b)} [(u,v)\cdot(1,q_o)] \\ &= F(a,b,q_o) - S^o(a,b;1,q_o) \\ &\leq F(a,b,q_o) - (u_o,v_o)\cdot(1,q_o) \\ &= 0 ; \end{aligned}$$

and in consideration of (a) in the theorem,

$$F(a,b,q_0) - S^o(a,b;1,q_o) = 0$$

completing the proof.

Notice that "equation 2" in the proof of theorem I provides a generalization to Caratheodory's fundamental lemma ([7,§5.1]) : under the hypotheses of the theorem, if y is optimal and y' is continuous at a point x with $(x,y(x))$ within η , then conditions (i) and (ii) at the beginning of part II of the proof hold, and so according to "equation 2" ;

$$\begin{aligned} &\underset{q}{\mathrm{MIN}}\ [F(x,y(x),q) - S^{o}(x,y(x);1,q)] \\ &\quad = F(x,y(x),y'(x)) - S^{o}(x,y(x),y'(x)) \\ &\quad = 0 . \end{aligned}$$

If S is continuously differentiable as is assumed by Caratheodory's lemma, then S^{o} coincides with the classical directional derivative.

Although the value function will not be the only solution to the G.H.-J.Ineq., any solution can be used to establish a lower bound on the value function, as we see in the following lemma.

LEMMA II LOWER BOUNDS ON THE VALUE FUNCTION

Let $F(x,y,q)$ *be an integrand yielding the Hamiltonian* $H(x,y,p)$. *Let* $R(x,y)$ *be a locally lipschitz function defined on an open region* Ω *and satisfying the generalized Hamilton-Jacobi inequality*

$$H(a,b,\lambda_2) + \lambda_1 \leq 0 \qquad (\lambda_1,\lambda_2) \in \partial R(a,b)$$

at each (a,b) *in* Ω .

<u>If</u> z <u>is any lipschitz curve joining</u> $(a_1,b_1) \varepsilon \Omega$ <u>and</u> $(a_2,b_2) \varepsilon \Omega$ <u>with</u> $a_1 < a_2$ <u>and</u> $(x,z(x))$ <u>in</u> Ω <u>for almost all</u> $x \varepsilon [a_1,a_2]$, <u>it follows that</u>

$$\int_{a_1}^{a_2} F(x,z(x),z'(x))dx \geq R(a_2,b_2) - R(a_1,b_1) .$$

<u>Proof of lemma II</u>

For almost all $x \varepsilon [a_1,a_2]$ $z'(x)$ exists, the G.H.-J.Ineq. holds at $(x,z(x))$, and $\frac{d}{dx}R(x,z(x))$ exists. The equivalent integrand $F^*(x,y,q) = F(x,y,q) - R^o(x,y;1.q)$ is positive on Ω, and so by property 3, chapter II:

$$F(x,z(x),z'(x)) - \frac{d}{dx}R(x,z(x)) \geq 0$$

for almost all $x \varepsilon [a_1,a_2]$. Integrating, we get the desired result:

$$\int_{a_1}^{a_2} F(x,z(x),z'(x))dx - R(a_2,b_2) + R(a_1,b_1) \geq 0 .$$

Under the hypotheses of lemma II, let the point (a_o,b_o) lie in the closure of Ω; $\overline{\Omega}$. Assume also that we can define $R(a_o,b_o)$ as:

$$R(a_o,b_o) \equiv \underset{(a,b)\to(a_o,b_o)}{\mathrm{LIM'}} R(a,b)$$

where LIM' indicates the limit with (a,b) restricted to Ω.
If we let Ω coincide with the domain of the problem; that is, for (a,b) in Ω let

$$A(a,b) = \{\text{curves } y \mid y(a)=b \text{ and } y \text{ lies in } \Omega \text{ almost everywhere}\},$$

then we have

$$\underset{y\varepsilon A(a,b)}{\mathrm{INF}} \int_{a_o}^{a} F(x,y,y')dx \geq R(a,b) - R(a_o,b_o).$$

We see then, that the value function is the largest (majorizing) locally lipschitz solution to the G.H.-J.Ineq. which can be extended continuously by setting $S(a_o,b_o) = 0$. Often however, the following question is of more practical interest: 'given a locally lipschitz function R (which we suspect of being the value function), under what conditions are we assured that it is in fact the value function?'. A set of sufficient conditions is provided by lemma III below.

<u>LEMMA III</u> <u>SUFFICIENT CONDITIONS FOR THE VALUE FUNCTION</u>

<u>Let</u> $\Omega \varepsilon \mathbb{R}^2$ <u>be a set of termination and let</u> U , <u>the domain of the problem, coincide with</u> Ω . <u>If the integrand</u> F <u>is continuous then a function</u> R <u>defined on</u> $\Omega \cup \{(a_o,b_o)\}$ <u>which is locally lipschitz on</u> Ω <u>will be the value function for the problem if and only if, for</u> (a,b) <u>restricted to</u> Ω:

(i) R satisfies the G.H.-J.Ineq. on Ω ,

(ii) $R(a_o,b_o) = \underset{(a,b)\to(a_o,b_o)}{\mathrm{LIM}} R(a,b) = 0$, and

(iii) $\forall\ (a,b,) \in \Omega\ \ \exists\ \{y_n\}_{n=1} \subset A(a,b)$ such that:

$$\underset{n\to\infty}{\mathrm{LIM}} \int_{a_o}^{a} F(x,y_n,y_n')dx = R(a,b).$$

Proof of lemma III

Notice that since $U \equiv \Omega$, (a_o,b_o) must lie in the closure of Ω otherwise $A(a,b) = \emptyset$ for all (a,b).

For necessity; (i) follows from part (a) of theorem I, while (ii) and (iii) are consequences of the definition of the value function.

Consider now the sufficiency of the conditions. Lemma II and conditions (i) and (ii) provide that:

$$\underset{y\in A(a,b)}{\mathrm{INF}} \int_{a_o}^{a} F(x,y,y')dx \geq R(a,b) - R(a_o,b_o)$$
$$= R(a,b) ,$$

while condition (iii) provides the reverse inequality. This completes the proof.

Very often the basic problem is posed, not to find the minimum itself, but rather, to obtain the optimal curve(s) which provide the minimum. Having imbedded the problem in a family of problems, the theory involving the solutions for the entire family should help solve the original basic problem. In theorem II below we find pointwise sufficiency conditions for optimality in a basic problem.

THEOREM II SUFFICIENCY CONDITIONS FOR OPTIMALITY

Let $R(x,y)$ be a locally lipschitz solution to the generalized Hamilton-Jacobi inequality on a open set $\Omega \subset \mathbb{R}^2$. If y_o is a lipschitz curve to (a,b) lying in Ω almost everywhere and if

$$F(x,y_o(x),y_o'(x)) - \frac{d}{dx}R(x,y_o(x)) = 0$$

almost everywhere for $x \in [a_o,a]$, then y_o is an optimal curve for the basic problem: find

$$\inf_{y \in A(a,b)} \int_{a_o}^{a} F(x,y,y')dx$$

with

$$A(a,b) = \{\text{curves } y \mid y(a)=b \text{ and } y \text{ lies in } \Omega \text{ almost everywhere}\} .$$

Notice that Ω is both the set of termination and the domain of the problem. In practice it may be desired to find a solution over a domain U which properly includes the open set Ω. In this case, under the hypothesis of theorem II, the curve y_o may be considered a local solution. More precisely, if y_o lies entirely within Ω , then y_o is a <u>strong local solution</u> to the basic problem, and if the domain U coincides with Ω, then y_o is a <u>global solution</u> (see [7,§2.1]).

Since theorem II does not require that R be the value function, but only a solution to the G.H.-J.Ineq., solving a basic problem need not involve solving an entire family of problems.

<u>Proof of theorem II</u>

Integrating

$$F(x,y_o,y_o') - \frac{d}{dx}R(x,y_o(x)) = 0$$

we get:

$$\int_{a_o}^{a} F(x,y_o,y_o')dx - R(a,b) + R(a_o,b_o) = 0 .$$

From Lemma II, for all y in A(a,b),

$$\int_{a_o}^{a} F(x,y,y')dx - R(a,b) + R(a_o,b_o) \geq 0 ,$$

and so y_o is optimal in A(a,b), completing the proof.

Notice that since R satisfies the G.H.-J.Ineq.,

$$F(x,y_o(x),y_o'(x)) - R^o(x,y_o(x);1,y_o'(x)) \geq 0$$

wherever y_o' exists and $(x,y_o(x)) \varepsilon \Omega$. Since

$$\frac{d}{dx}R(x,y_o(x)) \leq R^o(x,y_o(x);1,y_o'(x))$$

where $y_o'(x)$ exists, we have

$$\begin{aligned} 0 &= F(x,y_o(x),y_0'(x)) - \frac{d}{dx}R(x,y_o(x)) \\ &\geq F(x,y_o(x),y_o'(x)) - R^o(x,y_o(x);1,y_o'(x)) \\ &\geq 0, \end{aligned}$$

and so

$$\frac{d}{dx}R(x,y_o(x)) = R^o(x,y_o(x);1,y_o'(x))$$

in Ω wherever y_o' exists.

CHAPTER IV EXAMPLES

The generalized gradient of the value function can be considered as a closed compact convex region in the x-y plane. If it is rotated 90° clockwise about the origin (that is, $\{(x,y)\} \to \{(-y,x)\}$) then the G.H.-J.Eq. can be expressed by saying that $\partial S(a,b)$ will lie above the graph of $y = H(a,b,x)$ and will touch the graph at one or more points.

EXAMPLE I

Let the integrand F be the following:

$$F(x,y,q) = \begin{cases} q^2-1 & \text{if } |q| \geq 1 , \\ 0 & \text{if } |q| \leq 1 . \end{cases}$$

Let the set of termination be $\Omega = \{(x,y) \mid x > 0 \}$, and the domain be $\mathbb{R}^2$. Let $(a_o,b_o) = (0,0)$ then for each $(a,b) \in \Omega$ the set of admissible curves is $A(a,b) = \{$lipschitz curves from $(0,0)$ to $(a,b)\}$. For each (a,b) in Ω consider the straight line curve $y(x) = bx/a$, which lies in $A(a,b)$. These satisfy the necessary condition of the Euler-Lagrange differential inclusion for optimality - see [4,theorem 2.4]. Define R, a lipschitz function on Ω as:

$$R(a,b) = \int_0^a F(x,bx/a,b/a)dx$$

$$= \begin{cases} (b^2- a^2)/a & \text{if } |b/a| \geq 1 , \\ 0 & \text{if } |b/a| \leq 1 . \end{cases}$$

The Hamiltonian for the integrand F is:

$$H(a,b,p) = \sup_q \begin{cases} pq - q^2 + 1 & \text{if } |q| \geq 1 , \\ pq & \text{if } |q| \leq 1 \end{cases}$$

$$= \begin{cases} p^2/4 + 1 & \text{if } |p| \geq 2 , \\ |p| & \text{if } |p| \leq 2 . \end{cases}$$

H(a,b,p) is independent of (a,b) $\varepsilon\ \Omega$, and a sketch of H as a function of p is given in figure I below.

The generalized gradient of R(a,b) is given by:

$$\partial R(a,b) \doteq \begin{cases} \{(-(a^2+b^2)/a^2 , 2b/a)\} & \text{if } |b/a| > 1 , \\ \{(0,0)\} & \text{if } |b/a| < 1 , \\ \{(-\lambda,\lambda) \mid 0 \leq \lambda \leq 2 \} & \text{if } b/a = 1 \\ \{(-\lambda,-\lambda) \mid 0 \leq \lambda \leq 2 \} & \text{if } b/a = -1 \end{cases}$$

This is represented in figure II below.

Notice that when $\partial R(a,b)$ is rotated clockwise it always lies <u>on</u> the graph of H(a,b,x) for any (a,b) in Ω, and so R satisfies the G.H.-J.Eq. on Ω. We can now verify each of the hypotheses of theorem II

(i) $y(x) = bx/a$ lies in Ω almost everywhere,

(ii) $F(x,y,y') = \frac{d}{dx} R(x,y,y')$,

(iii) R is lipschitz on Ω, and

(iv) R satisfies the G.H.-J.Ineq. on Ω;

hence each of the curves of the form $y(x) = bx/a$ are optimal curves to (a,b), and R coincides with the value function. These curves, however,

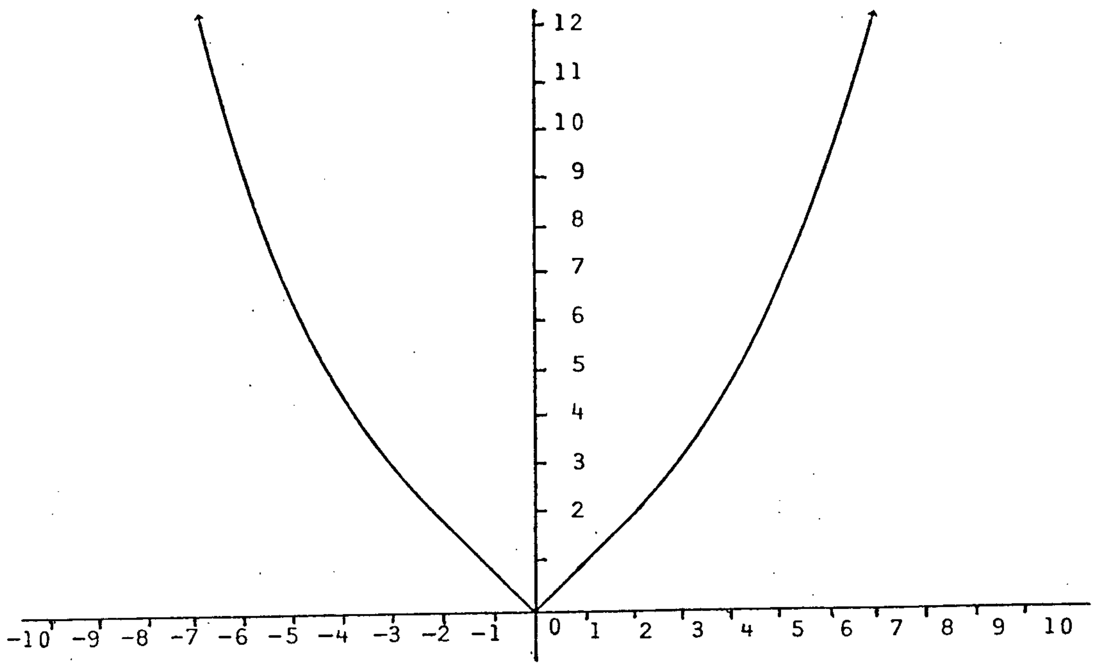


DIAGRAM I The Hamiltonian for Example I

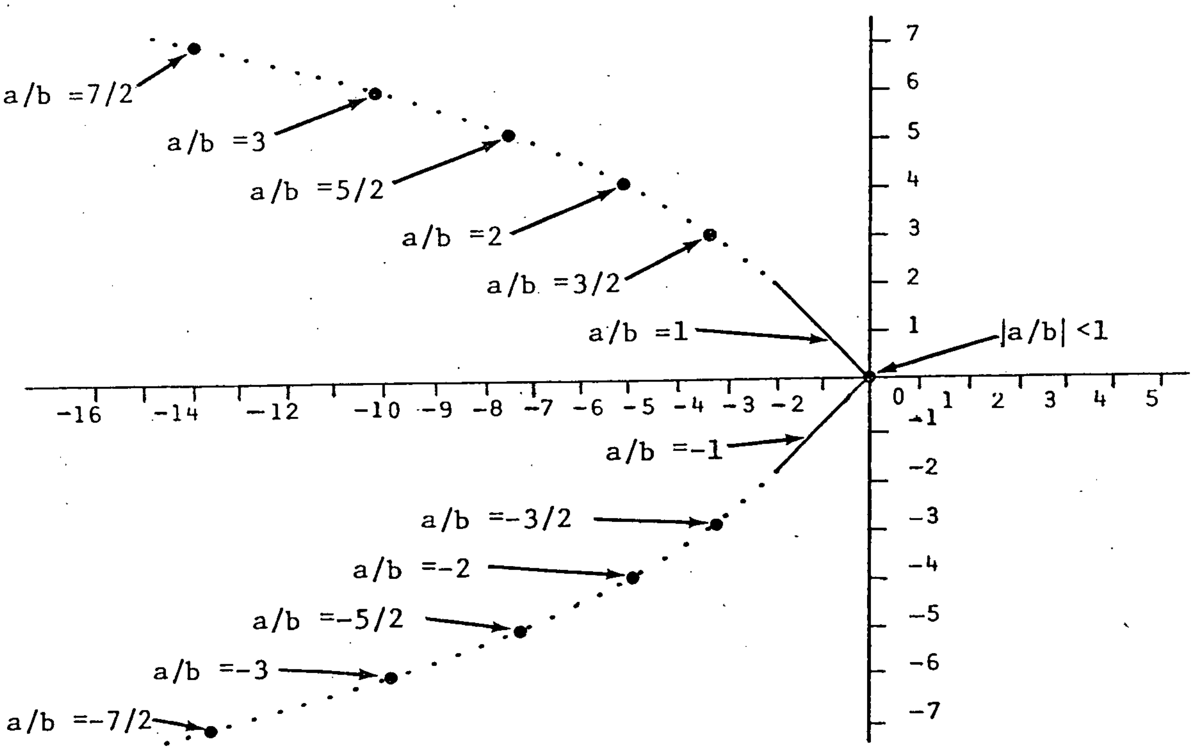


DIAGRAM II $\partial S(a,b)$ for Some Values of a/b, Example I are not unique optimal solutions since, for $0 < b/a < 1$, it is easy to construct others such as $y(x) = \mathrm{MIN}\{x,b\}$.

EXAMPLE II

Consider the integrand $F(x,y,q) = q^2 - y^2$. From classical theory we find that optimal curves, if they exist, should be of the form $y(x) = k \cdot \sin(x)$, $k \in \mathbb{R}$ (see [7,§2.3]). All such extremals pass through the point $(\pi,0)$ which is a conjugate point (again, see [7,§3.6]). Let Ω, the set of termination, be given by:

$$\Omega = \{(x, k\cdot\sin x) \mid 0 < x < 2\pi,\ x \neq \pi,\ |k| < 1\} .$$

Notice that Ω is open and disconnected. Let the domain be Ω and let $(a_o,b_o) = (0,0)$ then the set of admissible curves to $(a,b) \in \Omega$ will be:

$$A(a,b) = \{\text{lipschitz curves from } (0,0) \text{ to } (a,b) \text{ lying in } \Omega \text{ a.e.}\} .$$

For each (a,b) in Ω there is a unique $y \in A(a,b)$ of the form $k\cdot\sin x$: $y_{ab}(x) = k_{ab}\sin x$ with $k_{ab} = b/\sin a$. Define R on Ω as:

$$\begin{aligned} R(a,b) &= \int_0^a F(x, y_{ab}, y'_{ab})\,dx \\ &= b^2 \cot a . \end{aligned}$$

Let $R(0,0) = 0$; then R is continuous on $\Omega \cup \{(0,0)\}$ and R is seen to be lipschitz on Ω, with

$$\begin{aligned} ||\nabla R(a,b)|| &= ||(-b^2/\sin^2 a\ ,\ 2b\cot a)|| \\ &< 5 . \end{aligned}$$

The Hamiltonian for the problem is found to be

$$H(a,b,p) = p^2/4 + b^2$$

and it easily seen that R satisfies the G.H.-J.Eq. on Ω. By construction $\frac{d}{dx}R(x,y_{ab}(x),y'_{ab}(x)) = F(x,y_{ab}(x),y'_{ab}(x))$ so by corollary II y_{ab} is optimal in A(a,b) and hence R is the value function on Ω.

In a similar fashion, if we let

$$\Omega_m = \{(x,k\sin x) \mid 0 < x < 2\pi,\ x \neq \pi,\ |k| < m\},$$

and

$$A_m(a,b) = \{\text{ lipschitz curves from } (0,0) \text{ to } (a,b) \text{ within } \Omega_m \text{ a.e. }\},$$

then on Ω_m, y_{ab} will be optimal in $A_m(a,b)$. For $a > \pi$, any lipschitz curve from (0,0) to (a,b) which passes through $(\pi,0)$ must lie within Ω_m almost everywhere for some m, hence our curves y_{ab} are optimal over all lipschitz curves y(x) with $y(0) = 0$, $y(\pi) = 0$, and $y(a) = b$. Notice that if the point $(\pi,0)$ was included in the sets Ω or Ω_m, then our curves y_{ab} could no longer be optimal for $a > \pi$ by Jacobi's necessary condition (see [7, 3.6])

EXAMPLE III

Consider the smooth integrand

$$F(x,y,q) = y(1 + q^2)^{1/2}.$$

Let the set of termination be $\Omega = \{(x,y) \mid x > 0,\ y > 0\}$, let the domain coincide with Ω, and let $(a_o,b_o) = (0,1)$ then the set of admissible curves to $(a,b) \varepsilon \Omega$ is given by

$$A(a,b) = \{\text{lipschitz curves from } (0,1) \text{ to } (a,b) \text{ lying within } \Omega \text{ a.e.}\}.$$

This is a very old and often cited problem from the classical theory of the calculus of variations. It corresponds to minimizing the area of revolution of a positive function. Physically, the problem seeks the shape of a soap bubble spanning two concentric hoops. Experience indicates that as the two hoops are moved away from one another the bubble eventually breaks. We will see that this happens, not because of air currents or insufficient soap or any other accident but because eventually, locally optimal curves for the problem fail to exist, making the soap film unstable.

Of particular interest is the fact that the value function fails to be differentiable despite the fact that the integrand satisfies the classical requirements of being twice continuously differentiable and convex. In regions where there are no optimal curves we will be able to establish the value function, while with the same tools, in regions which have optimal curves we will find them and the value function. Where strong local solutions are found, a well defined locale within which they are optimal is also found.

The solutions to the Euler-Lagrange equation (2,§273) are of the form $y(x) = d\cdot\cosh(\frac{x-c}{d})$. To each point $(a,b) \varepsilon \Omega$ there are at most two curves of this form in $A(a,b)$ (see diagram III below). The ensemble of curves of the form $y(x) = d\cdot\cosh(\frac{x-c}{d})$ with $y(x) \varepsilon A(a,b)$

for some $(a,b) \in \Omega$ forms an envelope curve, 'E', as in diagram III. If we truncate each of the curves of the ensemble at the point where it touches the envelope (see [7, §A3.13]), there remains exactly one member of the ensemble lying in $A(a,b)$ for each $(a,b) \in \Omega$ lying above the envelope 'E'. The constants c and d are smooth functions of the coordinates (a,b) above 'E';

$$y_{ab}(x) = d(a,b) \cdot \cosh\left(\frac{x-c(a,b)}{d(a,b)}\right)$$

Define the locally lipschitz function R_1 on the region above the envelope 'E' as follows:

$$R_1(a,b) = \int_0^a F(x, y_{ab}, y'_{ab})\,dx$$

$$= \tfrac{1}{2}\left[\, da + d \cdot \sinh\left(\frac{a-c}{d}\right) \cdot b + d \cdot \sinh\left(\frac{c}{d}\right) \,\right],$$

with $c = c(a,b)$ and $d = d(a,b)$.

The Hamiltonian for the problem is given by:

$$H(x,y,q) = \sup_q \left[\, pq - F(x,y,q) \,\right]$$

$$= \begin{cases} -(y^2 - p^2)^{\frac{1}{2}} & \text{if } y \geq 0 \text{ and } |p| \geq y, \\ +\infty & \text{otherwise}. \end{cases}$$

The gradient of R_1 is found to be

$$\nabla R_1(a,b) = \left(d \,,\, \frac{|a-c|}{a-c} \cdot (b^2 - d^2)^{\frac{1}{2}} \right),$$

(again with $c = c(a,b)$ and $d = d(a,b)$) which satisfies the G.H.-J.Eq. on the region where R_1 is defined.

The hypotheses of the lemma III are satisfied on the region above the curve 'E', and so the curves y_{ab} are optimal in

$$A'(a,b) \equiv \{ y \in A(a,b) \mid y \text{ lies above 'E' almost everywhere } \} .$$

For the original problem however, we are assured only that the y_{ab} are strong local solutions. In search of the value function we will require a function defined on all of Ω .

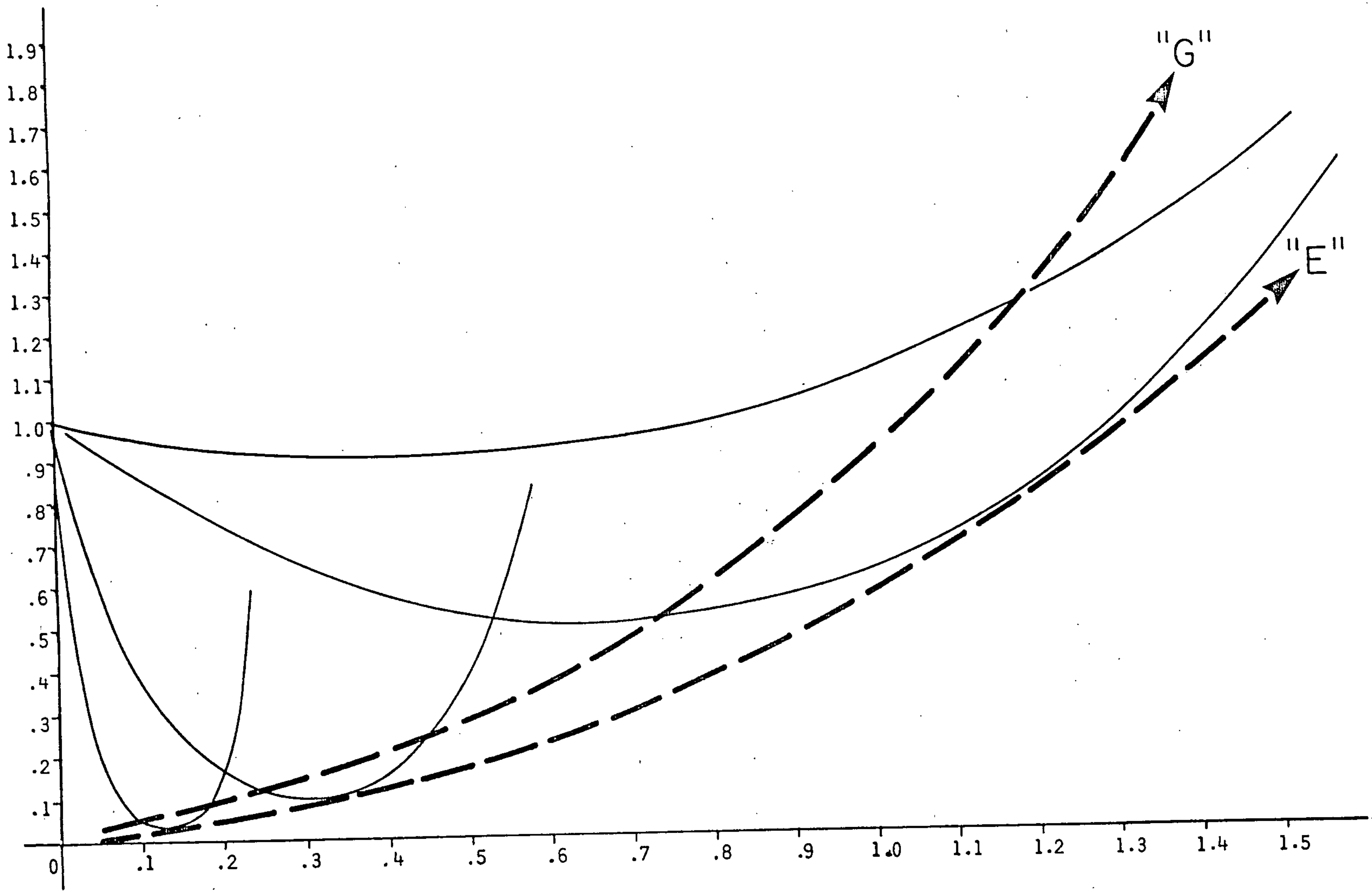
For each $(a,b) \in \Omega$ consider the sequence of curves $\{y_n\}_{n=1}^{\infty}$ defined as follows:

$$y_n(x) = \begin{cases} 1-nx & \text{if } 0 \le x \le (n-1)/n^2 \\ 1/n & \text{if } (n-1)/n^2 \le x \le a-(nb-1)/n^2 \\ b-n(a-x) & \text{if } a-(nb-1)/n^2 \le x \le a . \end{cases}$$

For sufficiently large n, y_n is an element of $A(a,b)$ and we find that

$$\lim_{n\to\infty} \int_0^a F(x,y_n,y_n')dx = (b^2+1)/2$$

Set $R_2(a,b)$ equal to this value, $(b^2+1)/2$, for each (a,b) in Ω. As computation verifies ([1]) $R_1 = R_2$ along a curve 'G' which lies above 'E' (see diagram III). Above the curve 'G' we have $R_1 < R_2$, while between 'G' and 'E' we have $R_1 > R_2$. Notice that R_2 is smooth and $\nabla R_2(a,b) = (0,b)$ on Ω, hence R_2 satisfies the G.H.-J.Eq. Notice however, that since $R_2(a_o,b_o) \neq 0$, R_2 cannot be the value



<u>DIAGRAM III</u> Extremal Curves for Example III

function. By combining R_1 and R_2 we will attempt to satisfy all the hypotheses of lemma III. Define R on Ω as follows:

$$R(a,b) = \begin{cases} \text{MIN } \{R_1, R_2\} & \text{on the region above 'E'} \\ R_2 & \text{elsewhere on } \Omega \ . \end{cases}$$

The function R is lipschitz on Ω, but fails to be differentiable along the curve 'G'. The generalized gradient of R is as follows:

$$\partial R(a,b) = \begin{cases} \{\nabla R_1(a,b)\} & \text{if } (a,b) \text{ is above 'G'} \\ \{\nabla R_2(a,b)\} & \text{if } (a,b) \text{ is below 'G'} \\ \text{co}\{\nabla R_1(a,b), \nabla R_2(a,b)\} & \text{if } (a,b) \text{ is on 'G'} \ . \end{cases}$$

Since; (i) R satisfies the G.H.-J.Eq. on Ω, (ii) $R(0,1) = 0$, and (iii) R is constructed from curves or limits of sequences of curves in $A(a,b)$, we conclude by lemma III that R is the value function.

We have seen that to the right of 'E', that is for large hoop separation, no film of minimal surface area exists, hence no bubble is expected to be observed. The minimum is obtained as the limit of a sequence which, loosely speaking, tends toward a situation in which each of the hoops has a flat film within its circumference. This is known as the Goldsmidt solution and it provides the minimal surface area in the parametrized form of the problem.

Between the curves 'G' and 'E' the catenaries defining R_1 exist, but are not optimal. In classical theory they are called strong local solutions. It can easily be seen that they are optimal in the smaller set of admissible functions:

$$A'(a,b) = \{ \ y \ \varepsilon \ A(a,b) \ | \ y \text{ lies above 'E'} \ \} \ .$$

We would expect that soap bubbles may exist for (a,b) between 'E' and 'G', but that they would not be stable except under sufficiently small perturbations.

EXAMPLE IV

Consider the problem with the integrand $F(x,y,q) = e^{-|y|} + e^{-|q|}$, domain $U = \{ \ (x,y) \ \varepsilon \ \mathbb{R} \ | \ x > 0 \ \}$ and set of termination $\Omega = U$. Let the initial point (a_o,b_o) be the origin. The integrand is neither convex nor smooth, so the classical Hamiltonian is not defined anywhere. The generalized Hamiltonian, however, is defined as follows:

$$H(x,y,q) = \begin{cases} -e^{-|y|} & \text{if } p = 0 \\ \infty & \text{otherwise} \ . \end{cases}$$

For a point $(a,b) \ \varepsilon \ \Omega$ let $\{y_n\}_{n=1}$ be the sequence of admissible curves to (a,b) given by

$$y_n(x) = \begin{cases} \frac{b}{a} nx & \text{for } 0 \le x \le \frac{n-1}{2n} \cdot a \\ (n-1) \cdot b - \frac{b}{a} \cdot nx & \text{for } \frac{n-1}{2n} \cdot a \le x \le a \ . \end{cases}$$

we find that:

$$\lim_{n \to \infty} \int_0^a F(x,y_n,y_n')dx = 0 \ .$$

Let $R(a,b) = 0$ for all (a,b) in $\Omega \cup \{(0,0)\}$, then R satisfies the G.H.-J.Ineq. on Ω since $H(a,b,0) = -e^{-|b|} \le 0$. Since R satisfies all the hypotheses of lemma III, R is the value function on Ω.

Notice however, that R fails to satisfy the G.H-J.Eq. Since existence of optimal curves would guarantee that R satisfies the G.H.-J.Eq. (by theorem I part (b)) we must conclude that no optimal curves exist.

REFERENCES

1) G.A.Bliss, 'Lectures on the Calculus of Variations', Chicago Ill., University of Chicago Press, 1946.

2) C.Caratheodory, 'Calculus of Variations and Partial Differential Equations of the 1st Order', (1937), translated by R.B.Dean, Holden-Day Inc., San Franciso Ca., 1967.

3) Frank H. Clarke, 'Generalized Gradients and Applications', Trans. Amer. Math. Soc. 205 (1975) pp247-262.

4) Frank H. Clarke, 'Euler Lagrange Differential Inclusion', J.Differential Equations Vol. 19 (1975) pp80-90.

5) I.M.Gel'Fand and S.V.Fomin, 'Calculus of Variations', translated by R.A.Silverman, Englewood Cliffs N.J., Prentice-Hall 1963.

6) H.Goldstein, 'Classical Mechanics', Addison Wesley, Reading Mass.,1950.

7) Hans Sagan, 'Introduction to the Calculus of Variations', N.Y.,McGraw-Hill 1969.

8) D.R.Snow, 'A Sufficiency Technique in the Calculus of Variations Using Caratheodory's Equivalent Problems Approach', J. of Math. Analysis and Applications Vol. 51 (1975) pp129-140.