THE UNIVERSITY OF ALBERTA

SUFFICIENCY CRITERIA FOR OPTIMAL CONTROL THEORY

BY

DAVID JOHN ORRELL

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS

EDMONTON, ALBERTA

SPRING 1986

ISBN 0-315-30158-9

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: David John Orrell

TITLE OF THESIS: Sufficiency Criteria for Optimal Control Theory

DEGREE: M.Sc.

YEAR DEGREE GRANTED: 1986

8520-104 Street,

Edmonton, Alberta

Date: 19

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled ....SUFFICIENCY CRITERIA FOR OPTIMAL CONTROL THEORY.............. submitted by ....DAVID JOHN ORRELL........................... in partial fulfilment of the requirements for the degree of Master of Science.

......Vera Zeidan......
(Supervisor)

......Jack W. Macki......

......V. G. Gowrisankaran......

......Jim McCleary......

Date: October 21. 1985

ABSTRACT

This thesis is concerned with three problems: the calculus of variations, optimal control and the generalized problem of Bolza. It presents known sufficiency conditions for each of the problems, and explains their development using the modified Hamilton-Jacobi approach. It gives a new sufficient condition for the optimal control problem where the control set is given by smooth functions. This criterion generalizes prior results of the same kind when the control set is polyhedral.

In the section dealing with the calculus of variations, an improvement is made on certain known results. Finally, the interrelationships between the various sufficiency conditions are studied.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Vera Zeidan, for the great amount of time and effort that she spent working with me on this thesis, and for introducing me to this interesting and enjoyable field of research.

My thanks also to Janina Barteluk without whose help this thesis would never have been completed on time.

TABLE OF CONTENTS

1. INTRODUCTION

Three areas of importance in the subject of optimization are the calculus of variations, optimal control, and the generalized problem of Bolza. These are ranked historically and in terms of generality; a calculus of variations problem may be written as an optimal control problem, which may in turn be posed as a generalized problem of Bolza. It is our aim to state sufficiency conditions for the three problems and show how these conditions interrelate. We will provide assumptions under which one result is a corollary to another, and find counterexamples when one result may be applied but another may not. Special attention is paid to the problem of optimal control, for which a new sufficiency condition has been developed. We show how and when the new condition improves on known ones.

The generalized problem of Bolza is:

$$\text{minimize } J(x) = l(x(a),x(b))+\int_a^b L(t,x(t),\dot{x}(t))dt$$

where $x:[a,b]\to\mathbb{R}^n$ is absolutely continuous with derivative $\dot{x}$ (almost everywhere) and where

$L:[a,b]\times\mathbb{R}^n\times\mathbb{R}^n\to(-\infty,+\infty]$ and $l:\mathbb{R}^n\times\mathbb{R}^n\to(-\infty,+\infty]$ are given. We call such a function x an arc. The Hamiltonian of this problem is defined to be:

$$H(t,x,p) = \sup\{\langle p,v\rangle - L(t,x,v):v\in\mathbb{R}^n\}.$$

We make the assumption that L is $\mathcal{L}\times\mathcal{B}$- measurable, where $\mathcal{L}$ is the collection of Lebesgue measurable subsets of $[a,b]$ and $\mathcal{B}$ the Borel subsets of $\mathbb{R}^n \times \mathbb{R}^n$.

In Chapter 2 we present known conditions guaranteeing the optimality of an arc x for this problem. These conditions are based on the modified Hamilton-Jacobi inequality of [22] and [24]. The results are written in terms of the Hamiltonian.

Chapter 3 is concerned with the problem of optimal control:

$$\text{minimize } \ell(x(b))+\int_a^b g(t,x(t),u(t))dt$$

over all absolutely continuous functions $x:[a,b]\to\mathbb{R}^n$ with derivative $\dot{x}$ (almost everywhere), and over all measurable functions $u:[a,b]\to\mathbb{R}^n$ satisfying:

$$\dot{x}(t) = f(t,x(t),u(t)) \text{ a.e., } x(a) = A, \text{ and } u(t)\in U \text{ a.e.}$$

The Hamiltonian is defined as follows:

$$H(t,x,p) = \sup\{\langle p,f(t,x,u)\rangle - g(t,x,u):u\in U\} .$$

It has been shown in [17] that the above problem may be written as a generalized problem of Bolza where the Hamiltonian is the same.

In this chapter we state known sufficiency conditions (see [23] and [25]) parallelling those of chapter 2, including second order results written both in terms of the Hamiltonian and in terms of the original data. The main contribution of this thesis lies in a new condition, adapted from these second order results for the special case when U is given by smooth functions, that is superior for U polyhedral.

The calculus of variations is a generalized problem of Bolza where L is a real-valued function, the boundary values $x(a) = A$ and $x(b) = B$ are given, and the minimum is taken over all piecewise smooth $x:[a,b] \to \mathbb{R}^n$. It is also easily seen to be a problem of optimal control.

In Chapter 4 we present the sufficency criteria of [25] for the calculus of variations. We also improve slightly on the second order result of that reference.

Chapter 5 is devoted to the interrelationships between the sufficiency conditions of the previous chapters. It includes new work that clarifies the situation for known results.

Various conditions necessary for optimality have been developed for the three problems, and often our results involve these conditions or strengthened versions of them. We will not state these conditions at the outset, but will mention them as they occur.

2. THE GENERALIZED PROBLEM OF BOLZA. KNOWN RESULTS

Consider the generalized problem of Bolza (P) minimize
$J(x) = l(x(a),x(b))+\int_a^b L(t,x(t),\dot{x}(t))dt$ over arcs x, where
$L:[a,b]\times\mathbb{R}^n\times\mathbb{R}^n\to(-\infty,+\infty]$ and $l:\mathbb{R}^n\times\mathbb{R}^n\to(-\infty,+\infty]$. We here present known sufficiency conditions, developed by Zeidan in [21], [22] and [24], ensuring the optimality of a certain candidate for the problem (P). We begin by defining optimality.

Def^n: Given a function $z:[a,b]\to\mathbb{R}^q$ and a positive number ε, we define the ε-tube about z, denoted $T(z;\varepsilon)$, to be
$T(z;\varepsilon) = \{(t,y):t\in[a,b],y\in\mathbb{R}^q,|y-z(t)|<\varepsilon\}$.
The projection of $T(z;\varepsilon)$ on $\mathbb{R}^q$ is denoted by $N(z;\varepsilon)$.
We say that a function $y\in T(z;\varepsilon)$ if $(t,y(t))\in T(z;\varepsilon)$ for $t\in[a,b]$ a.e.

Def^n: Given an arc $\hat{x}$ such that $J(\hat{x})$ is finite, we say that $\hat{x}$ is a strong local minimum for the problem (P) if we can find a positive number ε such that $\hat{x}$ minimizes $J(x)$ over all arcs x satisfying, for all $t\in[a,b]$,

$$(t,x(t))\in T(\hat{x};\varepsilon).$$

The first sufficiency condition, from which all the results of this section are ultimately derived, is the following.

Prop^{n}2.1 [Zeidan] Assume that L is $\mathcal{L}\times\mathcal{B}$-measurable and that $\hat{x}$ is a given arc with $J(\hat{x})$ finite. Suppose that there exist a positive number ε and a function $W(t,x)$ defined on $T(\hat{x};\varepsilon)$ such that, for all arcs $x \in T(\hat{x};\varepsilon)$, the function $W(\cdot,x(\cdot))$ is absolutely continuous and

(a) $\frac{d}{dt} W(t,x(t)) - L(t,x(t),\dot{x}(t)) \leq \frac{d}{dt} W(t,\hat{x}(t)) - L(t,\hat{x}(t),\dot{\hat{x}}(t))$ a.e.;

(b) for all c,d with $|c| < \varepsilon$ and $|d| < \varepsilon$ we have:

$$W(a,\hat{x}(a)+c) - W(a,\hat{x}(a)) + W(b,\hat{x}(b)) - W(b,\hat{x}(b)+d)$$
$$\leq l(\hat{x}(a)+c,\hat{x}(b)+d) - l(\hat{x}(a),\hat{x}(b)).$$

Then $J(x)$ is well defined (possibly $+\infty$) for x near $\hat{x}$, and $\hat{x}$ is a strong local minimum for (P). Moreover, if $\varepsilon = +\infty$ than $\hat{x}$ is a global minimum for (P).

Remark: If the function $W(\cdot,\cdot)$ is Lipschitz, then $W(\cdot,x(\cdot))$ is absolutely continuous for any arc x. In the case where the boundary values are fixed ($x(a) = A, x(b) = B$), the condition (b) is automatically satisfied for all c,d in $\mathbb{R}^n$ and any function W. Note that in this case

$$l(x_1,x_2) = \chi_{\{A\}}(x_1) + \chi_{\{B\}}(x_2),$$

where for any set C

$$\chi_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise.} \end{cases}$$

Proof: Refer to [24].

The Hamilton-Jacobi equation, given in [14] for the Calculus of Variations, has been adapted in the generalized problem of Bolza to yield two sufficient conditions, involving respectively the modified and the generalized Hamilton-Jacobi inequality. We now state these results as corollaries of Prop^{n}2.1.

Th^{m}2.1. [Zeidan] Let the function L be $\mathcal{L} \times \mathcal{B}$ - measurable and let $\hat{x}$ be an arc such that $J(\hat{x})$ is finite. Suppose that there exist a positive number ε and a Lipschitz function $W(t,x)$ defined on $T(\hat{x};\varepsilon)$ satisfying:

(1) for all c,d such that $|c| < \varepsilon$ and $|d| < \varepsilon$,

$$W(a,\hat{x}(a)+c)-W(a,\hat{x}(a)) + W(b,\hat{x}(b))-W(b,\hat{x}(b)+d)$$
$$\leq l(\hat{x}(a)+c,\hat{x}(b)+d) - l(\hat{x}(a),\hat{x}(b));$$

(2) for $Z(t,x) = \sup \{\alpha+H(t,x,\beta):(\alpha,\beta)\in\partial W(t,x)\}$,

$$Z(t,x) \leq Z(t,\hat{x}(t)) \quad t\in[a,b] \text{ a.e.}$$

$$Z(t,\hat{x}(t)) = \frac{d}{dt} W(t,\hat{x}(t)) - L(t,\hat{x}(t),\dot{\hat{x}}(t)) \text{ a.e.}$$

Then $J(x)$ is well defined (possibly $+\infty$) for x near $\hat{x}$, and $\hat{x}$ is a strong local minimum for (P). Moreover, if $\varepsilon = +\infty$ then $\hat{x}$ is a global minimum for (P).

Condition (2) is the modified Hamilton-Jacobi inequality. The proof is obtained by showing that a Lipschitz function $W(t,x)$

satisfying condition (2) of Th^{m}2.1 also satisfies condition (b) of Prop^{n}2.1; see [24] for details.

When we take $Z(t,\hat{x}(t)) \equiv 0$ condition (2) of Th^{m}2.1 becomes the generalized Hamilton-Jacobi inequality, which was developed in [14] for the problem (P) in Mayer's form, and is incorporated in the following result.

Cor 2.1. [Zeidan] In Th^{m}2.1, condition (2) may be replaced by the following:

For $Z(t,x) = \sup\{\alpha + H(t,x,\beta) : (\alpha,\beta) \in \partial W(t,x)\}$,

$Z(t,x) \leq 0$ for $t \in [a,b]$ a.e.,

and

$\frac{d}{dt} W(t,\hat{x}(t)) - L(t,\hat{x}(t),\dot{\hat{x}}(t)) = 0$ a.e.

The proof, which appears in [21], is obtained by showing that a Lipschitz function W satisfying the conditions of Corollary 2.1 also satisfies those of Th^{m}2.1.

We now state a sufficiency condition derived from Prop^{n}2.1 and calling for the existence of an absolutely continuous matrix function Q(t) satisfying a certain inequality.

Th^{m}2.2. [Zeidan] Assume that L is $\mathcal{L} \times \mathcal{B}$- measurable and that $\hat{x}$ is a given arc with $J(\hat{x})$ finite. Suppose that there exist a positive number ε, an arc $\hat{p}$ and an absolutely continuous symmetric matrix function Q(t) such that:

(1) $L(t,\hat{x}(t),\dot{\hat{x}}(t)+v)-L(t,\hat{x}(t),\dot{\hat{x}}(t)) \geq \langle \hat{p}(t),v\rangle$

for all v in $\mathbb{R}^n$ and almost all $t\epsilon[a,b]$;

(2) $H(t,x,\hat{p}(t)-Q(t)(x-\hat{x}(t))) - H(t,\hat{x}(t),\hat{p}(t))$

$\leq - \langle \dot{\hat{p}}(t),x - \hat{x}(t)\rangle - \langle \dot{\hat{x}}(t), Q(t)(x - \hat{x}(t))\rangle$

$+ 1/2 \langle x - \hat{x}(t),\dot{Q}(t)(x - \hat{x}(t))\rangle$

almost everywhere in t with $(t,x) \epsilon T(\hat{x};\varepsilon)$;

(3) for all c,d with $|c|< \varepsilon$ and $|d|< \varepsilon$,

$l(\hat{x}(a)+c,\hat{x}(b)+d)-l(\hat{x}(a),\hat{x}(b))$

$\geq \langle\hat{p}(a),c \rangle - \langle \hat{p}(b),d \rangle - 1/2 \langle c,Q(a)c \rangle + 1/2 \langle d,Q(b)d \rangle$.

Then J(x) is well-defined (possibly $+ \infty$) for x near $\hat{x}$, and $\hat{x}$ is a strong local minimum for (P). If $\varepsilon = + \infty$ then $\hat{x}$ is a global minimum.

<u>Remark</u>. The existence of a function $\hat{p}$ (not necessarily absolutely continuous) satisfying condition (1) is a necessary condition (see [2]).

The proof of Th^{m}2.2 appears in [24], where it is shown that, given the conditions of Th^{m}2.2, the function

$W(t,x) = \langle \hat{p}(t),x \rangle - 1/2\langle x-\hat{x}(t),Q(t)(x-\hat{x}(t))\rangle$, defined on $T(\hat{x};\varepsilon)$, satisfies the conditions of Prop^{n}1.1.

<u>Remark</u>. In the special case where $l(\cdot,\cdot)$ and $-H(t,\cdot,\hat{p}(t))$ are convex and $\hat{p}$ is an arc satisfying the Hamiltonian inclusions and the transversality condition, the matrix $Q(t)\equiv 0$ satisfies conditions (2) and (3) of our theorem.

If we define on $\bar{T}(\hat{x};\varepsilon)$ the function

$$V(t,x) = H\big(t,x,\hat{p}(t)-Q(t)(x-\hat{x}(t))\big)+\langle \dot{\hat{p}}(t),x-\hat{x}(t)\rangle + \langle\dot{\hat{x}}(t),Q(t)(x-\hat{x}(t))\rangle - 1/2\ \langle x-\hat{x}(t),\dot{Q}(t)(x-\hat{x}(t))\rangle$$

then condition (2) of the theorem is eqivalent to the condition that $V(t,x) \leq V(t,\hat{x}(t))$ almost everywhere in t with $(t,x)\epsilon T(\hat{x};\varepsilon)$. By imposing first order or second order conditions on the function $V(t,x)$, we may guarantee that this is the case. For the first order result see [24]; we prefer to pass to the second order result, which is proven in the same reference. We need, from [22] and [24], the following.

[illegible]: The Hamiltonian H is said to be C^{1+} near a given arc $\hat{z} = (\hat{x},\hat{p})$ if we can find some positive γ such that for each t in $[a,b]$, $H(t,\cdot)$ is C^1 with locally Lipschitz first derivatives on the set of x with $|x-\hat{x}(t)| < \gamma$.

If H is C^{1+} near $\hat{z}$ then the generalized Jacobian ([6]) $\partial_z H_z(t,\cdot)$ exists on x with $|x-\hat{x}(t)|< \gamma$, and it is defined at a point z as being the convex hull of all matrices M of the form

$$M = \lim_{i\to\infty} \{D_z H_z(t,z_i)\}$$

where z_i converges to z and the usual Jacobian $D_z H_z(t,z_i)$ exists for each i.

<u>Def^n</u>: Let the Hamiltonian H be C^{1+} near $\hat{z} = (\hat{x}, \hat{p})$. The <u>extended Jacobi condition</u> is said to be satisfied at $\hat{z}$ if there exists a Lipschitz matrix function $Q(\cdot)$ from $[a,b]$ to the space of nxn - matices such that, for all t in $[a,b]$, $Q(t)$ is symmetric and satisfies:

$$\eta(t) - Q(t)\gamma(t)Q(t) + Q(t)\beta(t) + \delta(t)Q(t) - \alpha(t) > 0$$

for all t in $[a,b]$, for all matrices

$$\begin{vmatrix} \alpha(t) & \delta(t) \\ \beta(t) & \gamma(t) \end{vmatrix} \in \partial_z H_z(t, \hat{z}(t)),$$

and for all $\eta(t) \in \partial Q(t)$.

Suppose we are given arcs $\hat{x}, \hat{p}$ from $[a,b]$ to $\mathbb{R}^n$. The following hypothesis will be made:

The Hamiltonian H is C^{1+} on $T(\hat{z};\gamma)$
and the map **(H)**

$$(t,z) \to \partial_z H_z(t,z)$$

is upper semicontinuous on $T(\hat{z};\gamma)$.

<u>**Th^{m}2.3.**</u> **[Zeidan] Assume that L is $\mathcal{L} \times \mathcal{B}$- measurable and that $\hat{z} = (\hat{x}, \hat{p})$ is a given arc such that $J(\hat{x})$ is finite and hypothesis (H) holds. In addition, assume that there exists a Lipschitz symmetric matrix**

function $Q(t)$ such that

(a) $L(t,\hat{x}(t),\dot{\hat{x}}(t)+v)-L(t,\hat{x}(t),\dot{\hat{x}}(t)) \geq \langle\hat{p}(t),v\rangle$

for all $v \in \mathbb{R}^n$ and almost all $t \in [a,b]$;

(b) $\left(-\dot{\hat{p}}(t),\dot{\hat{x}}(t)\right) = H_z\left(t,\hat{z}(t)\right)$ a.e.;

(c) for all c,d with $|c|<\varepsilon$ and $|d|<\varepsilon$,

$l(\hat{x}(a)+c,\hat{x}(b)+d)-l(\hat{x}(a),\hat{x}(b))$

$\geq \langle p(a), c\rangle - \langle\hat{p}(b),d\rangle - 1/2\langle c,Q(a)c\rangle+1/2\langle d,Q(b)d\rangle$;

(d) the extended Jacobi condition is satisfied by some matrix function Q.

Then $J(x)$ is well defined (possibly $+\infty$) near $\hat{x}$, and $\hat{x}$ provides a strong local minimum for (P).

<u>Remark:</u> The proof of this theorem is given in [24] using the modified Hamilton-Jacobi approach. Another proof is presented in [21] and [22] that uses canonical transformations of Hamiltonian equations.

<u>Remark:</u> Condition (a) is called the Weierstrass condition and condition (b) is called the Hamiltonian equations. Under some additional hypotheses, we know from [2], [3] and [5] that these conditions are necessary.

This completes our presentation of results concerning sufficient conditions for the Generalized Problem of Bolza. The following diagram illustrates how the theory has been developed; an arrow $A \to B$ means that B has been proven as a corollary of A.

Figure (1)

$$
\begin{tikzcd}
\text{Prop}^{n}2.1 \arrow[r] \arrow[d] & \text{Th}^{m}2.1 \ \text{(mod. H-J)} \arrow[r] & \text{Cor } 2.1 \ \text{(gen. H-J)} \\
\text{Th}^{m}2.2 \ \text{(zero order)} \arrow[d] & & \\
\text{Th}^{m}2.3 \ \text{(second order)} & &
\end{tikzcd}
$$

As the remainder of the thesis wil be concerned with problems in which at least one boundary value is fixed, we make the following remark.

<u>**Remark:**</u> Consider the generalized problem of Bolza (P) with a given boundary value $x(a) = A$. Then the function $l(x_1,x_2)$ takes the form

$$l(x_1,x_2) = \chi_{\{A\}}\ (x_1)+l^{\circ}(x_2)$$

where $l^{\circ}: \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$, and $\chi_{\{A\}}\{x\}$ is the indicator function of the set $\{A\}$,

that is,

$$\chi_{\{A\}}(x) = \begin{cases} 0 & \text{if } x = A \\ +\infty & \text{if } x \neq A \, . \end{cases}$$

3. OPTIMAL CONTROL. KNOWN AND NEW RESULTS

A problem of optimal control may be written as a generalized problem of Bolza, and may therefore be solved by similar means.

In this chapter we present three known sufficiency criteria that are adapted directly from results of Chapter 2, and then go on to develop a new sufficiency criterion in a more specialized context, namely when the control set is given by smooth functions. It is seen that when the control set is polyhedral, our results improve on prior results of the same kind.

Let f, g and l be given functions:

$f:[a,b]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$, $g:[a,b]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ and $l:\mathbb{R}^n\to\mathbb{R}$.

Let U be a closed subset of $\mathbb{R}^m$ and A a point of $\mathbb{R}^n$. The optimal control problem is defined to be

(C) minimize $J(x,u) = l(x(b)) + \int_a^b g(t,x(t),u(t))dt$

over all absolutely continous functions $x:[a,b]\to\mathbb{R}^n$ with derivative $\dot{x}$ (almost everywhere), and over all measurable functions $u:[a,b]\to\mathbb{R}^m$ satisfying

$\dot{x}(t) = f(t,x(t),u(t))$ a.e., (3.1)

$x(a) = A$ (3.2)

$u(t)\in U$ a.e. (3.3)

The Hamiltonian is defined as

$H(t,x,p) = \sup \{\langle p,f(t,x,u)\rangle - g(t,x,u) : u \in U\}$.

Def^n: An absolutely continuous function from $[a,b]$ to R^n is called an arc. A pair (x,u) is admissible for (C) if x is an arc, u is measurable, and (x,u) satisfies (3.1)-(3.3).

Def^n: An admissible pair $(\hat{x},\hat{u})$ is a weak local minimum for (C) if there exists a positive number ε such that $(\hat{x},\hat{u})$ minimizes $J(x,u)$ over all admissible pairs (x,u) satisfying, for all $t \in [a,b]$,

$$\big(t,x(t),u(t)\big) \in T\big((\hat{x},\hat{u});\varepsilon\big).$$

Def^n: An admissible pair $(\hat{x},\hat{u})$ is a strong local minimum for (C) if for some $\varepsilon > 0$, $(\hat{x},\hat{u})$ minimizes $J(x,u)$ over all admissible pairs (x,u) satisfying, for all $t \in [a,b]$,

$$\big(t,x(t)\big) \in T(\hat{x};\varepsilon).$$

The following assumption will be made:

f and g are measurable on $T(\hat{x};\varepsilon)xU$ and, for each $t, g(t,\cdot,\cdot)$ is lower semicontinuous and $f(t,\cdot,\cdot)$ is continuous on $N(\hat{x};\varepsilon)xU$. (H_1).

3.1 Known Results

The next three results parallel Prop^{n}2.1, Th^{m}2.1 and Th^{m}2.2, and are similarly proven.

Prop^{n}3.1. [Zeidan] Let $(\hat{x},\hat{u})$ be an admissible pair for (C) such that

$J(\hat{x},\hat{u})$ is finite. Assume that for some positive ε hypothesis (H_1) holds and there exists a function $W(t,x)$ defined on $T(\hat{x};\varepsilon)$ such that, for all admissible (x,u) with $|x(t)-\hat{x}(t)|<\varepsilon\ \forall t\in[a,b]$, the function $W(\cdot,x(\cdot))$ is absolutely continuous and

(a) $\frac{d}{dt}W(t,x(t))-g(t,x(t),u(t))$
$\leq \frac{d}{dt}W(t,\hat{x}(t))-g(t,\hat{x}(t),\hat{u}(t))$

(b) for all d with $|d|<\varepsilon$ we have
$W(b,\hat{x}(b))-W(b,\hat{x}(b)+d)\leq l(\hat{x}(b)+d)-l(\hat{x}(b))$.

Then $J(x,u)$ is well defined (possible $+\infty$) for x near $\hat{x}$, and $(\hat{x},\hat{u})$ is a global minimum for (C).

<u>Th^{m}3.1.</u> [Zeidan] Let $(\hat{x},\hat{u})$ be an admissible pair for (C) such that $J(\hat{x},\hat{u})$ is finite. Assume that for some positive ε hypothesis (H_1) holds and there exists a Lipschitz function $W(t,x)$ defined on $T(\hat{x};\varepsilon)$ such that, for all admissible (x,u), we have

(1) $W(b,\hat{x}(b))-W(b,\hat{x}(b)+d)\leq l(\hat{x}(b)+d)-l(\hat{x}(b))$ for all $|d|<\varepsilon$;

(2) for $Z(t,x)=\sup\{\alpha+H(t,x,\beta):(\alpha,\beta)\in\partial W(t,x)\}$,
$Z(t,x)\leq Z(t,\hat{x}(t))$ for $t\in[a,b]$ a.e. and
$Z(t,\hat{x}(t))=\frac{d}{dt}W(t,\hat{x}(t))-g(t,\hat{x}(t),\hat{u}(t))$ a.e.

Then $J(x,u)$ is well defined (possibly $+\infty$) for x near $\hat{x}$, and $(\hat{x},\hat{u})$ is a strong local minimum for (C). Moreover, if $\varepsilon = +\infty$ then $(\hat{x},\hat{u})$ is a global minimum for (C).

Th^{m}3.2. [Zeidan] Let $(\hat{x},\hat{u})$ be an admissible pair for (C) such that $J(\hat{x},\hat{u})$ is finite. Assume that there exist on arc $\hat{p}$, an absolutely continuous symmetric matrix function $Q(t)$, and a positive number ε that:

(i) $\langle \hat{p}(t), f(t,\hat{x}(t),u) \rangle - g(t,\hat{x}(t),u)$

$\leq \langle \hat{p}(t), f(t,\hat{x}(t),\hat{u}(t)) \rangle - g(t,\hat{x}(t),\hat{u}(t))$

for $t \in [a,b]$ a.e. and all $u \in U$;

(ii) $H(t,x,\hat{p}(t)-Q(t)(x-\hat{x}(t))) - H(t,\hat{x}(t),\hat{p}(t))$

$\leq - \langle \dot{\hat{p}}(t), x-\hat{x}(t) \rangle - \langle \dot{\hat{x}}(t), Q(t)(x-\hat{x}(t)) \rangle$

$+ 1/2 \langle x-\hat{x}(t), \dot{Q}(t)(x-\hat{x}(t)) \rangle$

for $t \in [a,b]$ a.e. with $(t,x) \in T(\hat{x};\varepsilon)$;

(iii) $l(\hat{x}(b)+d) - l(\hat{x}(b)) \geq - \langle \hat{p}(b), d \rangle + 1/2 \langle d, Q(b)d \rangle$

for all d with $|d| < \varepsilon$.

Then $(\hat{x},\hat{u})$ is a strong local minimum for the problem (C).

As with their counterparts of Chapter 2, Th^{m}3.1 and Th^{m}3.2 are corollaries of Prop^{n}3.1.

<u>Remark:</u> Assume that $l(\cdot)$ is C^2 near $\hat{x}(b)$ with $\hat{p}(b) = -l_x(\hat{x}(b))$, and that $Q(t)$ is an $n \times n$ matrix function satisfying $Q(b) < l_{xx}(\hat{x}(b))$.

Then for some $\zeta > 0$, we have, for all

d with $|d| < \zeta$, for all $c \in {}^n$,

$l(\hat{x}(a)+c, \hat{x}(b)+d) - l(\hat{x}(a), \hat{x}(b)) \geq \langle \hat{p}(a), c\rangle$

$-\langle p(b), d\rangle - 1/2\langle c, Q(a)c\rangle + 1/2\langle d, Q(b)d\rangle$.

If we define on $T(x;\varepsilon)xU$ the function

$$F(t,x,u) = \langle \hat{p}(t) - Q(t)(x-\hat{x}(t)), f(t,x,u)\rangle - g(t,x,u)$$
$$- 1/2 \langle x-\hat{x}(t), \dot{Q}(t)(x-\hat{x}(t))\rangle + \langle \dot{\hat{p}}(t), x-\hat{x}(t)\rangle \qquad (3.4)$$
$$+ \langle \dot{\hat{x}}(t), Q(t)(x-\hat{x}(t))\rangle$$

then condition (ii) of $\text{Th}^m 3.2$ may be expressed as the condition that

$$F(t,x,u) \leq F(t,\hat{x}(t),\hat{u}(t))$$

for almost all t in $[a,b]$ with $(t,x) \in T(\hat{x};\varepsilon)$, and all $u \in U$. This formulation suggests the following corollary of $\text{Th}^m 3.2$.

<u>Cor.3.1</u>. [Zeidan] Let $(\hat{x},\hat{u})$ be admissible for (C) with $J(\hat{x},\hat{u})$ finite. Assume that U is compact and that, for some $\gamma > 0$, f and g are continuous on $T(\hat{x};\gamma)xU$. Suppose that there exist an arc $\hat{p}$ and an absolutely continuous matrix function $Q(t)$ such that condition (iii) of $\text{Th}^m 3.2$ holds and

(a) $F(t,x,u) \leq F(t,\hat{x}(t),\hat{u}(t))$ for $u \in U$ and $t \in [a,b]$ a.e. with $(t,x,u) \in T(\hat{x},\hat{u};\gamma)$;

(b) $\langle\hat{p}(t), f(t,\hat{x}(t),\hat{u}(t))\rangle - g(t,\hat{x}(t),\hat{u}(t))$

$> \langle\hat{p}(t), f(t,\hat{x}(t),u) - g(t,\hat{x}(t),u)$

for all $t\epsilon[a,b]$ and all $u\epsilon U$ with $u \neq \hat{u}(t)$;

(c) for all $t\epsilon[a,b]$, the function

$Z(t,x,u) = \langle\hat{p}(t)-Q(t)(x-\hat{x}(t)),\ f(t,x,u)\rangle - g(t,x,u)$

is strictly concave in u for $u\epsilon U$ and $(t,x,u)\epsilon T(\hat{x},\hat{u};\gamma)$.

Then $(\hat{x},\hat{u})$ is a strong local minimum for (C).

Proof: Given in reference [25].

Other conditions exist to guarantee that the function $F(t,x,u)$ achieve its maximum at $F(t,\hat{x}(t),\hat{u}(t))$ for x near $\hat{x}(t)$ and $u\epsilon U$; by imposing, for example, first and second order conditions on the function F. A first order result appears in [25]. We pass here to a second order result, from the same reference, expressed in terms of the data f and g.

Suppose we are given a pair $(\hat{x},\hat{u})$ and a positive number ε. The following assumption is required:

f and g and their partial derivatives up to second order with respect to (x,u) exist and are continuous on $T(\hat{x};\varepsilon)xU$. (H_2)

We define the following quantities:

$$R(t) = g_{uu}(t,\hat{x}(t),\hat{u}(t)) - D_u(f_u^T(t,\hat{x}(t),\hat{u}(t))\hat{p}(t)),$$

$$D(t) = g_{xx}(t,\hat{x}(t),\hat{u}(t)) - D_x(f_x^T(t,x,\hat{u}(t))\hat{p}(t))|_{x=\hat{x}(t)}$$

$$A(t) = f_x(t,\hat{x}(t),\hat{u}(t)),$$

$$S(t) = g_{ux}(t,\hat{x}(t),\hat{u}(t)) - D_x(f_u^T(t,x,\hat{u}(t))\hat{p}(t)))|_{x} = \hat{x}(t),$$

and

$$B(t) = f_u(t,\hat{x}(t),\hat{u}(t)).$$

Cor. 3.2. [Zeidan] Let $(\hat{x},\hat{u})$ be an admissible pair for (C). Assume that Hypothesis (H_2) holds for some $\varepsilon > 0$, and

(1) U is convex and compact;

(2) there exists an arc $\hat{p}$ satisfying

$$-\dot{\hat{p}}(t) = f_x(t,\hat{x}(t),\hat{u}(t))^T\hat{p}(t) - g_x(t,\hat{x}(t),\hat{u}(t)) \quad \text{a.e.},$$

with

$$\hat{p}(b) = -l_x(\hat{x}(b));$$

(3) for all $t\in[a,b]$ and for all $u\in U$ with $u\neq\hat{u}(t)$,

$$\langle\hat{p}(t),f(t,\hat{x}(t),\hat{u}(t))\rangle - g(t,\hat{x}(t),\hat{u}(t))$$

$$>\langle\hat{p}(t),f(t,\hat{x}(t),u)\rangle - g(t,\hat{x}(t),u);$$

(4) $R(t) > 0$ for all $t\in[a,b]$;

(5) there exists a Lipschitz symmetric matrix function $Q(t)$ on $[a,b]$

satisfying

$Q(b) < l_{xx}(\hat{x}(b)),$

and, for all $t \in [a,b]$ and all $\eta(t) \in \partial Q(t)$,

$\bar{M}(t,\eta(t)) := \eta(t) + Q(t)A(t) + A^T(t)Q(t) + D(t) - \bar{K}^T(t)R(t)\bar{K}(t) > 0$

where $\bar{K}(t) = R^{-1}(t)(S(t) + B^T(t)Q(t))$.

Then $(\hat{x},\hat{u})$ is a strong local minimum for (C).

Remark: In the classical calculus of variations, condition (5) is seen to reduce to the Jacobi condition (see [25] and [26]).

Remark: Assume $\hat{u}$ is in $L^\infty[a,b]$. By replacing condition (3) of the theorem with the weaker condition

(3') $\langle \hat{p}(t), f(t,\hat{x}(t),\hat{u}(t))\rangle - g(t,\hat{x}(t),\hat{u}(t))$

$\geq \langle \hat{p}(t), f(t,\hat{x}(t),u)\rangle - g(t,\hat{x}(t),u)$

for all $(t,u) \in T(\hat{u};\varepsilon)$

we obtain a suficiency criterion for weak local optimality.

Remark: Let $(\hat{x},\hat{u})$ be a candidate for strong local optimality. Then if the problem is normal the pair $(\hat{x},\hat{u})$ must satisfy conditions (2) and (3') for some arc $\hat{p}$. These conditions together make up the maximum principle (see [15]).

Cor^{s}3.1 and 3.2 provide ways of testing whether the function $F(t,x,u)$ achieves its maximum at $(\hat{x}(t),\hat{u}(t))$ when x is near $\hat{x}(t)$ and u is in U, which is equivalent to conditions (i) and (ii) of Th^{m}3.2. If we now define on $T(\hat{x};\varepsilon)$ the function

$$V(t,x) = H(t,x,\hat{p}(t)-Q(t)(x-\hat{x}(t))) + \langle \dot{\hat{p}}(t), x-\hat{x}(t)\rangle$$
$$+ \langle f(t,\hat{x}(t),\hat{u}(t)), Q(t)(x-\hat{x}(t))\rangle$$
$$- 1/2 \langle x-\hat{x}(t), \dot{Q}(t)(x-\hat{x}(t))\rangle \qquad (3.6)$$

then we may express condition (ii) of Th^{m}3.2 as the condition that

$$V(t,x) \leq V(t,\hat{x}(t))$$

almost everywhere in t with $(t,x)\epsilon T(\hat{x};\varepsilon)$. Again, first and second order conditions exist to guarantee this to be the case. See [25] for the first order conditions; we pass to the second order result, appearing in [26], which applies to the autonomous control problem and is now expressed in terms of the Hamiltonian.

Def^n: Suppose the Hamiltonian H is C^{1+} near a given arc $\hat{z} = (\hat{x},\hat{p})$. We say that the extended Jacobi condition is satisfied at $\hat{z}$ if there exists a Lipschitz symmetric matrix function $Q(\cdot)$ on [a,b] such that

$$\eta(t)-Q(t)\gamma(t)Q(t)+Q(t)\beta(t)+\delta(t)Q(t)-\alpha(t) > 0$$

for all t in [a,b], all matrices

$$\begin{vmatrix} \alpha(t) & \delta(t) \\ \beta(t) & \gamma(t) \end{vmatrix} \epsilon\,\partial\nabla H(\hat{z})(t)$$

and all $\eta(t)\epsilon\partial Q(t)$.

Cor 3.3. [Zeidan] Let $(\hat{x},\hat{u})$ be an admissible pair for the autonomous control problem. Assume that Hypothesis (H_2) holds for some $\varepsilon>0$, and

(a) U is a nonempty convex compact polyhedron in $\mathbb{R}^n$;

(b) there exists an arc $\hat{p}$ from [a,b] to $\mathbb{R}^n$ satisfying

$$-\dot{\hat{p}}(t) = f_x(\hat{x}(t),\hat{u}(t))^T\hat{p}(t) - g_x(\hat{x}(t),\hat{u}(t)) \qquad \text{a.e.},$$

with

$$\hat{p}(b) = -l_x(\hat{x}(b));$$

(c) for all $t\in[a,b]$ and all $u\in U$ such that $u\neq\hat{u}(t)$,

$$\langle\hat{p}(t),f(\hat{x}(t),\hat{u}(t))\rangle - g(\hat{x}(t),\hat{u}(t))$$

$$> \langle\hat{p}(t),f(\hat{x}(t),u)\rangle - g(\hat{x}(t),u);$$

(d) $R(t)>0$ for all $t\in[a,b]$;

(e) the extended Jacobi condition is satisfied by a matrix function Q such that $Q(b)<l_{xx}(\hat{x}(b))$.

Then the pair $(\hat{x},\hat{u})$ provides a strong local minimum for the autonomous control problem.

Remark: In the classical setting, the Hamiltonian H is C^2 and the extended Jacobi condition reduces to the Jacobi condition written in terms of the Hamiltonian. Thus, the extended Jacobi condition and condition (5) of Cor. 3.2 are equivalent when H is C^2, otherwise they differ. The assumption that U is a polyhedron is only needed to have a C^{1+} Hamiltonian. Therefore, when H is given to be C^2, Cors. 3.2 and 3.3 are equivalent. In particular, when $\hat{u}(t)$ is in the interior of the control set U, then H is C^2 and Cor. 3.3 reduces to the result of Mayne [13, Th^{m}3.2].

3.2 New Results.

We now narrow our attention to the problem (C) where the control set U is of the form

$$U = \{u \in \mathbb{R}^m : h(u) = 0, d(u) \leq 0\} \qquad (3.7)$$

for functions

$$h : \mathbb{R}^m \to \mathbb{R}^p \text{ and } d : \mathbb{R}^m \to \mathbb{R}^q$$

We present new second order sufficiency criteria for weak and strong local optimality involving, like the criterion of Cor. 3.2, an equality expressed in terms of the data f and g which reduces to the Jacobi condition in the classical case. It is seen that when the control set U is polyhedral our results generalize Cors. 3.2 and 3.3. Moreover, we give an example to which our result for weak local optimality applies while the known one does not.

Let $(\hat{x},\hat{u})$ be a given pair. We will have call for the following hypotheses:

There exists a positive number ε such that f and g and their partial derivatives up to second order with respect to (x,u) exist and are continuous on $T((\hat{x},\hat{u});\varepsilon)$; h and d are C^2 on $N(\hat{u};\varepsilon)$; and l is C^2 on the ε - neighbourhood of $\hat{x}(b)$. (H_3)

There exists $\varepsilon>0$ such that f and g and their partial derivatives up to second order with respect to (x,u) exist and are continuous on $T(\hat{x};\varepsilon)xU$; h and d are C^2 on U, and l is C^2 on the ε - neighbourhood of $\hat{x}(b)$. (H_4)

Let $(\hat{x},\hat{u})$ be a candidate for weak (resp. strong) local optimality. Then if the problem is normal the pair $(\hat{x},\hat{u})$ must satisfy the maximum principle:

There exists an arc $\hat{p}:[a,b]\to\mathbb{R}^n$ such that

$$-\dot{\hat{p}}(t) = f_x(t,\hat{x}(t),\hat{u}(t))\hat{p}(t) - g_x(t,\hat{x}(t),\hat{u}(t)) \quad \text{a.e.} \tag{3.8}$$

$$\hat{p}(b) = -\nabla\ell(\hat{x}(b));$$

$$\langle\hat{p}(t),f(t,\hat{x}(t),\hat{u}(t))\rangle - g(t,\hat{x}(t),\hat{u}(t))$$

$$\geq \langle p(t),f(t,\hat{x}(t),u)\rangle-g(t,\hat{x}(t),u) \tag{3.9}$$

for almost all $t\in[a,b]$ and all $u\in N(\hat{x};\varepsilon)\cap U$ (respectively all $u\in U$).

If $\hat{u}(t)$ is a regular point for the constraints (3.7), by [11] inequality (3.9) implies the existence of Lagrange multipliers $\lambda(t)\in{}^p$ and $\mu(t)\in\mathbb{R}^q$ such that $\mu(t)\leq 0$ and

$$f_u^T(t,\hat{x}(t),\hat{u}(t))\hat{p}(t) - g_u(t,\hat{x}(t),\hat{u}(t)) \tag{3.10}$$

$$+\lambda^T(t)\nabla h(\hat{u}(t))+\mu^T(t)\nabla d(\hat{u}(t)) = 0;$$

$$\mu^T(t)d(\hat{u}(t)) = 0;$$

and

$$L(t) = g_{uu}(t,\hat{x}(t),\hat{u}(t))-D_u[f_u^T(t,\hat{x}(t),u)\hat{p}(t)]\big|_{u=\hat{u}(t)}$$

$$-\lambda^T(t)\nabla^2 h(\hat{u}(t))-\mu^T(t)\nabla^2 d(\hat{u}(t)) \tag{3.11}$$

is positive semidefinite on the tangent subspace of the active constraints at $\hat{u}(t)$.

For $t\in[a,b]$, define the subspace

$$T(t) = \{u \in \mathbb{R}^m : \nabla h(\hat{u}(t))u = 0, \mu_j(t)\nabla d_j(\hat{u}(t))u = 0 \quad j = 1,\dots,q\}. \tag{3.12}$$

When the Lagrange multipliers associated with active constraints are all non-zero, $T(t)$ coincides with the tangent subspace of the active constraints at $\hat{u}(t)$.

Let $Y(t)$ be a matrix whose columns form a basis for $T(t)$.

Suppose $(\hat{x},\hat{u})$ is a pair for which there exist an arc $\hat{p}$ and Lagrange multipliers λ and μ satisfying (3.10) and, for $t \in [a,b]$ a.e.,

$$Y^T(t)L(t)Y(t) > 0 \quad \text{or} \quad Y(t) = 0.$$

We say that the <u>modified Jacobi condition</u> is satisfied if there exists an absolutely continuous symmetric matrix function $Q(t)$ with $Q(b) < \nabla^2 l(\hat{x}(b))$ and, for $t \in [a,b]$ a.e.,

$$\begin{aligned} M(t,\dot{Q}(t)) := \dot{Q}(t) &+ Q(t)A(t) + A^T(t)Q(t) \\ &+ D(t) - K^T(t)L(t)K(t) > 0 \end{aligned} \tag{3.13}$$

where

$$K(t) = Z(t)[S(t) + B^T(t)Q(t)] \tag{3.14}$$

and

$$Z(t) = \begin{cases} Y(t)[Y^T(t)L(t)Y(t)]^{-1}Y^T(t) & \text{if } Y(t) \neq 0 \\ 0 & \text{if } Y(t) = 0 \end{cases} \tag{3.15}$$

The mxm - matrix $Z(t)$ is the generalized inverse of $L(t)$.

We note that when $R(t)$ (as opposed to $Y^T(t)L(t)Y(t)$) is assumed to be positive definite and $Z(t)$ in (3.15) is replaced by $R^{-1}(t)$ we obtain the known Jacobi condition of Cor. 3.2.

Th^{m}3.3: Let $(\hat{x},\hat{u})$ with $\hat{u}\in L^\infty[a,b]$ be an admissible pair for (C) satisfying (H_3). Assume that, for an arc $\hat{p}$ and for integrable multipliers $\lambda(t),\mu(t)\leq 0$, equations (3.8) and (3.10) hold for $t\in[a,b]$ a.e., and

(1) for $t\in[a,b]$ a.e., $Y^T(t)L(t)Y(t)>0$ or $Y(t) = 0$;

(2) the modified Jacobi condition is satisfied for some $Q(t)$;

(3) U is convex.

Then $(\hat{x},\hat{u})$ is a weak local minimum for (C).

Remark: Equations (3.8) and (3.10) are necessary for optimality. Condition (1) is a strengthening of the positive semi-definiteness of $L(t)$ on the tangent subspace of the active constraints at $\hat{u}(t)$. Note that we do not require $R(t)$ to be positive definite.

Th^{m}3.4: Let $(\hat{x},\hat{u})$ with $\hat{u}\in L^\infty[a,b]$ be admissible for (C) and let (H_4) be satisfied. Assume that there exists an arc $\hat{p}$ and integrable multipliers $\lambda(t)$ and $\mu(t)\leq 0$ such that (3.8), (3.10) and conditions (1) and (2) of Th^{m}3.3 hold, and

(i) U is convex and compact;

(ii) $R(t)>0$ for all $t\in[a,b]$;

(iii) $t\in[a,b]$ and $\forall$ $u\in U$ with $u\neq\hat{u}(t)$,

$$\langle\hat{p}(t),f(t,\hat{x}(t),\hat{u}(t))\rangle - g(t,\hat{x}(t),\hat{u}(t))$$

$$> \langle\hat{p}(t),f(t,\hat{x}(t),u)\rangle - g(t,\hat{x}(t),u).$$

Then $(\hat{x},\hat{u})$ is a strong local minimum for (C).

Remark: Condition (iii) is a strengthening of the necessary condition given by (3.9) and it implies that equation (3.10) holds for some $\lambda(t)$ and $\mu(t)$.

Remark: Consider the case when $\hat{u}(t)$ is in the interior of the control set ($h\equiv 0, d(\hat{u}(t))<0$). Then $L(t)\equiv R(t)$, $T(t)\equiv\mathbb{R}^m$ and (3.10) implies $\mu(t)\equiv 0$, so our criteria for weak and strong local optimality, Theorems 3.3. and 3.4, coincide with the known Jacobi sufficiency criterion Cor. 3.2, and its weaker version in which condition (3) is replaced by (3').

Remark: When $U=\mathbb{R}^m$, Th^{m}3.4 cannot be directly applied because condition (i) demands that U be compact. However Th^{m}3.4 may be applied for $u\in V_X$, where

$$V_X = \{u=(u_1,\ldots,u_m):d_X(u_1,\ldots,u_m) = \sum_{i=1}^m u_i^2 - X\leq 0\}.$$

If X is so large that $\hat{u}(t)$ is in the interior of V_X, Th^{m}3.4 reduces to Cor.3.2. Since optimality over $u\in\mathbb{R}^m$ is equivalent to optimality over $u\in V_X$ for all large X, we deduce that in this case Th^{m}3.4 again reduces to Cor. 3.2, where the requirement that U be compact is omitted.

The proofs of Th^{m}s. 3.3 and 3.4 are based on the following result, which is a consequence of the Remark following Th^{m}3.2.

Lemma 3.1. Let $(\hat{x},\hat{u})$ be admissible for (C) and let l be C^2 near $\hat{x}(b)$. Assume that there exist a positive number ε, an arc $\hat{p}$ and a Lipschitz symmetric matrix function $Q(t)$ satisfying

(a) $\hat{p}(b) = -\nabla l(\hat{x}(b))$, $Q(b) < \nabla^2 l(\hat{x}(b))$;

(b) $F(t,x(t),u(t)) \leq F(t,\hat{x}(t),\hat{u}(t))\ \forall (x,u)$ such that $u \in U, (t,x,u) \in T(\hat{x},\hat{u};\varepsilon)$, and for $t \in [a,b]$ a.e.

Then $(\hat{x},\hat{u})$ is a weak local minimum for (C). If in addition condition (b) is satisfied for all $u(t) \in U$, then $(\hat{x},\hat{u})$ is a strong local minimum.

Proof of Theorem 3.3.: The proof is done by showing that the conditions of Lemma 3.1 are satisfied. Let $\lambda(t),\mu(t)$ and $Q(t)$ be the functions given in Th^{m}3.3. Then equations (3.8) and (3.10) yield, for $t \in [a,b]$ a.e.,

$$\nabla_{x,u}F(t,\hat{x}(t),\hat{u}(t)) + (0,\lambda^T(t))\nabla_{x,u}h(\hat{u}(t)) + (0,\mu^T(t))\nabla_{x,u}d(\hat{u}(t)) = 0, \qquad (3.16)$$

where $F(t,x,u)$ is given by (3.4).

Let $T(t)$ be defined by (3.12) and let $Y(t)$ be its basis matrix.

We will first show that, for $t\epsilon[a,b]$ a.e.,

$$\bar{L}(t) = -\nabla^2_{x,u}F(t,\hat{x}(t),\hat{u}(t))-(0,\lambda^T(t))\nabla^2_{x,u}h(\hat{u}(t))$$
$$- (0,\mu^T(t))\nabla^2_{x,u}d(\hat{u}(t))$$

is positive definite on the subspace

$$\bar{T}(t) = \left\{\binom{x}{u}:x\epsilon\ \mathbb{R}^n,\ u\epsilon T(t)\right\}.$$

This is equivalent to showing

$$\bar{Y}^T(t)\bar{L}(t)\bar{Y}(t)>0 \text{ for } t\epsilon[a,b] \text{ a.e.,}$$

where

$$\bar{Y}(t) = \begin{bmatrix} I & 0 \\ 0 & Y(t) \end{bmatrix}.$$

Suppose that $Y(t) \neq 0$ on a set of positive measure. Then on that set we find

$$\bar{Y}^T(t)\bar{L}(t)\bar{Y}(t) = M(t,\dot{Q}(t))>0 \ .$$

Suppose that $Y(t) = 0$ on a set of positive measure. For such t, define

$$N(t) = \begin{bmatrix} M(t,\dot{Q}(t)) & 0 \\ 0 & L(t) \end{bmatrix}$$

and

$$C(t) = \begin{bmatrix} I & 0 \\ K(t) & I \end{bmatrix},$$

where $M(t,\dot{Q}(t))$ and $K(t)$ are as defined in (3.13) and (3.15). From conditions (1) and (2) of $\text{Th}^{m}3.3$ we have

$$\bar{Y}^T(t)N(t)\bar{Y}(t)>0$$

and hence

$$C^T(t)\bar{Y}^T(t)N(t)\bar{Y}(t)C(t) = \bar{Y}^T(t)\bar{L}(t)\bar{Y}(t)>0 \ .$$

Thus for $t\epsilon[a,b]$ a.e.

$\bar{L}(t)>0$ on the subspace $\bar{T}(t)$. (3.17)

The proof now proceeds by contradiction; suppose that condition (b) of Lemma 4.1 is not satisfied. Then there exist a sequence A_k of subsets of $[a,b]$ with positive measure and functions $(\tilde{x}_k,\tilde{u}_k), \tilde{x}_k:A_k \to \mathbb{R}^n$ and $u_k:A_k \to U$, such that:

$\forall t\epsilon A_k$, $(t,\tilde{x}_k(t),\tilde{u}_k(t))\epsilon T(\hat{x},\hat{u};1/k)$ and $F(t,\tilde{x}_k(t),\tilde{u}_k(t))>$
$F(t,\hat{x}(t),\hat{u}(t))$.

$$x_k(t) = \begin{cases} \tilde{x}_k(t) & \text{for } t\epsilon A_k \\ \hat{x}(t) & \text{for } t\epsilon A_k \end{cases} ,$$

and

$$u_k(t) = \begin{cases} \tilde{u}_k(t) & \text{for } t\epsilon A_k \\ \hat{u}_k(t) & \text{for } t\epsilon A_k \end{cases} .$$

Then, $\forall t\epsilon[a,b]$, we have $(x_k,u_k)\epsilon T(\hat{x},\hat{u};1/k), u_k(t)\epsilon U$, and

$$\int_a^b F(t,x_k(t),u_k(t))dt > \int_a^b F(t,\hat{x}(t),\hat{u}(t))dt \ . \tag{3.18}$$

Since $F(t,x_k(t),u_k(t)) =$

$$\langle\hat{p}(t)-Q(t)(x_k(t)-\hat{x}(t)),f(t,x_k(t),u_k(t))\rangle - g(t,x_k(t),u_k(t))$$

$$-\frac{d}{dt}\,1/2\langle x_k(t)-\hat{x}(t),Q(t)(x_k(t)-\hat{x}(t))\rangle \quad \text{a.e. } t$$

and since

$$||u_k - \hat{u}||_\infty < 1/k \ ,$$

we have that $u_k \epsilon L^\infty[a,b]$ and $\int_a^b F(t,x_k(t),u_k(t))dt$ is finite.

Now, for each k write

$(x_k(t),u_k(t)) = (\hat{x}(t),\hat{u}(t)) + \delta_k(M_k(t),N_k(t))$

where $\delta_k = ||(x_k-\hat{x},u_k-\hat{u})||_\infty$,

$$M_k(t) = \frac{x_k(t)-\hat{x}(t)}{\delta_k} \text{ and } N_k(t) = \frac{u_k(t)-\hat{u}(t)}{\delta_k} .$$

Then $\delta_k > 0$, $M_k(\cdot) \epsilon L^\infty[a,b]$, $N_k(\cdot)\epsilon L^\infty[a,b]$ and $||(M_k,N_k)||_\infty = 1$. There is a subsequence of (M_k,N_k), which we do not relable, converging to a function $S_o = (M_o,N_o)$ with $||(M_o,N_o)||_\infty = 1$. We first show that $S_o(t)\epsilon\overline{T}(t)$ a.e., i.e. $N_o(t)\epsilon T(t)$ a.e. We have

$0 = h(u_k(t))-h(\hat{u}(t)) = \nabla h(\overline{u}_k(t))\delta_k N_k(t)$ a.e. t, where $\overline{u}_k(t)$ is between $u_k(t)$ and $\hat{u}(t)$ and $||\overline{u}_k(t)-\hat{u}(t)||_\infty \leq ||u_k(t)-\hat{u}(t)||_\infty < 1/k \ \forall k$.

Therefore $\overline{u}_k \to \hat{u}$ in $L^\infty[a,b]$. Since $\delta_k > 0$, we get at the limit that

$0 = \nabla h(\hat{u}(t))N_o(t)$.

Let j be any index of $d = (d_j)$. For $t\epsilon[a,b]$ a.e.

$0 \leq \mu_j(t)[d_j(u_k(t))-d_j(\hat{u}(t))] = \delta_k\mu_j(t)\nabla d_j(\overline{\overline{u}}_k(t))\cdot N_k(t)$

where $\overline{\overline{u}}_k(t)$ is between $u_k(t)$ and $\hat{u}(t)$. Again we have $\overline{\overline{u}}_k \to \hat{u}$ in $L^\infty[a,b]$, and since $\delta_k > 0$ we get at the limit that $0 \leq \mu_j(t)\nabla d_j(\hat{u}(t))\cdot N_o(t)$ a.e. t, all j. It remains to show that for all j

$\mu_j(t)\nabla d_j(\hat{u}(t))\cdot N_o(t) = 0$ a.e. t .

If not, there exist an index j and a subset $E\subset[a,b]$ of positive Lebesgue measure such that:

$\mu_j(t)\nabla d_j(\hat{u}(t))\cdot N_o(t)>0$ for all $t\in E$.

Since $N_o,\hat{u}$ are $L^\infty[a,b]$, $\mu_j(t)$ is integrable and d_j is C^2 we have

$\int_a^b \mu_j(t)\nabla d_j(\hat{u}(t)).N_o(t)dt>0$ and so $\int_a^b \mu_j(t)\nabla d(\hat{u}(t))\cdot N_o(t)dt>0$.

Using (3.16) we obtain

$$\int_a^b \nabla_{x,u}F(t,\hat{x}(t),\hat{u}(t))\cdot(M_o(t),N_o(t))dt<0 \ . \tag{3.19}$$

But by (3.18) we have for all k that

$$0 < \int_a^b \{F(t,x_k(t),u_k(t))-F(t,\hat{x}(t),\hat{u}(t))\}dt$$

$$= \int_a^b \nabla_{x,u}F(t,\bar{\bar{x}}_k(t),\bar{\bar{u}}_k(t))\cdot\delta_k(M_k(t),N_k(t))dt,$$

for $(\bar{\bar{x}}_k,\bar{\bar{u}}_k)$ between (x_k,u_k) and $(\hat{x},\hat{u})$. Since $(\bar{\bar{x}}_k,\bar{\bar{u}}_k)\to(\hat{x},\hat{u})$ in $L^\infty[a,b]$ and $\delta_k>0$, at the limit we get (here we use the Lebesgue dominated convergence theorem)

$$0\leq \int_a^b \nabla_{x,u}F(t,\hat{x}(t),\hat{u}(t))\cdot(M_o(t),N_o(t))dt,$$

which contradicts (3.19).

It therefore follows that $N_o(t)\in T(t)$ a.e. and hence $(M_o(t),N_o(t))\in\overline{T}(t)$ a.e.

Now, we want to obtain a contradiction of our assumption (3.18). For all k the following statments are true:

$0 = \lambda(t)[h(u_k(t))-h(\hat{u}(t))]$

$= \lambda(t)\nabla h(\hat{u}(t))\delta_k N_k(t) + \frac{\lambda(t)}{2}\,\delta_k^2 N_k^T(t)\nabla^2 h(u_k'(t))N_k(t)$ a.e.

for u_k' between u_k and $\hat{u}$, $u_k' \to \hat{u}$ in L^∞ ;

$0 \leq \mu(t)[d(u_k(t)) - d(\hat{u}(t))]$

$= \mu(t)\cdot\nabla d(\hat{u}(t)\delta_k N_k(t) + \frac{\mu(t)}{2}\,\delta_k^2 N_k^T(t)\nabla^2 d(u_k''(t))N_k(t)$ a.e.

where $u_k'' \to \hat{u}$ in L^∞ ; and,

$$0 < \int_a^b \{F(t,x_k(t),u_k(t)) - F(t,\hat{x}(t),\hat{u}(t))\}\,dt$$

$$= \int_a^b \{\nabla_{x,u}F(t,\hat{x}(t),\hat{u}(t))\delta_k(M_k(t),N_k(t))$$

$$+ 1/2\ \delta_k^2(M_k(t),N_k(t))\nabla_{x,u}^2 F(t,x_k\ (t),u_k\ (t))\begin{bmatrix} M_k\ (t) \\ N_k\ (t)\end{bmatrix}dt$$

where $x_k \to \hat{x}$ and $u_k \to \hat{u}$ in L^∞..

Using the integrability of λ and μ and the fact that $\hat{u}$ and u_k are $L^\infty[a,b]$, we integrate the first two expressions and then add them to the third to obtain:

$$0 < \int_a^b (M_k(t),N_k(t))[\nabla_{x,u}^2 F(t,x_k\ (t),\mu_k\ (t)) +$$

$$(0,\lambda^T(t))\nabla_{x,u}^2 h(u_k'(t)) + (0,\mu^T(t))\nabla_{x,u}^2 d(u_k''(t))\begin{bmatrix} M_k\ (t) \\ N_k\ (t)\end{bmatrix}dt.$$

By the Lebesgue dominated convergence theorem we get

$$0 \leq \int_a^b (M_o(t),N_o(t))[\nabla_{x,u}^2 F(t,\hat{x}(t),\hat{u}(t)) + (0,\lambda^T(t))\nabla_{x,u}^2 h(\hat{u}(t))$$

$$+(0,\mu^T(t))\nabla_{x,u}^2 d(\hat{u}(t))]\begin{bmatrix} M_o(t) \\ N_o(t)\end{bmatrix}\ dt$$

But $(M_o(t),N_o(t))\in\overline{T}(t)$ a.e., and we know that

$$(M_o(t),N_o(t))\bar{L}(t)\begin{bmatrix} M_o(t) \\ N_o(t) \end{bmatrix} > 0 \quad \underline{a.e.}\ t.$$

Thus we have a contradiction, and the theorem is proved.

Proof of $Th^m 3.4$: From the proof of $Th^m 3.3$, we know that condition (b) of Lemma 4.1 holds for $(t,x(t),u(t))\epsilon T(\hat{x},\hat{u};\varepsilon)$ with $u(t)\epsilon U$. Since $R(t)>0$ for all $t\epsilon[a,b]$, there exists a positive number δ $(\delta\leq\varepsilon)$ such that $Z(t,x,u) = \langle\hat{p}(t)-Q(t)(x-\hat{x}(t)),f(t,x,u)\rangle-g(t,x,u)$ is strictly concave in u for $(t,x,u)\epsilon T(\hat{x},\hat{u};\delta)$ with $u\epsilon U$. By [21], this fact together with condition (iii) implies that there exists $\alpha > 0$ $(\alpha \leq \delta)$ such that

$\forall(t,x)\epsilon T(\hat{x};\alpha)$ and $u\epsilon U$,

$$F(t,x,u) \leq F(t,x,u(t,x)) \quad \text{a.e.},$$

where $(t,x,u(t,x))\epsilon T(\hat{x},\hat{u};\varepsilon)$.

Therefore, given a pair (x,u) with $(t,x(t),u(t))\epsilon T(\hat{x};\alpha)xU$, we have

$$F(t,x(t),u(t)) \leq F(t,x(t),u(t,x(t)))$$
$$\leq F(t,\hat{x}(t),\hat{u}(t)) .$$

The result then folows by Lemma 3.1.

Remark: When in the problem (C) the value x(b) is prescribed, the conditions $\hat{p}(b) = -\nabla l(\hat{x}(b))$ and $Q(b)<\nabla^2 l(\hat{x}(b))$ are omitted from $Th^m s$ 3.3 and 3.4.

Remark: The conditions of $Th^m 3.4$ imply, by Lemma 3.1, that condition

(a) of Cor. 3.1 is satisfied. Therefore $Th^m 3.4$ may be viewed as a corollary of Cor. 3.1.

3.3 The Polyhedral Case.

We will now examine the specific case in which the control set U is polyhedral. This is the case considered in Cor. 3.3. We will show here that $Th^m 3.4$ generalizes both the Jacobi criterion expressed in terms of the data (Cor. 3.2) and the one expressed in terms of the Hamiltonian (Cor. 3.3), in which Q is assumed ~~to~~ be C^1. We will also show that $Th^m 3.3$ generalizes the weak version of Cor. 3.2.

Important remark: since the functions h and d in (3.7) are affine, (3.11) yields that $L(t) = R(t)$.

Thus, condition (1) of $Th^m 3.3$, which is now

$$\dot{Y}^T(t)R(t)Y(t) > 0 \qquad \text{or} \qquad Y(t) = 0,$$

can be omitted from $Th^m 3.4$, since it is implied by the assumption $R(t) > 0$.

Prop^{n}3.2: Assume that U is a polyhedron. Then the weak version of Cor. 3.2 (in which condition (3) is replaced by condition (3') or (3.9)) is a special case of $Th^m 3.3$.

Proof: Inequality (3.9) yields that, for some Lagrange multipliers $\lambda_o(t)$ and $\mu_o(t)$, equations (3.10) hold. To $\lambda_o(t)$ and $\mu_o(t)$ there correspond via (3.12) a subspace $T_o(t)$ and its basis matrix $Y_o(t)$.

Condition (4) of Cor. 3.2 implies condition (1) of $Th^m 3.3$, and

$$R^{-1}(t) \geq Y_o[Y_o^T(t)R(t)Y_o(t)]^{-1}Y_o^T(t) \quad t\epsilon[a,b] \qquad (3.20).$$

Therefore the matrix function Q(t) in condition (5) of Cor.3.2 satisfies the modified Jacobi condition (3.13).

To see (3.20) fix t and let $w \neq 0$ be an eigenvector of the matrix

$$A = R^{-1}(t) - Y_o(t)[Y_o^T(t)R(t)Y_o(t)]^{-1}Y_o^T(t)$$

and let λ be the eigenvalue corresponding to w, so that

$$Aw = \lambda w \qquad (3.21)$$

Assume $\lambda \neq 0$. Since $Y_o^T(t)R(t)\ A = 0$, we have by (3.21) that

$$Y_o(t)^T R(t)w = 0 \qquad (3.22)$$

By (3.21) we have $\lambda w^T = w^T A$, and so $\lambda w^T(R(t)w) = w^T A(R(t)w)$
$= w^T(w - Y_o(t)[Y_o^T(t)R(t)Y_o(t)]^{-1}Y_o^T(t)R(t)w) = w^T w$ by (3.22).

The positive definiteness of R(t) then yields that λ is positive, and therefore eigenvalues of A are greater than or equal to zero and A is positive semidefinite, as required.

<u>Prop^{n}3.3</u>: Assume that U is a polyhedron. Then Cor. 3.2 is a special case of $Th^m 3.4$.

<u>Proof</u>: Since Cor. 3.2 differs from its weaker version only in

condition (3), and since this condition is duplicated in $Th^m 3.4$, it is clear that the conditions of Cor. 3.2 imply those of $Th^m 3.4$.

The next proposition shows how Cor. 3.3 (which incorporates the Jacobi condition expressed in terms of the Hamiltonian) relates to $Th^m 3.4$ when the problem is autonomous.

<u>$Prop^n 3.4$:</u> Assume that the Lagrange multipliers associated with active constraints are all non-zero, and that $\hat{u}$ is a regular point for the constraints. Assume also that f and g are independent of t and U is polyhedral. If the function Q(t) satisfying the extended Jacobi conditon in Cor. 3.2 is assumed to be C^1, then Cor. 3.3 is a special case of $Th^m 3.4$.

<u>Proof</u>: It was mentioned (see 3.6) that in the proof of Cor. 3.3, which appears in [26], it is shown that the conditions of Cor. 3.3 imply that for some positive ε, and all $x \epsilon T_{\varepsilon}(\hat{x};\varepsilon)$,

$$V(t,x) \leq V(t,\hat{x}(t)),$$

where

$$\begin{aligned} V(t,x) = {} & H(t,x,\hat{p}(t)-Q(t)(x-\hat{x}(t))) + \langle\dot{\hat{p}}(t),x-\hat{x}(t)\rangle \\ & + \langle f(t,\hat{x}(t),\hat{u}(t)),Q(t)(x-\hat{x}(t))\rangle \\ & - 1/2 \langle x-\hat{x}(t),\dot{Q}(t)(x-\hat{x}(t))\rangle . \end{aligned}$$

By the definition of the Hamiltonian, this is equivalent to saying that, for all $(t,x,\hat{u})\in T(\hat{x};\varepsilon)\times U$

$$F(t,x,u) \leq F(t,\hat{x}(t),\hat{u}(t))$$

where, as in (3.4),

$$F(t,x,u) = \langle\hat{p}(t)-Q(t)(x-\hat{x}(t)),f(t,x,u)\rangle$$
$$-g(t,x,u)-1/2\ \langle x-\hat{x}(t),\dot{Q}(t)(x-\hat{x}(t))\rangle$$
$$+\langle\dot{\hat{p}}(t),x-\hat{x}(t)\rangle + \langle\dot{\hat{x}}(t),Q(t)(x-\hat{x}(t))\rangle .$$

Thus, the theory of Lagrange multipliers ([11],pg 316) and the affine property of U yield the existence of $\lambda(t),\mu(t)\leq 0$ such that (3.10) holds and

$$\overline{R}(t) := -\nabla^2_{x,u}F(t,\hat{x}(t),\hat{u}(t))$$

is positive semidefinite on

$\overline{T}(t) = \left\{\binom{x}{u}:x\in\mathbb{R}^n,u\in T(t)\right\}$, where $T(t)$

is the tangent subspace of the active constraints at $\hat{u}(t)$ and is equal to T as defined in (3.12). Let $Y(t)$ be the basis matrix of $T(t)$, and define

$$\overline{Y}(t) = \begin{bmatrix} I & 0 \\ 0 & Y(t) \end{bmatrix},$$

where I is the nxn identity matrix. The above then implies that

$$\overline{Y}^T(t)\overline{R}(t)\overline{Y}(t) = \begin{bmatrix} \dot{Q}(t)+Q(t)A(t) +A^T(t)Q(t)+B(t) & \{S^T(t)+Q(t))B(t)\}Y(t) \\ Y^T(t)\{S(t)+B^T(t)Q(t)\} & Y^T(t)R(t)Y(t) \end{bmatrix} \geq 0$$

Define

$$N(t) = \begin{bmatrix} M(t,\dot{Q}(t)) & 0 \\ 0 & R(t) \end{bmatrix}$$

where $M(t,\dot{Q}(t))$ is defined by (3.13). If $Y(t) = 0$ then from the above we have

$$M(t,\dot{Q}(t)) \geq 0 .$$

If $Y(t)=0$, define

$$C(t) = \begin{bmatrix} I & 0 \\ K(t) & I \end{bmatrix},$$

where $K(t)$ is defined by (3.14) and (3.15).

After computation we get

$$[C^T(t)]^{-1}\bar{Y}^T(t)\bar{R}(t)\bar{Y}(t)[C(t)]^{-1} = \bar{Y}^T(t)N(t)\bar{Y}(t) \geq 0,$$

from which it follows that $M(t,\dot{Q}(t)) \geq 0$.

We conclude that for all $t \in [a,b]$

$$M(t,\dot{Q}(t)) \geq 0 .$$

Using the embedding theorem of differential equations (see [9]), we can find a function $Q(t)$ satisfying $M(t,\dot{Q}(t)) > 0$ for all $t \in [a,b]$.

We now present a numerical example illustrating the utility of the assumption (1) of Th^{m}3.3 as compared to the usual assumption: $R(t)>0$ for all $t \in [a,b]$. In this example we will see that the weak version of

Cor. 3.2 cannot be used and $\mathrm{Th}^{\mathrm{m}}3.3$ is the only criterion which can be applied to conclude weak optimality of our candidate.

Consider the optimal control problem (C)

minimize

$$\int_0^1 (u_1^3-(1/2)u_2^2-(1/8)xu_1)dt$$

subject to $\dot{x} = u_1-(1/8)x, x(0) = x(1) = 0$

$$(u_1(t),u_2(t))\epsilon U = \{(u_1,u_2):(1/2)u_1^2+(1/2)u_2^2 \leq 2\} .$$

Take the candidate $\hat{x} \equiv 0$, $\hat{u} = (0,2)$, and let $\hat{p} \equiv 0$. Then $(\hat{x},\hat{u})$ and $\hat{p}$ satisfy (3.8) and (3.10) for $\mu(t) \equiv -1$.

The matrices

$R(t) \equiv \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$ and $L(t) \equiv \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ are not positive definite. Thus, we cannot replace the conditions of $\mathrm{Th}^{\mathrm{m}}3.3$ by those of Cor. 3.2. However, from (3.12) we have $T(t) \equiv \{\binom{u}{o}:u\epsilon \mathbb{R}\}$, and so $Y(t) \equiv \binom{1}{0}$, for which $Y^T(t)L(t)Y(t) = 1 > 0$.

From (3.13) we obtain

$$\begin{aligned} M(t,\dot{Q}(t)) &= \dot{Q}(t)-(1/4)Q(t)-(Q(t)-1/8)^2 \\ &= \dot{Q}(t)-Q^2(t)-1/64. \end{aligned}$$

Taking $Q(t) = (1/2)t$, we have for all $t\epsilon[0,1]$

$$\dot{Q}(t)-Q^2(t)-1/64 = 1/2-(1/4)t^2-1/64 > 0 .$$

Therefore, by $\mathrm{Th}^{\mathrm{m}}3.3$, $\hat{x} \equiv 0$ and $\hat{u} = (0,2)$ provide a weak local minimum for (C).

4. THE CALCULUS OF VARIATIONS

The calculus of variations is the least general of the three problems. Results developed for the generalized problem of Bolza or optimal control may be applied. In this section we state the zero order condition of [25] and give a second order result that improves on that of [25].

Let $L:[a,b]\times\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be given. The calculus of variations problem is:

$$\text{Min } J(x) = \int_a^b L(t,x(t),\dot{x}(t))dt \qquad\qquad (V)$$

subject to

$x(a) = A, x(b) = B$

over all absolutely continuous $x:[a,b]\to\mathbb{R}^n$.

The Hamiltonian corresponding to (V) is given by the conjugacy formula

$$H(t,x,p) = \sup\{\langle p,v\rangle - L(t,x,v) : v\in\mathbb{R}^n\}.$$

Def^n: An absolutely continuous function $x:[a,b]\to\mathbb{R}^n$ is admissible for (V) if $x(a) = A$ and $x(b) = B$.

We present the proof of the next result so as to see how it follows from Th^{m}3.2. For a given arc x, L(t) designates $L(t,\hat{x}(t),\dot{\hat{x}}(t))$.

Th^{m}4.1. [Zeidan] Let $\hat{x}$ be admissible for (V). Assume that $L(\cdot,\cdot,\cdot)$ is C^1, and that there exist $\varepsilon > 0$ and an absolutely continuous symmetric matrix function Q(t) such that

(1) there exists a constant $c \in \mathbb{R}^n$ satisfying $\hat{L}_v(t) = \int_a^b \hat{L}_x(s)\, ds + c$ a.e.;

(2) for almost all t, for x with $(t,x) \in T(\hat{x};\varepsilon)$ and for all $v \in \mathbb{R}^n$,

$$L(t,x,v)-\hat{L}(t) \geq \langle \hat{L}_x(t), x-\hat{x}(t)\rangle + \langle \hat{L}_v(t)-Q(t)(x-\hat{x}(t)),\ v-\dot{\hat{x}}(t)\rangle$$
$$-1/2\ \langle x-\hat{x}(t), \dot{Q}(t)(x-\hat{x}(t))\rangle .$$

Then $\hat{x}$ is a strong local minimum for (V). If $\varepsilon = +\infty$, then $\hat{x}$ is a global minimum.

Proof: It suffices to show that conditions (i) and (ii) of Th^{m}3.2 are satisfied, where the functions f and g are as given in (V). Letting $\hat{p}(t) = \hat{L}_v(t), f(t,x,v) = v$ and $g(t,x,v) = L(t,x,v)$, condition (2) becomes

$$g(t,x,v)-g(t,\hat{x}(t),\dot{\hat{x}}(t)) \geq \langle \dot{\hat{p}}(t), x-\hat{x}(t)\rangle$$
$$+\langle \hat{p}(t)-Q(t)(x-\hat{x}(t)), f(t,x,v)-\dot{\hat{x}}(t)\rangle$$

$-1/2\ \langle x-\hat{x}(t), \dot{Q}(t)(x-\hat{x}(t))\rangle$.

for $t\in[a,b]$ a.e. and all $v\in\mathbb{R}^n$. Rearrangement then shows that

$$F(t,x,v) \leq F(t,\hat{x}(t),\dot{\hat{x}}(t))$$

for $t\in[a,b]$ a.e. with $(t,x)\in T(\hat{x};\varepsilon)$, and for all $v\in\mathbb{R}^n$, where the function $F(t,x,v)$ is as defined in (3.4). This guarantees condition (ii) of Th^{m}3.2. Setting $x = \hat{x}(t)$ in the above gives condition (i).

Remark: If $L(t,\cdot,\cdot)$ is convex, then $Q(t) \equiv 0$ satisfies condition (2) of Th^{m}4.1.

Cor.4.1: Let $\hat{x}$ be smooth and admissible for (V) and let $L(t,x,v)$ be C^1 in (t,x) and C^2 in v. Suppose that there exist $\varepsilon > 0$ and an absolutely symmetric matrix function $Q(t)$ such that condition (1) of Th^{m}4.1 holds and

(i) $\mathcal{L}_{vv}(t) > 0$

(ii) $L(t,\hat{x}(t),v)-\mathcal{L}(t) > \langle\mathcal{L}_v(t), v-\dot{\hat{x}}(t)\rangle$, all $v\in\mathbb{R}^n$ with $v \neq \dot{\hat{x}}(t)$, and all $t\in[a,b]$

(iii) condition (2) of Th^{m}4.1 holds for $(t,x,v)\in T(\hat{x},\dot{\hat{x}};\varepsilon)$.

Then $\hat{x}$ is a strong local minimum for (V).

Proof: Since Th^{m}3.2 reduces in this problem to Th^{m}4.1, and since Cor. 3.1 follows from Th^{m}3.2, we may prove that the above result is a corollary of Th^{m}4.1 by showing that it follows from Cor. 3.1. Define $f(t,x,v) = v$, $g(t,x,v) = L(t,x,v)$ and $\hat{p}(t) = \mathcal{L}_v(t)$. Condition (ii) is then exactly condition (b) of Cor. 3.1 and condition (iii) is condition (a) of Cor. 3.1. To check condition (c) of Cor. 3.1, we have

$Z(t,x,v) = \langle\hat{p}(t)-Q(t)(x-\hat{x}(t)),v\rangle-L(t,x,v)$, so $Z_{vv}(t,\hat{x}(t),\dot{\hat{x}}(t)) = -\hat{L}_{vv}(t) < 0$.

Therefore, we may find $\gamma>0$ such that $Z(t,x,v)$ is strictly concave in v for $(t,x,v)\in T(\hat{x},\dot{\hat{x}};\gamma)$, as desired.

Now Cor.3.1 demands that U be compact. However, from our conditions we may deduce strong local optimality of the pair $(\hat{x},\dot{\hat{x}})$ over $v\in U_X$, where

$$U_X = \{v=(v_1,\ldots,v_n): \sum_{i=1}^{n} v_i^2-X \le 0\}$$

for X sufficiently large. Since strong local optimality over $U = \mathbb{R}^n$ is equivalent to strong local optimality over U_X for all large X, we have our result.

<u>Remark</u>: Assumption (i) is the strengthened Legendre condition.

<u>Remark</u>: Cor. 5.1 of [25] is our Cor. 4.1 with condition (ii) replaced by the strengthened Weierstrass condition:

$$L(t,x,v)-L(t,x,w) \ge \langle L_v(t,x,w),v-w\rangle \qquad (4.1)$$

for all $(t,x,w)\in T(\hat{x},\dot{\hat{x}};\varepsilon)$, all $v\in\mathbb{R}^n$.

We claim our condition is the weaker. For suppose (4.1) holds, and $\hat{L}_{vv}(t) > 0$. Using the fact that $t\ L(t,\hat{x}(t),\cdot)$ is convex at $\dot{\hat{x}}(t)$, we may find a positive ϵ such that for all $t \in [a,b]$ and w with $|w-\dot{\hat{x}}(t)| < \epsilon$ we have

$$\langle L_v(t,\hat{x}(t),w)-\hat{L}_v(t), w-\dot{\hat{x}}(t)\rangle > 0 \qquad (4.2)$$

Fix t and pick $v \in \mathbb{R}^n$, $v \neq \dot{\hat{x}}(t)$. Take w to be a point on the line segment joining v to $\dot{\hat{x}}(t)$. Then (4.2) holds and $v-w = \alpha(w-\dot{\hat{x}}(t))$, some $\alpha > 0$.

Now,

$$L(t,\hat{x}(t),v)-\hat{L}(t) = L(t,\hat{x}(t),v)-L(t,\hat{x}(t),w)+L(t,\hat{x}(t),w)-\hat{L}(t)$$

$$\geq \langle L_v(t,\hat{x}(t),w), v-w\rangle - \langle \hat{L}_v(t), \dot{\hat{x}}(t)-w\rangle \qquad \text{by (4.1)}$$

$$= \langle L_v(t,\hat{x}(t),w)-\hat{L}_v(t), v-w\rangle + \langle \hat{L}_v(t), v-\dot{\hat{x}}(t)\rangle$$

$$= \alpha \langle L_v(t,\hat{x}(t),w)-\hat{L}_v(t), w-\dot{\hat{x}}(t)\rangle$$

$$+ \langle \hat{L}_v(t), v-\dot{\hat{x}}(t)\rangle$$

$> \langle \mathcal{L}_v(t), v-\dot{\hat{x}}(t) \rangle$ by (4.2).

This is exactly condition (ii).

At the end of this chapter we give an example that satisfies condition (ii) of Cor.4.1 but not (4.1).

Th^{m}4.2: Let $L(\cdot,\cdot,\cdot)$ be C^2, $\hat{x}:[a,b]\to\mathbb{R}^n$ be C^1, and suppose there exist $\varepsilon > 0$ and an absolutely continuous symmetric matrix function $Q(t)$ such that

(a) there exists a constant $c \in \mathbb{R}^n$ satisfying $\mathcal{L}_v(t) = \int_a^b \mathcal{L}_x(s)ds + c$ a.e.;

(b) $\mathcal{L}_{vv}(t) > 0$;

(c) $L(t,\hat{x}(t),v)-\mathcal{L}(t) > \langle \mathcal{L}_v(t), v-\dot{\hat{x}}(t)\rangle$ for all $v \in \mathbb{R}^n$ with $v = \dot{\hat{x}}(t)$, and all $t \in [a,b]$;

(d) $$\bar{M}(t,\dot{Q}(t)) = \dot{Q}(t)-Q(t)\mathcal{L}_{vv}^{-1}(t)\mathcal{L}_{vx}(t)-\mathcal{L}_{xv}(t)\mathcal{L}_{vv}^{-1}(t)Q(t)$$
$$-Q(t)\mathcal{L}_{vv}^{-1}(t)Q(t)-\mathcal{L}_{xv}(t)\mathcal{L}_{vv}^{-1}(t)\mathcal{L}_{vx}(t)+\mathcal{L}_{xx}(t) > 0$$
for $t \in [a,b]$ a.e. ;

Then $\hat{x}$ is a strong local minimum for (V).

Proof: The proof consists of showing that the above conditions are the same as the conditions of $Th^m 3.4$ when applied to the problem (V). It was seen in the remark following $Th^m 3.4$ that when $U = {}^n$ the assumptions of $Th^m 3.4$ become those of Cor. 3.2, with the exception of the requirement that U be compact, which is dropped.

Set $f(t,x,v) = v$, $g(t,x,v) = L(t,x,v)$ and $U = \mathbb{R}^n$. It is then easily checked that conditions (a),(b) and (c) give conditions (2),(3) and (4) of Cor. 3.2. It remains to show that condition (5) of Cor. 3.2 is satisfied. With f and g as above, we calculate

$$R(t) = \mathcal{L}_{vv}(t), D(t) = \mathcal{L}_{xx}(t), A(t) \equiv 0 ,$$
$$S(t) = \mathcal{L}_{vx}(t) \text{ and } B(t) \equiv 1 .$$

The function $\overline{M}(t,\dot{Q}(t))$ of Cor.3.2 is then

$$\overline{M}(t,\dot{Q}(t)) = \dot{Q}(t)+\mathcal{L}_{xx}(t)-(\mathcal{L}_{xv}(t)+Q(t))\mathcal{L}_{vv}^{-1}(t)(\mathcal{L}_{vx}(t)+Q(t))$$

which is the same as the function appearing in condition (d) of the theorem. $\overline{M}(t,\dot{Q}(t))$ is therefore positive definite, which completes the proof.

Remark: If there exists a solution Q_o of $\overline{M}(t,\dot{Q}_o(t)) = 0$, then by the embedding theorem of differential equations ([9]) there exists a solution Q of the inequality $\overline{M}(t,\dot{Q}(t)) > 0$. Since the existence of no conjugate points to a in (a,b] is equivalent to the existence of a solution Q_o to $\bar{M}(t,\dot{Q}_o(t)) = 0$ for $t\epsilon[a,b]$, the sufficiency condition involving the Jacobi condition ([9], Chap. 3. Th^{m}6.1] is a special case of Th^{m}4.2.

A similar result to Th^{m}4.2 appears in [25] in which condition (c) is replaced by the strengthened Weirstrass condition. We have already seen that our condition is weaker than the other. We now provide an example in which the conditions of Th^{m}4.2 are satisfied but the function L does not satisfy the extended Weierstrass condition.

Example: Consider the calculus of variations problem:

$$\text{(V) Min } J(x) = \int_a^b \frac{\dot{x}(t)^2}{\dot{x}(t)^2+1}\,dt$$

subject to $x(0) = x(1) = 0$.

In this problem we have $L(t,x,v) = \dfrac{v^2}{v^2+1}$.

Consider the arc $\hat{x}(t) \equiv 0$ for $t\epsilon[0,1]$.

Then $\hat{L}_x(t) \equiv \hat{L}_v(t) \equiv \hat{L}_{xv}(t) \equiv \hat{L}_{xx}(t) \equiv 0$, and $\hat{L}_{vv}(t) = 2$.

Condition (a) is therefore satisfied for $c = 0$, and $\hat{L}_{vv}(t) > 0$ which is condition (b). Pick $v \neq \dot{\hat{x}}(t) = 0$; then

$$L(t,\hat{x}(t),v)-\hat{L}(t) = \frac{v^2}{v^2+1} > \langle \mathcal{L}_v(t), v-\dot{\hat{x}}(t)\rangle = 0 ,$$

so condition (c) is satisfied. Condition (d) becomes

$\bar{M}(t,\dot{Q}(t)) = \dot{Q}(t) - 1/2Q^2(t)$ which is positive definite when $Q(t) \equiv t$ for $t \in [a,b]$. By Th^{m}4.2 we therefore conclude that $\hat{x}$ gives a strong local minimum for the problem. However, the result of [25] may not be applied, because $L(t,x,v)$ does not satisfy the strengthened Weierstrass condition. For suppose there existed an $\varepsilon > 0$ with

$$L(t,x,v)-L(t,x,w) \geq \langle L_v(t,x,w), v-w\rangle$$

for $(t,x,w) \in T(\hat{x},\dot{\hat{x}};\varepsilon)$ and all $v \in \mathbb{R}$.

Then

$$\frac{v^2}{v^2+1} - \frac{w^2}{w^2+1} \geq \frac{2w}{(w^2+1)^2}(v-w)$$

for w with $|w| < \varepsilon$ and all $v \in \mathbb{R}^n$.

Assume $\varepsilon < 1$, $w = 0$ and choose $v = 1/w$.

The above becomes

$$\frac{1-w^2}{w^2+1} \geq \frac{2(1-w^2)}{(w^2+1)^2}$$

or $1 \geq \dfrac{2}{w^2+1} > 1$ which is a contradiction.

Thus Th^{m}4.2 may be applied while the other result may not.

5. Interrelationships Between the Sufficient Conditions

In this chapter we examine how the different conditions developed for the generalized problem of Bolza, the problem of optimal control and the calculus of variations relate to one another. From the way the results were developed, we may construct the following diagram.

Figure (2)

$$\begin{tikzcd}
\begin{array}{l}\text{Prop}^{n}\text{s} \quad 2.1\\ \qquad\qquad 3.1\end{array} \arrow[r] \arrow[d] & \begin{array}{c}\text{Th}^{m}\text{s } 2.1\\ 3.1\\ (\text{Mod. H-J})\end{array} \arrow[r] & \begin{array}{l}\text{Cor. } 2.1\\ (\text{gen. H-J})\end{array} \\
\begin{array}{l}\text{Th}^{m}\ 2.2\\ \text{Cor. } 3.1\ (\text{zero order})\\ \text{Cor. } 4.1\end{array} \arrow[d] & & \\
\begin{array}{l}\text{Th}^{m}\text{s } 2.3\\ \quad 3.4\ (\text{second order})\\ \quad 4.2\end{array} & &
\end{tikzcd}$$

In figure (2) the arrow between two sets of results A and B means that a result for one of the three problems in group B is a corollary of the result dealing with the same problem in group A. We now go on to present conditions under which certain results are equivalent, and provide counterexamples to show when they are not equivalent.

The following result of Zeidan shows that under some conditions $\text{Th}^{m}2.1$ and Cor. 2.1 are equivalent.

<u>$\text{Prop}^{n}5.1$.</u> [Zeidan] Let $\hat{x}$ be an admissible arc for the generalized problem of Bolza (P). If the Hamiltonian (t,x,p) is continuous with respect to (t,x,p), then the hypotheses of $\text{Th}^{m}2.1$ and Cor. 2.1 are equivalent.

<u>Proof</u>: See [21].

<u>Remark</u>: The continuity assumption on H in $\text{Prop}^{n}5.1$ cannot be removed. This is shown in the following example, from [21].

<u>Example</u>: Consider the generalized problem of Bolza:

(P) minimize $J(x) = \chi_{\{0\}}(x(0)) + \chi_{\{0\}}(x(1)) + \int_0^1 \frac{1}{t^{1/2}}(x^2 + \frac{\dot{x}^2}{4} + 1)dt$,

where $\chi_A(\cdot)$ denotes the indicator function of the set A.

In this problem we have:

$$L(t,x,v) = \frac{1}{t^{1/2}}(x^2 + \frac{v^2}{4} + 1)$$

and

$$l(x_1,x_2) = \chi_{\{0\}}(x_1)+\chi_{\{0\}}(x_2) .$$

It is clear that L is $\mathcal{L}\times\mathcal{B}$- measurable.

Consider the arc

$$\hat{x}(t) \equiv 0 \text{ for } t\in[0,1].$$

Then,

$$J(\hat{x}) = l(\hat{x}(0),\hat{x}(1))+\int_0^1 L(t,\hat{x}(t),\dot{\hat{x}}(t))dt = \int_0^1 \frac{dt}{t^{1/2}} = 2$$

hence $J(\hat{x})$ is finite and $\hat{x}$ is admissible for (P).

The Hamiltonian corresponding to our problem (P) is:

$$H(t,x,p) = \sup\{\langle p,v\rangle-L(t,x,v): v\in \mathbb{R}\}$$

$$= \sup\{pv-\frac{1}{t^{1/2}}(x^2+\frac{v^2}{4}+1): v\in \mathbb{R}\}$$

$$= t^{1/2}p^2-\frac{1}{t^{1/2}}(x^2+1).$$

Thus, $H(0,x,p) = -\infty$ and the Hamiltonian is not continuous.

Define

$$W(t,x)\equiv 0 .$$

Then,

$$Z(t,x) = \max\{\alpha + H(t,x,\beta):(\alpha,\beta)\in\partial W(t,x)\}$$

$$= -\frac{1}{t^{1/2}}(x^2+1).$$

Clearly, for all $t\in[a,b]$ and $x\in \mathbb{R}$ we have:

$$Z(t,x) = -\frac{1}{t^{1/2}}(x^2+1)\leq -\frac{1}{t^{1/2}} = Z(t,\hat{x}(t)),$$

and

$$Z(t,\hat{x}(t)) = \frac{-1}{t^{1/2}} = \frac{d}{dt} W(t,\hat{x}(t)) - L(t,\hat{x}(t),\dot{\hat{x}}(t)).$$

Thus, all the conditions of $Th^m 2.1$ are satisfied and $\hat{x} \equiv 0$ is a strong local minimum for (P).

Now suppose that there exists a locally Lipschitz function $\overline{W}$ satisfying the conditions of corollory 2.1. In particular, $\overline{W}$ must satisfy

$$\frac{d}{dt} \overline{W} (t,0) = L(t,0,0) = \frac{1}{t^{1/2}} \quad \text{a.e.},$$

whence,

$$\overline{W}_t(t,0) = \frac{1}{t^{1/2}} \text{ for } t \in [0,1] \quad \text{a.e.}$$

Thus, $\overline{W}(\cdot,0)$ is not Lipschitz, and there is no Lipschitz function $\overline{W}$ satisfying the hypotheses of Cor. 2.1. However, $W(t,x) \equiv 0$ satisfies the conditions of $Th^m 2.1$.

We now present some new results along these lines.

<u>Claim</u>: Th^ms 2.1 and 3.1 are corollaries of $Prop^n 2.1$ and 3.1, but not vice versa. Proof is by example.

<u>Example</u>: Consider the problem .

$$\min J(x) = \int_0^1 \left\{(\sin 1/\sqrt{t})(x+tv) - \frac{x}{2\sqrt{t}} \cos 1/\sqrt{t}\right\} dt \ ,$$

subject to: $x(0) = x(1) = 0$.

This problem may be viewed as a problem of optimal control, in which

$$g(t,x,v) = (\sin 1/\sqrt{t})(x+tv) - \frac{x}{2\sqrt{t}} \cos 1/\sqrt{t} \ ,$$

$f(t,x,v) = v$, and

$x(0) = x(1) = 0$.

Let $\hat{x}(t) \equiv 0$. Then $J(\hat{x}) = \int_0^1 0dt = 0$, so $J(\hat{x})$ is finite.

Suppose that the cond^{n}s of Th^{m}3.1 are satisfied for some Lipschitz W. Then for $t \in [0,1]$ we have

$$Z(t,x) = \sup \{\alpha + H(t,x,\beta) : (\alpha,\beta) \in \partial W(t,x)\}$$

$$= \sup \{\alpha + \sup\{(\beta - t \sin 1/\sqrt{t})v - x \sin 1/\sqrt{t} + \frac{x}{2\sqrt{t}} \cos 1/\sqrt{t}$$

$$: v \in \mathbb{R}\} : (\alpha,B) \in \partial W(t,x)\}$$

$$= \sup \{\alpha : (\alpha,\beta) \in \partial W(t,x)\} - x \sin 1/\sqrt{t} + \frac{x}{2\sqrt{t}} \cos 1/\sqrt{t} \ \}$$

if $\beta = t \sin 1/\sqrt{t} \ \forall (\alpha,\beta) \in \partial W(t,x)$

$+ \infty$ if there exists $\beta \neq t \sin 1/\sqrt{t}$ with $(\alpha,\beta) \in \partial W(t,x)$, some $\alpha \in \mathbb{R}$.

Now, the conditions of Th^{m}3.1 and the fact that $L(t,\hat{x}(t),\dot{\hat{x}}(t)) = 0$ for $t \in [0,1]$ a.e. imply

$Z(t,x) \leq Z(t,\hat{x}(t)) = \frac{d}{dt} W(t,\hat{x}(t))$ for $t \in [0,1]$ a.e.

Since W is Lipschitz, we have $Z(t,x) < \infty$ for $t \in [0,1]$ a.e., from which we deduce that $(\alpha,\beta) \in \partial W(t,x) \Rightarrow \beta = t \sin 1/\sqrt{t}$ for $[0,1]$ a.e.

But if W is Lipschitz then for $t \in [0,1]$ $\partial W(t,x)$ is compact, so $\partial W(t,x)$ is uniformly bounded on $[0,1]$. Choosing $x > 0$ and $t = \frac{1}{(2n\Pi)^2}$ for $n \in N$, we then have that

$$Z(t,x) = \sup \{\alpha \colon (\alpha,\beta) \in \partial W(t,x)\} - x \sin 1/\sqrt{t} + \frac{x}{2\sqrt{t}} \cos 1/\sqrt{t}$$

$$= \sup \{\alpha \colon (\alpha,\beta) \in \partial W(t,x)\} + n\Pi x$$

which is unbounded for large n contradicting the observation that $Z(t,x) < \infty$.

Now let $\bar{W}(t,x) = xt \sin 1/\sqrt{t}$, which is easily seen to be absolutely continuous. For $t \in [0,1]$, we then have

$$\frac{d}{dt} \bar{W}(t,x(t)) = x(t)\left[\frac{-1}{2\sqrt{t}} \cos 1/\sqrt{t} + \sin 1/\sqrt{t}\right] + \dot{x}(t) t \sin 1/\sqrt{t} =$$

$L(t,x(t),\dot{x}(t))$ so condition (a) of Prop^n 3.1 is satisfied. Thus the conditions of Th^{m}s 2.1 (resp. 3.1) are stronger than those of Prop^{n}2.1 (resp. 3.1).

<u>**Claim:**</u> Th^{m}s 2.2, 3.2 and 4.1 are corollaries of Prop^{n}s 2.1 and 3.1, but not vice-versa. Proof is by example.

Example: Consider the problem

$$\min J(\cdot x) = \int_0^1 x^2(t)\dot{x}(t)dt,$$

subject to $x(0) = x(1) = 0$.

This is a calculus of variations problem in which

$$L(t,x,v) = x^2 v.$$

Let $\hat{x}(t) \equiv 0$. Then $J(\hat{x}) = 0$, so $\hat{x}$ is admissible.

If $W(t,x(t)) = \frac{x^3(t)}{3}$, then

$$\frac{d}{dt} W(t,x(t))-L(t,x(t),\dot{x}(t)) = x^2(t)\dot{x}(t)-x^2(t)\dot{x}(t) = 0$$

$$\leq \frac{d}{dt} W(t,\hat{x}(t))-L(t,\hat{x}(t),\dot{\hat{x}}(t)) = 0$$

so the conditions of Prop^{n}2.1 are satisfied.

Condition (2) of Th^{m}4.1 is in this case:

$$x^2 v \geq -Q(t)xv-1/2\dot{Q}(t)x^2 \qquad t\epsilon[a,b] \text{ a.e., } v\epsilon \mathbb{R}^n$$

for some absolutely continuous $Q(t)$.

We may assume $Q(t)\equiv 0$. Pick $0<t_1<t_2<1$ such that $|Q(t_1)| = 2\varepsilon > 0$ and $|Q(t_1)-Q(t)| \leq \varepsilon \quad t\epsilon[t_1,t_2]$.

Integrating with respect to t on $[t_1,t_2]$ gives

$$\int_{t_1}^{t_2} \left\{x^2 v+Q(t)\dot{x}v+1/2\dot{Q}(t)x^2\right\}dt$$

$$= x^2 v\Delta t + xv\int_{t_1}^{t_2} Q(t)dt + 1/2x^2(Q(t_2)-Q(t_1)) \geq 0,$$

where $\Delta t = t_2 - t_1$. Since $Q(t_2)-Q(t_1) \leq \varepsilon$,

we have

$$vx(x\Delta t + \int_{t_1}^{t_2} Q(t)dt) + 1/2x^2\varepsilon \geq 0 \ .$$

If $Q(t_1) = 2\varepsilon$, pick $x>0$ with $|x|<\varepsilon$.

Then

$$\varepsilon\Delta t < |x|\Delta t + \int_{t_1}^{t_2} Q(t)dt < 4\varepsilon\Delta t$$

and

$$v \geq \frac{-1/2|x|\varepsilon}{|x|\Delta t + \int_{t_1}^{t_2} Q(t)dt} \ .$$

If $Q(t_1) = -2\varepsilon$, pick $x < 0$ with $|x| < \varepsilon$.

Then

$$-4\varepsilon\Delta t < -|x|\Delta t + \int_{t_1}^{t_2} Q(t)dt < -\varepsilon\Delta t$$

and

$$v \geq \frac{-1/2(-|x|)\varepsilon}{-|x|\Delta t + \int_{t_1}^{t_2} Q(t)dt} \ .$$

In either case this is a contradiction when we choose

$$v < \frac{-1/2|x|\varepsilon}{\varepsilon\Delta t} = \frac{-|x|}{2\Delta t} \ .$$

Note that such points v exist in any neighbourhood of $\dot{\hat{x}}=0$, so condition (2) does not hold even locally in v.

In Figure (2), Th^{m}s 2.3 and 4.2 are shown as the respective corollaries of Th^{m}2.2 and Cor. 4.1.

<u>Prop^{n}5.2</u>. [Zeidan] Consider the problem of Bolza (P) with a fixed boundary value $x(a) = A$. Assume that the hypotheses of Th 2.1 are satisfied by a Lipschitz function W.

Suppose additionally that

(i) $W(t,\cdot)$ is C^3, the functions $W(\cdot,x), W_x(\cdot,x)$ and $W_{xx}(\cdot,x)$ are C^1, $W_t(t,\cdot)$ is C^2, and $W_{tx}(t,\hat{x}(t)) = W_{xt}(t,\hat{x}(t)), W_{txx}(t,\hat{x}(t)) = W_{xxt}(t,\hat{x}(t))$;

(ii) for $\hat{p}(t) = W_x(t,\hat{x}(t))$, the function $H(t,\cdot)$ is C^2 near $\hat{z} = (\hat{x},\hat{p})$ and the functions $H_z(\cdot,z)$ and $D_zH_z(\cdot,z)$ are continuous on $[a,b]$;

(iii) $\dot{\hat{x}}(t) = H_p(t,\hat{x}(t),\hat{p}(t))$ for $t\in[a,b]$.

Then the hypotheses of Th^{m}2.3 are satisfied for some C^1 - function Q.

<u>Remark</u>: In the case where the hypotheses of Th^{m}2.2 hold, the conditions of Th^m 2.1 and condition (i) of Prop^{n}5.3 are automatically satisfied.

<u>Proof of Prop^{n}5.2</u>: See reference [21].

The next proposition, for the problem of optimal control, shows that under certain conditions $Th^m3.1$ may be regarded as a corollary of $Th^m3.4$. We note here that the conditions $\hat{p}(b) = -\nabla l(\hat{x}(b))$ and $Q(b) < l_{xx}(\hat{x}(b))$, which appear in $Th^m3.4$, serve only to guarantee that condition (iii) of $Th^m3.2$ is satisfied, and are not essential to the thrust of the result.

<u>$Prop^n5.3$:</u> Consider the problem of optimal control (C) where U is convex, compact and polyhedral of the form (3.7). We will assume that the Lagrange multipliers associated with active constraints are all nonzero, and that $\hat{u}$ is a regular point for the constraints. Suppose there exists a function W satisfying the requirements of $Th^m3.1$ as well as conditions (i) and (ii) of $Prop^n5.2$. Suppose also that Hypothesis (H_4) holds, that R(t) as defined in Chapter 3 is positive definite, and that for $\hat{p}(t) = W_x(t,\hat{x}(t))$, we have

$$\langle \hat{p}(t), f(t,\hat{x}(t),\hat{u}(t))\rangle - g(t,\hat{x}(t),\hat{u}(t)) \qquad (5.20)$$

$$> \langle \hat{p}(t), f(t,\hat{x}(t),u)\rangle - g(t,\hat{x}(t),u)$$

for all $t \in [a,b]$ and all $u \in U$ with $u \neq \hat{u}(t)$. Then the hypotheses of $Th^m3.4$ are satisfied, where the conditions $\hat{p}(b) = -\nabla l(\hat{x}(b))$ and $Q(b) < l_{xx}(\hat{x}(b))$ are replaced by condition (iii) of $Th^m3.2$.

Remark: It is seen in the proof that $\hat{p}(t)$ as chosen satisfies (5.20) in which strict inequality is replaced by inequality.

Proof: Let $(\hat{x},\hat{u})$ be the admissible pair given in Th^{m}3.1. Following the proof of Prop^{n}5.2 we have

$$Z(t,x) = W_t(t,x)+H(t,x,W_x(t,x)) \tag{5.21}$$

and , for all $t\in[a,b]$,

$$Z(t,\hat{x}(t)) = W_t(t,\hat{x}(t)) + \sup\{\langle \hat{p}(t),f(t,\hat{x}(t),u)\rangle - g(t,\hat{x}(t),u):u\in\mathbb{R}^m\} \tag{5.22}$$

But condition (2) of Th^{m}3.1 gives

$$Z(t,\hat{x}(t)) = W_t(t,\hat{x}(t)) + \langle \hat{p}(t),f(t,\hat{x}(t),\hat{u}(t))\rangle - g(t,\hat{x}(t),\hat{u}(t)) .$$

From these last two equalities we obtain the weakened version of (5.20), as noted in the remark.

Now, from condition (2) of Th^{m}3.1 we have, for $t\in[a,b]$ and x near $\hat{x}(t)$, that

$$\max Z(t,x) = Z(t,\hat{x}(t)) .$$

If we define

$$P(t,x,u) = W_t(t,x) + \langle W_x(t,x),f(t,x,u)\rangle - g(t,x,u) \tag{5.23}$$

this is equivalent to saying that, for $t\epsilon[a,b]$, x near $\hat{x}(t)$ and $u\epsilon U$,

$$\max P(t,x,u) = P(t,\hat{x}(t),\hat{u}(t)) .$$

Consequently, for all t in [a,b], there exist Lagrange multiplier functions $\lambda(t)$ and $\mu(t)\leq 0$ such that

$$\nabla_x P(t,\hat{x}(t),\hat{u}(t)) = 0, \qquad (5.24)$$

$$\nabla_u P(t,\hat{x}(t),\hat{u}(t)) + \lambda^T(t)\nabla_u h(\hat{u}(t)) + \mu^T(t)\nabla_u d(\hat{u}(t)) = 0, \qquad (5.25)$$

$$\mu^T(t)\nabla_u d(\hat{u}(t)) = 0, \text{ and} \qquad (5.26)$$

$$\nabla^2_{x,u} P(t,\hat{x}(t),\hat{u}(t)) \qquad (5.27)$$

is negative semidefinite on the subspace

$\bar{T}(t) = \{\binom{x}{u}: x\epsilon\mathbb{R}^n, u\epsilon T(t)\}$

where $T(t)$ is as defined in (3.12).

Using (5.23), equation (5.24) becomes

$$W_{tx}(t,\hat{x}(t)) + f(t,\hat{x}(t),\hat{u}(t))W_{xx}(t,\hat{x}(t)) + f_x^T(t,\hat{x}(t),\hat{u}(t))W_x(t,\hat{x}(t))$$

$$- g_x(t,\hat{x}(t),\hat{u}(t)) = 0 .$$

Since $\hat{p}(t) = W_x(t,\hat{x}(t))$, we have

$$\dot{\hat{p}}(t) = W_{xt}(t,\hat{x}(t)) + f(t,\hat{x}(t),\hat{u}(t))W_{xx}(t,\hat{x}(t)),$$

and the above equality may be written

$$-\dot{\hat{p}}(t) = f_x^T(t,\hat{x}(t),\hat{u}(t))\hat{p}(t) - g_x(t,\hat{x}(t),\hat{u}(t)),$$

which is equation (3.8) of Th^{m}3.4.

Equation (5.25) gives

$$f_u^T(t,\hat{x}(t),\hat{u}(t))\hat{p}(t) - g_u(t,\hat{x}(t),\hat{u}(t))$$

$$+ \lambda^T(t)\nabla_u h(\hat{u}(t)) + \mu^T(t)\nabla_u d(\hat{u}(t)) = 0$$

which together with (5.26) gives equations (3.10).

Using (5.23), we find

$$\nabla^2_{x,u}P(t,x,u) =$$

$$\begin{bmatrix} \begin{array}{l} W_{xxt}(t,\hat{x}(t)) + W_{xxx}(t,\hat{x}(t))f(t,\hat{x}(t),\hat{u}(t)) \\ - \; + W_{xx}(t,\hat{x}(t))A(t) + A^T(t)W_{xx}(t,\hat{x}(t)) \\ - D(t) \end{array} & S^T(t)-W_{xx}(t,\hat{x}(t))B(t) \\ S(t) - B^T(t)W_{xx}(t,\hat{x}(t)) & R(t) \end{bmatrix}$$

where $A(t)$, $B(t)$, $D(t)$ and $S(t)$ are as in Chapter 3.

Define $Q_o(t) = - W_{xx}(t,\hat{x}(t))$. Then

$\dot{Q}_o(t) = - W_{xxt}(t,\hat{x}(t)) - W_{xxx}(t,\hat{x}(t))f(t,\hat{x}(t),\hat{u}(t))$,

and (5.27) yields that the matrix

$$\begin{bmatrix} \dot{Q}_o(t)+Q_o(t)A(t) & S^T(t)+Q(t)B(t) \\ +A^T(t)Q_o(t)+D(t) & \\ & \\ S(t)+B^T(t)Q(t) & R(t) \end{bmatrix}$$

is positive semidefinite on the subspace $\bar{T}(t)$.

Following the steps in the proof of Prop^{n}3.4, we deduce in the same way that $M(t,\dot{Q}_o(t)) \geq 0$ where $M(t,\dot{Q}(t))$ is defined by (3.13). Using the imbedding theorem of differential equations and the fact that $Q_o(b) = -W_{xx}(b,\hat{x}(b))$, we may find a C^1 matrix function $Q(t)$ with $M(t,\dot{Q}(t)) > 0$ and $Q_o(b) < - W_{xx}(b,\hat{x}(b))$.

From the proof of Prop^{n}5.2 we may then find some $\alpha > 0$ such that, for $|d| < \alpha$,

$l(\hat{x}(b)+d)-l(\hat{x}(b)) \geq - \langle\hat{p}(b),d\rangle + 1/2 \langle d,Q(b)d\rangle$,

which is condition (iii) of Th^{m}3.2.

Prop^{n}s 5.2 and 5.3 may now be used to show that, under certain conditions, the zero order results of Th^m 2.2, Cor. 3.1 and Cor. 4.1 are in fact equivalent to the second order results of Th^{m}s 2.3, 3.4 and 4.2. We first examine the case for the generalized problem of Bolza.

<u>Prop^{n}5.4</u>: Consider the generalized problem of Bolza (P) with a fixed boundary value $x(a) = A$. For the arcs $\hat{x}$ and $\hat{p}$, assume that the function $H(t,\cdot)$ is C^2 near $\hat{z} = (\hat{x},\hat{p})$, that the functions $H_z(\cdot,z)$ ~~and~~ $D_z(\cdot,z)$ are continuous on $[a,b]$, and that

$$\dot{\hat{x}}(t) = H_p(t,\hat{x}(t),\hat{p}(t)) \text{ for } t\in[a,b].$$

Suppose also that the function Q appearing in Th^{m}s 2.2 and 2.3 is C^1. Then the two theorems are equivalent.

<u>Proof</u>: We have already seen that Th^{m}2.3 is a corollary of Th^{m}2.2. It remains to show that the converse is true. Assume the conditions of Th^{m}2.2 are satisfied, and define the function

$$W(t,x) = \langle \hat{p}(t),x\rangle - 1/2 \ \langle x-\hat{x}(t),Q(t)(x-\hat{x}(t))\rangle \qquad (5.28)$$

on $T(\hat{x};\varepsilon)$, where ε is the number supplied in Th^{m}2.2. It is shown in [21] that $W(t,x)$ satisfies Th^{m}2.1. Using Prop^{n}5.3, we than deduce that the conditions of Th^{m}2.3 are satisfied.

<u>Prop^{n}5.5</u>. Consider the problem of optimal control (C) where U is polyhedral of the form (3.7). We will assume that the Lagrange

multipliers associated with active constraints are all nonzero, and that $\hat{u}$ is a regular point for the constraints. Assume that Hypothesis (H_4) holds, and that $R(t)$ is positive definite.

Suppose also that the function Q appearing in Cor. 3.1 and Th^{m}3.4 is C^1. Then the two results are equivalent, where the conditions $\hat{p}(b) = -\nabla l((\hat{x}(b))$ and $Q(b) < l_{xx}(\hat{x}(b))$ are replaced by condition (iii) of Th^{m}3.2.

Proof: We have already seen that Th^{m}3.4 is a corollary of Cor.3.1. To prove the converse, suppose the conditions of Cor. 3.1 are satisfied and let $W(t,x)$ be as defined in (5.28) on the set $T(\hat{x};\varepsilon)$. Our chain of reasoning will be as follows: if the conditions of Cor. 3.1 are satisfied, then, by the proof of that result, the conditions of Th^{m}3.2 are satisfied. Given this, the function $W(t,x)$ will be seen to satisfy the conditions of Th^{m}3.1. Prop^{n}5.3 will then be applied to deduce that the conditions of Th^{m}3.4 hold for some C^1 function Q.

So, let us assume that the conditions listed in Th^{m}3.2 hold. As in the proof of Prop^{n}5.4, we find, after using (5.28), that condition (1) of Th^{m}3.1 is just condition (iii) of Th^{m}3.2. Using [24] we then see that condition (2) of Th^{m}3.1 is also satisfied.

This shows that the function $W(t,x)$ satisfies the conditions of Th^{m}3.1. It is then easy to see that $W(t,x)$ satisfies the hypotheses of

Prop^{n}5.2, and we conclude that the conditions of Th^{m}3.4 are satisfied by some C^1 function Q, where the requirement that $\hat{p}(b) = -\nabla l(\hat{x}(b))$ and $Q(b) < l_{xx}(\hat{x}(b))$ is replaced by condition (iii) of Th^{m}3.2.

<u>Prop^{n}5.6</u>: Consider the calculus of variations problem (V). Let $\hat{x}:[a,b] \to \mathbb{R}^n$ be C^1, $L(\cdot,\cdot,\cdot)$ be C^2 and $Q(\cdot)$ be C^1 where Q is the function appearing in Cor. 4.1 and Th^{m}4.2. Then the conditions of Cor. 4.1 and Th^{m}4.2 are equivalent.

<u>Proof</u>: Since Th^{m}4.2 is derived from Th^{m}3.4 and Cor. 4.1 is Cor. 3.1 when applied to the problem (V), it follows that Th^{m}4.2 is a corollary of Cor. 4.1. Prop^{n}5.5 then shows the converse.

BIBLIOGRAPHY

[1] V.G. Boltyanskii, Sufficient conditions for optimality and the justification of the dynamic programming method, SIAM J. Control, (1966), pp. 326-361.

[2] F.H. Clarke, Admissible relaxation in variational and control problems, J. Math. Anal. Appal., 51 (1975), pp. 557-576.

[3] ———————, Necessary conditions for a general control problem, in Proceedings of the Symposium on the Calculus of Variations and Optimal Control, D.L. Russell, Editor (Mathematics Research Center, University of Wisconsin-Madison), Academic Press, N.Y. (1976).

[4] ———————, The general problem of Bolza, SIAM J. Control Optimization, 14 (1976), pp. 682-699.

[5] ———————, Extremal arcs and extended Hamiltonian systems, Trans. Amer. Math. Soc., 231 (1977), pp. 349-367.

[6] ———————, Generalized gradients of Lipschitz functionals, Adv. Math, 40 (1981), pp. 52-67.

[7] ———————, Optimization and Nonsmooth analysis, Wiley Interscience, N.Y. (1983).

[8] P.E. Gill, W. Murray and M.H. Wright, Practical Optimization, Academic Press, London (1981).

[9] M.R. Hestenes, Calculus of Variations and Optimal Control Theory, J. Wiley, N.Y. (1966).

[10] V.F. Krotov, Methods for solving variational problems on the basis of sufficient conditions for an absolute minimum, I, II,

III, Automat. Remote Control, (1962) pp. 1473-1484; (1963) pp. 539-553, (1964) pp. 924-933.

[11] D.G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley (1973).

[12] O.L. Maugasarian, Sufficient conditions for the optimal control of linear systems, SIAM J. Control, (1966), pp. 139-152.

[13] D.Q. Mayne, Sufficient conditions for a control to be a strong minimum, J. Opt. Theory Appl., 21 (1977), pp. 339-351.

[14] D.C. Offin, A. Hamiltonian-Jacobi approach to the differential inclusion problem, Thesis, University of British Columbia, Vancouver, B.C. (1979).

[15] L.S. Pontryagain, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko, The Mathematical Theory of Optimal Processes, J. Wiley, N.Y. (1965).

[16] R.T. Rockafellar, Convex integral functionals and duality. In: Contributions to Nonlinear Functional Analysis, Academic Press, N.Y., (1971), pp. 215-236.

[17] ——————, Optimal arcs and the minimum value function in problems of Lagrange, Trans. Amer. Math. Soc., 180 (1973), pp. 53-83.

[18] R.B. Vinter and R.M. Lewis, The equivalence of strong and weak formulations for certain problems in optimal control, SIAM J. Control Optimization, 16 (1978), pp. 546-570.

[19] ——————, A verification theorem which provides a necessary and sufficient condition for optimality, IEEE Trans. on Automatic Control, Vol. AC-25, No. 1 (1980).

[20] R.B. Vinter, New global optimality conditions in optimal control theory, SIAM J. Control Optimization, 21 (1980), pp. 235-245.

[21] V. Zeidan, Sufficient conditions for optimal control and generalized problem of Bolza (Ph.d. thesis), University of British Columbia, Vancouver, B.C. (1982).

[22] ————, Sufficient conditions for the generalized problem of Bolza, Trans. Amer. Math. Soc. Vol. 275, No. 2 (1983).

[23] ————, Extended Jacobi sufficiency criterion for optimal control, SIAM J. Control Optimization, Vol. 22, No. 2 (1984).

[24] ————, A Modified Hamilton-Jacobi approach in the generalized problem of Bolza, Appl. Math. Optimization, 11 (1984), pp. 97-109.

[25] ————, First and second order sufficent conditions for optimal control and the calculus of variations, 11 (1984), pp. 2099-226.

[26] ————, Conjugate points criterion for optimal control, Proceedings of 22nd IEEE Conference on Decision and Control (1983), pp. 379-383.