ON FIRST ORDER NECESSARY CONDITIONS FOR VARIATIONAL AND OPTIMAL CONTROL PROBLEMS

A dissertation submitted in partial
satisfaction of the requirements
for the degree Doctor of Philosophy
in Mathematics
at the
University of California, Los Angeles

by

Theodore Guinn

A TECHNICAL REPORT

The preparation of this dissertation was sponsored in part by the U.S. Army Research Office (Durham). Reproduction in whole or in part is permitted for any purpose of the United States Government.

Numerical Analysis Research
Department of Mathematics
University of California
Los Angeles 24, California

June 1964

TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION

On First Order Necessary Conditions for
Variational and Optimal Control Problems

by

Theodore Guinn

University of California, Los Angeles, 1964
Professor Magnus R. Hestenes, Chairman

A general variational problem is considered which contains most control problems as special cases. Included are differential constraints as well as isoparametric and finite inequalities on both state and control vairables.

The method of M. R. Hestenes for proving first order necessary conditions is used to obtain similar conditions under weaker hypothesis, principally requiring Lebegue integrability of all functions with respect to time. It is then shown these results can be easily extended to include the problem with inequality constraints on the space variables independent of control variables.

1. INTRODUCTION

In the first comprehensive study of optimal control problems from a variational calculus viewpoint, Hestenes in 1948 [1]* derived necessary conditions for a broad range of control problems by showing them to be a special case of the problem of Bolza. His formulation included not only the classical Euler-Lagrange equations and the Weierstrass condition, but also a formulation in terms of a Hamiltonian-type function - the latter has since become known as "Pontryagin's maximum principle." By use of a device of Valentine's [2], he was able to include problems with inequality constraints which are functions of both state and control variables.

Some ten years later, Pontryagin and his students published a series of papers [10, 11, 12, 13, 14] and principally [3], which developed the "maximum principle," a combination of the Euler-Lagrange equations and the Weierstrass condition using a Hamiltonian-type formulation. The proofs employ a variation of the method used by McShane [15] to derive the Weierstrass condition without normality assumptions. Where McShane used both weak and strong variations, however, Pontryagin was able to obtain his results using only strong variations and under somewhat weaker

* Numbers in square brackets refer to the bibliography.

hypotheses. The limitation which originally permitted constraints only on control variables was partially removed by Gamkrelidze [4] to include a combination of inequality constraints on the control variables and on the state variables, but not joint constraints. Because his method requires separating an optimizing arc into subarcs on which constraints on state variables are equalities and on which they are strict inequalities he has difficulties with continuity of the Lagrange multipliers. Berkovitz [8] using the classical variational theory from [9] reproduced all but one of Gamkrelidze's results but has the same difficulties. The papers of Pontrgagin et. al. were finally incorporated in [5].

Hestenes in 1963 [6] using a simplification of McShane's method obtained all of the results in [1] using only strong variations. This paper considers isoperimetric inequality constraints combined with inequality constraints which are functions of time, the state variables and control variables simultaneously, and obtains necessary conditions under only slightly stronger hypotheses than required by the Pontryagin school.

In Sections 2-6 of this paper the method used by Hestenes is shown to produce all of the results in [3, 5, 10-14] and many of those in [6] under the assumption that the functions appearing in differential and isoparametric constraints are Lebesque integrable with respect to time for any admissible control. The hypotheses are essentially those required for existance, uniqueness and embedding theorems for a system of ordinary differential equations.

In Section 7 the problem considered by Gamkrelidze [4] and Berkovitz [8] is shown to be a simple extension of the problem considered by Hestenes. Not only are their results obtained but also it is further shown that the Lagrange multipliers here used can be chosen to have less discontinuities than their's and an additional inequality for one is obtained.

We here only consider first order necessary conditions that a given function be minimizing. The problems of second order necessary conditions, sufficiency conditions and existence are not discussed.

Section 2 contains basis definitions as well as results for functions of a finite number of vairables which are fundamental for later theorems. This includes a very special form of an implicit function theorem for continuous functions which have a differential at a single point.

Sections 3-5 deal with control-type problems involving differential and integral constraints in the case that the control functions do not depend explicitly on the state variable. Section 3 considers problems in the state vari les are absent. In Section 4, the case in which the state variables appear linearly is reduced to the previous case by a standard transformation. Section 5 removes the restriction of linearity in any variable.

The problem in which the control variables are dependent on state variables is considered in Section 6 for a broad class of problems. This problem is extended in Section 7.

2. FUNCTIONS OF A FINITE NUMBER OF VARIABLES

The problem considered here will be that of minimizing a function $f_o(x)$ for $x = (x^1, x^2, \ldots, x^n)$ in some set X subject to the condition that

$$(2.1) \qquad f_\alpha(x) \geqq 0 \qquad (\alpha = 1, \ldots, m)$$

The main result is contained in theorem 2.2 which is used extensively in th remaining sections.

Let X be a set of points in n-dimensional Euclidean space R^n and let $f(x)$ be a function defined on X. If $f'(x_o, h)$ is a linear function in h for fixed x_o and if

$$(2.2) \qquad \lim_{x = x_o} \frac{f(x) - f(x_o) - f'(x_o, x - x_o)}{|x - x_o|} = 0 \qquad (x \text{ in } X)$$

then we will call $f'(x_o, h)$ the <u>differential</u> of $f(x)$ at x_o.

A sequence $\{x_q\}$ in X will be said to <u>converge directionally</u> to x_o if it is distinct from x_o, converges to x_o and if the sequence

$$(2.3) \qquad h_q = \frac{x_q - x_o}{|x_q - x_o|}$$

converges to a vector h. If $f(x)$ has a differential at x_o and $\{x_q\}$ converges directionally to h it follows from (2.2) that

$$(2.4) \qquad \lim_{q = \infty} \frac{f(x_q) - f(x_o)}{|x_q - x_o|} = f'(x_o, h)$$

Let x_o be a point of X and let $\{x_q\}$ converge to x_o in the direction h. For such vectors h, the set of all vectors $\alpha h (\alpha \geqq 0)$ form a closed cone which we shall call the tangent cone at x_o.

Theorem 2.1. Let f(x) be a function continuous on X and having a differential at x_o. If x_o affords a local minimum to f(x) on X then

$$f'(x_o, h) \geqq 0$$

for every vector h in the tangent cone C of X at x_o. If C is a linear space then $f'(x_o, h) = 0$ on C.

This follows from (2.4).

Now we consider the problem of minimizing a function f subject to conditions (2.1). To prove the main theorem we need some lemmas.

Lemma 2.1. Let $F_\alpha(x, t)$ $(\alpha = 1, \ldots, m)$ be continuous for $|x| \leqq \rho$ and $0 \leqq t \leqq \epsilon$. Let

$$(2.5) \qquad F_\alpha(x, o) = A_{\alpha i} x^i$$

where the matrix $(A_{\alpha i})$ of constants has rank m. Then there exist functions $x^i(t)$ $(0 \leq t \leq \delta;\ \delta \leq \epsilon)$ and a constant M such that

(2.6) $\quad F_\alpha(x(t), t) = 0, \ |x(t)| \leqq MN(t) \leqq \rho, \ \lim_{t = 0^+} x(t) = 0$

<u>where</u>

(2.7) $$|N(t)| = \sup_{|x| \leqq \rho} |F(x, t) - F(x, 0)|$$

To prove this result let $A_{\beta i}$ ($\beta = m+1, \ldots n;\ i = 1, \ldots, n$) be such that

$$|A_{ji}| \neq 0$$

Set

(2.8) $$F_\beta(x, t) = A_{\beta i}x^i$$

The functions $F_i(x, t)$ satisfy the hypothesis of the theorem. Let (A^{ij}) be the inverse of (A_{ji}) and set

(2.9) $$G^i(x, t) = x^i - A^{ij}F_j(x, t) = A^{ij}(F_j(x, 0) - F_j(x, t))$$

There is accordingly a constant M such that

(2.10) $$|G(x, t)| \leqq MN(t) \qquad (0 \leqq t \leqq \varepsilon; |x| \leqq \rho)$$

Choose δ such that

$$MN(t) \leq \rho \qquad (0 \leqq t \leqq \delta)$$

The transformation

(2.11) $$y^i = G^i(x, t)$$

for fixed t on $0 \leq t \leq \delta$ maps the sphere

$$(2.12) \qquad |x| \leq MN(t)$$

onto itself. By the Browder fixed point theorem [16] there is a point $x^i(t)$ in this sphere such that

$$(2.13) \qquad x^i(t) = G^i(x(t), t) = x^i(t) - A^{ij}F_j(x(t), t)$$

Hence, $F_j(x(t), t) = 0$ as was to be proved.

We shall now introduce a particular convex cone K^- of vectors $k = (k_o, k, \ldots, k_p)$ defined by the relations $k_o < 0$, $k_\alpha > 0$ ($\alpha = 1, \ldots, p'$), k_β arbitrary, ($\beta = p' + 1, \ldots, p''$) and $k_\gamma = 0$ ($\gamma = p'' + 1, \ldots, p$). For this cone we have

<u>Lemma 2.2</u>. <u>Given a linear functional</u>*

$$L(k) = \lambda_\rho k_\rho$$

<u>the inequality</u> $L(k) \leq 0$ <u>holds on</u> K^- <u>if and only if</u>

$$(2.14) \qquad \lambda_o \geq 0, \ \lambda_\alpha \leq 0 \ (\alpha = 1, \ldots, p') \qquad \lambda_\beta = 0 \ (\beta = p' + 1, \ldots, p'')$$

In order to prove this observe that $\bar{k} = (-1, 0, \ldots, 0)$

---

*Here and following the summation convention holds for repeated indices unless otherwise noted.

is in the closure of K^-. If $L(k) \leq 0$ on K^- we have $-\lambda_0 = L(\bar{k}) \leq 0$. Similarly for σ on the range 1, ..., p' let $\bar{k}$ be the vector having $\bar{k}_\sigma = 1$ and $\bar{k}_\rho = 0$ otherwise. Again, $\bar{k}$ is in the closure of K^- and $\lambda_\sigma = L(\bar{k}) \leq 0$ Finally, for σ on the range p' + 1, ..., p choose $\bar{k}$ such that $\bar{k}_\sigma = \pm 1$ and $\bar{k}_\rho = 0$ otherwise. In this event $\pm\lambda_\sigma = L(\bar{k}) \leq 0$. Hence, (2.14) holds. The converse is immediate. This proves the lemma.

A second convex cone K^+ will be said to be separated from K^- if there is a non-null linear functional $L(k)$ such that $L(k) \leq 0$ on K^- and $L(k) \geq 0$ on K^+

Lemma 2.3. Given a convex cone K^+ the following statements are equivalent: (1) The cone K^+ is separated from K^-; (2) No vector in K^- is interior to the convex cone $K^+ - K^-$; (3) The cone $K^+ - K^-$ does not contain all vectors k.

Suppose first that K^+ is separated from K^- and let $L(k)$ be a non-null linear functional such that $L(k) \leq 0$ on K^- and $L(k) \geq 0$ on K^+. Then $L(k) \geq 0$ on $K^+ - K^-$ Since $L(k) > 0$ on the interior it follows that (2) and (3) hold.

Suppose next that there is a vector k' in K^- interior to $K^+ - K^-$. Then there is a constant $\varepsilon > 0$ such that $k' + k$ is interior to $K^+ - K^-$ for all k such that $|k| \leq \varepsilon$. Given k such that $|k| \leq \varepsilon$ we may choose vectors k^+, k^- such that $k' + k = k^+ - k^-$ Hence

$$k = k^+ - (k^- + k')$$

is in $K^+ - K^-$, that is, every vector k is in $K^+ - K^-$. Consequently (1) and (3) hold.

Let K be the closure of $K^+ - K^-$. If K contains all vectors k, then a vector k^- in K^- is interior to K and hence also interior to $K^+ - K^-$, since $K^+ - K^-$ is convex. Hence, if (3) holds there is a vector k' which is not in K. Choose k'' in K such that $|k' - k''|$ is a minimum. The numbers

$$\lambda_\rho = k''_\rho - k'_\rho \qquad (\rho = 0, \ldots, p)$$

are not all zero and

$$L(k) = \lambda_\rho k_\rho$$

has the property that $L(k) \geq 0$ on K Hence, $L(k) \geq 0$ on K^+ and $L(k) \leq 0$ on K^-.

<u>Lemma 2.4</u>. <u>If a convex cone</u> K^+ <u>is not separated from</u> K^- <u>there is a vector</u> $\bar{k}$ <u>in</u> K^- <u>which is the sum of</u> $r = p - p'' + 1$ <u>vectors</u> $k_1, \ldots, k_r$ <u>in</u> K^+ <u>whose matrix</u>

$$(2.15) \qquad (k_{\gamma\sigma}) \qquad (\gamma = p'' + 1, \ldots, p;\ \sigma = 1, \ldots, r)$$

<u>has rank</u> $r - 1$.

In order to prove this result observe that $K^+ - K^-$ contains all vectors k, by Lemma 2.3. A vector k' in K^- is therefore expressible as the sum $k' = k^+ - k^-$, where k^+ is in K^+ and k^- is in K^-. The vector $k^+ = k^- + k'$ is therefore in K^- also.

Let $k^*_{\rho\sigma}(\sigma = 1, \ldots, r)$ be chosen so that the matrix

$$(k^*_{\gamma\sigma}) \qquad (\gamma = p'' + 1, \ldots, p; \sigma = 1, \ldots, r)$$

has rank $r - 1$ and $k^*_{\gamma 1} + \ldots + k^*_{\gamma r} = 0$. The vectors $k^*_\sigma = (k^*_{o\sigma}, \ldots, k^*_{\rho\sigma})$ are expressible in the form

$$k^*_\sigma = k^+_\sigma - k^-_\sigma$$

where k^+_σ are in K^+ and k^-_σ are in K^-. Since $k^*_{\gamma\sigma} = k^+_{\gamma\sigma}$ we may replace k^*_σ by k^+_σ. The vectors k^*_σ are then in K^+ and the vector

$$k^* = k^*_1 + \ldots + k^*_r$$

has

$$k^*_\gamma = 0 \qquad (\gamma = p'' + 1, \ldots, p).$$

The vector

$$\bar{k} = k^+ + \epsilon k^*$$

is in K^- and K^+ if ϵ is small and positive. The vecotrs

$$k_\sigma = \frac{1}{r} k^+ + \epsilon k^*_\sigma$$

are in K^+, have $\bar{k}$ as their sum, and its matrix (2.15) of rank $r - 1$. This proves the lemma.

Consider now a set of $p + 1$ real valued functions $J_o, \ldots, J_p$ defined on a space S. Let x_o be a point of S. A cone K^+ will be said to be a <u>derived cone</u> for $J_o, \ldots, J_p$ at x_o if it is convex and if given any set of vectors

$$k_1, \ldots, k_q$$

in K^+, there is a function

$$x(b_1, \ldots, b_q) \qquad\qquad 0 \leq b_j \leq \epsilon$$

whose range is in S, containing x_o for $b = 0$, and having the property that the functions

$$f_\rho(b) = J_\rho(x(b)) - J_\rho(x_o)$$

are continuous and have

$$f'_\rho(0, h) = k_{\rho\sigma} h_\sigma$$

as differentials at $b = 0$. If K' is a basis for K^+, and every set of vectors $k_1, \ldots, k_q$ in K' has the property described above then K^+ is a derived cone for $J_o, \ldots, J_p$ at $x = x_o$.

<u>Theorem 2.2</u>. <u>Let</u> K^+ <u>be a derived cone for</u> $J_o, \ldots, J_p$ <u>at a point</u> x_o <u>in</u> S. <u>Suppose further that</u>

(2.16) $\quad J_\rho(x_o) = 0 \ (1 \leq \rho \leq p', \ p'' < \rho \leq p), \ J_\rho(x_o) > 0$

$(p' < \rho \leq p'')$.

•

<u>If</u> x_o <u>minimizes</u> J_o <u>on</u> S <u>subject to the conditions</u>

$$(2.17) \qquad J_\rho(x) \geq 0 \ (1 \leq \rho \leq p''), \ J_\rho(x) = 0 \ (p'' < \rho \leq p)$$

<u>then</u> K^+ <u>is separated from</u> K^-.

Suppose that K^+ is not separated from K^-. Then there exist vectors $\bar{k}$, k_1, ..., k_r having the properties described in Lemma 2.4. Let

$$x(b_1, \ldots, b_r) \qquad\qquad 0 \leq b_j \leq \epsilon$$

be points in S, such that $x(0) = x_o$, and

$$f_\rho(b) = J_\rho(x(b)) - J_\rho(x_o)$$

are continuous and have

$$f'_\rho(0, h) = k_{\rho\sigma}h_\sigma$$

as differentials at $b = 0$. Let $\bar{h}_\rho = 1$. The vector $\bar{h}$ has the property that

$$f'_\rho(0, \bar{h}) = \bar{k}_\rho.$$

Set

$$(2.18) \qquad F_\gamma(y, t) = \frac{1}{t} f_\gamma(t\bar{h} + ty) \quad t > 0 \quad (\gamma = p'' + 1, \ldots, p)$$

$$= k_{\gamma\sigma}y_\sigma \qquad t = 0$$

The functions $F_\gamma(y, t)$ are continuous for

$$|y| \leq \delta, \quad 0 \leq t \leq \delta$$

if δ is sufficiently small. By Lemma 2.1 there are functions $y(t)$ $0 \leq t \leq \delta'$ such

$$(2.19) \qquad F_\gamma(y(t), t) = 0 \qquad\qquad \lim_{t=0^+} y(t) = y(0) = 0$$

Setting

$$(2.20) \qquad b(t) = t\bar{h} + ty(t)$$

It is seen that

$$f_\gamma(b(t)) = 0 \qquad (0 \leq t \leq \delta'; \; \gamma = p'' + 1, \ldots, p)$$

and that

$$g_\rho(t) = f_\rho(b(t)) \qquad (\rho = 0, 1, \ldots, p)$$

has $\bar{k}_\rho$ as its right hand derivative at $t = 0$. Since

$$\bar{k}_0 < 0, \quad \bar{k}_\rho > 0 \qquad (1 < \rho \leq p')$$

it follows that on an interval $0 < t \leq \delta''$ we have

$$f_0(b(t)) < 0, \quad f_\rho(b(t)) > 0 \qquad (1 < \rho \leq p')$$

which is impossible. This proves the theorem. The proof of this theorem and the proceeding four lemmas was repeated from the appendix of [6] for completeness.

We now consider the problem of minimizing the function $f_o(x)$ subject to the conditions that

(2.21) $f_\alpha(x) \geq 0$ $(\alpha = 1, \ldots, m)$

Here we take X to be R^n, in which case the tangent cone at a point x_o is again all of R^n. Then K^+ is given by

(2.22) $k_\rho = f'(x_o, h)$ $(h \text{ in } R^n)$

With this we have the following theorem.

Theorem 2.3. Let $f_o(x), f_1(x), \ldots, f_m(x)$ be $m + 1$ functions continuous on R^n and having a differential at the point x_o which minimizes $f_o(x)$ subject to (2.21).

There then exist constants $\lambda_o \geq 0, \lambda_1, \ldots, \lambda_m$ not all zero such that

(2.23) $\lambda_o f'_o(x_o, h) - \lambda_\alpha f'_\alpha(x_o, h) \geq 0$ $(\text{all } h \text{ in } R^n)$

For $\alpha = 1, \ldots, m$, $\lambda_\alpha \geq 0$ and if $f_\alpha(x_o) > 0$ the equality holds. If the functions are linear then λ_o can be chosen as 1 and the λ_α chosen so that

(2.24) $\lambda_o f'_o(x_o, h) - \lambda_\alpha f'_\alpha(x_o, h) = 0$ $(\text{all } h \text{ in } R^n)$

The first conclusion follows from Theorem 2.2. For the second, since K^+ is the image of a linear space it is linear.

The vector $(-1, 0, \ldots, 0)$ is not in K^{+} since if it were by linearity so also would be $(-t, 0, \ldots, 0)$ for $0 \leq t \leq \delta$ and x_o would not minimize $f_o(x)$ subject to $f_\alpha(x) \geq 0$ $(\alpha = 1, \ldots, 0)$ which proves the theorem.

3. ELEMENTARY MINIMUM PROBLEMS INVOLVING INTEGRALS

In this section minimization problems are considered which involve integral constraints independent of the space variables. The main results are contained in Theorem 3.2 for the fixed endpoint case and Theorem 3.3 for the variable endpoint case. The results obtained here are fundamental for proofs of theorems in later sections.

The problem considered is that of minimizing a function

$$(3.1) \qquad I(u) = \int_{t^0}^{t^1} f(t, u(t))dt$$

in a class of functions

$$u^k(t) \qquad (t^0 \leq t \leq t^1;\ k = 1, \ldots, m)$$

having certain prescribed properties.

Let R be an open set in (t, u) - space. Assume that

$$f(t, u) = f(t, u^1, \ldots, u^m)$$

is integrable with respect to t on $t^0 \leq t \leq t^1$ for fixed u. Let u(t) be a function with (t, u(t)) contained in some subset of R and such that f(t, u(t)) is integrable for t on the range for which u(t) is defined. Then a point τ will be called

<u>ordinary</u> for u(t) if τ is a point of definition for u(t) and if

$$(3.2) \qquad \frac{d}{d\tau}\int_a^\tau f(t, u(t))dt = f(\tau, u(\tau)) < \infty$$

For brevity, a point t will be called ordinary if the particular function u(t) for which it is an ordinary point is clear from the context. If t is an ordinary point for u(t), then (t, u(t)) will be called an ordinary point for f(t, u(t)). A point t which is not an ordinary point will be called a singular point.

Let U be a subset of R with the property that if $(\bar{t}, \bar{u})$ is in U except possible for $\bar{t}$ in a set $\mathcal{T}$ of linear measure zero there is a function

$$u^k(t) \qquad (\bar{t} - \delta \leq t \leq \bar{t} + \delta)$$

for some $\delta > 0$ such that the following hold for $|\bar{t} - t| \leq \delta$

i) (t, u(t)) is in U

ii) f(t, u(t)) is an integrable function of t

iii) $u(\bar{t}) = \bar{u}$

iv) $\bar{t}$ is an ordinary point for f(t, u(t)).

Then an element $(\bar{t}, \bar{u})$ in U will be called <u>admissible</u>. Denote by U(t) the set of all u for which (t, u) is admissible.

By an <u>admissible function</u> will be meant a function whose elements (t, u(t)) are admissible and such that f(t, u(t)) is integrable. Note that for any admissible function u(t), the set of ordinary points t has full measure.

Remark 1. The set T of measure zero will, in general, be the set of points t for which $f(t, u(t))$ is singular for any choice of $u(t)$. Without excluding such a set the theory being developed would not even apply to functions $f(t, u)$ which are piecewise continuous in t.

Remark 2. Since integrable functions can be varied on arbitrary sets of measure zero without effecting the value of the integral we will avoid repetition by making two conventions. The statements "an ordinary point t" and "all ordinary points," will mean "an ordinary point not contained in a possible set of measure zero" and "all ordinary points except possibly a set of measure zero" respectively.

Theorem 3.1. An admissible function $u_o(t)$ affords a minimum to $I(u)$ in the class of all admissible functions u if and only if for almost all points t on $t^o \leq t \leq t^1$, the elements $(t, u_o(t))$ minimize $f(t, u)$ in the class of admissible elements (t, u).

Suppose that u_o minimizes $I(u)$ on the class of admissible elements u. Let $(s, u_o(s))$ be an ordinary point for f with $t^o < s < t^1$. Let $(s, \bar{u})$ be an admissible element and let

$$u(t) \qquad (s \leq t \leq s + \delta,\ \delta > 0)$$

be an admissible function such that $u(s) = \bar{u}$. Let

$$u(t) = u_o(t) \qquad (t^o \leq t < s;\ s + \delta < t \leq t^1)$$

The function u(t) so defined is admissible for all $\delta > 0$ and sufficiently small. Then

$$(3.3) \qquad 0 \leq \frac{I(u) - I(u_o)}{\delta} = \frac{1}{\delta} \int_s^{s+\delta} (f(t, u(t)) - f(t, u_o(t)))dt$$

Taking the limit as δ tends to zero we have that

$$(3.4) \qquad o \leq f(s, \bar{u}) - f(s, u_o)$$

for all ordinary points s. Thus, $(t, u_o(t))$ minimizes $f(t, u)$ at almost all points t on $t^o < t < t^1$.

The converse is immediate.

We next consider the case in which the admissible functions u are subjected to additional constraints

$$(3.5a) \qquad I_\alpha(u) \geq 0(\alpha = 1, \ldots, q'), \quad I_\alpha(u) = 0(\alpha = q' + 1, \ldots, q)$$

of the form

$$(3.5b) \qquad I_\alpha(u) = \int_{t^o}^{t^1} f_\alpha(t, u(t))dt - c_\alpha$$

Here it is assumed the functions f_α have the same properties as f, and admissible functions u(t) have the same properties with respect to f_α as before with respect to f.

Theorem 3.2. *Suppose that* u_o *minimizes* $I(u)$ *subject to conditions* (3.5). *Then there exist multipliers* $\lambda_o \geq 0, \lambda_1, \ldots, \lambda_q$ *such that for almost all points* t *on* $t^o \leq t \leq t^1$ *the element* $u_o(t)$ *minimizes the function*

$$F = \lambda_o f - \lambda_\alpha f_\alpha \tag{3.6}$$

on $U(t)$. *Further*, $\lambda_1 \geq 0, \ldots, \lambda_{q'} \geq 0$ *and the equality holds for all* $\alpha(\alpha = 1, \ldots, q')$ *for which* $J_\alpha(u_o) > 0$. *Conversely, if such a function exists with* $\lambda_o > 0$ *then* u_o *minimizes* $I(u)$ *on the class of admissible functions satisfying* (3.5).

Assume u_o is a minimizing function for I subject to (3.5). Set $f = f_o$ and let K' be the set of all $(m + 1)$ - dimensional vectors of the form

$$k_\rho = f_\rho(t, u) - f_\rho(t, u_o(t)) \qquad (\rho = 0, 1, \ldots, q) \tag{3.7}$$

where (t, u) is admissible and t is an ordinary point for u_o. Let K be the set of all vectors of the form

$$k = h_1 k_1 + \ldots + h_p k_p \qquad (h_1 \geq 0, \ldots, h_p \geq 0) \tag{3.8}$$

for any finite number of vectors $k_1, \ldots, k_p$ in K'. Then K is convex.

Lemma 3.1. *The cone* K *is a derived cone for* $I_o, \ldots, I_q$ *at* x_o.

Assume that the lemma has been proven. Then by Theorem 2.2 there is a linear functional

$$(3.9) \qquad L(k) = \lambda_o k_o - \lambda_\alpha k_\alpha$$

with properties given in Lemma 2.2.

We also have that $L(k) \geq 0$ for all k in K, hence for all k in K'. Setting

$$(3.10) \qquad F(t, u) = \lambda_o f_o - \lambda_\alpha f_\alpha$$

gives

$$(3.11) \qquad \lambda_o k_o - \lambda_\alpha k_\alpha = F(t, u) - F(t, u_o(t)) \geq 0$$

for all u in U(t) where t is an ordinary point on $t^o < t < t^1$. Therefore, u_o minimizes F(t, u) for almost all t on $t^o \leq t \leq t^1$.

The converse follows from the fact that

$$I(u) - \lambda_\alpha I_\alpha(u) \geq I(u_o) - \lambda_\alpha I_\alpha(u_o) = I(u_o)$$

Hence

$$I(u) \geq I(u_o) + \lambda_\alpha I_\alpha(u)$$

and by assumption, $\lambda_\alpha I_\alpha(u) \geq 0$.

To prove Lemma 3.1 it is sufficient to consider vectors of the form

$$k = h_1 k_1 + \dots + h_p k_p \qquad (h_\alpha \geq 0, \ \alpha = 1, \dots, p)$$

where

$$(3.12) \qquad k_{\rho j} = f_\rho(t_j, u_j) - f_\rho(t_j, u_o(t_j)) \qquad (\rho = 0, 1, \dots, q)$$

and t_j is an ordinary point for f_ρ.

We can assume that

$$t_1 \leq t_2 \leq \dots \leq t_p$$

Choose admissible functions

$$u_j(t) \qquad (t_j \leq t \leq t_j + \varepsilon)$$

such that $u_j(t_j) = u_j$. Choose $\delta > 0$ ($\delta < \varepsilon$) so small that

$$t_i \leq t \leq t_i + p\delta, \quad t_j \leq t \leq t_j + p\delta$$

are disjoint if $i \neq j$. Let $b = (b_1, \dots, b_p)$ be a point on an interval

$$B: \qquad 0 \leq b_j \leq \delta \qquad (j = 1, \dots, p)$$

Set:

$$c_o = 0, \ c_j = c_{j-1} + b_j \qquad (j = 1, \dots, p)$$

Let:

$u(b)$: $\quad u(t, b)$ $\qquad\qquad\qquad\qquad\qquad (t^0 \le t \le t^1)$

be the admissible function defined by

$$u(t, b) = u_j(t), \quad t_j + c_{j-1} \le t \le t + c_j$$

$$u(t, b) = u_0(t), \quad \text{otherwise}$$

Note that $u(0) = u_0$. The functions

$$(3.13) \qquad \phi_\rho(b) = I_\rho(u(b)) - I_\rho(u_0)$$

are continuous on B. The tangent cone C to B at $b = 0$ is defined by $h_j \ge 0$ $(j = 1, \ldots, p)$.

We will now show that $\phi_\rho(b)$ has a differential at $b = 0$ given by $k_{\rho j} h_j$. Note that $\phi_\rho(0) = 0$ and consider

$$(3.14) \qquad |\phi_\rho(b) - k_{\rho j} b_j| / |b|$$

$$\le \sum_{j=1}^{p} \left| \int_{t_j + c_{j-1}}^{t_j + c_j} (f_\rho(t, u_j(t)) - f_\rho(t, u_0(t)))dt - k_{\rho j} b_j \right| / |b|$$

where j is not summed in $k_{\rho j} b_j$ on the right hand side only. A typical term of the sum is majorized by

$$(3.15)\quad \int_{t_j}^{t_j+c_j} |f_\rho(t, u_j(t)) - f_\rho(t, u_o(t)$$

$$- (f_\rho(t_j, u_j) - f_\rho(t_j, u_o(t_j))|dt/|b|$$

$$\leq \int_{t_j}^{t_j+c_j} |f_\rho(t, u_j(t)) - f_\rho(t_j, u_j)|dt/|b|$$

$$+ \int_{t_j}^{t_j+c_j} |f_\rho(t, u_o(t)) - f_\rho(t_j, u_o(t_j))|dt/|b|$$

where both quantities in the second line tend to zero with $|b|$ since t_j is an ordinary point for $u_j(t)$ and $u_o(t)$.

Therefore:

$$\phi'_\rho(o, h) = k_{\rho j} h_j$$

and K is a derived cone as to be shown.

We now extend the results to problems involving parameters.

Theorem 3.3. Suppose that the functions $g_\rho(w)$, $T^o(w)$, $T^1(w)$ are continuous for w in a neighborhood N and have a differential at w_o in $(w, \ldots, w^r)$ - space. Suppose that an admissible function

$$u_o: \qquad u_o^k(t), w_o^q \qquad\qquad (T^o(w_o) \leq t \leq T(w_o))$$

affords a minimum to

$$I_o(u) = g_o(v) + \int_{T^o(v)}^{T^1(v)} f_o(t, u(t))dt$$

in a class of admissible functions

u: $\qquad u^k(t), v^\sigma \qquad\qquad (T^o(v) \leq t \leq T^1(v))$

v in N and satisfying

(3.16a) $\qquad I_a(u) \geq 0 \quad (a = 1, \ldots q'), \quad I_a = 0 \quad (a = q' + 1, \ldots, q)$

where

(3.16b) $\qquad I_a(u) = g_a(v) + \int_{T^o(v)}^{T^1(v)} f_a(t, u(t))dt$

There exist multipliers $\lambda_o \geq 0, \lambda_1, \ldots, \lambda_q$ not all zero such that for almost all t on

$$t^o \leq t \leq t^1, \quad t^o = T^o(v_o), \; t^1 = T^1(v_o)$$

the inequality

(3.17) $\qquad F(t, u) \geq F(t, u_o(t)), \quad F = \lambda_o f_o - \lambda_a f_a$

holds for admissible elements (t, u). For $a = 1, \ldots, q'$, $\lambda_a \geq 0$ where the equality holds if $I_a(u_o) > 0$. Moreover, if $T^o(v_o)$, $T^1(v_o)$ are ordinary points then

$$(3.18)\qquad \frac{\partial G}{\partial w^\sigma} + F^1 \frac{\partial T^1}{\partial w^\sigma} - F^0 \frac{\partial T^0}{\partial w^\sigma} = 0,\ G = \lambda_0 g_0 - \lambda_\alpha g_\alpha$$

$$(\alpha = 1, \ldots, r)$$

<u>holds at</u> $v = v_0$, <u>where</u>

$$F^s = F(t^s, u_0(t^s)) \qquad (s = 0, 1)$$

We construct the cone K' as in Theorem 3.2 and to it adjoin all vectors of the form

$$(3.19)\qquad k'_\rho = [\frac{\partial g_\rho}{\partial w^\sigma} + f^1_\rho \frac{\partial T^1}{\partial w^\sigma} - f^0_\rho \frac{\partial T^0}{\partial w^\sigma}]h'_\rho = k'_{\rho\sigma}h'_\sigma$$

$$(h'_\sigma \text{ arbitrary})$$

where the partial derivatives are evaluated at $w = w_0$ and

$$f^s_\rho = f_\rho(t^s, u_0(t^s)) \qquad (s = 0, 1).$$

Call the result K" and let K be the convex hull of K". K is a derived cone for $I_0, \ldots, I_q$ at x_0. For the proof we need only consider vectors of the form

$$k = h_1 k_1 + \ldots + h_p k_p + k' \qquad (h_j \geq 0,\ j = 1, \ldots, p)$$

Extend the definition of the functions $u^k_0(t)$ to a larger interval

$$t^0 - \varepsilon \leq t \leq t^1 + \varepsilon \qquad (\varepsilon > 0;\ t^s = T^s(w_0);\ s = 0, 1)$$

Choose δ as in Theorem 3.2 and such that $T^s(w)$ is in the extended interval if v is in a δ- neighborhood of v_o. If v is so restricted then the function

$$u(v, b): \qquad u(t, b) \qquad (T^0(v) \leq t \leq T^1(v))$$

as defined in Theorem 3.2 is admissible and

$$\phi_\rho(v, b) = J_\rho(u(v, b)) - J_\rho(u_o)$$

is continuous on

$$B: \qquad 0 \leq b_j \leq \delta, \quad -\delta \leq v^\sigma \leq \delta$$

for δ sufficiently small. Also, $\phi_\rho(v, b)$ has a differential at $(v, h) = (v_o, 0)$ given by

$$k_{\rho j}h_j + k'_{\rho\sigma}h'_\sigma$$

as we show by considering

(3.20)

$$|\phi_\rho(b, v) - \phi_\rho(0, v_o) - k_{\rho j}b_j - k'_{\rho\sigma}(v^\sigma - v^\sigma_o)|/(|b| + |v - v_o|)$$

$$\leq \sum_{j=1}^{p} |\int_{t_j+c_{j-1}}^{t_j+c_j} (f_\rho(t, u_j(t)) - f_\rho(t, u_o(t)))dt - k_{\rho j}b_j|/|b|$$

$$+ |g_\rho(v) - g_\rho(v_o) - \frac{\partial g_\rho}{\partial v_\sigma}(v^\sigma - v^\sigma_o)|/|v - v_o|$$

$$+ \sum_{s=0}^{s=1} | \int_{T^s(v_o)}^{T^s(v)} (f_\rho(t, u_o(t))dt -$$

$$- f_\rho(t^s, u_o(t^s))T^s_{v\sigma}(v^\sigma - v^\sigma_o)|/|v - v_o|$$

The first sum is as in (3.14) and the estimate (3.15) holds. The remaining terms tend to zero as v tends to v_o since $g_\alpha(v)$, $T^\alpha(v)$, $T^1(v)$ have a differential at v_o and the points $T^o(v_o)$, $T^1(v_o)$ are ordinary for $u_o(t)$. Hence, $\phi_\rho(b, v)$ has a differential at $(b, v) = (0, v_o)$.

Just as in Theorem 3.2 we see that K is a derived cone. Thus, there is a linear functional

$$L(k) = \lambda_o k_o - \lambda_\alpha k_\alpha$$

with which conclusions regarding the λ's and the inequalify for F follow as in Theorem 3.2.

Setting

$$G = \lambda_o g - \lambda_\alpha g_\alpha$$

gives for ordinary points t^o, t^1 that

$$dG + [F^s dT^s]_{s=0}^{s=0} \geq 0$$

for all dv , which gives (3.18).

Now assume that $F(t, u)$ has a bounded derivative with respect to t almost everywhere on R which is integrable along $u_o(t)$. Also, assume that if $(\bar{t}, \bar{u})$ is in U, except possible for $\bar{t}$ in T there is a constant $\delta > 0$ such that $(t, \bar{u})$ is in U for $\bar{t} - \delta \leq t \leq \bar{t} + \delta$. Then we have the following additional result.

Theorem 3.4. Under the assumptions of Theorem 3.3 and those just given, if

u_0: $u^k_0(t)$ $(t^0 \leq t \leq t^1)$

is an admissible function for which at almost all points t the inequality

$$F(t, u) \geq F(t, u_0(t))$$

holds for all u in $U(t)$ then

$$(3.18) \quad F(t, u_0(t)) = \int_{t^0}^{t} F_t(s, u_0(s))ds + c$$

for c a constant and for almost all t on $t^0 \leq t \leq t^1$. That is, $F(t, u_0(t))$ is equal almost everywhere to an absolutely continuous function of t.

Let Y be the class of all Lipschitzian functions $\eta(t)$ on $t^0 \leq t \leq t^1$ such that $\eta(t^0) = \eta(t^1) = 0$. Given $\eta(t)$ in Y, the function

$$t = s + \epsilon\eta(s) \qquad (t^0 \leq t \leq t^1)$$

is monotone for ϵ sufficiently small, hence has a solution

$$s = S(t, \epsilon) \qquad (t^0 \leq t \leq t^1)$$

such that $S(t^0, \epsilon) = t^0$, $S(t^1, \epsilon) = t^1$. By assumption

$$u_0(S(t, \epsilon)) \qquad (t^0 \leq t \leq t^1)$$

is admissible for say $|\epsilon| \leq \delta$. The function

$$J(\varepsilon) = \int_{t^0}^{t^1} F(t, u_o(S(t, \varepsilon)))dt$$

$$= \int_{t^0}^{t^1} F(s + \varepsilon\eta(s), u_o(s))(1 + \varepsilon\dot{\eta}(s))ds$$

has a minimum at $\varepsilon = 0$. Hence

$$(3.19) \qquad J'(0) = \int_{t^0}^{t^1} (M(t)\eta(t) + N(t)\dot{\eta}(t))dt = 0$$

where

$$(3.20) \qquad M(t) = F_t(t, u_o(t)), \; N(t) = F(t, u_o(t))$$

Let a and t be two ordinary points for $t^0 \leq a < t \leq t^1$ and let

$$\begin{aligned} \eta(s) &= 0 && (t^0 \leq s \leq a, \; t \leq s \leq t^1) \\ &= \frac{1}{\varepsilon}(s - a) && (a \leq s \leq a + \varepsilon) \\ &= 1 && (a + \varepsilon \leq s \leq t - \varepsilon) \\ &= \frac{1}{\varepsilon}(t - s) && (t - \varepsilon \leq s \leq t). \end{aligned}$$

for $0 < \varepsilon < (t - a)/2$. For this function

$$\int_a^t M(s)\eta(s)\,ds + \frac{1}{\varepsilon}\int_a^{a+\varepsilon} N(s)\,ds - \frac{1}{\varepsilon}\int_{t-\varepsilon}^{t} N(s)ds = 0.$$

Taking the limit as ε tends to zero gives, since a and t are ordinary points that

$$\int_a^t M(s)ds + N(a) = N(t).$$

Setting

$$\alpha = -\int_{t^0}^{a} M(s)ds + N(a)$$

gives the desired result for all ordinary points t.

We note in passing the following result known as the <u>fundamental lemma of the calculus of variations</u> was proved.

<u>Lemma 3.2</u>. <u>Let</u> M(t) <u>and</u> N(t) <u>be integrable on</u> $a \leq t \leq b$. <u>Let</u> Y <u>be the class of all Lipschitzian functions on</u> $a \leq t \leq b$ <u>which vanish at</u> a <u>and</u> b. <u>Then</u>

$$\int_{a}^{b} (M(t)\eta(t) + N(t)\dot{\eta}(t))dt = 0$$

<u>holds for all</u> η <u>in</u> Y <u>if and only if</u>

$$N(t) = \int_{t^0}^{t} M(s)ds + \alpha \qquad \text{a.e.}$$

<u>where</u> α <u>is a constant</u>.

4. EXTENSION TO PROBLEMS INVOLVING LINEAR EQUATIONS

In this section the results of Section 3 are extended to problems involving linear differential constraints.

Let $f(t, u)$, $f^i(t, u)$ be as in Section 3. Then for each admissible function

$$u: \quad u^k(t),\ v^\sigma \qquad (T^0(v) \leq t \leq T^1(v);\ k = 1, \ldots, m;\ \sigma = 1, \ldots r)$$

a solution

$$x: \quad x^i(t) \qquad (T^0(v) \leq t \leq T^1(v);\ i = 1, \ldots, n)$$

of the differential equations

$$(4.1) \qquad \dot{x}^i = A^i_j(t)x^j + f^i(t, u)$$

where $A^i_j (i = 1, \ldots, n, j = 1, \ldots, n)$ are integrable, is an absolutely continuous function which transfers an initial point $(t^0, x(t^0))$ to a final point $(t^1, x(t^1))$. Let A be the class of all admissible u for which there is a solution x which transfers the point $(t^0, x(t^0)) = (T^0(v), X^0(v))$ to a final point $(t^1, x(t^1)) = (T^1(v), X^1(v))$. The problem is to find a function u_0 in which minimizes a functional of the form

$$I(u) = g(v) + \int_{T^0(v)}^{T^1(v)} f(t, u(t))dt$$

relative to all u in A. Here it is assumed that $g(v)$, $T^s(v)$ have the same properties as in Theorem 3.3.

Theorem 4.1. Suppose that

u_o: $u_o^k(t), v_o^\sigma$ $(t^o \leq t \leq t^1)$

is a solution of the above problem with

x_o: $x^i(t)$ $(T^o(v_o) \leq t \leq T^1(v_o))$

as the corresponding solution of (4.1). Then there exists a constant $\lambda_o \geq 0$ **and a solution**

p: $p^i(t)$ $(t^o \leq t \leq t^1;\ i = 1, \ldots, n)$

of the differential equation

$$\dot{p}_i = - p_j A_i^j(t)$$

such that

(1) $\lambda o, p_1(t), \ldots, p_n(t)$ **do not vanish simultaneously at any point** t **on**

$$t^o \leq t \leq t^1,\ t^o = T^o(v_o),\ t^1 = T^1(v_o)$$

(2) **the function**

$$F(t, u) = \lambda_o f - p_i f^i$$

$$H(t,x,u,p) = -\lambda_o f + p_i A_j^i x^j + p_i f^i = p_i A_j^i x^j - F$$

have the property that for almost all points t **on** $t^o \leq t \leq t^1$

(4.2a) $F(t, u) \geq F(t, u_o(t))$

(4.2b) $H(t, x_o(t), u, p(t)) \leq H(t, x_o(t), u_o(t), p(t))$

for all u **in** U(t,.

(3) **The transversality condition**

(4.3a) $\lambda_o dg + [(F^s - p_j A_i^j X^s) dT^s + p_i dX^s]_{s=o}^{s=1} = 0$

(4.3b) $\lambda_o dg + [-H^s dT^s + p_i dX^s]_{s=o}^{s=1} = 0$

holds if t^s **is an ordinary point for** s=o,1.

(4) If in addition F and H have bounded derivatives with respect to t almost everywhere on R which are integrable along $x_o(t)$ and U has the property that if $(\bar{t}, \bar{u})$ are in U, except possibly for $\bar{t}$ in T there is a $\delta > 0$ such that $(t, \bar{u})$ is in U for $|t-\bar{t}| < \delta$ then also

$$(4.4a) \qquad F(t, u_o(t)) = \int_{t^o}^{t} F_t(s, u_o(s))ds + \alpha$$

$$(4.4b) \qquad H(t, x_o(t), u_o(t), p(t)) = \int_{t^o}^{t} H_t(s, x_o(s), u_o(s), p(s))ds + \beta$$

for almost all points t on $t^o \leq t \leq t^1$ where α and β are constants.

Let $P_{ij}(t^o \leq t \leq t^1;\ i, j = 1, \ldots, n)$ be solutions of the system

$$\dot{P}_{ij} = -P_{ik}A^k_j \qquad P_{ij}(t^o) = \delta_{ij}$$

These functions are absolutely continuous.

Set

$$f_i(t, u(t)) = P_{ij}(t)f^j(t, u(t))$$

Then if $u(t)$ is an admissible function and x is the corresponding solution to (4.1) we have that

$$(4.5 \qquad \frac{d}{dt} P_{ij}x^j = \dot{P}_{ij}x^j + P_{ij}\dot{x}^j = P_{ij}(\dot{x}^j - A^j_k x^k) = P_{ij}f^j$$

Hence

$$(4.6) \qquad \int_{t^o}^{t^1} f_i(t, u(t))dt = P_{ij}(t^1)x^j(t^1) - P_{ij}(t^o)x^j(t^o)$$

Now set

$$g_i(v) = -[P_{ij}(T^s)X^{js}]_{s=o}^{s=1}$$

(4.7) $$J_i(u) = g_i(v) + \int_{t^o}^{t^1} f_i(t, u(t))dt$$

For every admissible function u with corresponding solution x to (4.1) we have that

$$J_i(u) = [P_{ij}(t^s)(x^j(t^s) - X^{js})]_{s=o}^{s=1}$$

where the determinant of P_{ij} does not vanish. Therefore if $x^i(t^o) = X^{io}(v)$, $J_i(u) = 0$ if and only if $x^i(t^1) = X^{i1}(v)$, that is if and only if u is in A.

Thus the problem posed is equivalent to minimizing I(u) in the class of all u such that $J_i(u) = 0$. Hence by Theorem (3.3), there exist functions F and G of the form*

$$F = \lambda_o f + \lambda_i f_i, \quad G = \lambda_o g + \lambda_i g_i$$

where $\lambda_o \geq o$, $\lambda_1, \ldots, \lambda_n$ are constants not all zero such that

(4.7) $$F(t, u) \geq F(t, u_o(t))$$

for almost all points t. Setting

$$p_j(t) = -\lambda_i P_{ij}(t)$$

gives

$$F = \lambda_o f + \lambda_i P_{ij} f_j = \lambda_o f - p_j f^j$$

so that (4.2a) and hence (4.2b) follow from (4.7).

* We adopt the convention here of choosing the plus sign for multipliers resulting from differential constraints and the minus sign for multipliers resulting from integral constraints.

Now

$$dG = \lambda_o dg + \lambda_i dg_i = \lambda_o dg - [\lambda_i d(P_{ij}(t^s)X^{js})]_{s=o}^{s=1}$$
$$= \lambda_o dg - [\lambda_i \dot{P}_{ij} X^{js} dT^s + \lambda_i P_{ij} dX^{js}]_{s=o}^{s=1}$$
$$= \lambda_o dg - [-\lambda_i P_{ik} A_j^k X^{js} dT^s - p_j dX^{js}]_{s=o}^{s=1}$$
$$= \lambda_o dg - [p_k A_j^k X^{js} dT^s - p_j dX^{js}]_{s=o}^{s=1}$$

Writing the transversality condition (3.18) in differential form gives that

$$dG + [F(t^s, u_o(t)^s))\, dT^s]_{s=o}^{s=1} =$$
$$\lambda_o dg + [(F - p_k A_j^k X^{js}) dT^s + p_i dX^{js}]_{s=o}^{s=1}$$
$$= \lambda_o dg + [-H dT^s + p_i dX^{is}]_{s=o}^{s=1} = 0$$

so (4.3) holds.

Also (4.4a) follows immediately from Theorem (3.4). When $x = x_o(t)$, $u = u_o(t)$ and $p = p(t)$,

$$F_t = \lambda_o f_t(t, u) - p_i f_t^i - p_i f^i$$

$$\frac{d}{dt}(p_i A_j^i x^j) = -\dot{p}_i f^i + p_i \dot{A}_j^i x^i \qquad \text{a.e.}$$

Hence by definition of H

$$H_t = -F_t + \frac{d}{dt}(p_i A_j^i x^j) \qquad \text{a.e.}$$

This together with (4.4a) gives 4.4b).

Since further generalizations of this theorem are special cases of later results we proceed immediately to problems which also are non-linear in the space variables.

5. EXTENSIONS TO NON-LINEAR PROBLEMS

We now consider the case in which all functions appearing are permitted to be non-linear in all variables. Here still it is assumed that the choice of admissible control functions u does not depend explicitly on x.

The problem considered is that of finding in a class of arcs

x: $x^i(t)$, $u^k(t)$, v^σ $(t^0 \leq t \leq t^1;\ i=1,\ldots,n;\ k=1,\ldots,m;\ \sigma=1,\ldots,r)$

satisfying conditions of the form

$$(5.1) \qquad \dot{x}^i = f^i(t, x, u, v)$$

$$(5.2) \qquad t^s = T^s(v),\ x^i(t^s) = X^{is}(v)$$

$$(5.3) \qquad I_\gamma(x) \geq 0\ (\gamma = 1, \ldots, q'),\quad I_\gamma = 0\ (\gamma = q' + 1, \ldots, q)$$

where

$$(5.3b) \qquad I_\gamma(x) = g_\gamma(v) + \int_{t^0}^{t^1} f_\gamma(t, x(t), u(t), v)dt$$

one which minimizes

$$I(x) = g(v) + \int_{t^0}^{t^1} f(t, x(t), u(t), v)dt$$

We assume that $T^s(v)$, $X^{is}(v)$, $g(v)$, $g_\gamma(v)$ have continuous partial derivatives with respect to v^σ.

Let R be a region in (t, x, u, w) - space which is convex in x and w for each t. Set $f^o = f$, $f^{n+\gamma} = f_\gamma$ and assume the following hold on R for j = 0, ..., n + q:

a) $f^j(t, x, u, w)$ is defined

b) f^j is continuous in x, u, and w for fixed t and locally

(5.4) integrable in t for fixed x, u, and w

c) $f^j_{x^i}$, $f^j_{w^\sigma}$ (i = 1, ..., n; σ = 1, ..., r) exist and satisfy b)

d) for each function u(t) for which $f^j(t, x, u, w)$ is locally integrable for fixed x and w there is a locally integrable function S(t) such that

$$|f^j_{x^i}(t, x, u(t), w)| \leq S(t)$$

$$|f^j_{w^\sigma}(t, x, u(t), w)| \leq S(t)$$

holds for (t, x, u(t), w) on R.

That under these assumptions the system (5.1) has solutions with properties used in proving subsequent theorems is shown in Theorem A-1 of the appendix. It should be noted that any set of hypotheses for which the conclusions of Theorem A-1 hold may be substituted for (5.4). We now assume that f, f_α, f^i do not depend on w, a restriction that will be removed later.

Let R_o be a subregion of R with the property that if $(\bar{t}, \bar{x}, \bar{u})$ is in R_o, except possibly for $\bar{t}$ in T there is a function

(5.5) $u^k(t)$ $(\bar{t} - \delta \leq t \leq \bar{t} + \delta)$

for some $\delta > 0$ such that for $|t - \bar{t}| \leq \delta$ the following hold for $j = 0, 1, \ldots, n + q$:

i) $(t, x, u(t))$ is in R_o for $|x - \bar{x}| < \delta$

ii) for fixed x, $f^j(t, x, u(t))$, $f^j{}_{x^i}(t, x, u(t))$ are integrable functions of t

iii) $u(\bar{t}) = \bar{u}$

iv) $\bar{t}$ is an ordinary point for $f^j(t, \bar{x}, u(t))$.

An element (t, x, u) will be called <u>admissible</u> if it is in R_o. A function u(t) will be called <u>admissible</u> if it is defined for $t^o \leq t \leq t^1$, $f^i(t, x, u(t))$ is integrable for fixed x and solutions x(t) of

$$\dot{x}^i = f^i(t, x, u(t)) \qquad (t^o \leq t \leq t^1)$$

are such that $(t, x(t), u(t))$ are in R_o.

If x(t) is a solution of (5.1) for admissilbe $u = u(t)$ such that $x(\bar{t}) = x$ then iv) implies that $\bar{t}$ is an ordinary point for $f_j(t, x(t), u(t))$ since

$$\left|\frac{1}{\delta}\int_{\bar{t}}^{\bar{t}+\delta} (f(t, x(t), u(t)) - f(\bar{t}, \bar{x}, \bar{u}))dt\right|$$

$$\leq \left|\frac{1}{\delta}\int_{\bar{t}}^{\bar{t}+\delta} (f(t, x(t), u(t)) - f(t, \bar{x}, u(t)))dt\right|$$

$$+ \left|\frac{1}{\delta}\int_{\bar{t}}^{\bar{t}+\delta} (f(t, \bar{x}, u(t)) - f(\bar{t}, \bar{x}, \bar{u})\, dt\right|$$

The first term on the right hand side tends to zero with δ because of (5.3c) and the second also because of iv).

By a differentially admissible arc

$$x: \qquad x^i(t), \quad u^k(t), \quad v^\sigma \qquad\qquad (t^0 \leq t \leq t^1)$$

will be meant one such that u(t) is admissible. A totally admissible arc is a differentially admissible arc which satisfies (5.2) and (5.3).

Let A be the class of totally admissible arcs. Let x_o be an arc in A such that

$$x_o: \qquad x^i_o(t), \; u^k_o(t), \; v^\sigma_o \qquad\qquad (t^0 \leq t \leq t^1)$$

and

$$I(x) \geq I(x_o)$$

for all x in A satisfying

$$|x(t) - x_o(t)| < \varepsilon \qquad\qquad (t^0 \leq t \leq t^1)$$

for some $\varepsilon > 0$. We say that an arc of this type affords a strong relative minimum to I in the class of totally admissible arcs.

Theorem 5.1. Let x_o afford a strong relative minimum to I in the class of totally admissible arcs. Then there exist multipliers λ_o, λ_γ, $p_i(t)$ such that if we set

$$H(t, x, u, p, v) = p_i f^i + \lambda_\gamma f_\gamma - \lambda_o f$$

$$G(v) = \lambda_o g - \lambda_\gamma g_\gamma$$

then

(1) The multipliers $\lambda_o, \lambda_1, \ldots, \lambda_q$, are non-negative constants. Further, $\lambda_\gamma = 0$ for each $\gamma \geq 1$ for which $I_\gamma(x_o) > 0$. The multipliers $\lambda_o, \lambda_\gamma, p_i(t)$ do not vanish simultaneously at any point of $t^o \leq t \leq t^1$.

(2) There exist constants c_i such that on x_o

$$p_i = - \int_{t^o}^{t} H_{x^i} ds + c_i \tag{5.6}$$

(3) The inequality

$$H(t, x_o(t), u, p(t), v_o) \leq H(t, x_o(t), u_o(t), p(t), v_o) \tag{5.7}$$

holds for almost all t on $t^o \leq t \leq t^1$ and for all u such that (t, x_o, u, v_o) is in R_o.

(4) If t^o, t^1 are ordinary points, the transversality condition

$$dG + [-HdT^s + p_i(t)dX^{is}]_{s=0}^{s=1} - \int_{t^o}^{t^1} H_{v^\sigma} dv^\sigma dt = 0 \tag{5.8}$$

holds on x_o for all dv^σ.

We will first prove the theorem for the case where v appears only in $g(v)$, $g_\gamma(v)$, $X^{is}(v)$ in which case (5.8) is replaced by

$$\text{(5.9)} \qquad dG + [p_i(t^s)dx^{is}]_{s=0}^{s=1} = 0$$

and all totally admissible arcs are defined on the fixed interval $t^0 \leq t \leq t^1$.

Set $f_o = f$ and define

$$\text{(5.10)} \qquad A^i_{\ j}(t) = f^i_{\ x^j}(t, x_o(t), u_o(t)) \qquad (i, j = 1, \ldots, n)$$

$$B_{\beta j}(t) = f_{\beta x^j}(t, x_o(t), u_o(t)) \qquad (\beta = 0, 1, \ldots, q)$$

By hypothesis each $A^i_{\ j}$, $B_{\beta j}$ is integrable on $t^0 \leq t \leq t^1$. Let $q_\beta(t)$ be solutions on $t^0 \leq t \leq t^1$ of

$$\text{(5.11)} \qquad \dot{q}_{\beta i}(t) = - q_{\beta j}A^j_{\ i}(t) + B_{\beta i}(t), \qquad q_{\beta i}(t^0) = 0$$

Then $q_\beta(t)$ is an absolutely continuous function and $\dot{q}_\beta(t)$ satisfies (5.11) almost everywhere. If x is a differentially admissible arc

$$\frac{d}{dt}(q_{\beta i}x^i) = B_{\beta i}x^i + q_{\beta i}(f^i - A^i_{\ j}x^j) \qquad \text{a.e.}$$

Setting

$$\text{(5.12a)} \qquad F_\beta(t, x, u) = f_\beta + B_{\beta i}x^i + q_{\beta i}(f^i - A^i_{\ j}x^j)$$

$$\text{(5.12b)} \qquad G_\beta(w) = g_\beta - q_{\beta i}(t^1)x^{i1} + q_{\beta i}(t^0)x^{i0}$$

we note that by assumption F_β is integrable.

Then

$$I_\beta(x) = g_\beta - [q_{\beta i}x^i]_{t^o}^{t^1} + \int_{t^o}^{t^1}[f_\beta + \frac{d}{dt}(q_{\beta i}x^i)]dt$$

(5.13) $$= G_\beta(v) + \int_{t^o}^{t^1} F_\beta(t, x(t), u(t))dt$$

$$\equiv J_\beta(x)$$

for all totally admissible arcs x. Hence,

$$J_\beta(x) - J_\beta(x_o) = I_\beta(x) - I_\beta(x_o)$$

for all x joining $x_o(t^o)$ and $x_o(t^1)$. This representation has been chosen so that along x_o

(5.14) $$F_{\beta x^i} = 0$$

Let P_{ij} be solutions of the system

(5.15) $$\dot{P}_{ij} = -P_{ik}A^k_{\ j} \qquad P_{ij}(t^o) = \delta_{ij}$$

Then $P_{ij}(t)$ are absolutely continuous functions on $t^o \leq t \leq t^1$ and for any differentially admissible arc x

$$\frac{d}{dt} P_{ij}x^j = P_{ij}(f^j - A^j_{\ k}x^k) \qquad \text{a.e.}$$

Now set

$$F_{q+i}(t, x, u) = P_{ij}(f^j - A^j_k x^k)$$

(5.16) $$G_{q+i}(w) = -[P_{ij}(t^s)X^{is}]_{s=o}^{s=1}$$

$$J_{q+i}(x) = G_{q+i}(w) + \int_{t^o}^{t^1} F_{q+i}(t, x(t), u(t))dt.$$

For every totally admissible arc x

$$J_{q+i}(x) = G_{q+i}(w) + \int_{t^o}^{t^1} \frac{d}{dt}(P_{ij}x^j)dt$$

$$= [P_{ij}(t^s)(x^j(t^s) - X^{js})]_{s=o}^{s=1}$$

Therefore, if $x^j(t^o) = X^{jo}(w)$, $J_{q+i} = 0$ if and only if $x^j(t^1) = X^{j1}(w)$.

Let K' be the set of all vectors k_ρ of the form

(5.17) $k_\rho = F_\rho(t, x_o(t), u) - F_\rho(t, x_o(t), u_o(t))$, $(\rho=0,1,\ldots,q+n)$

for all ordinary points t and all u for which $(t, x_o(t), u)$ is in R_o. Let K" be obtained by adjoining to K' all vectors k'_ρ of the form

(5.18) $$k'_\rho = G_{\rho w^\sigma}(w_o)h'_\sigma$$

where h'_σ is arbitrary. Let K be the convex hull of K".

<u>Lemma 5.1</u>. <u>The cone</u> K <u>is a derived cone for the functions</u> $J_o, \ldots, J_{q+n}$ <u>at</u> x_o.

Assume the lemma has been proven. By Theorem 2.2 there is a linear functional

$$L(k) = \lambda_o k_o - \lambda_\gamma k_\gamma + \lambda_{q+i} k_{q+i} \qquad (\gamma = 1,\ldots,q;\ i=1,\ldots,n)$$

with properties given there. Setting

$$(5.19) \qquad F(t, u) = \lambda_o F_o - \lambda_\gamma F_\gamma + \lambda_{q+i} F_{q+i}$$

and using that every vector of the form (5.17) is in K it follows that

$$(5.20) \qquad F(t, x_o(t), u) \geq F(t, x_o(t), u_o(t))$$

for all ordinary points t and all u for which $(t, x_o(t), u)$ is in R_o. Setting

$$(5.21) \qquad H(t, x, u, p) = p_i f^i + \lambda_\gamma f_\gamma - \lambda_o f_o$$

where $p_i(t)$ is given by

$$(5.22) \qquad p_i(t) = -\lambda_{q+j} P_{ji} + \lambda_\gamma q_{\gamma i} - \lambda_o q_{oi} \qquad (i = 1, \ldots, n)$$

we have that

$$(5.23) \qquad H = (p_i A^i_j - \lambda_\gamma B_{\gamma j} + \lambda_o B_{oj}) x^j - F$$

Hence, along x_o

(5.24) $$H_{x^i} = p_j A^j{}_i - \lambda_\gamma B_{\gamma i} + \lambda_o B_{oi}$$

But from (5.22)

$$\dot{p}_i = - \lambda_{q+j}\dot{P}_{ji} + \lambda_\gamma \dot{q}_{\gamma i} - \lambda_o \dot{q}_{oi}$$

$$= (\lambda_{q+j}P_{jk} - \lambda_\gamma q_{\gamma k} + \lambda_o q_{ok})A^k{}_i + \lambda_\gamma B_{\gamma i} - \lambda_o B_{oi}$$

which with (5.24) gives (2). Applying (5.20) to (5.23) gives (3).

Now setting

$$G = \lambda_o g - \lambda_\gamma g_\gamma \qquad (\gamma = 1, \ldots, q)$$

gives for ordinary points t^o, t^1 that

(5.25) $$\lambda_o G_o - \lambda_\gamma G_\gamma + \lambda_{q+i} G_{q+i} = G + [p_i X^{is}]_{s=0}^{s=1}$$

But, since all vectors of the form (5.18) are in K

(5.26) $$dG + [p_i dX^{is}]_{s=0}^{s=1} \geq 0$$

for all dw^σ, hence the equality must hold.

Returning to Lemma 5.1, we need only prove it for vectors of the form

(5.27) $$k_\rho = h_j k_{\rho j} + h'_\sigma k'_{\rho\sigma} \qquad (j=1,\ldots,N;\ h_j \geq 0;\ h'_\sigma \text{ arbitrary})$$

where

(5.28) $$k'_{\rho\sigma} = G_{\rho w^\sigma}(w_o)$$

$$(5.29)\qquad k_{\rho j} = F_\rho(t_j, x_o(t_j), u_j) - F_\rho(t_j, x_o(t_j), u_o(t_j))$$

and $t_1, \ldots, t_n$ are ordinary points for $u_o(t)$. We may assume $t_1 \leq t_2 \leq \cdots \leq t_N$.

Choose $\varepsilon > 0$ and functions

$$u_j(t) \qquad (t_j - \varepsilon \leq t \leq t_j + \varepsilon)$$

as in (5.5). Choose $\delta > 0$ such that $N\delta < \varepsilon$ and the intervals are on $t^0 < t < t^1$ and disjoint if $i \neq j$. Let b_j be restricted to the set

$$B: \quad 0 \leq b_j \leq \delta \qquad (j = 1, \ldots, N)$$

and set

$$c_o = 0,\ c_j = c_{j-1} + b_j \qquad (j = 1, \ldots, N)$$

Set

$$\begin{aligned} u(t, b) &= u_j(t) && (t_j + c_{j-1} \leq t \leq t_j + c_j) \\ &= u_o(t) && \text{otherwise.} \end{aligned}$$

If δ' is sufficiently small then

$$(5.30)\qquad \dot{x}^i = f^i(t, x(t), u(t, b)),\ x^i(t^o) = X^{io}(v)$$

has an absolutely continuous solution

$$x(b, v):\quad x^i(t, b, v) \qquad (t^o \leq t \leq t^1;\ 0 \leq b_j \leq \delta')$$

where $x(t, 0) = x_o(t)$.

Lemma 5.2. The functions $x^i(t, b, v)$ have a bounded difference quotient with respect to b and v at $(b, v) = (o, v_o)$.

We will use the lemma here and prove it later. The functions

$$\phi_\rho(b, v) = J_\rho(x(b, v)) - J_\rho(x_o) \quad (\rho = 0, \ldots, q + n) \tag{5.31}$$

are continuous on $0 \leq b_j \leq \delta$, $|v - v_o| \leq \delta$. In what follows, to simplify notations we will assume $v_o = 0$. To show that $\phi_\rho(b, v)$ has a differential at $(0, v_o)$ given by $k_{\rho j}h_j + k_{\rho\sigma}h'_\sigma$, consider

$$\begin{aligned}
&|\phi_\rho(b, v) - \phi_\rho(0, 0) - k_{\rho j}b_j - k'_{\rho\sigma}v^\sigma|/(|b| + |v|) \\
&\leq |G_\rho(v) - G_\rho(0) - k'_{\rho\sigma}v^\sigma|/(|b| + |v|) \\
&+ \sum_{j=1}^{N} |\int_{t_j+c_{j-1}}^{t_j+c_j} (F_\rho(t, x(t, b, v), u_j(t)) \\
&\qquad - F_\rho(t, x_o(t), u_o(t)))dt \\
&\qquad - k_{\rho j}b_j|/(|b| + |v|) \\
&+ \sum_{j=1}^{N} |\int_{t_j+c_j}^{t_{j+1}+c_j} (F_\rho(t, x(t, b, v), u_o(t)) \\
&\qquad - F_\rho(t, x_o(t), u_o(t)))dt|/(|b| + |v|)
\end{aligned} \tag{5.32}$$

where $k_{\rho j}b_j$ in the second term on the right hand side is not summed on j and $t_{N+1}+c_N = t^1$. On the interval $(t_j + c_{j-1}, t_j+c_j)$,

$x(t, b, v)$ is the solution of

$$\dot{x}^i = f^i(t, x, u_j(t))$$

and on the interval $(t_j + c_j, t_{j+1}+c_j)$ is the solution of

$$\dot{x}^i = f^i(t, x, u_o(t))$$

where $x(t^o, b, v) = X^o(v)$ and initial conditions for each interval are chosen to be the final value of the solution on the previous interval.

We estimate a term from the first sum by

$$(5.33a)\quad \Big|\int_{t_j+c_{j-1}}^{t_j+c_j} (F_\rho(t, x(t, b, v), u_j(t))$$

$$- F_\rho(t, x_o(t), u_o(t)))dt - k_{\rho j}b_j\Big|/(|b| + |v|)$$

$$\leq \int_{t_j}^{t_j+c_j} |F_\rho(t, x(t, b, v), u_j(t))$$

$$- F_\rho(t, x_o(t), u_j(t))\,dt|/(|b| + |v|)$$

$$+ \int_{t_j}^{t_j+c_j} |F_\rho(t, x_o(t), u_j(t)) - F_\rho(t_j, x_o(t_j), u_j)|dt/|b|$$

$$+ \int_{t_j}^{t_j+c_j} |F_\rho(t, x_o(t), u_o(t)) - F_\rho(t_j, x_o(t_j), u_o(t_j))|dt/|b|$$

$$\equiv I_1 + I_2 + I_3$$

Next, we estimate a term from the second sum by

$$(5.33b)\quad \left|\int_{t_j+c_j}^{t_{j+1}+c_j}(F_\rho(t,\, x(t,\, b,\, v),\, u_o(t)) - F_\rho(t,\, x_o(t),\, u_o(t)))dt\right|/(|b| + |v|)$$

$$\leq \int_{t_j}^{t^1} \left|\int_0^1 F_{\rho x^i}(t,\, x_o(t) + \theta(x(t,\, b,\, v) - x_o(t)),\, u_o(t))d\theta\, (x^i(t,\, b,\, v) - x^i_o(t))/(|b| + |v|)\right|dt$$

$$\equiv I_4$$

where we use that $x(t, 0, 0) = x_o(t)$. Since by Lemma 5.2 $x(t, b, v)$ has a bounded difference quotient at $(b, v) = (0, v_o)$ and $F_{\rho x^i}(t, x_o(t), u_o(t)) = 0$ we have immediately that $I_4 = 0$ at $(0, v_o)$. Now I_1 can be represented by an integral of the same form as I_4 with an upper limit of $t_j + c_j$ and $u_o(t)$ replaced by $u_j(t)$. Estimating $F_{\rho x^i}(t, x_o(t), u_j(t))$ by the function $S(t)$ in (5.3d) which corresponds to $u_j(t)$ and using Lemma 5.2 again gives that $I_4 = 0$ at $(0, v_o)$. Also, $I_2 = I_3 = 0$ since t_j is an ordinary point for both $u_o(t)$, $u_j(t)$ and the integrands are independent of b.

Hence, $\phi(b)$ has a differential at $(b, v) = (0, v_o)$ and K is a derived cone as to be shown.

We now prove Lemma 5.2. For this purpose again let $v_o = 0$ and set $b_{N+\sigma} = v^\sigma$ $(\sigma = 1, \ldots, r)$. With this notation consider

$$(5.34) \qquad y^i(t, b) = (x^i(t, b) - x^i_o(t))/|b| \equiv \Delta x^i(t, b)/|b|$$

We note that

$$(5.35) \qquad y^i(t_j + c_j, b) = x^i(t_j + c_j, b) - x^i_o(t_j + c_j))/|b|$$

Now consider the system of equations

$$(5.36) \qquad \dot{y}^i(t, b) = (f^i(t, x(t, b), u_o(t)) - f^i(t, x_o(t), u_o(t)))/|b|$$

$$= \int_o^1 f^i_{x^k}(t, x_o(t) + \theta\Delta x(t, b), u_o(t))d\theta\Delta x^k(t, b)/|b|$$

$$\equiv A^i_k(t, b)y^k(t, b)$$

where for each $b \neq 0$ A^i_k are integrable functions of t which satisfy the hypotheses of Theorem A-1. Hence, there is a unique solution with initial conditions (5.35) which clearly is given by (5.34). We must now show y(t, b) has a limit at b = 0 which is bounded on $t_j \leq t \leq t_{j+1}$. By Theorem A-1 it will suffice to show that (5.35) has a limit at b = 0 which is bounded. We proceed inductively. Since $y^i(t^o, b) = X^{io}(v)$ is assumed to have a differential at $v = v_o$, $y^i(t^o, 0)$ is bounded. Hence, by Theorem A-1 so is the solution of (5.36) with this initial condition when evaluated at $t_1 + c_o$. Let

$$(5.37)\qquad z^i_j(t, b) = \frac{1}{|b|}\int_{t_j+c_{j-1}}^{t} (f^i(t, x_j(t), u_j(t)) - f^i(t, x_o(t), u_o(t)))dt, \quad (t_j + c_{j-1} \leq t \leq t_j + c_j)$$

Since for each $j (j = 1, \ldots, N)$, t_j is an ordinary point for $u_o(t)$ and $u_j(t)$, $z^i_j(t, b)$ is bounded at $b = 0$. Hence, so is

$$(5.38)\qquad y^i(t_1 + c_1, b) = y^i(t_1 + c_o, b) + z^i(t_1 + c_1, b)$$

In general, using the notation of Theorem A-1,

$$(5.39)\qquad y^i(t_{k+1} + c_{k+1}, b) = y^i(t_{k+1} + c_k; t_k + c_k, z^i_k(t_k + c_k, b)) + z^i_{k+1}(t_{k+1} + c_{k+1}, b).$$

By inductive hypotheses, $y^i(t_k + c_k, b)$ is bounded for $b = 0$. Hence, as above, so is the solution of (5.36) with this initial condition evaluated at $t_{k+1} + c_k$ given by the value of y^i on the right hand side. This concludes the proof of Lemma 5.2.

To remove the restrictions in the proof of Theorem 5.1 we use the technique in [6] of introducing new variables $x^o, y^1, \ldots, y^r, u^o, \bar{v}^1, \ldots, \bar{v}^r$ subject to the condition that $u^o > 0$. Now we consider the equivalent problem of finding in the class of arcs

$$x: \qquad x^i(t), y^\sigma(t), u^k(t), v^\sigma, \bar{v}^\sigma \qquad (t^o \leq t \leq t^1)$$

satisfying

$$(5.40) \qquad \dot{x}^0 = u^0,\ \dot{x}^i = f^i(x^0, x, u, y)u^0,\ \dot{y}^\sigma = 0$$

$$x^0(t^s) = T^{is}(v),\ x^i(t^s) = X^{is}(v)$$

$$y^\sigma(t^0) = v^\sigma,\ y^\sigma(t^1) = \bar{v}^\sigma$$

$$J_\gamma(x) \geq 0(\gamma = 1, \ldots, q'),\ J_\gamma(x) = 0(\gamma = q' + 1, \ldots, q)$$

where

$$J_\gamma(x) = g_\gamma(v) + \int_{t^0}^{t^1} f_\gamma(x_0, x, u, y)u^0 dt$$

one which minimizes

$$J(x) = g(v) + \int_{t^0}^{t^1} f(x_0, x, u, y)\, u^0 dt$$

For this statement of the problem, the minimizing arc x_0 is given by

$$x_0: \qquad x^0 = t,\ x^i{}_0(t),\ y^\sigma = v^\sigma,\ u^0 = 1,\ u^k{}_0(t),\ v^\sigma{}_0,\ \bar{v}^\sigma{}_0 = v^\sigma{}_0,$$

$$(t^0 \leq t \leq t^1)$$

Setting

$$\bar{H} = (p_0 + p_i f^i + \lambda_\gamma f_\gamma - \lambda_0 f)u_0 = p_0 u_0 + Hu_0$$

$$G = \lambda_0 g - \lambda_\gamma g_\gamma$$

and applying the conclusions of Theorem 5.1 in the form just proved gives

$$\dot{p}_0 = -\bar{H}_{x^0},\ p_i = -\bar{H}_{x^i},\ \dot{p}_{n+\sigma} = -\bar{H}_{y^\sigma} = -H_{v^\sigma} u^0$$

and since $u^o = 1$ maximizes

$$\bar{H}(t, x_o(t), u_o(t), u^o, v_o, p, q)$$

and is an interior point in u^o - space

$$\bar{H}_{u^o} = p_o + H = 0$$

so

$$p_o = - H$$

Therefore, the transversality condition becomes

$$dG + [p_o(t^s)dx^o(t^s) + p_i(t^s)dX^{is} + p_{n+\sigma}(t^s)dv]_{s=0}^{s=1} = 0$$

or

$$dG + [-HdT^s + p_i(t^s)dX^{is}]_{s=0}^{s=1} - \int_{t^o}^{t^1} H_{v^\sigma} dv^\sigma dt = 0$$

The inequality for H follows from the inequality for $\bar{H}$ and the proof of Theorem 5.1 is complete.

Now assume that $H(t, x, u, \hat{p}, v)$ has a bounded derivative with respect to t on R which is integrable along $x_o(t)$ and that the functions $u^k(t)$ in (5.5) can be chosen as constants. Then we have the following additional result.

Theorem 5.2. Under the assumptions of Theorem 5.1 and those just given, if, $u_o(t)$ is an admissible function for which at almost all points t on $(t^o \leq t \leq t^1)$ the inequality

$$H(t, x_o(t), u, p(t), v_o) \leq H(t, x_o(t), u_o(t), p(t), v_o)$$

<u>holds for all</u> u <u>such that</u> $(t, x_o(t), u, v_o)$ <u>is in</u> R_o <u>then</u>

(5.41) $$H(t, x_o(t), u_o(t), v_o) = \int_{t_o}^{t} H_t(s, x_o(s), u_o(s), v_o)ds + c$$

<u>for almost all</u> t. <u>That is</u>, H <u>is equal almost everywhere to an absolutely continuous function</u>.

By (5.23)

$$F = (p_i A^i{}_j - \lambda_\gamma B_{\gamma j} + \lambda_o B_{oj})x^j - H$$

Hence,

$$F_t = \frac{d}{dt}[(p_i A^i{}_j - \lambda_\gamma B_{\gamma j} + \lambda_o B_{oj})x^j] - H_t - H_{x^i}\dot{x}^i - H_{p_i}\dot{p}_i$$

But,

$$H_{x^i}\dot{x}^i - H_{p_i}\dot{p}_i = 0$$

and we now apply Theorem 3.4 to this representation of F.

We now consider a particular case in which the set R_o is defined by certain inequalities.

<u>Lemma 5.3</u>. <u>Suppose the functions</u>

(5.42) $$\phi_\alpha(t, u) \geq 0$$

<u>are continuous and have continuous partial derivatives with respect to</u> u <u>on some region</u> U. <u>Suppose further, that at each point</u> $(\bar{t}, \bar{u})$

of U the matrix

$$(5.43) \qquad \left(\frac{\partial \phi_\alpha}{\partial u^k}\right)$$

has rank h where $\alpha_1, \ldots, \alpha_h$ is the set of values of α for which $\phi_\alpha(t, u) = 0$. Then there exists continuous functions

$$\bar{u}^k(t) \qquad (\bar{t} - \delta \leq t \leq \bar{t} + \delta)$$

for some $\delta > 0$ such that $(t, u(t))$ is in U and $u(t) = u$.

A proof of this lemma is given in [6].

For (5.4b) we now substitute

(5.4') b) f^j is continuous in t, x, and u

while retaining the remaining assumptions. Then we have the following result.

Theorem 5.3. Suppose R_0 is the class of all elements (t, x, u) in R satisfying (5.42) under the hypotheses of Lemma 5.3. Then the conclusions of Theorem 5.1 hold.

The proof consists of applying Lemma 5.3 to obtain functions $u^k(t)$ which because of (5.4') clearly are the functions (5.5). Hence, the hypotheses of Theorem 5.1 are satisfied.

6 THE PROBLEM OF BOLZA AS FORMULATED BY HESTENES

In this section we consider a generalization of the problem of Bolza, investigated extensively by Hestenes [6], in which the variables (t, x, u) are jointly restricted.

The problem considered is that of minimizing I(x) subject to (5.1), (5.2), and (5.3). The functions f, f^i, f_γ will be assumed not only to satisfy (5.4), but also to have partial derivatives with respect to u which are locally integrable in t for fixed x, u, and v and continuous in x, u, and v for fixed t, for (t, x, u, v) in R.

We now distinguish one arc in R as

$$x_o: \quad x^i{}_o(t),\ u^i{}_o(t),\ v^\sigma{}_o \qquad (t^0 \leq t \leq t^1)$$

where $u_o(t)$ is a function defined on $t^0 \leq t \leq t^1$ for which

$$f^j(t,\ x,\ u_o(t),\ v_o) \qquad (j = 0,\ 1,\ \ldots,\ n + q)$$

satisfy (5.4) as functions of t and x and $x^i{}_o(t)$ is the solution of

$$\dot{x}^i = f^i(t,\ x,\ u_o(t),\ v_o),\ t^s = T^s(v_o) \qquad (s = 0,\ 1)$$

Let R_o be a subset of R containing x_o. Let N be a neighborhood of x_o in R and let M_o be the intersection of the projections of R_o and N on (t, x, v) - space. We assume that R_o has the property that there exists a set of functions $U^k_o(t, x, v)$ defined for every point in M_o such that

i) $U^k_o(t, x_o(t), v_o) = u_o(t)$

ii) $(t, x, U_o(t, x, v), v)$ is in R_o

iii) The functions

(6.1a) $r^k_i = U^k_{ox^i}(t, x, v)$

$r^k_{n+\sigma} = U^k_{ov^\sigma}(t, x, v) \qquad (\sigma = 1, \ldots, r)$

(6.1b) $A^i_j = f^i_{x^j} + f^i_{u^k} r^k_j \qquad (j = 1, \ldots, n)$

$B_{\beta j} = f_{\beta x^j} + f_{\beta u^k} r^k_j \qquad (\beta = 0, \ldots, q)$

(6.1c) $A^i_{n+\sigma} = f^i_{v^\sigma} + f^i_{u^k} r^k_{n+\sigma}$

$B_{\beta(n+\sigma)} = f_{\beta v^\sigma} + f_{\beta u^k} r^k_{n+\sigma}$

exist, and together with f, f^i, f_γ - where the unspecified arguments are all $(t, x, U_o(t, x, v), v)$ - are integrable in t for fixed x and v and are continuous in x and v for fixed t, for all (t, x, v) in M_o except possibly in t in T. Further, we require that for every point $(\bar{t}, \bar{x}, \bar{u}, \bar{v})$ in R_o, except possibly for $\bar{t}$ in T, there is a neighborhood M of $(\bar{t}, \bar{x}, \bar{v})$ in (t, x, v) - space and a set of functions U(t, x, v) which have property ii) and have property iii) on M. Also, $U^k(\bar{t}, \bar{x}, \bar{v}) = \bar{u}^k$, and $\bar{t}$ is a regular point for U^k.

The definitions of admissible function, differentially and totally admissible arcs remain unchanged with R_o as here defined.

Theorem 6.1. The conclusions of Theorem 5.1 hold under these assumptions, providing that H_{x^i}, H_{v^σ} are replaced by $H_{x^i} + H_{\psi^k} r^k_i$, $H_{v^\sigma} + H_{\psi^k} r^k_{n+\sigma}$ in equations (5.6), (5.8) respectively.

As in the proof of Theorem 5.1 we will set $x^{n+\sigma} = v$. In this case $x^{n+\sigma}$ satisfies the differential equations

$$\dot{x}^{n+\sigma} = 0, \quad x^{n+\sigma}(t^o) = v^\sigma$$

Hence, for convenience we can assume that f, f^i, f_γ, and the set of function U^k_o associated with x_o are independent of v. Under this assumption the functions defined in (6.1c) do not appear.

With the functions defined in (6.1b) evaluated at $(t, x_o(t), u_o(t))$ and substituted for those in (5.10), the proof in Theorem 5.1 goes through without change to equation (5.19). Here, since U_o depends on x we must write

(6.3) $$F(t, x, u) = \lambda_o F_o - \lambda_\gamma F_\gamma + \lambda_{q+i} F_{q+i}$$

again obtaining that

$$(6.4)\qquad F(t, x_o(t), u) \geq F(t, x_o(t), u_o(t))$$

for almost all t and all u such that $(t, x_o(t), u)$ is in R_o. Defining H as in (5.21) gives

$$(6.5)\qquad H = (p_i A^i{}_j - \lambda_\gamma B_{\gamma_j} + \lambda_o B_{oj})x^j - F(t, x, u)$$

and since

$$(6.6)\qquad F_{\rho x^i} + F_{\rho u^k} r^k{}_i = 0 \qquad (\rho = 0, \ldots, q + n)$$

the remainder of the proof is the same, except for the changes noted in the statement of the theorem.

<u>Lemma 6.1</u>. <u>The cone</u> K <u>is a derived cone for</u> $J_o, \ldots, J_{q+n}$ <u>at</u> x_o.

We obtain points $t_1, \ldots, t_N$ as in Lemma 5.1. With each point $(t_j, x_o(t_j), u_j)$ we associate the set of functions $U^k{}_j(t, x)$ which by hypothesis exist for every point in R_o. Define $c_o, \ldots, c_N$ as in Lemma 5.1 and set

$$(6.7)\qquad \begin{aligned} U(t, x, b) &= U_j(t, x) && (t_j + c_{j-1} \leq t \leq t_j + c_j) \\ &= U_o(t, x) && \text{otherwise} \end{aligned}$$

For δ' sufficiently small the system

$$\dot{x}^i = f^i(t, x(t), U(t, x, b)),\ x^i(t^o) = X^{io}(v)$$

has an absolutely continuous solution

$x(b, v)$: $x^i(t, b, v)$, $(t^0 \leq t \leq t^1;\ 0 \leq b_j \leq \delta';\ |v - v_0| \leq \delta')$

with bounded difference quotient with respect to b and v at $(b, v) = (0, v_0)$ by Lemma 5.2. Define $\phi_\rho(b, v)$ as in 5.31. To prove $\phi_\rho(b, v)$ has a differential at $(0, v_0)$ we substitute $U_j(t, x)$ for $u_j(t)$ in (5.18). Since the resulting functions F_ρ again only depend on t and x and (6.6) holds, the proof is quite similar to that in 5.1. The removal of restrictions is also the same.

We now consider a particular case in which there are additional constraints of the form

$$(6.7) \qquad \phi_\alpha(t, x, u, v) \geq 0 \qquad (\alpha = 1, \ldots, p)$$

when ϕ_α, $\phi_{\alpha_{u^k}}$ are continuous. Given a point $(\bar{t}, \bar{x}, \bar{u}, \bar{v})$ for which (6.7) holds, let $\alpha_1, \ldots, \alpha_h$ be the set of values of for which $\phi_\alpha(\bar{t}, \bar{x}, \bar{u}, \bar{v}) = 0$. We assume the matrix

$$(6.8) \qquad \left(\frac{\partial\phi_\alpha}{\partial u^k}\right) \qquad (\alpha = \alpha_1, \ldots, \alpha_h)$$

has rank h at $(\bar{t}, \bar{x}, \bar{u}, \bar{v})$.

We now adopt Hestenes' definition of a <u>broken field</u>* as

* We note that this definition of a field differs from that usually made in variational theory. See Bliss [9].

a region in (t, x, v) - space and a set of functions $U^k(t, x, v)$ such that (1) the functions $U^k(t, x, v)$, $U^k_{x_j}(t, x, v)$ are continuous in (t, x, v) except for a finite set of values t , ..., t_N, at which points they have left and right hand limits which are continuous in x and v and (2) the elements $(t, x, U(t, x, v), v)$ are in R_o for all (t, x, v) in $\mathfrak{F}$. By an *arc* in $\mathfrak{F}$ we will mean a solution

$$x: \qquad x^i(t), u^k(t), v^\sigma \qquad (t^0 \leq t \leq t^1)$$

of the system

$$\dot{x}^i = f^i(t, x, U(t, x, v), v), \; u^k = U^k(t, x, v)$$

For the remainder of Section 6 we will assume, in addition to the hypotheses of Theorem 6.1, that all functions appearing are continuous and have continuous partial derivatives with respect to x^i, v^σ and u^k for (t, x, u, v) in R, that R_o is defined by (6.7) and that all admissible functions $u^k(t)$ are piecewise continuous. We note that all but a finite number of points t are ordinary points.

Lemma 6.2. *A differentially admissible arc*

$$x_o: \qquad x^i_o(t), u^k_o(t), v^\sigma_o \qquad (t^0 \leq t \leq t^1)$$

is an arc of a broken field $\mathfrak{F}_o$ *with control functions* $U_o(t, x, v)$. *Moreover, if* $(\bar{t}, \bar{x}, \bar{u}, \bar{v})$ *is in* R_o, *there is a broken field* $\mathfrak{F}_1$ *defined over a neighborhood of* $(\bar{t}, \bar{x}, \bar{v})$ *with control functions*

$\overline{U}(t, x, v)$ *such that*

$$U^k(\overline{t}, \overline{x}, \overline{v}) = \overline{u}^k$$

This is a trivial modification to Lemma 16 in [6].

Theorem 6.2. With R_o *given by* (6.7), *the conclusions of Theorem 6.1 hold. Moreover, they hold at every point* t *on* $t^o \leq t \leq t^1$.

The first conclusion follows immediately from Lemma 6.2. The second follows by continuity.

Theorem 6.3. If the function H *appearing in Theorem 5.1 is replaced by*

$$\overline{H}(t, x, u, p, v, \mu) = H + \mu_\alpha \phi_\alpha$$

the conclusions of Theorem 5.1 hold for all t *on* $t^o \leq t \leq t^1$ *where in* (5.7) $\overline{H}$ *is evaluated at* $\mu = 0$. *In addition to* (5.6), $\overline{H}_{u^k} = 0$ $(t^o \leq t \leq t^1)$. *Further, the multipliers* $\mu_\alpha(t)$ *are non-negative functions, continuous except possibly at discontinuties of* $u_o(t)$. *Also* $\mu_\alpha(t) = 0$ $(\alpha = 1, \ldots, p)$ *when* $\phi_\alpha(t, x_o(t), u_o(t)) > 0$.

Since $H(t, x_o(t), u, p, v_o)$ is maximuzed by $u_o(t)$ for each t which is a point of continuity of $u_o(t)$ and all u satisfying (6.7), by Theorem 2.3 there exist multipliers $\mu_1(t), \ldots, \mu_p(t)$ such that

(6.9) $\quad \bar{H}_{u^k} = H_{u^k} + \mu_\alpha(t)\phi_{\alpha j}k = 0.$

The continuity properties of $\mu_\alpha(t)$ follow from those of H_{u^k}, $\phi_{\alpha u^k}$.

In [6] it is shown that under the assumptions on the functions appearing in (6.7), $U_o(t, x, w)$ can be chosen so that for some $\delta > 0$,

$$\phi_\alpha(t, U_o(t, x, w), w) = 0$$

for $|x - x_o| < \delta$. Hence for $u = U_o$

(6.10) $\quad \phi_{\alpha x^i} + \phi_{\alpha u^k} r^k{}_i = 0$

(6.11) $\quad \phi_{\alpha w^\sigma} + \phi_{\alpha u^k} r^k{}_{n+\sigma} = 0.$

But by Theorem 6.2

$$\dot{p}_i = -(H_{x^i} + H_{u^k} r^k{}_i).$$

Using (6.9) and (6.10)

$$\begin{aligned}\dot{p}_i &= -H_{x^i} + \mu_\alpha(t)\phi_{\alpha u^k} r^k{}_i \\ &= -H_{x^i} - \mu_\alpha(t)\phi_{\alpha x^i} \\ &= -\bar{H}_{x^i}\end{aligned}$$

The transversality condition follows similarly from (6.11).

7. AN EXTENSION OF THE PREVIOUS PROBLEM

In this section a generalization of a problem involving bounded phase coordinates considered by Gamkrelidze [4] and Berkovitz [8] is obtained as a corollary to Theorem 6.2. As a consequence more extensive results are obtained than in either reference. We here use a simplification of a devise employed in [4] to obtain certain results.

Consider the problem of finding in a class of arcs

$$x: \quad x^i(t),\ u^k(t),\ w^\sigma \qquad (t^0 \leq t \leq t^1;\ i = 1, \ldots, n;\ k = 1, \ldots, m;\ \sigma = 1, \ldots, r)$$

satisfying conditions of the form

$$(7.1) \qquad x^i = f^i(t, x, u, w)$$

$$(7.2) \qquad \psi_\alpha(t, x, w) \geq 0 \qquad (\alpha = 1, \ldots, s)$$

$$(7.3) \qquad \phi_\beta(t, x, u, w) \geq 0 \qquad (\beta = 1, \ldots, p)$$

$$(7.4) \qquad t^s = T^s(w), \quad x^i(t^s) = X^{is}(w)$$

$$(7.5) \qquad I_\gamma = g_\gamma(w) + \int_{t^0}^{t^1} f_\gamma(t, x(t), u(t), w) \geq 0 \qquad (\gamma = 1, \ldots, q)$$

one which minimizes

$$I = g(w) + \int_{t^0}^{t^1} f(t, x(t), u(t), w)dt$$

We assume that all functions of (t, x, u, v) are of class C' in a region R, that $\psi_\alpha(t, x, v)$ is of class C'' and that grad $\psi_\alpha(t, x, v)$ does not vanish. Let R_0 be the set of all elements (t, x, u, v) satisfying (7.2) and (7.3). Let S be the set of elements in (t, x, v) - space which satisfy (7.2) with boundary S^*. Assume that if (t, x, v) is in S^* then the vectors grad $\psi_\alpha(t, x, v)$ are linearly independent for all α such that $\psi_\alpha(t, x, v) = 0$.

Define an <u>admissible control</u> $u(t)$ to be a piecewise continuous function such that if $x(t)$ is a solution of (7.1) for $u = u(t)$ then $(t, x(t), u(t)\ v)$ is in R_0 for $t^0 \leq t \leq t^1$. The corresponding arc

$$x: \qquad x^i(t),\ u^k(t),\ v^\sigma \qquad (t^0 \leq t \leq t^1)$$

will be called <u>differentially admissible</u> and if (7.4) and (7.5) are also satisfied, <u>totally admissible</u>.

Assume an arc x_0 affords a strong relative minimum to I in the class of totally admissible arcs. Appropriate necessary conditions for a special problem will be first derived, then stated in Theorem 7.1. These results will then be extended to a more general problem in Theorem 7.2.

Without loss of generality we can assume that only X^{is}, g, g_α depend on v since the removal of this restriction is just as in Theorem 5.1.

Suppose that for some $a(1 \leq a \leq s)$, $\Psi_a(t, x_o(t)) = 0$ for $t_1 \leq t \leq t_2$. Then since $\Psi_a(t, x)$ is of class C", we can choose a neighborhood M_a of $(t, x_o(t))$ and a vector $N_a(t, x)$ of class C' on M_a such that the inner product

(7.6) $\quad (N_a(t, x), \text{grad } \Psi_a(t, x)) \geq 0 \qquad ((t, x) \text{ in } M_a).$

The magnitude of $N_a(t, x)$ is at our disposal. For convenience we denote by $y = (y^o, y^1, \ldots, y^n)$, the $(n + 1)$ - dimentional vector with components $y^o = t$, $y^i = x^i(i = 1, \ldots, n)$. Then for $|\nu| \leq \delta'$ and δ' sufficiently small, if y satisfies $\Psi_a(y + \nu N_a(y)) = 0$, then y is in M_a. Further, if $\nu \leq 0$ the definition of $N_a(y)$ assures that y is in S. Define $f^o = 1$ and set

$$\Phi_{p+\alpha}(t, x, u, \nu) = \Psi_{\alpha y^i}(y + \nu N(y))f^i(t, x, u).$$
$$(\alpha = 1, \ldots, s)$$

Then for any point (t, x) in S* we must have that

(7.7) $\quad \Phi_{p+\alpha}(t, x, u, 0) \geq 0$

for all α such that $\Psi_\alpha (t, x) = 0$.

We first consider the case where $(t, x_o(t))$ is in S^* for $t^o \leq t \leq t_1$ and is interior to S for $t_1 < t \leq t^1$. By a suitable change in parameter we may take $t^o = 0$, $t^1 = 1$. We assume that on $0 \leq t \leq t_1$, the matrix

$$(7.8) \qquad \left(\frac{\partial \phi_\rho}{\partial u^k}\right)$$

has maximum rank for all subscripts $\rho = p + \alpha_h$ such that $\phi_{p+\alpha}(t, x, u, \nu) = 0$ and all subscripts $\rho = \beta_k$ such that $\phi_\beta(t, x, u) = 0$.

Set

$$\phi_{p+\alpha}(t, x_o(t), u_o(t), 0) = \phi_{p+\alpha}(t) \qquad (\alpha = 1, \ldots, s)$$

Then the system of equations

$$\phi_{p+\alpha}(t,x,u,\nu) - \phi_{p+\alpha}(t) = 0 \qquad (\alpha = \alpha_1, \alpha_2, \ldots)$$

$$\phi_\beta(t, x, u) = 0 \qquad (\beta = \beta_1, \beta_2, \ldots)$$

has a solution for each point t on the extended interval $0 \leq t \leq t_2 + \epsilon$ for some $\epsilon > 0$. By Lemma 6.2, x_o is an arc in a broken field $\mathfrak{F}_1$ on the larger interval. For $\nu \leq 0$, an arc $x(t)$ in $\mathfrak{F}_1$ is in R_o provided that $\psi_\alpha(0, x(0)) \geq 0$. Hence, in addition to (7.7) we must require that

$$(7.9) \qquad I_{q+\alpha} = \psi_\alpha(0, X^o(\nu)) \geq 0$$

$$I_{q+s+1} = -\nu \geq 0$$

Next we assume that on $t_1 \leq t \leq 1$ the matrix

$$(7.10) \qquad \left(\frac{\partial \phi_\beta}{\partial u^k}\right)$$

has maximum rank for all subscripts β_k as above. Then, again by Lemma 6.2, x_o is an arc in a broken field $\mathfrak{F}_2$ on $t_1 \leq t \leq 1$.

We extend the arcs in $\mathfrak{F}_1 \cap R_o$ to the interval $0 \leq t \leq 1$ using those in $\mathfrak{F}_2$ for $t + \epsilon \leq t \leq t^1$. The resulting arcs lie in R_o and the corresponding control functions $U(t, x)$ satisfy the hypotheses of Theorem 6.2 since we have at most introduced a discontinuity at $t_1 + \epsilon$. Hence, the conclusions of Theorem 6.2 hold, where because of (7.9) the function G in the transversality condition corresponding to (5.9) becomes

$$(7.11) \qquad G = \lambda_o g - \lambda_\gamma g_\gamma - \lambda_{q+\alpha}\Psi_\alpha(0, X^o(w)) + \lambda_{q+s+1}\nu = 0$$

The function of (t, x, u, ν) given by

$$F(t, x_o(t), u, 0)$$

corresponding to (6.4) is minimized by $u_o(t)$ for each t which is a point of continuity of $u_o(t)$ and for all u such that $(t, x_o(t), u)$ is in R_o. Hence, by Theorem 2.3 there exist multipliers $\mu_1(t), \ldots, \mu_{p+s}(t)$ all non-negative such that

$$F_{u^k}(t, x_o(t), u, 0) = \mu_\rho \phi_{\rho u^k} \qquad (\rho = 1, \ldots, p + s)$$

for all such u. For t such that $\phi_\beta(t, x_o(t), u_o(t)) > 0$,

$\mu_\beta = 0 (\beta = 1, \ldots, p)$. Further, since the sign of $\phi_{p+\alpha}(t, x_o(t), u, 0)$ is abritrary when $\psi_\alpha(t, x_o(t)) > 0$, $\mu_{p+\alpha}(t) = 0$ for such points. Set

(7.12a) $$v^k_\rho(t) = \phi_{\rho u^k}(t, x_o(t), u_o(t)), \qquad (\rho = 1, \ldots, p)$$

$$\phi_\rho(t) = \phi_\rho(t, x_o(t), u_o(t))$$

(7.12b) $$v^k_\rho = \phi_{\rho u^k}(t, x_o(t), u_o(t), 0), (\rho = p + 1, \ldots, p + s)$$

$$\phi_\rho(t) = \phi_\rho(t, x_o(t), u_o(t), 0)$$

and let $\rho_1, \ldots, \rho_h$ be the values of ρ for which $\phi_\rho(t_o) = 0$ where t_o is a point of continuity of $u_o(t)$. If $\rho \neq \rho_j$ $(j = 1, \ldots, h)$ then $\mu_\rho = 0$. Hence,

(7.13) $$F_{u^k}(t, x_o(t), u_o(t), 0)v^k_\beta = \mu_\rho v^k_\rho(t) v^k_\beta(t)$$

$$(\beta, \rho = \rho_1, \ldots, \rho_h)$$

in some neighborhood of t_o. Since the determinant of the coefficients of $\mu_{\rho_1}, \ldots, \mu_{\rho_h}$ is non-zero by hypothesis, the functions $\mu_{\rho_1}(t), \ldots, \mu_{\rho_h}(t)$ are continuous and have at least as many continuous derivatives as the lesser of F_{u^k}, $\phi_{\rho u^k}$. In particular, we will assume that they are of class C' and thus that the discontinuities of $\mu_\rho(t)$ and $\mu'_\rho(t)$ $(\rho = 1, \ldots, p + s)$ are at most those of $u_o(t)$.

Now set

(7.14) $$H(t, x, u, p, \mu, v) = p_i f^i + \lambda_\gamma f_\gamma - \lambda_o f + \mu_\rho \phi_\rho$$

where $p_i(t)$ is given by (5.21). Using (7.13) we have that

(7.15) $$H_{u^k}(t, x_o(t), u_o(t), p(t), \mu, 0) = 0$$

and as before

(7.16) $$H(t, x_o(t), u(t), p(t),0,0) \leq H(t, x_o(t), u_o(t), p(t),0,0).$$

The fact that for this H

$$\dot{p}_i = - H_{x^i}$$

along x_o is obtained by differentiating the identity

(7.17) $$H(t, x, U_o(t, x, 0), p, \mu, 0)$$

$$= (p_i A^i_j - \lambda_\gamma B_{\gamma j} + \lambda_o B_{oj})x^j - F(t, x, U_o(t, x, 0))$$

$$+ \mu_\sigma \phi_\rho(t, x, U_o(t, x, 0))$$

with respect to x_i and using the definition of p_i.

Now, since $(t, x_o(t))$ is in S^* for $0 \leq t \leq t_1$, for at least one value of α, $(p + 1 \leq \alpha \leq p + s)$, $\phi_\alpha(t, x_o(t), u_o(t), 0) = 0$ at each point t. For simplicity assume that $\alpha = a$, a constant. Take $N_\alpha(t, x) = 0$ $(\alpha \neq a)$ and $N_a(t, x) = 0$ at 0 ant t_1. By considering ν to be an additional parameter w, noting that ν appears in H only in ϕ_a and applying the transversality condition (5.8), we have that

(7.18) $\lambda_{q+s+1} - \int_0^1 H_\nu dt = \lambda_{q+s+1} - \int_0^{t_1} \mu_a \phi_{a\nu} dt = 0, \ \lambda_{q+s+1} \geq 0$

along x_0. To evaluate $\phi_{a\nu}$ along x_0 we define y as before and set $\eta^i = y^i + \nu N_a^i(y)$. Then

(7.19)
$$\phi_a(t, x, y, \nu) = \psi_{ay^i}(\eta)\dot{y}^i = \psi_{a\eta^j}\eta^j_{y^i}\dot{y}^i$$
$$= \psi_{a\eta^i}\dot{y}^i + \nu\psi_{a\eta^j}N^j_{ay^i}\dot{y}^i$$

where a is not summed. Hence, along x_0

$$\phi_{a\nu}(t, x_0(t), u_0(t), 0) = [\psi_{a\eta^i\eta^j}\eta^j_\nu\dot{y}^i]_{\nu=0} + \psi_{ay^j}N^j_{ay^i}\dot{y}^i$$
$$= \psi_{ay^iy^j}N^j_a\dot{y}^i + \psi_{ay^j}N^j_{ay^i}\dot{y}^i$$

where $\dot{y}^0 = 1$, $\dot{y}^i = f^i(t, x_0(t), u_0(t))$ $(i = 1, \ldots, n)$. But also

(7.20)
$$\frac{d}{dt}(\psi_{ay^i}N^i_a(y)) = \psi_{ay^iy^j}\dot{y}^jN^i_a + \psi_{ay^i}N^i_{ay^j}\dot{y}^j = \phi_{a\nu}$$

along x_0. Integrating (7.18) by parts we have that

$$-\int_{t_1}^{t_2}\mu_a\phi_{a\nu}dt = -\psi_{ay^i}(y)N^i_a(y)\Big|_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2}\frac{d\mu_a}{dt}\psi_{ay^i}N^i_a dt \leq 0$$

The boundary terms vanish by choice of $N_a(t, x)$. Since also $\psi_{ay^i}N^i_a \geq 0$ by (7.6), $N_a(t, x)$ is arbitrary on $0 < t < t_1$, and $\frac{d\mu_a}{dt}$ is continuous except possibly at discontinuities of $u_0(t)$ we have that $\frac{d\mu_a}{dt} \leq 0$ except at these points. With this we have proved

Theorem 7.1. Under the foregoing assumptions, let $x_o(t)$ afford a strong relative minimum to I in the class of totally admissible arcs. There exist multipliers λ_o, λ_γ, $\lambda_{q+\alpha}$, $u(t)$, $p_i(t)$ such that if we set

(7.21) $$H(t, x, p, w, \mu, \nu) = p_i f^i + \lambda_\gamma f_\gamma - \lambda_o f + \mu_\rho \phi_\rho$$

(7.22) $$G = \lambda_o g - \lambda_\gamma g_\gamma - \lambda_{q+\alpha} \Psi_\alpha(t^o, X^o(w))$$

the following conditions hold.

(1) The multipliers λ_o, λ_1, ..., λ_{q+s} are non-negative constants. Moreover, $\lambda_\gamma = 0$ for each $\gamma (1 \leq \gamma \leq q + s)$ for which $J_\gamma(x_o) > 0$. The multipliers λ_o, λ_γ, $p_i(t)$ do not vanish simultaneously. The $\mu_\rho(t)$ (ρ=1,...,p+s) are non-negative functions continuous on $t^o \leq t \leq t^1$ except possibly at discontinuities of $u_o(t)$. Also, when $\phi_\beta(t, x_o(t), u_o(t)) > 0$, $\mu_\beta(t) = 0$ (β = 1,..., p), $\mu_{p+\alpha}(t) = 0$ (α = 1, ..., s) when $\Psi_\alpha(t, x_o(t), w_o) > 0$. Further, $\frac{d\mu_{p+\alpha}}{dt}$ is a non-positive function with the same continuity properties as $u_o(t)$.

(2) Along x_o

(7.23) $$\dot{p}_i = - H_{x^i}, \quad H_{u^k} = 0$$

(3) The inequality

(7.24) $$H(t, x_o(t), u, p(t), w_o, 0, 0) \leq H(t, x_o(t), u_o(t), p(t), w_o, 0, 0)$$

holds along x_o for all u such that $(t, x_o(t), u, w_o)$ is in R_o.

(4) <u>The transversality condition</u>

$$(7.25) \qquad dG + [-H^s dT^s + p_i(t^s) dX^{is}]_{s=0}^{s=1} - \int_{t^0}^{t^1} H_{w^\sigma} dw^\sigma dt = 0$$

<u>holds along</u> x_0 <u>for all</u> dw^σ.

Next suppose that for $t^0 \leq t < t_1$ and $t_2 < t \leq t^1$, x_0 is interior to S while in S* for $t_1 \leq t \leq t_2$, where $t_1 \leq t_2$. By a suitable change of parameter we can take $t^0 = -1$, $t_1 = 0$, and $t^1 = 1$. Assume that for $0 \leq t \leq t_2$, the condition corresponding to (7.8) holds and similarly for (7.10) on $-1 \leq t \leq 0$ and $t^2 \leq t \leq 1$.

Let

$$\bar{x}(t) = x(-t) \qquad (0 \leq t \leq 1)$$

$$\bar{u}(t) = u(-t)$$

Set

$$F(t, x, u) = f(t, x, u) \qquad (0 \leq t \leq 1)$$

$$= 0 \qquad \text{otherwise}$$

$$\bar{F}(t, \bar{x}, u) = f(-t, \bar{x}, \bar{u}) \qquad (0 \leq t \leq 1)$$

$$= 0 \qquad \text{otherwise}$$

and similarly for other functions of (t, x, u) and (t, x).

Consider the problem of finding in a class of arcs

$$x: \quad x^i, \bar{x}^i, u^k, \bar{u}^k, w^\sigma \qquad (i=1,\ldots,n;\ k=1,\ldots,m;\ \sigma=1,\ldots,r+n)$$

satisfying conditions of the form

(7.1') $\dot{x}^i = F^i(t, x, u)$

$\dot{\overline{x}}^i = -\overline{F}^i(t, \overline{x}, \overline{u})$

(7.2') $\Psi_\alpha(t, x) \geq 0$

$\Psi_\alpha(t, \overline{x}) \geq 0$

(7.3') $\phi_\beta(t, x, u) \geq 0$

$\phi_\beta(t, \overline{x}, \overline{u}) \geq 0$

(7.4') $t^o = 0,\ t^1 = 1$ $\qquad x^i(1) = X^{i1}(v)$

$x^i(0) = v^{r+i}$ $\qquad \overline{x}^i(1) = X^{io}(v)$

(7.5') $I_\gamma = g_\gamma(v) + \int_o^1 (F(t, x, u) + \overline{F}(t, \overline{x}, \overline{u}))dt \geq 0$

$I_{q+\alpha} = \Psi_\alpha(0, x(0)) \geq 0$

$I_{q+s+1} = -\nu \geq 0$

one which minimizes

$$I = g(v) + \int_o^1 (F(t, x, u) + \overline{F}(t, \overline{x}, \overline{u}))dt$$

This problem corresponds to the previous case. Hence, if x_o minimizes the original problem, for the transformed problem there exist multipliers $\lambda_o, \lambda_1, \ldots, \lambda_{q+s+1}$, $\mu_1(t), \ldots, \mu_{p+s}(t), \overline{\mu}_1(t), \ldots, \overline{\mu}_{p+s}(t), p_1(t), \ldots, p_n(t)$, $\overline{p}_1(t), \ldots, \overline{p}_n(t)$, and functions

(7.26) $H = p_i F^i - \overline{p}_i \overline{F}^i + \lambda_\gamma (F_\gamma + \overline{F}_\gamma)$

$\qquad - \lambda_o (F + \overline{F}) + \mu_\rho \phi_\rho + \overline{\mu}_\rho \overline{\phi}_\rho$

(7.27) $\quad G = \lambda_o g - \lambda_\gamma g_\gamma - \lambda_{q+\alpha} \psi_\alpha(0, x(0)) + \lambda_{q+s+1} \nu$

such that the conclusions of Theorem 7.1 hold along x_o given by

x_o: $x_o^i(t)$, $\bar{x}_o^i(t)$, $u_o(t)$, $\bar{u}_o(t)$, v_o^σ

$(i = 1, \ldots, n;\ j = 1, \ldots, k;\ \sigma = 1, \ldots, r + n)$

Applying the transversality condition (7.25) gives

(7.28) $\quad \lambda_o dg - \lambda_\gamma dg_\gamma + p_i(1)X^{i1}(v) + \bar{p}_i(1)X^{io}(v) = 0$

(7.29) $\quad p_i(0) + \bar{p}_i(0) + \lambda_{q+\alpha} \psi_{\alpha x^i}(0, x_o(0)) = 0$

We note also that the multipliers $\bar{\mu}_{p+\alpha}(t) \equiv 0$ $(\alpha=1,\ldots,s)$ since $\bar{\psi}_\alpha(t, x_o(t)) > 0$.

On $0 \leq t < 1$ set

$\mu_\beta(t) = \bar{\mu}_\beta(-t)$ $\qquad (\beta = 1, \ldots, p)$

$p_i(t) = - \bar{p}_i(-t)$

and observe that

$\dot{x}(t) = - \dot{\bar{x}}(-t).$

Substituting these in (7.1') through (7.5'), (7.26) and (7.27) we have the conclusions of Theorem 7.5 hold except that now $\mu_\beta(t)$, $p_i(t)$ may be discontinuous at $t = 0$ with the discontinuity for $p_i(t)$ of the form (7.29).

To generalize this result suppose that for $t = 0$, 1 and 2

x_o is interior to S while in S* on one subinterval of $0 < t < 1$ and one subinterval of $1 < t < 2$.

Let

$$\bar{x}(t) = x(t - 1) \qquad (1 \leq t \leq 2)$$

$$\bar{u}(t) = u(t - 1).$$

Set

$$F(t, x, u) = f(t, x, u) \qquad (0 \leq t \leq 1)$$

$$= 0 \qquad \text{otherwise}$$

$$\bar{F}(t, x, u) = f(t - 1, \bar{x}, \bar{u}) \qquad (0 \leq t \leq 1)$$

$$= 0 \qquad \text{otherwise}$$

and similarly for other functions of (t, x) and (t, x, u). This reduces to the previous case where now

$$I_{q+\alpha} = \Psi_\alpha(0, x_o(0)) > 0$$

so $\lambda_{q+1}, \ldots, \lambda_{q+s}$ are all zero. Hence, the transversality condition gives

$$(7.30) \qquad p^i(0) = \bar{p}^i(0)$$

Setting

$$\mu_\rho(t + 1) = \bar{\mu}_\rho(t)$$

$$p_i(t + 1) = \bar{p}_i(t)$$

we have that because of (7.30) the functions $p_i(t)$ are continuous at $t = 0$.

We define a point $x_0(\bar{t})$ to be a contact point if for some $\delta > 0$, $x_0(t)$ is interior to S for $\bar{t} - \delta \leq t < \bar{t}$ while $x_0(\bar{t})$ is in S*. We call $\bar{t}$ a contact time. Then since Theorem 7.1 as proved includes the case where x_0 lies entirely in S*, the generalization to where x_0 has a finite number of contact points is clear. Hence we have proved

Theorem 7.2. Under the above assumptions, Theorem 7.1 holds except that the multipliers $p_i(t)$, $\mu_\rho(t)$ may be discontinuous at contact times, where discontinuities in $p_i(t)$ are of the form (7.29).

The conclusion that $\frac{d\mu_\alpha}{dt} \leq 0$ ($\alpha = p + 1, \ldots, p + s$) was first obtained by Gamkrelidze [4] for the problem where (7.1) is independent of v, (7.2) depends on x alone, (7.3) on u alone, and (7.4) and (7.5) are not given. The method for this result used here is an adaptation of his. He does not determine that μ_α itself is non-negative. Berkovitz [8], using methods based on [1], derives the condition on μ_α but not its derivative and for a problem where (7.5) does not appear.

Gamkrelidze's definition of "jump point" includes both a contact point as here defined and also a point at which x_0 leaves S*. The methods of [4] and [8] give a possible discontinuity of $p_i(t)$ at jump points rather than at just contact points. Neither author shows that λ_0 can be chosen continuous under these hypotheses.

REFERENCES

1. Hestenes, M. R., A general problem in the calculus of variations with applications to paths of least time. RM-100, The Rand Corp., See also ASTIA Document No. 112382.

2. Valentine, F. A., The problem of Lagrange with differential inequalities as added side conditions. Dissertation, University of Chicago, 1937.

3. Boltyanskii, V. G., Gamkrelidze, R. V., and Pontryagin, L. S., The theory of optimal processes I: The maximum principle. Izvestia Akad. Nauk. SSSR, Ser. Mat. 24, 3-24 (1960).

4. Gamkrelidze, R. V., Optimal processes with bounded phase coordinates. Izvestia Akad. Nauk. SSSR, Ser. Mat., 24, 315-356 (1960).

5. Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., Mishchenko, E. F., "The mathematical theory of optimal processes." Interscience Publishers, New York, 1962.

6. Hestenes, M. R., Variational theory and optimal control theory. Memographed lecture notes, UCLA, 1963.

7. Graves, L. M., On the Weierstrass condition for the problem of Bolza in the calculus of variations. Annals of Math. 33, 747-52 (1932).

8. Berkovitz, L. D., On control problems with bounded state variables. Jour. Math. Anal. Appl. 5, 488-489 (1962).

9. Bliss, G. A., "Lectures in the calculus of variations." University of Chicago Press, Chicago, Illinois, 1946.

10. Boltyanskii, V. G., Gamkrelidze, R. V., Pontryagin, L. S., On the theory of optimal processes. Dolady Akad. Nauk. SSSR 110, 7-10 (1957).

11. Gamkrelidze, R. V., On the theory of optimal processes in linear systems. Doklady Akad. Nauk. SSSR 116, 9-12 (1957).

12. Gamkrelidze, R. V., On the theory of time optimal processes in linear systems. Izvestia Akad. Nauk SSSR, Ser. Mat. 22, 449-474 (1958).

13. Gamkrelidze, R. V., On the general theory of optimal processes. Doklady Akad. Nauk. SSSR 123, 223-226 (1958).

14. Boltyanskii, V. G., The maximum principle in the theory of optimal processes. Doklady Akad. Nauk SSSR 119, 1070-1073 (1958).

15. McShane, E. J., On multipliers for Lagrange problems. American Journal Mathematics 61, 809-819 (1939).

16. Dunford, N., and Schwartz, J. T., "Linear Operators." Interscience Publishers, Inc., New York, 1958.

17. McShane, E. J., "Integration." Princeton University Press, Princeton, New Jersey, 1947.

APPENDIX A

In Sections 3 - 7 it has been assumed that under the stated hypotheses, solutions of a system of ordinary differential equations have certain properties. The following theorem justifies this assumption.

Theorem A-1. Let R *be a region in* (t, x, w) - *space and let* $t^o \leq t \leq t^1$ *be any interval in the projection of* R *on* t - *space. Consider the system*

$$x^i = f^i(t, x, w) \qquad (i = 1, \ldots, n) \tag{A-1}$$

Let R *and the functions* f^i *have the following properties.*

(1) R *is convex in* x *and* w *for each fixed* t.

(2) f^i *is defined on* R.

(3) f^i *is continuous in* x *and* w *for fixed* t, *and is integrable in* t *for fixed* x *and* w.

(4) $f^i_{x^j}$, $f^i_{w^\sigma}$ *exist for* (t, x, w) *in* R *and there is a function* S(t) *integrable on* $t^o \leq t \leq t^1$ *such that*

$$|f^i_{x^j}(t, x, w)| \leq S(t) \qquad (t^o \leq t \leq t^1) \tag{A-3a}$$

$$|f^i_{w^\sigma}(t, x, w)| \leq S(t) \tag{A-3b}$$

for $i = 1, \ldots, n;\ j = 1, \ldots, n;\ \sigma = 1, \ldots, r.$

Then, for any τ on $t^0 \le t \le t^1$ and any ξ, v for which (τ, ξ, v) is in R, there is a $\delta > 0$ and a function $x(t; \tau, \xi, v)$ such that

(a) $(t, x(t; \tau, \xi, v), v)$ is in R for $|t - \tau| \le \delta$.

(b) $x(t; \tau, \xi, v)$ is a unique function, absolutely continuous in t, which satisfies (A-1) almost everywhere on $|t - \tau| \le \delta$ and

$$x^i(\tau; \tau, \xi, v) = \xi^i \tag{A-4}$$

(c) The functions $x^i(t; \tau, \xi, v)$ are continuous in t, τ, ξ, v for $t^0 \le \tau \le t^1$, $|t - \tau| \le \delta$, (t, ξ, v) in R.

(d) For almost all τ on $t^0 \le \tau \le t^1$, $x^i(t; \tau, \xi, v)$ has a difference quotient with respect to (τ, ξ, v) which is bounded for $|t - \tau| \le \delta$.

(e) Assume further that there is a solution

$$x_0: \quad x_0^{\,i}(t), \; v_0^{\,a} \qquad (t^0 \le t \le t^1)$$

of (A-1) for which $(t, x_0(t), v_0)$ is in R. Then there is a neighborhood R' of x_0 such that for any point (τ, ξ, v) in R' there is a unique solution $x(t; \tau, \xi, v)$ of (A-1) with $(t, x(t; \tau, \xi, v), v)$ in R for t in the projection of R' on t - space.

Without loss of generality we can consider the case in which v does not appear. To show this, define

$$x^{n+\sigma} = v^{\sigma} \qquad (\sigma = 1, \ldots, r)$$

Then for the system of equations

$$\dot{x}^{i} = f^{i}(t, x \ \ldots, x^{n+r}) \qquad (i = 1, \ldots, n)$$

$$\dot{x}^{n+\sigma} = f^{n+\sigma} = 0$$

all hypotheses are satisfied for $f^{j}(j = 1, \ldots, n + r)$ and we need only define

$$\xi_{o}^{n+\sigma} = v_{o}^{\sigma}.$$

Further (a), (b), (c) follow from the proof of (e) by taking $t^{o} = t^{1}$ and $x^{i}_{o}(t) = \xi^{i}_{o}$. Hence, we will first prove (e).

Since R is convex in x for each t, it follows immediately from (A-2) that

(A-5) $\quad |f(t, x) - f(t, y)| \leq S(t)|x - y|$

Fixing y, f(t, y) is integrable and

(A-6) $\quad |f(t, x)| \leq |f(t, y)| + S(t)|y| + S(t)|x|$
$\qquad\qquad \leq \bar{S}(t)(1 + |x|)$

where

$$\bar{S}(t) = \max\ (|f(t, y)| + S(t)|y|, S(t))$$

is also integrable on the same range. Since $\bar{S}(t) \geq S(t)$ we can replace S(t) by $\bar{S}(t)$ in (A-2) and (A-5).

Let $\delta > 0$ be such that $(t, x_0(t))$ is in R for $t^0 - \delta \leq t \leq t^1 + \delta$, and extend $x_0(t)$ to be continuous on the larger interval. We note that given $0 < \varepsilon < 1/2$, there is a $\delta > 0$ such that for any set A on $t^0 - \delta \leq t \leq t^1 + \delta$ with measure less than δ,

$$\text{(A-7)} \qquad \int_A S(t)dt < \varepsilon/2, \quad \left|\int_A f^i(t, x_0(t))dt\right| < \varepsilon/2 \quad (i = 1, \ldots, n).$$

Now let $|\tau - \tau_0| \leq \delta/2$, $|\xi - \xi_0| \leq \varepsilon/2$, where ε is yet to be chosen and the choose of δ depends on ε so that (A-7) holds. We define for $|t - \tau| \leq \delta$,

$$x^i_{n+1}(t) = \xi^i + \int_\tau^t f^i_n(s)\, ds$$

where

$$f^i_n(s) = f^i(s, x_n(s)).$$

Then

$$|x^i_{n+1}(t) - x^i_0(t)| \leq |\xi^i - \xi^i_0| + \left|\int_\tau^t f^i_n(s)\, ds - \int_{\tau_0}^t f^i_0(s)ds\right|$$

$$\leq |\xi^i - \xi^i_0| + \left|\int_{\tau_0}^{\tau} f^i(s, x_0(s))ds\right|$$

$$+ \left|\int_\tau^t |f^i(s, x_n(s)) - f^i(s, x_0(s))|ds\right|.$$

Hence, taking maxima over $|\tau - t| \leq \delta$, and using (A-7) and (A-5),

$$\max |x^i_{n+1}(t) - x^i_0(t)| < \varepsilon + \varepsilon \max |x^i_n(t) - x^i_0(t)|$$

$$< \sum_{j=1}^{n+1} \varepsilon^j < \varepsilon/(1-\varepsilon) < 2\varepsilon$$

since

$$|x_1^i(t) - x_o(t)| \leq |\xi^i - \xi_o^i| + | \int_{\tau_o}^{\tau} f^i(s, x_o(s))ds| < \varepsilon.$$

So for ε sufficiently small, $(t, x_n(t))$ is in R for $|t - \tau| \leq \delta$ and for all n.

Again, taking maxima over $|t - \tau| \leq \delta$

$$|x_{n+1}^i(t) - x_n^i(t)|^2 \leq 2| \int_{\tau}^{t} |x_{n+1}(s) - x_n^i(s)||f_n^i(s) - f_{n-1}^i(s)|ds|$$

$$\leq 2| \int_{\tau}^{t} |x_{n+1}(s) - x_n^i(x)|$$

$$|x_n^i(s) - x_{n-1}^i(s)|S(s)ds|$$

$$< \varepsilon \max|x_{n+1}(s) - x_n^i(s)||x_n^i(s) - x_{n-1}^i(s)|.$$

Or

$$\max|x_{n+1}^i(t) - x_n^i(t)| < \varepsilon \max|x_n^i(t) - x_{n-1}^i(t)|$$

$$< \varepsilon^n$$

Therefore, $x_n(t)$ converges uniformly on $|t - \tau| \leq \delta$ to a continuous function $x(t)$ and by tne previous result, $(t, x\ (t))$ is in R for this interval. Now

$$|x^i(t) - \xi^i - \int_\tau^t f^i(t, x(t)dt|$$

$$\leq |x^i(t) - x_n^i(t)| + |\int_\tau^t |f^i(s, x(s)) - f_{n-1}^i(s)|ds|$$

$$\leq |x^i(t) - x_n^i(t)| + \varepsilon \max|x^i(s) - x_{n-1}^i(s)|.$$

Since the left hand side is independent of n and convergence is uniform, $x(t)$ satisfies (A-1) for $|t - \tau| \leq \delta$ and $x(\tau) = \xi$.

Suppose $y(t)$ also satisfies (A-1) and $y(\tau) = \xi$. Then

$$|x(t) - y(t)| \leq |\int_\tau^t S(s)|x(s) - y(s)|ds$$

$$\leq \varepsilon/2 \max|x(s) - y(s)|$$

and clearly this can hold only if $x(t) = y(t)$ for $|t - \tau| \leq \delta$.

Let τ_1, ξ_1 be such that $|\tau_1 - \tau_o| \leq \delta/2$, $|\xi_1 - \xi_o| < \varepsilon/2$. Then

$$|x(t; \tau_1; \xi_1) - x(t; \tau, \xi)|$$

$$\leq |\xi_1 - \xi_o| + |\int_{\tau_1}^{\tau} f^i(s\ x(s; \tau_1, \xi_1))ds|$$

$$+ |\int_\tau^t |f(s, x(s, \tau_1, \xi_1) - f(s\ x(s, \tau, \xi)|ds|$$

$$< \varepsilon + \varepsilon \max|x(t, \tau_1, \xi_1) - x(t; \tau, \xi)|$$

or

$$\max|x(t;\ \tau_1,\ \xi_1) - x(t;\ \tau,\ \xi)| < 2\varepsilon$$

so a solution $x(t;\ \tau,\ \xi)$ is continuous in τ, ξ near τ_o, ξ_o uniformly in t for $|t - \tau| < \delta$. Therefore, by Osgood's theorem $x(t;\ \tau,\ \xi)$ is continuous in t, τ, ξ.

The interval $t^o - \delta \leq t \leq t^1 + \delta$ can be covered by a finite number of such subintervals For each there is a neighborhood of $x_o(t)$ through every point of which there is a unique continuous solution $x(t)$ of (A-1) such that $(t, x(t))$ remains in R. Hence, (e) follows by continuity.

To prove (d) we define

$$\text{(A-8)} \qquad y^i(t,\ b) = (x^i(t;\tau + b_1,\ \xi + b_2) - x^i(t;\ \tau,\ \xi))/|b|$$

$$\equiv \Delta x^i(t,\ b)/|b|$$

and note that

$$\text{(A-9)} \qquad y^i(\tau,\ b) = (x^i(\tau;\ \tau + b_1,\ \xi + b_2) - x^i(\tau;\ \tau,\ \xi))/|b|$$

Now consider the system of equations

$$\text{(A-10)} \quad \dot{y}^i(t,\ b) = (f^i(t,\ x(t;\ \tau + b_1,\ \xi + b_2)) - f^i(t,\ x(t;\ \tau,\ \xi)))/|b|$$

$$= \int_o^1 f^i_{x^j}(t,\ x(t;\ \tau,\ \xi) + \theta\Delta x(t,\ b))d\theta\Delta x^j(t,\ b)/|b|$$

$$= A^i_j(t,\ b)y^j(t,\ b)$$

By hypothesis $|A^i_j(t, b)|$ is bounded by $S(t)$ so (A-10) has a solution. Since for $b \neq 0$, $y^i(t, b)$ given by (A-8) is a solution of (A-10) with initial conditions (A-9), it is the unique solution.

Let τ be an ordinary point for $S(t)$. Then

$$\text{(A-11)} \qquad |y^i(\tau, b)| \leq \frac{|b_2|}{|b|} + \frac{1}{|b|}\left| \int_\tau^{\tau+b_1} (f^i(t, x(t; \tau + b_1, \xi + b_2)) - f^i(t, x(t, \tau, \xi))dt \right|$$

$$\leq 1 + \frac{1}{|b|}\left| \int_\tau^{\tau+b_1} S(t)|x(t; \tau + b_1, \xi + b_2) - x(t; \tau, \xi))|dt \right|$$

But by (b) $x(t; \tau, \xi)$ is uniformly continuous on $|t - \tau| \leq \delta$, $|b_1| \leq \delta$, $|b_2| \leq \delta$ hence assumes its maximum M. Therefore,

$$\text{(A-12)} \qquad |y^i(\tau, b)| \leq 1 + \frac{M}{|b|}\left| \int_\tau^{\tau+b} S(t)dt \right|$$

So

$$\text{(A-13)} \qquad |y^i(\tau, 0)| \leq 1 + MS(\tau) < \infty$$

for almost all τ on $t^0 \leq \tau \leq t^1$.

By (A-10) and (A-2), for $b \neq 0$ and taking maxima over $|t - \tau| \leq \delta$

$$\text{(A-14)} \qquad |y^i(t, b)| \leq |y^i(\tau, b)| + \left| \int_\tau^t S(t)|y^i(t, b)|dt \right|$$

$$\leq |y^i(\tau, b)| + \epsilon \max|y^i(t, b)|$$

So

$$\max|y^i(t, b)| \leq |y^i(\tau, b)|/(1 - \epsilon)$$

Hence, by (A-12), $y^i(t, 0)$ is bounded on $|t - \tau| \leq \delta$ and as before is then bounded on $t^0 \leq t \leq t^1$.

Corollary. Let $\tau(b)$ have partial derivatives with respect to b_j $(j = 1, \ldots, p)$ at $b = 0$, and $\xi(b)$ have a bounded difference quotient at $b = 0$. Then $x(t, \tau(b), \xi(b))$ has a difference quotient with respect to b at $b = 0$ which is bounded on $t^0 \leq t \leq t^1$.

With appropriate changes in notations in (A-8) and (A-9) we have that

$$|y^i(\tau(b), b)| \leq \frac{|\xi(b) - \xi(0)|}{|b|} + \frac{1}{|b|}\left|\int_{\tau(0)}^{\tau(b)} S(t)|x(t,\tau(b),\xi(b)) - x(t,\tau(0),\xi(0))|dt\right|$$

$$\leq \frac{|\xi(b) - \xi(0)|}{|b|} + \frac{M}{|b|}\left|\int_{\tau(0)}^{\tau(b)} S(t)dt\right|$$

Using the hypotheses on $\tau(b)$ and $\xi(b)$ gives that $y^i(\tau, 0)$ is bounded for almost all τ on $t^0 \leq \tau \leq t^1$. Equation (A-14) and resulting conclusion remain unchanged.