A NEW TYPE OF VARIATIONAL THEORY SUFFICIENCY THEOREM

Russell D. Rupp

This paper contains a sufficiency theorem appropriate to applications of Hestenes' method of multipliers. It is a generalization of the lemma of Hahn, which has been used by previous authors to prove variable end point sufficiency theorems. The classical and control formulations of the problems of Mayer, Lagrange, and Bolza in both parametric and nonparametric form are included. The method of proof is indirect.

1. **Introduction.** A large class of optimization problems in the calculus of variations and optimal control theory may be solved via the reduction of constraints. One such technique is Hestenes' method of multipliers [5], 303 et seq. The sufficiency theory here given is appropriate to applications of Hestenes' method of multipliers. It includes the classical and control formulations of the problems of Mayer, Lagrange, and Bolza in both parametric and nonparametric form. The sufficiency theorem itself is a generalization of the lemma of Hahn, which has been used by previous authors to prove variable end point sufficiency theorems. The method of proof is indirect. This technique was developed by Hestenes [4, 6, 7] and McShane [10]. More recently Mookini [11] has used the method on an optimal control problem, and Pennisi [13] to consider differential inequality constraints. At the present time Nathanson [12] has extended the results of Mookini and Pennisi to a multistage optimal control problem.

The problem to be studied in this paper is that of minimizing a function of integrals. Like the functional the class of arcs contains parameters in a metric space. As shown in the author's dissertation [14], 85–134, such generality permits application to the existence and convergence of solutions for perturbed problems. Let $\mathfrak{D}$ be a metric space with metric $d(z, z^*)$. We assume that an open subset $\mathfrak{R}$ of $E^{2n+r+1} \times \mathfrak{D}$ is given, and we denote elements of $\mathfrak{R}$ by $(b, t, x, \dot{x}, z)$ where

$$b = (b^1, b^2, \cdots, b^r) \in E^r, \quad t \in E^1, \quad z \in \mathfrak{D}$$
$$x = (x^1, x^2, \cdots, x^n) \in E^n, \quad (\dot{x}^1, \dot{x}^2, \cdots, \dot{x}^n) \in E^n.$$

Let $T^s(b, z)$ and $X^{is}(b, z)$, $s = 1, 2$; $i = 1, 2, \cdots, n$ be real-valued functions with domain

$$\mathrm{proj}_{bz}(\mathfrak{R}) = \text{the projection of } \mathfrak{R} \text{ into } bz\text{-space}.$$

By an arc x we mean a collection of constants, real-valued functions, and an element of $\mathfrak{D}$

$$x\colon b^h, x^i(t), z \quad h = 1, 2, \cdots, r; \quad i = 1, 2, \cdots, n$$

such that

$$\mathrm{Dom}(x^i(t)) = \text{the domain of } x^i(t) = [T^1(b, z), T^2(b, z)]$$

and $x^i(t)$ has a piecewise continuous derivative. Let b and $x(t)$ denote the vectors whose components are given by b^h and $x^i(t)$. An arc $x\colon b, x(t), z$ is called admissible and said to belong to the class $\mathfrak{A}$ of admissible arcs if

$$x^i(T^s(b, z)) = X^{is}(b, z) \quad s = 1, 2; \; i = 1, 2, \cdots, n$$

and $(b, t, x(t), \dot{x}(t), z) \in \mathfrak{R}$ for all $T^1(b, z) \leqq t \leqq T^2(b, z)$. We extend if necessary the domain of the function $x(t)$ which is associated with an arc $x\colon b, x(t), z$ in $\mathfrak{A}$ by making it continuous and constant where it is undefined. In those cases where all the functions in a particular collection have been so extended, we denote their common domain by $T' \leqq t \leqq T''$.

Abnormal problems such as these considered by Hestenes [4, 7] and McShane [10] require another parameter y in a compact metric space $\mathfrak{D}^*$ with metric $d^*(y, y^*)$. Given y in $\mathfrak{D}^*$ and an admissible arc $x\colon b, x(t), z$, we define the collection x_y to be

$$x_y\colon b, x(t), z, y.$$

The functional

$$G(x_y) = G(I_0(x_y), I_1(x_y), \cdots, I_m(x_y), z, y)$$

is assumed to be well-defined on the Cartesian product $\mathfrak{A} \times \mathfrak{D}^*$ of $\mathfrak{A}$ and $\mathfrak{D}^*$ where

$$I_\sigma(x_y) = g_\sigma(b, z, y) + \int_{T^1(b,z)}^{T^2(b,z)} f_\sigma(b, t, x(t), \dot{x}(t), z, y)dt$$

and $g_\sigma(b, z, y), f_\sigma(b, t, x, \dot{x}, z, y), \sigma = 0, 1, 2, \cdots, m$ are real-valued functions of the indicated variables. In addition, we suppose that there are nonnegative integers N and R, $0 \leqq r - R \leqq n - N$, such that the arguments

$$x^{N+1}, x^{N+2}, \cdots, x^n \text{ and } b^{R+1}, b^{R+2}, \cdots, b^r$$

do not appear in f_σ or g_σ, $\sigma = 0, 1, \cdots, m$. Given vectors $x = (x^1, x^2, \cdots, x^n)$ and $b = (b^1, b^2, \cdots, b^r)$, $\hat{x}$ and $\hat{b}$ are defined to be the vectors

$$(x^1, x^2, \cdots, x^N) \text{ and } (b^1, b^2 \cdots, b^R).$$

We also assume that for $k = 1, 2, \cdots, r - R$,

$$X^{n+R-r+k2} = b^{R+k}$$

and that the b^{R+k} appear only in that manner. Brady [1] has applied the theory of the problem of Bolza to a less general function of integrals.

The indices σ, h, s, and i always have the ranges indicated above. The subscript σ is used to indicate the partial derivative of $G(I_0, I_1, \cdots, I_m, z)$ with respect to the variable I_σ, and the subscript h to indicate a partial derivative with respect to b^h. We further simplify notation by defining for an admissible arc $x: b, x(t), z$ and y in $\mathfrak{D}^*$,

$$I(x_y) = (I_0(x_y), I_1(x_y), \cdots, I_m(x_y)).$$

Unless otherwise indicated, the repeated indices summation convention is used. For example given a l^*-dimensional vector u with components u^l, its norm is denoted by

$$|u| = (u^l u^l)^{1/2}.$$

For a real number this norm would be its absolute value. By u-space we mean the space of points

$$(u^1, u^2, \cdots, u^{l*}).$$

Thus we use the term u-neighborhood to mean a neighborhood in u-space. As a measure of the distance between two arcs $x: b, x(t), z$ and $x^*: b^*, x^*(t), z^*$, we define the pseudometric

$$\rho(x, x^*) = |\hat{b} - \hat{b}^*| + \sup_t |\hat{x}(t) - \hat{x}^*(t)|.$$

Furthermore x is said to belong to a $\hat{b}t\hat{x}\dot{x}z$-neighborhood $\mathfrak{N}$ if $(\hat{b}, t, \hat{x}(t), \dot{x}(t), z)$ is always in $\mathfrak{N}$. The similar convention is made with respect to the other combinations of variables. An admissible arc of the form $x: b, x(t), z$ is called z-admissible. If $\dot{x}(t)$ is continuous, we say that x is smooth or has no corners.

Before proceeding to the sufficiency criteria in the next section, we pause to remark that a larger class of arcs of the form

$$x_y: b, x(t), z, y$$

may actually be considered. The variables $b, t, x(t), \dot{x}(t), z, y$ need only be in $\mathfrak{N} \times \mathfrak{D}^*$ for almost all $T^1(b, z) \leqq t \leqq T^2(b, z)$. Similarly we need only that $x^i(t)$ be absolutely continuous, $f_\sigma(b, t, x(t), \dot{x}(t), z, y)$ be integrable, and $G(I(x_y), z, y)$ be defined. For example this is usually the case if $\dot{x}(t)$ is essentially bounded. One still assumes that the admissible end

conditions are satisfied. It is easy to verify that our analysis is not
altered in this situation.

2. **Two sufficiency theorems.** The sufficient conditions are
strengthened forms of the first and second order necessary conditions.
The purpose of this section is to state sufficient conditions that an
admissible arc of the form $x_0: b_0, x_0(t), z_0$ be minimizing. As mentioned
in the preceding section, application to abnormal problems requires an
additional y-parameter. For normal problems and some abnormal cases
the parameter y is not needed. When that happens, we are free to
consider the space $\mathfrak{D}^*$ to be the trivial metric space of one element.
For the situation in abnormal problems of Bolza, see Hestenes [4, 7]
and McShane [6]. With this in mind, we fix the admissible arc
$x_0: b_0, x_0(t), z_0$.

The functions $T^s(b, z)$, $X^{is}(b, z)$, $g_\sigma(b, z)$ are assumed to have two
continuous b-derivatives in a neighborhood of b_0, z_0 and $G(I, z, y)$ to
have two continuous I-derivatives in a neighborhood of $\{(I(x_{0y}), z_0, y): y$
in $\mathfrak{D}^*\}$. We continue by assuming $f_\sigma(b, t, x, \dot{x}, z, y)$ to have a continuous
derivative with respect to its Euclidean arguments in a neighborhood
of $\{(b_0, t, x_0(t), \dot{x}_0(t), z_0, y): T^1(b_0, z_0) \leqq t \leqq T^2(b_0, z_0)$ and y in $\mathfrak{D}^*\}$. Finally
the same assumption is made concerning the functions $f_{\sigma h}(b, t, x, \dot{x}, z, y)$,
$f_{\sigma x^i}(b, t, x, \dot{x}, z, y)$, and $f_{\sigma \dot{x}^i}(b, t, x, \dot{x}, z, y)$. The lack of any hypothesis
concerning $f_{\sigma tt}(b, t, x, \dot{x}, z, y)$ is useful in applications. Given multipliers
$\lambda = (\lambda_0, \lambda_1, \cdots, \lambda_m)$, we define

$$f(b, t, x, \dot{x}, z, y, \lambda) = \lambda_\sigma f_\sigma(b, t, x, \dot{x}, z, y).$$

Here and later on, repeated indices denote summation with respect to
that index unless otherwise indicated.

A variation γ which is admissible with respect to an arc $x: b, x(t), z$
in $\mathfrak{A}$ is given by a set of constants and absolutely continuous functions

$$\gamma: \beta^h, \eta^i(t)$$

such that

$$\text{Dom}(\eta^i(t)) = [T^1(b, z), T^2(b, z)]$$
$$\dot{\eta}^i(t) \in L_2(T^1(b, z), T(b, z))$$
$$\eta^i(T^s(b, z)) = [X_h^{is}(b, z) - \dot{x}^i(T^s(b, z))T^s(b, z)]\beta^h.$$

Given an admissible arc $x: b, x(t), z$ in $\mathfrak{A}$ and y in $\mathfrak{D}^*$, we also define

$$G(x_y) = G(I(x_y), z, y)$$

and

$$\lambda_{x_y} = (G_0(x_y), G_1(x_y), \cdots, G_m(x_y))$$

where as indicated in the preceding section, the subscript σ on $G(I, y, z)$ denotes a partial derivative with respect to I_σ so that

$$\lambda_{xy\sigma} = \frac{\partial}{\partial I_\sigma} G(x_y) = G_\sigma(x_y).$$

Finally let $\lambda_{0y} = \lambda_{x_0 y}$. With these notations in mind, we enumerate some of the sufficiency criteria.

(2.1) transversality:

Along x_{0y}, λ_{0y} we have for every y in $\mathfrak{D}^*$,

$$\lambda_\sigma g_{\sigma h} + [(f - \dot{x}^i f_{\dot{x}i}) T_h^s + f_{\dot{x}i} X_h^{is}]_{s=1}^{s=2} + \int_{T^1}^{T^2} f_h ds = 0.$$

(2.2) Euler equations (integral form):

For every y in $\mathfrak{D}^*$ there exist constants c_{iy} such that along x_{0y}, λ_{0y},

$$f_{\dot{x}j} = \int_{T^1}^{t} f_{xj} ds + c_{jy}, \quad j = 1, 2, \cdots, N$$
$$f_{\dot{x}j} = c_{jy}, \quad j = N+1, N+2, \cdots, n-r+R$$
$$f_{\dot{x}j} = 0, \quad j = n-r+R+1, n-r+R+2, \cdots, n.$$

Furthermore if an arbitrary admissible arc x satisfies the Euler equations and transversality condition for one y in $\mathfrak{D}^*$, then x_y satisfies the Euler equations and transversality condition for every y in $\mathfrak{D}^*$.

To verify that this last statement is satisfied in applications, one may check Hestenes [4, 7] and McShane [10]. This hypothesis is used to obtain Theorem 4.1, which is used in the sufficiency theorem proof.

(2.3) positiveness of the second variation:

Let $\gamma: \beta^h$, $\eta^i(t)$ be an admissible variation with respect to x_0. There is some y_γ in $\mathfrak{D}^*$ such that along x_{0y_γ}, λ_{0y_γ},

$$G''_{y_\gamma}(\gamma) = Q(\beta) + \int_{T^1}^{T^2} 2\omega dt + I'_{y_\gamma\sigma} I'_{y_\gamma\mu} G_{\sigma\mu} > 0$$

where

$$Q(\beta) = \{\lambda_\sigma g_{\sigma h l} + [(f - \dot{x}^i f_{\dot{x}i}) T_{hl}^s + f_{\dot{x}i} X_{hl}^{is}]_{s=1}^{s=2}$$
$$+ [(f_t - \dot{x}^i f_{\dot{x}i}) T_h^s T_l^s + 2 f_{\dot{x}i} T_h^s X_l^{is} + 2 f_h T_l^s]_{s=1}^{s=2}\} \beta^h \beta^l$$
$$2\omega = f_{x^i x j} \eta^i \eta^j + f_{\dot{x}^i \dot{x} j} \dot{\eta}^i \dot{\eta}^j + f_{hl} \beta^h \beta^l + 2 f_{x^i \dot{x} j} \eta^i \dot{\eta}^j$$
$$+ 2 f_{x^i h} \eta^i \beta^h + 2 f_{\dot{x}^i h} \dot{\eta}^i \beta^h$$
$$I'_{y_\gamma \sigma}(\gamma) = g_{\sigma h} \beta^h + f_\sigma T_h^s |_{s=1}^{s=2} \beta^h + \int_{T^1}^{T^2} (f_{\sigma h} \beta^h + f_{\sigma x^i} \eta^i + f_{\sigma \dot{x} i} \dot{\eta}^i) dt.$$

(2.4) Weierstrass condition $\mathfrak{N}$:

Define $L(x) = (1 + |\dot{x}|^2)^{1/2} - 1$. There are a positive constant τ and a $bt x \dot{x} z \lambda$-neighborhood $\mathfrak{N}_0$ of x_{0y}, λ_{0y} such that for every y in $\mathfrak{D}^*$, $(b, t, x, \dot{x}, z, \lambda)$ in $\mathfrak{N}_0$ and $(b, t, x, \tilde{x}, z)$ in $\mathfrak{N}$,

$$E_f(b, t, x, \dot{x}, \dot{\tilde{x}}, z, y, \lambda) \geqq \tau E_L(x, \dot{\tilde{x}})$$

where

$$E_f(b, t, x, \dot{x}, \dot{\tilde{x}}, z, y, \lambda) = f(b, t, x, \dot{\tilde{x}}, z, y, \lambda) - f(b, t, x, \dot{x}, z, y, \lambda)$$
$$- (\dot{\tilde{x}}^i - \dot{x}^i)f_{\dot{x}^i}(b, t, x, \dot{x}, z, y, \lambda)$$

and

$$E_L(\dot{x}, \dot{\tilde{x}}) = L(\dot{\tilde{x}}) - L(\dot{x}) - (\dot{\tilde{x}}^i - \dot{x}^i)L_{\dot{x}^i}(\dot{x})$$
$$= L(\dot{\tilde{x}})(1 - \cos\theta)$$

where θ is the angle between the vectors $1, \dot{\tilde{x}}$ and $1, \dot{x}$.

It is not difficult to verify that (2.4) implies

(2.5) Legendre condition (strengthened form):

For any y in $\mathfrak{D}^*$ and $\Pi \neq 0$, we have along x_{0y}, λ_{0y},

(2.6) $$f_{\dot{x}^i \dot{x}^j}\Pi^i\Pi^j > 0.$$

If (2.6) holds with $>$ replaced by $\neq$ and an arbitrary admissible arc x, we say that x is nonsingular. An arc satisfying the Euler equations for all y in $\mathfrak{D}^*$ and which has no corners is called an extremal. For given z, an extremal of the form x: b, $x(t)$, z is called a z-extremal. Similarly an extremal which satisfies the admissible end conditions is called a terminally admissible extremal. For example if x_0 had no corners it would be described as a terminally admissible z_0-extremal. By an accessory extremal, one means an extremal for the second variation along a minimizing arc.

The first sufficiency theorem may now be stated.

THEOREM 2.1 *Let the smooth admissible arc x_0: b_0, $x_0(t)$, z_0 satisfy (2.1) thru (2.4). There exist positive constants ε, ε^*, δ such that given z having $d(z, z_0) < \delta$, there is a terminally admissible z-extremal*

$$x_z: b_z, x_z(t), z$$

such that $\rho(x_z, x_0) < \varepsilon$. For every z-admissible arc

$$x: b, x(t), z \neq x_z: b_z, x_z(t), z$$

having $\rho(x, x_0) < \varepsilon$ and $|I(x_y) - I(x_{0y})| < \varepsilon^*$ for all y in $\mathfrak{D}^*$, there is y_x in $\mathfrak{D}^*$ such that

(2.7) $$G(I(x_{y_x}), z, y_x) > G(I(x_{zy_x}), z, y_x).$$

If for each z, y in a neighborhood of $\{(z, y): z = z_0$ and y in $\mathfrak{D}^*\}$, $G(I, z, y)$ is convex as a function of $I(x)$ in a neighborhood of $I(x_0)$, we may take $\varepsilon^* = \infty$.

The proof of this theorem depends on results in §§ 4, 5, 6, and 8. It is actually carried out in § 7.

The second sufficiency theorem is a special case for which the conclusion of Theorem 2.1 is true under less restrictive hypotheses. There is no change in the proof. We simply do not use some of the hypotheses. This second theorem is Theorem 2.2.

THEOREM 2.2   *Let $T^1 < T^2$ be constants and make no assumption concerning partial derivatives with respect to t.   The conclusions of Theorem 2.1 still hold.*

3. **Corollaries.** In this section we obtain some immediate consequences of Theorem 2.1.   The admissible arc x_0: b_0, $x_0(t)$, z_0 continues to satisfy that theorem's hypotheses.

Let γ: β, $\eta(t)$ and γ^*: β^*, $\eta^*(t)$ be admissible variations.   Define the inner product $\langle \gamma, \gamma^* \rangle$ by

$$\langle \gamma, \gamma^* \rangle = \beta^h \beta^{*h} + \int_{T_0^1}^{T_0^2} \dot{\eta}^i(t) \dot{\eta}^{*i}(t) dt$$

where $T_0^s = T^s(b_0, z_0)$.   The associated norm is

(3.1)
$$\|\gamma\|^2 = |\beta|^2 + \int_{T_0^1}^{T_0^2} |\eta(t)|^2 dt \; .$$

Another inner product of this type is given by

$$\langle \gamma, \gamma^* \rangle + \int_{T_0^1}^{T_0^2} \eta^i(t) \eta^{*i}(t) dt \; .$$

The space $\mathfrak{B}$ of admissible variations, being complete under the norm (3.1), is a Hilbert space.   For each y in $\mathfrak{D}^*$ the second variation $G''_y(\gamma)$ of $G(I(x_y), z_0, y)$ along x_0 is a quadratic form.   In particular there is a positive constant C such that for all y in the compact set $\mathfrak{D}^*$ and all admissible variations γ,

$$|G''_y(\gamma)| < C\|\gamma\|^2 \; .$$

By Theorem 11.1 in Hestenes [2] there is also a positive constant C^* such that for all admissible variations γ and y in $\mathfrak{D}^*$,

(3.2)
$$G''_y(\gamma) \geqq C^*\|\gamma\|^2 \; .$$

Let x: b, $x(t)$, z be a z-admissible arc and let x_z: b_z, $x_z(t)$, z be the associated terminally admissible z-extremal satisfying the transversality condition.   Its existence follows from Theorem 2.1, and its continuity and differentiability properties are given in Theorem 5.1.   Define $\mathfrak{K}(x, x_z)$ *by*

$$\Re(x, x_z) = |b - b_z|^2 + \int_{T'}^{T''} [(1 + |\dot{x}(t) - \dot{x}_z(t)|^2)^{1/2} - 1] dt .$$

Thus by the Weierstrass condition $\Re$, there exists another positive constant τ such that

(3.3) $G(x_y) - \tau\Re(x, x_z)$

satisfies the same hypotheses as $G(x_y)$. We apply Theorem 2.1 or 2.2 to (3.3) and obtain Theorem 3.1.

THEOREM 3.1   *There is a positive constant τ such that the sufficiency theorem holds with the inequality*

$$G(x_y) > G(x_{zy})$$

replaced by

$$G(x_y) > G(x_{zy}) + \tau\Re(x, x_z) .$$

Now consider the space $\mathfrak{B}$ of admissible variations to be a collection of arcs and the collection $\mathfrak{A}$ of admissible arcs to be a metric space with the metric for x: b, $x(t)$, z and x^*: b^*, $x^*(t)$, z^* in $\mathfrak{A}$ defined by

(3.4) $|b - b^*| + \sup_t |\dot{x}(t) - \dot{x}^*(t)| + d(z, z^*) .$

The pair 0, x_0 plays the role of x_0. In other words the collection of admissible arcs is the space $\mathfrak{B}$ of admissible variations with parameters from $\mathfrak{A}$ and metric given by (3.4). Given an admissible arc x in $\mathfrak{A}$ and y in $\mathfrak{D}^*$, let $G''_{yx}(\gamma)$ be the second variation of $G(x_y)$ along x. We verify from (3.2) and the strong Legendre condition that there is a positive constant c such that the functional

$$G''_{yx}(\gamma) - c\|\gamma\|^2$$

satisfies the sufficiency theorem hypotheses. This proves Theorem 3.2.

THEOREM 3.2   *There is a $\hat{b}t\hat{x}\dot{x}z$-neighborhood $\Re$ of x_0 and a positive constant τ such that given an admissible variation γ, there is y_τ in $\mathfrak{D}^*$ such that for x in $\Re \cap \mathfrak{A}$,*

$$G''_{y_\tau x}(\gamma) \geqq \tau\|\gamma\|^2 .$$

The next theorem shows that the assumption on $G''_y(\gamma)$ may be stated in slightly weaker manner.

THEOREM 3.3   *Suppose that g_σ, T^s, X^{is}, f_σ, and G have $k \geq 2$ continuous derivatives with respect to their Euclidean arguments. The*

preceding results hold if we only assume that given γ: β, $\eta(t)$ *in* $\mathfrak{B}$ *for which* $\eta(t)$ *has* $k-1$ *continuous derivatives, there is* y_τ *in* $\mathfrak{D}^*$ *such that*

$$G''_{y_\tau}(\gamma) > 0 \;.$$

This corollary is proved by showing that the original assumption on the second variation holds. Let y_q, γ_q: β_q, $\eta_q(t)$ be a minimizing sequence for $G''_y(\gamma)$ over the Cartesian product of $\mathfrak{D}^*$ and the unit sphere of $\mathfrak{B}$. The sequence $y_q\gamma_q$: β_q, $\eta_q(t)$ may be selected so that y_q converges to y_0 in $\mathfrak{D}^*$ and γ_q: β_q, $\eta_q(t)$ converges weakly to λ_0: β_0, $\eta_0(t)$ in $\mathfrak{B}$. If $\gamma_0 = 0$, the strong Legendre condition implies that there is a positive constant τ such that

$$
(3.5) \qquad
\begin{aligned}
\lim G''_{y_q}(\gamma_q) &= \lim \int_{T^1}^{T^2} f_{\dot{x}^i \dot{x}^j} \dot{\eta}^i_q \dot{\eta}^j_q \, dt \\
&\geq \lim \tau \int_{T^1}^{T^2} |\dot{\eta}_q|^2 dt = \tau > 0 \;.
\end{aligned}
$$

If $\gamma_0 \neq 0$, the Legendre condition implies

$$(3.6) \qquad\qquad \lim G''_{y_q}(\gamma_q) \geq G''_{y_0}(\gamma_0) \;.$$

Thus y_0, γ_0 is a minimum which satisfies the Euler equations (integral form)

$$\omega_{\dot{\eta}i} = \int_{T^1}^{t} \omega_{\eta i} ds + c_i \;.$$

The implicit function theorem and the strong Legendre condition imply that $\eta^i_0(t)$ has $k-1$ continuous derivatives. The proof now follows from (3.5) and (3.6).

4. The Weierstrass condition $\mathfrak{N}$. One object of this section is to analyze the Weierstrass condition $\mathfrak{N}$ to the point where it may be used to ensure the proper convergence for the sequences of arcs which are used in §7. The other is to prove a theorem which permits the parametric theory to be derived from the results here given in the manner of Hestenes [8], 86–87. This latter theorem is the analogue of Theorem 4.1 in the parametric case.

In the nonparametric case the role of the Weierstrass E-function

$$E_L(\dot{x}, \tilde{\dot{x}}) = L(\tilde{\dot{x}}) - L(\dot{x}) - (\tilde{\dot{x}}^i - \dot{x}^i) L_{\dot{x}^i}(\dot{x})$$

of the arc length integrand minus one

$$L(\dot{x}) = (1 + |\dot{x}|^2)^{1/2} - 1$$

is essential. One also finds the formulas

$$E_L(\dot{x}, \dot{\tilde{x}}) = \int_0^1 (1 - \theta) L(w)^{-3} |\mathit{\Pi}|^2 (1 + |w|^2 \sin^2\phi) d\theta$$

where $\mathit{\Pi} = \dot{\tilde{x}} - \dot{x}$, $w(\theta) = \dot{x} + \theta\mathit{\Pi}$, $\phi(\theta)$ is the angle between $\mathit{\Pi}$ and $w(\theta)$ and

$$E_L(\dot{x}, \dot{\tilde{x}}) = L(\dot{\tilde{x}})(1 - \cos\theta)$$

where θ is the angle between the vectors 1, $\dot{x}$ and 1, $\dot{\tilde{x}}$ to be useful. In the parametric case the analogue of $L(\dot{x})$ is $|\dot{x}|$. Let

$$x: b, x(t), z \quad \text{and} \quad \tilde{x}: \tilde{b}, x(t), \tilde{z}$$

be admissible arcs and y be an arbitrary element in $\mathfrak{D}^*$. If it is necessary to extend $\tilde{x}(t)$ so that $x(t)$ and $\tilde{x}(t)$ have a common domain $T' \leq t \leq T''$, we do so by keeping $x(t)$ continuous and requiring that $\dot{x}(t) = \dot{\tilde{x}}(t)$ on the new domain. Using this extension, we make the following definitions, which also occur later:

$$\tilde{\lambda}_{y\sigma} = G_\sigma(\tilde{x}_y) \qquad \lambda_{y\sigma} = G_\sigma(x_y)$$
$$g_\sigma = g_\sigma(b, z, y) \qquad T^s = T^s(b, z)$$
$$J_y^*(x, \tilde{x}) = \tilde{\lambda}_{y\sigma} g_\sigma + \int_{T^1}^{T^2} [f(b, t, x(t), \tilde{x}(t), z, y, \tilde{\lambda}_y)$$
$$+ (\dot{x}^i(t) - \dot{\tilde{x}}^i(t)) f_{\dot{x}^i}(b, t, x(t), \tilde{x}(t), z, y, \tilde{\lambda}_y)] dt$$
$$E_y^*(x, \tilde{x}) = \int_{T^1}^{T^2} E_f(b, t, x(t), \tilde{x}(t), \dot{x}(t), z, y, \tilde{\lambda}_y) dt$$
$$J_y(x, \tilde{x}) = J_y^*(x, \tilde{x}) + E_y^*(x, \tilde{x})$$
$$E_L^*(x, \tilde{x}) = \int_{T'}^{T''} L(\dot{x}(t) - \dot{\tilde{x}}(t)) dt$$
$$\mathfrak{K}(x, \tilde{x}) = |\hat{b} - \hat{\tilde{b}}|^2 + E_L^*(x, \tilde{x}) .$$

THEOREM 4.1   *Let* x_0: b_0, $x_0(t)$, z_0 *be a terminally admissible extremal along which the Weierstrass condition* $\mathfrak{R}$ *holds. There is a positive constant* τ *such that given a positive constant* ε, *there are a* $\hat{b}t\hat{\tilde{x}}z$-*neighborhood* $\mathfrak{F}$ *and a* $\hat{b}t\hat{x}\hat{\tilde{x}}z$-*neighborhood* $\mathfrak{R}_0$ *of* x_0 *such that* $x \in \mathfrak{F} \cap \mathfrak{A}$ *and* $\tilde{x} \in \mathfrak{R}_0 \cap \mathfrak{A}$ *imply*

$$J_y(x, \tilde{x}) - J_y(\tilde{x}, \tilde{x}) \geq \tau[\mathfrak{R}(x, \tilde{x}) - \varepsilon]$$

for all y *in* $\mathfrak{D}^*$.

We begin the proof by observing that the Weierstrass condition $\mathfrak{R}$ implies that there are a $\hat{b}t\hat{\tilde{x}}z$-neighborhood $\mathfrak{F}$ of x_0, a $\hat{b}t\hat{x}\hat{\tilde{x}}z$-neighborhood $\mathfrak{R}_0$ of x_0, and a positive constant τ such that x in $\mathfrak{F} \cap \mathfrak{A}$ and $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$ imply for any y in $\mathfrak{D}^*$,

(4.1) $$E_y^*(x, \tilde{x}) \geq 2\tau E_L^*(x, \tilde{x}) .$$

Given a positive constant ε, diminish $\mathfrak{F}$ and $\mathfrak{R}_0$ so that

$$|b - \widehat{\widetilde{b}}|^2 < \frac{1}{2}\varepsilon\tau$$

when $\hat{b}, \widehat{\widetilde{b}} \in \operatorname{proj}_{\hat{b}}(\mathfrak{F} \cup \mathfrak{R}_0) = $ the projection of $\mathfrak{F} \cup \mathfrak{R}_0$ into $\hat{b}$-space.

Suppose we have shown that $\mathfrak{F}$ and $\mathfrak{R}_0$ may be diminished if necessary so that for $\delta = \min((1/2)\varepsilon\tau, \tau)$, x in $\mathfrak{F} \cap \mathfrak{A}$, $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$, and y in $\mathfrak{D}^*$,

(4.2) $$|J_y^*(x, \tilde{x}) - J_y^*(x, \tilde{x})| < \delta[1 + E_L^*(x, \tilde{x})] .$$

By (4.1) and (4.2) we have for x in $\mathfrak{F} \cap \mathfrak{A}$, $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$, and y in $\mathfrak{D}^*$,

$$J_y(x, \tilde{x}) - J_y(\tilde{x}, \tilde{x}) - \tau\mathfrak{R}(x, \tilde{x})$$
$$= J_y^*(x, \tilde{x}) - J_y^*(\tilde{x}, \tilde{x}) + E_y^*(x, \tilde{x}) - \tau|\hat{b} - \widehat{\widetilde{b}}|^2 - \tau E_L^*(x, \tilde{x})$$
$$\geqq -\delta[1 + E_L^*(x, \tilde{x})] + \tau E_L^*(x, \tilde{x}) - \frac{1}{2}\tau\varepsilon \geqq -\varepsilon\tau .$$

This implies

$$J_y(x, \tilde{x}) - J_y(\tilde{x}, \tilde{x}) \geqq \tau[\mathfrak{R}(x, \tilde{x}) - \varepsilon] .$$

We now prove (4.2). We first note that a $\hat{b}t\hat{x}z$-neighborhood $\mathfrak{F}$ of x_0 and a $\hat{b}t\hat{x}\hat{x}z$-neighborhood $\mathfrak{R}_0$ of x_0 may be diminished so that for x in $\mathfrak{F}$, x in $\mathfrak{R}_0$, and y in $\mathfrak{D}^*$,

(4.3)
$$\left| [g_\sigma - g_\sigma(\tilde{b}, \tilde{z}, x, y)\widetilde{\lambda}_{y\sigma} | \right.$$
$$+ \left| \int_{T^1}^{T^2} f(b, t, x, \dot{\tilde{x}}, z, y, \widetilde{\lambda}_y)dt \right.$$
$$\left. - \int_{T^1(\tilde{b},\tilde{z})}^{T^2(\tilde{b},\tilde{z})} f(\tilde{b}, t, \tilde{x}, \dot{\tilde{x}}, \tilde{z}, y, \widetilde{\lambda}_y)dt \right| < \frac{1}{4}\delta .$$

Now recall that x_0 is an extremal. Setting

$$\widetilde{f}_{\dot{x}^i} = f_{\dot{x}^i}(b, t, x, \dot{\tilde{x}}, z, y, \widetilde{\lambda}_y)$$

and

$$f_{\dot{x}^i}^0 = f_{\dot{x}^i}(b_0, t, x_0, \dot{x}_0, z_0, y, \lambda_{0y}) ,$$

we may again diminish $\mathfrak{F}$ and $\mathfrak{R}_0$ so that for x in $\mathfrak{F}$, $\tilde{x}$ in $\mathfrak{R}_0$, and y $\mathfrak{D}^*$,

$$\sup (\widetilde{f}_{\dot{x}^i} - f_{\dot{x}^i}^0)(\widetilde{f}_{\dot{x}^i} - f_{\dot{x}^i}^0) < \frac{1}{2}\delta^2 \min (1, |T^2 - T^1|^{-2})$$

(4.4)
$$\left| f_{\dot{x}^i}^0(x^i - \tilde{x}^i)\Big|_{T^1}^{T^2} \right| + \left| \int_{T^1}^{T^2}\Big(\frac{d}{dt}f_{\dot{x}^i}^0\Big)(x^i - \tilde{x}^i)dt \right| < \frac{1}{4}\delta .$$

Setting

$$a_i = \dot{x}^i - \dot{\tilde{x}}_i \text{ and } c_i = \tilde{f}_{\dot{x}^i} - f^0_{\dot{x}^i},$$

we add and subtract the same terms and integrate by parts to obtain from (4.4) for x in $\mathfrak{F}$, $\tilde{x}$ in $\mathfrak{R}_0$, and y in $\mathfrak{D}^*$,

$$\left| \int_{T^1}^{T^2} (\dot{x}^i - \dot{\tilde{x}}^i) \tilde{f}_{\dot{x}^i} dt \right|$$

$$\leqq \left| \int_{T^1}^{T^2} (a_i c_i + |c| - |c|) dt \right| + \left| \int_{T^1}^{T^2} (\dot{x}^i - \dot{\tilde{x}}^i) f^0_{\dot{x}^i} dt \right|$$

$$\leqq \left| \int_{T^1}^{T^2} \{ [(|a|^2 + 1)(|c|^2 + |c|^2)]^{1/2} - \sqrt{2}\,|c| + \sqrt{2}\,|c| - |c| \} dt \right.$$

$$+ \left| f^0_{\dot{x}^i}(x^i - \tilde{x}^i) \Big|_{T^1}^{T^2} \right| + \left| \int_{T^1}^{T^2} \left(\frac{d}{dt} f^0_{\dot{x}^i} \right) (x^i - \tilde{x}^i) dt \right|$$

$$\leqq \delta E_L^*(x, \tilde{x}) + \frac{3}{4} \delta .$$

Combining (4.3) and (4.5), we prove (4.2) and hence the theorem.

We continue by giving the analogue of Theorem 4.1 in the parametric case. The appropriate Weierstrass condition $\mathfrak{R}$ is the same as the one we have been discussing with $\hat{L}(\dot{x}) = |\dot{x}|$ playing the role of $L(\dot{x}) = [1 + |\dot{x}|^2]^{1/2} - 1$. The corresponding E-function is represented by the formulas

$$E_{\hat{L}}(u, v) = \int_0^1 (1 - \theta) |w| |\mathit{\Pi}|^2 \sin^2 \phi\, d\theta$$

where $w(\theta) = u + \theta(v - u)$, $\mathit{\Pi} = v - u$, and $\phi(\theta)$ is

the angle between $\mathit{\Pi}$ and $w(\theta)$;

(4.6)
$$E_{\hat{L}}(u, v) = |v|[1 - \cos \theta]$$

where θ is the angle between v and u,

which are useful in the proof of Theorem 4.2. It is convenient to fix the parameterization of the admissible arcs $x\colon b, x(t), z$ so that

$$0 \leqq t \leqq 1 \text{ and } |\dot{x}(t)| = \text{length of } x(t) .$$

This is the parameterization which is referred to when a $\hat{b}\hat{x}\hat{x}z$-neighborhood of x_0 is specified. Invariance of the functionals under reparameterization implies some well-known identities. For their description we refer the reader to Hestenes [3], 78–82.

Let X_0^1 and X_0^2 be the end points of the admissible arc $x_0\colon b_0, x_0(t), z_0$. We assume that $0 < N = n - R$ and that

$$(X_0^{11}, X_0^{12}, \cdots X_0^{1N}) \neq (X_0^{21}, X_0^{22}, \cdots, X_0^{2N}).$$

In the applications this corresponds to the assumption of distinct end conditions. Similarly the nonintersection hypothesis of the applica-

tions corresponds to the supposition that $(x_0^1(t),\, x_0^2(t),\, \cdots,\, x_0^N(t))$ does not intersect itself, which we now make. Given admissible arcs $x\colon b,\, x(t),\, z;\, \tilde{x}\colon b,\, \tilde{x}(t),\, \tilde{z},$ and y in $\mathfrak{D}^*$, we define

$$\hat{J}_y^*(x,\, \tilde{x}) = \tilde{\lambda}_{y\sigma} g_\sigma + \int_0^1 \dot{x}^i(t) f_{\dot{x}i}(b,\, x(t),\, \dot{\tilde{x}}(t),\, z,\, y,\, \tilde{\lambda}_y) dt$$

$$\hat{E}_y^*(x,\, \tilde{x}) = \int_0^1 f(b,\, x(t),\, \dot{x}(t),\, z,\, y,\, \tilde{\lambda}) dt - \hat{J}_y^*(x,\, \tilde{x})$$

$$J_y(x,\, \tilde{x}) = \hat{J}_y^*(x,\, \tilde{x}) + \hat{E}_y^*(x,\, \tilde{x})$$

$$E_{\hat{L}}^*(x,\, \tilde{x}) = \int_0^1 |\dot{x}(t)|\, dt - \int_0^1 \frac{\dot{x}^i(t)\dot{\tilde{x}}^i(t)}{|\dot{\tilde{x}}(t)|} dt\, .$$

THEOREM 4.2 *Let $x_0\colon b_0,\, x_0(t),\, z_0$ be a terminally admissible extremal along which the parametric Weierstrass condition $\mathfrak{N}$ holds. Given a positive constant ε, there are a $\hat{b}\hat{x}$-neighborhood $\mathfrak{F}$ of x_0, a $\hat{b}\hat{x}\hat{z}$-neighborhood $\mathfrak{R}_0$ of x_0, and a positive constant η such that given x in $\mathfrak{F} \cap \mathfrak{A}$, $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$, and y in $\mathfrak{D}^*$, we have either*

$$\left| \int_0^t |\dot{x}(s)|\, ds - \int_0^t |\dot{\tilde{x}}(s)|\, ds \right| < \varepsilon$$

or

$$J_y(x,\, \tilde{x}) - J_y(\tilde{x},\, \tilde{x}) > \eta\, .$$

In order to prove this theorem, let a positive constant ε be given. Suppose we have shown that given $\varepsilon^* > 0$, there are a $\hat{b}\hat{x}$-neighborhood $\mathfrak{F}$ of x_0 and a $\hat{b}\hat{x}\dot{x}$-neighborhood $\mathfrak{R}_0$ of x_0 such that given x in $\mathfrak{F} \cap \mathfrak{A}$, $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$, and y in $\mathfrak{D}^*$,

(4.7) $$|\hat{J}_y^*(x,\, \tilde{x}) - \hat{J}_y^*(\tilde{x},\, \tilde{x})| < \varepsilon^*[1 + E_{\hat{L}}^*(x,\, \tilde{x})]\, .$$

The Weierstrass condition $\mathfrak{N}$ (See (4.6).) and our choice of parametrization imply that given $\varepsilon_1 > 0$, the neighborhoods $\mathfrak{F}$ and $\mathfrak{R}_0$ may be diminished so that there is a positive constant C such that for x in $\mathfrak{F} \cap \mathfrak{A}$, $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$, and y in $\mathfrak{D}^*$,

(4.8) $$\int_0^1 |\dot{x}(t) - \dot{\tilde{x}}(t)|\, dt \leqq \varepsilon_1 + C E_{\hat{L}}^*(x,\, \tilde{x})\, .$$

Furthermore the constant C may be taken independent of sufficiently small ε_1, $\mathfrak{F}$, $\mathfrak{R}_0$, as we now agree to do.

Let τ be the positive constant occuring in the statement of the Weierstrass condition $\mathfrak{N}$, and let ε_2 be a positive constant such that $\varepsilon_2 C < (1/2)\varepsilon$. We choose the neighborhoods $\mathfrak{F}$ and $\mathfrak{R}_0$ so small that (4.8) holds with $\varepsilon_1 = (1/2)\varepsilon$. Diminish the neighborhoods $\mathfrak{F}$ and $\mathfrak{R}_0$ if

necessary so that (4.7) holds with $\varepsilon^* = \min((1/2)\tau, (1/4)\tau\varepsilon_2)$. Suppose that $E_{\hat{L}}^*(x, x) < \varepsilon_2$. By (4.8) one has

$$\left| \int_0^t |\dot{x}(s)| \, ds - \int_0^t |\dot{x}(s)| \, ds \right| \leqq \varepsilon_1 + \varepsilon_2 C < \varepsilon .$$

On the other hand, suppose $E_{\hat{L}}^*(x, \hat{x}) \geqq \varepsilon_2$. By (4.7) and the parametric Weierstrass condition $\mathfrak{N}$ one has

$$\begin{aligned}
J_y(x, \tilde{x}) - J_y(\tilde{x}, \tilde{x}) &= \hat{J}_y^*(x, \tilde{x}) - \hat{J}_y^*(\tilde{x}, \tilde{x}) + \hat{E}_y^*(x, \tilde{x}) \\
&> -\varepsilon^*[1 + E_{\hat{L}}^*(x, \tilde{x})] + \tau E_{\hat{L}}^*(x, \tilde{x}) \\
&\geqq (\tau - \varepsilon^*) E_{\hat{L}}^*(x, \tilde{x}) - \varepsilon^* \geqq (\tau - \varepsilon^*)\varepsilon_2 - \varepsilon^* \\
&\geqq \frac{1}{4}\tau\varepsilon_2 = \eta > 0 .
\end{aligned}$$

With this in mind, we complete the proof by proving (4.7). Let a positive constant ε^* be given. Fix y_0 in $\mathfrak{D}^*$. By the compactness of $\mathfrak{D}^*$ it suffices to prove (4.7) for a neighborhood $\mathfrak{N}^*$ of y_0. Let ε_1 and C be positive constants such that (4.8) holds with neighborhoods $\mathfrak{F}$ and $\mathfrak{R}_0$. We may diminish these neighborhoods if necessary so that there are continuous functions $C_i(b, x)$ such that:

For x in $\mathfrak{F} \cap \mathfrak{A}$, $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$, and y in $\mathfrak{N}^*$,

(i) $\displaystyle\sup_t |f_{\dot{x}^j}(b, x(t), \tilde{x}(t), z, y, \tilde{\lambda}) - C_j(b, x(t) - \tilde{x}(t))| < \varepsilon_3$

where $\varepsilon_3 = \dfrac{\varepsilon^*}{2n} \min(\varepsilon_1^{-1}, C^{-1}), \quad j = 1, 2, \cdots, N$.

and

(ii) The integral $\sum_{j=1}^{N} \int_0^1 C_j(b, x) dx^j$ depends only on its end points.

(Recall our parameterization convention. To verify that such functions exist, one may check the proof of Lemma 5.1, Hestenes [9], 62.) Further diminish these neighborhoods so that for x in $\mathfrak{F} \cap \mathfrak{A}$, $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$, and y in $\mathfrak{N}^*$,

$$\left| \sum_{j=1}^{N} \int_0^1 (\dot{x}^j - \dot{\tilde{x}}^j) C_j(b, x - \tilde{x}) dt \right| < \frac{1}{4}\varepsilon^*$$

and

$$|[g_\sigma - g_\sigma(\tilde{b}, \tilde{z}, y)]\tilde{\lambda}_{y\sigma}| < \frac{1}{4}\varepsilon^* .$$

Finally by the Euler equations the neighborhoods may again be

diminished so that for x in $\mathfrak{F} \cap \mathfrak{A}$, $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$, and y in $\mathfrak{N}^*$,

$$\sup_t |f_{\tilde{x}j}(b, x(t), \tilde{\tilde{x}}(t), z, y, \tilde{\lambda})| < \varepsilon_3,$$

$j = N + 1, N + 2, \cdots, n$. Thus for x in $\mathfrak{F} \cap \mathfrak{A}$, $\tilde{x}$ in $\mathfrak{R}_0 \cap \mathfrak{A}$, and y in $\mathfrak{N}^*$, one has upon setting

$$A_i = f_{\tilde{x}i}(b, x(t), \tilde{\tilde{x}}(t), z, y, \tilde{\lambda}),$$

$$|\hat{J}_y^*(x, \tilde{x}) - \hat{J}_y^*(\tilde{x}, \tilde{x})| = \left| [g_\sigma - g_\sigma(\tilde{b}, \tilde{z}, y)] + \int_0^1 (\dot{x}^i - \dot{\tilde{x}}^i) A_i dt \right|$$

$$< \frac{1}{4}\varepsilon^* + \left| \sum_{j=1}^N \int_0^1 (\dot{x}^j - \dot{\tilde{x}}^j) C_j(b, x - \tilde{x}) dt \right.$$

$$+ \sum_{j=1}^N \int_0^1 (\dot{x}^j - \dot{\tilde{x}}^j)[A_j - C_j(b, x - \tilde{x})] dt$$

$$+ \left. \sum_{j=N+1}^n \int_0^1 (\dot{x}^j - \dot{\tilde{x}}^j) A_j dt \right|$$

$$\leqq \frac{1}{2}\varepsilon^* + n\varepsilon_3 E_{\hat{L}}^*(x, \tilde{x})$$

$$\leqq \varepsilon^*[1 + E_{\hat{L}}^*(x, \tilde{x})] .$$

This proves (4.7) and hence the theorem.

5. **The extremal family.** Given an arc x_0 satisfying certain hypotheses, the purpose of this section is to prove the existence of a suitable family of nearby extremals. The sufficiency theorem shows these arcs to be solutions of a family of perturbed problems. The special properties and uniform nature of these solutions are essential to the proofs in § 8. We require an implicit function theorem in a slightly more general context than is usually necessary. The proof (given in the author's dissertation [14], Theorem 1–5.1) is a consequence of standard techniques.

THEOREM 5.1. *Let (z, x, u) be variables in an open set $\mathfrak{R}$ contained in $\mathfrak{D} \times E^m \times E^n$, and let $\mathfrak{K}$ be a compact set in $\mathfrak{D} \times E^m$. Suppose there exist continuous functions $u_0^i(z, x)$, $i = 1, 2, \cdots, n$ defined on $\mathfrak{K}$ and such that $(z, x, u(z, x))$ is in $\mathfrak{R}$ for all $(z, x) \in \mathfrak{K}$. Let $f_j(z, x, u)$ be realvalued functions defined on $\mathfrak{R}$ which are continuous and have continuous partial derivatives with respect to the u-variables and are such that on $\mathfrak{K}$,*

$$f(z, x, u_0(z, x)) \equiv 0$$

and

$$\det \|f_{iuj}\| \neq 0 .$$

There are a neighborhood $\mathfrak{N}$ of $\mathfrak{K}$ and continuous functions $u^i(z, x)$ defined on $\mathfrak{N}$ such that

$$u(z, x) \equiv u_0(z, x) \text{ on } \mathfrak{K}$$

and

$$f(z, x, u(z, x)) \equiv 0 \text{ on } \mathfrak{N}.$$

Also $u(z, x)$ is unique in the sense that there is a positive constant ε such that $f(z, x, u) = 0$ and $|u - u(z, x)| < \varepsilon$ imply $u = u(z, x)$. If $f_i(z, x, u)$, $i = 1, 2, \cdots, n$ has k continuous xu-derivatives, then $u_i(z, x)$, $i = 1, 2, \cdots, n$ has k continuous x-derivatives.

THEOREM 5.2. *Let the smooth, nonsingular admissible arc x_0: b_0, x_0, (t) z_0 satisfy (2.1), (2.2), and (2.3). There is a unique z-parameter family*

$$(5.1) \qquad\qquad x(z): b(z), x(t, z), z$$

of smooth, nonsingular terminally admissible z-extremals satisfying the transversality condition and containing x_0 at $z = z_0$. There are positive constants δ and δ^* such that $b(z)$, $x(t, z)$, and $\dot{x}(t, z)$ are continuous in all their arguments and such that (5.1) satisfies the Euler equations for $T^1(b_0, z_0) - \delta \leq t \leq T^2(b_0, z_0) + \delta$, $d(z, z_0) < \delta^*$ and y in $\mathfrak{D}^*$.

In order to prove this theorem, we introduce the cannonical variables $btxpzy\lambda$ instead of $btx\dot{x}zy\lambda$ by setting

$$(5.2) \qquad\qquad p_i = f_{\dot{x}^i}(b, t, x, \dot{x}, z, y, \lambda).$$

Since x_0 is nonsingular and has no corners, the implicit function theorem implies that (5.2) has a solution

$$\dot{x} = P(b, t, x, p, z, y, \lambda)$$

on a $btxpzy\lambda$-neighborhood of the elements $btxpzy\lambda$ along x_{0y}. Furthermore $P = (P^1, P^2, \cdots, P^n)$ is continuously differentiable in its Euclidean arguments and is unique in the sense described in the implicit function theorem. Now define

$$H(b, t, x, p, z, y, \lambda) = p_i P^i(b, t, x, p, z, y, \lambda) - f^i(b, t, x, P, z, y, \lambda)$$

and note that the Euler equations are equivalent to the Hamiltonian equations

$$\dot{x}^i = H_{p_i}(b, t, x, P, z, y, \lambda)$$

and

$$\dot{p}_i = -H_{x^i}(b, t, x, P, z, y, \lambda).$$

The proof of the embedding theorem for differential equations given

in the appendix of Hestenes [3] uses only the fact that the parameter is in a metric space. We may consequently apply this theorem to obtain for $\beta = (\beta_1, \beta_2, \cdots, \beta_{2n})$ a unique $b\beta zy\lambda$-family

$$x(t, b, \beta, z, y, \lambda), \quad \dot{x}(t, b, \beta, z, y, \lambda)$$

of nonsingular extremals. This family contains $x_0(t), \dot{x}_0(t)$ for $(b, \beta, z, y, \lambda) = (b_0, \beta_0, z_0, y, \lambda_{0y})$ and $T^1(b_0, z_0) \leqq t \leqq T^2(b_0, z_0)$. The family is also continuous and has continuous partial derivatives with respect to its Euclidean arguments. Moreover the matrix of partial derivatives with respect to the β-variables is nonsingular along x_{0y}, y in $\mathfrak{D}^*$.

We wish to apply the implicit function theorem to the system of equations

$$\lambda_{x_y\sigma} = G_\sigma(x_y)$$

(5.3)
$$x^i(T^s) = X^{is}$$
$$A_h = g_h + [(f - \dot{x}^i f_{\dot{x}^i}) T_h^s + f_{\dot{x}^i} X_h^{is}]_{s=1}^{s=2} + \int_{T_1}^{T_2} f_h dt = 0$$

at $b_0, \beta_0, z_0, y_0, \lambda_{0y}$ for some y_0 in $\mathfrak{D}^*$. Suppose we have shown that the matrix

(5.4)
$$\left\| \begin{array}{ccc} -\delta_{\mu\sigma} + \dfrac{\partial}{\partial\lambda_\mu}G_\sigma & \dfrac{\partial}{\partial\beta_k}G_\sigma & \dfrac{\partial}{\partial b^l}G_\sigma \\[2ex] x^i_{\lambda_\mu}(T^s) & x^i_{\beta_k}(T^s) & x^i_l(T^s) + \dot{x}^i T_l^s - X_l^{is} \\[2ex] \dfrac{\partial}{\partial\lambda_\mu}A_h & \dfrac{\partial}{\partial\beta_k}A_h & \dfrac{\partial}{\partial b^l}A_h \end{array} \right\|$$

of partial derivatives with respect to the $\lambda\beta b$-variables is nonsingular at x_{0y} for some y_0 in $\mathfrak{D}^*$. By the implicit function theorem there is a continuous solution $b(z, y), \beta(z, y), \lambda(z, y)$ of (5.3) in a neighborhood of z_0, y_0. Fixing $y = y_0$, we see that the hypothesis in (2.2) implies that

$$x(z): b(z, y_0), x(t, b(z, y_0), \beta(z, y_0), z, y_0, \lambda(z, y_0)), z$$

is the required family of nonsingular extremals.

It is therefore sufficient to show that the supposition that the matrix (5.4) is singular for all y in $\mathfrak{D}^*$ leads to a contradiction. Let y be a fixed but arbitrary element in $\mathfrak{D}^*$. Suppose there are constants a_μ, c_k, and $d^l, \mu = 0, 1, \cdots, m; k = 1, 2, \cdots, 2n; l = 1, 2, \cdots, r$ not all zero such that at x_{0y},

$$-a_\sigma + a_\mu\frac{\partial}{\partial\lambda_\mu}G_\sigma + c_k\frac{\partial}{\partial\beta_k}G_\sigma + d^l\frac{\partial}{\partial b^l}G_\sigma = 0$$

(5.5)
$$a_\mu x^i_{\lambda_\mu}(T^s) + c_k x^i_{\beta_k}(T^s) + d^l x^i_l(T^s) = [X_l^{is} - \dot{x}^i T_l^s]d^l$$

$$a_\mu\frac{\partial}{\partial\lambda_\mu}A_h + c_k\frac{\partial}{\partial\beta_k}A_h + d^l\frac{\partial}{\partial b^l}A_h = 0 \ .$$

Let the variation $\gamma: d^h$, $\eta^i(t)$ be defined by

$$\eta^i(t) = a_\sigma x^i_{\lambda_\sigma}(t) + c_k x^i_{\beta_k}(t) + d^l x^i_l(t) \, .$$

The second set of equations above imply that γ is an admissible varia-
tion. Multiply the equations involving partial derivatives of G_σ by
$I'_\sigma(\gamma)$ evaluated along x_{0y} and multiply the equations involving A_h by d^h.
Adding the result and setting at x_{0y}, γ,

$$S_1 = I'_\sigma\Big[a_\mu\frac{\partial}{\partial\lambda_\mu}G_\sigma + c_k\frac{\partial}{\partial\beta_k}G_\sigma + d^l\frac{\partial}{\partial b^l}G_\sigma\Big]$$

$$S_2 = -a_\sigma I'_\sigma + d^h\Big[a_\sigma\frac{\partial}{\partial\lambda_\sigma}A_h + c_k\frac{\partial}{\partial\beta_k}A_h\Big]$$

$$S_3 = d^h d^l\frac{\partial}{\partial b^l}A_h \, ,$$

we obtain

(5.6) $0 = S_1 + S_2 + S_3 \, .$

After collecting terms, we have at x_{0y}, γ,

(5.7) $S_1 = I'_\sigma I'_\mu G_{\sigma\mu} \, .$

On the other hand, using the Euler equations and accessory end con-
ditions, one obtains at x_{0y} after setting

$$\eta^{*j} = a_\sigma x^j_{\lambda_\sigma} + c_k x^j_{\beta_k} \, ,$$

(5.8)
$$S_2 = [f_{x^i}\eta^{*i}T^s_h d^h + f_{x^i x_j}\eta^i\dot\eta^{*j} + f_{\dot x^i x_j}\eta^i\dot\eta^{*j}]^{s=2}_{s=1}$$
$$+ \int_{T_1}^{T_2}(f_{hx^j}d^h\eta^{*j} + f_{h\dot x^j}d^h\dot\eta^{*j})dt$$

and

$$S_3 = g_{hl}d^h d^l + \{f_{x^i}d^l\dot x^i_l T^s_h d^h + f_{x^i x_j}\eta^i d^l x^j_l + f_{\dot x^i x_j}\eta^i d^l\dot x^j_l + f_{\dot x^i l}\eta^i d^l$$
$$+ [(f - \dot x^i f_{x^i})T^s_{hl} + f_{x^i}X^{is}_{hl} + f_t T^s_h T^s_l$$
(5.9)
$$+ f_{x^i}X^{is}_h T^s_l - \dot x^i f_{x^i}T^s_h T^s_l + 2f_h T^s_l]d^h d^l\}^{s=2}_{s=1}$$
$$+ \int_{T_1}^{T_2}(f_{hl}d^h d^l + f_{hx^i}d^h d^l x^i_l + f_{h\dot x^i}d^h d^l\dot x^i_l)dt \, .$$

Combining (5.7), (5.8), and (5.9) with (5.6), we use the accessory Euler
equations and accessory end conditions to obtain along x_{0y},

$$0 = S_1 + S_2 + S_3$$
(5.10)
$$= Q(\beta) + \omega_{\dot\gamma i}\eta^i\Big|^{T2}_{T1} + \int_{T_1}^{T_2}[f_{hl}d^h d^l + f_{hx^i}d^h\eta^i + f_{h\dot x^i}d^h\dot\eta^i]dt$$
$$+ I'_\sigma I'_\mu G_{\sigma\mu} = G''(\gamma) \, .$$

It now remains to show that γ cannot be zero. Suppose then that $\gamma = 0$. This means that $d^h = 0$ and $\eta^i(t) \equiv 0$. Collecting terms in the equations (5.5) which involve G_σ, we consequently obtain at x_{0y},

$$0 = - a_\sigma + G_{\sigma\mu} \int_{T^1}^{T^2} (f_{\sigma x^i} \eta^i + f_{\sigma \dot{x}^i} \dot{\eta}^i) dt = - a_\sigma \ .$$

This implies

$$c_k x^i_{\beta_k}(t) \equiv 0 \ .$$

Since not all of the a_σ, c_k, and d^h are zero, some of the c_k must not be zero. This contradicts the selection of the family $x(t, b, \beta, z, y, \lambda)$, $\dot{x}(t, b, \beta, z, y, \lambda)$ to have a nonsingular matrix of partial derivatives with respect to the β-variables along x_{0y}, y in $\mathfrak{D}^*$. Consequently, γ cannot be zero.

Since γ is not zero, (5.10) is a contradiction to the assumption that the matrix (5.4) is singular at x_{0y} for all y in $\mathfrak{D}^*$. This proves the theorem.

The exact statement of the uniqueness of the family $x(z)$ of extremals is given here as a corollary. This result holds because the family $x(z)$ was obtained by an application of the implicit function theorem and the embedding theorem for differential equations.

COROLLARY 5.1. *The family $x(z)$: $b(z)$, $x(t, z)$, z is unique in the sense that there is a positive constant ε such that if $\tilde{x}$: $\tilde{b}$, $\tilde{x}(t)$, z satisfies the transversality condition and Euler equations and*

$$| b(z) - \tilde{b} | + \sup_{T' \leq t \leq T''} [\, | x(t, z) - \tilde{x}(t) | + | \dot{x}(t, z) - \dot{\tilde{x}}(t) | \,] < \varepsilon \ ,$$

then $x(z) = \tilde{x}$.

This next corollary is proved by checking that the required continuity and differentiability is present at each step of the preceding theorem.

COROLLARY 5.2. *Let z be of the form (z^*, z) where z^* is Euclidean and z is in an arbitrary metric space. If g_σ, T^s, X^{is}, and f_σ have k continuous $btx\dot{x}z^*$-derivatives ($k \geq 2$) near x_{0y} and if G has k continuous ($k \geq 2$) Iz^*-derivatives near $(I(x_{0y}), z_0, y)$, then $x(t, z^*, z)$ and $\dot{x}(t, z^*, z)$ have $k - 1$ continuous tz^*-derivatives.*

Similarly it is easy to check that Corollary 5.3 below holds in the case where $T^1 < T^2$ are constants and no assumption is made concerning partial derivatives with respect to t.

COROLLARY 5.3. *Let the additional assumptions indicated above hold, and let z be of the same form as in Corollary 5.1. If g_σ, X^{is},*

and f_σ have $k \geq 2$ continuous $bx\dot{x}z^*$-derivatives near x_{0y} and if $G(I, z, y)$ has k continuous Iz^*-derivatives near $(I(x_{0y}), z_0, y)$, then $x(t, z^*, z)$ and $\dot{x}(t, z^*, z)$ have $k - 1$ continuous z^*-derivatives.

6. **An admissible variation.** The existence of a suitable admissible variation is fundamental to the indirect method. In this section we obtain such a variation.

Let $x_q: b_q, x_q(t), z_q$ and $x_{0q}: b_{0q}, x_{0q}(t), z_q$ be sequences of admissible arcs. Suppose there is a positive constant δ such that $x_0(t)$ and $x_{0q}(t)$ have two continuous derivatives on

$$[T', T''] = [T^1(b_0, z_0) - \delta, T^2(b_0, z_0) + \delta] .$$

We also assume that

(6.1) $$\lim \dot{x}_{0q}(t) = \dot{x}_0(t)$$

uniformly in $T' \leq t \leq T''$ and that

(6.2) $$\lim \Re(x_q, x_{0q}) = \lim |\, \hat{b}_q - \hat{b}_0 \,| = 0 .$$

By (6.1), (6.2), and the admissible end conditions at $T^1(b, z)$, the sequence of functions $x_q(t)$ converges uniformly to $x_0(t)$, $T' \leq t \leq T''$. We are motivated by the method of Hestenes [3], 152–159. We consider a vector like $\hat{b}_q - \hat{b}_{0q}$ as an n-vector by adding zero components if necessary. Thus we may define

(6.3)
$$\beta_q^h = (2^{1/2}d_q)^{-1}(b_q^h - b_{0q}^h)$$
$$\eta_q^i(t) = (2^{1/2}d_q)^{-1}(x_q^i(t) - x_{0q}^i(t))$$

where $d_q = L(\hat{b}_q - \hat{b}_{0q}) + \int_{T'}^{T''} L(\dot{x}_q - \dot{x}_{0q})dt$. Because we are dealing with the parametric form of the accessory minimum problem, we define for $\gamma: \beta, \eta(t)$ in $\mathfrak{B}$,

$$\mathscr{D}(\gamma) = L(\hat{\beta}) + \int_{T'}^{T''} L(\hat{\eta})dt ,$$

which differs from the analogous equation on p. 152 of Hestenes [3]. The changes in the statements of the results and the proofs are obvious. Thus we may read off Lemmas 6.1 and 6.2 below from pp. 154, 156, 158 of Hestenes [3]. Some of the limits in Lemmas 6.1 and 6.2 may hold only in subsequence, which for convenience we again denote by the original sequence.

LEMMA 6.1. *Let* $w_q^{*2} = (1/2)L(\hat{b}_q - \hat{b}_{0q})$ *and* $w_q(t)^2 = (1/2)L(\dot{x}_q(t) - \dot{x}_{0q}(t))$. *We have*

$$\left|\frac{\widehat{\beta}_q}{w_q^*}\right|^2 + \int_{T'}^{T''}\left|\frac{\dot{\eta}_q(t)}{w_q(t)}\right|^2 dt = 1$$

$$\lim d_q = 0$$

$$\lim \dot{x}_q(t) = \dot{x}_0(t) \ \ almost \ \ uniformly \ \ in \ \ T' \leqq t \leqq T'' \ .$$

There are absolutely continuous functions $\eta_0^i(t)$ *having* $\dot{\eta}_0^i(t)$ *in* $L_2(T', T'')$ *and constants* β_0^h, $h = 1, 2, \cdots, R$ *such that*

$$\lim \beta_q^h = \beta_0^h \qquad h = 1, 2, \cdots, R$$

$$\lim \eta_q^i(t) = \eta_0^i(t) \ \ uniformly \ \ in \ \ T' \leqq t \leqq T''$$

LEMMA 6.2. *If* $h_q^{ij}(t)$ *are continuous functions on* $T' \leqq t \leqq T''$ *converging uniformly to* $h_0^{ij}(t)$, *then we have*

$$\lim \int_{T'}^{T''} h_q^{ij}\dot{\eta}_q^i dt = \int_{T'}^{T''} h_0^{ij}\dot{\eta}_0^i dt \ .$$

If $\dot{x}_q$ *converges uniformly to* $\dot{x}_0$ *on a measurable subset* $\mathfrak{M}$ *of* $T' \leqq t \leqq T''$ *and the matrix* $\| h_0^{ij}(t) \|$ *is nonnegative on* $\mathfrak{M}$, *then we have*

$$\lim \int_{\mathfrak{M}} h_q^{ij}\dot{\eta}_q^i\dot{\eta}_q^j dt \geqq \int_{\mathfrak{M}} h_0^{ij}\dot{\eta}_0^i\dot{\eta}_0^j dt \ .$$

It remains to check that there are additional constants β_0^h, $h = R+1, R+2, \cdots, r$, such that $\eta_0(t)$ satisfies the accessory end conditions. Let

$$T_q^s = T^s(b_q, z_q) \qquad T_{0q}^s = T^s(b_{0q}, z_q)$$

and

$$X_q^{is} = X^{is}(b_q, z_q) \qquad X_{0q}^{is} = X^{is}(b_{0q}, z_q) \ .$$

At $T^1(b_0, z_0)$ the accessory end conditions are an immediate consequence of the calculation

$$\lim \eta_q^i(T_q^1) = \lim \frac{x_q^i(T_q^1) - x_{0q}^i(T_q^1)}{\sqrt{2}\,d_q}$$

$$= \lim\left\{\frac{X_q^{i1} - X_{0q}^{i1}}{\sqrt{2}\,d_q} - \frac{x_{0q}^i(T_q^1) - x_{0q}^i(T_{0q}^1)}{\sqrt{2}\,d_q}\right\}$$

$$= \lim\left\{X_h^{i1}(\theta_q)\frac{b_q^h - b_{0q}^h}{\sqrt{2}\,d_q} - \dot{x}_{0q}^i(T^1(\theta_q))T_h^1(\theta_q)\frac{b_q^h - b_{0q}^h}{\sqrt{2}\,d_q}\right\}$$

$$= [X_h^{i1}(b_0, z_0) - \dot{x}_0^i T_h^1(b_0, z_0)]\beta_0^h$$

where $\theta_q = (b_q + \theta(b_{0q} - b_q), z_q)$ and $\theta \in (0, 1)$ is a generic constant. The similar procedure holds at $T^2(b_0, z_0)$ after we first verify that $\lim \beta_q^h = \beta_0^h$ exists, $h = R+1, R+2, \cdots, r$. Hence,

$$\gamma_0 \colon \beta_0, \ \eta_0(t)$$

is an admissible variation.

7. **The proof of Theorem 2.1.** In order to prove this theorem, we first observe that there is a unique family of extremals having the properties described in Theorem 5.1. This means that given a positive constant ε, there is a positive constant δ such that for every z with $d(z, z_0) < \delta$, there is a terminally admissible z-extremal $x_z\colon b_z,\, x_z(t),\, z$ of the above family with $\rho(x_z, z_0) < \varepsilon$. The proof is by contradiction. Consequently suppose the theorem is false. This means that there is a sequence z_q tending to z_0 for which there are z_q-admissible arcs $x_q\colon b_q,\, x_q(t),\, z_q$ and terminally admissible z_q-extremals $x_{0q}\colon b_{0q},\, x_{0q}(t),\, z_q$ such that for every y in $\mathfrak{D}^*$,

$$0 \geqq G(x_{qy}) - G(x_{0qy}) \qquad \rho(x_q, x_0) < \frac{1}{q}$$

$$|\, b_{0q} - b_0 \,| + \sup_t |\, x_{0q}(t) - x_0(t) \,| + \sup_t |\, \dot{x}_{0q}(t) - \dot{x}_0(t) \,| < \frac{1}{q}$$

and either the convexity assumption holds or $|\, I(x_{qy}) - I(x_{0qy}) \,| < 1/q$.

Given y in $\mathfrak{D}^*$, let

$$T_q^s = T^s(b_q, z_q) \qquad T_{0q}^s = T^s(b_{0q}, z_q)$$
$$\lambda_{0y\sigma} = G_\sigma(x_{0y}) \qquad \lambda_{0qy\sigma} = G_\sigma(x_{0qy})\,,$$

and define $J_y''(\gamma)$ to be the second variation of

$$\lambda_{0y\sigma} g_\sigma + \int_{T^1}^{T^2} f(b, t, x, \dot{x}, z_0, y, \lambda_{0y})dt$$

along x_0. We also make the following convention concerning our arcs. It is useful later on and gives their function parts a common domain. First note that by Theorem 5.1, there is a positive constant δ such that the extremals satisfy the Euler equations on $[T', T''] = [T_0^1 - \delta, T_0^2 + \delta]$. By proper selection of a subsequence if necessary we may assume that $[T_{0q}^1, T_{0q}^2] \subseteqq [T', T'']$ and $[T_q^1, T_q^2] \subseteqq [T', T'']$. The functions $x_{0q}(t)$ are already defined on $[T', T'']$ as solutions of the Euler equations. The convention is to extend if necessary the functions $x_q(t)$ to $[T', T'']$ by keeping then continuous and requiring that $\dot{x}_q(t) = \dot{x}_{0q}(t)$.

Our proof uses the following three lemmas, which are proved in Section 8.

LEMMA 7.1.

$$\lim \mathfrak{R}(x_q, x_{0q}) = 0\,.$$

This lemma permits us to replace x_q and x_{0q} by the convergent subsequences which were described in § 6 and which we again denote by

x_q and x_{0q}. In particular the variations γ_q defined by (6.3) converge uniformly on $T' \leqq t \leqq T''$ to an admissible variation $\gamma_0 : \beta_0, \eta_0(t)$.

LEMMA 7.2. *For each y in $\mathfrak{D}^*$ one has*

$$\lim d_q^{-2}[J_y^*(x_q, x_{0q}) - J_y^*(x_{0q}, x_{0q})] = J_y''(\gamma_0) - \int_{T_0^1}^{T_0^2} f_{\dot{x}^i \dot{x}^j} \dot{\eta}_0^i \dot{\eta}_0^j dt$$

where the arguments of the last integral are taken along x_{0y}.

LEMMA 7.3. *For each y in $\mathfrak{D}^*$ one has*

$$\lim \frac{I_\sigma(x_{qy}) - I_\sigma(x_{0qy})}{d_q} = 2^{1/2} I'_{y\sigma}(\gamma_0)$$

and

$$\lim [I_\sigma(x_{qy}) - I_\sigma(x_{0qy})] = 0 .$$

We now expand $G(I, z, y)$ by Taylor's theorem to obtain for large q, y in $\mathfrak{D}^*$,

$$0 \geqq G(I(x_{qy}), z_q, y) - G(I(x_{0qy}), z_q, y)$$
$$= [I_\sigma(x_{qy}) - I_\sigma(x_{0qy})]G_\sigma(I(x_{0qy}), z_q, y) + 2^{-1}[I_\sigma(x_{qy}) - I_\sigma(x_{0qy})]G_{\sigma\mu}$$

where the arguments of $G_{\sigma\mu}$ are $(I(x_{0qy}) + \theta[I(x_{qy}) - I(x_{0qy})], z_q, y)$, $\theta \in (0, 1)$ a generic constant. Setting for y in $\mathfrak{D}^*$,

$$f_{\dot{x}^i \dot{x}^j}^y = f_{\dot{x}^i \dot{x}^j}(b_0, t, x_0, \dot{x}_0, z_0, y, \lambda_{0y}) ,$$

we divide the above inequality by d_q^2 and observe that by Lemmas 7.2 and 7.3, we obtain in the limit

$$(7.2) \qquad 0 \geqq G_y''(\gamma_0) - \int_{T_0^1}^{T_0^2} f_{\dot{x}^i \dot{x}^j}^y \dot{\eta}_0^i \dot{\eta}_0^j dt + \liminf d_q^{-2} E_y^*(x_q, x_{0q}) .$$

We wish to show

$$(7.3) \qquad \liminf d_q^{-2} E_y^*(x_q, x_{0q}) \geqq \int_{T_0^1}^{T_0^2} f_{\dot{x}^i \dot{x}^j}^y \dot{\eta}_0^i \dot{\eta}_0^j dt .$$

Let $\mathfrak{M}$ be a measurable subset of $T' \leqq t \leqq T''$ on which $\dot{x}_q$ converges uniformly to $\dot{x}_0$. Setting $\phi_q = (b_q, t, x_q, \dot{x}_{0q} + \theta[\dot{x}_q - \dot{x}_{0q}], z_q, y, \lambda_{0qy})$, by Taylor's theorem we obtain on $\mathfrak{M}$ for large q,

$$d_q^{-2} E_f(b_q, t, x_q, \dot{x}_{0q}, \dot{x}_q, z_q, y, \lambda_{0qy}) = 2 \int_0^1 (1 - \theta) f_{\dot{x}^i \dot{x}^j}(\phi_q) \dot{\eta}_q^i \dot{\eta}_q^j d\theta .$$

Consequently Lemma 7.3 and the Legendre condition imply

$$(7.4) \qquad \begin{aligned} \liminf d_q^{-2} & \int_{\mathfrak{M}} E_f(b_q, t, x_q, \dot{x}_{0q}, \dot{x}_q, y, z_q, \lambda_{0qy}) dt \geqq \\ & \int_{\mathfrak{M}} f_{\dot{x}^i \dot{x}^j}^y \dot{\eta}_0^i \dot{\eta}_0^j dt . \end{aligned}$$

By our convention in extending the domain of x_q to $T' \leqq t \leqq T''$, η_q is constant on that portion of $T' \leqq t \leqq T''$ to which x_q has been extended. Thus the almost uniform convergence of $\dot{x}_q$ to $\dot{x}_0$ on $T' \leqq t \leqq T''$, (7.4), and the Weierstrass condition $\mathfrak{N}$ prove (7.3).

By combining (7.2) and (7.3) we show that

$$0 \geqq G_y''(\gamma_0) \text{ for all } y \text{ in } \mathfrak{D}^*$$

and hence that $\gamma_0 = 0$. This along with the Weierstrass condition and (7.2) imply

$$0 \geqq \liminf d_q^{-2} E_y^*(x_q, x_{0q}) \geqq \liminf \tau d_q^{-2} \int_{T_q^1}^{T_q^2} E_L(\dot{x}_{0q}, \dot{x}_q)dt$$

where τ is a positive constant. Thus Lemma 6.1 implies

$$0 \geqq \liminf \tau\tau^* d_q^{-2} E_L^*(x_q, x_{0q}) = \frac{1}{2}\tau\tau^* \liminf \int_{T'}^{T''} \left| \frac{\dot{\eta}_q}{w_q} \right|^2 dt = \frac{\tau\tau^*}{2}$$

where τ^* is another positive constant. This contradiction proves the theorem.

8. **Auxiliary lemmas.** In this section we complete the proof of the sufficiency theorem by proving Lemmas 7.1, 7.2, and 7.3.

LEMMA 7.1.

$$\lim \mathfrak{K}(x_q, x_{0q}) = 0 .$$

The proof of this lemma has two cases. In the first case, we have

$$\lim [I_\sigma(x_{qy}) - I_\sigma(x_{0qy})] = 0 \qquad y \text{ in } \mathfrak{D}^* ,$$

and in the second case, we have the convexity assumption occurring in the statement of the sufficiency theorem.

To prove the first case, we define x_q^*: b_q, $x_q^*(t)$, z_q by setting $x_q^*(t)$ equal to $x_q(t)$ on $T_q^1 \leqq t \leqq T_q^2$, by requiring $x_q^*(t)$ to be continuous, and by requiring $\dot{x}_q^*(t) = \dot{x}_0(t)$ on the remainder of $T' \leqq t \leqq T''$. This means

$$0 = [\lim (I_\sigma(x_{qy}^*) - I_\sigma(x_{0y}))]\lambda_{0y\sigma}$$
$$= \lim [I_\sigma(x_{qy}^*) - I_\sigma(x_{0y})]\lambda_{0y\sigma}$$
$$= \lim [J_y(x_q^*, x_0) - J_y(x_0, x_0)] .$$

By Theorem 4.1 one has

$$\limsup \mathfrak{K}(x_q^*, x_0) \leqq 0 ,$$

which means that

$$\lim \Re(x_q^*, x_0) = 0$$

since $\Re(x, x^*)$ is nonnegative. Thus the inequality

$$0 \leqq \Re(x_q^*, x_{0q}) \leqq 2[\Re(x_q^*, x_0) + \Re(x_0, x_{0q})]$$

and the fact that $\dot{x}_{0q}$ and b_{0q} tend uniformly to $\dot{x}_0$ and b_0 imply that

$$\lim \Re(x_q^*, x_{0q}) = 0 \ .$$

Since by our definition of x_q^*

$$0 \leqq \Re(x_q, x_{0q}) \leqq \Re(x_q^*, x_{0q}) \ ,$$

this proves

$$\lim \Re(x_q, x_{0q}) = 0 \ .$$

In the second case, the convexity assumption implies for all y in $\mathfrak{D}^*$,

$$0 \geqq G(I(x_{qy}), z_q, y) - G(I(x_{0qy}), z_q, y)$$
$$\geqq [I_\sigma(x_{qy}) - I_\sigma(x_{0qy})]G_\sigma(I(x_{0qy}), z_q, y) \ .$$

In the limit one has

$$0 \geqq \lim \sup [I_\sigma(x_{qy}) - I_\sigma(x_{0qy})]G_\sigma(x_{0qy})$$
$$= \lim \sup [J_y(x_q, x_{0q}) - J_y(x_{0q}, x_{0q})] \ .$$

By Theorem 4.1 this implies

$$\lim \sup \Re(x_q, x_{0q}) \leqq 0 \ ,$$

which means that

$$\lim \Re(x_q, x_{0q}) = 0$$

since $\Re(x, x^*)$ is nonnegative. This proves the lemma.

LEMMA 7.2. *For each y in $\mathfrak{D}^*$ one has*

$$\lim d_q^{-2}[J_y^*(x_q, x_{0q}) - J_y^*(x_{0q}, x_{0q})] = J_y''(\gamma_0) - \int_{T_0^1}^{T_0^2} f_{\dot{x}^i\dot{x}^j}\dot{\eta}_0^i\dot{\eta}_0^j dt$$

where the arguments of the last integral are taken along x_{0y}.

In order to prove this lemma, we set

$$X_q^{is} = X^{is}(b_q, z_q) \qquad X_{0q}^{is} = X^{is}(b_{0q}, z_q)$$
$$\varDelta X_q^{is} = X_q^{is} - X_{0q}^{is} \qquad \varDelta x_q^{is} = x_{0q}^i(T_q^s) - x_{0q}^i(T_{0q}^s)$$

and consider the difference

$$x_q^i(T_q^s) - x_{0q}^i(T_q^s) = \Delta X_q^{is} - \Delta x_q^{is} \ .$$

We continue to use $\theta \in (0, 1)$ as a generic constant. Applying Taylor's theorem to ΔX_q^{is} and Δx_q^{is}, we obtain for large q,

$$
\begin{aligned}
(8.1) \quad \Delta X_q^{is} &= 2^{1/2} d_q X_h^{is}(b_{0q}, z_q)\beta_q^h + d_q^2 X_{hl}^{is}\beta_q^h\beta_q^l \\
\Delta x_q^{is} &= 2^{1/2} d_q \dot{x}_{0q}^i(T_{0q}^s)T_h^s(b_{0q}, z_q)\beta_q^h + d_q^2[\ddot{x}_{0q}^i T_h^s T_l^s + \dot{x}_{0q}^i T_{hl}^s]\beta_q^h\beta_q^l
\end{aligned}
$$

where the arguments of the coefficients of d_q^2 are of the form $(b_{0q} + \theta[b_q - b_{0q}], z_q)$. Consequently the coefficients of d_q^2 have limits

$$X_{hl}^{is}\beta_0^h\beta_0^l \text{ and } \ddot{x}_0^i T_h^s T_l^s \beta_0^h\beta_0^l + \dot{x}_0^i T_{hl}^s\beta_0^h\beta_0^l$$

evaluated at x_0.

We continue the proof by using Taylor's theorem to obtain for large q, y in $\mathfrak{D}^*$,

$$
\begin{aligned}
\Delta g_q^y &= [g_\sigma(b_q, z_q) - g_\sigma(b_{0q}, z_q)]\lambda_{0qy\sigma} \\
&= 2^{1/2} d_q \lambda_{0qy\sigma}(b_{0q}, z_q) + d_q^2 \lambda_{0qy\sigma} g_{\sigma hl}\beta_q^h\beta_q^l
\end{aligned}
$$

where $g_{\sigma hl}$ has an argument of the form $(b_{0q} + \theta[b_q - b_{0q}], z_q)$. This means that setting

$$T_{1q}^y = 2^{1/2} d_q \lambda_{0qy\sigma} g_{\sigma h}(b_{0q}, z_q)\beta_q^h,$$

we have

$$(8.2) \qquad \lim \,]d_q^{-2}\Delta g_q^y - T_{1q}^y] = \lambda_{0y\sigma} g_{hl}(b_0, z_0)\beta_0^h\beta_0^l \ .$$

Given y in $\mathfrak{D}^*$, we define

$$F_{sq}^y(b) = \int_{T_{0q}^s}^{T^s(b, z_q)} f(b_{0q}, t, x_{0q}, \dot{x}_{0q}, z_q, y, \lambda_{0qy})dt$$

and expand by Taylor's theorem to obtain for large q,

$$
\begin{aligned}
F_{sq}^y(b_q) \,|_{s=1}^{s=2} &= 2^{1/2} d_q f T_h^s \beta_q^h + d_q^2[f T_{hl}^s + f_t T_h^s T_l^s \\
&\quad + f_{x^i} T_h^s X_l^{is} + \ddot{x}_{0q}^i f_{\dot{x}^i} T_h^s T_l^s]_{s=1}^{s=2}\beta_q^h\beta_q^l
\end{aligned}
$$

where the arguments of the coefficients of d_q^2 are at x_{0qy} with the end conditions determined by $(b_{0q} + \theta[b_q - b_{0q}], z_q)$. Thus setting

$$T_{2q}^y = 2^{1/2} d_q^{-1} f T_h^s \beta_q^h$$

evaluated at x_{0qy}, we obtain

$$
\begin{aligned}
(8.3) \quad &\lim \, [d_q^{-2} F_{sq}^y(b_q) \,|_{s=1}^{s=2} - T_{2q}^y] \\
&= [f T_{hl}^s + f_t T_h^s T_l^s + f_{x^i} T_h^s X_l^{is} + \ddot{x}_0^i f_{\dot{x}^i} T_h^s T_l^s]_{s=1}^{s=2}\beta_0^h\beta_0^l
\end{aligned}
$$

evaluated at the end points of x_{0y}.

Given y in $\mathfrak{D}^*$, we define

$$a_q^y = f(b_q, t, x_q, \dot{x}_{0q}, z_q, y, \lambda_{0qy})$$

$$d_q^y = f(b_{0q}, t, x_{0q}, \dot{x}_{0q}, z_q, y, \lambda_{0qy})$$

$$e_{qy}^i = f_{\dot{x}^i}(b_q, t, x_q, \dot{x}_{0q}, z_q, y, \lambda_{0qy}) \ .$$

For large q we use Taylor's theorem to obtain

$$d_q^{-2}\Delta f_q^y = d_q^{-2}\int_{T_q^1}^{T_q^2} [a_q^y - d_q^y + (\dot{x}_q^i - \dot{x}_{0q}^i)e_{qy}^i]dt$$

(8.4)
$$= \int_{T_q^1}^{T_q^2} [f_{hl}\beta_q^h\beta_q^l + f_{x^ix^j}\eta_q^i\eta_q^j + 2f_{hx^i}\beta_q^h\eta_q^i$$

$$+ 2f_{x^i\dot{x}^j}\eta_q^i\dot{\eta}_q^i + 2f_{h\dot{x}^i}\beta_q^h\dot{\eta}_q^i]dt$$

$$+ 2^{1/2}d_q^{-1}\int_{T_q^1}^{T_q^2} [f_h\beta_q^h + f_{x^i}\eta_q^i + f_{\dot{x}^i}\dot{\eta}_q^i]dt$$

where the arguments of the first order partial derivatives are along x_{0qy} while those of the second order partial derivatives are of the form $(b_{0q} + \theta[b_q - b_{0q}], t, x_{0q} + \theta[x_q - x_{0q}], \dot{x}_{0q}, z_q, y, \lambda_{0qy})$.
First define along x_{0qy},

$$I_{1q}^y = 2^{1/2}d_q^{-1}\int_{T_{0q}^1}^{T_{0q}^2} f_h dt \beta_q^h + 2^{1/2}d_q^{-1}\int_{T_{0q}^s}^{T_q^s} f_h dt \mid_{s=1}^{s=2} \beta_q^h$$

$$I_{2q}^y = \left[\frac{\Delta X_q^{is}}{d_q^2} - \frac{\Delta x_q^{is}}{d_q^2}\right] f_{\dot{x}^i} \mid_{T_{0q}^1}^{T_{0q}^2}$$

$$I_{3q}^y = \left[\frac{\Delta X_q^{is}}{d_q} - \frac{\Delta x_q^{is}}{d_q}\right]\left[d_q^{-1}f_{\dot{x}^i} \mid_{T_q^s}^{T_q^s}\right] \mid_{s=1}^{s=2}$$

$$T_{3q}^y = 2^{1/2}d_q^{-1}f_{\dot{x}^i}[X_h^{is} - \dot{x}_{0q}^i T_h^s]_{s=1}^{s=2}\beta_q^h + \int_{T_{0q}^1}^{T_{0q}^2} f_h dt \beta_q^h \ .$$

By adding and subtracting the same terms and using the Euler equations, one sees that the last integral in (8.4) is equal to

$$I_{1q}^y + I_{2q}^y + I_{3q}^y \ .$$

Again using the Euler equations to evalute I_{3q}^y in the limit, we obtain from Lemmas 6.1 and 6.2 along x_{0y},

$$\lim [d_q^{-2}\Delta f_q^y - T_{3q}^y]$$

(8.5)
$$= \int_{T_0^1}^{T_0^2} [f_{hl}\beta_0^h\beta_0^l + f_{x^ix^j}\eta_0^i\eta_0^j + 2f_{x^i\dot{x}^j}\eta_0^i\dot{\eta}_0^j + 2f_{hx^i}\beta_0^h\eta_0^i + 2f_{h\dot{x}^i}\beta_0^h\dot{\eta}_0^i]dt$$

$$+ [f_{\dot{x}^i}X_{hl}^{is} - \ddot{x}_0^i f_{\dot{x}^i}T_h^s T_l^s - \dot{x}_0^i f_{\dot{x}^i}T_{hl}^s + f_{x^i}X_h^{is}T_l^s$$

$$- \dot{x}_0^i f_{x^i}T_h^s T_l^s + 2f_h T_l^s]_{s=1}^{s=2}\beta_0^h\beta_0^l \ .$$

For each y in $\mathfrak{D}^*$ one now sees that

$$J_y^*(x_q, x_{0q}) - J_y^*(x_{0q}, x_{0q}) = \varDelta g_q^y + F_{sq}^y(b_q)\,|_{s=1}^{s=2} + \varDelta f_q^y$$

and that by the transversality condition,

$$T_{1q}^y + T_{2q}^y + T_{3q}^y = 0 \ .$$

Hence, adding (8.2), (8.3), and (8.5), we prove the lemma.

LEMMA 7.3. *For each y in $\mathfrak{D}^*$ one has*

$$\lim \frac{I_\sigma(x_{qy}) - I_\sigma(x_{0qy})}{d_q} = 2^{1/2} I_{y\sigma}'(\gamma_0)$$

and

$$\lim [I_\sigma(x_{qy}) - I_\sigma(x_{0qy})] = 0 \ .$$

It suffices to prove the first limit since the second follows from it by the usual product of limits theorem. Given y in $\mathfrak{D}^*$, we set

$$T_{1q}^y = g_\sigma(b_q, z_q, y) - g_\sigma(b_{0q}, z_q, y) + \int_{T_{0q}^s}^{T_q^s} f_\sigma(b_{0q}, t, x_{0q}, \dot{x}_{0q}, z_q, y)dt\,|_{s=1}^{s=2}$$

$$T_{2q}^y = \int_{T_q^1}^{T_q^2} [f_\sigma(b_q, t, x_q, \dot{x}_{0q}, z_q, y) - f_\sigma(b_{0q}, t, x_{0q}, \dot{x}_{0q}, z_q, y)$$

$$+ (\dot{x}_q^i - \dot{x}_{0q}^i) f_{\sigma \dot{x}^i}(b_q, t, x_q, \dot{x}_{0q}, z_q, y)]dt$$

$$\phi_q^y = (b_{0q} + \theta[b_q - b_{0q}], t, x_{0q} + \theta[x_q - x_{0q}], \dot{x}_{0q}, z_q, y)$$

$$\phi_{0q}^y = (b_{0q}, t, x_q, \dot{x}_{0q}, z_q, y)$$

$$\phi_q^* = (b_{0q} + \theta[b_q - b_{0q}], z_q)$$

and define $E_{\sigma y}^*(x_q, x_{0q})$ like $E_y^*(x_q, x_{0q})$ with f replaced by f_σ. For large q one has by Taylor's theorem,

$$T_{1q}^y = 2^{1/2} d_q[g_{\sigma h}(\phi_q^*, y)\beta_q^h + f_\sigma(b_{0q}, t, x_{0q}, \dot{x}_{0q}, z_q, y) T_k^s(\phi_q^*)\beta_q^h\,|_{s=1}^{s=2}]$$

and

$$T_{2q}^y = 2^{1/2} d_q \int_{T_q^1}^{T_q^2} [f_{\sigma h}(\phi_q^y)\beta_q^h + f_{\sigma x^i}(\phi_q^y)\eta_q^i + f_{\sigma \dot{x}^i}(\phi_{0q}^y)\dot{\eta}_q^i]dt \ .$$

From the identity

$$I_\sigma(x_{qy}) - I_\sigma(x_{0qy}) = T_{1q}^y + T_{2q}^y + E_{\sigma y}^*(x_q, x_{0q})$$

and Lemmas 6.1 and 6.2, we now obtain

(8.6)
$$\lim \begin{Bmatrix} \text{dsu} \\ \inf \end{Bmatrix} d_q^{-1} [I_\sigma(x_{qy}) - I_\sigma(x_{0qy})]$$

$$= 2^{1/2} I_{y\sigma}'(\gamma_0) + \lim \begin{Bmatrix} \sup \\ \inf \end{Bmatrix} d_q^{-1} E_{\sigma y}^*(x_q, x_{0q}) \ .$$

By our selection of the arcs x_q, x_{0q} we have

$$(8.7) \quad \begin{aligned} 0 &\geqq G(x_{qy}) - G(x_{0qy}) \geqq [I_\sigma(x_{qy}) - I_\sigma(x_{0qy})][G_\sigma(x_{0qy}) - \varepsilon^y_{\sigma q}] \\ &= J^*_y(x_q, x_{0q}) - J^*_y(x_{0q}, x_{0q}) + E^*_y(x_q, x_{0q}) + [I_\sigma(x_{qy}) - I_\sigma(x_{0qy})]\varepsilon^y_{\sigma q} \end{aligned}$$

where $\lim \varepsilon^y_{\sigma q} = 0$. Combining (8.6), (8.7), and Lemma 7.2, we obtain

$$(8.8) \quad 0 \geqq o(1) + d_q^{-1} E^*_y(x_q, x_{0q}) + [o(1) + d_q^{-1} E^*_\sigma(x_q, x_{0q})]\varepsilon^y_{\sigma q} .$$

Setting $\phi^y_q = (b_q, t, x_q, \mathring{x}_{0q}, \mathring{x}_q, z_q, y, \lambda_{0qy1} + [I_1(x_{qy}) - I_1(x_{0qy})]\varepsilon^y_{1q}, \cdots)$, one has by (8.8) and the Weierstrass condition $\mathfrak{N}$,

$$(8.9) \quad \begin{aligned} 0 &\geqq \tau \lim \left\{ {\sup \atop \inf} \right\} d_q^{-1} \int_{T^1_q}^{T^2_q} E_f(\phi^y_q) dt \\ &\geqq \tau \lim \left\{ {\sup \atop \inf} \right\} d_q^{-1} \mid E^*_{\sigma y}(x_q, x_{0q}) \mid \; \geqq 0 , \end{aligned}$$

where τ is a positive constant.   The proof is now completed with (8.6) and (8.9).

REFERENCES

1. C. P. Brady, *The Minimum of Functions of Integrals in the Calculus of Variations*, Dissertation, University of Chicago, 1938.

2. M. R. Hestenes, *Applications of the theory of quadratic forms in Hilbert space to the calculus of variations.* Pacific J. Math., **1** (1951), 525-581.

3. ———, *Calculus of Variations and Optimal Control Theory*, John Wiley and Sons, Inc., New York, 1966.

4. ———, *An indirect sufficiency proof for the problem of Bolza in nonparametric form*, Trans. Amer. Math. Soc., **62** (1947), 509-535.

5. ———, *Multiplier and gradient methods*, J. Optimization Theory Appl. **4** (1969), 303-320.

6. ———, *Sufficient conditions for multiple integral problems in the calculus of variations*, Trans. Amer. Math. Soc., **70** (1948), 239-276.

7. ———, *Sufficient conditions for the isoperimetric problem of Bolza in the calculus of variations*, Trans. Amer. Math. Soc., **60** (1946), 93-118.

8. ———, *Theorem of Lindeberg in the calculus of variations*, Trans. Amer. Math. Soc., **60** (1946), 72-92.

9. ———, *The Weierstrass E-function in the calculus of variations*, Trans. Amer. Math. Soc., **60** (1946), 51-71.

10. E. J. McShane, *Sufficient conditions for a weak relative minimum in the problem of Bolza*, Trans. Amer. Math. Soc., **52** (1942), 344-379.

11. E. H. Mookini, *Necessary and Sufficient Conditions for an Optimal Control Problem in the Calculus of Variations*, Dissertation, University of California, Los Angeles, 1964.

12. W. I. Nathanson, *A Control Theory Bolza Problem with Intermediate Constraints*, Dissertation, University of California, Los Angeles, 1970.

13. L. L. Pennisi, *An indirect sufficiency proof for the problem of Lagrange with differential inequalities as added side conditions*, Trans. Amer. Math. Soc., **74** (1953), 177-198.

14. R. D. Rupp, *A Unified Sufficiency Theory in the Calculus of Variations*, Dissertation, University of California, Los Angeles, 1970.

UNIVERSITY OF CALIFORNIA, LOS ANGELES
AND
STATE UNIVERSITY OF NEW YORK AT ALBANY