

A NONLINEAR OPTIMAL CONTROL MINIMIZATION TECHNIQUE

BY

RUSSELL D. RUPP⁽¹⁾

ABSTRACT. Hestenes' method of multipliers is applied to a nonlinear optimal control problem. This requires that a differentially constrained problem be embedded in a family of unconstrained problems so as to preserve standard sufficiency criteria. Given an initial estimate of the Lagrange multipliers, a convergent sequence of arcs is generated. They are minimizing with respect to members of the above family, and their limit is the solution to the differentially constrained problem.

1. **Introduction.** The problem here considered is to minimize

$$I(x) = g(b) + \int_1^{T^2} L(b, t, x(t), u(t)) dt$$

subject to the differential equation

$$\dot{x}(t) = f(b, t, x(t), u(t)).$$

Our technique is the method of multipliers suggested by Hestenes [3, pp. 305 et seq.]. This method requires neither the solution of the dynamics nor the presence of a singularity in the penalty terms. Such a singularity may be insignificant if the Lagrange multipliers are near zero. However in general, this is not the case.

Hestenes' method of multipliers is to embed the differentially constrained problem in a family of unconstrained problems so as to preserve standard sufficiency criteria. Given an initial estimate of the Lagrange multipliers, a convergent sequence of arcs is generated. They are minimizing with respect to members of the above family, and their limit is the solution to the differentially constrained problem.

Hestenes [4, p. 309], and [3, p. 514], has modified constrained problems to unconstrained ones. In [4, pp. 308–310], he suggests algorithms which may approximate the minimizing arc. The author [7] has obtained a sufficiency theorem

Received by the editors February 8, 1971 and, in revised form, April 27, 1972.

AMS (MOS) subject classifications (1970). Primary 49D30, 49B10; Secondary 49D20, 65D99.

Key words and phrases. Calculus of variations, optimal control theory, method of multipliers, algorithm.

(1) Partially supported by National Science Foundation Grant GU 3171 and a State University of New York Faculty Fellowship.

suitable to applications of Hestenes' method of multipliers. The family of unconstrained problems is obtained via results in Hestenes [1] and the author [8], [9]. In [5] and [6], the author has applied this technique to a classical isoperimetric problem and a quadratic optimal control problem.

Let \mathfrak{R} be an open set in the $(2n + m + r + 1)$ -dimensional space of points (b, t, x, \dot{x}, u) . The (real-number) components of the vectors b, x, \dot{x}, u are denoted by b^b, x^i, \dot{x}^i, u^k ; $b = 1, 2, \dots, r$; $i = 1, 2, \dots, n$; $k = 1, 2, \dots, m$. Unless otherwise indicated, the indices b, i, k always have the ranges indicated above. Similarly, the index s below always has the range $s = 1, 2$. By b -space, we mean the space of points (b^1, b^2, \dots, b^r) . The similar statement holds with respect to the other combinations of variables. Let $X^{is}(b)$ be real-valued functions defined on the projection of \mathfrak{R} into b -space.

Let constants $T^1 < T^2$ be given. By an arc x , we mean a collection of constants and real-valued functions

$$(1.1) \quad x: b^b, x^i(t), u^k(t)$$

such that: The functions $x^i(t)$ and $u^k(t)$ have domain $T^1 \leq t \leq T^2$. The functions $x^i(t)$ are differentiable, and the derivatives $\dot{x}^i(t)$ are piecewise continuous. The functions $u^k(t)$ are piecewise continuous. Let $b, x(t)$, and $u(t)$ be the vectors whose components are given by (1.1) above. The arc x in (1.1) is usually written $x: b, x(t), u(t)$. An arc $x: b, x(t), u(t)$ is called admissible if $(b, t, x(t), \dot{x}(t), u(t))$ is in \mathfrak{R} for $T^1 \leq t \leq T^2$. An admissible arc $x: b, x(t), u(t)$ is called terminally admissible and is said to belong to the class \mathfrak{U} if $x^i(T^s) = X^{is}(b)$. We say that an arc $x: b, x(t), u(t)$ is smooth or has no corners if $\dot{x}(t)$ and $u(t)$ are continuous.

By a $bt\dot{x}u$ -neighborhood of an arc $x: b, x(t), u(t)$, we mean a neighborhood \mathfrak{N} in $bt\dot{x}u$ -space such that $b, t, x(t), \dot{x}(t), u(t)$ is in \mathfrak{N} for $T^1 \leq t \leq T^2$. The similar definition is made with respect to the other combinations of variables such as $bt\dot{x}$. Unless otherwise indicated, repeated indices denote summation with respect to that index. For example, given an l^* -dimensional vector v with components v^l , the norm of v is $|v| = (v^l v^l)^{1/2}$. If v were a real number, this norm would be its absolute value.

Let the functions $L(b, t, x, u)$ and $f^i(b, t, x, u)$ be continuous on the projection of \mathfrak{R} into $bt\dot{x}u$ -space, and let the function $g(b)$ be continuous on the projection of \mathfrak{R} into b -space. Let $f(b, t, x, u)$ be the vector-valued function whose components are given by $f^i(b, t, x, u)$. An admissible arc $x: b, x(t), u(t)$ is said to be differentially admissible if

$$(1.2) \quad \dot{x}(t) = f(b, t, x(t), u(t)).$$

The functional $I(x)$ is defined by

$$I(x) = g(b) + \int_{T^1}^{T^2} L(b, t, x(t), u(t)) dt$$

where $x: b, x(t), u(t)$ is an admissible arc. The problem here considered is to minimize $I(x)$ on the class of differentially admissible arcs in \mathfrak{U} .

The assumption that an arc $x: b, x(t), u(t)$ be such that $\dot{x}^i(t)$ and $u^k(t)$ are piecewise continuous is not necessary. Essentially the same results could have been obtained in the same manner if $\dot{x}^i(t)$ and $u^k(t)$ were only square integrable. On the other hand, the assumption that T^1 and T^2 are constants seems to be technically necessary. Suppose, however, that $T^1 = T^1(b)$ and $T^2 = T^2(b)$ were functions of the parameter b . In this case, the change of variables defined by $t \rightarrow 0 \leq s \leq 1$ where

$$s = (t - T^1(b))/(T^2(b) - T^1(b)) \quad \text{and} \quad T^1(b) < T^2(b)$$

puts the problem into the form we are considering.

2. Sufficiency criteria. In this section, we give sufficient conditions for the smooth, terminally and differentially admissible arc $x_0: b_0, x_0(t), u_0(t)$ to minimize $I(x)$ over the set of all terminally and differentially admissible arcs which lie in a btu -neighborhood \mathfrak{N} of x_0 . Necessary conditions are given in Hestenes [2, pp. 346–351]. The other purpose of this section is to show the relation between the class of arcs considered here and by the author in [7].

We begin with some continuity and differentiability assumptions. In a b -neighborhood of the arc x_0 , we assume that $g(b)$ and $X^{is}(b)$ have two continuous b -derivatives. In a btu -neighborhood of the arc x_0 , we assume that $L(b, t, x, u)$ and $f^i(b, t, x, u)$ have two continuous btu -derivatives. Furthermore, we agree to denote differentiation with respect to b^b by the subscript b . For example, we write $\partial/\partial b^b X^{is}(b) = X_b^{is}(b)$.

Let $p_{0i}(t)$, $T^1 \leq t \leq T^2$, be continuous Lagrange multipliers (costate variables) associated with the arc x_0 . We usually write such multipliers as a vector function $p_0(t)$ whose components are given by $p_{0i}(t)$. Let

$$H(b, t, x, u, p) = p_i f^i(b, t, x, u) - L(b, t, x, u).$$

Sufficient conditions that the arc x_0 with multipliers $p_0(t)$ be minimizing are that (2.1) through (2.4) below hold. If all of the arguments in (2.1) through (2.4) are not written in, we assume that they are taken with respect to $b_0, x_0(t), \dot{x}_0(t), u_0(t), p_0(t)$.

$$(2.1) \text{ Transversality. } g_b + X_b^{is} p_i \Big|_{T^1}^{T^2} + \int_{T^1}^{T^2} H_b dt = 0.$$

$$(2.2) \text{ Euler equations (integral form). } p_i(t) = \int_{t^*}^t H_{x_i} dt + p_i(t^*), \quad T^1 \leq t^* \leq T^2, \text{ and } H_{u^k} = 0.$$

(2.3) *Weierstrass condition.* Let $\mathfrak{Q}(\dot{x}, u) = [1 + |\dot{x}|^2 + |u|^2]^{1/2}$, and let \mathfrak{D} be the set of all points (b, t, x, \dot{x}, u, p) such that $\dot{x} = f(b, t, x, u)$ and (b, t, x, \dot{x}, u) is in \mathfrak{R} . There is a *bt \dot{x} up*-neighborhood \mathfrak{N} of x_0 and a positive constant r such that: For $(b, t, x, \dot{x}_0(t), u_0(t), p_0(t))$ in $\mathfrak{N} \cap \mathfrak{D}$ and $(b, t, x, \tilde{x}, \tilde{u}, p_0(t))$ in \mathfrak{D} , one has

$$\begin{aligned} & -H(b, t, x, \tilde{u}, p_0(t)) + H(b, t, x, u_0(t), p_0(t)) \\ & \quad + (\tilde{u}^k - u_0^k(t)) H_{u^k}(b, t, x, u_0(t), p_0(t)) \\ & \geq r [\mathfrak{Q}(\tilde{x} - \dot{x}_0(t), \tilde{u} - u_0(t))]. \end{aligned}$$

(2.4) *Positiveness of the second variation.* An admissible variation η is a collection of constants and functions $\eta: \beta^b, \eta^i(t), \mu^k(t)$ such that: The β^b are constants. The $\eta^i(t)$ are absolutely continuous functions on $T^1 \leq t \leq T^2$ which have a square integrable derivative. The $\mu^k(t)$ are square integrable functions on $T^1 \leq t \leq T^2$. The admissible variation η is called terminally admissible with respect to the arc x_0 if the secondary end conditions hold: $\eta^i(T^2) = X_b^{is} \beta^b$. Let $\beta, \eta(t)$, and $\mu(t)$ be the vectors whose components are given by η above. The admissible variation η is usually written in the vector form $\eta: \beta, \eta(t), \mu(t)$. This admissible variation $\eta: \beta, \eta(t), \mu(t)$ is said to be differentially admissible if, almost everywhere in $T^1 \leq t \leq T^2$,

$$\dot{\eta}^i(t) = f_b^i \beta^b + f_{x^j}^i \eta^j + f_{u^k}^i \mu^k.$$

Let $\eta: \beta, \eta(t), \mu(t)$ be a nonzero differentially and terminally admissible variation with respect to x_0, p_0 . Let

$$Q(\beta) = g_{bl} \beta^b \beta^l + X_{bl}^{is} p_i \bigg|_{T^1}^{T^2} \beta^b \beta^l$$

and

$$-2\omega = H_{x^i x^j} \eta^i \eta^j + H_{u^k u^l} \mu^k \mu^l + H_{bl} \beta^b \beta^l + 2H_{x^i u^k} \eta^i \mu^k + 2H_{x^i b} \eta^i \beta^b + 2H_{u^k b} \mu^k \beta^b.$$

We assume that the second variation $I_2(\eta) = Q(\beta) + \int_{T^1}^{T^2} 2\omega dt$ is positive.

An arc $x: b, x(t), u(t)$ which satisfies the Euler equations with respect to multipliers $p(t)$ is called an extremal x with multipliers p . We sometimes refer to x, p as an extremal pair. An extremal which is also a terminally admissible arc is called a terminally admissible extremal. An accessory extremal is an extremal for the second variation with respect to a minimizing arc and its multipliers. Finally, we remark that in view of Theorem 3.3 (Rupp [7]), the assumption (2.4) of positive second variation could actually have been made with respect to a slightly

smaller class of terminally admissible variations.

We now demonstrate how the class of arcs here considered may be put in the form considered by the author in [7]. Let $x: b, x(t), u(t)$ be a terminally admissible arc. We introduce new state vectors $x^{n+k}(t)$ and end condition parameters b^{r+k} by defining

$$x^{n+k}(t) = \int_{T^1}^t u^k(s) ds \quad \text{and} \quad b^{r+k} = \int_{T^1}^{T^2} u^k(s) ds.$$

The new end conditions are

$$x^{n+k}(T^1) = 0 \quad \text{and} \quad x^{n+k}(T^2) = b^{r+k}.$$

Thus the terminally admissible arc $x: b, x(t), u(t)$ is equivalent to a terminally admissible arc without any control variables.

It now only remains to associate metric space parameters with this new class of arcs. The proper parameters for our purposes are the continuous Lagrange multipliers $p(t)$ on $T^1 \leq t \leq T^2$ and the real numbers ξ . The components of $p(t)$ are given by $p_i(t)$. We often write these parameters in the form (p, ξ) . They are referred to as parameters (p, ξ) or as multipliers p with parameter ξ . They form a normed linear space with norm given by $\sup_t |p(t)| + |\xi|$. The associated metric is $\sup_t |p(t) - p^*(t)| + |\xi - \xi^*|$.

Thus in the obvious manner, a terminally admissible arc x with parameters (p, ξ) is equivalent to an arc in the class considered in [7]. The terminally admissible variations considered here and those considered in [7] correspond in the analogous manner. Thus the theory developed in [7] is applicable to the present problem.

3. A sufficiency theorem. This section contains the sufficiency results which are necessary to an application of Hestenes' method of multipliers.

Let the smooth terminally admissible arc $x_0: b_0, x_0(t), u_0(t)$ and the continuous Lagrange multipliers $p_0(t)$ satisfy (2.1) through (2.4). As was demonstrated in §2, the arc x_0 with parameters $(p, \xi) = (p_0, 0)$ is equivalent to an arc in the class considered by the author in [7]. Define the function $F(b, t, x, \dot{x}, u, p)$ by

$$F(b, t, x, \dot{x}, u, p) = p_i \dot{x}^i - H(b, t, x, u, p).$$

This first theorem establishes the existence of a family of unconstrained problems to which the sufficiency theory in [7] applies.

Theorem 3.1. *There is an infinitely differentiable function $C(b, t, x, \dot{x}, u)$ with the following property: Let*

$$\mathfrak{F}(b, t, x, \dot{x}, u, p, \xi) = F(b, t, x, \dot{x}, u, p) + \frac{1}{2}[C(b, t, x, \dot{x}, u) + \xi]|\dot{x} - f(b, t, x, u)|^2.$$

The arc x_0 with parameters $(p_0, 0)$ and the functional

$$J(x, p, \xi) = g(b) + \int_{T^1}^{T^2} \mathfrak{G}(b, t, x, \dot{x}, u, p, \xi) dt$$

satisfy the hypothesis of Theorem 3.1 (Rupp [7]).

In order to prove this theorem, we observe that (2.4) implies the following: Whenever $\eta: \beta, \eta(t), \mu(t)$ is a nonnull terminally admissible variation with respect to x_0 and

$$Q(\eta) = \int_{T^1}^{T^2} |\dot{\eta} - f_b \beta^b - f_{x^i} \eta^i - f_u \mu^k|^2 dt = 0,$$

one has $I_2(\eta) > 0$. Furthermore, (2.3) implies by Lemma 2.3 (Rupp [9]) that along the arc x_0 with parameters $(p_0, 0)$ the following holds: Whenever for $(\Pi_1, \Pi_2) \neq 0$, $Q_1(\Pi_1, \Pi_2) = |\Pi_1 - f_{u^l} \Pi_2^l|^2 = 0$, one has

$$Q_2(\Pi_1, \Pi_2) = 2F_{\dot{x}^i u^k} \Pi_1^i \Pi_2^k + F_{u^k u^l} \Pi_2^k \Pi_2^l > 0.$$

This means that Lemma 1-5.1 (Rupp [8]) applies. Thus there is a positive constant C_1 such that, when evaluated at x_0 with parameters $(p_0, 0)$, the integral functional

$$(3.1) \quad \int_{T^1}^{T^2} [F(b, t, x, \dot{x}, u) + \frac{1}{2} C_1 |\dot{x} - f(b, t, x, u)|^2] dt$$

has positive second variation and such that the integrand in (3.1) has a positive definite matrix of second partials with respect to the $\dot{x}u$ -variables.

Our object is now to ensure a suitable Weierstrass condition. Let the compact set \mathfrak{R} and the closed set \mathfrak{D} be defined by

$$\mathfrak{R} = \{(b, t, x, \dot{x}, u, p, 0): b = b_0, T^1 \leq t \leq T^2,$$

$$x = x_0(t), \dot{x} = \dot{x}_0(t), u = u_0(t), \text{ and } p = p_0(t)\},$$

$$\mathfrak{D} = \{(b, t, x, \dot{x}, u, p, \xi): \dot{x} = f(b, t, x, u)\}.$$

By Lemmas 2.2 and 2.4 (Rupp [9]) there is a positive constant τ_1 such that for $(b, t, x, \dot{x}, u, p, 0)$ in \mathfrak{R} , and $(b, t, x, \dot{x}, \tilde{u}, p, 0)$ in \mathfrak{D} ,

$$[\mathfrak{L}(\tilde{x} - \dot{x}, \tilde{u} - u) - 1] \geq \tau_1 E_{\dot{x}^i - f^i}(b, t, x, \dot{x}, u, \tilde{x}, \tilde{u})$$

where

$$E_{\dot{x}^i - f^i}(b, t, x, \dot{x}, u, \tilde{x}, \tilde{u}) = -f^i(b, t, x, \tilde{u}) + f^i(b, t, x, u) - (\tilde{u}^k - u^k) f_{u^k}^i(b, t, x, u).$$

Thus by the Weierstrass condition (2.3), there are positive constants δ and τ_2 such that, upon defining the compact set \mathfrak{R}^* by

$$\mathfrak{R}^* = \{(b, t, x, \dot{x}, u, p, \xi): |b - b_0| + |x - x_0(t)| + |p - p_0(t)| + |\xi| \leq \delta, \\ \dot{x} = \dot{x}_0(t), u = u_0(t), T^1 \leq t \leq T^2, \text{ and } \dot{x} = f(b, t, x, u)\},$$

one has for $(b, t, x, \dot{x}, u, p, \xi)$ in \mathfrak{R}^* and $(b, t, x, \tilde{x}, \tilde{u}, p, \xi)$ in \mathfrak{D} ,

$$E_F(b, t, x, \dot{x}, u, \tilde{x}, \tilde{u}, p) \geq \tau_2 [\mathfrak{L}(\tilde{x} - \dot{x}, \tilde{u} - u) - 1]$$

where

$$E_F(b, t, x, \dot{x}, u, \tilde{x}, \tilde{u}, p) = F(b, t, x, \tilde{x}, u, p) - F(b, t, x, \dot{x}, u, p) \\ - (\tilde{x}^i - \dot{x}^i) F_{x^i}(b, t, x, \dot{x}, u, p) \\ - (\tilde{u}^k - u^k) F_{u^k}(b, t, x, \dot{x}, u, p).$$

Hence Theorem 3.3 (Rupp [9]) implies that there is a nonnegative infinitely differentiable function $C_2(b, t, x, \dot{x}, u)$, a positive constant τ_3 , and a neighborhood \mathfrak{N}_1 of \mathfrak{R} such that the following holds: Let

$$\mathfrak{H}(b, t, x, \dot{x}, u, p, \xi) = F(b, t, x, \dot{x}, u, p) \\ + \frac{1}{2} [C_1 + C_2(b, t, x, \dot{x}, u) + \xi] |\dot{x} - f(b, t, x, u)|^2.$$

One has for $(b, t, x, \dot{x}, u, p, \xi)$ in $\mathfrak{N}_1 \cap \mathfrak{R}^*$ and $(b, t, x, \tilde{x}, \tilde{u})$ in \mathfrak{R} ,

$$E_{\mathfrak{H}}(b, t, x, \dot{x}, u, \tilde{x}, \tilde{u}, p, \xi) \geq \tau_3 [\mathfrak{L}(\tilde{x} - \dot{x}, \tilde{u} - u) - 1]$$

where

$$E_{\mathfrak{H}}(b, t, x, \dot{x}, u, \tilde{x}, \tilde{u}, p, \xi) = \mathfrak{H}(b, t, x, \tilde{x}, \tilde{u}, p, \xi) - \mathfrak{H}(b, t, x, \dot{x}, u, p, \xi) \\ - (\tilde{x}^i - \dot{x}^i) \mathfrak{H}_{x^i}(b, t, x, \dot{x}, u, p, \xi) \\ - (\tilde{u}^k - u^k) \mathfrak{H}_{u^k}(b, t, x, \dot{x}, u, p, \xi).$$

Now note that $\mathfrak{N}_1 \cap \mathfrak{R}^*$ plays the role of \mathfrak{D}^* with respect to the compact set \mathfrak{R} in Theorem 3.1 (Rupp [9]). Thus there is a neighborhood \mathfrak{N}_2 of \mathfrak{R} and a positive constant τ_4 such that for $(b, t, x, \dot{x}, u, p, \xi)$ in \mathfrak{N}_2 and (b, t, x, \dot{x}, u) in \mathfrak{R} ,

$$E_{\mathfrak{H}}(b, t, x, \dot{x}, u, \tilde{x}, \tilde{u}, p, \xi) \geq \tau_4 [\mathfrak{L}(\tilde{x} - \dot{x}, \tilde{u} - u)].$$

It is now easy to verify that $C(b, t, x, \dot{x}, u) = C_1 + C_2(b, t, x, \dot{x}, u)$ is the infinitely differentiable function which we sought to construct. This proves the theorem.

Under appropriate circumstances the function $C(b, t, x, \dot{x}, u)$ may depend only on some of the indicated variables or may be a constant. See Rupp [9, §3]. Let $x: b, x(t), u(t)$ and $x^*: b^*, x^*(t), u^*(t)$ be terminally admissible arcs. We define

$$\mathcal{D}(x, x^*) := |b - b^*|^2 + \int_{T_1}^{T_2} [\mathcal{Q}(\dot{x} - \dot{x}^*, u - u^*) - 1]^{1/2} dt.$$

As a corollary of Theorem 3.1 and Theorem 3.1 (Rupp [7]), we have the sufficiency theorem below.

Theorem 3.2. *Let the smooth, differentially and terminally admissible arc $x_0: b_0, x_0(t), u_0(t)$ and the parameters $(p(t), \xi) = (p_0(t), 0)$ satisfy (2.1) through (2.4). There are positive constants $\tilde{\tau}, \tau_1, \tau_2$, a neighborhood \mathfrak{P} of p_0 , and a btx-neighborhood \mathfrak{N} of x_0 with the following property: Let parameters p in \mathfrak{P} and $-\tau_1 \leq \xi \leq \tau_2$ be given. There is a corresponding smooth terminally admissible minimizing arc $x(p, \xi) = x^*$ in \mathfrak{N} . That is, given a terminally admissible arc $x: b, x(t), u(t)$ in \mathfrak{N} , one has*

$$(3.2) \quad J(x, p, \xi) \geq J(x^*, p, \xi) + \tilde{\tau} \mathcal{D}(x, x^*).$$

Thus for each p in \mathfrak{P} and $-\tau_1 \leq \xi \leq \tau_2$, we have a simple (in other words unconstrained) minimization problem. When $(p, \xi) = (p_0, 0)$ this simple problem is equivalent to the original one since on the class of differentially admissible arcs,

$$(3.3) \quad J(x, p, \xi) = I(x).$$

This proves Corollary 3.1.

Corollary 3.1. *Let the smooth, differentially and terminally admissible arc $x_0: b_0, x_0(t), u_0(t)$ and its continuous multipliers $p_0(t)$ satisfy (2.1) through (2.4). There are a btx-neighborhood \mathfrak{N} of x_0 and a positive constant $\tilde{\tau}$ such that: Given a differentially and terminally admissible arc $x: b, x(t), u(t)$ in \mathfrak{N} , one has $I(x) \geq I(x_0) + \tilde{\tau} \mathcal{D}(x, x_0)$.*

This next corollary is a maxmin principle.

Corollary 3.2. *Let the situation in Theorem 3.2 hold. One has that*

$$I(x_0) = \max_{\mathfrak{P}^*} \min_{\mathfrak{A}^*} J(x, p, \xi)$$

where \mathfrak{P}^* is the Cartesian product of \mathfrak{P} with $-\tau_1 \leq \xi \leq \tau_2$ and \mathfrak{A}^* is the class of terminally admissible arcs in \mathfrak{N} .

In order to prove this corollary, we observe that by Theorem 3.2, one has for (p, ξ) in \mathfrak{P}^* and $x^* = x(p, \xi)$,

$$(3.4) \quad J(x_0, p_0, 0) \geq J(x^*, p, \xi) + \tilde{\tau} \mathcal{D}(x^*, x_0).$$

Thus the corollary follows from (3.4) and (3.3).

The family of minimizing arcs $x(p, \xi)$ which is described in Theorem 3.2 has some important continuity and differentiability properties. We state them

below as Theorem 3.3.

Theorem 3.3. *Let $x(p, \xi)$: $b(p, \xi)$, $x(t, p, \xi)$, $u(t, p, \xi)$ be the family of minimizing arcs described in Theorem 3.2. There are positive constants τ_1, τ_2 and a neighborhood \mathfrak{P} of p_0 such that: For p in \mathfrak{P} and $-\tau_1 \leq \xi \leq \tau_2$, the functions*

$$b(p, \xi), \quad x(t, p, \xi), \quad \dot{x}(t, p, \xi), \quad u(t, p, \xi)$$

are continuous and have two continuous ξ -derivatives.

In order to prove this theorem, we choose positive constants τ_1, τ_2 , and a neighborhood \mathfrak{P} of p_0 such that Theorem 3.2, Corollary 5.1 (Rupp [7]), and Corollary 5.3 (Rupp [7]) all hold for p in \mathfrak{P} and $-\tau_1 \leq \xi \leq \tau_2$. Let p be in \mathfrak{P} and $-\tau_1 \leq \xi \leq \tau_2$. By Theorem 3.2, there is an associated terminally admissible minimizing arc $x(p, \xi)$. This arc $x(p, \xi)$ satisfies the Euler equations and transversality condition with respect to the parameters (p, ξ) . Hence Corollary 5.1 (Rupp [7]) and Corollary 5.3 (Rupp [7]) imply that the family $x(p, \xi)$ has the required properties. This proves the theorem.

The next theorem shows that the arc x_0 is minimizing in a slightly stronger sense than that indicated in Theorem 3.2.

Theorem 3.4. *Let $x(p, \xi)$: $b(p, \xi)$, $x(t, p, \xi)$, $u(t, p, \xi)$ be the family of minimizing arcs described in Theorem 3.2. There are positive constants τ_1 and τ_2 such that, for $-\tau_1 \leq \xi \leq \tau_2$, $x(p_0, \xi) = x_0$.*

In order to prove this theorem, we choose positive constants τ_1, τ_2 and a neighborhood \mathfrak{P} of p_0 such that Theorem 3.2, and Corollary 5.1 (Rupp [7]) both hold for p in \mathfrak{P} and $-\tau_1 \leq \xi \leq \tau_2$. Let $-\tau_1 \leq \xi \leq \tau_2$ be given, and let $x^* = x(p_0, \xi)$. By Theorem 3.2, one has $J(x, p_0, \xi) \geq J(x^*, p_0, \xi)$ for all terminally admissible arcs x in some btx -neighborhood \mathfrak{N} of x_0 . Hence x^* and x_0 both satisfy the same Euler equations and transversality condition. Thus Corollary 5.1 (Rupp [7]) implies $x^* = x_0$. This proves the theorem.

4. The algorithm. The first part of this section establishes some notation and states an additional hypothesis. The latter part contains the algorithm, Theorem 4.1.

We begin with an application of Theorem 3.1. The functional which we consider is

$$J(x, p, \xi) = \int_{T_1}^{T_2} \mathfrak{H}(b, t, x, \dot{x}, u, p, \xi) dt.$$

When $\xi = 0$, we write $J(x, p, 0) = J(x, p)$. Theorem 3.2 implies the existence of a neighborhood \mathfrak{P}_1 of p_0 , \mathfrak{N}_1 of x_0 , and positive constants $\tilde{\tau}, \tau_1, \tau_2$ with the

following property: Given multipliers p in \mathfrak{P}_1 and $-\tau_1 \leq \xi \leq \tau_2$, there is a corresponding terminally admissible extremal $x(p, \xi) = x^*$ such that for a terminally admissible arc x in \mathfrak{N}_1 ,

$$J(x, p, \xi) \geq J(x^*, p, \xi) + \tilde{\tau} \mathcal{D}(x, x^*).$$

The integral form of the Euler equations associated with the state variables is

$$(4.1) \quad p_i + (C + \xi)(\dot{x}^i - f^i) = \int_{t^*}^t \tilde{\xi}_{x_i} dt + C_i(t^*).$$

The constant

$$(4.2) \quad C_i(t^*) = [p_i + (C + \xi)(\dot{x}^i - f^i)]_{t=t^*}$$

depends on the selection of the point $T^1 \leq t^* \leq T^2$.

We now agree that the constants $\tilde{\tau}$, τ_1 , τ_2 , and neighborhoods \mathfrak{P}_1 , \mathfrak{N}_1 are so small that all the theorems and corollaries in §3 apply. The additional assumption which we wish to make is that the function $C(b, t, x, \dot{x}, u)$ is a constant C . This assumption may be avoided in the following manner: We are dealing with extremals which may be taken to lie in an arbitrarily small $btx\dot{x}u$ -neighborhood \mathfrak{N} of x_0 . When \mathfrak{N} is sufficiently small, there are positive constants C_1 and C_2 such that $C_1 \geq C(b, t, x, \dot{x}, u) \geq C_2$ on \mathfrak{N} . Our present results can easily be modified in terms of such a pair of constants. Finally we remark that Theorem 3.4 (Rupp [9]) implies that by starting with a slightly smaller and bounded subregion \mathfrak{N}^* of \mathfrak{N} instead of \mathfrak{N} , the function $C(b, t, x, \dot{x}, u)$ could originally have been chosen to be a constant.

Now suppose that multipliers p_q in \mathfrak{P}_1 and real numbers $0 \leq \tau \leq \xi_q \leq \tau^* \leq \tau_2$ for $q = 1, 2, \dots, Q-1$ have been given. Theorem 3.2 implies the existence of terminally admissible minimizing extremals $x_q = x(p_q, 0)$ for $q = 1, 2, \dots, Q-1$. Suppose that, for $q = 1, 2, \dots, Q-2$, $p_{q+1} = p_q + \xi_q(\dot{x} - f)_{q+1}$ where $(\dot{x} - f)_{q+1}$ is the function $\dot{x} - f(b, t, x, u)$ with arguments given by the arc x_{q+1} . According to Theorem 3.2, there is an arc x_Q in \mathfrak{N} that minimizes the integral $J(x, p_{Q-1}, \xi_{Q-1})$ on \mathfrak{N} . We shall show in the next theorem that $p_Q = p_{Q-1} + \xi_{Q-1}(\dot{x} - f)_Q$ is in \mathfrak{P} . It follows (Corollary 5.1, Rupp [7]) that x_Q minimizes $J(x, p_Q, 0)$ on \mathfrak{N} and hence that $x_Q = x(p_Q, 0)$.

Theorem 4.1. *There are positive constants $\tau < \tau^*$, δ , K , a neighborhood \mathfrak{P} of p_0 , and a btx -neighborhood \mathfrak{N} of x_0 with the following properties: Let multipliers p_1 in \mathfrak{P} be chosen such that $\sup |p_1 - p_0| < \delta$. Given constants $\tau \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots$, one may define multipliers in \mathfrak{P} by*

$$(4.3) \quad p_{q+1} = p_q + \xi_q(\dot{x} - f)_{q+1}.$$

Furthermore, the minimizing arcs $x_q = x(p_q, 0)$ are in \mathfrak{N} and converge as indicated below,

$$(4.4) \quad \lim \left\{ |b_q - b_0|^2 + \int_{T_1}^{T^2} [|x_q - x_0|^2 + |\dot{x}_q - \dot{x}_0|^2 + |u_q - u_0|^2] dt \right\} = 0.$$

The proof of Theorem 4.1 is an application of Theorems 3.2, 3.3 and the two lemmas below, 4.1 and 4.2. The proof of Lemmas 4.1 and 4.2 is given in §7.

Lemma 4.1. *Let multipliers p_q for $q = 1, 2, \dots, Q-1$ be defined in \mathfrak{P} by (4.3), and let $\tilde{\tau}$ be the positive constant in (3.2). The inequality below is valid for $q = 1, 2, \dots, Q-1$,*

$$(4.5) \quad \begin{aligned} & \int_{T_1}^{T^2} |p_q - p_0|^2 dt \\ & \geq \int_{T_1}^{T^2} |p_{q+1} - p_0|^2 dt + 4\tilde{\tau} \mathcal{D}(x_{q+1}, x_0) + \tau^2 \int_{T_1}^{T^2} |(\dot{x} - f)_{q+1}|^2 dt. \end{aligned}$$

Lemma 4.2. *There is a neighborhood \mathfrak{P} of p_0 and positive constants τ^* and K such that the following holds: Suppose that the multipliers p_q are defined in \mathfrak{P} and that $0 \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots, Q-1$. The inequality below is valid for $q = 1, 2, \dots, Q$,*

$$(4.6) \quad \sup_t |p_q - p_0 + C(\dot{x} - f)_q| \leq K \left\{ [\mathcal{D}(x_q, x_0)]^{1/2} + \left[\int_{T_1}^{T^2} |p_q - p_0|^2 dt \right]^{1/2} \right\}.$$

Let the positive constants $\tau_1 < \tau_2$, $\text{blx-neighborhood } \mathfrak{N}$ of x_0 , and neighborhood \mathfrak{P}_1 of p_0 be chosen so that Theorem 3.2 applies. Let the positive constants τ_3 , K_1 and the neighborhood \mathfrak{P}_2 of p_0 be chosen such that Lemma 4.2 applies. That is, whenever multipliers p_q in \mathfrak{P}_2 and constants $0 \leq \xi_q \leq \tau_3$ for $q = 1, 2, \dots, Q-1$ are given, we have for $q = 2, 3, \dots, Q$,

$$(4.7) \quad \sup |(p_q - p_0) + C(\dot{x} - f)_q| \leq K_1 [\mathcal{D}(x_q, x_0)]^{1/2} + \left[\int_{T_1}^{T^2} |p_q - p_0|^2 dt \right]^{1/2}.$$

Choose positive constants τ and τ^* such that $0 < \tau \leq \tau^* \leq \min \{\tau_2, \tau_3\}$. We note that the positive constant $R = C(\tau + C)^{-1}$ is less than one. Let $\mathfrak{P} = \mathfrak{P}_1 \cap \mathfrak{P}_2$, and let $\tilde{\tau}$ be the positive constant in (3.2). Furthermore, let $K_2 = 2\tau^* K_1 (\tau^* + C)^{-1} \max \{1, (4\tau\tilde{\tau})^{-1}\}$. Now let the positive constant δ be chosen such that the neighborhood of p_0 of radius

$$(1 + K_2 |T^2 - T^1|^{1/2} / (1 - R)) \delta$$

is contained in \mathfrak{P} .

Let multipliers p_1 in \mathfrak{P} such that $\sup |p_1 - p_0| < \delta$ be given, and let constants $\tau \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots$ be given. Suppose that multipliers p_q for $q = 1, 2, \dots, Q-1$ have been defined in \mathfrak{P} for $Q-1 \geq 1$. Furthermore, define for $q = 1, 2, \dots, Q$

$$\mathcal{D}_q = [\mathcal{D}(x_q, x_0)]^{1/2} \quad \text{and} \quad P_q = \left[\int_{T^1}^{T^2} |p_q - p_0|^2 dt \right]^{1/2}.$$

Now observe that for $q = 2, 3, \dots, Q$, one has

$$(4.8) \quad C(\dot{x} - f)_q = C\xi_{q-1}^{-1}(p_q - p_{q-1}).$$

Substituting (4.8) into (4.7), we obtain for $q = 2, 3, \dots, Q$

$$(4.9) \quad \sup |p_q - p_0| \leq R \sup |p_{q-1} - p_0| + \tau^* K_1 (\tau^* + C)^{-1} (\mathcal{D}_q + P_q).$$

Thus Lemma 4.1 implies for $q = 2, 3, \dots, Q$

$$(4.10) \quad \sup |p_q - p_0| \leq R \sup |p_{q-1} - p_0| + K_2 P_1$$

where $K_2 = 2\tau^* K_1 (\tau^* + C)^{-1} \max \{1, (4\tau\tilde{\tau})^{-1}\}$. Hence (4.10) implies

$$(4.11) \quad \sup |p_Q - p_0| \leq R^{Q-1} \sup |p_1 - p_0| + K_2 \left(\sum_{j=0}^{Q-2} R^j \right) P_1.$$

However (4.11) implies

$$(4.12) \quad \sup |p_Q - p_0| \leq (1 + K_2 |T^2 - T^1|^{1/2} / (1 - R)) \sup |p_1 - p_0|.$$

Thus (4.12) implies that the multipliers p_Q are in \mathfrak{B} . This proves that the sequence p_q of multipliers is defined in \mathfrak{B} and that the sequence x_q of minimizing arcs remains in \mathfrak{N} .

In order to prove (4.4), we obtain by Lemma 4.1 for $q = 1, 2, \dots$

$$(4.13) \quad \begin{aligned} & \int_{T^1}^{T^2} |p_q - p_0|^2 dt - \int_{T^1}^{T^2} |p_{q+1} - p_0|^2 dt, \\ & \geq 4\tau\tilde{\tau} \mathcal{D}(x_{q+1}, x_0) + \tau^2 \int_{T^1}^{T^2} |(\dot{x} - f)_{q+1}|^2 dt. \end{aligned}$$

Summing (4.13) over $q = 1, 2, \dots$, we have

$$(4.14) \quad \begin{aligned} & \int_{T^1}^{T^2} |p_1 - p_0|^2 dt - \liminf \int_{T^1}^{T^2} |p_q - p_0|^2 dt \\ & \geq 4\tau\tilde{\tau} \sum_{q=2}^{\infty} \mathcal{D}(x_q, x_0) + \tau^2 \sum_{q=2}^{\infty} \int_{T^1}^{T^2} |(\dot{x} - f)_q|^2 dt. \end{aligned}$$

Thus (4.14) implies

$$(4.15) \quad \lim \mathcal{D}(x_q, x_0) = 0.$$

Hence (4.15), Theorem 3.3, and Lemma 2.1 (Rupp [9]) imply (4.4). This proves the theorem.

The last inequality (4.14) is interesting in its own right as it gives a rate of convergence estimate. In §6, we consider this topic in more detail.

5. **Examples.** Free problems from the calculus of variations may be used to illustrate the preceding theory: Suppose one wished to minimize

$$I(x) = \int_{T^1}^{T^2} f(t, x(t), \dot{x}(t)) dt$$

over a suitable class of arcs. When the differential equation

$$(5.1) \quad \dot{x}(t) = u(t)$$

is adjoined, this yields the problem of minimizing

$$I(x, u) = \int_{T^1}^{T^2} f(t, x(t), u(t)) dt$$

subject to (5.1).

It is instructive to consider the case when $f(t, x, \dot{x}) = \frac{1}{4} (\dot{x}^2 - 1)^2$ or $(\dot{x}^2 - x^2)$. In the first instance, the problem would be to minimize

$$I(x, u) = \frac{1}{4} \int_0^1 (u^2 - 1)^2 dt$$

over the class of arcs determined by

$$\dot{x} - u = 0, \quad x(0) = 0, \quad x(1) = 1.$$

One verifies that the arc and multipliers

$$x_0 = t, \quad u_0 = 1, \quad p_0 = 0,$$

satisfy the sufficiency hypotheses for a minimum in a $x\dot{x}u$ -neighborhood. The algorithm (4.3) can be explicitly verified when the initial multiplier estimate is an arbitrary constant. In the second case, the problem is to minimize $I(x, u) = \int_0^1 (u^2 - x^2) dt$ subject to $x(0) = x(1) = 0$, $\dot{x} - u = 0$. The sufficiency criteria applies to $x_0 = u_0 = p_0 = 0$, and explicit verification of the algorithm (4.3) is again possible when the initial multiplier estimate is an arbitrary constant.

Of course, the algorithm (4.3) is of most value when it is not a simple matter to solve the constraints (1.2) for u as a function of the other variables. Examples of that type are easily concocted: Let (1.2) be given by $\dot{x} = u^2 - x^2$. Over $0 < T^1 \leq t \leq T^2$, a suitable integrand would be $L(t, x, u) = x^2 + u^2(e^{2t} - 1)$. One verifies that the sufficiency criteria is satisfied by $x_0 = 1$, $u_0 = 1$, $p_0 = e^{2t} - 1$, where the end conditions are $x(T^1) = x(T^2) = 1$.

6. **Convergence estimates.** In this section, we estimate how fast and in what manner the algorithm given by (4.1) converges. Throughout this section, we assume that the situation described in Theorem 4.1 holds. This first theorem is an

immediate consequence of the inequality (4.7), which was used in the proof of Theorem 4.1.

Theorem 6.1. *We have the following convergence estimates on $\mathcal{D}(x_q, x_0)$ and $\int_{T-1}^{T^2} |(\dot{x} - f)_q|^2 dt$:*

$$(6.1) \quad \sum_{q=2}^{\infty} \mathcal{D}(x_q, x_0) \leq (4\tau\tilde{\tau})^{-1} \int_{T-1}^{T^2} |p_1 - p_0|^2 dt,$$

$$(6.2) \quad \sum_{q=2}^{\infty} \int_{T-1}^{T^2} |(\dot{x} - f)_q|^2 dt \leq \tau^{-2} \int_{T-1}^{T^2} |p_1 - p_0|^2 dt.$$

Theorem 6.2. *Suppose that $\xi_q = \tau$ for $q = 1, 2, \dots$. The inequality below holds for $q, k = 1, 2, \dots$.*

$$(6.3) \quad \begin{aligned} \int_{T-1}^{T^2} |p_{q-1} - p_{k-1}|^2 dt &\geq \int_{T-1}^{T^2} |p_q - p_k|^2 dt + 4\tau\tilde{\tau} \mathcal{D}(x_q, x_k) \\ &\quad + \tau^2 \int_{T-1}^{T^2} |(\dot{x} - f)_q - (\dot{x} - f)_k|^2 dt. \end{aligned}$$

In order to prove this theorem, we first verify the identity, for $q, k = 1, 2, \dots$,

$$(6.4) \quad \begin{aligned} \int_{T-1}^{T^2} (p_{qi} - p_{ki}) [(\dot{x} - f)_k^i - (\dot{x} - f)_q^i] dt \\ = [J(x_q, p_k) - J(x_k, p_k)] + [J(x_k, p_q) - J(x_q, p_q)]. \end{aligned}$$

Let $\tilde{\tau}$ be the positive constant in (3.2). Thus (6.4) and Theorem 3.2 imply for $q, k = 1, 2, \dots$

$$(6.5) \quad - \int_{T-1}^{T^2} (p_{qi} - p_{ki}) [(\dot{x} - f)_q^i - (\dot{x} - f)_k^i] dt \geq 2\tilde{\tau} \mathcal{D}(x_q, x_k).$$

Hence for $q, k = 1, 2, \dots$, one may expand the right-hand side of

$$\int_{T-1}^{T^2} |p_{q-1} - p_{k-1}|^2 dt = \int_{T-1}^{T^2} |p_q - p_k - \tau[(\dot{x} - f)_q - (\dot{x} - f)_k]|^2 dt$$

to obtain (6.3), which proves the theorem.

Setting $k = q - 1$ in (6.3) and summing, we obtain the following corollary.

Corollary 6.1. *The estimate below is valid for $k \geq 2$.*

$$(6.6) \quad \sum_{q=k}^{\infty} \mathcal{D}(x_q, x_{q-1}) \leq (4\tau\tilde{\tau})^{-1} \int_{T-1}^{T^2} |p_k - p_{k-1}|^2 dt.$$

In order to state the next lemma, we need some more notation. Let the neighborhood \mathfrak{P} of p_0 and positive constants τ_1, τ_2 be such that Theorems 3.2 and 3.3 apply. Let multipliers p in \mathfrak{P} and the parameter $-\tau_1 \leq \xi \leq \tau_2$ be given. By Theorem 3.2, this defines a ξ -parameter family of minimizing arcs and multipliers associated with the functional $J(x, p) = J(x, p, 0)$. We write this family

$$(6.7) \quad x(\xi): b(\xi), x(t, \xi), u(t, \xi),$$

$$(6.8) \quad p(\xi) = p + \xi[\dot{x} - f](\xi),$$

where $[\dot{x} - f](\xi)$ is the function $\dot{x} - f$ with arguments given by $x(\xi)$.

Theorem 3.3 implies that the functions $b(\xi), x(t, \xi), \dot{x}(t, \xi), u(t, \xi)$ have two continuous ξ -derivatives. Thus the family of ξ -derivatives of $x(\xi)$ can be seen to be a family

$$(6.9) \quad \eta(\xi): \beta(\xi), \eta(t, \xi), \mu(t, \xi)$$

of terminally admissible variations with respect to $x(\xi)$. One need only differentiate the terminal conditions

$$(6.10) \quad x^i(T^s(b(\xi))) = X^{is}(b(\xi))$$

with respect to ξ to obtain the secondary terminal conditions

$$(6.11) \quad \eta^i(T^s) = X_b^{is} \beta^b.$$

Differentiating (6.11) again with respect to ξ , we obtain an additional secondary terminal condition,

$$(6.12) \quad d\eta^i(T^s, \xi)/d\xi = X_{b^l}^{is} \beta^b \beta^l + X_b^{is} (d\beta^b/d\xi).$$

This condition (6.12) is used later on in the proof of Lemma 6.1.

We continue by defining

$$\psi(\xi) = \frac{1}{2} \int_{T_1}^{T^2} |\dot{x} - f(\xi)|^2 dt \quad \text{and} \quad J(\xi) = J(x(\xi), p(\xi)).$$

Furthermore, we define $J_1(\xi)$ and $J_2(\xi)$ to be the first and second variations of $J(x, p)$ with respect to $(x(\xi), p(\xi))$ and the terminally admissible variation $\eta(\xi)$.

Lemma 6.1. *The relations (6.13) hold.*

$$(6.13) \quad dJ(\xi)/d\xi = \psi(\xi), \quad d^2 J(\xi)/d\xi^2 = d\psi(\xi)/d\xi = -J_2(\xi).$$

In order to prove this lemma, we observe that $J_1(\xi)$, being the first variation associated with a minimizing arc and its Lagrange multipliers, vanishes identically as a function of ξ . That is, one has

$$(6.14) \quad g_b \beta^b + \int_{T_1}^{T_2} [\mathfrak{H}_b \beta^b + \mathfrak{H}_{x_i} \eta^i + \mathfrak{H}_{\dot{x}_i} \dot{\eta}^i + \mathfrak{H}_u \mu^k] dt = 0$$

Differentiating (6.14) with respect to ξ , one obtains

$$(6.15) \quad J_2(\xi) + T - \mathfrak{H}_{x_i} X_{hl}^{is} \beta^b \beta^l + d\psi(\xi)/d\xi = 0$$

where

$$T = \int_{T_1}^{T_2} [\mathfrak{H}_b d\beta^b/d\xi + \mathfrak{H}_{x_i} d\eta^i/d\xi + \mathfrak{H}_{\dot{x}_i} d\dot{\eta}^i/d\xi + \mathfrak{H}_u d\mu^k/d\xi] dt.$$

However, the first order necessary conditions (Euler equations, transversality) and (6.12) imply that $T = \mathfrak{H}_{x_i} X_{hl}^{is} \beta^b \beta^l$. Thus (6.15) implies

$$(6.16) \quad d\psi(\xi)/d\xi = -J_2(\xi).$$

On the other hand,

$$(6.17) \quad dJ(\xi)/d\xi = \psi(\xi)$$

follows by differentiating $J(\xi)$ with respect to ξ and applying (6.14). Thus (6.16) and (6.17) imply (6.13). This proves the lemma.

Theorem 3.2 (Rupp [7]) implies that the second variation of $J(x, p, \xi)$ is positive with respect to arcs x in a $bt\dot{x}u$ -neighborhood of x_0 and parameters (p, ξ) near $(p_0, 0)$. Thus by diminishing the neighborhood \mathfrak{B} and positive constants $\tau < \tau^*$ if necessary, we may assume that $J_2(\xi)$ is positive.

Theorem 6.3. *The inequality below holds for $q = 2, 3, \dots$,*

$$(6.18) \quad J(x_q, p_q) - J(x_{q-1}, p_{q-1}) \geq \tau \int_{T_1}^{T_2} |(\dot{x} - f)_{q-1}|^2 dt.$$

This theorem is proved by applying the mean value theorem to the function $J(\xi)$ which is obtained by letting $p = p_{q-1}$ in (6.8) for $q = 2, 3, \dots$. Thus one has for fixed $q \geq 2$

$$(6.19) \quad J(\xi_{q-1}) - J(0) = \xi_{q-1} \frac{d}{d\xi} J(0) + \xi_{q-1}^2 \int_{T_1}^{T_2} (1 - \theta) \frac{d^2}{d\xi^2} J(\theta\xi) d\theta.$$

Hence by (6.19) and Lemma 6.1, we have for fixed $q \geq 2$

$$(6.20) \quad \begin{aligned} & J(x_q, p_q) - J(x_{q-1}, p_{q-1}) \\ &= \xi_{q-1} \int_{T_1}^{T_2} |(\dot{x} - f)_{q-1}|^2 - \xi_{q-1}^2 \int_{T_1}^{T_2} (1 - \theta) J_2(\theta\xi) d\theta. \end{aligned}$$

Hence (6.20) and the nonnegativity of $J_2(\xi)$ imply (6.18). This proves the theorem.

The corollary below is an immediate consequence of Corollary 3.2 and Theorem 6.3.

Corollary 6.2. *Let the situation in Theorem 6.3 hold. One has*

$$\sum_{q=2}^{\infty} \int_{T^1}^{T^2} |(\dot{x} - f)_q|^2 dt \leq \tau^{-1} [I(x_0) - I(x_1, p_1)].$$

If the differential equation constraints (1.2) are normal relative to the problem here considered, then one can show

$$\limsup_t |p_q(t) - p_0(t)| = 0.$$

For the definition and properties of normality relative to the problem here considered, the reader is referred to Hestenes [2, pp. 270–283]. We begin with the following lemma.

Lemma 6.2. *If the differential equation constraints (1.2) are normal relative to the problem here considered, then there are terminally admissible variations $\eta_q; \beta_q = 0, \eta_q(t), \mu_q(t)$ such that*

$$(6.21) \quad \Phi(\eta_q) = p_q(t) - p_0(t)$$

where $\Phi(\eta_q)$ is the vector whose i th component is $\dot{\eta}_q^i - f_b^i \beta_q^b - f_{x_i}^i \eta_q^j - f_{u^k}^i \mu_q^k$. Furthermore, there is a positive constant K which is independent of q such that for $q = 1, 2, \dots$

$$(6.22) \quad \sup_t [|\eta_q(t)| + |\dot{\eta}_q(t)| + |\mu_q(t)|] \leq K.$$

In order to prove this lemma, we first note that by the normality assumption, there are solutions $(y_1, v_1), (y_2, v_2), \dots, (y_n, v_n)$ of the system of equations

$$(6.23) \quad \dot{y}^i(t) - f_{x_j}^i y^j(t) - f_{u^k}^i v^k(t) = 0$$

where the arguments of $f_{x_j}^i$ and $f_{u^k}^i$ are along x_q for some fixed $q \geq 1$. (If necessary, we agree to diminish the neighborhood \mathfrak{P} of p_0 and the positive constants $\delta, \tau < \tau^*$ which occur in the statement of Theorem 4.1.) These solutions $(y_1, v_1), (y_2, v_2), \dots, (y_n, v_n)$ may also be taken such that

$$(6.24) \quad y_j^i(T^1) = 0$$

and such that the matrix

$$(6.25) \quad \|y_j^i(T^2)\|$$

has rank n . They may also be normalized so that

$$(6.26) \quad \sup_t [|\dot{y}_i(t)| + |y_i(t)| + |v_i(t)|] = 1.$$

In particular, (6.26) implies that the entries in the matrix (6.25) are uniformly bounded.

Fix $q \geq 1$. Let $y^*(t)$ be a solution of

$$(6.27) \quad \dot{y}^{*i}(t) - f_{x_i}^i y^{*j}(t) = p_{q_i}(t) - p_0(t)$$

which satisfies the initial condition

$$(6.28) \quad y^{*i}(T^1) = 0.$$

Select constants K_1, K_2, \dots, K_n such that

$$(6.29) \quad y^{*i}(T^2) + K_j y_j^i(T^2) = 0.$$

By (6.24), (6.28), and (6.29), the variation $\eta_q: \beta_q, \eta_q(t), \mu_q(t)$ defined by

$$\beta_q = 0, \quad \eta_q(t) = y^*(t) + K_i y_i(t), \quad \mu_q(t) = K_i v_i(t)$$

is terminally admissible. By (6.23) and (6.27), the equation (6.21) is satisfied.

By Theorems 3.3 and 4.1, (6.26), (6.27), and (6.28), there is a positive constant K which does not depend on q such that (6.22) holds. This proves the lemma.

Lemma 6.3. *If the differential equation constraints (1.2) are normal relative to the problem here considered, then we have*

$$\lim \int_{T^1}^{T^2} |p_q - p_0|^2 dt = 0.$$

In order to prove this lemma, we need that along the pairs (x_q, p_q) for $q = 0, 1, \dots$, the terms $g_b, \tilde{\mathfrak{H}}_b, \tilde{\mathfrak{H}}_{x_i}, \tilde{\mathfrak{H}}_{\dot{x}_i}, \tilde{\mathfrak{H}}_{u^k}, f_{u^k}$, and L_{u^k} are uniformly bounded. This condition can be satisfied by diminishing if necessary the neighborhood \mathfrak{B} of p_0 and the positive constants $\tau < \tau^*$ and δ , which we now agree to do.

Let the terminally admissible variations $\eta_q: \beta_q = 0, \eta_q(t), \mu_q(t)$ and positive constant K be as described in Lemma 6.2. In particular, if one defines the inner product

$$\langle \eta, \eta^* \rangle = \beta^b \beta^{*b} + \int_{T^1}^{T^2} [\eta^i \eta^{*i} + \dot{\eta}^i \dot{\eta}^{*i} + \mu^k \mu^{*k}] dt,$$

then $\|\eta_q\|^2 = \langle \eta_q, \eta_q \rangle$ is uniformly bounded. Thus there is a terminally admissible variation $\eta_0: \beta_0 = 0, \eta_0(t), \mu_0(t)$ with respect to x_0 such that given any subsequence $\{q_1\}$ of $\{q\}$, there is a subsubsequence $\{q_2\}$ such that

$$(6.30) \quad \limsup_t |\eta_{q_2}(t) - \eta_0(t)| = 0$$

and such that $\dot{\eta}_{q_2}^i$ and $\mu_{q_2}^k$ converge weakly in $L_2(T^1, T^2)$ to $\dot{\eta}_0^i$ and μ_0^k .

Furthermore, by (4.4), there is a subsequence $\{q_3\}$ of $\{q_2\}$ such that

$$\lim b_{q_3} = b_0, \quad \limsup_t |x_{q_3}(t) - x_0(t)| = 0,$$

$$(6.31) \quad \lim \dot{x}_{q_3}(t) = \dot{x}_0(t) \quad \text{almost everywhere in } T^1 \leq t \leq T^2.$$

In particular, Egorov's theorem implies that the convergence in (6.31) is almost

uniform. Next, we observe that the first order necessary conditions imply that along (x_q, p_q) for $q = 1, 2, \dots$

$$(6.32) \quad g_b \beta_q^b + \int_{T_1}^{T_2} (\zeta_b \beta_q^b + \zeta_{x_i} \eta_q^i + \zeta_{\dot{x}_i} \dot{\eta}_q^i + \zeta_u \mu_q^k) dt = 0.$$

Finally, we have that along (x_0, p_0)

$$(6.33) \quad g_b \beta_0^b + \int_{T_1}^{T_2} (\zeta_b \beta_0^b + \zeta_{x_i} \eta_0^i + \zeta_{\dot{x}_i} \dot{\eta}_0^i + \zeta_u \mu_0^k) dt = 0.$$

Let a positive constant ϵ be given. Taking into account the convergence and uniform boundedness properties which we have indicated above, we see the following: There is a positive integer Q depending on ϵ such that, when we subtract (6.33) from (6.32), we obtain for $q_3 \geq Q$,

$$(6.34) \quad \int_{T_1}^{T_2} |(p_{q_3} - p_0)(\dot{\eta}_{q_3} - \int_b \beta_{q_3}^b - \int_{x_i} \eta_{q_3}^i - \int_u \mu_{q_3}^k)| dt < \epsilon.$$

Thus (6.34) and (6.21) imply $\lim \int_{T_1}^{T_2} |p_{q_3} - p_0|^2 dt = 0$. Since the subsequence $\{q_1\}$ was arbitrary, this proves $\lim \int_{T_1}^{T_2} |p_q - p_0|^2 dt = 0$. This proves the lemma.

Theorem 6.4. *If the arc x_0 is normal relative to the problem here considered, then*

$$(6.35) \quad \lim \left\{ |b_q - b_0| + \sup_t [|x_q - x_0| + |\dot{x}_q - \dot{x}_0| + |u_q - u_0|] \right\} = 0,$$

$$\lim \sup_t |p_q(t) - p_0(t)| = 0.$$

First note that the second limit implies the first by Theorem 3.3. In order to prove the second, we observe that by inequality (4.7), which was established in the proof of Theorem 4.1, for $q \geq 2$,

$$(6.36) \quad \sup |p_q - p_0| \leq R \sup |p_{q-1} - p_0| + K \left\{ [\mathcal{D}(x_q, x_0)]^{1/2} + \left[\int_{T_1}^{T_2} |p_q - p_0|^2 dt \right]^{1/2} \right\}$$

where $0 < R < 1$ and K are constants which are independent of q . Let τ and $\tilde{\tau}$ be the positive constants in (4.5). By (6.36) and Lemma 4.1, we have for $q \geq 2$,

$$(6.37) \quad \sup |p_q - p_0| \leq R \sup |p_{q-1} - p_0| + [1 + (4\tau\tilde{\tau})^{-1}] K \left[\int_{T_1}^{T_2} |p_{q-1} - p_0|^2 dt \right]^{1/2}.$$

Let $K^* = [1 + (4\tau\tilde{\tau})^{-1}] K$. Thus (6.36) implies for $q \geq 2$,

$$(6.38) \quad \begin{aligned} & \sup |p_q - p_0| \\ & \leq R^{q-1} \sup |p_1 - p_0| + K^* \sum_{j=0}^{q-2} R^j \left[\int_{T^1}^{T^2} |p_{q-1-j} - p_0|^2 dt \right]^{1/2}. \end{aligned}$$

Hence by (6.38), we have for $q \geq 2$,

$$(6.39) \quad \begin{aligned} \sup |p_{2q} - p_0| & \leq R^{2q-1} \sup |p_1 - p_0| + K^*(1-R)^{-1} \left[\int_{T^1}^{T^2} |p_{q-1} - p_0|^2 dt \right]^{1/2} \\ & + K^* R^q (q-1) \left[\int_{T^1}^{T^2} |p_1 - p_0|^2 dt \right]^{1/2}. \end{aligned}$$

An inequality similar to (6.39) holds for $2q+1$. Thus Lemma 6.3 implies (6.35). This proves the theorem.

7. Auxiliary lemmas. In this section, we complete the proof of Theorem 4.1 by proving Lemmas 4.1 and 4.2. Lemma 4.1 is contained in Lemma 7.1 while Lemma 4.2 is contained in Lemma 7.4.

We assume the notation and hypotheses of §4. In particular, the positive constants $\tilde{\tau}$, τ_1 , τ_2 , a *bt* x -neighborhood \mathfrak{N}_1 of x_0 , and a neighborhood \mathfrak{P}_1 of p_0 have been chosen so that Theorem 3.2 as well as the other theorems and corollaries in §3 apply. Multipliers p_q in \mathfrak{P}_1 and constants $0 \leq \tau \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots, Q-1$ are given. Theorem 3.2 implies the existence of minimizing arcs $x_q = x(p_q, 0)$ for $q = 1, 2, \dots, Q-1$. We suppose that for $q = 1, 2, \dots, Q-2$, $p_{q+1} = p_q + \xi_q (\dot{x} - f)_{q+1}$. Theorem 3.2 implies that new multipliers p_Q can be defined by $p_Q = p_{Q-1} + \xi_{Q-1} (\dot{x} - f)_Q$ where $(\dot{x} - f)_Q$ is the function $\dot{x} - f(b, t, x, u)$ with arguments given by the arc x_Q . By the uniqueness of a solution to the Euler equations and transversality condition (Corollary 5.1, Rupp [7]), the arc x_Q is also minimizing with respect to the multipliers p_Q , that is, $x_Q = x(p_Q, 0)$.

Lemma 7.1. *Let $\tilde{\tau}$ be the positive constant in (3.2), and let τ be the nonnegative constant described above. The inequality below is valid for $q = 1, 2, \dots, Q-1$.*

$$(7.1) \quad \begin{aligned} & \int_{T^1}^{T^2} |p_q - p_0|^2 dt \\ & \geq \int_{T^1}^{T^2} |p_{q+1} - p_0|^2 dt + 4\tilde{\tau} \mathcal{D}(x_{q+1}, x_0) + \tau^2 \int_{T^1}^{T^2} |(\dot{x} - f)_{q+1}|^2 dt. \end{aligned}$$

The proof of this lemma requires the identity, for $q = 1, 2, \dots, Q-1$,

$$(7.2) \quad \begin{aligned} & [J(x_0, p_{q+1}) - J(x_{q+1}, p_{q+1})] + [J(x_{q+1}, p_0) - J(x_0, p_0)] \\ & = - \int_{T^1}^{T^2} (\dot{x} - f)_{q+1}^i (p_{q+1}^i - p_0^i) dt. \end{aligned}$$

By (7.2) and Theorem 3.2, one has for $q = 1, 2, \dots, Q-1$

$$(7.3) \quad - \int_{T^1}^{T^2} (\dot{x} - f)_{q+1}^i (p_{q+1}^i - p_0^i) dt \geq 2\tilde{\tau} \mathcal{D}(x_{q+1}, x_0).$$

Now observe that for $q = 1, 2, \dots, Q-1$

$$(7.4) \quad \begin{aligned} \int_{T^1}^{T^2} |p_q - p_0|^2 dt &= \int_{T^1}^{T^2} |(p_{q+1} - p_0) - \xi_q(\dot{x} - f)_{q+1}|^2 dt \\ &= \int_{T^1}^{T^2} |p_{q+1} - p_0|^2 dt - 2\xi_q \int_{T^1}^{T^2} (p_{q+1}^i - p_0^i)(\dot{x} - f)_{q+1}^i dt \\ &\quad + \xi_q^2 \int_{T^1}^{T^2} |(\dot{x} - f)_{q+1}|^2 dt. \end{aligned}$$

Thus (7.1) follows from (7.3) and (7.4). This proves the lemma.

Lemma 7.2. *There are positive constants τ^* and K and a neighborhood \mathfrak{P} of p_0 such that the following holds: If p_q is in \mathfrak{P} and $0 \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots, Q-1$, then for $q = 1, 2, \dots, Q$*

$$(7.5) \quad \int_{T^1}^{T^2} [|b_q - b_0| + |x_q - x_0| + |\dot{x}_q - \dot{x}_0| + |u_q - u_0|] dt \leq K[\mathcal{D}(x_q, x_0)]^{1/2}.$$

In order to prove this lemma, we apply Lemma 2.1 (Rupp [9]). This yields a $btx\dot{x}u$ -neighborhood \mathfrak{N} of x_0 and a positive constant K_1 such that whenever $x : b, x(t), u(t)$ is in \mathfrak{N} , one has

$$(7.6) \quad \int_{T^1}^{T^2} [|b - b_0|^2 + |x - x_0|^2 + |\dot{x} - \dot{x}_0|^2 + |u - u_0|^2] dt \leq K_1 \mathcal{D}(x, x_0).$$

By Theorem 3.3 we choose the positive constant τ^* and a neighborhood \mathfrak{P} of p_0 so small that x_q is in \mathfrak{N} for $q = 1, 2, \dots, Q$ whenever p_q is in \mathfrak{P} and $0 \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots, Q-1$. Thus (7.6) implies the inequality below for $q = 1, 2, \dots, Q$,

$$(7.7) \quad \int_{T^1}^{T^2} [|b_q^* - b_0|^2 + |x_q^* - x_0|^2 + |\dot{x}_q^* - \dot{x}_0|^2 + |u_q^* - u_0|^2] dt \leq K_1 \mathcal{D}(x_q^*, x_0).$$

Thus (7.7) and Hölder's inequality imply (7.5) with $K = K_1 |T^2 - T^1|^{-1/2}$. This proves the lemma.

We now consider the constants $C_i(t^*)$,

$$(7.8) \quad C_i(t^*) = [p_i + (C + \xi)(\dot{x}^i - f^i)]_{t=t^*},$$

which occur in the Euler equations (4.1) associated with the state variables. Let $C(t^*)$ be the vector whose components are given by $C_i(t^*)$. Let $C_q(t^*)$ be the vector $C(t^*)$ associated with the arc x_q and parameters $(p, \xi) = (p_q, 0)$ for $q = 0, 1, \dots, Q$.

Lemma 7.3. *There are positive constants τ^* and K and a neighborhood \mathfrak{P}*

of p_0 such that the following holds: If p_q is in \mathfrak{P} and $0 \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots, Q-1$, then for some $T^1 \leq t_q \leq T^2$, one has for $q = 1, 2, \dots, Q$

$$(7.9) \quad |C_q(t_q) - C_0(t_q)| \leq K \left\{ [\mathcal{D}(x_q, x_0)]^{1/2} + \left[\int_{T^1}^{T^2} |p_q - p_0|^2 dt \right]^{1/2} \right\}.$$

In order to prove this lemma, we observe that $\dot{x} - f(b, t, x, u)$ has a continuous $b\dot{x}\dot{x}u$ -derivative in a $b\dot{x}\dot{x}u$ -neighborhood of x_0 . Thus we may choose a $b\dot{x}\dot{x}u$ -neighborhood \mathfrak{N} of x_0 and a constant K_1 such that whenever the arc x^* : $b^*, x^*(t), u^*(t)$ is in \mathfrak{N} , one has

$$(7.10) \quad |\dot{x}^* - f(b^*, t, x^*, u^*)| \leq K_1[|b^* - b_0| + |x^* - x_0| + |\dot{x}^* - \dot{x}_0| + |u^* - u_0|].$$

By Theorem 3.2, we choose a positive constant τ_1 and a neighborhood \mathfrak{P}_1 of p_0 so small that x_q is \mathfrak{N} for $q = 1, 2, \dots, Q$ whenever p_q is in \mathfrak{P}_1 and $0 \leq \xi_q \leq \tau_1$ for $q = 1, 2, \dots, Q-1$. Thus (7.10) implies the inequality below for any $T^1 \leq t \leq T^2$, $q = 1, 2, \dots, Q$,

$$(7.11) \quad |(\dot{x} - f)_q| \leq K_1[|b_q - b_0| + |x_q - x_0| + |\dot{x}_q - \dot{x}_0| + |u_q - u_0|].$$

From (7.10) and (7.11), we now obtain for $q = 1, 2, \dots, Q$

$$(7.12) \quad \begin{aligned} & \int_{T^1}^{T^2} |C_q(t) - C_0(t)| dt \\ & \leq CK_1 \int_{T^1}^{T^2} [|b_q - b_0| + |x_q - x_0| + |\dot{x}_q - \dot{x}_0| + |u_q - u_0|] dt \\ & \quad + \int_{T^1}^{T^2} |p_q - p_0| dt. \end{aligned}$$

Let the positive constants τ_2 and K_2 and the neighborhood \mathfrak{P}_2 of p_0 be such that Lemma 7.2 applies. Let $\mathfrak{P} = \mathfrak{P}_1 \cap \mathfrak{P}_2$ and let $\tau^* = \min\{\tau_1, \tau_2\}$. By (7.12), Lemma 7.2, and Hölder's inequality we obtain when p_q is in \mathfrak{P} and $0 \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots, Q-1$

$$(7.13) \quad \begin{aligned} & \int_{T^1}^{T^2} |C_q(t) - C_0(t)| dt \\ & \leq CK_1 K_2 [\mathcal{D}(x_q, x_0)]^{1/2} + |T^2 - T^1|^{1/2} \left[\int_{T^1}^{T^2} |p_q - p_0|^2 dt \right]^{1/2}. \end{aligned}$$

By the mean value theorem for integrals, there are $T^1 \leq t_q \leq T^2$ for $q = 1, 2, \dots, Q$ such that

$$(7.14) \quad |T^2 - T^1| |C_q(t_q) - C_0(t_q)| = \int_{T^1}^{T^2} |C_q(t) - C_0(t)| dt.$$

Thus (7.13) and (7.14) imply (7.9) with $K = |T^2 - T^1|^{-1} \max\{|T^2 - T^1|^{1/2}, CK_1 K_2\}$. This proves the lemma.

Lemma 7.4. *There are positive constants τ^* and K , and a neighborhood \mathfrak{P} of p_0 such that the following holds: If p_q is in \mathfrak{P} and $0 \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots, Q-1$, then one has the inequality below for $q = 1, 2, \dots, Q$.*

$$(7.15) \quad \sup |p_q - p_0 + C(\dot{x} - f)_q| \leq K \left\{ [\mathcal{D}(x_q, x_0)]^{1/2} + \left[\int_{T^1}^{T^2} |p_q - p_0|^2 dt \right]^{1/4} \right\}.$$

In order to prove this lemma, we recall that the Euler equations (4.1) associated with the state variables are

$$(7.16) \quad p_i + (C + \xi)(\dot{x}^i - f^i) = \int_{t^*}^t \mathfrak{H}_{x_i} dt + C_i(t^*).$$

The function $\mathfrak{H}_{x_i}(b, t, x, \dot{x}, u, p, \xi)$ has a continuous *bt \dot{x} up*-derivative in a *bt \dot{x} up*-neighborhood of (x_0, p_0) . Thus there is a *bt \dot{x} up*-neighborhood \mathfrak{N} of (x_0, p_0) and a constant K_1 such that on \mathfrak{N} ,

$$(7.17) \quad |\mathfrak{H}_{x_i}(b, t, x, \dot{x}, u, p, 0) - \mathfrak{H}_{x_i}(b_0, t, x_0, \dot{x}_0, u_0, p_0, 0)| \leq K_1[|b - b_0| + |x - x_0| + |\dot{x} - \dot{x}_0| + |u - u_0| + |p - p_0|].$$

Hence by Theorem 3.2, we choose a positive constant τ_1 and a neighborhood \mathfrak{P}_1 of p_0 so small that (x_q, p_q) for $q = 1, 2, \dots, Q$ is in \mathfrak{N} whenever p_q is in \mathfrak{P}_2 and $0 \leq \xi_q \leq \tau_1$ for $q = 1, 2, \dots, Q-1$. Thus (7.17) implies, for $q = 1, 2, \dots, Q$,

$$(7.18) \quad \int_{T^1}^{T^2} |\mathfrak{H}_{x_i q} - \mathfrak{H}_{x_i 0}| dt \leq K_1 \int_{T^1}^{T^2} [|b_q - b_0| + |x_q - x_0| + |\dot{x}_q - \dot{x}_0| + |u_q - u_0| + |p_q - p_0|] dt$$

where $\mathfrak{H}_{x_i q}$ and $\mathfrak{H}_{x_i 0}$ denote the function $\mathfrak{H}_{x_i}(b, t, x, \dot{x}, u, p, \xi)$ with arguments given by $(x, p, \xi) = (x_q, p_q, 0)$ and $(x, p, \xi) = (x_0, p_0, 0)$.

Let the positive constants τ_2 and K_2 and the neighborhood \mathfrak{P}_2 of p_0 be such that Lemma 7.2 applies. Let $\tau_3 = \min\{\tau_1, \tau_2\}$ and $\mathfrak{P}_3 = \mathfrak{P}_1 \cap \mathfrak{P}_2$. Thus when p_q is in \mathfrak{P}_3 and $0 \leq \xi_q \leq \tau_3$ for $q = 1, 2, \dots, Q-1$, one has by (7.18), Lemma 7.2, and Hölder's inequality, for $q = 1, 2, \dots, Q$,

$$(7.19) \quad \int_{T^1}^{T^2} |\mathfrak{H}_{x_i q} - \mathfrak{H}_{x_i 0}| dt \leq K_1 K_2 [\mathcal{D}(x_q, x_0)]^{1/2} + |T^2 - T^1|^{1/2} K_1 \left[\int_{T^1}^{T^2} |p_q - p_0|^2 dt \right]^{1/2}.$$

Let the positive constants τ_4 and K_3 , the neighborhood \mathfrak{P}_4 of p_0 , and $T^1 \leq t_q \leq T^2$ be such that Lemma 7.3 applies. Let $\tau^* = \min\{\tau_3, \tau_4\}$, and $\mathfrak{P} = \mathfrak{P}_3 \cap \mathfrak{P}_4$. Thus when p_q is in \mathfrak{P} and $0 \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots, Q-1$, one has by (7.19) and Lemma 7.3, for $q = 1, 2, \dots, Q$,

$$(7.20) \quad \int_{T^1}^{T^2} |\xi_{x_q} - \xi_{x_0}| dt + |C_q(t_q) - C_0(t_q)| \\ \leq K \left\{ [\mathbb{D}(x_q, x_0)]^{1/2} + \left[\int_{T^1}^{T^2} |p_q - p_0|^2 dt \right]^{1/2} \right\}$$

where $K = 2 \max\{|T^2 - T^1|^{1/2} K_1, K_1 K_2, K_3\}$. Hence (7.20), (7.16), and the definition $p_q = p_{q-1} + \xi_{q-1}(\dot{x} - f)_q$ imply (7.15) for $q = 2, 3, \dots, Q$. This proves the lemma.

8. The finite dimensional case. This section illustrates the application of the preceding theory to the problem of minimizing a real-valued function f of several real variables x^i , $i = 1, 2, \dots, n$, subject to equality constraints $g_\alpha = 0$, $\alpha = 1, 2, \dots, m$. In many cases, calculus analogies lead nowhere. The problem considered here is an exception. As mentioned in Hestenes [4, p. 307], it is not hard to show results of this nature directly.

Let x be a vector in n -dimensional space whose components are given by x^i , $i = 1, 2, \dots, n$. Let $f(x)$ and $g_\alpha(x)$, $\alpha = 1, 2, \dots, m$, be real-valued functions of x . The indices i and α are always assumed to have the above ranges. Let $f(x)$ and $g_\alpha(x)$ be twice continuously differentiable in a neighborhood of the vector x_0 . Let $\lambda_{0\alpha}$ be the Lagrange multipliers associated with x_0 . That is, one has at x_0 ,

$$(8.1) \quad f_{x^i} + \lambda_{0\alpha} g_{\alpha x^i} = 0.$$

We also assume the strong condition of Legendre. That is, for any $\Pi \neq 0$ such that $g_{\alpha x^i}(x_0)\Pi^i = 0$, one has at x_0 ,

$$(8.2) \quad f_{x^i x^j} \Pi^i \Pi^j + \lambda_{0\alpha} g_{\alpha x^i x^j} \Pi^i \Pi^j > 0.$$

In the present case, (8.1) and (8.2) constitute the sufficiency criteria (2.1) through (2.4). Thus Theorems 3.1 and 3.2 imply the theorem below.

Theorem 8.1. *There are positive constants $\tilde{\tau}$, τ_1 , τ_2 , C , a neighborhood Λ of λ_0 , and a neighborhood \mathfrak{N} of x_0 such that the following holds: Let*

$$F(x, \lambda, \xi) = f(x) + \lambda_\alpha g_\alpha(x) + \frac{1}{2}(C + \xi)g_\alpha(x)g_\alpha(x).$$

Let multipliers λ in Λ and a constant $-\tau_1 \leq \xi \leq \tau_2$ be given. There is $x(\lambda, \xi) = x^$ in \mathfrak{N} such that given x in \mathfrak{N} , one has*

$$(8.3) \quad F(x, \lambda, \xi) - F(x^*, \lambda, \xi) \geq \tilde{\tau} |x - x^*|^2.$$

Similarly, by Theorem 3.3, one has

Theorem 8.2. *The family $x(\lambda, \xi)$ has two continuous $\lambda\xi$ -derivatives.*

As a consequence of Theorem 4.1, we have

Theorem 8.3. *There are positive constants $\tilde{\tau}$, $\tau < \tau^*$, δ , a neighborhood Λ of λ_0 , and a neighborhood \mathcal{N} of x_0 with the following property: Let multipliers $|\lambda_1 - \lambda_0| < \delta$ be given, and let constants $\tau \leq \xi_q \leq \tau^*$ for $q = 1, 2, \dots$ be given. The algorithm below is convergent in Λ ,*

$$(8.4) \quad \lambda_{q+1} = \lambda_q + \xi_q g(x_{q+1})$$

where $x_q = x(\lambda_q, 0)$ and $g(x_{q+1}) = g_{q+1}$ is the vector whose components are given by $g_\alpha(x_{q+1})$. That is, one has

$$(8.5) \quad \lim x_q = x_0.$$

Analogues of the convergence estimates in §6 and Lemma 4.1 are valid. In particular, one can prove

Theorem 8.4. *There are positive constants τ_1 and τ_2 such that for $q = 1, 2, \dots$,*

$$|\lambda_q - \lambda_0|^2 \geq |\lambda_{q+1} - \lambda_0|^2 + \tau_1 |x_{q+1} - x_0|^2 + \tau_2 |g_{q+1}|^2.$$

Furthermore, if the constraints are normal, one has $\lim |\lambda_q - \lambda_0| = 0$.

REFERENCES

1. 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. MR 13, 759.
2. ———, *Calculus of variations and optimal control theory*, Wiley, New York, 1966. MR 34 #3390.
3. ———, *An indirect sufficiency proof for the problem of Bolza in nonparametric form*, Trans. Amer. Math. Soc. **62** (1947), 509–535. MR 9, 360.
4. ———, *Multiplier and gradient methods*, J. Optimization Theory Appl. **4** (1969), 303–320. MR 42 #6690.
5. R. D. Rupp, *Approximation of the classical isoperimetric problem*, J. Optimization Theory Appl. **4** (1972), 251–264.
6. ———, *A method for solving a quadratic optimal control problem*, J. Optimization Theory Appl. **4** (1972), 238–250.
7. ———, *A new type of variational theory sufficiency theorem*, Pacific J. Math. **40** (1972), 415–444.
8. ———, *A unified sufficiency theory in the calculus of variations*, Dissertation, University of California, Los Angeles, Calif., 1970.
9. ———, *The Weierstrass excess function*, Pacific J. Math. **41** (1972), 529–536.