

# Second-Order Necessary Conditions in Optimal Control: Accessory-Problem Results without Normality Conditions<sup>1</sup>

E. G. GILBERT<sup>2</sup> AND D. S. BERNSTEIN<sup>3</sup>

Communicated by W. F. Powers

Dedicated to L. Cesari

**Abstract.** An optimal control problem, which includes restrictions on the controls and equality/inequality constraints on the terminal states, is formulated. Second-order necessary conditions of the accessory-problem type are obtained in the absence of normality conditions. It is shown that the necessary conditions generalize and simplify prior results due to Hestenes (Ref. 5) and Warga (Refs. 6 and 7).

**Key Words.** Optimal control, second-order necessary conditions, accessory problem, control constraints, normality condition.

## 1. Introduction

Second-order theory for problems in optimal control has been studied for some time. The earliest results were motivated by the work of Bliss (Ref. 1) in the calculus of variations. They concerned problems with open control sets and depended on heuristic arguments to obtain both necessary conditions and local sufficiency conditions; see, for example, Refs. 2-4, which include some additional references. Mathematically rigorous treatments of second-order necessary conditions seem to be limited (Refs. 5-8).

---

<sup>1</sup> This research was sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-77-3158.

<sup>2</sup> Professor, Computer, Information, and Control Engineering Program and Department of Aerospace Engineering, University of Michigan, Ann Arbor, Michigan.

<sup>3</sup> Staff Member, Control Systems Engineering Group, MIT, Lincoln Laboratory, Lexington, Massachusetts.

Hestenes, whose work is the earliest, considered a fairly general optimal control problem, but made the standard assumption that the control set is open. His main result (Theorem 9.1, Chapter 6, Ref. 5) states that the second variation of a suitably defined function is nonnegative on a set of admissible variations related to the first-order necessary conditions. More recently, Warga (Ref. 6) obtained a similar result, stated in a somewhat different way, for problems where the controls are restricted to a convex, not necessarily open, constraint set. A more radical departure is considered in Refs. 7 and 8, where relaxed controls are introduced. The resulting second-order conditions may be interpreted as exploiting strong (Weierstrass-Pontryagin) variations instead of the weak (Lagrange) variations used in Refs. 5 and 6. It should be emphasized that all of the above conditions are of the accessory-problem type and are global over time. Thus, they are quite different from the pointwise-in-time, second-order conditions considered in the theory of singular control (Refs. 9-12). See Ref. 7 for some additional remarks on these matters.

The approaches to the proofs in Refs. 5-8 appear philosophically different. Hestenes works directly in the space of state-control pairs, while Warga obtains necessary conditions for an abstract optimization problem, which he then applies to the optimal control problem. There is, however, a common ground. Both modes of proof rely on the implicit function theorem to generate a one-parameter family of admissible elements. This is done in such a way that the range of a variational operator, which is associated with problem constraints, must have full dimension. Conditions which imply the full-range property are called normality conditions, because they imply (among other things) that the multiplier associated with the cost in the necessary conditions is nonzero.

Normality conditions, such as those found in Refs. 5-8, are unpleasant because they are usually difficult to verify and may, on occasion, fail to be satisfied. This is why first-order conditions which do not require normality, such as those due to Fritz John (in mathematical programming) and Pontryagin (in optimal control), are popular. In the absence of normality conditions, there are some second-order necessary conditions in the mathematical programming literature (Refs. 13-16); but, apart from the singular control results (Refs. 9, 11, 12), there seem to be none in the theory of optimal control. Our objective in this paper is to make significant progress in filling this gap.

We begin in Section 2 by stating a rather general optimal control problem which requires weak (not necessarily convex) assumptions on the control set and includes mixed equality and inequality constraints on the initial and terminal states. This problem subsumes the problems in Refs. 5-8 as special cases. Our initial attention (Section 3) is on the case of weak

variations. Theorem 3.1 contains the necessary conditions which are stated concisely in terms of Lagrangian and Hamiltonian functions and a multiplier-free characterization of admissible variations. The lack of a normality condition leads to an unusual min-max formulation of an accessory minimum problem. The necessary conditions take on a simpler and more familiar form in the presence of normality and regularity. Precise definitions of these concepts are given and they are characterized concretely [by Conditions (C1) and (C2)]. For open control sets, normality and regularity together are equivalent to natural generalizations of Hestenes' normality conditions [our Condition (C4) or (C5)]. In the presence of control constraints, the natural generalization of Warga's (Ref. 6) normality condition [our Condition (C6)] is sufficient, but not necessary for normality and regularity. Thus, Warga's second-order necessary conditions may hold under weaker conditions than he states. Specifically, our Condition (C3), which implies that his second-order conditions hold, may be satisfied under his alternative (a), Theorem 1.1, Ref. 6.

The proof of Theorem 3.1 is based on necessary conditions for an abstract optimization problem obtained recently by the second author (Ref. 17). These results are stated without proof in Section 4 and are used in Section 5 to prove Theorem 3.1.

The theory developed in Sections 3–5 for weak variations of ordinary controls can be extended with little change to relaxed controls (Ref. 7) and mixed relaxed-ordinary controls (Ref. 8). This is done for relaxed controls in Section 6. The results generalize those in Ref. 7 in much the same way the results in Section 3 generalize those in Ref. 6. Under certain assumptions, the necessary conditions for the relaxed problem become necessary conditions which apply to an optimum over the class of ordinary controls. This result leads to a second-order necessary condition for ordinary controls (Theorem 6.4), which is expressed entirely in terms of ordinary controls and applies under weaker assumptions than those used in Theorem 3.1. Even when the assumptions for Theorem 3.1 hold, the second-order condition in Theorem 6.4 is quite different from the one in Theorem 3.1.

## 2. Statement of the Optimal Control Problem

We need the following notation. Let  $T \triangleq [t_1, t_2]$  be a closed interval of the real numbers  $\mathbb{R}$  and  $S \subset \mathbb{R}^s$ . Use  $M(T, S)$ ,  $L_\infty(T, S)$ ,  $PC(T, S)$ , and  $AC(T, S)$  to denote, respectively, the set of all functions  $g: T \rightarrow S$  which are measurable, measurable and essentially bounded, piecewise continuous, and absolutely continuous. Let  $m$  and  $n$  be positive integers,  $j$  and  $k$  be

nonnegative integers,  $X \subset \mathbb{R}^n$  and  $U \subset \mathbb{R}^m$  be open sets,

$$\mathcal{X} \triangleq \text{AC}(T, X), \quad \mathcal{U} \subset \text{M}(T, U),$$

$$f: T \times X \times U \rightarrow \mathbb{R}^n, \quad \phi_i: X \times X \rightarrow \mathbb{R}, \quad \text{for } i \in \{0, \dots, j\},$$

$$\psi: X \times X \rightarrow \mathbb{R}^k, \quad \text{if } k > 0.$$

**Optimal Control Problem (OCP).** Find  $(x(\cdot), u(\cdot)) \in \mathcal{X} \times \mathcal{U}$  which minimizes

$$J(x(\cdot), u(\cdot)) \triangleq \phi_0(x(t_1), x(t_2)), \quad (1)$$

subject to

$$\phi_i(x(t_1), x(t_2)) \leq 0, \quad i \in \{1, \dots, j\} \quad (\text{omit if } j = 0), \quad (2)$$

$$\psi(x(t_1), x(t_2)) = 0 \quad (\text{omit if } k = 0), \quad (3)$$

$$\dot{x}(t) = f(t, x(t), u(t)), \quad \text{a.a. } t \in T. \quad (4)$$

This problem generalizes in several ways the optimal control problems treated by Hestenes and Warga. In Ref. 5, Ineq. (2) is omitted, (3) is replaced by

$$x(t_1) = \bar{x}_1, \quad x(t_2) = \bar{x}_2,$$

where  $\bar{x}_1, \bar{x}_2 \in \mathbb{R}^n$  are fixed, and there are isoperimetric equality and inequality constraints involving integrals of  $x(t)$  and  $u(t)$  on  $T$ . The isoperimetric constraints can be written in the form of (2) and (3) if additional components are added to  $f$  and  $x(t)$  in (4). In Ref. 6, Ineq. (2) is omitted, and (3) assumes the form

$$x(t_1) = 0, \quad \hat{\psi}(x(t_2)) = 0.$$

We impose a very loose condition on the set of admissible controls  $\mathcal{U}$ ; it need only be a subset of  $\text{M}(T, U)$ . Thus, it may incorporate a variety of interesting constraints beyond the classic situation of Ref. 5, where

$$\mathcal{U} \triangleq \text{PC}(T, U)$$

is open. For instance, it includes the case considered in Ref. 6, where

$$\mathcal{U} \triangleq \{u(\cdot) \in L_\infty(T, U_0) : u(t) \in \hat{U}(t), t \in T\},$$

$$\hat{U}(t) \subset U_0, \quad \text{for all } t \in T,$$

$U_0 \subset U$  is compact, and  $\hat{U}(t)$  is closed and convex for all  $t \in T$ .

To obtain necessary conditions for optimality, notation and assumptions concerning the functions  $f, \phi_i, \psi$  must be introduced. We begin with some notation for the derivatives of functions of several variables.

Let  $g: A \subset \mathbb{R}^r \rightarrow \mathbb{R}^s$ , where  $A$  is open and  $g$  is  $C^1$ . For  $\bar{a} \in A$ ,

$$g'(\bar{a}): \mathbb{R}^r \rightarrow \mathbb{R}^s$$

denotes the first (Fréchet) derivative of  $g$  at  $\bar{a}$ . The linear map  $g'(\bar{a})$  is written as

$$y = g'(\bar{a})x,$$

where  $g'(\bar{a})$  is identified with the Jacobian matrix of  $g$ . Suppose that  $g$  is  $C^2$ . For  $\bar{a} \in A$ ,

$$g''(\bar{a}): \mathbb{R}^r \times \mathbb{R}^r \rightarrow \mathbb{R}^s$$

denotes the second (Fréchet) derivative of  $g$  at  $\bar{a}$ . The bilinear map  $g''(\bar{a})$  is written as

$$y = g''(\bar{a})(x_1, x_2),$$

where  $g''(\bar{a})$  is identified with the Hessian matrices of the components of  $g$ . Specifically, the  $i$ th component of  $y$  is written as

$$y_i = x_1^T g_i''(\bar{a})x_2,$$

where  $T$  denotes transpose and  $g_i''(\bar{a})$  is the symmetric  $r \times r$  matrix of second partial derivatives of the  $i$ th component of  $g$  at  $\bar{a}$ . For notational simplicity,

$$g''(\bar{a})(x)^2 \triangleq g''(\bar{a})(x, x).$$

Suppose that

$$r = \sum_{i=1}^{\nu} r_i, \quad A = A_1 \times A_2 \times \cdots \times A_{\nu}, \quad \bar{a} = (\bar{a}_1, \dots, \bar{a}_{\nu}),$$

where

$$A_i \subset \mathbb{R}^{r_i} \quad \text{and} \quad \bar{a}_i \in \mathbb{R}^{r_i}.$$

Then, the first partial derivative of  $g$  with respect to the  $p$ th component of  $a$  at  $\bar{a}$ , denoted by  $g_{a_p}(\bar{a})$ , is the first derivative of the map

$$\hat{g}: A_p \rightarrow \mathbb{R}^s,$$

where

$$\hat{g}(a_p) = g(\bar{a}_1, \dots, \bar{a}_{p-1}, a_p, \bar{a}_{p+1}, \dots, \bar{a}_{\nu}).$$

Similarly, the second partial derivative

$$g_{a_p a_q}(\bar{a}): \mathbb{R}^{r_p} \times \mathbb{R}^{r_q} \rightarrow \mathbb{R}^s$$

may be defined. It is a bilinear map, which may be identified with  $s \times r_p \times r_q$  matrices of second partial derivatives.

For  $y \in \mathbb{R}^s$ , linear maps

$$\mathcal{A}: \mathbb{R}^r \rightarrow \mathbb{R}^s,$$

and bilinear maps

$$\mathcal{B}: \mathbb{R}^{r_1} \times \mathbb{R}^{r_2} \rightarrow \mathbb{R}^s,$$

let  $|y|$ ,  $|\mathcal{A}|$ ,  $|\mathcal{B}|$  denote sup norms over the elements of the characterizing matrices.

Depending on whether we want first-order or second-order necessary conditions, the first or both of the following assumptions are required.

**Assumption 2.1.** For all  $t \in T$ ,  $f(t, \cdot, \cdot)$  is  $C^1$  on  $X \times U$ ; for all  $(x, u) \in X \times U$ , the components of  $f(\cdot, x, u)$  and the elements of  $f_x(\cdot, x, u)$  and  $f_u(\cdot, x, u)$  are measurable on  $T$ ; there exists an integrable function  $\zeta: T \rightarrow \mathbb{R}$  such that, for all  $(t, x, u) \in T \times X \times U$ ,

$$|f(t, x, u)| + |f_x(t, x, u)| + |f_u(t, x, u)| \leq \zeta(t);$$

for  $i \in \{0, \dots, j\}$ ,  $\phi_i$  is  $C^1$  on  $X \times X$ ;

if  $k > 0$ ,  $\psi$  is  $C^1$  on  $X \times X$ .

**Assumption 2.2.** For all  $t \in T$ ,  $f(t, \cdot, \cdot)$  is  $C^2$  on  $X \times U$ ; for all  $(x, u) \in X \times U$ , the elements of the matrices which define  $f_{xx}(\cdot, x, u)$ ,  $f_{xu}(\cdot, x, u)$ ,  $f_{uu}(\cdot, x, u)$  are measurable on  $T$ ; there exists an integrable function  $\zeta: T \rightarrow \mathbb{R}$  such that, for all  $(t, x, u) \in T \times X \times U$ ,

$$|f_{xx}(t, x, u)| + |f_{xu}(t, x, u)| + |f_{uu}(t, x, u)| \leq \zeta(t);$$

for  $i \in \{0, \dots, j\}$ ,  $\phi_i$  is  $C^2$  on  $X \times X$ ;

if  $k > 0$ ,  $\psi$  is  $C^2$  on  $X \times X$ .

Apart from the fact that  $U$  need not be bounded, these assumptions are essentially those used in Ref. 6. They are certainly weaker than those appearing in Ref. 5.

### 3. Necessary Conditions

This section contains the main results of the paper. We begin by introducing some special notation. Then, we state the necessary conditions (Theorem 3.1) and elaborate on them through a series of theorems and remarks. The longer proofs appear in Section 5.

In the statement of the necessary conditions,  $(\bar{x}(\cdot), \bar{u}(\cdot)) \in \mathcal{X} \times \mathcal{U}$  designates a solution of OCP. Since many functions must be evaluated on  $(\bar{x}(t), \bar{u}(t))$ , this evaluation is represented compactly by a superbar. For example,

$$\bar{f}_x(t) \triangleq f_x(t, \bar{x}(t), \bar{u}(t)) \quad \text{and} \quad \bar{\psi} \triangleq \psi(\bar{x}(t_1), \bar{x}(t_2)).$$

We require sets of state and control variations. The state variations are elements of

$$\tilde{\mathcal{X}} \triangleq \text{AC}(T, \mathbb{R}^n).$$

The set of control variations is any set  $\tilde{\mathcal{U}}$  which has the following properties:

$$\tilde{\mathcal{U}} \subset L_\infty(T, \mathbb{R}^m), \tag{5}$$

$$\tilde{\mathcal{U}} \subset \mathcal{U} - \bar{u}(\cdot) \triangleq \{\omega(\cdot) : \omega(\cdot) = u(\cdot) - \bar{u}(\cdot), u(\cdot) \in \mathcal{U}\}, \tag{6}$$

$$0 \in \tilde{\mathcal{U}}, \tag{7}$$

$$\tilde{\mathcal{U}} \text{ is convex.} \tag{8}$$

As will be discussed later, the freedom in the choice of  $\tilde{\mathcal{U}}$  has certain advantages. The set of critical directions, i.e., the elements of  $\tilde{\mathcal{X}} \times \tilde{\mathcal{U}}$  which satisfy both the linearized equations of motion and the linearized active terminal constraints and have inferior linearized cost, is also required. It is given by

$$\mathcal{D} \triangleq \{(\eta(\cdot), \omega(\cdot)) \in \tilde{\mathcal{X}} \times \tilde{\mathcal{U}} : (\eta(\cdot), \omega(\cdot)) \text{ satisfies (10)–(12)}\}, \tag{9}$$

where

$$\bar{\phi}_{ix_1}\eta(t_1) + \bar{\phi}_{ix_2}\eta(t_2) \leq 0, \quad i \in I_{A_0}, \tag{10}$$

$$\bar{\psi}_{x_1}\eta(t_1) + \bar{\psi}_{x_2}\eta(t_2) = 0 \quad (\text{omit if } k = 0), \tag{11}$$

$$\dot{\eta}(t) = \bar{f}_x(t)\eta(t) + \bar{f}_u(t)\omega(t), \quad \text{a.a. } t \in T, \tag{12}$$

and

$$I_{A_0} \triangleq \{0\}, \quad j = 0, \tag{13a}$$

$$I_{A_0} \triangleq \{0\} \cup \{i \in \{1, \dots, j\} : \bar{\phi}_i = 0\}, \quad j > 0. \tag{13b}$$

Finally, it is useful to introduce a Lagrangian function

$$\mathcal{L} : X \times X \times \mathbb{R}^{1+j+k} \rightarrow \mathbb{R},$$

defined by

$$\mathcal{L}(x_1, x_2, l) \triangleq \sum_{i=0}^j l_i \phi_i(x_1, x_2) + l_\psi^T \psi(x_1, x_2),$$

$$l = (l_0, \dots, l_j, l_\psi), \quad k > 0, \tag{14a}$$

$$\mathcal{L}(x_1, x_2, l) \triangleq \sum_{i=0}^j l_i \phi_i(x_1, x_2), \quad l = (l_0, \dots, l_j), \quad k = 0, \quad (14b)$$

and a Hamiltonian function

$$\mathcal{H}: T \times X \times U \times \mathbb{R}^n \rightarrow \mathbb{R},$$

defined by

$$\mathcal{H}(t, x, u, p) \triangleq p^T f(t, x, u). \quad (15)$$

**Theorem 3.1.** (i) If  $(\bar{x}(\cdot), \bar{u}(\cdot))$  solves OCP and Assumption 2.1 is satisfied, then there exist  $l \in \mathbb{R}^{1+j+k}$  and  $p(\cdot) \in \tilde{\mathcal{H}}$ , such that

$$l \neq 0, \quad (16)$$

$$l_i \geq 0, \quad i \in \{0, \dots, j\}, \quad (17)$$

$$l_i = 0, \quad i \in \{0, \dots, j\} / I_{A0}, \quad (18)$$

$$p(t_1) = -\bar{\mathcal{L}}_{x_1}(l)^T, \quad p(t_2) = \bar{\mathcal{L}}_{x_2}(l)^T, \quad (19)$$

$$\dot{p}(t) = -\bar{\mathcal{H}}_x(t, p(t))^T = -\bar{f}_x(t)^T p(t), \quad \text{a.a. } t \in T, \quad (20)$$

$$J_1(\omega(\cdot), l) \geq 0, \quad \text{for all } \omega(\cdot) \in \tilde{\mathcal{U}}, \quad (21)$$

where

$$J_1(\omega(\cdot), l) \triangleq \int_{t_1}^{t_2} \bar{\mathcal{H}}_u(t, p(t)) \omega(t) dt. \quad (22)$$

(ii) If, in addition, Assumption 2.2 is satisfied, then for each  $(\tilde{\eta}(\cdot), \tilde{\omega}(\cdot)) \in \mathcal{D}$ , there exist  $l \in \mathbb{R}^{1+j+k}$  and  $p(\cdot) \in \tilde{\mathcal{H}}$ , such that (16)–(21) are satisfied and

$$J_2(\tilde{\eta}(\cdot), \tilde{\omega}(\cdot), l) \geq 0, \quad (23)$$

where

$$\begin{aligned} J_2(\eta(\cdot), \omega(\cdot), l) &\triangleq \bar{\mathcal{L}}_{x_1 x_1}(l)(\eta(t_1))^2 + 2\bar{\mathcal{L}}_{x_1 x_2}(l)(\eta(t_1), \eta(t_2)) \\ &+ \bar{\mathcal{L}}_{x_2 x_2}(l)(\eta(t_2))^2 + \int_{t_1}^{t_2} [\bar{\mathcal{H}}_{xx}(t, p(t))(\eta(t))^2 \\ &+ 2\bar{\mathcal{H}}_{xu}(t, p(t))(\eta(t), \omega(t)) + \bar{\mathcal{H}}_{uu}(t, p(t))(\omega(t))^2] dt. \end{aligned} \quad (24)$$

Our second-order condition is less complex than the one in Ref. 6. First, as in Ref. 5,  $J_2$  is written in terms of the Hamiltonian function instead of a second-variation differential equation of order  $n$  (see Form 1.2 in Ref. 6). Second, a multiplier-dependent condition, which in our notation is equivalent to

$$\bar{\mathcal{H}}_u(t, p(t))\omega(t) = 0, \quad \text{a.a. } t \in T,$$

does not appear. As will be seen, its role is served by (10) with  $i = 0$ . Lastly, and this simplification holds with respect to Ref. 5 as well, there are no difficult-to-verify conditions related to normality. Additional assumptions do produce simplifications in statement (ii). This will be discussed shortly.

Let  $S$  be a subset of a vector space, and define

$$\text{cone } S \triangleq \{\alpha x : x \in S, \alpha > 0\}.$$

Then, the following remark is easily verified.

**Remark 3.1.** If  $\tilde{\mathcal{U}}$  is replaced by cone  $\tilde{\mathcal{U}}$  in (9) and (21), Theorem 3.1 remains valid.

Remark 3.1 does not strengthen the necessary conditions, but it does make them positively homogeneous in the variations  $(\eta(\cdot), \omega(\cdot))$ . This is consistent with the way second-order necessary conditions are usually written in the mathematical programming literature. The remark may make Theorem 3.1 easier to apply. For example, when

$$\mathcal{U} = \text{PC}(T, U)$$

and

$$\tilde{\mathcal{U}} = \text{PC}(T, U) - \bar{u}(\cdot),$$

as in Ref. 5,

$$\text{cone } \tilde{\mathcal{U}} = \text{PC}(T, \mathbb{R}^m).$$

Also, cone  $\tilde{\mathcal{U}}$  may be viewed as a conical approximation to  $\mathcal{U}$  at  $\bar{u}(\cdot)$  (Refs. 16 and 18). In Theorem 3.1 (i), the remark may be extended so that  $\tilde{\mathcal{U}}$  is replaced by the closure of cone  $\tilde{\mathcal{U}}$ . However, simple examples show that this extension does not apply to (ii).

Several other points deserve attention. For each  $l$ , (19) and (20) assure that  $p(\cdot)$  is uniquely determined. This fact justifies writing  $J_1$  and  $J_2$  as functions of  $l$ , rather than of  $l$  and  $p(\cdot)$ . Also, it implies that the family of pairs  $(l, p(\cdot))$  which satisfy the necessary conditions [either (i) or (ii)] is completely characterized by the corresponding family of multipliers  $l$ . The family of multipliers  $l$  which satisfy (ii) is *dependent* on  $(\tilde{\eta}(\cdot), \tilde{\omega}(\cdot)) \in \mathcal{D}$ . Unlike Refs. 5 and 6, we do not know if there exists an  $l$  and corresponding  $p(\cdot)$  such that (23) is satisfied for all  $(\tilde{\eta}(\cdot), \tilde{\omega}(\cdot)) \in \mathcal{D}$ . Consequently, (ii) requires some modifications if it is to be cast as an accessory minimum problem (Refs. 1–3 and 5).

Let the normalized multipliers associated with the first-order conditions be given by

$$\Omega \triangleq \{l : l \text{ satisfies conditions (16)–(21) and } |l| = 1\}. \tag{25}$$

From Remark 4.1 and the connections established in Section 5, it follows that  $\Omega$  is compact. This and the continuity of  $J_2$  in  $l$  yield the following obvious paraphrasing of (ii).

**Theorem 3.2.** Suppose that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  solves OCP and that Assumptions 2.1 and 2.2 are satisfied. Then,

$$\hat{J}_2(\eta(\cdot), \omega(\cdot)) \triangleq \max_{l \in \Omega} J_2(\eta(\cdot), \omega(\cdot), l) \quad (26)$$

is defined for all  $(\eta(\cdot), \omega(\cdot)) \in \mathcal{D}$ , and

$$(\eta(\cdot), \omega(\cdot)) = 0$$

solves the *accessory minimum problem*: Find  $(\eta(\cdot), \omega(\cdot)) \in \mathcal{D}$  which minimizes  $\hat{J}_2(\eta(\cdot), \omega(\cdot))$ .

Second-order necessary conditions having this min-max structure have appeared recently in the mathematical programming literature Refs. 14 and 15, but are new in the theory of optimal control. The min-max structure disappears if additional assumptions, similar to those used in Refs. 5 and 6, are imposed.

Let us pursue this issue in greater detail. If

$$l_0 > 0, \quad \text{for some } l \in \Omega,$$

we say that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  is *normal*. This is a weaker definition of normality than the more common one:

$$l_0 > 0, \quad \text{for all } l \in \Omega.$$

If  $\Omega$  has a single element  $\bar{l}$ , we say that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  is *regular*.

**Remark 3.2.** If  $(\bar{x}(\cdot), \bar{u}(\cdot))$  is a regular solution of OCP, the conclusion of Theorem 3.1(ii) is strengthened as follows: there exists  $l \in \mathbb{R}^{1+j+k}$  and  $p(\cdot) \in \mathcal{X}$ , such that (16)–(21) are satisfied and

$$J_2(\eta(\cdot), \omega(\cdot), l) \geq 0, \quad \text{for all } (\eta(\cdot), \omega(\cdot)) \in \mathcal{D}. \quad (27)$$

Moreover,

$$\hat{J}_2(\eta(\cdot), \omega(\cdot)) = J_2(\eta(\cdot), \omega(\cdot), \bar{l}),$$

and the accessory minimum problem loses its min-max structure. If  $(\bar{x}(\cdot), \bar{u}(\cdot))$  is both a regular and normal solution of OCP, there is a unique  $l$ ,  $l_0 = 1$ , such that (27) holds. This consequence occurs in both Refs. 5 and 6; it is noted in Ref. 5, but not in Ref. 6.

Normality also allows the set  $\mathcal{D}$  in Theorem 3.1 to be expressed in alternative ways. Use

$$I_A \triangleq I_{A0}/\{0\} \tag{28}$$

to denote the set of active inequality constraints. Let

$$\mathcal{D}_\delta \triangleq \{(\eta(\cdot), \omega(\cdot)) \in \tilde{\mathcal{X}} \times \tilde{\mathcal{U}} : (\eta(\cdot), \omega(\cdot)) \text{ satisfies } (\delta), (11), (12)\}, \tag{29a}$$

where  $\delta = a, b, c$  is given by

$$(a) \quad \bar{\phi}_{0x_1}\eta(t_1) + \bar{\phi}_{0x_2}\eta(t_2) = 0, \quad \bar{\phi}_{ix_1}\eta(t_1) + \bar{\phi}_{ix_2}\eta(t_2) \leq 0, \quad i \in I_A; \tag{29b}$$

$$(b) \quad J_1(\omega(\cdot), l) = 0; \tag{29c}$$

$$(c) \quad \bar{\phi}_{ix_1}\eta(t_1) + \bar{\phi}_{ix_2}\eta(t_2) \leq 0, \quad i \in \{p \in I_A : l_p = 0\}, \tag{29d}$$

$$\bar{\phi}_{ix_1}\eta(t_1) + \bar{\phi}_{ix_2}\eta(t_2) = 0, \quad i \in \{p \in I_A : l_p > 0\}. \tag{29e}$$

If  $\tilde{u} \in \tilde{\mathcal{U}}$  implies  $-\alpha\tilde{u} \in \tilde{\mathcal{U}}$  for some  $\alpha > 0$ , we say that  $\tilde{\mathcal{U}}$  is *two-sided*.

**Theorem 3.3.** Suppose that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  is a normal solution of OCP and Assumption 2.1 holds.

(i) Then,  $\mathcal{D}_a = \mathcal{D}$ .

(ii) Assume that  $I_A = \emptyset$ , and let  $l, l_0 > 0$ , satisfy Theorem 3.1(i). Then,  $\mathcal{D}_b = \mathcal{D}$ .

(iii) Suppose that  $\tilde{\mathcal{U}}$  is two-sided, and let  $l, l_0 > 0$ , satisfy Theorem 3.1(i). Then,  $\mathcal{D}_c = \mathcal{D}$ .

This theorem is useful in the application of the second-order necessary conditions, because certain inequality conditions in the definition of  $\mathcal{D}$  are replaced by more easily treated equality conditions. It also establishes connections with Refs. 5 and 6. When OCP is specialized to Warga's problem, Theorem 3.3(ii) applies, because there are no inequality constraints. His second-order condition is essentially expressed in terms of  $\mathcal{D}_b$  [the condition  $\tilde{\mathcal{H}}_u(t, p(t))\omega(t) = 0$ , a.a.  $t \in T$ , implies (b)]. When OCP is specialized to Hestenes' problem, Theorem 3.3(iii) applies, because

$$\tilde{\mathcal{U}} = \text{PC}(T, U) - \bar{u}(\cdot)$$

is two-sided. His second-order condition is expressed in terms of  $\mathcal{D}_c$  (see Theorem 9.1, Chapter 6, Ref. 5). Note that (c) does not include  $\bar{\phi}_{0x_1}$  and  $\bar{\phi}_{0x_2}$ . This is the traditional way of writing the second-order necessary conditions (Refs. 2-4).

Since normality and regularity have significant implications, it is of interest to determine conditions which imply them. In seeking such conditions, we have more freedom than prior researchers, because (unlike them) we have separated the issues of normality and regularity from the proof

of the necessary conditions. In fact, some simple observations yield conditions which are equivalent to regularity and normality. After stating these conditions, we will introduce stronger conditions and briefly explore their relationship to the conditions in Refs. 5 and 6.

We need to use the derivatives  $\bar{\psi}_{x_1}, \bar{\psi}_{x_2}, \bar{\phi}_{ix_1}, \bar{\phi}_{ix_2}, i \in I$ , for several index sets  $I \subset \{0, \dots, j\}$ . To ease the notation, let  $i_1 < i_2 < \dots < i_s$  be the elements of  $I$  when  $I \neq \emptyset$ , and let  $s = 0$  when  $I = \emptyset$ . If  $s + k > 0$ , introduce

$$\Phi^I : X \times X \rightarrow \mathbb{R}^{s+k}$$

by

$$\Phi^I \triangleq (\phi_{i_1}, \dots, \phi_{i_s}, \psi), \quad \text{for } I \neq \emptyset, k > 0, \tag{30a}$$

$$\Phi^I \triangleq (\phi_{i_1}, \dots, \phi_{i_s}), \quad \text{for } I \neq \emptyset, k = 0, \tag{30b}$$

$$\Phi^I \triangleq \psi, \quad \text{for } I = \emptyset, k > 0. \tag{30c}$$

Now, define

$$\Gamma^I : \mathbb{R}^n \times \tilde{\mathcal{U}} \rightarrow \mathbb{R}^{s+k}$$

by

$$\Gamma^I(\xi, \omega(\cdot)) \triangleq \bar{\Phi}_{x_1}^I \xi + \bar{\Phi}_{x_2}^I \eta(t_2), \tag{31a}$$

$$\eta(\cdot) \text{ solves (12) with } \eta(t_1) = \xi. \tag{31b}$$

From Section 5, it follows that, for  $I = \{0, \dots, j\}$ , the condition

$$l^T \Gamma^I(\xi, \omega(\cdot)) \geq 0, \quad \text{for all } (\xi, \omega(\cdot)) \in \mathbb{R}^n \times \tilde{\mathcal{U}}, \tag{32}$$

is equivalent to (19)–(21). Moreover,  $|l| = 1$ , (17), and (18) are equivalent to

$$|l| = 1, \tag{33a}$$

$$l_i \geq 0, \quad \text{for } i \in I_{A0}, \tag{33b}$$

$$l_i = 0, \quad \text{for } i \in \{0, \dots, j\} / I_{A0}. \tag{33c}$$

Thus,  $\Omega$  is the set of  $l$  satisfying both (32) and (33). We may omit from our consideration the components of  $l$  which must automatically be zero. Let

$$\nu \triangleq \text{cardinality of } I_A.$$

Then, normality and regularity can be characterized by the system

$$\rho = (\rho_0, \dots, \rho_{\nu+k}), \quad \rho_i \geq 0, i \in \{0, \dots, \nu\}, \tag{34}$$

$$|\rho| = 1, \tag{35}$$

$$\rho^T \Gamma^{I_{A0}}(\xi, \omega(\cdot)) \geq 0, \quad \text{for all } (\xi, \omega(\cdot)) \in \mathbb{R}^n \times \tilde{\mathcal{U}}. \tag{36}$$

Specifically, we can make the following statement.

**Remark 3.3.** Suppose that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  solves OCP. Then, (34)–(36) has a solution  $\rho$ . The pair  $(\bar{x}(\cdot), \bar{u}(\cdot))$  is regular if and only if the following condition holds:

*Condition (C1).* There is a unique  $\rho$  which satisfies (34)–(36).

The pair  $(\bar{x}(\cdot), \bar{u}(\cdot))$  is normal if and only if the following condition holds:

*Condition (C2).* There is a  $\rho$  which satisfies (34)–(36) and  $\rho_0 > 0$ .

A variety of conditions exist which imply Conditions (C1) and (C2). A trivial example occurs when there are no equality constraints or active inequality constraints,

$$\nu + k = 0.$$

Hereafter, we assume

$$\nu + k \geq 1.$$

The following remark is easily verified.

**Remark 3.4.** Suppose that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  solves OCP. Then, there exists a  $\rho$  which satisfies (35), (36), and  $\rho_0 \geq 0$ . Conditions (C1) and (C2) are implied by the following condition:

*Condition (C3).* There is a unique  $\rho$  which satisfies (35), (36), and  $\rho_0 > 0$ .

If  $\nu = 0$ , condition C3 is equivalent to the joint satisfaction of Conditions (C1) and (C2).

Condition (C3) is closely related to traditional *normality* conditions. The connection with Ref. 5 is made by the following theorem.

**Theorem 3.4.** If cone  $\tilde{\mathcal{U}}$  is a linear space, then Condition (C3) is equivalent to the following condition:

*Condition (C4).*  $\Gamma^{I_A}(\mathbb{R}^n, \text{cone } \tilde{\mathcal{U}}) = \mathbb{R}^{\nu+k}$ .

If cone  $\tilde{\mathcal{U}} = L_\infty(T, \mathbb{R}^m)$ , Condition (C3) is equivalent to the following condition:

*Condition (C5).*  $\text{rank}[(\bar{\Phi}_{x_1}^{I_A} + \bar{\Phi}_{x_2}^{I_A} Z(t_2)) \bar{\Phi}_{x_2}^{I_A} W] = \nu + k$ , where  $Z(t_2)$  is determined by

$$\dot{Z}(t) = \bar{f}_x(t)Z(t), \quad \text{a.a. } t \in T, \tag{37}$$

$$Z(t_1) = I_n \triangleq n \times n \text{ identity matrix}, \tag{38}$$

and  $W$  is defined by

$$W \triangleq \int_{t_1}^{t_2} Z(t_2)Z^{-1}(t)\bar{f}_u(t)(Z(t_2)Z^{-1}(t)\bar{f}_u(t))^T dt. \tag{39}$$

**Proof.** Note that the determination of  $\rho$  by (35), (36), and  $\rho_0 > 0$  is not changed if  $\tilde{\mathcal{U}}$  is replaced by cone  $\tilde{\mathcal{U}}$ . Since cone  $\tilde{\mathcal{U}}$  is a linear space, so is

$$\mathcal{G} \triangleq \Gamma^{I \wedge_0}(\mathbb{R}^n, \text{cone } \tilde{\mathcal{U}}).$$

Condition (C3) holds if and only if  $\mathcal{G} \subset \mathbb{R}^{\nu+k+1}$  is a hyperplane with normal

$$\bar{\rho} = (\bar{\rho}_0, \dots, \bar{\rho}_{\nu+k}), \quad \bar{\rho}_0 > 0.$$

From this, it is easy to verify the equivalence of Conditions (C3) and (C4). For

$$\eta(t_1) = 0 \quad \text{and} \quad \omega(\cdot) \in L_\infty(T, \mathbb{R}^m),$$

the set of  $\eta(t_2)$  generated by (12) is the range of  $W$  (see, e.g., Ref. 19). The equivalence of Conditions (C4) and (C5) follows from this.

For the problem treated in Ref. 5, Conditions (C4) and (C5) are similar to those appearing in Sections 7 and 8 of Chapter 6. See also Ref. 20. When cone  $\tilde{\mathcal{U}}$  is not a linear space, simple, equivalent characterizations of Condition (C3) are not available. However, there are conditions which imply Condition (C3).

**Theorem 3.5.** Suppose that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  solves OCP. Condition (C3) is implied by the following condition:

*Condition (C6).* There is a  $\rho \in \mathbb{R}^{\nu+k+1}$ ,  $\rho \neq 0$ , such that

$$0 \in \text{int}\{\Gamma^{I \wedge}(\xi, \omega(\cdot)) : \xi \in \mathbb{R}^n, \omega(\cdot) \in \tilde{\mathcal{U}}, \rho^T \Gamma^{I \wedge_0}(\xi, \omega(\cdot)) = 0\}.$$

Condition C6 is implied by the following condition:

*Condition (C7).*  $\text{rank}(\bar{\Phi}_{x_1}^{I \wedge} + \bar{\Phi}_{x_2}^{I \wedge} Z(t_2)) = \nu + k$ .

Though simple, Condition (C7) has limited value, because it often fails to hold. Consider, for instance, the problem of Ref. 6, written in the form OCP. Then,  $x(t_1)$  is specified and

$$\hat{\psi}(x(t_2)) = 0,$$

where

$$\hat{\psi}: X \rightarrow \mathbb{R}^k \quad \text{and} \quad j > 0.$$

Specifying these constraints in terms of  $\psi$  requires

$$k = \hat{k} + n > n.$$

Since the matrix in Condition (C7) has  $n$  columns, Condition (C7) cannot possibly hold. Condition (C6) is the natural generalization of the interior point condition given in Theorem 1.1 of Ref. 6. This can be seen from the equivalence noted in Theorem 3.3(ii) and the fact that

$$J_1(\omega(\cdot), l) = 0$$

is equivalent to

$$\rho^T \Gamma^{I_{A0}}(\xi, \omega(\cdot)) = 0.$$

All of the conclusions of Ref. 6 hold under Condition (C3). Moreover, it is easy to find examples where Condition (C3) holds and Condition (C6) does not. Thus, when it can be verified, Condition (C3) is the preferred condition.

A few additional facts concerning Theorem 3.1 are worth noting.

**Remark 3.5.** For  $(\eta(\cdot), \omega(\cdot)) \in \mathcal{D}$ , define

$$\mathcal{F}(\eta(\cdot), \omega(\cdot)) \triangleq \{i \in I_{A0} : \bar{\phi}_{ix_1} \eta(t_1) + \bar{\phi}_{ix_2} \eta(t_2) < 0\}, \tag{40}$$

$$\mathcal{J} \triangleq \bigcup_{(\eta(\cdot), \omega(\cdot)) \in \mathcal{D}} \mathcal{F}(\eta(\cdot), \omega(\cdot)). \tag{41}$$

If  $l$  satisfies (16)–(21), then it must also satisfy

$$l_i = 0, \quad i \in \mathcal{J}. \tag{42}$$

This remark follows from Remark 4.2, which shows that it is a consequence of conditions corresponding to the conditions in Theorem 3.1(i). While Remark 3.4 adds nothing new to (16)–(21), it is a way of deducing that additional components of  $l$  are zero. In applications, it may be difficult to obtain  $\mathcal{J}$ . However, if  $(\eta(\cdot), \omega(\cdot)) \in \mathcal{D}$  is found such that (12) holds as a strict inequality for  $i = \hat{i}$ , then  $l_{\hat{i}} = 0$ . Such determinations are useful in testing for normality and regularity, because they allow  $I_{A0}$  in (36) to be replaced by a proper subset of  $I_{A0}$ .

In general, (21) may not be replaced by a condition which holds pointwise in  $T$ . One exception of obvious practical value is the following.

**Remark 3.6.** Suppose that

$$\tilde{\mathcal{U}} = \{\omega(\cdot) \in L_\infty(T, \mathbb{R}^m) : \omega(t) \in \tilde{U}(t), t \in T\},$$

where

$$\tilde{U}(t) \triangleq \hat{U} - \hat{u}(t),$$

$\hat{U}$  is a convex subset of  $\mathbb{R}^m$ , and

$$\hat{u}(\cdot) \in L_\infty(T, \hat{U}).$$

Then, the integral condition (21) is equivalent to the pointwise condition:

$$\bar{\mathcal{H}}_u(t, p(t))\nu \geq 0, \quad \text{for all } \nu \in \tilde{U}(t), \text{ a.a. } t \in T. \tag{43}$$

The result (43) follows easily. Write

$$\omega(t) = v(t) - \hat{u}(t).$$

Then, (21) is equivalent to

$$\int_{t_1}^{t_2} \bar{\mathcal{H}}_u(t, p(t))\hat{u}(t) dt \leq \int_{t_1}^{t_2} \bar{\mathcal{H}}_u(t, p(t))v(t) dt, \quad \text{for all } v(\cdot) \in L_\infty(T, \tilde{U}). \tag{44}$$

By a standard argument used in the proof of the maximum principle (see, e.g., Ref. 21, p. 79), Ineq. (44) is equivalent to

$$\bar{\mathcal{H}}_u(t, p(t))\hat{u}(t) \leq \bar{\mathcal{H}}_u(t, p(t))\mu, \quad \mu \in \tilde{U}, \text{ a.a. } t \in T,$$

which is equivalent to (43).

A careful reading of Section 5 shows that Assumptions 2.1 and 2.2 are somewhat stronger than they need to be. In Assumption 2.1, it is sufficient to require the following of  $\phi_i$  and  $\psi$ :  $\phi_i$ , for  $i \in I_{A_0}$ , and  $\psi$  have a Fréchet derivative at  $(\bar{x}(t_1), \bar{x}(t_2))$ , and  $\psi$  is continuous in a neighborhood of  $(\bar{x}(t_1), \bar{x}(t_2))$ . In Assumption 2.2, it is sufficient to require that  $\phi_i$ ,  $i \in I_{A_0}$ , and  $\psi$  are twice Fréchet differentiable at  $(\bar{x}(t_1), \bar{x}(t_2))$ .

#### 4. Necessary Conditions for an Abstract Optimization Problem

We now state necessary conditions for an abstract optimization problem. These conditions, which are an obvious specialization of results proved and discussed further in Ref. 17, are the basis for the proof of Theorem 3.1.

Let  $j$  and  $k$  be nonnegative integers,  $\mathcal{E}$  be a vector space,

$$\begin{aligned} E_0 &\subset E \subset \mathcal{E}, \\ \hat{\phi}_i &: E \rightarrow \mathbb{R}, \quad \text{for } i \in \{0, \dots, j\}, \\ \hat{\psi} &: E \rightarrow \mathbb{R}^k, \quad \text{if } k > 0. \end{aligned}$$

**Abstract Optimization Problem (AOP).** Find  $e \in E$  which minimizes

$$\hat{J}(e) \triangleq \hat{\phi}_0(e), \tag{45}$$

subject to

$$e \in E_0, \tag{46}$$

$$\hat{\phi}_i(e) \leq 0, \quad i \in \{1, \dots, j\} \quad (\text{omit if } j = 0), \quad (47)$$

$$\hat{\psi}(e) = 0 \quad (\text{omit if } k = 0). \quad (48)$$

To state assumptions and necessary conditions for AOP, it is necessary to generalize the notion of the Fréchet derivative, so that it applies at boundary points of a function's domain (see Ref. 22, p. 167). The mapping

$$g: A \subset \mathbb{R}^r \rightarrow \mathbb{R}^s$$

is  $\tilde{F}$ -differentiable at  $\bar{a} \in A$  if  $\bar{a}$  is contained in a convex subset of  $A$  having nonempty interior and there exists an  $r \times s$  matrix  $g'(\bar{a})$  satisfying

$$\lim_{\substack{a \rightarrow \bar{a} \\ a \in A \setminus \{\bar{a}\}}} |a - \bar{a}|^{-1} [g(a) - g(\bar{a}) - g'(\bar{a})(a - \bar{a})] = 0. \quad (49)$$

The second-order  $\tilde{F}$ -derivative is defined by application of the above definition to the map  $g': A \rightarrow \mathbb{R}^s$ .

One-sided directional differentials are obtained as a simple application of the  $\tilde{F}$ -derivative. Let  $\mathcal{W}$  be a vector space,

$$W \subset \mathcal{W}, \quad G: W \rightarrow \mathbb{R}^s, \quad \bar{w} \in W, \quad h \in \mathcal{W},$$

and assume that there exists  $\bar{\alpha} > 0$  such that

$$\bar{w} + \alpha h \in W, \quad \alpha \in [0, \bar{\alpha}).$$

The  $\tilde{F}$ -derivative at  $\alpha = 0$  (if it exists) of the map

$$\alpha \rightarrow G(\bar{w} + \alpha h): [0, \bar{\alpha}) \rightarrow \mathbb{R}^s$$

is  $DG(\bar{w}; h)$ , the one-sided directional differential of  $G$  at  $\bar{w}$  in the direction  $h$ . The second-order  $\tilde{F}$ -derivative of the map

$$\alpha \rightarrow G(\bar{w} + \alpha h): [0, \bar{\alpha}) \rightarrow \mathbb{R}^s$$

at  $\alpha = 0$  (if it exists) is  $D^2G(\bar{w}; h)$ , the second-order one-sided directional differential of  $G$  at  $\bar{w}$  in the direction  $h$ .

Elements of  $\mathbb{R}^r$  are written as

$$\alpha = (\alpha_1, \dots, \alpha_r).$$

Define

$$\mathbb{P}^r(\sigma) \triangleq \left\{ \alpha \in \mathbb{R}^r : \alpha_i \geq 0, i = 1, \dots, r, \sum_{i=1}^r \alpha_i < \sigma \right\},$$

$$\Delta^r \triangleq \left\{ \alpha \in \mathbb{R}^r : \alpha_i \geq 0, i = 1, \dots, r, \sum_{i=1}^r \alpha_i = 1 \right\},$$

and let  $\text{co } W$  be the convex hull of  $W$ .

In the following assumptions, let  $\bar{e} \in E$  satisfy (46)–(48), and let  $M \subset \mathcal{E}$  be a convex set such that

$$0 \in M \quad \text{and} \quad E \cap (M + \bar{e}) \subset E_0.$$

Define

$$\hat{\Phi} \triangleq (\hat{\phi}_0, \dots, \hat{\phi}_j), \quad k = 0, \tag{50a}$$

$$\hat{\Phi} \triangleq (\hat{\phi}_0, \dots, \hat{\phi}_j, \hat{\psi}), \quad k > 0. \tag{50b}$$

**Assumption 4.1.** For each

$$\tilde{E} \triangleq \text{co}\{h_1, \dots, h_{k+1}\} \subset M,$$

there exists  $\bar{\sigma} > 0$  such that

$$\bar{e} + \sigma \tilde{E} \subset E, \quad \text{for all } \sigma \in [0, \bar{\sigma}), \tag{51}$$

$$\alpha \rightarrow \hat{\psi} \left( \bar{e} + \sigma \sum_{i=1}^{k+1} \alpha_i h_i \right) : \Delta^k \rightarrow \mathbb{R}^k \text{ is continuous for all } \sigma \in (0, \bar{\sigma}) \quad (\text{omit if } k = 0), \tag{52}$$

$$\alpha \rightarrow \hat{\Phi} \left( \bar{e} + \sum_{i=1}^{k+1} \alpha_i h_i \right) : \mathbb{P}^{k+1}(\bar{\sigma}) \rightarrow \mathbb{R}^{1+j+k} \text{ is } \tilde{F}\text{-differentiable at } \alpha = 0, \tag{53}$$

$$h \rightarrow D\phi_i(\bar{e}; h) : M \rightarrow \mathbb{R} \text{ is convex, } i \in \{0, \dots, j\}, \quad \text{if } k = 0. \tag{54}$$

**Assumption 4.2.** For each

$$\tilde{E} \triangleq \text{co}\{h_1, \dots, h_{k+2}\} \subset M,$$

there exists  $\bar{\sigma} > 0$  such that (51) is satisfied and

$$\alpha \rightarrow \hat{\Phi} \left( \bar{e} + \sum_{i=1}^{k+2} \alpha_i h_i \right) : \mathbb{P}^{k+2}(\bar{\sigma}) \rightarrow \mathbb{R}^{1+j+k} \text{ is twice } \tilde{F}\text{-differentiable at } \alpha = 0. \tag{55}$$

Finally, we require the following notation:

$$\hat{\mathcal{D}} \triangleq \{y \in M : y \text{ satisfies (57) and (58)}\}, \tag{56}$$

where

$$D\hat{\phi}_i(\bar{e}; y) \leq 0, \quad i \in \hat{I}_{A_0}, \tag{57}$$

$$D\hat{\psi}(\bar{e}; y) = 0 \quad (\text{omit if } k = 0), \tag{58}$$

and

$$\hat{I}_{A_0} \triangleq \{0\}, \quad j = 0, \tag{59a}$$

$$\hat{I}_{A_0} \triangleq \{0\} \cup \{i \in \{1, \dots, j\} : \hat{\phi}_i(\bar{e}) = 0\}, \quad j > 0. \tag{59b}$$

**Theorem 4.1.** (i) Suppose that  $\bar{e}$  solves AOP and that Assumption 4.1 is satisfied. Then, there exists

$$l \triangleq (l_0, \dots, l_{j+k}) \in \mathbb{R}^{1+j+k},$$

such that

$$l \neq 0, \tag{60}$$

$$l_i \geq 0, \quad i \in \{0, \dots, j\}, \tag{61}$$

$$l_i = 0, \quad i \in \{0, \dots, j\} / \hat{I}_{A0}, \tag{62}$$

$$l^T D \hat{\Phi}(\bar{e}; h) \geq 0, \quad h \in M. \tag{63}$$

(ii) If, in addition, Assumption 4.2 is satisfied, then, for each  $y \in \hat{\mathcal{D}}$ , there exists

$$l \triangleq (l_0, \dots, l_{j+k}) \in \mathbb{R}^{1+j+k}$$

satisfying (60)–(63) and

$$l^T D^2 \hat{\Phi}(\bar{e}; y) \geq 0. \tag{64}$$

**Remark 4.1.** The set

$$\{l : l \text{ satisfies (63)}\}$$

is closed, because it is the dual cone (Ref. 21, p. 34) of  $D \hat{\Phi}(\bar{e}; M)$ . This and the form of (61) and (62) imply that

$$\hat{\Omega} \triangleq \{l : l \text{ satisfies (61)–(63) and } |l| = 1\}$$

is compact.

**Remark 4.2.** Let  $l$  satisfy Theorem 4.1(i). Suppose that  $y \in \mathcal{D}$  and  $i \in \hat{I}_{A0}$  are such that

$$D \hat{\phi}_i(\bar{e}; y) < 0.$$

Putting  $h = y$  in (63) shows that

$$l_i = 0.$$

More generally,

$$l_i = 0, \quad i \in \hat{\mathcal{J}}, \tag{65}$$

where

$$\hat{\mathcal{J}} \triangleq \bigcup_{y \in \mathcal{D}} \{i \in \hat{I}_{A0} : D \hat{\phi}_i(\bar{e}; y) < 0\}. \tag{66}$$

**Remark 4.3.** Suppose that (53) holds with  $k = 1$ . Then, for  $h_1, h_2, \alpha_1 \geq 0, \alpha_2 \geq 0$ , it is easy to see that

$$D\hat{\Phi}(\bar{e}; \alpha_1 h_1 + \alpha_2 h_2) = \alpha_1 D\hat{\Phi}(\bar{e}; h_1) + \alpha_2 D\hat{\Phi}(\bar{e}; h_2).$$

Thus, (54) holds.

We now consider the abstract counterpart of Theorem 3.3. Let

$$\hat{\mathcal{D}}_{\hat{\sigma}} \triangleq \{y \in M : y \text{ satisfies } (\hat{\sigma}) \text{ and (58)}\}, \tag{67a}$$

where  $\hat{\sigma} = \hat{a}, \hat{b}, \hat{c}$  is given by:

$$(\hat{a}) \quad D\hat{\phi}_0(\bar{e}; y) = 0, \quad D\hat{\phi}_i(\bar{e}; y) \leq 0, \quad i \in \hat{I}_A, \tag{67b}$$

$$(\hat{b}) \quad l^T D\hat{\Phi}(\bar{e}; y) = 0, \tag{67c}$$

$$(\hat{c}) \quad D\hat{\phi}_i(\bar{e}; y) \leq 0, \quad i \in \{p \in \hat{I}_A : l_p = 0\}, \tag{67d}$$

$$D\hat{\phi}_i(\bar{e}; y) = 0, \quad i \in \{p \in \hat{I}_A : l_p > 0\}, \tag{67e}$$

and

$$\hat{I}_A \triangleq \hat{I}_{A0} / \{0\}. \tag{68}$$

**Theorem 4.2.** Suppose that Assumption 4.1 is satisfied, that  $\bar{e}$  solves AOP, and that there exists  $l$  satisfying (60)–(63) such that  $l_0 > 0$ .

- (i) Then,  $\hat{\mathcal{D}}_{\hat{a}} = \hat{\mathcal{D}}$ .
- (ii) Assume that  $\hat{I}_A = \emptyset$ , and let  $l, l_0 > 0$ , satisfy Theorem 4.1(i). Then,  $\hat{\mathcal{D}}_{\hat{b}} = \hat{\mathcal{D}}$ .
- (iii) Suppose that  $M$  is two-sided, and let  $l, l_0 > 0$ , satisfy Theorem 4.1(i). Then,  $\hat{\mathcal{D}}_{\hat{c}} = \hat{\mathcal{D}}$ .

**Proof.** (i) Clearly,

$$\hat{\mathcal{D}}_{\hat{a}} \subset \hat{\mathcal{D}}.$$

Conversely, let  $y \in \hat{\mathcal{D}}$ , and suppose that

$$D\hat{\phi}(\bar{e}; y) < 0.$$

If  $l$  satisfies (60)–(63), then by Remark 4.2  $l_0$  must be zero. Thus,

$$D\hat{\phi}_0(\bar{e}; y) = 0.$$

- (ii) If  $y \in \hat{\mathcal{D}}_{\hat{b}}$ , then  $(\hat{b})$  and  $l_0 > 0$  imply that

$$D\hat{\phi}_0(\bar{e}; y) = 0.$$

Since  $y$  satisfies (58) and  $\hat{I}_A = \emptyset$ ,

$$y \in \hat{\mathcal{D}}.$$

Conversely, if  $y \in \hat{\mathcal{D}}$ , then by (i),

$$D\hat{\phi}_0(\bar{e}; y) = 0.$$

Since  $\hat{I}_A = \emptyset$  and  $y$  satisfies (58),  $(\hat{b})$  must hold.

(iii) Since  $M$  is two-sided, it is easy to see that (63) is equivalent to

$$l^T D\hat{\Phi}(\bar{e}; y) = 0, \quad y \in M. \tag{69}$$

If  $y \in \hat{\mathcal{D}}_{\hat{e}}$ , then it follows from (69) and  $l_0 > 0$  that

$$D\hat{\phi}_0(\bar{e}; y) = 0.$$

This and  $(\hat{c})$  imply that

$$\hat{\mathcal{D}}_{\hat{e}} \subset \hat{\mathcal{D}}_{\hat{a}}.$$

Thus,

$$y \in \hat{\mathcal{D}}.$$

Conversely, if  $y \in \hat{\mathcal{D}}$ , then (58) and (67d) must hold. By Remark 4.2,

$$l_i = 0, \quad \text{for all } i \in \hat{I}_A \text{ such that } D\hat{\phi}_i(\bar{e}; y) < 0.$$

This last statement is equivalent to (67e), which completes the proof.

### 5. Proofs for Section 3

Our principal task is to prove Theorem 3.1 using Theorem 4.1. The correspondence between AOP and OCP is established in the following way. Let

$$\mathcal{E} \triangleq \mathbb{R}^n \times M(T, \mathbb{R}^m), \tag{70}$$

$$E \triangleq \{(x_1, u(\cdot)) \in X \times M(T, U) : x(\cdot) \in \mathcal{E} \text{ exists which satisfies (4) and } x(t_1) = x_1\}, \tag{71}$$

$$E_0 \triangleq E \cap (\mathbb{R}^n \times \mathcal{U}), \tag{72}$$

$$\hat{\Phi}(x_1, u(\cdot)) \triangleq \Phi(x_1, x(x_1, u(\cdot))(t_1)), \tag{73}$$

where

$$\Phi \triangleq (\phi_0, \dots, \phi_j), \quad k = 0, \tag{74a}$$

$$\Phi \triangleq (\phi_0, \dots, \phi_p, \psi), \quad k > 0, \tag{74b}$$

and  $x(x_1, u(\cdot))(\cdot)$  denotes the unique solution of (4) satisfying  $x(t_1) = x_1$ . Let

$$\bar{e} \triangleq (\bar{x}(t_1), \bar{u}(\cdot)),$$

where  $(\bar{x}(\cdot), \bar{u}(\cdot))$  solves OCP, and define  $M$  by

$$M \triangleq \mathbb{R}^n \times \tilde{U}. \tag{75}$$

From (6)–(8), it follows that  $M$  is convex,

$$0 \in M \quad \text{and} \quad E \cap (M + \bar{e}) \subset E_0.$$

Using (70)–(73) and (75), we now show that Assumption 2.1 implies Assumption 4.1. Let

$$\tilde{E} \triangleq \text{co}\{h_1, \dots, h_{k+1}\} \subset M,$$

where

$$h_i \triangleq (\xi_i, \omega_i(\cdot)) \in \mathbb{R}^n \times \tilde{U}, \quad i \in \{1, \dots, k+1\}.$$

Conditions (6)–(8) imply that

$$\bar{u}(t) + \sum_{i=1}^{k+1} \alpha_i \omega_i(t) \in U, \quad t \in T, \alpha \in \mathbb{P}^{k+1}(1).$$

Thus, if  $\bar{\sigma}' \in (0, 1]$  is chosen so that

$$\bar{x}(t_1) + \sum_{i=1}^{k+1} \alpha_i \xi_i \in X, \quad \alpha \in \mathbb{P}^{k+1}(\bar{\sigma}'),$$

we can consider solutions  $x(\cdot, \alpha)$  of the system

$$\dot{x}(t, \alpha) = F(t, x(t, \alpha), \alpha), \quad \text{a.a. } t \in T, \tag{76}$$

$$x(t_1, \alpha) = \bar{x}(t_1) + \sum_{i=1}^{k+1} \alpha_i \xi_i, \tag{77}$$

where  $\alpha \in \mathbb{P}^{k+1}(\bar{\sigma}')$  and

$$F(t, x, \alpha) \triangleq f\left(t, x, \bar{u}(t) + \sum_{i=1}^{k+1} \alpha_i \omega_i(t)\right). \tag{78}$$

We want to show that there is a  $\bar{\sigma} \in (0, \bar{\sigma}']$  such that  $x(\cdot, \alpha) \in \mathcal{X}$  exists for  $\alpha \in \mathbb{P}^{k+1}(\bar{\sigma})$  and that  $x(\cdot, \alpha)$  has an  $\tilde{F}$ -derivative at  $\alpha = 0$ . If, for all  $t \in T$ , the function  $F(t, \cdot, \cdot)$  is  $C^1$  on  $X \times A$ , where  $A \subset \mathbb{R}^{k+1}$  is open and  $0 \in A$ , these results follow immediately from Assumption 2.1 and standard theorems on the parametric dependence of solutions of differential equations. Unfortunately,

$$\bar{u}(t) + \sum_{i=1}^{k+1} \alpha_i \omega_i(t) \in U, \quad t \in T,$$

may imply that there is no such set  $A$ , even though  $U$  is open. Thus, we rely on Theorem II.4.11 of Ref. 22, in which  $F(t, \cdot, \cdot)$  is  $\tilde{F}$ -differentiable

on a set which is not necessarily open. This theorem [with  $F(t, \cdot, \cdot)$  defined on  $X \times \mathbb{P}^{k+1}(\bar{\sigma}')$ ] guarantees our desired existence result, which proves (51).

If  $C(T, X)$  denotes the Banach space of bounded continuous functions  $g: T \rightarrow X$  with the uniform topology, then it also follows from Theorem II.4.11 of Ref. 22 that the map

$$\alpha \rightarrow x(\cdot, \alpha): \mathbb{P}^{k+1}(\bar{\sigma}) \rightarrow C(T, X)$$

has a continuous  $\tilde{F}$ -derivative. This fact and the  $C^1$ -differentiability of  $\Phi$  imply by simple arguments that the mapping

$$\alpha \rightarrow \Phi(x(t_1, \alpha), x(t_2, \alpha)): \mathbb{P}^{k+1}(\bar{\sigma}) \rightarrow \mathbb{R}^{1+j+k} \tag{79}$$

is  $\tilde{F}$ -differentiable at  $\alpha = 0$ , which proves (52). Similarly, the map in (79), with  $\Phi$  replaced by  $\psi$ , is continuous on  $\mathbb{P}^{k+1}(\bar{\sigma})$ , which proves (53). To show (54), note that, if  $k = 0$ , then (53) can still be verified with  $k = 1$  which, by Remark 4.3, implies (54).

To prove Theorem 3.1(i), it remains to be shown that (60)–(63) imply (16)–(21). Clearly, (60)–(62) are equivalent to (16)–(18). By setting

$$\xi_1 = \xi, \quad \omega_1(\cdot) = \omega(\cdot), \quad \eta(\cdot) = x_{\alpha_1}(\cdot, 0),$$

it follows from (76) and (77) that the first-order directional differential of the map

$$(x_1, u(\cdot)) \rightarrow x(x_1, u(\cdot))(\cdot): E \rightarrow C(T, X) \tag{80}$$

at  $(\bar{x}(t_1), \bar{u}(\cdot))$  in the direction  $(\xi, \omega(\cdot)) \in M$  is given by the solution  $\eta(\cdot)$  of (12) with  $\eta(t_1) = \xi$ . From the chain rule for  $\tilde{F}$ -derivatives (see Ref. 22, p. 172), it follows that

$$D\hat{\Phi}(\bar{e}; h) = \bar{\Phi}_{x_1}\eta(t_1) + \bar{\Phi}_{x_2}\eta(t_2), \tag{81}$$

where

$$h \triangleq (\xi, \omega(\cdot)).$$

Thus, using (12) and (81), (63) becomes

$$l^T \Gamma^I(\xi, \omega(\cdot)) \geq 0, \quad (\xi, \omega(\cdot)) \in \mathbb{R}^n \times \tilde{\mathcal{U}}, \tag{82}$$

where

$$\Gamma^I(\xi, \omega(\cdot)) \triangleq (\bar{\Phi}_{x_1} + \bar{\Phi}_{x_2}Z(t_2))\xi + \bar{\Phi}_{x_2}Z(t_2) \int_{t_1}^{t_2} Z^{-1}(t)\bar{f}_u(t)\omega(t) dt, \tag{83}$$

and  $Z(\cdot)$  is given by (37) and (38). Obviously, (82) implies that

$$l^T [\bar{\Phi}_{x_1} + \bar{\Phi}_{x_2}Z(t_2)] = 0. \tag{84}$$

If  $p: T \rightarrow \mathbb{R}^n$  is defined by

$$p^T(t) = l^T \bar{\Phi}_{x_2} Z(t_2) Z^{-1}(t), \tag{85}$$

it is clear that (84) is equivalent to

$$l^T \bar{\Phi}_{x_1} Z^{-1}(t) + p(t) = 0, \quad t \in T.$$

From this, it is easy to verify that  $p(\cdot)$  satisfies (19) and (20). Using (84) and (85) in (82) yields (21).

To prove Theorem 3.1(ii), we apply a second-order generalization of Theorem II.4.11 of Ref. 22 to (76) and (77), with  $k + 1$  replaced by  $k + 2$ . This shows that Assumption 4.2 is satisfied. By setting

$$\xi_1 = \xi, \quad \omega_1(\cdot) = \omega(\cdot), \quad \eta(\cdot) = x_{\alpha_1}(\cdot, 0), \quad \theta(\cdot) = x_{\alpha_1 \alpha_1}(\cdot, 0),$$

it follows from (76) and (77) that the second-order directional differential of the map in (80) at  $(\bar{x}(t_1), \bar{u}(\cdot))$  in the direction  $(\xi, \omega(\cdot)) \in M$  is given by the solution  $\theta(\cdot) \in \mathcal{X}$  of

$$\begin{aligned} \dot{\theta}(t) &= \bar{f}_x(t)\theta(t) + \bar{f}_{xx}(t)(\eta(t))^2 + 2\bar{f}_{xu}(t)(\eta(t), \omega(t)) + \bar{f}_{uu}(t)(\omega(t))^2, \\ &\text{a.a. } t \in T, \end{aligned} \tag{86}$$

$$\theta(t_1) = 0, \tag{87}$$

where  $\eta(\cdot)$  is given by (12) with  $\eta(t_1) = \xi$ . For

$$h \triangleq (\xi, \omega(\cdot)) \in M,$$

we have

$$\begin{aligned} D^2 \hat{\Phi}(\bar{e}; h) &= \bar{\Phi}_{x_1 x_1}(\eta(t_1))^2 + 2\bar{\Phi}_{x_1 x_2}(\eta(t_1), \eta(t_2)) \\ &\quad + \bar{\Phi}_{x_2 x_2}(\eta(t_2))^2 + \bar{\Phi}_{x_2} \theta(t_2). \end{aligned} \tag{88}$$

By using (85)–(87) and (20), we observe that

$$\begin{aligned} l^T \bar{\Phi}_{x_2} \theta(t_2) &= p^T(t_2)\theta(t_2) - p^T(t_1)\theta(t_1) \\ &= \int_{t_1}^{t_2} [ \dot{p}^T(t)\theta(t) + p^T(t)\dot{\theta}(t) ] dt \\ &= \int_{t_1}^{t_2} [ \bar{\mathcal{L}}_{xx}(t, p(t))(\eta(t))^2 + 2\bar{\mathcal{L}}_{xu}(t, p(t))(\eta(t), \omega(t)) \\ &\quad + \bar{\mathcal{L}}_{uu}(t, p(t))(\omega(t))^2 ] dt. \end{aligned} \tag{89}$$

Hence, (88) and (24) show that

$$l^T D^2 \hat{\Phi}(\bar{e}; h) = J_2(\eta(\cdot), \omega(\cdot), l).$$

From (81), (56), and (9), it follows that  $(\xi, \omega(\cdot)) \in \hat{\mathcal{D}}$  is equivalent to

$$(\eta(\cdot), \omega(\cdot)) \in \mathcal{D},$$

with  $\xi = \eta(t_1)$ . Thus, (64) with  $y = (\tilde{\eta}(t_1), \tilde{\omega}(\cdot))$  implies (23), which completes the proof of Theorem 3.1.

The proof of Theorem 3.3 is a consequence of Theorem 4.2 and the correspondence between  $\hat{\mathcal{D}}$  and  $\mathcal{D}$ .

We now consider the proof of Theorem 3.5. To prove that Condition (C6) implies Condition (C3), let

$$\rho = (\rho_0, \dots, \rho_{\nu+k}),$$

and assume without loss of generality that

$$|\rho| = 1.$$

If  $\rho_0 = 0$ , then  $\rho$  satisfies

$$\begin{aligned} &(\rho_1, \dots, \rho_{\nu+k})^T \Gamma^{I\wedge}(\xi, \omega(\cdot)) \geq 0, \\ &(\xi, \omega(\cdot)) \in \Lambda \triangleq \{(\xi, \omega(\cdot)) \in \mathbb{R}^n \times \mathcal{Q}\tilde{U} : \rho^T \Gamma^{I\wedge}(\xi, \omega(\cdot)) = 0\}. \end{aligned}$$

It follows from Condition (C6) that

$$(\rho_1, \dots, \rho_{\nu+k}) = 0,$$

which contradicts  $\rho \neq 0$ . Thus, we assume without loss of generality that  $\rho_0 > 0$ . Now, let

$$\tilde{\rho} = (\tilde{\rho}_0, \dots, \tilde{\rho}_{\nu+k})$$

be any solution of (35) and (36). Since, by Condition (C6),

$$0 \in \text{int } \Gamma^{I\wedge}(\mathbb{R}^n, \mathcal{Q}\tilde{U}),$$

we must have  $\tilde{\rho}_0 > 0$ . This proves existence in Condition (C3). To prove uniqueness, we show that  $\tilde{\rho} = \rho$ . Since  $\tilde{\rho}_0 > 0$ , it can be seen that

$$((\rho_0/\tilde{\rho}_0)\tilde{\rho} - \rho)^T \Gamma^{I\wedge}(\xi, \omega(\cdot)) \geq 0, \quad (\xi, \omega(\cdot)) \in \Lambda.$$

Thus, Condition (C6) implies

$$(\rho_0/\tilde{\rho}_0)(\tilde{\rho}_1, \dots, \tilde{\rho}_{\nu+k}) = (\rho_1, \dots, \rho_{\nu+k}),$$

and hence

$$\tilde{\rho} = (\tilde{\rho}_0/\rho_0)\rho.$$

Since

$$|\rho| = |\tilde{\rho}| = 1 \quad \text{and} \quad \tilde{\rho}_0/\rho_0 > 0,$$

we have that  $\rho = \tilde{\rho}$ , as desired. To prove that Condition (C6) follows from Condition (C7), note that (83) and (84) imply that  $\mathbb{R}^n \times \{0\}$  is a subset of  $\Lambda$ . By Condition (C7),

$$\Gamma^{\lambda}(\mathbb{R}^n \times \{0\}) = \mathbb{R}^{\nu+k},$$

which implies Condition (C6).

## 6. Necessary Conditions for a Relaxed Optimal Control Problem

This section contains first-order and second-order necessary conditions for a relaxed optimal control problem, expressed in a form (Theorem 6.1) analogous to Theorem 3.1. The general approach follows Warga (Ref. 22), in that relaxed controls are represented by probability measures. Since the measures belong to a convex set, the arguments used to obtain Theorem 6.1 are similar to those used in the proof of Theorem 3.1. The relation of Theorem 6.1 to solutions of OCP is also explored. By specializing measures in the necessary conditions to Dirac measures, we obtain, for special cases of OCP, an interesting second-order necessary condition involving ordinary controls, which is unlike the one of Section 3.

First, we state the relaxed optimal control problem. It is based on OCP with  $\mathcal{U} = \hat{\mathcal{U}}$ , where

$$\hat{\mathcal{U}} \triangleq \{u(\cdot) \in L_{\infty}(T, U_0) : u(t) \in \hat{U}(t), t \in T\}, \quad (90)$$

$$\hat{U}(t) \subset U_0, \quad \text{for all } t \in T,$$

$U_0 \subset U$  is compact, and  $\hat{U}(\cdot)$  is a measurable set-valued mapping on  $T$ , in the sense described in Ref. 22, Section 1.7, p. 146. The relaxed controls corresponding to  $\hat{\mathcal{U}}$  are defined as follows. Let  $\text{rpm}(U_0)$  be the set of Radon probability measures on  $U_0$  (see Ref. 22, p. 50), and let

$$f^R : T \times X \times \text{rpm}(U_0) \rightarrow \mathbb{R}^n$$

be given by

$$f^R(t, x, \sigma) \triangleq \int_{U_0} f(t, x, u) \sigma(du). \quad (91)$$

If  $f(t, x, \cdot)$  is continuous on  $U_0$  for all  $(t, x) \in T \times X$ , then  $f^R$  can be regarded as an extension of  $f$ , in the sense that

$$f^R(t, x, \delta_u) = f(t, x, u),$$

where  $\delta_u$  denotes the Dirac measure at  $u$  (see Ref. 22, p. 266). The relaxed

controls are elements of the set

$$\hat{\mathcal{P}} \triangleq \{\sigma(\cdot) : \sigma : T \rightarrow \text{rpm}(U_0), \sigma(t)(\hat{U}(t)) = 1, \text{ a.a. } t \in T\}. \tag{92}$$

**Relaxed Optimal Control Problem (ROCP).** Find  $(x(\cdot), \sigma(\cdot)) \in \mathcal{X} \times \hat{\mathcal{P}}$  which minimizes

$$J^R(x(\cdot), \sigma(\cdot)) \triangleq \phi_0(x(t_1), x(t_2)), \tag{93}$$

subject to (2), (3), and

$$\dot{x}(t) = f^R(t, x(t), \sigma(t)), \quad \text{a.a. } t \in T. \tag{94}$$

This problem is closely related to the relaxed optimal control problem treated in Ref. 7. There, (2) is omitted, (3) is of the form

$$x(t_1) = 0, \quad \hat{\psi}(x(t_2)) = 0,$$

and  $U_0$  is a compact metric space.

To obtain first-order and second-order necessary conditions for ROCP, we require the following assumptions. Note that they are weaker than Assumptions 2.1 and 2.2, in that  $u \in U_0$  and  $f(t, x, u)$  need not be differentiable in  $u$ .

**Assumption 6.1.** For all  $(t, u) \in T \times U_0$ ,  $f(t, \cdot, \cdot)$  is continuous on  $X \times U_0$  and  $f(t, \cdot, u)$  is  $C^1$  on  $X$ ; for all  $(x, u) \in X \times U_0$ , the components of  $f(\cdot, x, u)$  and the elements of  $f_x(\cdot, x, u)$  are measurable on  $T$ ; there exists an integrable function  $\zeta : T \rightarrow \mathbb{R}$  such that, for all  $(t, x, u) \in T \times X \times U_0$ ,

$$\begin{aligned} |f(t, x, u)| + |f_x(t, x, u)| &\leq \zeta(t); \\ \text{for } i \in \{0, \dots, j\}, \phi_i &\text{ is } C^1 \text{ on } X \times X; \\ \text{if } k > 0, \psi &\text{ is } C^1 \text{ on } X \times X. \end{aligned}$$

**Assumption 6.2.** For all  $(t, u) \in T \times U$ ,  $f(t, \cdot, u)$  is  $C^2$  on  $X$ ; for all  $(x, u) \in X \times U_0$ , the elements of the matrices which define  $f_{xx}(\cdot, x, u)$  are measurable on  $T$ ; there exists an integrable function  $\zeta : T \rightarrow \mathbb{R}$  such that, for all  $(t, x, u) \in T \times X \times U_0$ ,

$$\begin{aligned} |f_{xx}(t, x, u)| &\leq \zeta(t); \\ \text{for } i \in \{0, \dots, j\}, \phi_i &\text{ is } C^2 \text{ on } X \times X; \\ \text{if } k > 0, \psi &\text{ is } C^2 \text{ on } X \times X. \end{aligned}$$

As in Section 3, we employ a superbar notation for a solution  $(\bar{x}(\cdot), \bar{\sigma}(\cdot))$  of ROCP. We work directly with the convex set  $\hat{\mathcal{P}}$ , instead of introducing a set of control variations corresponding to  $\hat{\mathcal{U}}$  in Section 3. In the proof of Theorem 6.1,  $\hat{\mathcal{P}} - \bar{\sigma}(\cdot)$  plays the role of  $\hat{\mathcal{U}}$ . The set of

relaxed critical directions, which is the counterpart of  $\mathcal{D}$  in (9), is given by

$$\mathcal{D}^R \triangleq \{(\eta(\cdot), \sigma(\cdot)) \in \tilde{\mathcal{X}} \times \tilde{\mathcal{F}} : (\eta(\cdot), \sigma(\cdot)) \text{ satisfies (10), (11), and (96)}\}, \quad (95)$$

where

$$\dot{\eta}(t) = \bar{f}_x^R(t)\eta(t) + f^R(t, \bar{x}(t), \sigma(t)) - \bar{f}^R(t), \quad \text{a.a. } t \in T. \quad (96)$$

Finally, we introduce a relaxed Hamiltonian function

$$\mathcal{H}^R : T \times X \times \text{rpm}(U_0) \times \mathbb{R}^n \rightarrow \mathbb{R},$$

defined by

$$\mathcal{H}^R(t, x, \sigma, p) \triangleq p^T f^R(t, x, \sigma). \quad (97)$$

The following theorem is proved by arguments exactly parallel to those used in the proof of Theorem 3.1. The additional technical details concerning variations with respect to relaxed controls can be found in Ref. 7.

**Theorem 6.1.** (i) If  $(\bar{x}(\cdot), \bar{\sigma}(\cdot))$  solves ROCP and Assumption 6.1 is satisfied, then there exists  $l \in \mathbb{R}^{1+j+k}$  and  $p(\cdot) \in \tilde{\mathcal{X}}$ , such that (16)–(19) are satisfied and

$$\dot{p}(t) = -\bar{\mathcal{H}}_x^R(t, p(t))^T = -\bar{f}_x^R(t)^T p(t), \quad \text{a.a. } t \in T, \quad (98)$$

$$J_1^R(\sigma(\cdot), l) \geq 0, \quad \text{for all } \sigma(\cdot) \in \tilde{\mathcal{F}}, \quad (99)$$

where

$$J_1^R(\sigma(\cdot), l) \triangleq \int_{t_1}^{t_2} [\mathcal{H}^R(t, \bar{x}(t), \sigma(t), p(t)) - \bar{\mathcal{H}}^R(t, p(t))] dt. \quad (100)$$

(ii) If, in addition, Assumption 6.2 is satisfied, then, for each  $(\tilde{\eta}(\cdot), \tilde{\sigma}(\cdot)) \in \mathcal{D}^R$ , there exists  $l \in \mathbb{R}^{1+j+k}$  and  $p(\cdot) \in \tilde{\mathcal{X}}$ , such that (16)–(19), (98), and (99) are satisfied and

$$J_2^R(\tilde{\eta}(\cdot), \tilde{\sigma}(\cdot), l) \geq 0, \quad (101)$$

where

$$\begin{aligned} J_2^R(\eta(\cdot), \sigma(\cdot), l) &\triangleq \bar{\mathcal{L}}_{x_1 x_1}(l)(\eta(t_1))^2 + 2\bar{\mathcal{L}}_{x_1 x_2}(l)(\eta(t_1), \eta(t_2)) + \bar{\mathcal{L}}_{x_2 x_2}(l)(\eta(t_2))^2 \\ &+ \int_{t_1}^{t_2} \{\bar{\mathcal{H}}_{xx}^R(t, p(t))(\eta(t))^2 \\ &+ 2[\mathcal{H}_x^R(t, \bar{x}(t), \sigma(t), p(t)) - \bar{\mathcal{H}}_x^R(t, p(t))]\eta(t)\} dt. \end{aligned} \quad (102)$$

**Remark 6.1.** Because of the assumptions on  $\hat{U}(\cdot)$ , it can be shown (see Ref. 22, p. 361) that (99) is equivalent to

$$\bar{\mathcal{H}}^R(t, p(t)) = \min_{u \in \hat{U}(t)} \mathcal{H}(t, \bar{x}(t), u, p(t)), \quad \text{a.a. } t \in T. \quad (103)$$

Thus, Theorem 6.1(i) is a pointwise minimum principle for relaxed controls.

Since most applications involve *ordinary controls*, it is of interest to examine situations where Theorem 6.1 can be applied to OCP. The arguments used in the proof of Theorem VI.2.3 of Ref. 22 show that (i) yields the familiar minimum principle.

**Theorem 6.2.** Suppose that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  solves OCP,  $\mathcal{U} = \hat{\mathcal{U}}$ , and Assumption 6.1 is satisfied. Then, there exist  $l \in \mathbb{R}^{1+j+k}$  and  $p(\cdot) \in \tilde{\mathcal{X}}$ , such that (16)–(20) are satisfied and

$$\bar{\mathcal{H}}(t, p(t)) = \min_{u \in \hat{U}(t)} \mathcal{H}(t, \bar{x}(t), u, p(t)), \quad \text{a.a. } t \in T. \quad (104)$$

Extending Theorem 6.1(ii) to solutions of OCP is more complex.

**Theorem 6.3.** Suppose that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  solves OCP,  $\mathcal{U} = \hat{\mathcal{U}}$ , and Assumptions 6.1 and 6.2 are satisfied. Assume that one or both of the following assumptions hold:

*Assumption (A1).*  $f(t, x, \hat{U}(t))$  is convex for all  $(t, x) \in T \times X$ ;

*Assumption (A2).*  $l_0 > 0$ , for all  $l$  which satisfy the minimum principle [(16)–(20) and (104)].

In (96) and (102), make the following substitutions:

$$\bar{f}^R(t) = \bar{f}(t), \bar{f}_x^R(t) = \bar{f}_x(t),$$

$$\bar{\mathcal{H}}_x^R(t, p(t)) = \bar{\mathcal{H}}_x(t, p(t)), \bar{\mathcal{H}}_{xx}^R(t, p(t))(\eta(t))^2 = \bar{\mathcal{H}}_{xx}(t, p(t))(\eta(t))^2.$$

Then, for each  $(\tilde{\eta}(\cdot), \tilde{\sigma}(\cdot)) \in \mathcal{D}^R$ , there exist  $l \in \mathbb{R}^{1+j+k}$  and  $p(\cdot) \in \tilde{\mathcal{X}}$ , such that (16)–(20), (104), and (101) are satisfied.

**Proof.** Assumption (A1) implies that the set of relaxed controls is no richer than the set of ordinary controls (see Ref. 22, Theorem VI.3.2). Thus, if  $(\bar{x}(\cdot), \bar{u}(\cdot))$  is a solution of OCP,  $(\bar{x}(\cdot), \delta_{\bar{u}(\cdot)})$  is a solution of ROCP. Hence, Theorem 6.1 applies to

$$(\bar{x}(\cdot), \tilde{\sigma}(\cdot)) = (\bar{x}(\cdot), \delta_{\bar{u}(\cdot)}).$$

Noting that

$$f^R(t, \cdot, \bar{\sigma}(t)) = f(t, \cdot, \bar{u}(t))$$

gives the desired results. Under Assumption (A2), the arguments used in Ref. 7, p. 306, extend directly to the present situation. They show that  $(\bar{x}(\cdot), \delta_{\bar{u}(\cdot)})$  solves ROCP, with  $\hat{\mathcal{F}}$  replaced by  $\hat{\mathcal{F}} \cap N$ , where  $N$  is a sufficiently small neighborhood of  $\bar{\sigma}(\cdot)$ . The preceding argument then applies, because the necessary conditions for this modified problem coincide with the necessary conditions for ROCP.

Theorems 6.1–6.3 are closely related to Theorem 2.2 of Ref. 7. They simplify and generalize its results in essentially the same way that Theorem 3.1 simplifies and generalizes the results of Ref. 22. Most of the conclusions in Section 3 concerning normality and regularity can be adapted with little change to the present context. For example, suppose that  $(\bar{x}(\cdot), \bar{u}(\cdot))$  solves OCP, that (16)–(20) and (104) are satisfied by only one  $l$ , with  $|l| = 1$ , and that, for this  $l$ ,  $l_0 > 0$ . Then, Assumption (A2) is satisfied, and Theorem 6.3 becomes: there exists  $l \in \mathbb{R}^{1+j+k}$  and  $p(\cdot) \in \mathcal{F}$ , such that (16)–(20) and (104) are satisfied and

$$J_2^R(\eta(\cdot), \sigma(\cdot), l) \geq 0, \quad \text{for all } (\eta(\cdot), \sigma(\cdot)) \in \mathcal{D}^R. \quad (105)$$

This strengthens Theorem 2.2 of Ref. 7, because the normality assumption there is stronger than the above assumption on  $l$ .

In applying Theorem 6.3 to OCP, it may be inconvenient to work with relaxed controls expressed as measures. Warga (Ref. 7), in part (c) of his Theorem 2.2, gives a second-order necessary condition which avoids both ordinary controls and measures. Here, we give a second-order necessary condition which involves ordinary controls. Since it is obtained (trivially) by setting  $\bar{\sigma}(\cdot) = \delta_{\bar{u}(\cdot)}$  in Theorem 6.3, it is weaker than the one in Theorem 6.3.

Let

$$\mathcal{D}^+ \triangleq \{(\eta(\cdot), u(\cdot)) \in \mathcal{F} \times \mathcal{U} : (\eta(\cdot), u(\cdot)) \text{ satisfies (10), (11), (107)}\}, \quad (106)$$

where

$$\dot{\eta}(t) = \bar{f}_x(t)\eta(t) + f(t, \bar{x}(t), u(t)) - \bar{f}(t), \quad \text{a.a. } t \in T. \quad (107)$$

**Theorem 6.4.** Suppose that the hypotheses of Theorem 6.3 are satisfied. Then, for each  $(\tilde{\eta}(\cdot), \tilde{u}(\cdot)) \in \mathcal{D}^+$ , there exist  $l \in \mathbb{R}^{1+j+k}$  and  $p(\cdot) \in \mathcal{F}$ , such that (16)–(20) and (104) are satisfied and

$$J_2^+(\tilde{\eta}(\cdot), \tilde{u}(\cdot), l) \geq 0, \quad (108)$$

where

$$\begin{aligned}
 J_2^+(\eta(\cdot), u(\cdot), l) & \\
 & \triangleq \bar{\mathcal{L}}_{x_1x_1}(l)(\eta(t_1))^2 + 2\bar{\mathcal{L}}_{x_1x_2}(l)(\eta(t_1), \eta(t_2)) + \bar{\mathcal{L}}_{x_2x_2}(l)(\eta(t_2))^2 \\
 & + \int_{t_1}^{t_2} \{\bar{\mathcal{H}}_{xx}(t, p(t))(\eta(t))^2 \\
 & + 2[\bar{\mathcal{H}}_x(t, \bar{x}(t), u(t), p(t)) - \bar{\mathcal{H}}_x(t, p(t))]\eta(t)\} dt. \quad (109)
 \end{aligned}$$

When the requirements for Theorem 3.1 are met, condition (108) is quite different from (23). In fact, when Theorems 3.1 and 6.4 are applied to the examples of Ref. 7, pp. 287–288, it is seen that neither theorem is stronger than the other.

## References

1. BLISS, G. A., *Lectures on the Calculus of Variations*, University of Chicago Press, Chicago, Illinois, 1946.
2. BREAKWELL, J. V., and HO, Y. C., *On the Conjugate Point Condition for the Control Problem*, International Journal of Engineering Science, Vol. 2, pp. 565–579, 1965.
3. BRYSON, A. E., and HO, Y. C., *Applied Optimal Control*, Blaisdell, Waltham, Massachusetts, 1969.
4. WOOD, L. J., and BRYSON, A. E., *Second-Order Optimality Conditions for Variable End-Time Terminal Control Problems*, AIAA Journal, Vol. 11, pp. 1241–1246, 1973.
5. HESTENES, M., *Calculus of Variations and Optimal Control Theory*, John Wiley, New York, New York, 1964.
6. WARGA, J., *A Second-Order Lagrangian Condition for Restricted Control Problems*, Journal of Optimization Theory and Applications, Vol. 24, pp. 475–483, 1978.
7. WARGA, J., *A Second-Order Condition that Strengthens Pontryagin's Maximum Principle*, Journal of Differential Equations, Vol. 28, pp. 284–307, 1979.
8. WARGA, J., *A Hybrid Relaxed-Lagrangian Second-Order Condition for Minimum*, Differential Games and Control Theory, Proceedings of Third Kingston Conference, Part A, Marcel Dekker, New York, New York, pp. 77–94, 1979.
9. AGRACHEV, A. A., and GAMKRELIDZE, R. V., *A Second-Order Optimality Principle for a Time-Optimal Problem*, Mathematics USSR Sbornik, Vol. 29, pp. 547–576, 1976.
10. GABASOV, R., and KIRILLOVA, F. M., *High Order Necessary Conditions for Optimality*, SIAM Journal on Control, Vol. 10, pp. 127–168, 1972.
11. KNOBLOCH, H. W., *Higher Order Necessary Conditions in Optimal Control Theory*, Springer-Verlag, New York, New York, 1981.

12. KRENER, A. J., *The High Order Maximal Principle and Its Application to Singular Extremals*, SIAM Journal on Control and Optimization, Vol. 15, pp. 256–293, 1977.
13. BEN-TAL, A., *Second-Order and Related Extremality Conditions in Nonlinear Programming*, Journal of Optimization Theory and Applications, Vol. 31, pp. 143–165, 1980.
14. IOFFE, A. D., *Necessary and Sufficient Conditions for a Local Minimum 3: Second-Order Conditions and Augmented Duality*, SIAM Journal on Control and Optimization, Vol. 17, pp. 266–288, 1979.
15. LEVITIN, E. S., MILJUTIN, A. A., and OSMOLOVSKII, N. P., *On Conditions for a Local Minimum in a Problem with Constraints*, Mathematical Economics and Functional Analysis, Edited by B. S. Mitjagin, Nauka, Moscow, 1974 (in Russian).
16. MESSERLI, E. J., and POLAK, E., *On Second-Order Necessary Conditions for Optimality*, SIAM Journal on Control and Optimization, Vol. 7, pp. 272–291, 1969.
17. BERNSTEIN, D. S., *A Systematic Approach to Higher-Order Necessary Conditions in Optimization Theory*, SIAM Journal on Control and Optimization (to appear).
18. CANON, M., CULLUM, C., and POLAK, E., *Theory of Optimal Control and Mathematical Programming*, McGraw-Hill, New York, New York, 1970.
19. BROCKETT, R. W., *Finite-Dimensional Linear Systems*, John Wiley, New York, New York, 1970.
20. HAAS, V. B., *Normality and Controllability for the Problem of Bolza*, International Journal on Control, Vol. 30, pp. 331–338, 1979.
21. NEUSTADT, L. W., *Optimization, A Theory of Necessary Conditions*, Princeton University Press, Princeton, New Jersey, 1976.
22. WARGA, J., *Optimal Control of Differential and Functional Equations*, Academic Press, New York, New York, 1972.