SIAM J. CONTROL AND OPTIMIZATION
Vol. 22, No. 2, March 1984

© 1984 Society for Industrial and Applied Mathematics
0363-0129/84/2202-0004 $01.25/0

A SYSTEMATIC APPROACH TO HIGHER-ORDER NECESSARY CONDITIONS IN OPTIMIZATION THEORY*

DENNIS S. BERNSTEIN†

Abstract. Necessary conditions for an abstract optimization problem are derived under weak assumptions. The presence of a generalized critical direction in these conditions is the basis for deriving necessary conditions of arbitrary order for various concrete problems. Two applications are considered in detail. The first concerns first- and second-order necessary conditions for a constrained optimization problem in an infinite-dimensional vector space where the cost, equality and inequality functions possess differentials of a finite-dimensional one-sided character. The second application concerns first-, second- and third-order necessary conditions for a constrained optimization problem in a Banach space with Fréchet differentiability hypotheses. In both applications normality conditions are not required. Several well-known results are generalized.

Key words. higher-order necessary conditions, normality condition, convex constraint set, mathematical programming

1. Introduction. We consider the following optimization problem (OP). Minimize $\phi_0(e)$ subject to

$$e \in E, \tag{1.1}$$

$$\tilde{\phi}(e) \leqq 0, \tag{1.2}$$

$$\psi(e) = 0, \tag{1.3}$$

where: $\mathscr{E}$ is a set, $E \subset \mathscr{E}$, $k \in \{1, 2, \cdots\}$, $\psi: \mathscr{E} \to \mathbb{R}^k$, $\mathscr{Z}_0$ and $\tilde{\mathscr{Z}}$ are topological vector spaces, $\phi_0: \mathscr{E} \to \mathscr{Z}_0$ and $\tilde{\phi}: \mathscr{E} \to \tilde{\mathscr{Z}}$. Let $Z_0 \subset \mathscr{Z}_0$ and $\tilde{Z} \subset \tilde{\mathscr{Z}}$ be closed convex cones with nonempty interior such that $Z_0 \neq \mathscr{Z}_0$ and $\tilde{Z} \neq \tilde{\mathscr{Z}}$. For $z, \hat{z} \in \mathscr{Z}_0$, $z < z$ means $z - \hat{z} \in \operatorname{int} Z_0$ and $z \leqq z$ means $z - \hat{z} \in Z_0$. Identical notation applies if $z, \hat{z} \in \tilde{\mathscr{Z}}$. The vector spaces are defined over the real field. The element e is feasible if (1.1)–(1.3) are satisfied and a feasible element $\bar{e}$ solves OP if there is no feasible element e such that $\phi_0(e) < \phi_0(\bar{e})$.

The purpose of this paper is to present a systematic approach for deriving higher-order necessary conditions for a solution of OP. By first deriving necessary conditions for OP under weak assumptions, we obtain results for more specialized problems by successively incorporating stronger hypotheses. New results are obtained and several well-known results are generalized.

This work was motivated by two factors. The first of these was the appearance in the literature of second-order necessary conditions with significantly different assumptions. In particular, [35, Thm. 2.3] involves an infinite-dimensional constraint space without a topology, a convex constraint set, directional differentials and a full-range normality assumption. On the other hand, [28, Thm. 6] involves a finite-dimensional Euclidean constraint space, a conical approximation to the constraint set, continuous differentiability and no normality assumption. The second motivating factor was the series of papers [34], [35] and [36] which are based entirely

* Received by the editors November 30, 1981, and in revised form December 20, 1982. This research was performed at the University of Michigan, Ann Arbor, Michigan, Program in Computer, Information and Control Engineering. Support was provided by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under grant AFOSR-77-3158.

† Control Systems Engineering Group, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02173.

on [35, Thm. 2.3]. Weak, strong and "hybrid" variations are systematically exploited in these papers to derive a trio of second-order necessary conditions for optimal control. Thus, the usefulness of [35, Thm. 2.3] for optimal control and the possibility of its being generalized were evident.

The approach of the present paper follows in the spirit of [29]. [29, Thm. 3.1] contains the essential features of a wide variety of optimization problems and yields several well-known first-order necessary conditions as special cases. In the present paper, the formulation of [29] is expanded in the Main Theorem (§ 2) to include a generalized critical direction. Although the Main Theorem itself involves no differentiability hypotheses, this extra feature is the key ingredient for obtaining necessary conditions of arbitrary order when differentiability assumptions are present. Specifically, the generalized critical direction Y and accompanying convex set K correspond respectively to the intermediate- and highest-order terms in a power series expansion.

To illustrate the role of Y and K in deriving higher-order necessary conditions, let $\mathscr{E}$ be a subset of $\mathbb{R}^\nu$, $\mathscr{Z}_0 = \mathbb{R}$, $\tilde{\mathscr{Z}} = \mathbb{R}^m$, $Z_0 = \bar{\mathbb{R}}_-$ and $\tilde{Z} = \bar{\mathbb{R}}_-^m$, where $\bar{\mathbb{R}}_-$ denotes the nonpositive real numbers. Also, let $\bar{e} \in E$, $y \in E - \bar{e}$ and define $\Phi = (\phi_0, \phi, \psi)$. Then in the second-order expansion

$$\Phi(\bar{e} + \alpha y + \alpha^2 x) \approx \Phi(\bar{e}) + \alpha \Phi'(\bar{e}) y + \frac{\alpha^2}{2}[2\Phi'(\bar{e})x + y^T \Phi''(\bar{e}) y],$$

where $x \in E - \bar{e}$ and $\alpha > 0$, Y and K are given by

$$Y = (\phi_0, \tilde{\phi})'(\bar{e}) y,$$

$$K = \{\Phi'(\bar{e})x + \tfrac{1}{2} y^T \Phi''(\bar{e}) y : x \in E - \bar{e}\}.$$

Since we seek directions y which have inferior cost and which approximately satisfy (1.2), we require that $Y \in \bar{\mathbb{R}}_-^{m+1}$. Also, because of (1.3) the term $\psi'(\bar{e})y$ is required to be zero and thus does not appear in the definition of Y. As will be seen, Y leads to the generalized complementary slackness condition $l_\phi(Y) = 0$, where $l = (l_\phi, l_\psi)$ is the Lagrange multiplier, and K yields the Lagrangian condition $l(h) \geqq 0$, $h \in K$. Note that first-order necessary conditions are obtained by setting $y = 0$ (and hence $Y = 0$). Third- and higher-order conditions can be obtained in a straightforward manner. For example, we can consider a pair of directions y and $\hat{y}$ in $E - \bar{e}$ to obtain the third-order expansion

$$\Phi(\bar{e} + \alpha y + \alpha^2 \hat{y} + \alpha^3 x) \approx \Phi(\bar{e}) + \alpha \Phi'(\bar{e}) y + \frac{\alpha^2}{2}[2\Phi'(\bar{e})\hat{y} + y^T \Phi''(\bar{e}) y]$$

$$+ \frac{\alpha^3}{6}[6\Phi'(\bar{e})x + 6y^T \Phi''(\bar{e})\hat{y} + \Phi'''(\bar{e})(y)^3].$$

Now Y and K are given by

$$Y = (\phi_0, \tilde{\phi})'(\bar{e}) y + (\phi_0, \tilde{\phi})'(\bar{e})\hat{y} + \tfrac{1}{2} y^T (\phi_0, \tilde{\phi})''(\bar{e}) y,$$

$$K = \{\Phi'(\bar{e})x + y^T \Phi''(\bar{e})\hat{y} + \tfrac{1}{6}\Phi'''(\bar{e})(y)^3 : x \in E - \bar{e}\}.$$

In the Main Theorem ϕ_0, $\tilde{\phi}$ and ψ are assumed to possess weak approximation-like properties (see the Main Condition in § 2). As in [10, Thm. 13.1], these properties are stated solely in terms of the elements of the image spaces of ϕ_0, $\tilde{\phi}$ and ψ, and hence the set E requires neither topological nor algebraic structure. Because of this fundamental setting, the proof of the Main Theorem given in § 3 is quite simple and succinct. Further simplification is obtained by utilizing a refined separation theorem from [19].

By imposing additional structure (but still no differentiability hypotheses), our next result, Theorem 4.1, follows from the Main Theorem. In the hypotheses for Theorem 4.1 (Condition 4.1), an auxiliary vector space is introduced and the cost and constraint functions are assumed to possess nth-order polynomial expansions. Because Condition 4.1 retains much of the generality of the Main Condition, results from the literature involving conical approximations (see, e.g., [7], [28]) can be obtained as corollaries. In our development, Theorem 4.1 serves as a convenient intermediate step to the results of §§ 5 and 6. For example, the polynomial expansions of Condition 4.1 are given concrete realizations in terms of directional differentials in § 5 and in terms of Frechet derivatives in § 6.

Section 5 begins with a generalization of the Fréchet derivative (the $\tilde{F}$-derivative) due to Warga (Definition 5.1). This definition allows us to work with one-sided directional differentials when several directions appear simultaneously. Theorems 5.1 and 5.2 contain first- and second-order necessary conditions, respectively. The relationship between these results can be rather complex. Examples are given to illustrate the following points: 1) the first-order conditions of Theorem 5.1 may not follow from the second-order conditions of Theorem 5.2, and 2) Theorem 5.2 may yield necessary conditions that have the appearance of first-order necessary conditions but which are unobtainable from Theorem 5.1. The discussion generalizes and clarifies some remarks made in [28]. Further specialization leads to Theorem 5.4 which generalizes [35, Thm. 2.3]. It is shown that for each critical direction y there exists a multiplier l satisfying both the first-order necessary conditions and an additional second-order condition.

In § 6, E is a subset of Banach space and ϕ_0, $\tilde{\phi}$ and ψ are assumed to be Fréchet differentiable. OP now closely resembles a nonlinear programming problem (NP). First-, second- and third-order necessary conditions for this problem are given in Theorem 6.1. The third-order conditions are derived directly from Theorem 4.1 to illustrate the usefulness of the Main Theorem in obtaining higher-order necessary conditions. The relationship between Theorem 6.1 and various higher-order necessary conditions in the literature is discussed.

The appendix contains notation and results pertaining to Fréchet and $\tilde{F}$-derivatives. We state a version of Taylor's theorem for $\tilde{F}$-derivatives and a result on converting a multivariable expansion into a one-parameter expansion with remainder satisfying a uniform convergence condition. These results allow nth-order generalizations of the results of §§ 5 and 6.

It is important to point out that neither normality assumptions nor constraint qualifications appear in the statements of the necessary conditions. The absence of these hypotheses leads to nonuniqueness of the multiplier l and hence dependence of l on the critical direction y (or critical directions $y, \hat{y}, \cdots$ for third- and higher-order conditions). This idea seems to have first appeared (without proof) in [27] and was apparently rediscovered (in a more general version) in [28]. Subsequent related results can be found in [3], [4], [5], [14], [16], [18], [22], [23], [24], [32].

Before continuing, it is convenient to collect here general notation and definitions and some results concerning vector spaces and topological vector spaces. The empty set is denoted by $\varnothing$. If S_1 and S_2 are sets, then $S_1/S_2 \triangleq \{s \in S_1 : s \notin S_2\}$. If A and B are sets, $A_1 \subset A$, $B_1 \subset B$ and $f: A \to B$ then $f(A_1) \triangleq \{f(a): a \in A_1\}$ and $f^{-1}(B_1) \triangleq \{a \in A: f(a) \in B_1\}$. Let $\mathbb{N} \triangleq \{1, 2, \cdots\}$, $\mathbb{R} \triangleq$ real field, $\mathbb{R}_+ \triangleq \{\alpha \in \mathbb{R}: \alpha > 0\}$, $\mathbb{R}_- \triangleq \{\alpha \in \mathbb{R}: \alpha < 0\}$, $\bar{\mathbb{R}}_+ \triangleq \mathbb{R}_+ \cup \{0\}$ and $\bar{\mathbb{R}}_- \triangleq \mathbb{R}_- \cup \{0\}$. The results obtained here do not depend on the choice of norm for $\mathbb{R}^m$; for convenience, the norm of $\boldsymbol{\alpha} \triangleq (\alpha_1, \cdots, \alpha_m) \in \mathbb{R}^m$ is taken to be $|\boldsymbol{\alpha}| = \sum_{i=1}^m |\alpha_i|$. For $\beta > 0$, $\mathbb{B}^m(\beta) \triangleq \{\boldsymbol{\alpha} \in \mathbb{R}^m: |\boldsymbol{\alpha}| < \beta\}$ and $\mathbb{P}^m(\beta) \triangleq \{\boldsymbol{\alpha} \in \bar{\mathbb{R}}_+^m: |\alpha| < \beta\}$. Define $\Delta^m \triangleq \{\boldsymbol{\mu} \triangleq (\mu_1, \cdots, \mu_{m+1}) \in \bar{\mathbb{R}}_+^{m+1}: |\boldsymbol{\mu}| = 1\}$. When not implied by context, the origin of $\mathbb{R}^m$ is denoted by 0_m.

Let $\mathscr{V}$ be a vector space and V, $\hat{V} \subset \mathscr{V}$. Define co $V \triangleq$ convex hull of V, cone $V \triangleq \{\alpha v: \alpha > 0, v \in V\}$ and coco $V \triangleq$ co cone V. V is a cone if $V =$ cone V and a convex cone if $V =$ coco V. If $v \in V$ and $V =$ cone $\{v\}$ then V is a ray. Let $\alpha V \triangleq \{\alpha v: v \in V\}$, where $\alpha \in \mathbb{R}$, $V + \hat{V} \triangleq \{v + \hat{v}: v \in V, \hat{v} \in \hat{V}\}$ and $V - \hat{V} \triangleq V + (-1)\hat{V}$. For $v \in \mathscr{V}$, cone $v \triangleq$ cone $\{v\}$, $V + v \triangleq V + \{v\}$ and $V - v \triangleq V + (-v)$. If V is a cone then $\alpha V = V$, $\alpha > 0$; if V is a convex cone then $\alpha V + \beta V = V$, $\alpha > 0$, $\beta > 0$. If V is convex the dimension of V is dim V.

Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces, $V \subset \mathscr{V}$ and $f: V \to \mathscr{W}$. The vectors $v_1, \cdots, v_m \in \mathscr{V}$ are linearly independent if $\boldsymbol{\alpha} \in \mathbb{R}^m$ and $\sum_{i=1}^m \alpha_i v_i = 0$ imply that $\boldsymbol{\alpha} = 0$; they are affinely independent if $\boldsymbol{\alpha} \in \mathbb{R}^m$, $\sum_{i=1}^m \alpha_i v_i = 0$ and $\sum_{i=1}^m \alpha_i = 0$ imply that $\boldsymbol{\alpha} = 0$. V is an m-simplex if $m \in \{0, 1, 2, \cdots\}$ and $V = \text{co}\{v_1, \cdots, v_{m+1}\}$, where $v_1, \cdots, v_{m+1}$ are affinely independent. The points $v_1, \cdots, v_{m+1}$ are the (unique) vertices of V and, for $v \triangleq \sum_{i=1}^{m+1} \mu_i v_i \in V$, where $\boldsymbol{\mu} \in \Delta^m$, the numbers $\mu_1, \cdots, \mu_{m+1}$ are the (unique) barycentric coordinates of v. f is positively homogeneous if V is a cone and $f(\alpha v) = \alpha f(v)$, $\alpha > 0$, $v \in V$; f is affine if V is convex and for every $m \in \mathbb{N}$, $\boldsymbol{\mu} \in \Delta^{m-1}$ and $v_1, \cdots, v_m \in V$ it follows that $f(\sum_{i=1}^m \mu_i v_i) = \sum_{i=1}^m \mu_i f(v_i)$. Let W be a cone and for $w_1, w_2 \in \mathscr{W}$ let $w_1 \leqq w_2$ denote $w_1 - w_2 \in W$. f is W-convex if V is convex and for every $m \in \mathbb{N}$, $\boldsymbol{\mu} \in \Delta^{m-1}$ and $v_1, \cdots, v_m \in V$ it follows that $f(\sum_{i=1}^m \mu_i v_i) \leqq \sum_{i=1}^m \mu_i f(v_i)$.

Suppose that $\mathscr{V}$, $\mathscr{V}_1, \cdots, \mathscr{V}_n$ and $\mathscr{W}$ are topological vector spaces and $V \subset \mathscr{V}$. The closure of V is cl V, the interior of V is int V and the boundary of V is bd $V \triangleq$ (cl V)/(int V). V is solid if int $V \neq \varnothing$. $\mathscr{B}(\mathscr{V}_1, \cdots, \mathscr{V}_n; \mathscr{W})$ denotes the vector space of all continuous multilinear mappings from $\mathscr{V}_1 \times \cdots \times \mathscr{V}_n$ into $\mathscr{W}$. If $\mathscr{V}_1 = \cdots = \mathscr{V}_n$ then we write $\mathscr{B}_n(\mathscr{V}; \mathscr{W})$ for $\mathscr{B}(\mathscr{V}_1, \cdots, \mathscr{V}_n; \mathscr{W})$. Recall (see, e.g., [11, p. 318]) that if $\mathscr{V}_1, \cdots, \mathscr{V}_n$ and $\mathscr{W}$ are Banach spaces with norms $|\cdot|_{\mathscr{V}_1}, \cdots, |\cdot|_{\mathscr{V}_n}$ and $|\cdot|_{\mathscr{W}}$, respectively, then $\mathscr{B}(\mathscr{V}_1, \cdots, \mathscr{V}_n; \mathscr{W})$ is a Banach space with norm $|F| \triangleq \sup\{|F(v_1, \cdots, v_n)|_{\mathscr{W}}: v_i \in \mathscr{V}_i, |v_i|_{\mathscr{V}_i} = 1, i = 1, \cdots, n\}$. Define the dual space $\mathscr{V}^* \triangleq \mathscr{B}(\mathscr{V}; \mathbb{R})$ and the conjugate cone $V^{\ominus} \triangleq \{l \in \mathscr{V}^*: l(v) \leqq 0, v \in V\}$.

If $\mathscr{V}_1$ and $\mathscr{V}_2$ are topological vector spaces, then $\mathscr{V}_1 \times \mathscr{V}_2$ is assumed to be the topological vector space possessing the product topology. $(\mathscr{V}_1 \times \mathscr{V}_2)^*$ and $\mathscr{V}_1^* \times \mathscr{V}_2^*$ are in one-to-one correspondence in the sense that $l \in (\mathscr{V}_1 \times \mathscr{V}_2)^*$ if and only if there exist $l_1 \in \mathscr{V}_1^*$ and $l_2 \in \mathscr{V}_2^*$ such that $l(v) = l_1(v_1) + l_2(v_2)$, $v \triangleq (v_1, v_2) \in \mathscr{V}_1 \times \mathscr{V}_2$. Specifically, $l_1(v_1) \triangleq l(v_1, 0)$, $v_1 \in \mathscr{V}_1$, and $l_2(v_2) \triangleq l(0, v_2)$, $v_2 \in \mathscr{V}_2$. We denote this correspondence by $l = (l_1, l_2)$. If $\Omega \subset (\mathscr{V}_1 \times \mathscr{V}_2)^*$ and $\hat{\Omega} \subset \mathscr{V}_1^* \times \mathscr{V}_2^*$ then the relations $\Omega \subset \hat{\Omega}$, $\Omega = \hat{\Omega}$, etc., can be interpreted in this sense. Analogous remarks apply to $(\mathscr{V}_1 \times \cdots \times \mathscr{V}_n)^*$ and $\mathscr{V}_1^* \times \cdots \times \mathscr{V}_n^*$, where $\mathscr{V}_1, \cdots, \mathscr{V}_n$ are topological vector spaces.

For the following results let $\mathscr{X}$ and $\mathscr{W}$ be vector spaces, $M \subset \mathscr{X}$ and $f: M \to \mathscr{W}$.

PROPOSITION 1.1. *Assume M is a convex cone, f is positively homogeneous and affine and $x_1, \cdots, x_m \in M$. If $f(x_1), \cdots, f(x_m)$ are linearly independent then $x_1, \cdots, x_m$ are linearly independent.*

PROPOSITION 1.2. *Assume M is convex, f is affine and $x_1, \cdots, x_m \in M$. If $f(x_1), \cdots, f(x_m)$ are affinely independent then $x_1, \cdots, x_m$ are affinely independent.*

Now let $\mathscr{V}$, $\mathscr{V}_1$ and $\mathscr{V}_2$ be topological vector spaces, $V, \hat{V} \subset \mathscr{V}$, $V_1 \subset \mathscr{V}_1$ and $V_2 \subset \mathscr{V}_2$.

PROPOSITION 1.3. *If V is an m-simplex with vertices $v_1, \cdots, v_{m+1}$, then the map $\boldsymbol{\mu} \to \sum_{i=1}^{m+1} \mu_i v_i : \Delta^m \to V$ is a homeomorphism.*

PROPOSITION 1.4. *If V is solid and convex, then* $\operatorname{cl} \operatorname{int} V = \operatorname{cl} V$ *and* $\operatorname{int} \operatorname{cl} V = \operatorname{int} V$.

PROPOSITION 1.5. $V^{\ominus} = (\operatorname{cl} V)^{\ominus} = (\operatorname{co} V)^{\ominus} = (\operatorname{cone} V)^{\ominus}$.

PROPOSITION 1.6. *If V is solid and convex, then* $V^{\ominus} = (\operatorname{int} V)^{\ominus}$.

PROPOSITION 1.7. *If V is a solid closed convex cone, then* $V + \operatorname{int} V = \operatorname{int} V$.

PROPOSITION 1.8. *If V is open, $l \in V^{\ominus}$ and $0 \in l(V)$, then $l = 0$.*

PROPOSITION 1.9. $V^{\ominus} \cap \hat{V}^{\ominus} = (V + \operatorname{cone} \hat{V})^{\ominus}$.

PROPOSITION 1.10. $V_1^{\ominus} \times V_2^{\ominus} = (V_1 \times \operatorname{cone} V_2)^{\ominus}$.

Proofs. Propositions 1.1 and 1.2 follow from the definitions of linear and affine independence. Propositions 1.3 and 1.4 can be found in [30, pp. 25, 27] and [17, p. 59], respectively. Proposition 1.5 follows from the linearity and continuity of the elements of $V^{\ominus}$. Using Propositions 1.4 and 1.5 we have $V^{\ominus} = (\operatorname{cl} V)^{\ominus} = (\operatorname{cl} \operatorname{int} V)^{\ominus} = (\operatorname{int} V)^{\ominus}$, which proves Proposition 1.6. Proposition 1.7 is a consequence of [30, p. 28], and Proposition 1.4. Proposition 1.8 follows from [30, p. 34], Proposition 1.5 and the fact that cone V is open. To prove Proposition 1.9, let $l \in (V + \operatorname{cone} \hat{V})^{\ominus}$ and suppose $l \notin V^{\ominus}$, i.e., there exists $v_1 \in V$ such that $l(v_1) > 0$. For each $\hat{v} \in \hat{V}$ there exists $\alpha > 0$ (sufficiently small) such that $l(v_1 + \alpha\hat{v}) > 0$, which is a contradiction. Now suppose $l \notin \hat{V}^{\ominus}$, i.e., there exists $v_2 \in \hat{V}$ such that $l(v_2) > 0$. Then, for each $v \in V$ there exists $\alpha > 0$ (sufficiently large) such that $l(v + \alpha v_2) > 0$, which is also a contradiction. The reverse inclusion follows from Proposition 1.5 and the obvious fact $V^{\ominus} \cap \hat{V}^{\ominus} \subset (V + \hat{V})^{\ominus}$. Identical arguments can be used to prove Proposition 1.10. $\square$

We conclude this section with some comments about the orderings $\leqq$ and $<$ on $\mathscr{Z}_0$ (similar remarks apply to $\hat{\mathscr{Z}}$). Clearly, $\leqq$ is reflexive and both orderings are transitive. The ordering $\leqq$ is antisymmetric only when $Z_0 \cap (-Z_0) = \{0\}$, i.e., when Z_0 contains no lines. Because $\operatorname{int} Z_0$ does not contain any lines, the relations $z < \hat{z}$ and $\hat{z} < z$ are never both satisfied.

Note that both orderings are compatible with the linear structure on $\mathscr{Z}_0$ in the sense that if $z \leqq \hat{z}$ then $\alpha z \leqq \alpha\hat{z}$, $\alpha \geqq 0$, and if $z \leqq \hat{z}$ and $z' \leqq \hat{z}'$ then $z + z' \leqq \hat{z} + \hat{z}'$ (and similarly for $<$ with $\alpha = 0$ excluded). Finally, Proposition 1.7 leads to the fact that $z \leqq 0$ and $z < 0$ imply $z + \hat{z} < 0$.

2. The Main Theorem. First, we introduce some notation and conventions for OP which simplify the statement of the necessary conditions in this and subsequent sections. Without loss of generality we assume in this section and in §§ 4 and 5 that the solution $\bar{e}$ of OP satisfies $\phi_0(\bar{e}) = 0$. This convention, which simplifies the notation considerably, can be removed by replacing ϕ_0 by $\phi_0 - \phi_0(\bar{e})$ wherever it appears. In order to deal efficiently with various special cases of OP we adopt the convention that $\tilde{\mathscr{Z}} = \varnothing$ and $k = 0$ denote the absence of, respectively, (1.2) and (1.3). For the case $\tilde{\mathscr{Z}} \neq \varnothing$ (i.e., (1.2) is present) define $\mathscr{Z} \triangleq \mathscr{Z}_0 \times \tilde{\mathscr{Z}}$, $Z \triangleq Z_0 \times \tilde{Z}$ and $\phi \triangleq (\phi_0, \tilde{\phi})$; when $\tilde{\mathscr{Z}} = \varnothing$ let $\mathscr{Z} \triangleq \mathscr{Z}_0$, $Z \triangleq Z_0$ and $\phi \triangleq \phi_0$. For $k \neq 0$ (i.e., (1.3) is present) define $\mathscr{Z}' \triangleq \mathscr{Z} \times \mathbb{R}^k$ and for $k = 0$ let $\mathscr{Z}' \triangleq \mathscr{Z}$. For $h \in \mathscr{Z} \times \mathbb{R}^k$ let $\pi_\phi h \triangleq$ component of h in $\mathscr{Z}$ and $\pi_\phi \triangleq$ component of $\mathbb{R}^k$. If $H \subset \mathscr{Z} \times \mathbb{R}^k$ then $\pi_\phi H$ and $\pi_\psi H$ are defined in the obvious way.

The assumptions required for the Main Theorem are contained in the Main Condition (MC) which follows. Roughly speaking, this condition involves: a rep-

resentation of E (through a set K and a map Θ) "near" the feasible point $\bar{e}$, a generalized "critical direction" Y, which is the basis for higher-order necessary conditions, and some differentiability-like conditions ((2.4) and (2.6)) on ϕ_0, $\tilde{\phi}$ and ψ. It is shown in later sections that MC includes as special cases conditions commonly satisfied in optimization problems. The paper [29] contains closely related assumptions and considerable comment about them.

MAIN CONDITION (MC). There exist a feasible element $\bar{e}$, $Y \in Z$ and a nonempty convex set $K \subset \mathscr{Z}'$ such that the following property is satisfied. (Let σ and η be positive real numbers, γ be a nonzero real number, S be a k-simplex in K, N be a neighborhood of the origin in $\mathscr{Z}$, $A: \mathscr{Z} \to \mathscr{Z}$ and $\Theta: S \to E$.) For all N and η and for all S and σ satisfying

(2.1) $\quad 0 \in \operatorname{int} \pi_\psi S$ (omit if $k = 0$),

(2.2) $\quad \sigma(\phi(\bar{e}) + Y) + \pi_\phi S \subset \operatorname{int} Z$

there exist γ, A and Θ satisfying

(2.3) $\quad A^{-1}(\operatorname{int} Z) \subset \operatorname{int} Z$,

(2.4) $\quad A \circ \phi \circ \Theta(h) \in \sigma(\phi(\bar{e}) + Y) + \pi_\phi S + Z + N,\ h \in S$,

(2.5) $\quad \psi \circ \Theta: S \to \mathbb{R}^k$ is continuous (omit if $k = 0$),

(2.6) $\quad |\gamma\psi \circ \Theta(h) - \pi_\psi h| < \eta,\ h \in S$ (omit if $k = 0$).

MAIN THEOREM (MT). *Suppose* MC *is satisfied. If* $\bar{e}$ *solves* OP *then there exists* $l \triangleq (l_\phi, l_\psi) \in \mathscr{Z}^* \times \mathbb{R}^k$ *when* $k > 0$ *and* $l \triangleq l_\phi \in \mathscr{Z}^*$ *when* $k = 0$ *such that*

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

(2.8) $\quad l_\phi \in Z^\ominus$,

(2.9) $\quad l_\phi(\phi(\bar{e}) + Y) = 0$,

(2.10) $\quad l(h) \geqq 0,\ h \in K$.

Remark 2.1. When $k = 0$ the simplex S in MC consists of a single element h. In this case the role of Θ and S can be handled by a single element $e \triangleq \Theta(h) \in E$ and (2.4) can be replaced by

$$A \circ \phi(e) \in \sigma(\phi(\bar{e}) + Y) + h + Z + N. \tag{2.4$'$}$$

The distinction between (2.4) and (2.4)′ accounts for the pair of conditions [29, 3.1 and 3.2].

Remark 2.2. The focus in [29] is on first-order necessary conditions for a very general extremal problem which includes OP as a special case. MC and MT can be reformulated to apply to this problem although this is not pursued here. In the context of OP, MC both weakens and generalizes [29, conditions 3.1 and 3.2]. These conditions involve the introduction of an auxiliary vector space $\mathscr{Y}$, a convex set $M \subset \mathscr{Y}$ and a map $f: M \to \mathscr{Z}'$. Necessary conditions are then stated in terms of the elements of M. MC is weaker since $\mathscr{Y}$ does not appear. Instead MC involves elements of the set K which corresponds to the image of M under f. A similar idea appears in [10, Thm. 13.1], p. 46, which follows from MT with $\mathscr{Z}_0 = \mathbb{R}$, $\tilde{\mathscr{Z}} = \varnothing$, $Y = 0$ and K a cone. The term Y, which leads to higher-order necessary conditions, has no counterpart in either [10] or [29].

Remark 2.3. Because $\phi(\bar{e}) \in Z$, $Y \in Z$ and $l_\phi \in Z^\ominus$, it follows that (2.9) is equivalent to the pair of conditions

$$l_\phi(\phi(\bar{e})) = 0, \tag{2.11}$$

$$l_\phi(Y) = 0. \tag{2.12}$$

Note that (2.11) is a complementary slackness condition in that it yields additional information concerning l_ϕ. By virtue of the additional condition (2.12), (2.9) may be viewed as a generalized complementary slackness condition.

Sometimes MC is verified with a qualification of S which is weaker than (2.1) and (2.2). Although this results in a stronger version of MC it may be more suitable for applications. For example, either or both of the conditions (2.1) and (2.2) may be omitted. A common situation is the subject of the following easily proved result.

PROPOSITION 2.1. *Suppose* MC *is satisfied and*

$$\phi(\bar{e}) + Y \in \operatorname{bd} Z. \tag{2.13}$$

Then (2.1) *and* (2.2) *imply*

(2.14) *The vertices of S are linearly independent.*

Remark 2.4. Because $\phi_0(\bar{e}) = 0$, (2.13) is satisfied when $Y = 0$. This is the situation in [10, Thm. 13.1]. There (2.1) and (2.2) are omitted and the (weaker) linear independence condition (2.14) appears.

Remark 2.5. The set $\mathscr{E}$ plays no role in the Main Theorem or its proof. In §§ 5 and 6, $\mathscr{E}$ is used to define differentiability properties which cannot be stated solely in terms of the elements of E.

3. Proof of the Main Theorem. The following notation and definitions are needed. Let $\mathscr{V}$ be a vector space and $V \subset \mathscr{V}$. If $V + v$ is a subspace of $\mathscr{V}$ for some $v \in \mathscr{V}$ then V is an affine subset of $\mathscr{V}$. The codimension of an affine subset V, codim V, is the dimension of a subspace $\tilde{V} \subset \mathscr{V}$ such that the direct sum of $\tilde{V}$ and the subspace $V + v$ is $\mathscr{V}$. The affine hull of an arbitrary set V, aff V, is the smallest affine subset containing V.

Suppose now that $\mathscr{V}$ is a topological vector space and $V, A, B \subset \mathscr{V}$. The interior of V relative to aff V is denoted by ri V. $l \in \mathscr{V}^*$ separates A and B if $l \neq 0$ and there exists $\alpha \in \mathbb{R}$ such that $l(a) \leqq \alpha \leqq l(b)$, $a \in A$, $b \in B$. If either A or B is a cone then α can be chosen to be zero without loss of generality and thus $l \in A^\ominus \cap (-B)^\ominus$.

We will also require the following easily verified results. If $\mathscr{V}_1$ and $\mathscr{V}_2$ are vector spaces, $V_1 \subset \mathscr{V}_1$ and $V_2 \subset \mathscr{V}_2$ then aff $(V_1 \times V_2) = (\text{aff } V_1) \times (\text{aff } V_2)$. If in addition $\mathscr{V}_1$ and $\mathscr{V}_2$ are topological vector spaces then ri $(V_1 \times V_2) = (\text{ri } V_1) \times (\text{ri } V_2)$.

The proof of MT rests upon the following separation lemma. This result is a generalization of a well-known theorem which follows when int $A \neq \varnothing$ (see, e.g., [7, p. 63], or [31, p. 24]).

LEMMA 3.1. *Let A and B be convex subsets of a topological vector space $\mathscr{V}$ such that* aff A *is closed and has finite codimension,* ri $A \neq \varnothing$ *and* $(\text{ri } A) \cap B = \varnothing$. *Then there exists $l \in \mathscr{V}^*$ separating A and B.*

Lemma 3.1 is a corollary of an algebraic separation theorem stated without proof in the survey paper [19, p. 253]. In the algebraic setting $\mathscr{V}$ is assumed to be a vector space and the "intrinsic core" of A plays the role of ri A. The additional assumption that aff A is closed implies that l is continuous (see [19, pp. 240–1]). A proof of the algebraic separation theorem can be obtained by means of induction on n = codim aff A using a method similar to that used in [31, proof of Thm. 2.9].

We consider the full proof of the Main Theorem only for the case $\tilde{\mathscr{Z}} \neq \emptyset$ and $k > 0$ since the proofs of the remaining cases involve similar arguments. A large portion of the proof is contained in the following two lemmas. Suppose MC is satisfied and define $\hat{Z} \triangleq (\operatorname{int} Z) - \operatorname{cone}(\phi(\bar{e}) + Y)$. Note that $\hat{Z}$ is an open convex cone.

LEMMA 3.2. *If K and $\hat{Z} \times \{0\}$ are separated by a continuous linear functional then there exists $l = (l_\phi, l_\psi) \in \mathscr{Z}^* \times \mathbb{R}^k$ satisfying* (2.7)–(2.10).

Proof. Since $\hat{Z} \times \{0\}$ is a cone there exists $l = (l_\phi, l_\psi) \in \mathscr{Z}^* \times \mathbb{R}^k$ satisfying (2.7) and such that $l \in (-K)^\ominus \cap (\hat{Z} \times \{0\})^\ominus$. Thus l satisfies (2.10). By Proposition 1.10, $l_\phi \in \hat{Z}^\ominus$ and, by Proposition 1.9, $l_\phi \in (\operatorname{int} Z)^\ominus$ and $l_\phi(\phi(\bar{e}) + Y) \geqq 0$. By Proposition 1.6 $(\operatorname{int} Z)^\ominus = Z^\ominus$ and thus (2.8) holds. Since $\phi(\bar{e}) + Y \in Z$, $l_\phi(\phi(\bar{e}) + Y) \leqq 0$, which implies (2.9). □

LEMMA 3.3. *If K and $\hat{Z} \times \{0\}$ are not separated by a continuous linear functional then there exist a k-simplex $S \subset K$ and a positive real number σ satisfying* (2.1) *and* (2.2).

Proof. We first show that $0 \in \operatorname{int} \pi_\psi K$. If this is not true, then it follows (see, e.g., [30, Thm. I.5.19] that there exists $\xi \in \mathbb{R}^k$, $\xi \neq 0$, such that $\xi \cdot v \geqq 0$, $v \in \pi_\psi K$. Then $l = (0, \xi) \in \mathscr{Z}^* \times \mathbb{R}^k$ separates K and $\hat{Z} \times \{0\}$, which is a contradiction. Since $0 \in \operatorname{int} \pi_\psi K$ there exists a k-simplex $U \triangleq \operatorname{co}\{u_1, \cdots, u_{k+1}\} \subset \pi_\psi K$ such that $0 \in \operatorname{int} U$. For each $i \in \{1, \cdots, k+1\}$ let $s_i \in K$ satisfy $\pi_\psi s_i = u_i$.

Since $\operatorname{aff}(\hat{Z} \times \{0\}) = (\operatorname{aff} \hat{Z}) \times (\operatorname{aff}\{0\}) = \mathscr{Z} \times \{0\}$ is closed and has finite codimension k and since $\operatorname{ri}(\hat{Z} \times \{0\}) = (\operatorname{ri} \hat{Z}) \times (\operatorname{ri}\{0\}) = \hat{Z} \times \{0\} \neq \emptyset$, Lemma 3.1 implies that $(\hat{Z} \times \{0\}) \cap K \neq \emptyset$. Thus, there exists $s \triangleq (\hat{z} - \sigma'(\phi(\bar{e}) + Y), 0) \in K$, where $\hat{z} \in \operatorname{int} Z$ and $\sigma' > 0$. Since $\hat{z} \in \operatorname{int} Z$ we can choose $\lambda \in (0, 1)$ sufficiently close to 1 so that $\lambda \hat{z} + (1-\lambda)\pi_\phi s_i \in \operatorname{int} Z$, $i \in \{1, \cdots, k+1\}$. Define $h_i \triangleq \lambda s + (1-\lambda) s_i$, $i \in \{1, \cdots, k+1\}$, and $S \triangleq \operatorname{co}\{h_1, \cdots, h_{k+1}\}$. Letting $\sigma \triangleq \lambda \sigma'$, and $\tilde{z} \triangleq \sigma(\phi(\bar{e}) + Y)$ it is easy to see that $h_i = (\lambda \hat{z} + (1-\lambda)\pi_\phi s_i - \tilde{z}, (1-\lambda) u_i) \in K$, $S \subset K$ and $\pi_\phi h_i \in (\operatorname{int} Z) - \tilde{z}$. Since $\pi_\psi S = (1-\lambda) U$ and $0 \in \operatorname{int} U$, (2.1) must hold. Since $u_1, \cdots, u_{k+1}$ are affinely independent and $(1-\lambda)^{-1} \pi_\psi h_i = u_i$ it follows from Proposition 1.2 that $h_1, \cdots, h_{k+1}$ are affinely independent and hence S is a k-simplex. Finally, since $\tilde{z} + \pi_\phi h_i \in \operatorname{int} Z$, $i \in \{1, \cdots, k+1\}$, (2.2) is satisfied. □

We can now proceed with the proof of MT. Suppose that the theorem is false. By Lemmas 3.2 and 3.3 there exist a k-simplex $S = \operatorname{co}\{h_1, \cdots, h_{k+1}\} \subset K$ and a positive real number σ satisfying (2.1) and (2.2). Since $\sigma(\phi(\bar{e}) + Y) + \pi_\phi h_i \in \operatorname{int} Z$, $i \in \{1, \cdots, k+1\}$, the open set $N \triangleq \bigcap_{i=1}^{k+1} ((\operatorname{int} Z) - \sigma(\phi(\bar{e}) + Y) - \pi_\phi h_i)$ is a neighborhood of the origin in $\mathscr{Z}$. It follows easily that $\sigma(\phi(\bar{e}) + Y) + \pi_\phi S + N \subset \operatorname{int} Z$. Because of (2.1) we can choose $\eta > 0$ so that $\mathbb{B}^k(\eta) \subset \pi_\psi S$. For S, σ, N and η thus defined there exist γ, A and Θ with the properties specified in MC.

From (2.3), (2.4), Proposition 1.7 and the choice of N we have

$$\begin{aligned}\phi \circ \Theta(S) &\subset A^{-1}(\sigma(\phi(\bar{e}) + Y) + \pi_\phi S + Z + N) \\ &\subset A^{-1}((\operatorname{int} Z) + Z) = A^{-1}(\operatorname{int} Z) \subset \operatorname{int} Z.\end{aligned}$$

This implies that, for all $h \in S$, $\phi_0 \circ \Theta(h) < 0$ and $\tilde{\phi} \circ \Theta(h) < 0$.

Let $\tilde{\pi}_\psi : S \to \pi_\psi S$ be defined by $\tilde{\pi}_\psi(h) = \pi_\psi h$. Since S and $\pi_\psi S$ are k-simplexes, $\tilde{\pi}_\psi$ associates points with the same barycentric coordinates. Proposition 1.3 implies that $\tilde{\pi}_\psi^{-1} : \pi_\psi S \to S$ is continuous. Define $G : \pi_\psi S \to \mathbb{R}^k$ by $G(u) = -\gamma \psi \circ \Theta \circ \tilde{\pi}_\psi^{-1}(u) + u$. From (2.5) it follows that G is continuous and, from (2.6) and the choice of η, $G : \pi_\psi S \to \pi_\psi S$. Since $\pi_\psi S$ is compact and convex, the Brouwer fixed point theorem implies that there exists $u^* \in \pi_\psi S$ such that $G(u^*) = u^*$. Since $\gamma \neq 0$ it follows that $\psi(e^*) = 0$, where $e^* \triangleq \Theta(h^*) \in E$ and $h^* \triangleq \tilde{\pi}_\psi^{-1}(u^*) \in S$. Since e^* also satisfies $\phi_0(e^*) < 0$ and $\tilde{\phi}(e^*) < 0$, $\bar{e}$ does not solve OP, which is a contradiction.

The proof for $\tilde{\mathscr{X}} = \varnothing$ is almost identical to the above proof. It is only necessary to delete the details pertaining to $\tilde{\phi}$. In the case $k = 0$ the proof is considerably shorter. $\hat{Z}$ plays the role of $\hat{Z} \times \{0\}$ in Lemma 3.2 and, in Lemma 3.3, the k-simplex S is a singleton. The remainder of the proof proceeds along the same lines as the above proof except now the development relating to ψ is not needed.

4. A specialization of the Main Theorem. In this section we specialize the Main Theorem to obtain Theorem 4.1. For our purposes here, Theorem 4.1 may be regarded as a convenient intermediate step to the results of §§ 5 and 6. It may have independent interest, however, in other applications.

Theorem 4.1 is obtained by imposing additional structure on MT in three ways. First, we introduce an auxiliary vector space as discussed in Remark 2.2. This compensates for the lack of assumptions on $\mathscr{E}$. Second, we assume that the cost and constraint functions possess one-parameter polynomial expansions. These expansions may be regarded as a primitive form of the variational and power series required for the derivation of higher-order necessary conditions in §§ 5 and 6. Third, we impose additional structure on the inequality constraint (1.2). Specifically, we assume that there exists $j \in \mathbb{N}$ and, for each $i \in \{1, \cdots, j\}$, there exist a topological vector space $\mathscr{X}_i$, a mapping $\phi_i: \mathscr{E} \to \mathscr{X}_i$ and a solid closed convex cone $Z_i \subset \mathscr{X}_i$ not equal to $\mathscr{X}_i$ such that $\tilde{\mathscr{X}} = \mathscr{X}_1 \times \cdots \times \mathscr{X}_j$, $\tilde{\phi} = (\phi_1, \cdots, \phi_j)$ and $\tilde{Z} = Z_1 \times \cdots \times Z_j$. For $i \in \{1, \cdots, j\}$ and $z, \hat{z} \in \mathscr{X}_i$ we define $z \leqq \hat{z}$ and $z < \hat{z}$ in the obvious way. Now (1.2) becomes

$$\phi_i(e) \leqq 0, \qquad i \in \{1, \cdots, j\}. \tag{1.2$'$}$$

As discussed below, writing (1.2) in the form (1.2)′ extends the applicability of the Main Theorem.

Condition 4.1, which is used in Theorem 4.1, incorporates the above aspects. This rather complex condition may be motivated by the following comments. Roughly speaking, we assume that $\phi_0, \cdots, \phi_j$ have expansions of the form

$$\phi_i(e) = \phi_i(\bar{e}) + \sum_{r=1}^{m_i} \alpha^r Y_{ir} + o(\alpha^{m_i}), \tag{4.1}$$

where $m_i \in \mathbb{N} \cup \{0\}$, $Y_{ir} \in Z_i$ for all $r \in \{1, \cdots, m_i\}$, $\alpha > 0$ and $\alpha^{-m_i} o(\alpha^{m_i}) \to 0$ as $\alpha \to 0^+$. In the proof of Theorem 4.1 it turns out that if for all $\alpha > 0$ sufficiently small

$$\phi_i(\bar{e}) + \sum_{r=1}^{m_i} \alpha^r Y_{ir} < 0, \tag{4.2}$$

then the term $o(\alpha^{m_i})$ plays no role in the higher-order necessary conditions. This is advantageous if either: (1) ϕ_i possesses an expansion of order m_i and not of order $m_i + 1$, or (2) the term Y_{i,m_i+1} is not an element of Z_i. The form of the constraint (1.2)′ is used so that the order m_i of the expansion can depend on i.

By adding $\sum_{r=1}^{m_i} (1 - \alpha^r) Y_{ir}$ to the left side of (4.2) (assuming $\alpha < 1$) and using Proposition 1.7 it follows that (4.2) implies

$$\phi_i(\bar{e}) + \sum_{r=1}^{m_i} Y_{ir} < 0. \tag{4.3}$$

Similarly, it can be seen that the reverse implication is true. The equivalence of (4.2) and (4.3) accounts for the appearance of (4.3) in Condition 4.1 (via the set I' defined below). In the proof of Theorem 4.1 the ith component of Y in MT is given by $\sum_{r=1}^{m_i} Y_{ir}$.

Additional aspects of Condition 4.1 are: the "auxiliary" vector space $\mathscr{X}$, a convex set of "variations" $M \subset \mathscr{X}$, a mapping $V: M \to \mathscr{Z}'$ which provides a representation for K in MC (specifically, $K = \operatorname{co} V(M)$) and, for each k-simplex $X \subset M$, a mapping $\zeta: X \to E$ which corresponds to Θ in MC.

In what follows we use the conventions $\{a_i\}_{i=1}^0 = \varnothing$, $\sum_{i=1}^0 a_i = 0$ and let $j = 0$ when the inequality constraints (1.2)′ are absent. As in § 2 we assume for convenience that $\phi_0(\bar{e}) = 0$. For $\bar{e} \in E$ define the index sets $I' \triangleq \{i \in \{0, \cdots, j\}: \phi_i(\bar{e}) + \sum_{r=1}^{m_i} Y_{ir} < 0\}$ and $I'' \triangleq \{0, \cdots, j\}/I'$. Finally, let $l \in \mathscr{Z}'^*$ denote $l = (l_0, \cdots, l_j, l_\psi) \in \mathscr{Z}_0^* \times \cdots \times \mathscr{Z}_j^* \times \mathbb{R}^k$ when $k > 0$ and $l = (l_0, \cdots, l_j) \in \mathscr{Z}_0^* \times \cdots \times \mathscr{Z}_j^*$ when $k = 0$.

Condition 4.1. There exist a feasible element $\bar{e}$, a vector space $\mathscr{X}$, a nonempty convex set $M \subset \mathscr{X}$, $n \in \mathbb{N}$, $m_i \in \{0, \cdots, n-1\}$ and $\{Y_{ir}\}_{r=1}^{m_i} \subset Z_i$ for all $i \in \{0, \cdots, j\}$, and $V \triangleq (V_0, \cdots, V_j, V_\psi): M \to \mathscr{Z}'$ (omit V_ψ when $k = 0$) such that $V_i: M \to \mathscr{Z}_i$ is Z_i-convex for all $i \in I''$, $V_\psi: M \to \mathbb{R}^k$ is affine and such that the following property is satisfied. (Let $X = \operatorname{co}\{x_1, \cdots, x_{k+1}\} \subset M$ be a k-simplex, N_i be a neighborhood of the origin in $\mathscr{Z}_i$ for all $i \in \{0, \cdots, j\}$, $\eta > 0$, $\tau > 0$, $\alpha \in (0, \tau)$ and $\zeta: X \to E$.) For all $X, N_0, \cdots, N_j, \eta$ and τ there exist α and ζ satisfying

$$\alpha^{-m_i}\left[\phi_i \circ \zeta(X) - \phi_i(\bar{e}) - \sum_{r=1}^{m_i} \alpha^r Y_{ir}\right] \subset Z_i + N_i, \qquad i \in I', \tag{4.4}$$

$$\alpha^{-(m_i+1)}\left[\phi_i \circ \zeta(X) - \phi_i(\bar{e}) - \sum_{r=1}^{m_i} \alpha^r Y_{ir}\right] \subset (\operatorname{co} V_i(X)) + Z_i + N_i, \qquad i \in I'', \tag{4.5}$$

$$\boldsymbol{\mu} \to \psi \circ \zeta\left(\sum_{i=1}^{k+1} \mu_i x_i\right): \Delta^k \to \mathbb{R}^k \quad \text{is continuous} \quad (\text{omit if } k = 0), \tag{4.6}$$

$$|\alpha^{-n}\psi \circ \zeta(x) - V_\psi(x)| < \eta, \qquad x \in X \quad (\text{omit if } k = 0). \tag{4.7}$$

THEOREM 4.1. *Suppose Condition* 4.1 *is satisfied. If* $\bar{e}$ *solves* OP *then there exists* $l \in \mathscr{Z}'^*$ *such that*

$$l \neq 0, \tag{4.8}$$

$$l_i \in Z_i^{\ominus}, \qquad i \in \{0, \cdots, j\}, \tag{4.9}$$

$$l_i = 0, \qquad i \in I', \tag{4.10}$$

$$l_i\left(\phi_i(\bar{e}) + \sum_{r=1}^{m_i} Y_{ir}\right) = 0, \qquad i \in I'', \tag{4.11}$$

$$l(V(x)) \geqq 0, \qquad x \in M. \tag{4.12}$$

Remark 4.1. The additional structure of (1.2)′ accounts for the form of (4.10) and (4.11). Condition (4.10) generalizes the "complementary slackness" condition pertaining to the "inactive" inequality constraints ($\phi_i(\bar{e}) < 0$) in first-order necessary conditions. Note that $\phi_i(\bar{e}) + \sum_{r=1}^{m_i} Y_{ir} \neq 0$ for $i \in I''$ is possible when $\dim Z_i > 1$. In this way (4.11) may yield information regarding l_i.

Remark 4.2. By setting $n = 1$ in Condition 4.1 it is easy to see that Theorem 4.1 implies [29, Thm. 3.1] when this result is specialized to OP. Note that in Condition 4.1, X is a k-simplex whereas in [29, cond. 3.1], X is an i-simplex, where $i \in \{1, \cdots, k\}$. Furthermore, because inactive inequality constraints require only a trivial expansion of the form (4.4) (since $m_i = 0$), (4.4) and (4.5) allow a more general treatment of the inequality constraints than is possible by [29, (3.2)].

Proof of Theorem 4.1. We give the details for the case $k>0$ only since the proof for $k=0$ is similar. Since the functions V_i, $i\in I'$, play no role in Condition 4.1 or the results of the theorem, we define $V_i(x)=0$, $x\in M$, $i\in I'$. The main idea of the proof is to show that MC holds with

$$Y=\left(\sum_{r=1}^{m_0} Y_{0r},\cdots,\sum_{r=1}^{m_j} Y_{jr}\right) \tag{4.13}$$

and

$$K=\operatorname{co} V(M). \tag{4.14}$$

To show that (2.7)–(2.10) imply (4.8)–(4.12), note that (4.9) is merely a rewriting of (2.8) and that (4.11) and (4.12) follow from (2.9) and (2.10), respectively. Also, (4.10) follows from (2.9), (4.9) and Propositions 1.6 and 1.8.

We now show that Condition 4.1 implies MC. Let $S=\operatorname{co}\{h_1,\cdots,h_{k+1}\}$, σ, N and η have the properties specified in MC. Since $h_i\in K$, $i\in\{1,\cdots,k+1\}$, and $K=\operatorname{co} V(M)$, we have $h_i=\sum_{r=1}^{\nu_i}\mu_{ir}V(x_{ir})$, where $\nu_i\in\mathbb{N}$, $(\mu_{i1},\cdots,\mu_{i,\nu_i})\in\Delta^{\nu_i-1}$ and $x_{i1},\cdots,x_{i,\nu_i}\in M$. Define $x_i\triangleq\sum_{r=1}^{\nu_i}\mu_{ir}x_{ir}$, $i\in\{1,\cdots,k+1\}$, and $X\triangleq\operatorname{co}\{x_1,\cdots,x_{k+1}\}$. Since V_ψ is affine,

$$V_\psi(x_i)=\pi_\psi h_i,\qquad i\in\{1,\cdots,k+1\}. \tag{4.15}$$

From (2.1) it follows that $\pi_\psi h_1,\cdots,\pi_\psi h_{k+1}$ are affinely independent. Proposition 1.2 and (4.15) thus imply that X is a k-simplex. For each $i\in\{0,\cdots,j\}$ choose N_i to be a neighborhood of the origin in $\mathscr{Z}_i$ so that

$$\sigma(N_0\times\cdots\times N_j)\subset N. \tag{4.16}$$

Also, let $\tau\triangleq\sigma^{-1}$.

With X, $N_0,\cdots,N_j$, ζ and τ now specified Condition 4.1 implies that there exist $\alpha\in(0,\tau)$ and $\zeta:X\to E$ satisfying (4.4)–(4.7). We now exhibit γ, A and Θ so that (2.3)–(2.6) are satisfied.

Let $\Theta: S\to E$ be defined by $\Theta=f\circ g$ where $g:S\to\Delta^k$, $f:\Delta^k\to E$, $g(\sum_{i=1}^{k+1}\mu_i h_i)\triangleq\boldsymbol{\mu}$ and $f(\boldsymbol{\mu})\triangleq\zeta(\sum_{i=1}^{k+1}\mu_i x_i)$. From (4.6) it follows that $\psi\circ f:\Delta^k\to\mathbb{R}^k$ is continuous. Since, by Proposition 1.3, g is continuous, $\psi\circ\Theta=\psi\circ f\circ g: S\to\mathbb{R}^k$ is continuous. This proves (2.5). For use below note that

$$\Theta\left(\sum_{i=1}^{k+1}\mu_i h_i\right)=\zeta\left(\sum_{i=1}^{k+1}\mu_i x_i\right),\qquad \boldsymbol{\mu}\in\Delta^k, \tag{4.17}$$

$$\Theta(S)=\zeta(X). \tag{4.18}$$

For $h\in Z_0\times\cdots\times Z_j\times\mathbb{R}^k$, $H\subset Z_0\times\cdots\times Z_j\times\mathbb{R}^k$ and $i\in\{0,\cdots,j\}$ let $\pi_i h\triangleq$ component of h in Z_i and $\pi_i H\triangleq\{\pi_i h: h\in H\}$.

Since V_i is Z_i-convex, $i\in I''$, we have

$$V_i(x_p)\leqq\sum_{r=1}^{\nu_p}\mu_{pr}V_i(x_{pr})=\pi_i h_p,\qquad p\in\{1,\cdots,k+1\}, \tag{4.19}$$

which implies that, for all $\boldsymbol{\mu}\in\Delta^k$,

$$V_i\left(\sum_{p=1}^{k+1}\mu_p x_p\right)\leqq\sum_{p=1}^{k+1}\mu_p V_i(x_p)\leqq\sum_{p=1}^{k+1}\mu_p\pi_i h_p\in\pi_i S. \tag{4.20}$$

Consequently,

$$\text{(4.21)} \qquad \operatorname{co} V_i(X) \subset \pi_i S + Z_i, \qquad i \in I''.$$

To simplify what follows we assume without loss of generality that $\sigma > 1$. To see that this is possible note that if $\sigma \leqq 1$ then we can add $\rho(\phi(\bar{e}) + Y)$, where $\rho > 1 - \sigma$, to each side of (2.2). Since $\rho(\phi(\bar{e}) + Y) \in Z$, Proposition 1.7 can be used to replace $(\operatorname{int} Z) + \rho(\phi(\bar{e}) + Y)$ by $\operatorname{int} Z$. Then $\sigma + \rho$ can be relabelled as σ to yield the desired result. Note that $\sigma > 1$ implies $\alpha < 1$.

Since Z_i is a convex cone it follows that, for all $i \in \{0, \cdots, j\}$,

$$\text{(4.22)} \qquad \alpha_0 \phi_i(\bar{e}) + \sum_{r=1}^{m_i} \alpha_i Y_{ir} \in Z_i, \qquad \alpha_0, \cdots, \alpha_{m_i} > 0.$$

From (4.13) it follows that

$$\text{(4.23)} \qquad \pi_i Y = \sum_{r=1}^{m_i} Y_{ir}, \qquad i \in \{0, \cdots, j\},$$

and, since $V_i(x) = 0$, $i \in I'$, we have

$$\text{(4.24)} \qquad \pi_i S = \{0\}, \qquad i \in I'.$$

Since $\sigma\alpha^{-r} \geqq \sigma$, $r \in \{0, 1, \cdots\}$, (4.4), (4.18), (4.22), (4.23) and (4.24) imply

$$\text{(4.25)} \qquad \sigma\alpha^{-m_i}\phi_i \circ \Theta(S) \subset \sigma(\phi_i(\bar{e}) + \pi_i Y) + \pi_i S + Z_i + \sigma N_i, \qquad i \in I'.$$

Since $\alpha^{-r} > \sigma$, $r \in \mathbb{N}$, (4.5), (4.18), (4.21), (4.22) and (4.23) imply

$$\text{(4.26)} \qquad \alpha^{-(m_i+1)}\phi_i \circ \Theta(S) \subset \sigma(\phi_i(\bar{e}) + \pi_i Y) + \pi_i S + Z_i + N_i, \qquad i \in I''.$$

For $z \triangleq (z_0, \cdots, z_j) \in Z$ define $A(z) \triangleq (a_0 z_0, \cdots, a_j z_j)$, where $a_i \triangleq \sigma\alpha^{-m_i}$, $i \in I'$, and $a_i \triangleq \alpha^{-(m_i+1)}$, $i \in I''$, which satisfies (2.3). Condition (2.4) follows from (4.16), (4.25), (4.26) and the fact that $N_i \subset \sigma N_i$. Since V_ψ is affine, (4.15) implies

$$\text{(4.27)} \qquad V_\psi\left(\sum_{i=1}^{k+1} \mu_i x_i\right) = \pi_\psi \sum_{i=1}^{k+1} \mu_i h_i, \qquad \boldsymbol{\mu} \in \Delta^k.$$

Letting $\gamma \triangleq \alpha^{-n}$, (4.7), (4.17) and (4.27) imply (2.6), which completes the proof. □

Remark 4.3. If $k > 0$ and $\dim M < k$ then M does not contain a k-simplex and Condition 4.1 holds trivially. In this case the necessary conditions (4.8)–(4.12) can be satisfied by choosing $l = (0, \cdots, 0, l_\psi)$ where $l_\psi \in (-V_\psi(M))^\ominus$ and $l_\psi \neq 0$. Such l_ψ exists since by Proposition 1.2 $\dim V_\psi(M) < k$.

Remark 4.4. Suppose Condition 4.1 is satisfied and, in addition, $I'' \neq \varnothing$, M is a cone and V is positively homogeneous and affine. Then Theorem 4.1 remains valid with a slightly weakened version of Condition 4.1 in which $x_1, \cdots, x_{k+1}$ are assumed to be linearly independent. To see this, note first that $I'' \neq \varnothing$ is equivalent to (2.13) with (4.13). By (2.14) $h_1, \cdots, h_{k+1}$ are linearly independent. Thus Proposition 1.2 and the fact that $V(x_i) = h_i$, $i \in \{1, \cdots, k+1\}$, imply that $x_1, \cdots, x_{k+1}$ are linearly independent. This weakened version of Condition 4.1 is thus satisfied trivially when M does not contain $k+1$ linearly independent elements.

5. Applications involving directional differentials. We now assume that the mappings $\phi_0, \cdots, \phi_j$ and ψ satisfy certain first- and second-order one-sided differentiability conditions. These assumptions lead to the principal results of this section, Theorems 5.1 and 5.2, which contain the first- and second-order necessary conditions for OP. A rather extensive investigation of the relationship between these results leads to Theorem 5.4 which generalizes [35, Thm. 2.3]. The proof of Theorem 5.2 is then given along with remarks pointing out how the results of this section can be generalized.

In order to state the differentiability assumptions for the cost and constraint functions, it is necessary to introduce a generalization of the Fréchet derivative. This definition, which is due to Warga [33, p. 167] allows a function to have a derivative at a point which is not in the interior of its domain. Some consequences of this definition needed in the proofs of this section are discussed in the Appendix.

DEFINITION 5.1. Let $\hat{\mathscr{X}}$ be a Banach space with norm $|\cdot|$, $\mathscr{Y}$ be a topological vector space, $\hat{A} \subset \hat{\mathscr{X}}$ and $f: \hat{A} \to \mathscr{Y}$. f is $\tilde{F}$-differentiable at $\bar{x} \in \hat{A}$ and has the $\tilde{F}$-derivative $f^{(1)}(\bar{x}) \triangleq f'(\bar{x}) \in \mathscr{B}(\hat{\mathscr{X}}; \mathscr{Y})$ if $\bar{x}$ is contained in a solid convex subset of $\hat{A}$ and

$$\lim_{\substack{x \to \bar{x} \\ x \in \hat{A}/\{\bar{x}\}}} |x - \bar{x}|^{-1}[f(x) - f(\bar{x}) - f'(\bar{x})(x - \bar{x})] = 0. \tag{5.1}$$

If $f'(x)$ exists for all $x \in \hat{A}$ then f is $\tilde{F}$-differentiable.

The second derivative requires a topology on $\mathscr{B}(\hat{\mathscr{X}}; \mathscr{Y})$. This is handled by assuming that $\mathscr{Y}$ is a Banach space and defining a norm on $\mathscr{B}(\hat{\mathscr{X}}; \mathscr{Y})$ as in § 1.

DEFINITION 5.2. Let $\hat{\mathscr{X}}$, $\hat{A}$, $\mathscr{Y}$ and f be as in Definition 5.1 and assume further that $\mathscr{Y}$ is a Banach space. f is twice $\tilde{F}$-differentiable at $\bar{x} \in \hat{A}$ and has the second $\tilde{F}$-derivative $f^{(2)}(\bar{x}) \triangleq f''(\bar{x}) \in \mathscr{B}(\hat{\mathscr{X}}; \mathscr{B}(\hat{\mathscr{X}}; \mathscr{Y}))$ if f is $\tilde{F}$-differentiable and the mapping $x \to f'(x): \hat{A} \to \mathscr{B}(\hat{\mathscr{X}}; \mathscr{Y})$ is $\tilde{F}$-differentiable at $\bar{x}$.

First- and second-order one-sided directional differentials appear in the theorem statements. Their definition is a simple application of the preceding definitions.

DEFINITION 5.3. Let $\mathscr{X}$ be a vector space, $\mathscr{Y}$ be a topological vector space, $A \subset \mathscr{X}$, $F: A \to \mathscr{Y}$, $\bar{x} \in A$ and $h \in \mathscr{X}$. Suppose that there exists $\beta > 0$ such that $\bar{x} + \alpha h \in A$, $\alpha \in [0, \beta)$, and define $f: [0, \beta) \to \mathscr{Y}$ by $f(\alpha) = F(\bar{x} + \alpha h)$. If $f'(0)$ exists then $DF(\bar{x}; h) \triangleq f'(0)$ is the one-sided directional differential of f at $\bar{x}$ in the direction h. If $\mathscr{Y}$ is a Banach space and $f''(0)$ exists, then $D^2F(\bar{x}; h) \triangleq f''(0)$ is the second-order one-sided directional differential of F at $\bar{x}$ in the direction h.

Note that $\hat{A}$ in Definition 5.1 is given by $[0, \beta)$ in Definition 5.3. It can be seen that if $DF(\bar{x}; h)$ exists then $DF(\bar{x}; \alpha h)$ exists for all $\alpha > 0$. Thus, although $DF(\bar{x}; h)$ as defined is an element of $\mathscr{B}(\mathbb{R}; \mathscr{Y})$, we regard $DF(\bar{x}; \cdot)$ as a map from cone h into $\mathscr{Y}$.

The following notation concerning a feasible element $\bar{e}$ simplifies the statement of what follows. If $j > 0$ let $I_N \triangleq \{i \in \{1, \cdots, j\}: \phi_i(\bar{e}) < 0\}$, $I_A \triangleq \{1, \cdots, j\}/I_N$ and $I_{A0} \triangleq I_A \cup \{0\}$; if $j = 0$ let $I_N \triangleq I_A \triangleq \varnothing$ and $I_{A0} \triangleq \{0\}$. Note that if $j > 0$ and $\dim Z_i > 1$ for some $i \in \{1, \cdots, j\}$, then I_A does not necessarily coincide with the set $\{i \in \{1, \cdots, j\}: \phi_i(\bar{e}) = 0\}$. We define

$$\Phi \triangleq \begin{cases} (\phi_0, \cdots, \phi_j, \psi), & k > 0, \\ (\phi_0, \cdots, \phi_j), & k = 0. \end{cases} \tag{5.2}$$

Finally, we recall from § 4 the meaning of the notation $l \in \mathscr{L}'^*$ and the convention $\phi_0(\bar{e}) = 0$.

Condition 5.1. E is a subset of a vector space $\mathscr{X}$ and there exist a feasible element $\bar{e}$ and a nonempty convex subset M of $\mathscr{X}$ such that the following property is satisfied.

For each $X = \mathrm{co}\{x_1, \cdots, x_{k+1}\} \subset M$ there exists $\beta > 0$ such that

$$\bar{e} + \alpha X \subset E, \qquad \alpha \in [0, \beta), \tag{5.3}$$

$$\boldsymbol{\mu} \to \psi\left(\bar{e} + \alpha \sum_{i=1}^{k+1} \mu_i x_i\right) : \Delta^k \to \mathbb{R}^k \quad \text{is continuous,} \quad \alpha \in (0, \beta) \quad (\text{omit if } k = 0), \tag{5.4}$$

$$\boldsymbol{\alpha} \to \Phi\left(\bar{e} + \sum_{i=1}^{k+1} \alpha_i x_i\right) : \mathbb{P}^{k+1}(\beta) \to \mathscr{Z}' \quad \text{is } \tilde{F}\text{-differentiable at } \boldsymbol{\alpha} = 0. \tag{5.5}$$

Furthermore, if $k = 0$ then

$$x \to D\phi_i(\bar{e}; x) : M \to \mathscr{Z}_i \quad \text{is } Z_i\text{-convex,} \quad i \in \{0, \cdots, j\}. \tag{5.6}$$

THEOREM 5.1. *Suppose Condition* 5.1 *is satisfied. If* $\bar{e}$ *solves* OP *then there exists* $l \in \mathscr{Z}'^*$ *such that*

$$l \neq 0, \tag{5.7}$$

$$l_i \in Z_i^{\ominus}, \qquad i \in \{0, \cdots, j\}, \tag{5.8}$$

$$l_i = 0, \qquad i \in I_N, \tag{5.9}$$

$$l_i(\phi_i(\bar{e})) = 0, \qquad i \in I_A, \tag{5.10}$$

$$l(D\Phi(\bar{e}; x)) \geqq 0, \qquad x \in M. \tag{5.11}$$

Condition 5.2. $\mathscr{Z}_0, \cdots, \mathscr{Z}_j$ are Banach spaces, $\mathscr{E}$ is a subset of a vector space $\mathscr{X}$ and there exist a feasible element $\bar{e}$, a vector $y \in \mathscr{X}$ and a nonempty convex set $M' \subset \mathscr{X}$ such that the following property is satisfied. For each $X = \mathrm{co}\{x_1, \cdots, x_{k+1}\} \subset M'$ there exists $\beta > 0$ such that

$$\bar{e} + \alpha_1 y + \alpha_2 X \subset \mathscr{E}, \qquad (\alpha_1, \alpha_2) \in \mathbb{P}^2(\beta), \tag{5.12}$$

$$\bar{e} + \alpha y + \alpha^2 X \subset E, \qquad (\alpha, \alpha^2) \in \mathbb{P}^2(\beta), \tag{5.13}$$

$$\boldsymbol{\alpha} \to \Phi\left(\bar{e} + \alpha_1 y + \sum_{i=2}^{k+2} \alpha_i x_{i-1}\right) : \mathbb{P}^{k+2}(\beta) \to \mathscr{Z}' \tag{5.14}$$

is twice $\tilde{F}$-differentiable at $\boldsymbol{\alpha} = 0$.

Furthermore,

$$D\phi_i(\bar{e}; y) \leqq 0, \qquad i \in I_{A0}, \tag{5.15}$$

$$D\psi(\bar{e}; y) = 0 \quad (\text{omit if } k = 0) \tag{5.16}$$

and, if $k = 0$, (5.6) is satisfied with M replaced by M'.

THEOREM 5.2. *Suppose Condition* 5.2 *is satisfied. If* $\bar{e}$ *solves* OP *then there exists* $l \in \mathscr{Z}'^*$ *such that*

$$l \neq 0, \tag{5.17}$$

$$l_i \in Z_i^{\ominus}, \qquad i \in \{0, \cdots, j\}, \tag{5.18}$$

$$l_i = 0, \qquad i \in I_0' \triangleq I_N \cup \{i \in I_{A0} : \phi_i(\bar{e}) + D\phi_i(\bar{e}; y) < 0\}, \tag{5.19}$$

$$l_i(\phi_i(\bar{e}) + D\phi_i(\bar{e}; y)) = 0, \qquad i \in I_0'' \triangleq \{0, \cdots, j\}/I_0', \tag{5.20}$$

$$l(D\Phi(\bar{e}; x) + \tfrac{1}{2}D^2\Phi(\bar{e}; y)) \geqq 0, \qquad x \in M'. \tag{5.21}$$

The proofs of Theorems 5.1 and 5.2, given at the end of this section, are based on Theorem 4.1. In brief, the reasons for Conditions 5.1 and 5.2 are as follows. Items (5.3) and (5.12) set up the differentiability domains for Φ. Theorem A.1 (see Appendix), a version of Taylor's theorem, is used with (5.5) and (5.14) to obtain expansions for Φ. These expansions in the variable α can be reduced to one-parameter expansions by means of Proposition A.1. The correct properties for the intermediate terms of these expansions follow from (5.15) and (5.16). In this way (4.4), (4.5), and (4.7) are obtained. The purpose of (5.4) is to guarantee (4.6). In Theorem 5.2, (5.14) implies (4.6). Condition (5.6) is required so that when $k=0$ the map V in Condition 4.1, which is given by $V(x) \triangleq D\Phi(\bar{e}; x)$, is Z-convex. No assumption analogous to (5.6) is needed when $k>0$ since (5.5) and Theorem A.2 imply that V in Condition 4.1 is affine. Similar remarks apply to Condition 5.2. In the proof of Theorem 5.1 $\zeta(x) = e+\alpha x$ which, by (5.3), is an element of E. Theorem 5.2 involves feasible elements of the form $\zeta(x) = \bar{e}+\alpha y+\alpha^2 x$ which, according to (5.13), are in E.

Remark 5.1. Because the origin of $\mathbb{R}^n$ does not lie in the interior of $\mathbb{P}^n(\beta)$, (5.5) and (5.14) rely on Definitions 5.1 and 5.2 instead of the classical definition of the Fréchet derivatives. Since $\mathbb{P}^n(\beta) \subset \mathbb{B}^n(\beta)$, weaker versions of Theorems 5.1 and 5.2 are obtained by replacing "$\mathbb{P}$" by "$\mathbb{B}$" in (5.5) and (5.14). Proofs of these weaker results depend on the classical definition of the Fréchet derivative along with a classical Taylor theorem result. In this case, Theorem A.1 is not needed and a weakened version of Proposition A.1 suffices.

The necessary conditions of Theorems 5.1 and 5.2 are not necessarily satisfied by a common l. Thus, one should not jump to conclusions about the relationship between the first- and second-order results. Much of the development in the remainder of this section involves an examination of this issue. Of particular importance is the relationship between M and E in Theorem 5.1 and among y, M', E and $\mathscr{E}$ in Theorem 5.2. The latter case is more complex because of the "quadratic" nature of (5.13). For example, y and M' may be chosen to characterize features of E such as a curved boundary.

The following two examples will be useful for illustrative purposes. It is easy to see that (5.12) and (5.13) are satisfied in both cases.

Example 5.1. $\mathscr{X} = \mathscr{E} = \mathbb{R}^2$, $k \in \mathbb{N}$, $\bar{e} = (0, 0)$, $y = (1, 0)$, $M' = \{0\} \times [1, 2]$, $E = \{(s_1, s_2) \in \mathbb{R}^2 : 0 \leqq s_1 \leqq 1,\ s_1^2 \leqq s_2 \leqq 2s_1^2\}$.

Example 5.2. $\mathscr{X}$, $\mathscr{E}$, k, $\bar{e}$, y as in Example 5.1, $M' = \{(s_1, s_2) \in \mathbb{R}^2 : s_1 \leqq 0,\ 1-s_1 \leqq s_2 \leqq 2-2s_1\}$, $E = \{(s_1, s_2) \in \mathbb{R}^2 : s_1^2 \leqq s_2,\ s_2^2 \leqq s_1\}$.

Remark 5.2. Note that if Condition 5.1 holds then it remains valid if M is replaced by cone M. Thus, without loss of generality, M can be assumed to be a convex cone. Examples 5.1 and 5.2 show that it is not always possible to replace M' by cone M' in Condition 5.2 since (5.13) may not be satisfied.

Example 5.1 shows that Condition 5.2 does not imply that Condition 5.1 is meaningfully satisfied. For this example the only set M satisfying (5.3) is $M = \{(0, 0)\}$ which yields trivial necessary conditions. Even when Conditions 5.1 and 5.2 are both meaningfully satisfied, M and M' may be disjoint as in Example 5.2 where M must be a subset of $\mathbb{R}^2_+ \cup \{(0, 0)\}$. Thus, it is not surprising that there may be no l satisfying both (5.7)–(5.11) and (5.17)–(5.21). The following assumption about the structure of M' leads to the existence of a common l. Consider

$$M' = Q + R, \quad Q \text{ convex}, \quad R \text{ a convex cone}. \tag{5.22}$$

We use (5.22) to obtain the following result.

PROPOSITION 5.1. *Suppose the assumptions of Theorem* 5.2 *are valid with* M' *given by* (5.22). *If* $k>0$ *then* (5.21) *is equivalent to*

$$l(D\Phi(\bar{e};x))\geqq 0, \qquad x\in R, \tag{5.23}$$

$$l(D\Phi(\bar{e};x)+\tfrac{1}{2}D^2\Phi(\bar{e};y))\geqq 0, \qquad x\in Q. \tag{5.24}$$

If $k=0$ *then* (5.21) *implies* (5.23) *and* (5.24). *The converse in this case is valid if either* $Q=\{0\}$ *or the map* $x\to D\Phi(\bar{e};x):M'\to\mathscr{X}$ *is affine.*

Proof. Define $V_1\triangleq -(D\Phi(\bar{e};Q)+\frac{1}{2}D^2\Phi(\bar{e};y))$, $V_2\triangleq -(D\Phi(\bar{e};R))$ and $V_3\triangleq -(D\Phi(\bar{e};M')+\frac{1}{2}D^2\Phi(\bar{e};y))$ and note that (5.24), (5.23) and (5.21) are equivalent to $l\in V_1^{\ominus}$, $l\in V_2^{\ominus}$ and $l\in V_3^{\ominus}$, respectively. For the case $k>0$ we have $V_1+V_2=V_3$ since $D\Phi(\bar{e};\cdot)$ is positively homogeneous and affine (see Theorem A.2). Since V_2 is a cone, Proposition 1.9 implies $V_1^{\ominus}\cap V_2^{\ominus}=V_3^{\ominus}$, as desired. For the case $k=0$ it can be shown that (5.6) with M replaced by M' implies $V_1+V_2\subset V_3+Z$. Thus, $(V_3+Z)^{\ominus}\subset(V_1+V_2)^{\ominus}$ and, by Proposition 1.9, $V_3^{\ominus}\cap Z^{\ominus}\subset V_1^{\ominus}\cap V_2^{\ominus}$. Since (5.18) is satisfied, (5.21) implies (5.23) and (5.24). If either $Q=\{0\}$ or $D\Phi(\bar{e};\cdot)$ is affine, $V_1+V_2=V_3$ and thus $V_1^{\ominus}\cap V_2^{\ominus}=V_3^{\ominus}$, which completes the proof. □

Notice that (5.22) can be satisfied trivially with $Q=M'$ and $R=\{0\}$. However, this does not produce new results. Situations of general interest are: $Q=\{0\}$, $R=M'$ (i.e., when M' is a cone) and $Q\neq\{0\}$, $R\neq\{0\}$ (e.g., Example 5.2, where $Q=\{0\}\times[1,2]$ and $R=\{(s_1,s_2)\in\mathbb{R}^2: -s_1\leqq s_2\leqq -2s_1\}$). The following result is a consequence of Theorem 5.2 and Proposition 5.1.

THEOREM 5.3. *Suppose Condition* 5.2 *is satisfied with* M' *given by* (5.22) *and let* $M\subset R$. *If* $\bar{e}$ *solves* OP *then each* l *satisfying* (5.17)–(5.21) *also satisfies* (5.7)–(5.11).

Proof. By Proposition 5.1, (5.21) implies (5.23) and (5.24). Since $M\subset R$, (5.23) implies (5.11). Condition (5.9) follows immediately from (5.19); (5.10) follows from (5.19) and (5.20) using (5.15) and (5.18). □

Theorem 5.3 shows that if the set M' in Condition 5.2 is given by (5.22) and M is an arbitrary convex subset of R, then the second-order necessary conditions (5.17)–(5.20), (5.23) and (5.24) imply the first-order necessary conditions (5.7)–(5.11). The strongest version of (5.7)–(5.11) is obtained when $M=R$. However, these "first-order necessary conditions" may be unobtainable from Theorem 5.1 because Condition 5.1 may not hold for certain choices of $M\subset R$. For instance, in Example 5.2 it is necessary to choose $M=\{(0,0)\}$ in order to satisfy both (5.3) and the requirement $M\subset R$. The reason Theorem 5.2 gives stronger "first-order necessary conditions" than Theorem 5.1 is that (5.13) takes the curvature of E into account while (5.3) does not. Note that y must be nonzero since otherwise (5.13) and (5.3) are equivalent.

Remark 5.3. The observation that second-order necessary conditions may contain "first-order necessary conditions" that are stronger than actual first-order necessary conditions has been made previously in [28, Remark 2, p. 278]. In their treatment of OP $j=0$, $\mathscr{X}=\mathbb{R}^{\nu}$, ϕ_0 and ψ are continuously differentiable and it is assumed from the outset that (5.22) holds with $Q=\{q\}$.

Remark 5.4. The relationship between the first- and second-order necessary conditions can be thought of in the following way. If Condition 5.1 is satisfied, then the first-order necessary conditions involve finding an element l in $\mathscr{X}'^*$ satisfying (5.7)–(5.11). If, in addition, Condition 5.2 is satisfied with (5.22) and $M\subset R$, then the second-order necessary conditions are equivalent to the existence of an element l in $\mathscr{X}'^*$ satisfying both (5.7)–(5.11) and the additional conditions (5.19), (5.20), (5.23) and (5.24). In this case, Theorem 5.2 supplements Theorem 5.1 and verification of the necessary conditions can be thought of as a two-stage process.

Remark 5.5. Suppose Conditions 5.1 and 5.2 are satisfied with (5.22) and $M \subset R$ and define

$$\Gamma \triangleq \{l \in \mathscr{X}'^*: l \text{ satisfies } (5.7)\text{–}(5.11)\}. \tag{5.25}$$

Remark 5.4 shows that it is of particular interest to determine whether Γ is a ray. Specifically, if Γ is a ray then the search for l satisfying the second-order necessary conditions is simplified since l is determined uniquely to within a scalar multiple by the first-order necessary conditions alone.

We now consider a subset of M that is useful in verifying the first-order necessary conditions and understanding the relationship between Theorems 5.1 and 5.2. Define

$$\mathscr{D} \triangleq \{y \in M: y \text{ satisfies } (5.15) \text{ and } (5.16)\}, \tag{5.26}$$

$$\mathscr{I}(y) \triangleq \{i \in I_{A0}: \phi_i(\bar{e}) + D\phi_i(\bar{e}; y) < 0\}, \tag{5.27}$$

where $y \in \mathscr{D}$, and

$$\mathscr{I} \triangleq \bigcup_{y \in \mathscr{D}} \mathscr{I}(y). \tag{5.28}$$

PROPOSITION 5.2. *If l satisfies* (5.7)–(5.11) *then l also satisfies*

$$l_i = 0, \qquad i \in I_N \cup \mathscr{I}, \tag{5.29}$$

and

$$l_i(\phi_i(\bar{e}) + D\phi_i(\bar{e}; y)) = 0, \qquad i \in \{0, \cdots, j\}/(I_N \cup \mathscr{I}), \quad y \in \mathscr{D}. \tag{5.30}$$

Proof. Since $\mathscr{D} \subset M$, set $x = y \in \mathscr{D}$ in (5.11). From (5.16) we obtain

$$\sum_{i \in I_{A0}} l_i(D\phi_i(\bar{e}; y)) \geqq 0. \tag{5.31}$$

It follows from (5.31), (5.8) and (5.15) that $l_i(D\phi_i(\bar{e}; y)) = 0$, $i \in I_{A0}$. The result now follows from (5.10) and Proposition 1.8. □

Remark 5.6. Proposition 5.2 sharpens Remark 5.4 in the following way. Suppose Conditions 5.1 and 5.2 are satisfied with (5.22) and $M \subset R$. If y in Condition 5.2 is an element of M (which implies $y \in \mathscr{D}$) then (5.19) and (5.20) are a consequence of (5.29) and (5.30) and thus do not strengthen (5.7)–(5.11).

Remark 5.7. The ideas used in the proof of Proposition 5.2 could also be used in other contexts in the optimization literature to show that additional multiplier components are zero. However, this approach does not appear to have been used before.

We now consider some consequences of strengthening Condition 5.1. Sometimes the set E is sufficiently "large" near $\bar{e}$ so that

$$\bar{e} + \alpha_1 y + \alpha_2 X \subset E, \qquad (\alpha_1, \alpha_2) \in \mathbb{P}^2(\beta). \tag{5.32}$$

This strengthening of Condition 5.2 is formalized in Condition 5.3.

Condition 5.3. Condition 5.2 is satisfied with (5.12) replaced by (5.32) and with (5.13) omitted.

The following results are easily verified.

PROPOSITION 5.3. *If Condition* 5.3 *is satisfied then it is also satisfied with M' replaced by* cone M'.

PROPOSITION 5.4. *If Condition* 5.3 *is satisfied, then Condition* 5.1 *is satisfied with $M = M'$.*

The importance of Condition 5.3 lies in Proposition 5.4 which guarantees that nontrivial first-order necessary conditions can be obtained independently of the second-order necessary conditions. Proposition 5.3 shows that (5.22) is satisfied with $R = \operatorname{cone} M'$ and $Q = \{0\}$.

Sometimes Conditions 5.2 and 5.3 are equivalent. For example, this occurs if either $\mathscr{E} = E$ or $y = 0$. A more interesting case is contained in the following proposition which leads ultimately to Theorem 5.4, a simply stated specialization of Theorem 5.1 and 5.2.

PROPOSITION 5.5. *Suppose Condition* 5.2 *is satisfied with* (5.13) *deleted and with the additional conditions*

$$0 \in M', \tag{5.33}$$

$$y \in M', \tag{5.34}$$

$$\mathscr{E} \cap (M' + \bar{e}) \subset E. \tag{5.35}$$

Then Condition 5.3 *is satisfied.*

Proof. It suffices to show that (5.33)–(5.35) and (5.12) imply (5.32). Let $\beta \in (0, 1)$ and $X = \operatorname{co}\{x_1, \cdots, x_{k+1}\} \subset M'$. From (5.33) and (5.34) it follows that $\{\bar{e}, \bar{e} + \beta y, \bar{e} + \beta x_1, \cdots, \bar{e} + \beta x_{k+1}\} \subset M' + \bar{e}$. Thus $\operatorname{co}\{\bar{e}, \bar{e} + \beta y, \bar{e} + \beta x_1, \cdots, \bar{e} + \beta x_{k+1}\} \subset M' + \bar{e}$ which is equivalent to $\bar{e} + \alpha_1 y + \alpha_2 X \subset M' + \bar{e}$, $(\alpha_1, \alpha_2) \in \mathbb{P}^2(\beta)$. The desired result now follows from (5.12) and (5.35). □

It is easy to verify that the following conditions imply Conditions 5.1 and 5.2, respectively.

Condition 5.4. E is a subset of a vector space $\mathscr{X}$ and there exist a feasible element $\bar{e}$ and a nonempty convex subset M of $\mathscr{X}$ such that $0 \in M$ and $\mathscr{E} \cap (M + \bar{e}) \subset E$ and such that the following property is satisfied. For each $X = \operatorname{co}\{x_1, \cdots, x_{k+1}\} \subset M$ there exists $\beta > 0$ such that

$$\bar{e} + \alpha X \subset \mathscr{E}, \qquad \alpha \in [0, \beta), \tag{5.36}$$

and (5.4) and (5.5) are satisfied. Furthermore, if $k = 0$ then (5.6) holds.

Condition 5.5. Condition 5.4 is satisfied, $\mathscr{Z}_0, \cdots, \mathscr{Z}_j$ are Banach spaces and the following property is satisfied. For each $X = \operatorname{co}\{x_1, \cdots, x_{k+2}\} \subset M$ there exists $\beta > 0$ satisfying (5.36) and such that

$$\boldsymbol{\alpha} \to \Phi\left(\bar{e} + \sum_{i=1}^{k+2} \alpha_i x_i\right) : \mathbb{P}^{k+2}(\beta) \to \mathscr{Z}' \tag{5.37}$$

is twice $\tilde{F}$-differentiable at $\boldsymbol{\alpha} = 0$.

THEOREM 5.4. *Suppose $\bar{e}$ solves* OP. *If Condition* 5.4 *is satisfied then there exists $l \in \mathscr{Z}'^*$ satisfying* (5.7)–(5.11). *If, furthermore, Condition* 5.5 *is satisfied, then for each $y \in \mathscr{D}$ there exists $l \in \mathscr{Z}'^*$ satisfying* (5.7)–(5.11) *and*

$$l(D^2\Phi(\bar{e}; y)) \geqq 0. \tag{5.38}$$

Proof. Arguments similar to those used to prove Proposition 5.5 show that the conditions on M and (5.36) imply (5.3). Thus, the first part of the theorem follows from Theorem 5.1. To prove the second part of the theorem let $X = \operatorname{co}\{x_1, \cdots, x_{k+1}, y\} \subset M$, where $y \in \mathscr{D}$, and note that (5.36) with this choice of X is equivalent to (5.12). Thus, Condition 5.2 is satisfied with (5.13) deleted. Since also (5.33)–(5.35) (with $M' = M$) are satisfied, Proposition 5.5 implies that Condition 5.3

must hold with $M' = M$. By Proposition 5.3, Condition 5.3 is satisfied with M replaced by cone M. Since Condition 5.3 implies Condition 5.2, it follows from Theorem 5.2 and Proposition 5.1 (with $R = \text{cone}\, M$ and $Q = \{0\}$) that there exists $l \in \mathscr{Z}'^*$ satisfying (5.17)–(5.20), (5.11) and (5.38). Finally, Proposition 5.2 implies that l satisfying (5.7)–(5.11) must also satisfy (5.19) and (5.20). Thus, these last two conditions have been omitted. □

Remark 5.8. Remark 5.5 takes on added importance in the context of Theorem 5.4. This is because when Γ is a ray, l no longer depends on $y \in \mathscr{D}$. In this case Theorem 5.4 can be strengthened by deleting the phrase "for each $y \in \mathscr{D}$" and replacing (5.38) by

$$l(D^2\Phi(\bar{e}; y)) \geqq 0, \qquad y \in \mathscr{D}. \tag{5.39}$$

Remark 5.9. Theorem 5.4 generalizes Warga [35, Thm. 2.3]. To obtain his result, specialize OP by setting $j = 0$, $\mathscr{Z}_0 = \mathbb{R}$ and $Z_0 = \bar{\mathbb{R}}_-$. The hypotheses of Theorem 5.4 are weaker than those of [35, Thm. 2.3] in several important ways. In [35], $\mathscr{E} \cap (M + \bar{e}) \subset E$ is replaced by $\mathscr{E} \cap (M + \bar{e}) = E$, $\mathbb{P}$ is replaced by $\mathbb{B}$ in (5.37) (see Remark 5.1), the map in (5.37) is assumed to be twice continuously differentiable in a neighborhood of the origin and an additional normality-like condition is assumed. It is shown in [13] that this normality condition implies that Γ (specialized to the problem of [35]) is a ray.

We now prove Theorem 5.2. For brevity, we assume $k > 0$; the case $k = 0$ follows from similar arguments. For $I \subset \{0, \cdots, j\}$ let $\Phi_I \triangleq (\bar{\phi}_0, \cdots, \bar{\phi}_j, \psi)$, where $\bar{\phi}_i \triangleq \phi_i$, $i \in I$, and $\phi_i \triangleq 0$, $i \in \{0, \cdots, j\}/I$. We will show that Condition 4.1 is satisfied with $\bar{e}$ and $\mathscr{X}$ as specified, $M = M'$, $n = 2$, $m_i = 0$ for $i \in I_N$, $m_i = 1$ for $i \in I_{A0}$, $Y_{i1} = D\phi_i(\bar{e}; y)$ for $i \in I_{A0}$, and $V(x) = D\Phi_{I_0''}(\bar{e}; x) + \frac{1}{2}D^2\Phi_{I_0''}(\bar{e}; y)$. Since $I' = I_0'$ and $I'' = I_0''$, (5.19) and (5.20) follow from (4.10) and (4.11). Also, (4.12) and (5.19) imply (5.21).

We now show that Condition 5.2 implies Condition 4.1. Note that from (5.14) and Theorem A.2 (with $\nu = 2$) it follows that $V: M' \to \mathscr{Z}'$ is affine. Thus (see Condition 4.1 for notation), V_ψ is affine and V_i is Z_i-convex, $i \in I_0''$. Let $X = \text{co}\,\{x_1, \cdots, x_{k+1}\} \subset M'$ be a k-simplex and define $f: \mathbb{P}^{k+2}(\beta) \to \mathscr{Z}'$ by

$$f(\boldsymbol{\alpha}) = \Phi_{I_0''}\Big(\bar{e} + \alpha_1 y + \sum_{i=2}^{k+2} \alpha_i x_{i-1}\Big).$$

Also, let $\beta_0 > 0$ satisfy $\beta_0 + \beta_0^2 < \beta$. From (5.14), Proposition A.1.III, Remark A.1 and Theorem A.2, it follows that, for $\alpha \in [0, \beta_0)$, $\boldsymbol{\mu} \in \Delta^k$ and $x = \sum_{i=1}^{k+1} \mu_i x_i$,

$$\begin{aligned} \Phi_{I_0''}(\bar{e} + \alpha y + \alpha^2 x) &= f(\alpha, \alpha^2\mu_1, \cdots, \alpha^2\mu_{k+1}) \\ &= f(0) + \alpha f_{\alpha_1}(0) + \frac{1}{2}\alpha^2\Big[f_{\alpha_1\alpha_2}(0) + 2\sum_{i=1}^{k+1} \mu_i f_{\alpha_{i+1}}(0)\Big] + R(\alpha, \boldsymbol{\mu}) \\ &= \Phi_{I_0''}(\bar{e}) + \alpha D\Phi_{I_0''}(\bar{e}; y) + \alpha^2 V(x) + R(\alpha, \boldsymbol{\mu}), \end{aligned} \tag{5.40}$$

where $\alpha^{-2}R(\alpha, \boldsymbol{\mu}) \xrightarrow[\alpha \to 0^+]{} 0$ uniformly for $\boldsymbol{\mu} \in \Delta^k$. From (5.14) it follows that the mapping

$$\boldsymbol{\alpha} \to \phi_i\Big(\bar{e} + \alpha_1 y + \sum_{i=2}^{k+2} \alpha_i x_{i-1}\Big): \mathbb{P}^{k+2}(\beta) \to \mathscr{Z}_i$$

is continuous at $\boldsymbol{\alpha} = 0$, $i \in I_N$, and $\tilde{F}$-differentiable at $\boldsymbol{\alpha} = 0$, $i \in I_0'/I_N$. Thus, Proposition A.1.I implies

$$\phi_i(\bar{e} + \alpha y + \alpha^2 x) - \phi_i(\bar{e}) \xrightarrow[\alpha \to 0^+]{} 0 \tag{5.41}$$

uniformly for $\boldsymbol{\mu} \in \Delta^k$, $i \in I_N$, and Proposition A.1.II yields

$$\alpha^{-1}[\phi_i(\bar{e}+\alpha y+\alpha^2 x)-\phi_i(\bar{e})-\alpha D\phi_i(e\,;y)] \xrightarrow[\alpha\to 0^+]{} 0 \tag{5.42}$$

uniformly for $\boldsymbol{\mu} \in \Delta^k$, $i \in I_0'/I_N$. For $N_0, \cdots, N_j$, η and τ as specified in Condition 4.1, it is now easy to see from (5.40)–(5.42) that there exists $\alpha \in (0, \tau)$ such that (4.4), (4.5) (without need for Z_i) and (4.7) are satisfied with $\zeta(x) = \bar{e} + \alpha y + \alpha^2 x$. Finally, (4.6) is a consequence of (5.14).

To prove Theorem 5.1, let $n = 1$, $m_i = 0$ for $i \in I_N$, $m_i = 1$ for $i \in I_{A0}$, and $V(x) = D\Phi_{I_{A0}}(\bar{e}\,;x)$. Since $I' = I_N$ and $I'' = I_{A0}$, (5.9) and (5.10) follow from (4.10) and (4.11). Also, (5.11) follows from (4.12) and (5.9). The remainder of the proof follows from arguments that should by now be clear. □

We now point out several ways in which the results of this section can be generalized. First, it is possible to take advantage of the presence of Z_i in (4.4) and (4.5) by replacing continuity and differentiability conditions such as (5.51) and (5.52) by conditions involving semicontinuity and semidifferentiability. This approach appears in [20] for first-order necessary conditions but apparently has not been extended to higher-order necessary conditions.

Another approach to generalizing Theorems 5.1 and 5.2 is based on the concept of a conical approximation. This involves the existence of a map θ depending on X such that, instead of (5.3), (5.12) and (5.13), the following conditions hold:

$$\theta(\bar{e}+\alpha X) \subset E, \qquad \alpha \in [0, \beta), \tag{5.3*}$$

$$\theta(\bar{e}+\alpha_1 y+\alpha_2 X) \subset \mathscr{E}, \qquad (\alpha_1, \alpha_2) \in \mathbb{P}^2(\beta), \tag{5.12*}$$

$$\theta(\bar{e}+\alpha y+\alpha^2 X) \subset E, \qquad (\alpha, \alpha^2) \in \mathbb{P}^2(\beta). \tag{5.13*}$$

Moreover, in (5.5) and (5.14), $\Phi \circ \theta$ plays the role of Φ. Additional assumptions such as the following are then required: if X and $\hat{X}$ have corresponding maps θ and $\hat{\theta}$, then

$$D\Phi \circ \theta(\bar{e}\,;x) = D\Phi \circ \hat{\theta}(\bar{e}\,;x), \qquad x \in X \cap \hat{X}. \tag{5.43}$$

A closely related approach appears in [7, Def. B.1.3]. See [28] for a related notion of conical approximation in the context of second-order necessary conditions for a nonlinear programming problem.

6. Applications to nonlinear programming. In this section we further illustrate the use of Theorem 4.1 by deriving first-, second- and third-order necessary conditions for a nonlinear programming problem. The third-order conditions are new while the first- and second-order conditions sharpen results from the previous literature. To keep the conditions reasonably simple, the hypotheses of this section are considerably stronger than those of § 5.

To derive nth-order necessary conditions, the strengthened assumptions for OP are:

(6.1) $\mathscr{E}$ is a Banach space,

(6.2) $\mathscr{X}_i = \mathbb{R}$, $Z_i = \bar{\mathbb{R}}_-$, $i \in \{0, \cdots, j\}$,

(6.3) $\phi_0, \cdots, \phi_j$ and ψ are C^n.

A mapping from $\mathscr{E}$ into $\mathbb{R}^r$ is C^n if it is n-times Fréchet differentiable and its first- through nth-order derivatives are continuous. Let NP denote OP in the presence of (6.1)–(6.3). As before, let $j=0$ and $k=0$ denote the absence of (1.2)′ and (1.3), respectively. Define the (Lagrangian) function $L: \mathscr{E} \times \mathbb{R}^{1+j+k} \to \mathbb{R}$ by

$$(6.4) \qquad L(e, l) = \begin{cases} \sum_{i=0}^{j} l_i \phi_i(e) + l_\psi^T \psi(e), & l = (l_0, \cdots, l_j, l_\psi), \quad k > 0, \\ \sum_{i=0}^{j} l_i \phi_i(e), & l = (l_0, \cdots, l_j), \quad k = 0. \end{cases}$$

Notation for the Fréchet derivative and partial Fréchet derivative, such as in

$$L_{ee}(e, l)(y, \hat{y}) = \sum_{i=0}^{j} l_i \phi_i''(e)(y, \hat{y}) + l_\psi^T \psi''(e)(y, \hat{y}),$$

is made precise in the Appendix. We require some notation pertaining to a feasible element $\bar{e}$. Recall that if $j>0$ then $I_N = \{i \in \{1, \cdots, j\}: \phi_i(\bar{e}) < 0\}$ and, because of (6.2), $I_{A0} = \{0\} \cup \{i \in \{1, \cdots, j\}: \phi_i(\bar{e}) = 0\}$. If $j=0$ then $I_N = \varnothing$ and $I_{A0} = \{0\}$. Let the set of variations M be chosen to satisfy

$$(6.5) \qquad 0 \in M \subset E - \bar{e}, \qquad M \text{ is convex.}$$

For the following definitions let $\psi = 0$ when $k=0$. If $n=1$ define

$$(6.6) \qquad \mathscr{D}_1 \triangleq \{y \in M: \phi_i'(\bar{e})(y) \leqq 0, i \in I_{A0}, \psi'(\bar{e})(y) = 0\},$$

$$(6.7) \qquad \mathscr{I}_1(y) \triangleq \{i \in I_{A0}: \phi_i'(\bar{e})(y) < 0\},$$

where $y \in \mathscr{D}_1$, and

$$(6.8) \qquad \mathscr{I}_1 \triangleq \bigcup_{y \in \mathscr{D}_1} \mathscr{I}_1(y).$$

If $n=2$, then let

$$(6.9) \qquad \begin{aligned} \mathscr{D}_2(y) \triangleq \{\hat{y} \in M: & \tfrac{1}{2}\phi_i''(\bar{e})(y)^2 + \phi_i'(\bar{e})(\hat{y}) \leqq 0, \\ & i \in I_{A0}/\mathscr{I}_1(y), \tfrac{1}{2}\psi''(\bar{e})(y)^2 + \psi'(\bar{e})(\hat{y}) = 0\}, \end{aligned}$$

where $y \in \mathscr{D}_1$,

$$(6.10) \qquad \mathscr{D}_2 \triangleq \{(y, \hat{y}) \in M^2: y \in \mathscr{D}_1, \hat{y} \in \mathscr{D}_2(y)\}$$

and

$$(6.11) \qquad \mathscr{I}_2(y, \hat{y}) \triangleq \{i \in I_{A0}/\mathscr{I}_1(y): \tfrac{1}{2}\phi_i''(\bar{e})(y)^2 + \phi_i'(\bar{e})(\hat{y}) < 0\},$$

where $\hat{y} \in \mathscr{D}_2(y)$.

THEOREM 6.1. *Suppose $\bar{e}$ solves* NP. I. *If $n=1$ then there exists $l \in \mathbb{R}^{1+j+k}$ satisfying*

$$(6.12) \qquad |l| = 1,$$

$$(6.13) \qquad l_i \geqq 0, \qquad i \in \{0, \cdots, j\},$$

$$(6.14) \qquad l_i = 0, \qquad i \in I_N,$$

$$(6.15) \qquad L_e(\bar{e}, l)(x) \geqq 0, \qquad x \in M.$$

II. *If* $n=2$, *then for each* $y \in \mathscr{D}_1$ *there exists* $l \in \mathbb{R}^{1+j+k}$ *satisfying* (6.12)–(6.15) *and*

$$L_{ee}(\bar{e}, l)(y)^2 \geqq 0. \tag{6.16}$$

III. *If* $n=3$, *then for each* $(y, \hat{y}) \in \mathscr{D}_2$ *there exists* $l \in \mathbb{R}^{1+j+k}$ *satisfying* (6.12)–(6.15) *and*

$$l_i = 0, \qquad i \in \mathscr{I}_2(y, \hat{y}), \tag{6.17}$$

$$L_{ee}(\bar{e}, l)(y, \hat{y}) + \tfrac{1}{6} L_{eee}(\bar{e}, l)(y)^3 \geqq 0. \tag{6.18}$$

Remark 6.1. From Proposition 5.2 it follows that l satisfying (6.12)–(6.15) must also satisfy

$$l_i = 0, \qquad i \in \mathscr{I}_1. \tag{6.19}$$

Remark 6.2. If $0 \in \operatorname{int} M$ then (6.15) is clearly equivalent to the condition

$$L_e(\bar{e}, l) = 0. \tag{6.20}$$

Part I of Theorem 6.1 with $\mathscr{E} = \mathbb{R}^\nu$ is well known. When $E = \mathscr{E}$ (and thus $M = \mathscr{E}$) see, e.g., [25, Thm. 1]; when $0 \notin \operatorname{int} M$ see, e.g., [7, Thm. 2.3.12]. Second-order necessary conditions similar to those of part II can also be found in the literature. For example, part II follows from [4, Thm. 3.2] when $E = \mathscr{E} = \mathbb{R}^\nu$. There the condition $l_i = 0$, $i \in \mathscr{I}_1(y)$, is included. Remark 6.1 shows, however, that this adds no new information. After some manipulation, part II with $\mathscr{E} = \mathbb{R}^\nu$ and $j = 0$ can be obtained from [28, Thm. 6]. The only third-order necessary conditions from the literature which appear to be related to part III are [14, Thm. 2.5], [16, Thm. 5.1] and [22, Thm. 2.5]. Because of differing hypotheses, however, these results cannot be compared directly.

Note that l depends on $y \in \mathscr{D}_1$ in part II and on $(y, \hat{y}) \in \mathscr{D}_2$ in part III. Specific examples of second-order necessary conditions where this dependence actually occurs have been given in [4] and [22]. Note that for both parts II and III, l must belong to the set

$$\hat{\Gamma} \triangleq \{l \in \mathbb{R}^{1+j+k} : l \text{ satisfies (6.12)–(6.15)}\}. \tag{6.21}$$

Using $\hat{\Gamma}$, Theorem 6.1 can be written more compactly. This equivalent version of the theorem will help establish further connections with results from the literature.

COROLLARY 6.1. *Suppose* $\bar{e}$ *solves* NP. I. *If* $n = 1$, *then*

$$\hat{\Gamma} \neq \varnothing. \tag{6.22}$$

II. *If* $n = 2$, *then*

(6.23) *for each* $y \in \mathscr{D}_1$ *there exists* $l \in \hat{\Gamma}$ *satisfying* (6.16).

III. *If* $n = 3$, *then*

(6.24) *for each* $(y, \hat{y}) \in \mathscr{D}_2$ *there exists* $l \in \hat{\Gamma}$ *satisfying* (6.17) *and* (6.18).

Remark 6.3. As shown in Remark 5.10, (6.23) can be expressed in the equivalent form

$$\max_{l \in \hat{\Gamma}} L_{ee}(\bar{e}, l)(y)^2 \geqq 0, \qquad y \in \mathscr{D}_1, \tag{6.25}$$

which is similar to [18, Thm. 6]. His result is valid for infinite-dimensional equality constraints when $\psi'(\bar{e})$ has full range. A similar result was also obtained in [23].

Remark 6.4. Condition (6.19) clarifies some second-order necessary conditions

from the prior literature which involve a constraint qualification. In [26, p. 29] and [24, p. 102], for example, the element l satisfying the first-order necessary conditions (6.12)–(6.14) and (6.20) (since $M = \mathscr{E}$) is used to determine a set of critical directions

$$\begin{aligned} \mathscr{D}_1(l) \triangleq \{y \in \mathscr{E} : \phi_i'(\bar{e})(y) = 0 \text{ if } i \in I_{A0} \text{ and } l_i > 0, \\ \phi_i'(\bar{e})(y) \leqq 0 \text{ if } i \in I_{A0} \text{ and } l_i = 0, \\ \psi'(\bar{e})(y) = 0\}. \end{aligned} \tag{6.26}$$

This definition gives the impression that $\mathscr{D}_1(l)$ depends on l. However, using (6.19) it is easy to show that $\mathscr{D}_1(l) = \mathscr{D}_1$ for all l satisfying (6.12)–(6.14) and (6.20). Thus $\mathscr{D}_1(l)$ is independent of l.

Remark 6.5. Sometimes NP satisfies a constraint qualification which implies that (6.23) can be replaced by a stronger condition such as

There exists $l \in \hat{\Gamma}$ such that

$$L_{ee}(\bar{e}, l)(y)^2 \geqq 0, \qquad y \in \mathscr{D}_1, \tag{6.27}$$

or

$$L_{ee}(\bar{e}, l)(y)^2 \geqq 0, \qquad l \in \hat{\Gamma}, \quad y \in \mathscr{D}_1. \tag{6.28}$$

See [3], [4] for details of specific constraint qualifications in this context. Note that (6.28) implies (6.27) and (6.27) implies (6.23). If $\hat{\Gamma}$ is a singleton then (6.23), (6.27) and (6.28) are equivalent.

We now present an example illustrating the use of Theorem 6.1. Let $j = 3$, $k = 0$, $M = E = \mathscr{E} = \mathbb{R}^2$, $\bar{e} = (1, 0)$, $\phi_0(t_1, t_2) = t_1$, $\phi_1(t_1, t_2) = -(1 - t_1)^3 + t_2$, $\phi_2(t_1, t_2) = -t_1$ and $\phi_3(t_1, t_2) = -t_2$. This example is often discussed in connection with the Kuhn–Tucker constraint qualification (see, e.g., [9, p. 20]) which fails to hold at $\bar{e}$. Note that $I_N = \{2\}$, $I_{A0} = \{0, 1, 3\}$, $\phi_0'(\bar{e}) = [1\ 0]$ and $\phi_1'(\bar{e}) = -\phi_3'(\bar{e}) = [0\ 1]$. Thus, part I is satisfied uniquely with $l = (0, \frac{1}{2}, 0, \frac{1}{2})$. Because $\mathscr{D}_1 = \bar{\mathbb{R}}_- \times \{0\}$ and $\phi_1''(\bar{e}) = \phi_3''(\bar{e}) = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, l also satisfies (6.16) for all $y \in \mathscr{D}_1$. Thus, part II is satisfied. Define $y = (-1, 0)$ and, since $\mathscr{D}_2(y) = \mathbb{R} \times \{0\}$, let $\hat{y} = (0, 0)$. Note that $\mathscr{I}_2(y, \hat{y}) = \varnothing$. Since $\phi_3'''(\bar{e}) = 0$, $(\phi_1)_{t_1 t_1 t_1}(\bar{e}) = 6$ and all other components of $\phi_1'''(\bar{e})$ are zero, (6.18) implies that $l_1 \leqq 0$, which contradicts $l_1 = \frac{1}{2}$. Thus, part III is not satisfied and $\bar{e}$ is not optimal.

If in this example ϕ_0 is redefined to be $\phi_0(t_1, t_2) = -t_1$, then $\bar{e}$ is optimal and the necessary conditions are satisfied with l as given above. These examples show that even when $l_0 = 0$ (and thus ϕ_0 is absent from the Lagrangian) the higher-order necessary conditions yield useful information. The reason for this is that ϕ_0 still plays a role in the necessary conditions because of its appearance in the definition of $\mathscr{D}_1$.

We now prove Theorem 6.1. Parts I and II follow most easily from Theorem 5.4, but Theorem 4.1 can also be used without great difficulty. When $n = 1$, it is easy to see that (6.1)–(6.3) and (6.5) imply Condition 5.4 with $\mathscr{X} = \mathscr{E}$. Note that: (5.8) and (6.2) imply (6.13); (5.9) implies (6.14); and (5.11) implies (6.15). Note that (5.10) can be ignored since it yields no useful information. When $n = 2$, (6.1)–(6.3) imply Condition 5.5 with $\mathscr{X}$ as just defined. It remains only to note that (5.38) implies (6.16).

Part III follows from Theorem 4.1. The arguments used in the proof of Proposition 5.5 show that M can be replaced by cone M. We define $n = 3$, $I' = I_N \cup \mathscr{I}_1(y) \cup \mathscr{I}_2(y, \hat{y})$, $I'' = \{0, \cdots, j\}/I'$, $m_i = 0$ for $i \in I_N$, $m_i = 1$ for $i \in \mathscr{I}_1(y)$, $m_i = 2$ for $i \in I'' \cup \mathscr{I}_2(y, \hat{y})$, $Y_{i1} = \phi_i'(\bar{e})(y)$ for $i \in I_{A0}$, $Y_{i2} = \frac{1}{2}\phi_i''(\bar{e})(y)^2 + \phi_i'(\bar{e})(\hat{y})$ for $i \in I'' \cup \mathscr{I}_2(y, \hat{y})$ and $V(x) = \frac{1}{6}\Phi_{I''}'''(\bar{e})(y)^3 + \Phi_{I''}''(\bar{e})(y, \hat{y}) + \Phi_{I''}'(\bar{e})(x)$. (See the proof of Theorem 5.2 for the definition of $\Phi_{I''}$.)

Arguments similar to those used to prove Proposition 5.1 show that (4.12) yields (6.15) and (6.18). Conditions (6.14) and (6.17) follow from (4.10). The arguments required to show that Condition 4.1 is satisfied are a simple extension of those used in the proof of Theorem 5.2. We note only that $\zeta(x) = \bar{e} + \alpha y + \alpha^2 \hat{y} + \alpha^3 x$ and, because of (6.3), the only result needed from the Appendix is a weakened version of Proposition A.1.

7. Concluding remarks. It has been shown that various higher-order necessary conditions can be obtained systematically from a single result. Specific cases examined in detail include first- and second-order necessary conditions involving directional differentials and first-, second- and third-order necessary conditions for a nonlinear programming problem. With only minor extensions of the arguments used to obtain these results it is possible in both cases to obtain necessary conditions of arbitrary order.

Theorem 5.4, which generalizes [35, Thm. 2.3], contributes in several ways to generalizing the results of [34], [35] and [36]. First, since Theorem 5.4 involves finite- and infinite-dimensional inequality constraints, it allows the handling of both endpoint and state-space inequality constraints. Secondly, the cost criterion ϕ_0 in OP may be nonscalar (e.g., Pareto-type). And thirdly, Theorem 5.4 does not require a normality condition which is difficult to verify in practice. The results of [34] and [35] have been extended in this direction in [13]. Other second-order necessary conditions from the literature requiring such an assumption (e.g., [6] and [15]) can also benefit from this generalization. Finally, the third- and higher-order necessary conditions mentioned above lead to necessary conditions in optimal control theory of still higher order. It is hoped that these conditions will be useful in treating singular optimal control problems for which there is an extensive body of literature containing specialized higher-order necessary conditions (see, e.g., [1], [2], [12], [20], [21] and the references therein).

While this paper was being revised a theory of higher-order optimality conditions appeared in [22]. Although a direct comparison of their results to the present paper is rather complex because of differing assumptions, one interesting aspect of their development appears to be more general. In Condition 4.1 it is assumed that $\{Y_{ir}\}_{r=1}^{m_i} \subset Z_i$ so that Y as given by (4.13) is an element of Z. Their conditions require that $m_i = n$ for all i yet allow for a more general situation in which $Y_{i1}, \cdots, Y_{i,m_i}$ satisfy

$$Y_{ir} \in Z_i + \sum_{q=0}^{r-1} \mathbb{R} Y_{iq}, \qquad r = 1, \cdots, m_i,$$

where $Y_{i0} = \phi_i(\bar{e})$.

Appendix. This section collects together differentiation results needed for the proofs of Theorems 5.1, 5.2 and 6.1. Some of them concern the notion of $\tilde{F}$-differentiability given in Definitions 5.1 and 5.2 and extend well-known results from the literature. Also included is a restatement of a result from [33].

Let $\mathscr{X}$ be a Banach space with norm $|\cdot|$, $\mathscr{Y}$ be a topological vector space, $A \subset \mathscr{X}$ and $f: A \to \mathscr{Y}$. When $\mathscr{Y}$ is a Banach space higher-order $\tilde{F}$-derivatives $f^{(n)}(\bar{x}) \in \mathscr{B}(\mathscr{X}; \mathscr{B}(\mathscr{X}; \cdots, \mathscr{B}(\mathscr{X}; \mathscr{Y})) \cdots)$, $n \geqq 2$, may be defined inductively in the manner of Definition 5.2. Following others (e.g., [11, p. 192]) we view $f^{(n)}(\bar{x})$, $n \geqq 2$, as an element of $\mathscr{B}_n(\mathscr{X}; \mathscr{Y})$. Specifically, the values of $f^{(n)}(\bar{x})$ on $\mathscr{X}^n$ are given by $f^{(n)}(\bar{x})(h_1, \cdots, h_n) \triangleq (\cdots((f^{(n)}(\bar{x})(h_1))(h_2)) \cdots)(h_n)$. For notational simplicity $f^{(n)}(\bar{x})(h)^n \triangleq f^{(n)}(\bar{x})(h, \cdots, h)$. The following is an extension of Taylor's Theorem. Its proof is a slight modification of the proof of [11, Thm. 3.6.2].

THEOREM A.1. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces with the norm on $\mathscr{X}$ denoted by $|\cdot|$, $A \subset \mathscr{X}$ and $f: A \to \mathscr{Y}$. If f is m-times $\tilde{F}$-differentiable at $\bar{x} \in A$, where $m \in \mathbb{N}$, and A' is a convex set satisfying $\bar{x} \in A' \subset A$ then*

$$\lim_{\substack{x \to \bar{x} \\ x \in A'/\bar{x}}} |x-\bar{x}|^{-m}\left[f(x)-f(\bar{x})-\sum_{i=1}^{m}(i!)^{-1}f^{(i)}(\bar{x})(x-\bar{x})^i\right]=0. \tag{A.1}$$

Partial $\tilde{F}$-derivatives are introduced as follows (see [33, p. 168], and [30, pp. 58–59]).

DEFINITION A.1. Let $\mathscr{X}_1, \cdots, \mathscr{X}_r$ be Banach spaces, $\mathscr{Y}$ be a topological vector space, $A \subset \mathscr{X} \triangleq \mathscr{X}_1 \times \cdots \times \mathscr{X}_r$, $f: A \to \mathscr{Y}$ and $\bar{x} \triangleq (\bar{x}_1, \cdots, \bar{x}_r) \in A$. For $i \in \{1, \cdots, r\}$ define the set $A_i(\bar{x}) \triangleq \{x_i \in \mathscr{X}_i: (\bar{x}_1, \cdots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \cdots, \bar{x}_r) \in A\}$ and the mapping $x_i \to f_i(x_i): A_i(\bar{x}) \to \mathscr{Y}$ by $f_i(x_i) \triangleq f(\bar{x}_1, \cdots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \cdots, \bar{x}_r)$. If $f_i(x_i)$ is $\tilde{F}$-differentiable at $\bar{x}_i$ then $f_{x_i}(\bar{x}) \triangleq f_i'(\bar{x}_i) \in \mathscr{B}(\mathscr{X}_i; \mathscr{Y})$ is the partial $\tilde{F}$-derivative of f with respect to x_i at $\bar{x}$.

Note that ([33, p. 169]) if f is $\tilde{F}$-differentiable at $\bar{x}$, then $f_{x_i}(\bar{x})$ exists for all $i \in \{1, \cdots, r\}$ and

$$f'(\bar{x})(h)=\sum_{i=1}^{r} f_{x_i}(\bar{x})(h_i), \qquad h \triangleq (h_1, \cdots, h_r) \in \mathscr{X}_1 \times \cdots \times \mathscr{X}_r. \tag{A.2}$$

If $\mathscr{Y}$ is a Banach space then for $n \in \mathbb{N}$ and $\nu_1, \cdots, \nu_n \in \{1, \cdots, r\}$ the higher-order $\tilde{F}$-partial derivative $f_{x_{\nu_1} \cdots x_{\nu_n}}(\bar{x}) \in \mathscr{B}(\mathscr{X}_{\nu_n}, \cdots, \mathscr{X}_{\nu_1}; \mathscr{Y})$ can be defined inductively. If $\nu_1 = \cdots = \nu_n = \bar{\nu}$, then $f_{(x_{\bar{\nu}})^n}(\bar{x}) \triangleq f_{x_{\nu_1} \cdots x_{\nu_n}}(\bar{x})$. If $f^{(n)}(\bar{x})$ exists, then the following relation is valid [11, p. 197]

$$f^{(n)}(\bar{x})(h)^n = \sum f_{x_{\nu_1} \cdots x_{\nu_n}}(\bar{x})(h_{\nu_n}, \cdots, h_{\nu_1}), \qquad h \triangleq (h_1, \cdots, h_r) \in \mathscr{X}_1 \times \cdots \times \mathscr{X}_r, \tag{A.3}$$

where the summation is over all $(\nu_1, \cdots, \nu_n) \in \{1, \cdots, r\}^n$.

The following proposition concerns a one-parameter expansion of a mapping defined on $\mathbb{P}^\nu(\beta)$. The result, which follows from Theorem A.1, shows that the remainder term satisfies a uniform convergence condition.

PROPOSITION A.1. *Let κ, $n \in \mathbb{N}$, $\beta > 0$, $\mathscr{Y}$ be a topological vector space and $f: \mathbb{P}^{\kappa+n}(\beta) \to \mathscr{Y}$. For $\beta_0 > 0$ satisfying $\sum_{i=1}^{n} \beta_0^i \leqq \beta$ define $F: [0, \beta_0) \times \Delta^\kappa \to \mathscr{Y}$ by*

$$F(\alpha, \boldsymbol{\mu}) = \begin{cases} f(\alpha\mu_1, \cdots, \alpha\mu_{\kappa+1}), & n=1, \\ f(\alpha, \alpha^2, \cdots, \alpha^{n-1}, \alpha^n\mu_1, \cdots, \alpha^n\mu_{\kappa+1}), & n \geqq 2. \end{cases} \tag{A.4}$$

I. *If f is continuous at $\boldsymbol{\alpha} = 0$ then*

$$F(\alpha, \boldsymbol{\mu}) - f(0) \xrightarrow[\alpha \to 0^+]{} 0 \tag{A.5}$$

uniformly for $\boldsymbol{\mu} \in \Delta^\kappa$.

II. *If f is $\tilde{F}$-differentiable at $\boldsymbol{\alpha} = 0$, then*

$$\alpha^{-1}[F(\alpha, \boldsymbol{\mu}) - f(0) - \alpha F_\alpha(0, \boldsymbol{\mu})] \xrightarrow[\alpha \to 0^+]{} 0 \tag{A.6}$$

uniformly for $\boldsymbol{\mu} \in \Delta^\kappa$.

III. *If $\mathscr{Y}$ is a Banach space, $n \geqq 2$ and f is m-times $\tilde{F}$-differentiable at $\boldsymbol{\alpha} = 0$, where $m \in \{2, \cdots, n\}$, then*

$$\alpha^{-m}\left[F(\alpha, \boldsymbol{\mu}) - f(0) - \sum_{i=1}^{m}(i!)^{-1}\alpha^i F_{\alpha^i}(0, \boldsymbol{\mu})\right] \xrightarrow[\alpha \to 0^+]{} 0 \tag{A.7}$$

uniformly for $\boldsymbol{\mu} \in \Delta^\kappa$.

Proof. We prove part III only, since similar arguments apply in the other cases. Since $f^{(m)}(0)$ exists, Theorem A.1 can be applied with $\bar{x}=0\in A'=A=\mathbb{P}^{\kappa+n}(\beta)$ to obtain

$$\text{(A.8)}\qquad \lim_{\substack{\boldsymbol{\alpha}\to 0\\ \boldsymbol{\alpha}\in\mathbb{P}^{\kappa+n}(\beta)/\{0\}}} |\boldsymbol{\alpha}|^{-m}\left[f(\boldsymbol{\alpha})-f(0)-\sum_{i=1}^{m}(i!)^{-1}f^{(i)}(0)(\boldsymbol{\alpha})^i\right]=0.$$

Specialize (A.8) by replacing $\boldsymbol{\alpha}$ by $(\alpha,\alpha^2,\cdots,\alpha^{n-1},\alpha^n\mu_1,\cdots,\alpha^n\mu_{\kappa+1})$, where $\alpha\in(0,\beta_0)$ and $\boldsymbol{\mu}\in\Delta^\kappa$. Then (A.8) implies

$$\text{(A.9)}\qquad \lim_{\alpha\to 0^+}\alpha^{-m}[F(\alpha,\boldsymbol{\mu})-f(0)-\theta_m(\alpha,\boldsymbol{\mu})]=0$$

uniformly for all $\boldsymbol{\mu}\in\Delta^\kappa$, where

$$\text{(A.10)}\qquad \theta_m(\alpha,\boldsymbol{\mu})\triangleq\sum_{i=1}^{m}(i!)^{-1}f^{(i)}(0)(\alpha,\alpha^2,\cdots,\alpha^{n-1},\alpha^n\mu_1,\cdots,\alpha^n\mu_{\kappa+1})^i.$$

Using the chain rule [33, p. 172] to express $F_{\alpha^i}(0,\boldsymbol{\mu})$ in terms of the partial $\tilde{F}$-derivatives of f and applying (A.3) to (A.10) it can be shown that

$$\text{(A.11)}\qquad \lim_{\alpha\to 0^+}\alpha^{-m}\left[\theta_m(\alpha,\boldsymbol{\mu})-\sum_{i=1}^{m}(i!)^{-1}\alpha^iF_{\alpha^i}(0,\boldsymbol{\mu})\right]=0$$

uniformly for all $\boldsymbol{\mu}\in\Delta^\kappa$. The desired result (A.7) now follows from (A.9) and (A.11). □

Remark A.1. The $\tilde{F}$-partial derivative $F_{\alpha^i}(0,\boldsymbol{\mu})$ can be written explicitly in terms of the $\tilde{F}$-partial derivatives of f. For example:

$$n=1:\qquad F_\alpha(0,\boldsymbol{\mu})=\sum_{i=1}^{\kappa+1}\mu_if_{\alpha_i}(0);$$

$$n=2:\qquad F_\alpha(0,\boldsymbol{\mu})=f_{\alpha_1}(0),$$

$$F_{\alpha\alpha}(0,\boldsymbol{\mu})=f_{\alpha_1\alpha_1}(0)+2\sum_{i=1}^{\kappa+1}\mu_if_{\alpha_{i+1}}(0);$$

$$n=3:\qquad F_\alpha(0,\boldsymbol{\mu})=f_{\alpha_1}(0),$$

$$F_{\alpha\alpha}(0,\boldsymbol{\mu})=f_{\alpha_1\alpha_1}(0)+2f_{\alpha_2}(0),$$

$$F_{\alpha\alpha\alpha}(0,\boldsymbol{\mu})=f_{\alpha_1\alpha_1\alpha_1}(0)+6f_{\alpha_1\alpha_2}(0)+6\sum_{i=1}^{\kappa+1}\mu_if_{\alpha_{i+2}}(0).$$

The following result follows from [33, Thm. II.3.3].

THEOREM A.2. *Let $\mathscr{X}$ be a vector space, $\mathscr{Y}$ be a topological vector space, $A\subset\mathscr{X}$, $F:A\to\mathscr{Y}$, $\bar{x}\in A$, $\nu\in\mathbb{N}$ and $X\triangleq\operatorname{co}\{x_1,\cdots,x_\nu\}\subset\mathscr{X}$. Suppose there exists $\beta>0$ such that $\bar{x}+\alpha X\subset A$, $\alpha\in[0,\beta)$, and define $f:\mathbb{P}^\nu(\beta)\to\mathscr{Y}$ by $f(\boldsymbol{\alpha})=F(\bar{x}+\sum_{i=1}^{\nu}\alpha_ix_i)$. If $f'(0)$ exists then $DF(\bar{x};x)$ exists for all $x\in\operatorname{cone}X$ and the mapping $x\to DF(\bar{x};x):\operatorname{cone}X\to\mathscr{Y}$ is positively homogeneous and affine.*

Note that $f_{\alpha_i}(0)=DF(\bar{x};x_i)$. Also, if $f''(0)$ exists then $f_{\alpha_i\alpha_i}(0)=D^2F(\bar{x};x_i)$.

Acknowledgment. I wish to thank Elmer G. Gilbert for providing numerous helpful suggestions throughout the course of this work.

REFERENCES

[1] A. A. AGRACHEV AND R. V. GAMKRELIDZE, *A second-order optimality principle for a time-optimal problem*, Math. USSR-Sb., 29 (1976), pp. 547–576.

[2] D. J. BELL AND D. H. JACOBSON, *Singular Optimal Control Problems*, Academic Press, London, 1975.

[3] A. BEN-ISRAEL, A. BEN-TAL AND S. ZLOBEC, *Optimality in Nonlinear Programming*, John Wiley, New York, 1981.

[4] A. BEN-TAL, *Second-order and related extremality conditions in nonlinear programming*, J. Optim. Theory Appl., 31 (1980), pp. 143–165.

[5] A. BEN-TAL AND J. ZOWE, *A unified theory of first- and second-order conditions for extremum problems in topological vector spaces*, Math. Programming Stud., 19 (1982), pp. 39–76.

[6] D. S. BERNSTEIN AND E. G. GILBERT, *Optimal periodic control: The π test revisited*, IEEE Trans. Automat. Control, AC-25 (1980), pp. 673–684.

[7] M. CANON, G. CULLUM AND E. POLAK, *Theory of Optimal Control and Mathematical Programming*, McGraw-Hill, New York, 1970.

[8] A. J. DUBOVICKII AND A. A. MILYUTIN, *Second variations in extremal problems with constraints*, Soviet Math. Dokl., 6 (1965), pp. 12–16.

[9] A. FIACCO AND G. MCCORMICK, *Nonlinear Programming*, John Wiley, New York, 1968.

[10] W. H. FLEMING AND R. W. RISHEL, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, 1975.

[11] T. M. FLETT, *Differential Analysis*, Cambridge Univ. Press, Cambridge, 1980.

[12] R. GABASOV AND F. M. KIRILLOVA, *High-order necessary conditions for optimality*, this Journal, 10 (1972), pp. 127–168.

[13] E. G. GILBERT AND D. S. BERNSTEIN, *Second-order necessary conditions in optimal control: accessory-problem results without normality*, J. Optim. Theory Appl., 41 (1983), pp. 75–106.

[14] B. GOLLAN, *High-order necessary conditions for an abstract optimization problem*, Math. Programming Stud., 14 (1981), pp. 69–76.

[15] M. R. HESTENES, *Calculus of Variations and Optimal Control Theory*, John Wiley, New York, 1964.

[16] K. H. HOFFMANN AND H. J. KORNSTAEDT, *High-order necessary conditions in abstract mathematical programming*, J. Optim. Theory Appl., 26 (1978), pp. 533–568.

[17] R. B. HOLMES, *Geometric Functional Analysis and its Applications*, Springer-Verlag, New York, 1975.

[18] A. D. IOFFE, *Necessary and sufficient conditions for a local minimum 3: second-order conditions and augmented duality*, this Journal, 17 (1979), pp. 266–288.

[19] V. L. KLEE, *Separation and support properties of convex sets—A survey*, in Control Theory and the Calculus of Variations, A. V. Balakrishnan, ed., Academic Press, New York, 1969.

[20] H. W. KNOBLOCH, *Higher-order Necessary Conditions in Optimal Control Theory*, Springer-Verlag, New York, 1981.

[21] A. J. KRENER, *The high-order maximal principle and its application to singular extremals*, this Journal, 15 (1977), pp. 256–293.

[22] F. LEMPIO AND J. ZOWE, *High-order optimality conditions*, in Modern Applied Mathematics-Optimization and Operations Research, B. Korte, ed., North-Holland, Amsterdam, 1982, pp. 147–193.

[23] E. S. LEVITIN, A. A. MILJUTIN AND N. P. OSMOLOVSKII, *On conditions for a local minimum in a problem with constraints*, in Mathematical Economics and Functional Analysis, B. S. Mitjagin, ed., Nauka, Moscow, 1974, pp. 139–202. (In Russian.)

[24] H. MAURER AND J. ZOWE, *First- and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Programming, 16 (1979), pp. 98–110.

[25] G. P. MCCORMICK, *Second-order conditions for constrained minima*, SIAM J. Appl. Math., 15 (1967), pp. 641–652.

[26] ———, *Optimality criteria in nonlinear programming*, in Nonlinear Programming, SIAM-AMS Proceedings, Vol. 9, R. W. Cottle and C. E. Lemke, eds., American Mathematical Society, Providence, RI, 1976, pp. 27–38.

[27] E. J. MCSHANE, *Sufficient conditions for a weak relative minimum in the problem of Bolza*, Trans. Amer. Math. Soc., 52 (1942), pp. 344–379.

[28] E. J. MESSERLI AND E. POLAK, *On second-order necessary conditions for optimality*, this Journal, 7 (1969), pp. 272–291.

[29] L. W. NEUSTADT, *A general theory of extremals*, J. Comput. System Sci., 3 (1969), pp. 57–92.

[30] ———, *Optimization—A Theory of Necessary Conditions*, Princeton Univ. Press, Princeton, NJ, 1976.

[31] F. A. VALENTINE, *Convex Sets*, McGraw-Hill, New York, 1964.

[32] C. VIRSAN, *Necessary conditions of extremality of high order*, Revue Roumaine de Mathématiques Pures et Appliquées, 28 (1973), pp. 591–611.

[33] J. WARGA, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

[34] ———, *A second-order Lagrangian condition for restricted control problems*, J. Optim. Theory Appl., 24 (1978), pp. 475–483.

[35] ———, *A second-order condition that strengthens Pontryagin's maximum principle*, J. Differential Equations, 28 (1978), pp. 284–307.

[36] ———, *A hybrid relaxed-Lagrangian second-order condition for minimum*, Differential Games and Control Theory, Proc. 3rd Kingston Conference, Part A, Marcel Dekker, New York, 1979, pp. 77–94.