SIAM J. OPTIM.
Vol. 9, No. 2, pp. 466–492

SECOND ORDER OPTIMALITY CONDITIONS BASED ON PARABOLIC SECOND ORDER TANGENT SETS*

J. FRÉDÉRIC BONNANS†, ROBERTO COMINETTI‡, AND ALEXANDER SHAPIRO§

Abstract. In this paper we discuss second order optimality conditions in optimization problems subject to abstract constraints. Our analysis is based on various concepts of second order tangent sets and parametric duality. We introduce a condition, called second order regularity, under which there is no gap between the corresponding second order necessary and second order sufficient conditions. We show that the second order regularity condition always holds in the case of semidefinite programming.

Key words. second order optimality conditions, semidefinite programming, semi-infinite programming, tangent sets, Lagrange multipliers, cone constraints, duality

AMS subject classifications. 49K27, 90C30, 90C34

PII. S1052623496306760

1. Introduction.

In this paper we investigate *necessary* as well as *sufficient* second order optimality conditions for an optimization problem in the form

$$\operatorname*{Min}_{x \in X} f(x) \;\text{ subject to }\; G(x) \in K, \tag{1.1 (P)}$$

where X is a finite dimensional space, Y is a Banach space, K is a closed convex subset of Y, and the objective function $f : X \to \mathbb{R}$ as well as the constraint mapping $G : X \to Y$ are assumed to be twice continuously differentiable. By $\Phi := G^{-1}(K)$ we denote the feasible set of (P).

A number of optimization problems can be formulated in the form (1.1) in a natural way. When $Y = \mathbb{R}^p$ and $K = \{0\} \times \mathbb{R}^{p-q}_+$, the feasible set of (P) is defined by a finite number of equality and inequality constraints and (P) becomes a nonlinear programming problem. As another example, consider the space $Y = C(\Omega)$ of continuous functions $\psi : \Omega \to \mathbb{R}$, defined on a compact metric space Ω and equipped with the sup-norm $\|\psi\| := \sup_{\omega\in\Omega} |\psi(\omega)|$. Let $K := C_+(\Omega)$ be the cone of nonnegative valued functions, i.e.,

$$C_+(\Omega) := \{\psi \in C(\Omega) : \psi(\omega) \ge 0 \text{ for all } \omega \in \Omega\}.$$

In that case the abstract constraint $G(x) \in K$ corresponds to $g(x, \omega) \ge 0$ for all $\omega \in \Omega$, where $g(x, \cdot) := G(x)(\cdot)$. If the set Ω is infinite, this leads to an infinite number of constraints and (P) becomes a semi-infinite programming problem (cf. [18] and references therein). Yet another example is provided by semidefinite programming (see, e.g., [43]). There $Y = \mathcal{S}^n$ is the space of $n \times n$ symmetric matrices and $K = \mathcal{S}^n_+$

*Received by the editors July 16, 1996; accepted for publication (in revised form) February 4, 1998; published electronically March 17, 1999. This research was supported by the French-Chilean ECOS Program and by the European Community under contract 931091CL.

http://www.siam.org/journals/siopt/9-2/30676.html

†INRIA, B.P. 105, 78153 Rocquencourt, France (Frederic.Bonnans@inria.fr).

‡Universidad de Chile, Casilla 170/3 Correo 3, Santiago, Chile (rcominet@dim.uchile.cl). The work of this author was supported in part by Fondecyt 1961131.

§School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205 (ashapiro@isye.gatech.edu). The work of this author was supported in part by National Science Foundation grant DMI-9713878.

is the cone of positive semidefinite matrices. Note that $\mathcal{S}^n_+$ can be represented in the form

$$\mathcal{S}^n_+ = \left\{ Z \in \mathcal{S}^n : \omega^T Z \omega \geq 0, \ \omega \in \mathbb{R}^n, \|\omega\| = 1 \right\}$$

so that semidefinite programming can be considered in the framework of semi-infinite programming.

An alternative approach for studying abstract optimality conditions is to consider optimization problems of the form

$$\text{(1.2)} \qquad \operatorname*{Min}_{x \in X} g(F(x)),$$

where $g : Y \to \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous proper convex function and $F : X \to Y$. This problem, known as a *composite optimization* problem, is equivalent to (e.g., [27])

$$\text{(1.3)} \qquad \operatorname*{Min}_{(x,c) \in X \times \mathbb{R}} c \quad \text{subject to} \quad (F(x), c) \in \operatorname{epi}(g),$$

where $\operatorname{epi}(g) := \{(y, c) \in Y \times \mathbb{R} : g(y) \leq c\}$ is the epigraph of g and hence it can be considered as a particular case of the problem (1.1). The converse is also true, that is, problem (1.1) can be represented in the form (1.2) by taking $g(r, y) = r + I_K(y)$ and $F(x) = (f(x), G(x))$, where $I_K(y) = 0$ for $y \in K$ and $+\infty$ elsewhere (see [20, 27]), so that both approaches are essentially equivalent.

Second order optimality conditions have been studied in numerous publications. In order to give a general idea of that type of result, consider for the moment the simplest case when problem (P) is unconstrained. Let x_0 be a stationary point, i.e., it satisfies the first order optimality condition $\nabla f(x_0) = 0$. Then it is well known that the second order necessary condition for x_0 to be locally optimal is that the Hessian matrix $\nabla^2 f(x_0)$ should be positive semidefinite, i.e., $h^T \nabla^2 f(x_0) h \geq 0$ for all $h \in X$. The corresponding second order sufficient condition is that there exists $\alpha > 0$ such that $h^T \nabla^2 f(x_0) h > \alpha \|h\|^2$ for all $h \in X \setminus \{0\}$. Since X is finite dimensional, this is equivalent to $h^T \nabla^2 f(x_0) h > 0$ for all $h \in X \setminus \{0\}$, i.e., $\nabla^2 f(x_0)$ is positive definite. This condition is in fact necessary and sufficient for quadratic growth (3.13). The only difference between the second order necessary condition and the sufficient condition is the term $\alpha \|h\|^2$ in the right-hand side of the former. In such a case we say that there is *no gap* between the necessary and the sufficient second order conditions.

In the case of nonlinear programming (i.e., when the space Y is finite dimensional and the set K is polyhedral), "no gap" second order optimality conditions were already given, under somewhat restrictive assumptions, in [15]. In a sense, a complete description of no gap second order conditions for nonlinear programming was given in Ioffe [19], Ben-Tal [2], and Ben-Tal and Zowe [3].

In semi-infinite programming second order optimality conditions were first derived (under quite restrictive assumptions) by the so-called reduction method, e.g., [1, 16, 17, 37, 44] (see [18] for additional references). It was already clear in those papers that an additional term, representing the curvature of the set K, should appear in second order optimality conditions in order to obtain no gap second order conditions. An attempt to describe this additional term in an abstract way (in the case of semi-infinite programming) was made in Kawasaki [23]. This sparked an intensive investigation aimed at closing the gap between necessary and sufficient second order conditions [11, 12, 20, 21, 25, 26, 27, 34].

Second order optimality conditions for problem (P) may also be obtained by formulating it as a composite optimization problem in the form (1.2) and using the so-called second order (epi)subderivatives. That approach was investigated in Rockafellar [34] for twice epidifferentiable functions and further explored by Ioffe [21] and Cominetti [13]. (See also [36] for a detailed account of that approach.) In particular, in the case of the composition of a piecewise linear-quadratic convex function with a twice continuously differentiable mapping, no gap second order optimality conditions can be explicitly stated in terms of second order (epi)subderivatives.

An alternative approach developed in this paper, which goes back to Ben-Tal [2] and was later refined in Cominetti [12], is based on verification of optimality along curves that have a second order expansion (in that case we speak of a parabolic curve). This approach leads to more explicit second order optimality conditions involving the Hessian of the Lagrangian and the support function of a second order tangent approximation of the set K. Explicit expressions of this support function are known in various situations (see Cominetti and Penot [14]), including semidefinite programming (see Shapiro [41]). This approach is also convenient for sensitivity analysis of parameterized optimization problems [5, 9].

It is clear, however, that there is no reason a priori why optimality should be verified along parabolic curves only. Therefore, one may expect a gap between such necessary and corresponding sufficient second order optimality conditions. Nevertheless, one may search for classes of problems for which the "parabolic" estimates coincide with the estimates based on the second order lower epiderivatives approach. This was done, in the context of infinite dimensional sensitivity analysis, in [5] under an assumption of generalized polyhedricity (although second order lower epiderivatives are not explicitly mentioned in [5], all lower estimates in that paper, in fact, are lower epiderivative estimates).

The main purpose of this paper is to identify a wide class of sets K for which there is no gap between necessary and sufficient second order optimality conditions obtained via the parabolic curve approach. We argue that such sets, which we call *second order regular*, are natural for purposes of second order analysis. In particular we show that cones of positive semidefinite matrices are always second order regular. This complements results in [40, 41] and gives quite a complete description of no gap second order optimality conditions in semidefinite programming. It is possible to show that the epigraph of a piecewise linear-quadratic convex function is second order regular, and hence the suggested approach can be shown to cover the second order optimality conditions obtained for composite optimization in [34]. In the follow-up paper [6], we also show that the concept of second order regularity is useful in sensitivity analysis of parameterized optimization problems.

The organization of this paper is as follows. In the next section we introduce and discuss some concepts of second order tangent sets. Second order necessary and second order sufficient optimality conditions, for the problem (P) in the form (1.1), are given in section 3. Those conditions become no gap second order conditions under the assumption of second order regularity of the set K, which is discussed in section 4. In section 5 we translate the obtained results into the framework of composite optimization. Finally in section 6 some extensions to the case of nonisolated minima are presented.

Throughout this paper we use the following notation and terminology. Let $h : Y \to \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ be an extended real valued function. Assuming that $h(\cdot)$ is finite at a point $y \in Y$, we denote by $h'(y,d)$ its directional derivative

$$h'(y,d) := \lim_{t\downarrow 0} \frac{h(y+td)-h(y)}{t}$$

at the point y in the direction $d \in Y$. Recall that if $h(\cdot)$ is convex, finite valued, and continuous at y, then $h'(y,d)$ exists and is finite valued [31]. In order to deal with possibly discontinuous convex functions in composite optimization (see section 5), we also use the lower directional subderivative $h^{\downarrow}(y,d)$ (see [35])

$$h^{\downarrow}(y,d) := \liminf_{t\downarrow 0, d'\to d} \frac{h(y+td')-h(y)}{t}.$$

It is not difficult to show from the definitions that, provided $h(y)$ is finite, the epigraph of $\psi(\cdot) := h^{\downarrow}(y,\cdot)$ coincides with the contingent (Bouligand) cone (see (2.3) below) to the epigraph of h at the point $(y, h(y))$ (cf. [35]). Therefore the epigraph of $h^{\downarrow}(y,\cdot)$ is closed and hence $h^{\downarrow}(y,\cdot)$ is lower semicontinuous. Note that if $h(\cdot)$ is convex, finite valued, and continuous at y, and hence is Lipschitz continuous in a neighborhood of y, then $h^{\downarrow}(y,\cdot) \equiv h'(y,\cdot)$. In general, if h is a convex, possibly discontinuous function, then the topological closure of the epigraph of $h'(y,\cdot)$ coincides with the epigraph of $h^{\downarrow}(y,\cdot)$.

When $h'(y,d)$ exists and is finite, we denote by $h''_-(y;d,w)$ and $h''_+(y;d,w)$ its lower and upper second order parabolic derivatives [3], respectively, i.e.,

$$h''_-(y;d,w) := \liminf_{t\downarrow 0} \frac{h(y+td+\frac{1}{2}t^2w)-h(y)-th'(y,d)}{\frac{1}{2}t^2},$$

$$h''_+(y;d,w) := \limsup_{t\downarrow 0} \frac{h(y+td+\frac{1}{2}t^2w)-h(y)-th'(y,d)}{\frac{1}{2}t^2}.$$

We say that $h(\cdot)$ is second order directionally differentiable, at y in the direction d, if $h''_-(y;d,w)$ is equal to $h''_+(y;d,w)$ and is finite for all $w \in Y$. In that case the common value is denoted $h''(y;d,w)$. We also use, when $h(y)$ and $h^{\downarrow}(y,d)$ are finite, the following lower second order parabolic derivative:

$$h^{\downarrow\downarrow}_-(y;d,w) := \liminf_{t\downarrow 0,\, w'\to w} \frac{h(y+td+\frac{1}{2}t^2w')-h(y)-th^{\downarrow}(y,d)}{\frac{1}{2}t^2}.$$

Note that if $h(\cdot)$ is Lipschitz continuous near y, then $h^{\downarrow\downarrow}_-(y;d,w) \equiv h''_-(y;d,w)$. This holds, in particular, if $h(\cdot)$ is convex, finite, and continuous, and hence is Lipschitz continuous, at y.

By Y^* we denote the dual space of Y and by $\langle y^*, y\rangle$ the value $y^*(y)$ of the linear functional $y^* \in Y^*$ at $y \in Y$. For a linear continuous mapping $A : X \to Y$ we denote by $A^* : Y^* \to X^*$ its adjoint mapping, i.e., $\langle A^*y^*, x\rangle = \langle y^*, Ax\rangle$ for all $x \in X$ and $y^* \in Y^*$. For a set $T \subset Y$ we denote by $\sigma(\cdot, T)$ its support function, i.e., $\sigma(y^*, T) := \sup_{y\in T} \langle y^*, y\rangle$, and by $\text{dist}(\cdot, T)$ the distance $\text{dist}(y, T) := \inf_{z\in T} \|y - z\|$. By $Df(x)$ and $D^2f(x)$ we denote the first and second order derivatives, respectively, of a function $f(x)$. We denote by $B_Y := \{y \in Y : \|y\| \le 1\}$ the unit ball in Y. By $[\![y]\!] := \{ty : t \in \mathbb{R}\}$ we denote the linear space (one dimensional if $y \neq 0$) generated by vector y.

2. Tangent sets. In this section we discuss the notions of first and second order tangent sets on which our second order optimality conditions are based.

Let us first recall the notion of limits in the sense of Painlevé–Kuratowski for a multifunction $\Psi : X \to 2^Y$ from a normed space X into the set 2^Y of subsets of Y. The upper limit $\limsup_{x\to x_0} \Psi(x)$ is the set of points $y \in Y$ for which *there exists* a sequence $x_n \to x_0$ such that $y_n \to y$ for some $y_n \in \Psi(x_n)$. The lower limit $\liminf_{x\to x_0} \Psi(x)$ is the set of points $y \in Y$ such that for *every* sequence $x_n \to x_0$ it is possible to find $y_n \in \Psi(x_n)$ such that $y_n \to y$.

Let K be a closed subset of a Banach space Y. The (first order) tangent set (cone) to K at a point $y \in K$ can be defined as follows:

$$T_K(y) := \{h \in Y : \operatorname{dist}(y + th, K) = o(t),\ t \geq 0\}. \tag{2.1}$$

By the definition of lower set limits this can be written in the form

$$T_K(y) = \liminf_{t\downarrow 0} \frac{K - y}{t}. \tag{2.2}$$

It is well known that whenever K is convex, it is also true that

$$T_K(y) = \limsup_{t\downarrow 0} \frac{K - y}{t}. \tag{2.3}$$

Note that if K is a convex cone and $y \in K$, then $T_K(y) = \operatorname{cl}(K + [\![y]\!])$, where $[\![y]\!]$ denotes the linear space generated by vector y and "cl" stands for the topological closure in the norm topology of Y.

Similarly to (2.2) and (2.3) we consider second order variations of the set K at a point $y \in K$ in a direction d. That is,

$$T_K^2(y, d) := \liminf_{t\downarrow 0} \frac{K - y - td}{\frac{1}{2}t^2}, \tag{2.4}$$

$$O_K^2(y, d) := \limsup_{t\downarrow 0} \frac{K - y - td}{\frac{1}{2}t^2}. \tag{2.5}$$

We call $T_K^2(y, d)$ and $O_K^2(y, d)$ the *inner* and *outer* second order tangent sets, respectively. Alternatively these tangent sets can be written in the form

$$T_K^2(y, d) = \left\{w \in Y : \operatorname{dist}(y + td + \tfrac{1}{2}t^2 w, K) = o(t^2),\ t \geq 0\right\},$$

$$O_K^2(y, d) = \left\{w : \exists\, t_n \downarrow 0 \text{ such that } \operatorname{dist}(y + t_n d + \tfrac{1}{2}t_n^2 w, K) = o(t_n^2)\right\}.$$

It is clear from the above definitions that $T_K^2(y, d) \subset O_K^2(y, d)$ and that these second order tangent sets can be nonempty only if $d \in T_K(y)$. Also, both sets $T_K^2(y, d)$ and $O_K^2(y, d)$ are closed. If K is convex, then the set $T_K^2(y, d)$ is convex. On the other hand the outer second order tangent set $O_K^2(y, d)$ can be nonconvex. An example of a convex set K (in $\mathbb{R}^4$) for which $O_K^2(y, d)$ is nonconvex is constructed in the forthcoming book [10]. (That example is not trivial and will be not repeated here.)

The following example demonstrates that unlike the first order tangent variations, the second order inner and outer tangent sets can be different. (Other examples have been given in [14, 27].) It also shows that lower and upper second order directional derivatives can be different even for a convex continuous function of one variable.

EXAMPLE 2.1. *Let us first construct a convex piecewise linear function $y = \eta(x)$, $x \in \mathbb{R}$, oscillating between two parabolas $y = x^2$ and $y = 2x^2$. That is, we construct $\eta(x)$ in such a way that $\eta(x) = \eta(-x)$, $\eta(0) = 0$ and for some monotonically decreasing to zero sequence x_k, the function $\eta(x)$ is linear on every interval $[x_{k+1}, x_k]$, $\eta(x_k) = x_k^2$ and the straight line passing through the points $(x_k, \eta(x_k))$ and $(x_{k+1}, \eta(x_{k+1}))$ is tangent to the curve $y = 2x^2$. It is quite clear how such a function can be constructed. For a given point $x_k > 0$ consider the straight line passing through the point (x_k, x_k^2) and tangent to the curve $y = 2x^2$. This straight line intersects the curve $y = x^2$ at a point x_{k+1}. Clearly $x_k > x_{k+1} > 0$. One can proceed with the construction in an iterative way. It is easily proved that $x_k \to 0$.*

Define $K := \{(x, y) \in \mathbb{R}^2 : y \geq \eta(x)\}$. We have then that for the direction $d := (1, 0)$, $T_K^2(0, d) = \{(x, y) : y \geq 4\}$ and $O_K^2(0, d) = \{(x, y) : y \geq 2\}$. It also can be seen that for any $w \in \mathbb{R}$, $\eta''_-(0; 1, w) = 2$ and $\eta''_+(0; 1, w) = 4$ and hence $\eta(\cdot)$ is not second order directionally differentiable at zero.

We say that the set K is *second order directionally differentiable*, at $y \in K$ in a direction d, if $T_K^2(y, d) = O_K^2(y, d)$ (for various related concepts see [35]). This terminology is justified by the following result, which is an extension of [12, Proposition 4.1].

PROPOSITION 2.1. *Suppose that the set K is defined in the form $K = \{y \in Y : h(y) \leq 0\}$, where $h : Y \to \mathbb{R} \cup \{+\infty\}$ is a proper convex function. Let $h(y) = 0$ and $h^{\downarrow}(y, d) = 0$, and suppose that there exists $\bar{y}$ such that $h(\bar{y}) < 0$ (Slater condition). Then*

$$O_K^2(y, d) = \left\{w : h_-^{\downarrow\downarrow}(y; d, w) \leq 0\right\}. \tag{2.6}$$

If, in addition, $h(\cdot)$ is continuous at y, then

$$T_K^2(y, d) = \{w : h''_+(y; d, w) \leq 0\}. \tag{2.7}$$

Proof. We show only that (2.6) holds since the proof of (2.7) is similar. Consider $w \in O_K^2(y, d)$, and choose sequences $t_n \to 0^+$ and $w_n \to w$ such that $y + t_n d + \frac{1}{2} t_n^2 w_n \in K$, and hence $h(y + t_n d + \frac{1}{2} t_n^2 w_n) \leq 0$. Then

$$h_-^{\downarrow\downarrow}(y; d, w) \leq \liminf_{n\to\infty} \frac{h(y + t_n d + \frac{1}{2} t_n^2 w_n)}{\frac{1}{2} t_n^2} \leq 0.$$

Conversely, suppose first that $h_-^{\downarrow\downarrow}(y; d, w) < 0$. Then for some $t_n \to 0^+$ and $w_n \to w$, we have that

$$h(y + t_n d + \tfrac{1}{2} t_n^2 w_n) = \tfrac{1}{2} t_n^2 h_-^{\downarrow\downarrow}(y; d, w) + o(t_n^2),$$

and hence $h(y + t_n d + \frac{1}{2} t_n^2 w_n) < 0$ for n large enough. Consequently

$$y + t_n d + \tfrac{1}{2} t_n^2 w_n \in K,$$

which implies that $w \in O_K^2(y, d)$.

Suppose now that $h_-^{\downarrow\downarrow}(y; d, w) = 0$, and hence for some $t_n \to 0^+$ and $w_n \to w$, $h(y + t_n d + \frac{1}{2} t_n^2 w_n) = o(t_n^2)$. Given $\alpha > 0$ and $w' \in Y$, set $w'_\alpha := w' + \alpha(\bar{y} - y)$. Then, by convexity of h, we have that for $t' \geq 0$ small enough such that $1 - \frac{1}{2}\alpha {t'}^2 > 0$,

$$h(y + t' d + \tfrac{1}{2} {t'}^2 w'_\alpha) \leq (1 - \tfrac{1}{2}\alpha {t'}^2)\gamma(t', w') + \tfrac{1}{2}\alpha {t'}^2 h(\bar{y}), \tag{2.8}$$

where

$$\gamma(t', w') := h\left(y + t'(1 - \tfrac{1}{2}\alpha {t'}^2)^{-1}d + \tfrac{1}{2}{t'}^2(1 - \tfrac{1}{2}\alpha {t'}^2)^{-1}w\right).$$

Define t'_n and w'_n by the relations $t'_n(1-\frac{1}{2}\alpha {t'_n}^2)^{-1} = t_n$, i.e., $t'_n = 2t_n/(1+\sqrt{1+2\alpha t_n^2})$, and $(1 - \frac{1}{2}\alpha {t'_n}^2)w'_n = w_n$. Then

$$\gamma(t'_n, w'_n) = h(y + t_n d + \tfrac{1}{2}t_n^2 w_n) = o(t_n^2).$$

Since $t'_n \to 0^+$, $w'_n + \alpha(\bar{y} - y) \to w_\alpha$, and $h(\bar{y}) < 0$, it follows then by (2.8) that for any $\alpha > 0$,

$$h_-^{\downarrow\downarrow}(y; d, w_\alpha) \le \alpha h(\bar{y}) < 0$$

and hence $w_\alpha \in O_K^2(y, d)$. Since $O_K^2(y, d)$ is closed, letting $\alpha \to 0^+$ we obtain that $w \in O_K^2(y, d)$, which completes the proof of (2.6). □

If $h(\cdot)$ is convex and continuous at y, then second order derivatives $h_-^{\downarrow\downarrow}(y; d, \cdot)$ and $h''_-(y; d, \cdot)$ are the same. Then it follows from the above proposition, provided the Slater condition holds, that K is second order directionally differentiable, at the point y in the direction d, if and only if the level sets $\{w : h''_-(y; d, w) \le 0\}$ and $\{w : h''_+(y; d, w) \le 0\}$ coincide. In particular, K is second order directionally differentiable if $h(\cdot)$ is second order directionally differentiable.

To close this section we state two results, which extend Proposition 3.1 and Theorem 3.1 in [12] to the case of outer second order tangent sets. We omit the proofs, which are simple modifications of those in the cited reference.

PROPOSITION 2.2. *For all $y \in K, d \in T_K(y)$ one has*

$$T_K^2(y, d) + T_{T_K(y)}(d) \subset T_K^2(y, d) \subset T_{T_K(y)}(d), \tag{2.9}$$

$$O_K^2(y, d) + T_{T_K(y)}(d) \subset O_K^2(y, d) \subset T_{T_K(y)}(d). \tag{2.10}$$

In particular, it follows from the above proposition that $T_{T_K(y)}(d)$ is the recession cone of $T_K^2(y, d)$ and $O_K^2(y, d)$ whenever these sets are nonempty. Moreover, if $0 \in O_K^2(y, d)$, then $O_K^2(y, d) = T_{T_K(y)}(d)$ and when $0 \in T_K^2(y, d)$ all three sets coincide:

$$T_K^2(y, d) = O_K^2(y, d) = T_{T_K(y)}(d).$$

Note also that $T_{T_K(y)}(d) = \operatorname{cl}\{T_K(y) + [\![d]\!]\}$, provided $d \in T_K(y)$; $T_{T_K(y)}(d)$ is empty otherwise.

The following formulas (2.12) and (2.13) provide a rule for computing the second order tangent approximations of the feasible set $\Phi := G^{-1}(K)$ of (P) in terms of the second order tangent approximations of K. These formulas are valid under Robinson's constraint qualification [29]

$$0 \in \operatorname{int}\{G(x_0) + DG(x_0)X - K\} \tag{2.11}$$

and can be proved by using the Robinson–Ursescu [30, 42] stability theorem (see [12]).

PROPOSITION 2.3. *Let $x_0 \in \Phi := G^{-1}(K)$, and suppose that Robinson's constraint qualification* (2.11) *holds. Then, for all $h \in X$,*

$$T_\Phi^2(x_0, h) = DG(x_0)^{-1}\left[T_K^2(G(x_0), DG(x_0)h) - D^2G(x_0)(h, h)\right], \tag{2.12}$$

$$O_\Phi^2(x_0, h) = DG(x_0)^{-1}\left[O_K^2(G(x_0), DG(x_0)h) - D^2G(x_0)(h, h)\right]. \tag{2.13}$$

3. Second order optimality conditions. In this section we derive second order necessary and sufficient optimality conditions for a problem (P) given in the form (1.1). With problem (P) are associated the Lagrangian

$$L(x, \lambda) := f(x) + \langle \lambda, G(x) \rangle, \quad \lambda \in Y^*,$$

and the generalized Lagrangian

$$L^*(x, \alpha, \lambda) := \alpha f(x) + \langle \lambda, G(x) \rangle, \quad (\alpha, \lambda) \in \mathbb{R} \times Y^*.$$

Let x_0 be a locally optimal solution of problem (P). Then F. John–type (first order) optimality conditions can be written in the following form: there exists $(\alpha, \lambda) \in \mathbb{R} \times Y^*$, $(\alpha, \lambda) \neq (0, 0)$, such that

$$D_x L^*(x_0, \alpha, \lambda) = 0, \ \alpha \geq 0, \ \lambda \in N_K(G(x_0)). \tag{3.1}$$

Here $N_K(y) := \{y^* \in Y^* : \langle y^*, z - y \rangle \leq 0 \text{ for all } z \in K\}$ is the normal cone to K at y. We denote by $\Lambda^*(x_0)$ the set of generalized Lagrange multipliers $(\alpha, \lambda) \neq (0, 0)$ satisfying condition (3.1). It should be noted that for a general Banach space Y the set $\Lambda^*(x_0)$ can be empty. The above F. John optimality condition is necessary for local optimality, i.e., $\Lambda^*(x_0) \neq \emptyset$, in two important cases, namely, when the space Y is finite dimensional or when the set K has a nonempty interior [24, 45].

If the multiplier α in (3.1) is nonzero, then we can take $\alpha = 1$ and hence the corresponding first order necessary condition becomes

$$D_x L(x_0, \lambda) = 0, \ \lambda \in N_K(G(x_0)). \tag{3.2}$$

Under Robinson's constraint qualification (2.11) the set $\Lambda(x_0)$ of Lagrange multipliers satisfying (3.2) is nonempty and bounded [28, 45]. When the set K is a convex *cone* and $y \in K$, the normal cone $N_K(y)$ can be written in the form $N_K(y) = \{y^* \in K^- : \langle y^*, y \rangle = 0\}$, where

$$K^- := \{y^* \in Y^* : \langle y^*, y \rangle \leq 0 \text{ for all } y \in K\}$$

is the polar (negative dual) cone of the cone K. In that case condition $\lambda \in N_K(G(x_0))$ becomes $\lambda \in K^-$ and $\langle \lambda, G(x_0) \rangle = 0$.

Let us finally recall that the cone

$$C(x_0) := \{h \in X : DG(x_0)h \in T_K(G(x_0)), Df(x_0)h \leq 0\} \tag{3.3}$$

is called the *critical cone* of the problem (P) at the point x_0. It represents those directions for which a first order linearization of (P) does not provide information about the optimality of x_0. It may be noted that when the set $\Lambda(x_0)$ of Lagrange multipliers is nonempty, then $Df(x_0)h \geq 0$ for any $h \in X$ satisfying $DG(x_0)h \in T_K(G(x_0))$. In such a case the inequality $Df(x_0)h \leq 0$ in the definition of the critical cone can be replaced by the equation $Df(x_0)h = 0$, which in turn is equivalent to $\langle \lambda, DG(x_0)h \rangle = 0$ for any $\lambda \in \Lambda(x_0)$.

With these preliminaries we may now state a second order necessary condition for optimality, which is based on the analysis of feasible parabolic paths of the form

$$x(t) = x_0 + th + \tfrac{1}{2}t^2 w + o(t^2), \tag{3.4}$$

where $t \geq 0$. This necessary condition, combined with the sufficient condition given in Theorem 3.2, will lead to the notion of second order regularity (studied in the next section) under which they become no gap second order optimality conditions.

The following result improves [12, Theorem 4.2], where a similar theorem is stated based on the *inner* second order tangent set. We should mention here an alternative approach suggested by Penot [27] based on the notion of second order compound tangent set, which is a variant of the concept of outer second order tangent set specifically tailored to derive no gap optimality conditions. In this sense we observe that the following result is contained in [27, Corollary 3.6]. However, we will show that under second order regularity, a condition covering many interesting situations, there is no need to resort to the more complicated (and less explicit) concept of compound tangent set, and therefore the following result will suffice for our purpose of stating no gap second order optimality conditions. For the sake of completeness we provide a direct proof which follows the lines of [12, Theorem 4.2].

THEOREM 3.1. *Let x_0 be a locally optimal solution of the problem* (P). *Suppose that Robinson's constraint qualification* (2.11) *holds. Then for all $h \in C(x_0)$ and any convex set $\mathcal{T}(h) \subset O^2_K(G(x_0), DG(x_0)h)$,*

$$\sup_{\lambda \in \Lambda(x_0)} \left\{ D^2_{xx} L(x_0, \lambda)(h, h) - \sigma(\lambda, \mathcal{T}(h)) \right\} \geq 0. \tag{3.5}$$

Proof. Note that if $\mathcal{T}(h) = \emptyset$, then $\sigma(\cdot, \mathcal{T}(h)) = -\infty$ and (3.5) trivially holds. Therefore we assume that the set $\mathcal{T}(h)$, and hence the set $O^2_K(G(x_0), DG(x_0)h)$, is nonempty.

We claim that the optimal value of the optimization problem

$$\begin{array}{ll} \text{Min}_{w \in X} & Df(x_0)w + D^2 f(x_0)(h, h) \\ \text{subject to} & DG(x_0)w + D^2 G(x_0)(h, h) \in O^2_K(G(x_0), DG(x_0)h) \end{array} \tag{3.6}$$

is nonnegative. Indeed if w is feasible for this problem, then using Proposition 2.3 we obtain $w \in O^2_\Phi(x_0, h)$, where $\Phi := G^{-1}(K)$. Therefore we can find a sequence $t_k \downarrow 0$ such that $x_k := x_0 + t_k h + \frac{1}{2} t_k^2 w + o(t_k^2) \in \Phi$. The sequence x_k is feasible for (P) and converges to the local minimum x_0, consequently $f(x_k) \geq f(x_0)$ for all k sufficiently large. By using the second order Taylor expansion we have

$$f(x_0) \leq f(x_k) = f(x_0) + t_k Df(x_0)h + \tfrac{1}{2} t_k^2 [Df(x_0)w + D^2 f(x_0)(h, h)] + o(t_k^2),$$

and since $Df(x_0)h = 0$ for any $h \in C(x_0)$, we obtain

$$Df(x_0)w + D^2 f(x_0)(h, h) \geq 0,$$

establishing our claim.

Consider now the following set $T(h) := \text{cl}\{\mathcal{T}(h) + T_K(G(x_0))\}$. This set is the topological closure of the sum of two convex sets and hence is convex. Moreover, it follows from the first inclusion of (2.10), and the fact that second order outer tangent sets are closed, that $T(h) \subset O^2_K(G(x_0), DG(x_0)h)$. Clearly if we replace the outer second order tangent set in (3.6) by its subset $T(h)$, the optimal value of the obtained optimization problem will be greater than or equal to the optimal value of (3.6), and hence the optimal value of the problem

$$\begin{array}{ll} \text{Min}_{w \in X} & Df(x_0)w + D^2 f(x_0)(h, h) \\ \text{subject to} & DG(x_0)w + D^2 G(x_0)(h, h) \in T(h) \end{array} \tag{3.7}$$

is nonnegative as well.

The optimization problem (3.7) is convex and its (parametric) dual (cf. [32], [5]) is

$$\operatorname*{Max}_{\lambda\in\Lambda(x_0)} \left\{ D^2_{xx}L(x_0,\lambda)(h,h) - \sigma(\lambda, T(h)) \right\}. \tag{3.8}$$

Indeed, the Lagrangian of (3.7) is

$$\mathcal{L}(w,\lambda) = D_x L(x_0,\lambda)w + D^2_{xx}L(x_0,\lambda)(h,h).$$

Since for any $z \in T(h)$ we have that $z + T_K(G(x_0)) \subset T(h)$, it follows that $\sigma(\lambda, T(h)) = +\infty$ for any $\lambda \notin [T_K(G(x_0))]^- = N_K(G(x_0))$. Therefore the effective domain of the parametric dual of (3.7) is contained in $\Lambda(x_0)$. The duality then follows. Moreover, Robinson's constraint qualification (2.11) implies that

$$DG(x_0)X - T_K(G(x_0)) = Y.$$

Since for any $z \in T(h)$ it follows that $z + T_K(G(x_0)) \subset T(h)$, we have that

$$z + DG(x_0)X - T(h) = Y.$$

Therefore (3.7) has a feasible solution and Robinson's constraint qualification for the problem (3.7) holds as well. Consequently there is no duality gap between (3.7) and its dual (3.8) (cf. [5]).

We obtain that the optimal value of (3.8) is nonnegative. Since $\mathcal{T}(h) \subset T(h)$, we have that $\sigma(\lambda, \mathcal{T}(h)) \le \sigma(\lambda, T(h))$ and hence (3.5) follows, which completes the proof. $\square$

Remarks. (i) As we mentioned earlier, the outer second order tangent set

$$O^2_K(G(x_0), DG(x_0)h)$$

can be nonconvex. However, when it is convex, one can use this set in the second order condition (3.5), providing a sharper necessary condition. In any case one can take $\mathcal{T}(h)$ to be the *inner* second order tangent set $T^2_K(G(x_0), DG(x_0)h)$. For such a choice of $\mathcal{T}(h)$, (3.5) coincides with the second order necessary condition obtained in [12, Theorem 4.2]. In general, however, the set $\mathcal{T}(h)$ could be taken larger than $T^2_K(G(x_0), DG(x_0)d)$ and therefore Theorem 3.1 is stronger.

(ii) Note that in the second order necessary condition the optimal value of (3.6) is nonnegative, irrespective of whether $O^2_K(G(x_0), DG(x_0)h)$ is convex.

(iii) If

$$0 \in O^2_K(G(x_0), DG(x_0)h)$$

for every $h \in C(x_0)$, in particular if the set K is *polyhedral*, then

$$O^2_K(G(x_0), DG(x_0)h) = T_{T_K(G(x_0))}(DG(x_0)h)$$

and $\sigma(\lambda, \mathcal{T}(h)) = 0$ for every $\lambda \in \Lambda(x_0)$ and $\mathcal{T}(h) := O^2_K(G(x_0), DG(x_0)h)$. Therefore in that case the "sigma term" in (3.5) vanishes. This happens in the case of nonlinear programming.

(iv) Let Σ be the set of sequences $\{t_n\}$ of positive numbers converging to zero. With any $s = \{t_n\} \in \Sigma$, $y \in K$, and $d \in T_K(y)$ we can associate the following second order tangent set:

$$T^{2,s}_K(y,d) := \left\{ w : \operatorname{dist}(y + t_n d + \tfrac{1}{2}t_n^2 w, K) = o(t_n^2) \right\}.$$

For any $s \in \Sigma$ the set $T_K^{2,s}(y,d)$ is convex and closed. It is clear that the intersection of $T_K^{2,s}(y,d)$ over all $s \in \Sigma$ is $T_K^2(y,d)$ and the union of $T_K^{2,s}(y,d)$ over all $s \in \Sigma$ is $O_K^2(y,d)$. A possible choice for $\mathcal{T}(h)$ is then $T_K^{2,s}(G(x_0), DG(x_0)h)$ for any $s \in \Sigma$.

(v) We can formulate the second order necessary condition (3.5) in the form

$$\inf_{\mathcal{T}(h)\in\mathcal{O}(h)} \sup_{\lambda\in\Lambda(x_0)} \left\{ D_{xx}^2 L(x_0,\lambda)(h,h) - \sigma(\lambda, \mathcal{T}(h)) \right\} \geq 0, \tag{3.9}$$

where $\mathcal{O}(h)$ denotes the set of all convex subsets of $O_K^2(G(x_0), DG(x_0)h)$. In particular, if we take all singleton subsets of $O_K^2(G(x_0), DG(x_0)h)$ (i.e., consisting from one point), then condition (3.9) implies the following necessary condition:

$$\inf_{y\in O_K^2(G(x_0),DG(x_0)h)} \sup_{\lambda\in\Lambda(x_0)} \left\{ D_{xx}^2 L(x_0,\lambda)(h,h) - \langle \lambda, y\rangle \right\} \geq 0. \tag{3.10}$$

If $\Lambda(x_0)$ is a singleton, say, $\Lambda(x_0) = \{\lambda_0\}$, then condition (3.10) becomes

$$D_{xx}^2 L(x_0,\lambda_0)(h,h) - \sigma(\lambda_0, O_K^2(G(x_0), DG(x_0)h)) \geq 0, \tag{3.11}$$

irrespective of whether $O_K^2(G(x_0), DG(x_0)h)$ is convex.

DEFINITION 1. *Let $S \subset \Phi$ be a set of feasible points of the problem* (P) *such that $f(x) = f_0$ for all $x \in S$. It is said that the second order growth condition holds at S if there exist a constant $c > 0$ and a neighborhood N of S such that*

$$f(x) \geq f_0 + c\,[\operatorname{dist}(x,S)]^2 \text{ for all } x \in \Phi \cap N. \tag{3.12}$$

In particular, if $S = \{x_0\}$ is a singleton, the second order growth condition (3.12) takes the form

$$f(x) \geq f(x_0) + c\|x - x_0\|^2 \text{ for all } x \in \Phi \cap N, \tag{3.13}$$

which clearly implies that x_0 is a locally optimal solution of (P). Moreover, in this case (assuming always that Robinson's condition (2.11) holds) it follows easily that for any $h \in C(x_0)$ the optimal value of (3.6) is greater than or equal to $2c\|h\|^2$, so that the second order necessary condition (3.5) can be strengthened to strict inequality for all nonzero $h \in C(x_0)$.

The second order necessary condition (3.5) is based on *upper* estimates of the objective function along feasible parabolic curves of the form (3.4). In order to derive lower estimates, and hence to obtain second order *sufficient* conditions, we need an additional concept.

DEFINITION 2. *Let $y \in K$, $d \in T_K(y)$, and consider a continuous linear mapping $M : X \to Y$. We say that a closed set $\mathcal{A}_{K,M}(y,d) \subset Y$ is an upper second order approximation set for K at the point y in the direction d and with respect to M, if for any sequence $y_k \in K$ of the form $y_k := y + t_k d + \frac{1}{2}t_k^2 r_k$, where $t_k \downarrow 0$ and $r_k = Mw_k + a_k$ with $\{a_k\}$ being a convergent sequence in Y and $\{w_k\} \subset X$ satisfying $t_k w_k \to 0$, the following condition holds:*

$$\lim_{k\to\infty} \operatorname{dist}\left(r_k, \mathcal{A}_{K,M}(y,d)\right) = 0. \tag{3.14}$$

If the above holds for any X and M, i.e., (3.14) *is satisfied for any sequence*

$$y + t_k d + \tfrac{1}{2}t_k^2 r_k \in K$$

such that $t_k r_k \to 0$*, we omit* M *and say that the set* $\mathcal{A}_K(y,d)$ *is an upper second order approximation set for* K *at the point* y *in the direction* d*.*

Let us make the following observations. The above definition is aimed at providing a sufficiently large set $\mathcal{A}_K(y,d)$ such that if $y+td+\varepsilon(t)$ is a curve in K tangential to d with $\varepsilon(t)=o(t)$, then the second order remainder $r(t) := \varepsilon(t)/(\frac{1}{2}t^2)$ tends to $\mathcal{A}_K(y,d)$ as $t \downarrow 0$. Note that this remainder $r(t)$ and its sequential analogue $r_k = r(t_k)$ can be unbounded. The additional complication of considering the linear mapping M, etc. is needed for technical reasons, as is typically encountered in infinite dimensional functional spaces.

The upper second order approximation set $\mathcal{A}_K(y,d)$ is not unique. Clearly, if $\mathcal{A}_K(y,d) \subset B$, then B is also an upper second order approximation set. Since if $y \in K$, $d \in T_K(y)$ and $y+d+w \in K$ imply $d+w \in T_K(y)$ and hence $w \in T_{T_K(y)}(d)$, it follows that the set $T_{T_K(y)}(d)$ is always an upper second order approximation set. It is also not difficult to see from the definitions that the outer second order tangent set $O_K^2(y,d)$ is included in any upper second order approximation set $\mathcal{A}_K(y,d)$.

THEOREM 3.2. *Let* x_0 *be a feasible point of the problem* (P) *satisfying the first order (F. John–type) optimality condition* (3.1)*. Let every* $h \in C(x_0)$ *correspond to an upper second order approximation set* $\mathcal{A}(h) := \mathcal{A}_{K,M}(y_0,d)$ *for the set* K *at the point* $y_0 := G(x_0)$ *in the direction* $d := DG(x_0)h$ *and with respect to the linear mapping* $M := DG(x_0)$*, and suppose that the following second order condition is satisfied:*

$$\sup_{(\alpha,\lambda)\in\Lambda^*(x_0)} \left\{ D^2_{xx}L^*(x_0,\alpha,\lambda)(h,h) - \sigma\left(\lambda, \mathcal{A}(h)\right)\right\} > 0 \tag{3.15}$$

for all $h \in C(x_0) \setminus \{0\}$*. Then the second order growth condition* (3.13) *holds at* x_0*, and hence* x_0 *is a strict locally optimal solution of* (P)*.*

Proof. We argue by contradiction. Suppose that the second order growth condition does not hold at x_0. Then there exists a sequence of feasible points $x_k \in \Phi$, $x_k \neq x_0$, converging to x_0 and such that

$$f(x_k) \le f(x_0) + o(t_k^2), \tag{3.16}$$

where $t_k := \|x_k - x_0\|$. Since the space X is finite dimensional, and hence bounded closed sets in X are compact, we can assume that $h_k := (x_k - x_0)/t_k$ converges to a vector $h \in X$. Clearly $\|h\| = 1$ and hence $h \neq 0$. By using first order Taylor expansions, we obtain from $G(x_k) \in K$ that $DG(x_0)h \in T_K(G(x_0))$ and from (3.16) that $Df(x_0)h \le 0$. Therefore it follows that $h \in C(x_0)$.

By a second order Taylor expansion of $G(x_k)$ at x_0, we have that

$$G(x_k) = y_0 + t_k d + \tfrac{1}{2}t_k^2\left(DG(x_0)w_k + D^2G(x_0)(h,h)\right) + o(t_k^2),$$

where $y_0 := G(x_0)$, $d := DG(x_0)h$, and $w_k := 2t_k^{-2}(x_k - x_0 - t_k h)$. Note that $x_k - x_0 - t_k h = o(t_k)$ and hence $t_k w_k \to 0$. Together with the definition of upper second order approximation set this implies that

$$DG(x_0)w_k + D^2G(x_0)(h,h) \in \mathcal{A}(h) + o(1)B_Y. \tag{3.17}$$

We also have that

$$f(x_k) = f(x_0) + t_k Df(x_0)h + \tfrac{1}{2}t_k^2\left(Df(x_0)w_k + D^2f(x_0)(h,h)\right) + o(t_k^2)$$

so that using (3.16) and (3.17) one can find a sequence $\varepsilon_k \to 0$ such that

$$\begin{cases} 2t_k^{-1}Df(x_0)h + (Df(x_0)w_k + D^2f(x_0)(h,h)) \le \varepsilon_k, \\ DG(x_0)w_k + D^2G(x_0)(h,h) \in \mathcal{A}(h) + \varepsilon_k B_Y. \end{cases} \tag{3.18}$$

By (3.15) there exists $(\alpha, \lambda) \in \Lambda^*(x_0)$ such that

$$D^2_{xx} L^*(x_0, \alpha, \lambda)(h, h) - \sigma(\lambda, \mathcal{A}(h)) \geq \kappa \tag{3.19}$$

for some $\kappa > 0$. It follows from the second condition in (3.18) that

$$\langle \lambda, DG(x_0)w_k + D^2G(x_0)(h, h)\rangle \leq \sigma(\lambda, \mathcal{A}(h) + \varepsilon_k B_Y) = \sigma(\lambda, \mathcal{A}(h)) + \varepsilon_k \|\lambda\|.$$

Also $\alpha \geq 0$, and if $\alpha \neq 0$, then there exists a Lagrange multiplier and hence $Df(x_0)h = 0$. In any case $\alpha Df(x_0)h = 0$, and hence we obtain from (3.18) and (3.19) that

$$\begin{aligned} 0 &\geq \alpha(2t_k^{-1} Df(x_0)h + Df(x_0)w_k + D^2 f(x_0)(h, h) - \varepsilon_k) \\ &\quad + \langle \lambda, DG(x_0)w_k + D^2G(x_0)(h, h)\rangle - \sigma(\lambda, \mathcal{A}(h)) - \varepsilon_k \|\lambda\| \\ &= D^2_{xx} L^*(x_0, \alpha, \lambda)(h, h) - \sigma(\lambda, \mathcal{A}(h)) - \varepsilon_k(\alpha + \|\lambda\|) \\ &\geq \kappa - \varepsilon_k(\alpha + \|\lambda\|). \end{aligned}$$

Since $\varepsilon_k \to 0$ we obtain a contradiction which completes the proof. ◻

Let us first observe that *finite* dimensionality of the space X was used in the derivation of the above second order *sufficient* condition, while the corresponding second order necessary condition did not require that assumption.

If the set $\Lambda(x_0)$ of Lagrange multipliers is nonempty, then the second order sufficient condition (3.15) is equivalent to

$$\sup_{\lambda \in \Lambda(x_0)} \left\{ D^2_{xx} L(x_0, \lambda)(h, h) - \sigma(\lambda, \mathcal{A}(h)) \right\} > 0 \text{ for all } h \in C(x_0) \setminus \{0\}. \tag{3.20}$$

Also, as was mentioned earlier, the set $\mathcal{Z}(h) := T_{T_K(G(x_0))}(DG(x_0)h)$ is always an upper second order approximation set. Furthermore,

$$\sigma(\lambda, \mathcal{Z}(h)) = \begin{cases} 0 & \text{if } \lambda \in T_K(G(x_0)) \text{ and } \langle \lambda, DG(x_0)h\rangle = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Therefore for that choice of upper second order approximation set, the second order sufficient condition (3.15) takes the form

$$\sup_{(\alpha,\lambda) \in \Lambda^*(x_0)} D^2_{xx} L^*(x_0, \alpha, \lambda)(h, h) > 0 \text{ for all } h \in C(x_0) \setminus \{0\}. \tag{3.21}$$

We obtain the following result.

COROLLARY 3.3. *Let x_0 be a feasible point of the problem* (P) *satisfying the first order (F. John–type) optimality condition* (3.1). *Suppose that the second order sufficient condition* (3.21) *is satisfied. Then the second order growth condition* (3.13) *holds at x_0.*

If the set $\Lambda(x_0)$ of Lagrange multipliers is nonempty, then one can replace $\Lambda^*(x_0)$ in (3.21) by $\Lambda(x_0)$. In that form the second order sufficient condition (3.21) is well known [3, 19]. Moreover, if the set K is polyhedral, i.e., in the case of nonlinear programming, as we mentioned earlier the sigma term vanishes in the corresponding second order necessary condition, which leads to a pair of no gap second order conditions in that case.

4. Second order regularity. Comparing the necessary and sufficient conditions given in (3.5) and (3.20), respectively, one may observe that besides the change from weak to strict inequality, the set $\mathcal{T}(h) \subset O^2_K(G(x_0), DG(x_0)h)$ in the former was replaced by a possibly larger set $\mathcal{A}(h)$. Now, conditions (3.15) and (3.20) become stronger if one can take a smaller second order approximation set $\mathcal{A}(h)$. In particular, if $O^2_K(G(x_0), DG(x_0)h)$ is an upper second order approximation set, condition (3.20) becomes the strongest possible by taking $\mathcal{A}(h) = O^2_K(G(x_0), DG(x_0)h)$. In that case, provided $O^2_K(G(x_0), DG(x_0)h)$ is convex, the gap between (3.5) and (3.20) reduces to the difference between weak and strict inequality, and hence we obtain a pair of no gap second order conditions. This motivates the following definition.

DEFINITION 3. *We say that the set K is outer second order regular at a point $y \in K$ in a direction $d \in T_K(y)$ and with respect to a linear mapping $M : X \to Y$ if for any sequence $y_n \in K$ of the form $y_n := y + t_n d + \frac{1}{2}t_n^2 r_n$, where $t_n \downarrow 0$ and $r_n = Mw_n + a_n$ with $\{a_n\}$ being a convergent sequence in Y and $\{w_n\}$ being a sequence in X satisfying $t_n w_n \to 0$, the following condition holds:*

$$\lim_{n\to\infty} \operatorname{dist}\left(r_n, O^2_K(y,d)\right) = 0. \tag{4.1}$$

If K is outer second order regular at $y \in K$ in every direction $d \in T_K(y)$ and with respect to any X and M, we say that K is outer second order regular at y. If, in addition, $O^2_K(y,d) = T^2_K(y,d)$ for every $d \in T_K(y)$, we say that K is second order regular at y.

Outer second order regularity means that the outer second order tangent set $O^2_K(y,d)$ provides an upper second order approximation for K at y in direction d. If in addition the outer and inner second order tangent sets coincide, we simply talk about second order regularity. Second order regularity means that if $y + td + \varepsilon(t)$ is a curve in K tangential to d with $\varepsilon(t) = o(t)$, then $r(t) := \varepsilon(t)/(\frac{1}{2}t^2)$ is arbitrarily close to $T^2_K(y,d)$ as $t \downarrow 0$. Loosely speaking, second order regular sets are the appropriate ones for second order optimality conditions in the sense that there is no gap between the corresponding second order necessary and sufficient conditions; see the following theorem.

THEOREM 4.1. *Let x_0 be a feasible point of* (P) *satisfying the first order necessary condition* (3.2). *Suppose that Robinson's constraint qualification* (2.11) *holds, that for every $h \in C(x_0)$ the set K is outer second order regular at $G(x_0)$ in direction $DG(x_0)h$ and with respect to $M := DG(x_0)$, and that the outer second order tangent set $O^2_K(G(x_0), DG(x_0)h)$ is convex. Then the second order growth condition* (3.13) *holds if and only if the second order sufficient condition* (3.20) *is satisfied with $\mathcal{A}(h) = O^2_K(G(x_0), DG(x_0)h)$.*

Proof. The implication (3.20)⇒(3.13) follows from Theorem 3.2, while the converse is a consequence of Theorem 3.1 and the discussion following the statement of equation (3.13). □

Recall that the inner second order tangent sets are always convex, and hence in the case $O^2_K(G(x_0), DG(x_0)h) = T^2_K(G(x_0), DG(x_0)h)$ the assumed convexity of the outer second order tangent set automatically holds.

At first glance the second order regularity concept, introduced in Definition 3, may seem to be rather technical. Nevertheless it is possible to verify the second order regularity in a number of particular situations. It holds, for example, when $0 \in T^2_K(y,d)$ for every $d \in T_K(y)$, since then $T^2_K(y,d) = T_{T_K(y)}(d)$. This occurs, for instance, when K is a polyhedral set. We discuss in the next subsections several other situations where the second order regularity holds. In particular, we show that the cone $\mathcal{S}^n_+$ of $n \times n$ positive semidefinite matrices is second order regular (at every point $y \in \mathcal{S}^n_+$).

4.1. Sets defined by smooth and convex constraints. Second order regularity is preserved when taking inverse images through twice continuously differentiable mappings satisfying Robinson's constraint qualification.

PROPOSITION 4.2. *Let K be a closed convex subset of Y and $G : X \to Y$ be a twice continuously differentiable mapping. If Robinson's constraint qualification* (2.11) *holds and K is (outer) second order regular at $G(x_0)$ in the direction $DG(x_0)h$ with respect to the linear mapping $M := DG(x_0)$, then the set $G^{-1}(K)$ is (outer) second order regular at x_0 in the direction h.*

Proof. Let $x_k := x_0 + t_k h + \frac{1}{2}t_k^2 r_k \in G^{-1}(K)$ be such that $t_k \downarrow 0$ and $t_k r_k \to 0$. By Proposition 2.3 and the Robinson–Ursescu stability theorem, we obtain for some constant L and all k large enough

$$\begin{aligned}
&\operatorname{dist}\left(r_k, O^2_{G^{-1}(K)}(x_0,h)\right)\\
&\quad = \operatorname{dist}\left(r_k, DG(x_0)^{-1}\left[O^2_K(G(x_0), DG(x_0)h) - D^2G(x_0)(h,h)\right]\right)\\
&\quad \le L \operatorname{dist}\left(DG(x_0)r_k + D^2G(x_0)(h,h), O^2_K(G(x_0), DG(x_0)h)\right).
\end{aligned}$$

Now, a second order expansion of $G(x_k)$ gives

$$G(x_k) = G(x_0) + t_k DG(x_0)h + \tfrac{1}{2}t_k^2\left(DG(x_0)r_k + D^2G(x_0)(h,h)\right) + o(t_k^2).$$

Since $G(x_k) \in K$, the assumed (outer) second order regularity of K implies

$$\operatorname{dist}\left(DG(x_0)r_k + D^2G(x_0)(h,h), O^2_K(G(x_0), DG(x_0)h)\right) \to 0$$

and therefore $\operatorname{dist}(r_k, O^2_{G^{-1}(K)}(x_0,h)) \to 0$, as had to be proved. ∎

Consider the set

$$F := \{x \in X : g_i(x) \le 0,\ i = 1,\dots,p;\ \ h_j(x) = 0,\ j = 1,\dots,q\},$$

defined by a finite number of constraints. Suppose that the functions g_i and h_j are twice continuously differentiable. As a straightforward consequence of Proposition 4.2 and the fact that polyhedral sets are second order regular, we obtain that the set F is second order regular at every point $x_0 \in F$ satisfying the Mangasarian–Fromovitz constraint qualification. Another direct consequence of Proposition 4.2 is the following result.

COROLLARY 4.3. *Let $K_1, \dots, K_n$ be closed convex sets which are second order regular at a point $y_0 \in K_1 \cap \cdots \cap K_n$ in a direction $d \in T_{K_1}(y_0) \cap \cdots \cap T_{K_n}(y_0)$. If there exists a point in K_n which belongs to the interior of the remaining K_i's, $i = 1, \dots, n-1$, then the intersection $K_1 \cap \cdots \cap K_n$ is second order regular at y_0 in the direction d.*

Proof. It suffices to apply Proposition 4.2 with $G : Y \to Y^n$ given by $G(y) = (y, \dots, y)$ and $K = K_1 \times \cdots \times K_n$. It is easily seen that K is second order regular at $(y_0, \dots, y_0)$ in the direction $(d, \dots, d)$.

In order to check Robinson's constraint qualification we take $\bar{y} \in Y$ and $\varepsilon > 0$ such that $\bar{y} \in K_n$ and $\bar{y} + 2\varepsilon B_Y \subset K_1 \cap \cdots \cap K_{n-1}$. If $u_1, \dots, u_n \in \varepsilon B_Y$, letting $y = \bar{y} + u_n$ we have $k_i := y - u_i \in \bar{y} + 2\varepsilon B_Y \subset K_i$ for all $i = 1, \dots, n-1$. Therefore, if we set $k_n := \bar{y} \in K_n$ we have $u_i = y - k_i \in y - K_i$ for all $i = 1, \dots, n$ and then $[\varepsilon B_Y]^n \subset G(Y) - K$, which proves Robinson's constraint qualification. ∎

Returning to the case of sets defined by inequality constraints, we observe that when the constraint functions are convex one may relax the differentiability assumptions.

PROPOSITION 4.4. *Let* $K := \{y : h(y) \le 0\}$, *where* $h(\cdot)$ *is a convex function which is continuous at a point* y_0. *Suppose that the Slater condition holds and that* $h(y_0) = 0$. *Then* K *is outer second-order regular at* y_0 *if and only if, for any* $d \in T_K(y_0)$ *satisfying* $h'(y_0, d) = 0$ *and any path* $y(t) \in K$ *of the form* $y(t) = y_0 + td + \frac{1}{2}t^2 r(t)$, $t \ge 0$, *with* $tr(t) \to 0$ *as* $t \downarrow 0$, *the inequality*

$$\limsup_{t \downarrow 0} h''_-(y_0; d, r(t)) \le 0 \tag{4.2}$$

holds.

Proof. Since h is convex and continuous at y_0, it is directionally differentiable at y_0. Consider a direction $d \in T_K(y_0)$ and a sequence $y_k := y_0 + t_k d + \frac{1}{2}t_k^2 r_k \in K$ with $t_k \downarrow 0$ and $t_k r_k \to 0$. It follows from $d \in T_K(y_0)$ that $h'(y_0, d) \le 0$. Since $h'(y_0, d) < 0$ implies that $O_K^2(y_0, d) = Y$, we only need to consider the case $h'(y_0, d) = 0$.

Because of the Slater condition, there is a point $\bar{y} \in Y$ such that $h(\bar{y}) < 0$. By convexity of $h(\cdot)$ we then have that $h(y_0 + t(\bar{y} - y_0)) < 0$ for any $t \in (0, 1)$ and hence a point $\bar{y}$ where $h(\bar{y}) < 0$ can be chosen arbitrarily close to y_0. Therefore we can assume that $h(\cdot)$ is continuous at $\bar{y}$.

Assume that (4.2) holds. For $\alpha > 0$ let $w_\alpha := r_k + \alpha(\bar{y} - y_0)$. By convexity we get for all $t > 0$ small enough

$$h(y_0 + td + \tfrac{1}{2}t^2 w_\alpha) \le (1 - \tfrac{1}{2}\alpha t^2) h(y_0 + td + \tfrac{1}{2}t^2 r_k) + \tfrac{1}{2}\alpha t^2 h(\bar{y} + td + \tfrac{1}{2}t^2 r_k).$$

Since $h(y_0) = 0$ and $h'(y_0, d) = 0$, dividing by $\frac{1}{2}t^2$ and letting $t \to 0^+$ we deduce

$$h''_-(y_0; d, w_\alpha) \le h''_-(y_0; d, r_k) + \alpha h(\bar{y}),$$

and by virtue of (4.2) we deduce $h''_-(y_0; d, r_k + \alpha(\bar{y} - y_0)) < 0$ for all k sufficiently large. Proposition 2.1 implies that $r_k + \alpha(\bar{y} - y_0) \in O_K^2(y_0, d)$ so that

$$\limsup_{k \to \infty} \ \operatorname{dist}(r_k, O_K^2(y_0, d)) \le \alpha \|\bar{y} - y_0\|.$$

Since α can be made arbitrarily small, we obtain that K is second order regular.

Conversely, assume that K is second order regular. Let $t_k \downarrow 0$ be a sequence through which the upper limit (4.2) is attained as a limit, and let $r_k := r(t_k)$. Set $\varepsilon_k := \operatorname{dist}(r_k, O_K^2(y_0, d)) + 1/k$, so that $\varepsilon_k \to 0$, and choose $\tilde{r}_k \in O_K^2(y_0, d)$ such that $\|r_k - \tilde{r}_k\| < \varepsilon_k$. Since ε_k tends to 0, with no loss of generality we may assume that for all k we have $\bar{y} + 2\varepsilon_k \alpha^{-1} B_Y \subset K$. Choose a sequence $\tau_\ell \downarrow 0$ such that $y_0 + \tau_\ell d + \frac{1}{2}\tau_\ell^2 \tilde{r}_k + o(\tau_l^2) \in K$ and therefore $y_0 + \tau_\ell d + \frac{1}{2}\tau_\ell^2 r_k \in K + \frac{1}{2}\varepsilon_k \tau_\ell^2 B_Y$. Then, for all $\alpha > 0$ and $w_\alpha := r_k + \alpha(\bar{y} - y_0)$ we get

$$\begin{aligned} y_0 + \tau_\ell d + \tfrac{1}{2}\tau_\ell^2 w_\alpha &= (1 - \tfrac{1}{2}\alpha\tau_\ell^2)(y_0 + \tau_\ell d + \tfrac{1}{2}\tau_\ell^2 r_k) + \tfrac{1}{2}\alpha\tau_\ell^2(\bar{y} + \tau_\ell d + \tfrac{1}{2}\tau_\ell^2 r_k) \\ &\subset (1 - \tfrac{1}{2}\alpha\tau_\ell^2)(K + \tfrac{1}{2}\varepsilon_k\tau_\ell^2 B_Y) + \tfrac{1}{2}\alpha\tau_\ell^2(\bar{y} + \tau_\ell d + \tfrac{1}{2}\tau_\ell^2 r_k) \\ &= (1 - \tfrac{1}{2}\alpha\tau_\ell^2)K + \tfrac{1}{2}\alpha\tau_\ell^2\left[\bar{y} + \tau_\ell d + \tfrac{1}{2}\tau_\ell^2 r_k + (1 - \tfrac{1}{2}\alpha\tau_\ell^2)\frac{\varepsilon_k}{\alpha} B_Y\right]. \end{aligned}$$

Since $\bar{y} + 2\varepsilon_k\alpha^{-1}B_Y \subset K$ we deduce $y_0 + \tau_\ell d + \frac{1}{2}\tau_\ell^2 w_\alpha \in K$. By Proposition 2.1, $h''_-(y_0, d, w_\alpha) \le 0$. Since h is continuous at y_0, it is locally Lipschitz continuous.

Therefore, $h''_-(y_0, d, \cdot)$ is globally Lipschitz continuous with the same constant, say, L, and $h''_-(y_0, d, r_k) \le L\|w_\alpha - r_k\| = \alpha L\|\bar{y} - y_0\|$, from which

$$\limsup_{t\downarrow 0} h''_-(y_0, d, r(t)) = \lim_k h''_-(y_0, d, r_k) \le \alpha L\|\bar{y} - y_0\|.$$

Since α may be taken arbitrarily small, the conclusion follows. □

Let us derive now some criteria which allow us to check condition (4.2), assuming that h is convex and continuous at y_0. We first observe that this condition is satisfied whenever

$$h(y_0 + td + \tfrac{1}{2}t^2 r(t)) \ge h(y_0) + th'(y_0, d) + \tfrac{1}{2}t^2 h''_-(y_0; d, r(t)) + o(t^2) \tag{4.3}$$

for all $r(t)$ such that $\operatorname{tr}(t) \to 0$ as $t \downarrow 0$. This holds, for instance, when h is twice continuously differentiable.

A nondifferentiable (at zero) function satisfying (4.3) is the Euclidean norm $h(y) := \|y\|$. Many problems of robust optimization boil down to the minimization of a sum of Euclidean norms subject to linear constraints (see, e.g., [4]), say, $\sum_{i=1}^m \|A_i x\|$, where A_i are $q_i \times n$ matrices. Let us consider for simplicity the unconstrained problem. Introducing slack variables z_i, the problem reduces to the minimization of $\sum_{i=1}^m z_i$, subject to the constraints

$$\|A_i x\| - z_i \le 0, \quad 1 \le i \le m.$$

Set $h(y) = \|y\|$. Note that this is a twice continuously differentiable function at $y_0 \ne 0$. If $y_0 = 0$, we obtain $h'(0, d) = \|d\|$. If $d = 0$, then $h''(0; d, w) = \|w\|$, otherwise $h''(0; d, w) = \langle d, w\rangle/\|d\|$. In both cases (4.3) is easily checked. Therefore $h_i(x, z) = \|A_i x\| - z_i$ also satisfies (4.3) and the Slater condition is trivially satisfied.

Note also that if the functions $\{h_i : i = 1, ..., m\}$ are convex and second order directionally differentiable and satisfy (4.2), then $h := \sum_{i=1}^n h_i$ satisfies (4.2) as well. On the other hand it follows from Corollary 4.3 that the set

$$K = \{y \in Y : h_i(y) \le 0,\ i = 1, ..., m\}$$

is also second order regular, provided the Slater condition holds for the function $h(y) := \max\{h_i(y) : 1 \le i \le m\}$. This result can also be derived directly, as follows. Let y_0 be such that $h(y_0) = 0$. It is not difficult to show that $h(\cdot)$ is second order directionally differentiable with

$$h'(y_0, d) = \max\{h'_i(y_0, d) : i \in I_1(y_0)\},$$
$$h''(y_0; d, w) = \max\{h''_i(y_0; d, w) : i \in I_2(y_0, d)\},$$

where $I_1(y) := \{i : h_i(y) = h(y)\}$ and $I_2(y) := \{i \in I_1(y) : h'_i(y, d) = h'(y, d)\}$. Since $h_i(\cdot)$ satisfy (4.2), we have that for $y(t) := y_0 + td + \frac{1}{2}t^2 r(t)$, such that $h(y(t)) \le 0$, $\operatorname{tr}(t) \to 0$ and for d satisfying $h'(y_0, d) = 0$,

$$h''(y_0; d, r(t)) = \max_{i \in I_2(y_0, d)} h''_i(y_0; d, r(t)) \le o(t^2).$$

It then follows, assuming the Slater condition holds (i.e., there exists $\bar{y}$ such that $h_i(\bar{y}) < 0$, $i = 1, ..., m$), that the set K is second order regular with

$$T^2_K(y_0, d) = O^2_K(y_0, d) = \{w \in Y : h''_i(y_0, d, w) \le 0,\ i \in I_2(y_0, d)\}.$$

4.2. Semi-infinite and semidefinite programming. Let us consider now the case of semi-infinite programming with $Y := C(\Omega)$ and $K := C_+(\Omega)$ and with Ω being a compact metric space. For a function $y \in C_+(\Omega)$ its *contact set* is defined as

$$\Delta(y) := \{\omega \in \Omega : y(\omega) = 0\}. \tag{4.4}$$

It is well known that $d \in T_K(y)$ if and only if $d(\omega) \geq 0$ for all $\omega \in \Delta(y)$ (e.g., [39]). Denote

$$\Delta^*(y,d) := \{\omega \in \Delta(y) : d(\omega) = 0\}. \tag{4.5}$$

Note that if the set $\Delta^*(y,d)$ is empty, then d belongs to the interior of $T_K(y)$, and hence in that case $T_K^2(y,d) = Y$.

Suppose that Ω is a smooth compact manifold of finite dimension n. Consider a twice continuously differentiable function $y \in K$ with a nonempty contact set and a function $d \in T_K(y)$. A general formula for $T_K^2(y,d)$ is given in [14]. We derive now a particular case of that formula by direct arguments in the case where $\Delta(y)$ is a smooth submanifold of Ω. Moreover, we show that in such a case the second order regularity condition holds. These derivations are similar to the analyses in [38] and [5, Part III].

Since Ω is a smooth manifold, by using a local system of coordinates we identify an open neighborhood of a point $\bar{\omega} \in \Omega$ with an open subset of $\mathbb{R}^n$. Such an identification will not effect our local analysis and will simplify the presentation. Moreover, since $\Delta(y)$ is a smooth submanifold of Ω, for each $\bar{\omega} \in \Delta(y)$ we can choose such a local system of coordinates that $\Delta(y)$ is locally represented by a linear subspace of $\mathbb{R}^n$ in that system of coordinates. We denote by $T_{\Delta(y)}(\omega) \subset \mathbb{R}^n$ the tangent space to $\Delta(y)$ at $\omega \in \Delta(y)$ and by $N(\omega)$ its normal complement in $\mathbb{R}^n$, i.e., $N(\omega)$ is a linear space orthogonal to $T_{\Delta(y)}(\omega)$ and such that $T_{\Delta(y)}(\omega) + N(\omega) = \mathbb{R}^n$. Due to the above choice of local coordinates, these sets $T_{\Delta(y)}(\omega)$ and $N(\omega)$ are constant in the chosen system of local coordinates.

For a point $\omega \in \Omega$ we define its projection onto $\Delta(y)$ to be a point $\hat{\omega} \in \Delta(y)$ closest to ω with respect to the Euclidean distance in the chosen system of coordinates of Ω. This operation is well defined in the vicinity of $\bar{\omega}$ and of course depends on the choice of a local system of coordinates. Let $V(\omega)$ be a matrix whose columns form a basis of the linear space $N(\omega)$. Consider the following second order growth condition: for any $\bar{\omega} \in \Delta(y)$, there exists a local system of coordinates of the type described above such that

$$y(\omega) \geq c\,\mathrm{dist}(\omega, \Delta(y))^2 \text{ for all } \omega \in \Omega \cap \mathcal{N} \tag{4.6}$$

for some $c > 0$ and a neighborhood $\mathcal{N}$ of $\bar{\omega}$. Note that this condition does not depend on the system of coordinates (although the value of the constant c does of course) and is satisfied if and only if the matrix

$$U(\omega) := V(\omega)^T \nabla^2 y(\omega) V(\omega) \tag{4.7}$$

is positive definite for every $\omega \in \Delta(y)$ (see [38]).

THEOREM 4.5. *Let $y \in K := C_+(\Omega)$ be a twice continuously differentiable function, and let $d \in T_K(y)$ be continuously differentiable. Suppose that the set Ω is a smooth compact manifold, that $\Delta(y)$ is a smooth submanifold of Ω, and that the second order growth condition* (4.6) *holds for some $c > 0$ and with $\mathcal{N}$ being a neighborhood of*

$\Delta(y)$. Then the set K is second order directionally differentiable at y in the direction d with

$$T_K^2(y,d) = \left\{h \in C(\Omega) : h(\omega) \geq A(\omega)^T[U(\omega)]^{-1}A(\omega) \text{ for all } \omega \in \Delta^*(y,d)\right\}, \tag{4.8}$$

where $A(\omega) := V(\omega)^T \nabla d(\omega)$ and $U(\omega)$ is given in (4.7).

Moreover, let $M(x) := \sum_{i=1}^m x_i\psi_i(\cdot)$ be a linear mapping from $\mathbb{R}^m$ into $C(\Omega)$ such that the functions $\psi_i(\cdot)$, $i = 1, ..., m$, are Lipschitz continuous on Ω. Then the set K is second order regular at y in the direction d and with respect to M.

Proof. We already observed that when $\Delta^*(y,d)$ is empty we have $O_K^2(y,d) = T_K^2(y,d) = Y$ and the result holds trivially, so we may also assume that $\Delta^*(y,d) \neq \emptyset$.

Consider a path $\bar{y}_t(\cdot) := y(\cdot) + td(\cdot) + \frac{1}{2}t^2h(\cdot)$ and the corresponding min-function $\nu(t) := \min_{\omega\in\Omega} \bar{y}_t(\omega)$. Since $\text{dist}(\bar{y}_t, K) = \max\{0, -\nu(t)\}$, we have $h \in T_K^2(y,d)$ if and only if $\liminf_{t\downarrow 0} \nu(t)/t^2 \geq 0$ and $h \in O_K^2(y,d)$ if and only if $\limsup_{t\downarrow 0} \nu(t)/t^2 \geq 0$. We shall prove that in fact the limit $\lim_{t\downarrow 0} \nu(t)/t^2$ exists so that both second order tangent sets coincide.

Let $\bar{\omega}_t$ be a minimizer of $\bar{y}_t(\omega)$ over Ω, and let the sequence $t_n \to 0^+$ be such that $t^{-2}\nu(t)$ attains its lower limit. Let us denote $\bar{\omega}^n := \bar{\omega}_{t_n}$. Extracting if necessary a subsequence, we can assume that $\bar{\omega}^n \to \bar{\omega}^0 \in \Omega$. Since $\Delta^*(y,d) \neq \emptyset$, we have $\nu(t_n) \leq O(t_n^2)$, from which it follows that $\bar{\omega}^0 \in \Delta^*(y,d)$ (see [38]).

For n large enough, $\bar{\omega}^n$ can be described in terms of a local system of coordinates containing $\bar{\omega}^0$ in which the submanifold $\Delta(y)$ coincides with an affine space. To avoid heavy notation we will identify elements of Ω close to $\bar{\omega}^0$ with the corresponding vector of coordinates. Denote by $\hat{\omega}^n$ the projection of ω^n onto $\Delta(y)$ (in the given local system). Then $\delta^n := t_n^{-1}(\bar{\omega}^n - \hat{\omega}^n)$ is orthogonal to $\Delta(y)$ at the point $\hat{\omega}^n$, i.e., $\delta^n \in N(\hat{\omega}^n)$. Because of the second order growth condition (4.6) we get

$$\|\bar{\omega}^n - \hat{\omega}^n\| = \text{dist}(\bar{\omega}^n, \Delta(y)) = O(t_n). \tag{4.9}$$

By expanding $\bar{y}^n(\bar{\omega}^n)$ at $\hat{\omega}^n$, and since $y(\hat{\omega}^n) = 0$ and $\nabla y(\hat{\omega}^n) = 0$, we obtain

$$\nu(t_n) = \bar{y}^n(\bar{\omega}^n) = \bar{y}^n(\hat{\omega}^n) + \tfrac{1}{2}t_n^2\nabla^2 y(\hat{\omega}^n)(\delta^n, \delta^n) + t_n^2\nabla d(\hat{\omega}^n)\delta^n + o(t_n^2).$$

Since $y(\hat{\omega}^n) = 0$ and $d(\hat{\omega}^n) \geq 0$, it follows that

$$\nu(t_n) \geq \tfrac{1}{2}t_n^2 h(\hat{\omega}^n) + \tfrac{1}{2}t_n^2\nabla^2 y(\hat{\omega}^n)(\delta^n, \delta^n) + t_n^2\nabla d(\hat{\omega}^n)\delta^n + o(t_n^2). \tag{4.10}$$

Since $\hat{\omega}^n \to \hat{\omega}^0$, the continuity of the mapping

$$\omega \mapsto \min_{\delta\in N(\omega)} \{h(\omega) + \nabla^2 y(\omega)(\delta, \delta) + 2\nabla d(\omega)\delta\}$$

leads to

$$\liminf_{t\downarrow 0} \frac{\nu(t)}{t^2/2} \geq \min_{\omega\in\Delta^*(y,d)} \min_{\delta\in N(\omega)} \{h(\omega) + \nabla^2 y(\omega)(\delta, \delta) + 2\nabla d(\omega)\delta\}. \tag{4.11}$$

On the other hand, for any $\omega \in \Delta^*(y,d)$ and $\delta \in N(\omega)$ we have that $\nu(t) \leq \bar{y}_t(\omega + t\delta)$. Again using local coordinates, by expanding the right-hand side of this inequality, and since $y(\omega) = 0$, $\nabla y(\omega) = 0$, $d(\omega) = 0$, and $h(\omega + t\delta) = h(\omega) + o(1)$, we obtain that

$$\nu(t) \leq \tfrac{1}{2}t^2\{h(\omega) + \nabla^2 y(\omega)(\delta, \delta) + 2\nabla d(\omega)\delta\} + o(t^2), \tag{4.12}$$

which combined with (4.11) leads to

$$\lim_{t\downarrow 0}\frac{\nu(t)}{t^2/2}=\min_{\omega\in\Delta^*(y,d)}\min_{\delta\in N(\omega)}\{h(\omega)+\nabla^2 y(\omega)(\delta,\delta)+2\nabla d(\omega)\delta\}. \tag{4.13}$$

It follows that $O^2_K(y,d)=T^2_K(y,d)$ and $h\in T^2_K(y,d)$ if and only if for every $\omega\in\Delta^*(y,d)$,

$$h(\omega)+\min_{\delta\in N(\omega)}\{\nabla^2 y(\omega)(\delta,\delta)+2\nabla d(\omega)\delta\}\geq 0. \tag{4.14}$$

By calculating the minimum in (4.14) we obtain (4.8). We point out that, because of the second order growth condition (4.6), the minimum on $\delta\in N(\omega)$ is attained for $\|\delta\|\leq\|\nabla d\|_\infty/c$.

The proof of second order regularity involves similar arguments. Let $\Psi(\omega):=(\psi_1(\omega),...,\psi_m(\omega))^T$, and consider $t_k\downarrow 0$ and $y_k(\cdot):=y(\cdot)+t_kd(\cdot)+\frac{1}{2}t_k^2h_k(\cdot)\in K$, where $h_k\in C(\Omega)$ are such that $h_k(\cdot)=x_k^T\Psi(\cdot)+a_k(\cdot)$ with $C(\Omega)\ni a_k\to a$ and $t_kx_k\to 0$. Consider also $\nu_k:=\min_{\omega\in\Omega}y_k(\omega)$. Similarly to (4.12) we have that, given a system of local coordinates, for every $\omega\in\Delta^*(y,d)$ and $\delta\in N(\omega)$,

$$\nu_k\leq\tfrac{1}{2}t_k^2\left[h_k(\omega+t_k\delta)+\nabla^2 y(\omega)(\delta,\delta)+2\nabla d(\omega)\delta\right]+o(t_k^2). \tag{4.15}$$

Moreover, since $t_kx_k\to 0$ and $\Psi(\cdot)$ is Lipschitz continuous on Ω we have that

$$x_k^T\Psi(\omega+t_k\delta)=x_k^T\Psi(\omega)+o(1).$$

Also $a_k(\omega+t_k\delta)=a_k(\omega)+o(1)$ and hence

$$t_k^2h_k(\omega+t_k\delta)=t_k^2h_k(\omega)+o(t_k^2). \tag{4.16}$$

Since $y_k\in K$ and hence $\nu_k\geq 0$, we then obtain from (4.15) and (4.16) that

$$h_k(\omega)+\nabla^2 y(\omega)(\delta,\delta)+2\nabla d(\omega)\delta+o(1)\geq 0, \tag{4.17}$$

where (due to compactness of Ω) the term $o(1)$ can be taken uniformly in $\omega\in\Delta^*(y,d)$ and $\|\delta\|\leq\|\nabla d\|_\infty/c$. By using formula (4.14), we obtain from (4.17) that $h_k+o(1)\in O^2_K(y,d)$, which completes the proof. □

It follows from the above theorem that for semi-infinite programs with constraints of the form $g(x,\omega)\geq 0$, $\omega\in\Omega$, there is no gap between the corresponding second order necessary and sufficient conditions under the following conditions:

(i) $g(\cdot,\omega)$ is twice differentiable with $\nabla^2_{xx}g(x,\omega)$ being continuous on $X\times\Omega$,

(ii) Robinson's constraint qualification holds,

(iii) $g(x_0,\cdot)$ satisfies the second order growth condition (4.6),

(iv) Ω is a smooth compact manifold and $\Delta(g(x_0,\cdot))$ is a smooth submanifold of Ω,

(v) $g(x_0,\cdot)$ is twice continuously differentiable and the functions $\psi_i(\cdot)=\frac{\partial g}{\partial x_i}(x_0,\cdot)$ are continuously differentiable.

Note that since Ω is compact, the last assumption (v) implies that the functions $\psi_i(\cdot)$ are Lipschitz continuous on Ω. Also in the case of semi-infinite programming, Robinson's constraint qualification (postulated in the above condition (ii)) is equivalent to the extended Mangasarian–Fromovitz condition, that is, there exists $h\in X$ such that $h^T\nabla_x g(x_0,\omega)>0$ for all $\omega\in\Delta_0:=\Delta(g(x_0,\cdot))$ (e.g., [39]). We also observe that when the function $g(\cdot,\omega)$ is concave for every fixed $\omega\in\Omega$, the feasible set

$$\Phi:=\{x\in X: g(x,\omega)\geq 0 \text{ for all } \omega\in\Omega\}$$

is convex and Robinson's constraint qualification is equivalent to the Slater condition: there exists $\bar{x} \in X$ such that $g(\bar{x}, \omega) > 0$ for all $\omega \in \Omega$.

Combining Theorem 4.5 with Propositions 2.3 and 4.2 we deduce that, under assumptions (i)–(v) above, the set Φ is second order regular at x_0 and also second order directionally differentiable with

$$T^2_\Phi(x_0, h) = \{u \in X : \nabla_x g(x_0, \omega)u + \gamma(h, \omega) \geq 0 \text{ for all } \omega \in \Delta_1(h)\}, \tag{4.18}$$

where

$$\gamma(h, \omega) := \min_{\delta \in N(\omega)} \nabla^2 g(x_0, \omega)((h, \delta), (h, \delta)),$$

$\Delta_0 := \Delta(g(x_0, \cdot))$, and $\Delta_1(h) := \{\omega \in \Delta_0 : \nabla_x g(x_0, \omega)h = 0\}$. This formula may also be derived from Proposition 2.1 by using the characterization of second order directional derivatives of the min-function $\varphi(x) := \min_{\omega \in \Omega} g(x, \omega)$ given in [38, Theorem 4.1].

As an application, consider the example of semidefinite programming where

$$K = \mathcal{S}^n_+ := \{Z \in \mathcal{S}^n : g(Z, \omega) \geq 0 \text{ for all } \omega \in \Omega\}$$

with $g(Z, \omega) := \omega^T Z \omega$ and $\Omega := \{\omega \in \mathbb{R}^n : \|\omega\| = 1\}$. In this example the set Ω is a sphere, hence a smooth compact manifold. For a positive semidefinite matrix Z the corresponding contact set $\Delta(Z) := \{\omega \in \Omega : \omega^T Z \omega = 0\}$ is given by $\{\omega \in \Omega : Z\omega = 0\}$, which is a smooth submanifold of Ω. It is also not difficult to show that the corresponding second order growth condition holds (cf. [38]) and that the Lipschitz condition on functions ψ_i is automatically satisfied. Combining Theorem 4.5 and Proposition 2.3, we obtain the following result.

COROLLARY 4.6. *For any $n = 1, 2, ...,$ the cone $\mathcal{S}^n_+$ of symmetric positive semidefinite $n \times n$ matrices is second order directionally differentiable and second order regular at every point $Z \in \mathcal{S}^n_+$.*

An expression for the second order tangent sets and the corresponding second order optimality conditions for semidefinite optimization problems are given explicitly in [41].

5. Composite optimization. As we mentioned in the introduction, an alternative approach to derivation of second order optimality conditions is to consider composite functions as in the problem (1.2) and that such problems can be investigated in the form (1.3). In this transformation the corresponding convex function is replaced by its epigraph. In this section we translate results obtained in the previous sections into the framework of composite optimization and compare them with results discussed in some recent publications. We assume throughout this section that the function $g(\cdot)$ in (1.2) is convex, proper, and lower semicontinuous and the mapping $F : X \to Y$ is continuously differentiable.

Let $K := \text{epi}(g)$ and $G(x, c) := (F(x), c)$. Consider a point $(x_0, c_0) \in X \times \mathbb{R}$ such that $F(x_0) \in \text{dom}(g)$ and $c_0 = g(F(x_0))$, where $\text{dom}(g) := \{y \in Y : g(y) < +\infty\}$ is the domain of g. Note that $DG(x_0, c_0)(h, c) = (DF(x_0)h, c)$. Therefore Robinson's constraint qualification (2.11) at (x_0, c_0) becomes

$$0 \in \text{int}\{F(x_0) + DF(x_0)X - \text{dom}(g)\}. \tag{5.1}$$

Note that if $g(\cdot)$ is continuous at the point $y_0 := F(x_0)$, then $\text{dom}(g)$ contains a neighborhood of y_0, and hence in that case constraint qualification (5.1) holds. Robinson's constraint qualification (5.1) can be also written in the following equivalent form [45]:

$$Y = DF(x_0)X - \mathcal{R}_{\mathrm{dom}(g)}(F(x_0)), \tag{5.2}$$

where $\mathcal{R}_A(y) := \cup\{t(A-y) : t \geq 0\}$ denotes the radial cone to the convex set A at $y \in A$. By taking the polar cone of both sides of (5.2), and since the polar of $\mathcal{R}_A(y)$ is $N_A(y)$, we obtain that (5.2) implies the following condition:

$$\{0\} = [DF(x_0)X]^{\perp} \cap N_{\mathrm{dom}(g)}(F(x_0)). \tag{5.3}$$

If the space Y is finite dimensional, then (5.2) and (5.3) are equivalent. Constraint qualification (5.3) was used in [33] (in the finite dimensional case) and in [11], while (5.2) has been used, for instance, in [13, 27].

The Lagrangian of (1.3) is

$$L(x, c, \lambda, \gamma) := c + \langle \lambda, F(x) \rangle + \gamma c. \tag{5.4}$$

The tangent cone to epi(g) at the point $(F(x_0), c_0)$ is given by

$$T_{\mathrm{epi}(g)}(F(x_0), c_0) = \{(d, c) : g^{\downarrow}(F(x_0), d) \leq c\}. \tag{5.5}$$

Consequently the first order necessary condition (3.2) can be written in the form

$$[DF(x_0)]^* \lambda = 0, \ \gamma = -1, \ \langle \lambda, d \rangle \leq g^{\downarrow}(F(x_0), d) \ \text{ for all } d \in Y.$$

Since the epigraph of $g^{\downarrow}(F(x_0), \cdot)$ coincides with the topological closure of the epigraph of $g'(F(x_0), \cdot)$, we have that the condition $\langle \lambda, d \rangle \leq g^{\downarrow}(F(x_0), d)$ for all $d \in Y$ is equivalent to $\lambda \in \partial g(F(x_0))$, where $\partial g(F(x_0))$ is the subdifferential of $g(\cdot)$ at $F(x_0)$. Therefore the above first order necessary condition can be written in the following form: there exists $\lambda \in Y^*$ such that

$$[DF(x_0)]^* \lambda = 0, \quad \lambda \in \partial g(F(x_0)). \tag{5.6}$$

We obtain that if x_0 is a locally optimal solution of (1.2), then under constraint qualification (5.1) the set $\Lambda(x_0)$ of Lagrange multipliers satisfying (5.6) is nonempty and bounded. In the above form (5.6), first order necessary conditions in composite optimization were used in a number of publications [11, 13, 21, 26, 33].

DEFINITION 4. *Let $g(y)$ be a proper lower semicontinuous convex function with a finite value at a point $y_0 \in Y$. We say that $g(\cdot)$ is (outer) second order regular at y_0 if the set $K :=$ epi(g) is (outer) second order regular at the point $(y_0, g(y_0))$.*

The set epi(g) is defined by the constraint $h(y, c) \leq 0$, where $h(y, c) := g(y) - c$. Since g is proper, and hence its domain dom(g) is nonempty, we can find $\bar{y}$ and $\bar{c}$ such that $h(\bar{y}, \bar{c}) < 0$, i.e., the Slater condition always holds in the present situation. Now Proposition 4.4 implies the following result.

PROPOSITION 5.1. *Let $g(y)$ be a proper lower semicontinuous convex function. If g is finite and continuous at a point $y_0 \in Y$, then g is outer second order regular at y_0 if and only if, for every $d \in Y$ and every path $r : \mathbb{R}_+ \to Y$ satisfying* $\mathrm{t}r(t) \to 0$ *as $t \downarrow 0$, the inequality*

$$g\left(y_0 + td + \tfrac{1}{2}t^2 r(t)\right) \geq g(y_0) + tg'(y_0, d) + \tfrac{1}{2}t^2 g''_{-}(y_0; d, r(t)) + o(t^2) \tag{5.7}$$

holds.

Let $y \in \mathrm{dom}(g)$ be such that $g^{\downarrow}(y, d)$ is finite. Then it follows from Proposition 2.1 that

$$O^2_{\mathrm{epi}(g)}((y, g(y)), (d, g^{\downarrow}(y, d))) = \left\{(w, c) : g^{\downarrow\downarrow}_{-}(y; d, w) \le c\right\}. \tag{5.8}$$

Denote $\mathcal{T} := O^2_{\mathrm{epi}(g)}((y, g(y)), (d, g^{\downarrow}(y, d)))$, and $\psi(\cdot) := g^{\downarrow\downarrow}_{-}(y; d, \cdot)$. Then for $\lambda \in \Lambda(x_0)$ the corresponding sigma term becomes

$$\begin{array}{rcl} \sigma((\lambda, -1), \mathcal{T}) & = & \sup_{c,w}\{\langle \lambda, w\rangle - c : \psi(w) \le c\} \\ & = & \sup_{w}\{\langle \lambda, w\rangle - \psi(w)\} = \psi^*(\lambda), \end{array} \tag{5.9}$$

where ψ^* denotes the conjugate function of ψ.

Let us also note that the critical cone here can be written in the form

$$C(x_0, c_0) = \{(h, c) : g^{\downarrow}(F(x_0), DF(x_0)h) \le c,\ c = 0\},$$

provided constraint qualification (5.1) holds. Moreover, by the first order necessary conditions, $g^{\downarrow}(F(x_0), DF(x_0)h) \ge 0$ for any $h \in X$. Therefore this motivates us to consider the cone

$$\mathcal{C}(x_0) := \{h : g^{\downarrow}(F(x_0), DF(x_0)h) = 0\}. \tag{5.10}$$

Since $g^{\downarrow}(F(x_0), \cdot)$ is lower semicontinuous, this cone is closed. Combining Theorems 3.1 and 3.2 we get the following result.

THEOREM 5.2. *Suppose that $g(y)$ is a proper lower semicontinuous convex function, that $F : X \to Y$ is a twice continuously differentiable mapping, that $F(x_0) \in$* dom(g), *and that constraint qualification* (5.1) *holds. Then,*

(i) *(second order necessary condition) let x_0 be a locally optimal solution of* (1.2), *then for any $h \in \mathcal{C}(x_0)$ and any convex function $\phi(\cdot) \ge g^{\downarrow\downarrow}_{-}(F(x_0); DF(x_0)h, \cdot)$ the following inequality holds:*

$$\sup_{\lambda \in \Lambda(x_0)} \{\langle \lambda, D^2_{xx}F(x_0)(h, h)\rangle - \phi^*(\lambda)\} \ge 0; \tag{5.11}$$

(ii) *(second order sufficient condition) let x_0 be a stationary point of* (1.2), *i.e., it satisfies the first order necessary condition* (5.6), *and suppose that g is outer second order regular at $y_0 := F(x_0)$ and that*

$$\sup_{\lambda \in \Lambda(x_0)} \{\langle \lambda, D^2_{xx}F(x_0)(h, h)\rangle - \psi^*(\lambda)\} > 0 \text{ for all } h \in \mathcal{C}(x_0) \setminus \{0\}, \tag{5.12}$$

where $\psi(\cdot) := g^{\downarrow\downarrow}_{-}(F(x_0); DF(x_0)h, \cdot)$. Then for some $\alpha > 0$ and all x in a neighborhood of x_0,

$$g(F(x)) \ge g(F(x_0)) + \alpha\|x - x_0\|^2, \tag{5.13}$$

and hence x_0 is a locally optimal solution of (1.2).

It follows that if $g^{\downarrow\downarrow}_{-}(F(x_0); DF(x_0)h, \cdot)$ is convex and g is outer second order regular at $F(x_0)$, then there is no gap between second order necessary and sufficient conditions in the above theorem.

The second order optimality conditions of Theorem 5.2 are essentially equivalent to those obtained via second order (epi)subderivatives in [34, 13, 36], but they apply under different conditions.

For instance, in [34] (and subsequent work by the author) the function g is assumed to be piecewise linear-quadratic convex, a situation covered by Theorem 5.2 since such functions are second order regular. In order to check this we observe that any twice continuously differentiable convex function, in particular a quadratic convex function $g(y)$, is second order regular (see, e.g., Proposition 4.2 or 5.1). Now, if K is a polyhedral convex subset of Y and since the epigraph of the function $g(y) + I_K(y)$ is given by the intersection of the epigraph of g and $K \times \mathbb{R}$, it follows from Corollary 4.3 that $g(y)+I_K(y)$ is also second order regular. Finally, the epigraph of a piecewise linear-quadratic convex function is given by the union of a finite number of epigraphs of functions of the form $g(y)+I_K(y)$, with g being quadratic and K being polyhedral. It can be easily verified that union of a finite number of second order regular sets is also second order regular, from which the conclusion follows.

The results in [34] were extended in [13] beyond the class of piecewise linear-quadratic functions. The regularity condition used in that extension does not allow us to cover the second order regular case as in Theorem 5.2. (Among other things it requires some kind of local radiality of the domain of g, which is certainly not needed in our analysis.) However, it is also not clear whether second order regularity is weak enough to recover the results in [13].

The comparison with the results in [21] is more involved, since the optimality conditions are expressed in terms of lower second order epiderivatives instead of the parabolic ones as we express them. Of course, under suitable regularity assumptions both types of conditions can be shown to be equivalent thanks to the duality relation existing between both types of derivatives. However, in the general settings of [21] such duality relation cannot be ensured and the results are not comparable. The only exception concerns [21, Corollary 4], which is in fact a slight extension of results in [13], but, as in that paper, the regularity condition on which it is based is not comparable with second order regularity.

It is also possible to show that the second order regularity of g is a sufficient (but not necessary) condition for g to be parabolically regular. (See [36] for a discussion of the concept of parabolic regularity.) A detailed study of the relation between the concepts of second order regularity and parabolic regularity is given in the forthcoming book [10].

6. Extensions to nonisolated minima. Little is known about second order optimality conditions for nonisolated minima. A characterization of the second order growth condition is given in [8], under a constraint qualification, for smooth convex optimization problems with finitely many constraints. In [7] some sufficient conditions are stated for nonlinear programming problems. It is relatively easy to formulate a second order necessary condition that generalizes a result in [7].

Let $S \subset G^{-1}(K)$ be a set of optimal solutions (minimizers) of the problem (P), and let $\mathcal{T}_S(x) := \limsup_{t\downarrow 0} t^{-1}(S - x)$ be the contingent cone to S at x. It is easily checked that if $x \in S$ and $h \in X$, then $\operatorname{dist}(x + th, S) \geq t \operatorname{dist}(h, \mathcal{T}_S(x)) + o(t)$ for $t > 0$. Suppose that Robinson's constraint qualification holds at every point $x \in S$ and that S is compact. We have that if the second order growth condition holds at S, then for any feasible path $x(t)$ of the form (3.4) with $x_0 \in S$, $f(x(t)) \geq f(x_0) + ct^2 \operatorname{dist}(h, \mathcal{T}_S(x)) + o(t^2)$ for some $c > 0$ and $t > 0$ small enough. It then follows by the arguments used in the proof of Theorem 3.1 that a necessary condition for the second order growth (at S) is that there exists $c > 0$ such that for all $x \in S$ and $h \in C(x)$,

$$\sup_{\lambda \in \Lambda(x)} \{D_{xx}^2 L(x,\lambda)(h,h) - \sigma(\lambda, \mathcal{T}(x,h))\} \geq 2ct^2 \operatorname{dist}(h, \mathcal{T}_S(x)), \tag{6.1}$$

where $\mathcal{T}(x,h)$ is a convex subset of $O_K^2(G(x), DG(x)h)$. Recall that the set of proximal normals to S at $x \in S$ is defined as

$$\mathcal{N}_S(x) := \{h \in X;\ \operatorname{dist}(x + th, S) = t\|h\| \text{ for some } t > 0\},$$

and set $\mathcal{N}_S^\varepsilon(x) := \{h \in X;\ \operatorname{dist}(h, \mathcal{N}_S(x)) \leq \varepsilon\}$. As $\operatorname{dist}(h, \mathcal{T}_S(x)) = \|h\|$ whenever $h \in \mathcal{N}_S(x)$, a consequence of (6.1), and therefore a necessary condition for quadratic growth (see [7]), is that for $\varepsilon > 0$ small enough

$$\sup_{\lambda \in \Lambda(x)} \{D_{xx}^2 L(x,\lambda)(h,h) - \sigma(\lambda, \mathcal{T}(x,h))\} \geq c\|h\|^2 \text{ for all } h \in C(x) \cap \mathcal{N}_S^\varepsilon(x). \tag{6.2}$$

DEFINITION 5. *We say that the set S satisfies a property of uniform approximation of critical cones if for every $\varepsilon > 0$ there exists $\alpha > 0$ such that for all $x \in S$ and $h \in X$ satisfying $Df(x)h \leq \alpha\|h\|$ and $DG(x)h \in T_K(G(x)) + \alpha\|h\|B_Y$, we have* $\operatorname{dist}(h, C(x)) \leq \varepsilon\|h\|$.

DEFINITION 6. *We say that K is uniformly regular with respect to the set S and the mapping $G(x)$ if for $x \in S$ and $h \in C(x)$, $O_K^2(G(x), DG(x)h)$ is an upper second order approximation set for K at the point $G(x)$ in the direction $DG(x)h$ with respect to $DG(x)$ uniformly over S. That is, if $x_k \in S$, $h_k \in C(x_k)$, $t_k \downarrow 0$, and $r_k = DG(x_k)z_k + a_k$ are sequences such that $\{a_k\}$ is convergent, $t_k z_k \to 0$ and $G(x_k) + t_k DG(x_k)h_k + \frac{1}{2}t_k^2 r_k \in K$, then*

$$\lim_{k\to\infty} \operatorname{dist}(r_k, O_K^2(G(x_k), DG(x_k)h_k)) = 0. \tag{6.3}$$

THEOREM 6.1. *Let $S \subset G^{-1}(K)$ satisfy the property of uniform approximation of critical cones, and suppose that Robinson's constraint qualification holds at every point $x \in S$, that S is compact, that K is uniformly regular with respect to the set S and the mapping $G(x)$, and that $O_K^2(G(x), DG(x)h) = T_K^2(G(x), DG(x)h)$ for all $x \in S$ and $h \in C(x) \setminus \mathcal{T}_S(x)$. Then condition* (6.2) *is necessary and sufficient for the second order growth at S.*

Proof. We already observed that the condition is necessary. It suffices therefore to prove that it is sufficient. Let x_k be a sequence of feasible points $x_k \in G^{-1}(K)$ converging to a point $x_0 \in S$ and such that (3.16) holds. Let $\hat{x}_k$ be a projection of x_k onto S, i.e., $\hat{x}_k \in S$ and $\|x_k - \hat{x}_k\| = \operatorname{dist}(x_k, S)$. Set $t_k := \|x_k - \hat{x}_k\|$ and $\hat{h}_k := (x_k - \hat{x}_k)/t_k$. Then $\hat{h}_k \in \mathcal{N}_S(\hat{x}_k)$. From the property of uniform approximation of critical cones, there exists h_k such that $h_k \in C(\hat{x}_k)$ and $\|h_k - \hat{h}_k\| \to 0$, hence for all $\varepsilon > 0$, $h_k \in \mathcal{N}_S^\varepsilon(\hat{x}_k)$ for large enough k. Then $x_k = \hat{x}_k + t_k h_k + o(t_k)$. The remainder of the proof is similar to the one of Theorem 3.2. ∎

It was proved in [8] that the property of uniform approximation of critical cones is satisfied for finitely constrained convex optimization problems. Whether this property holds in more general settings is an open problem.

Acknowledgment. The authors thank the two referees for their useful remarks.

REFERENCES

[1] A. BEN-TAL, M. TEBOULLE, AND J. ZOWE, *Second order necessary optimality conditions for semi-infinite programming problems*, in Semi-infinite Programming, Lecture Notes in Control and Inform. Sci. 15, R. Hettich, ed., Springer-Verlag, Berlin, 1979, pp. 17–30.

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

[3] A. Ben-Tal and J. Zowe, *A unified theory of first and second order conditions for extremum problems in topological vector spaces*, Math. Programming Study, 19 (1982), pp. 39–76.

[4] A. Ben-Tal and A. Nemirovski, *Robust solutions of uncertain linear programs*, Oper. Res. Lett., to appear.

[5] J. F. Bonnans and R. Cominetti, *Perturbed optimization in Banach spaces* I: *A general theory based on a weak directional constraint qualification,* II: *A theory based on a strong directional qualification,* III: *Semi-infinite optimization*, SIAM J. Control Optim., 34 (1996), pp. 1151–1171, 1172–1189, and 1555–1567.

[6] J. F. Bonnans, R. Cominetti, and A. Shapiro, *Sensitivity analysis of optimization problems under second order regular constraints*, Math. Oper. Res., to appear.

[7] J. F. Bonnans and A. D. Ioffe, *Second-order sufficiency and quadratic growth for non isolated minima*, Math. Oper. Res., 20 (1995), pp. 801–817.

[8] J. F. Bonnans and A. D. Ioffe, *Quadratic growth and stability in convex programming problems with multiple solutions*, J. Convex Anal., 2 (1995), pp. 41–57.

[9] J. F. Bonnans and A. Shapiro, *Optimization problems with perturbations: A guided tour*, SIAM Rev., 40 (1998), pp. 228–264.

[10] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York, to appear.

[11] J. V. Burke and R. A. Poliquin, *Optimality conditions for non-finite valued convex composite functions*, Math. Programming, 57 (1992), pp. 103–120.

[12] R. Cominetti, *Metric regularity, tangent sets and second order optimality conditions*, Appl. Math. Optim., 21 (1990), pp. 265–287.

[13] R. Cominetti, *On pseudo-differentiability*, Trans. Amer. Math. Soc., 324 (1991), pp. 843–865.

[14] R. Cominetti and J-P. Penot, *Tangent sets to unilateral convex sets*, C. R. Acad. Sci. Sér. I Math., 321 (1995), pp. 1631–1636.

[15] A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968.

[16] R. P. Hettich and H. Th. Jongen, *Semi-infinite programming: Conditions of optimality and applications*, in Optimization Techniques, Proc. 8th IFIP Conf. on Optimization Techniques, Würzburg, Part 2, J. Stoer, ed., Springer-Verlag, New York, 1977.

[17] R. Hettich and P. Zencke, *Numerische Methoden der Approximation und Semi-infiniten Optimierung*, Teubner, Stuttgart, 1982.

[18] R. Hettich and K. O. Kortanek, *Semi-infinite programming: Theory, methods and applications*, SIAM Rev., 35 (1993), pp. 380–429.

[19] A. D. Ioffe, *Necessary and sufficient conditions for a local minimum.* 3: *Second order conditions and augmented duality*, SIAM J. Control Optim., 17 (1979), pp. 266–288.

[20] A. D. Ioffe, *On some recent developments in the theory of second order optimality conditions*, in Optimization, Lecture Notes in Math. 1405, S. Dolecki, ed., Springer-Verlag, Berlin, 1989, pp. 55–68.

[21] A. D. Ioffe, *Variational analysis of a composite function: A formula for the lower second order epi-derivative*, J. Math. Anal. Appl., 160 (1991), pp. 379–405.

[22] A. D. Ioffe, *On sensitivity analysis of nonlinear programs in Banach spaces: The approach via composite unconstrained optimization*, SIAM J. Optim., 4 (1994), pp. 1–43.

[23] H. Kawasaki, *An envelope-like effect of infinitely many inequality constraints on second order necessary conditions for minimization problems*, Math. Programming, 41 (1988), pp. 73–96.

[24] S. Kurcyusz, *On the existence and nonexistence of Lagrange multipliers in Banach spaces*, J. Optim. Theory Appl., 20 (1976), pp. 81–110.

[25] Zs. Pales and V. M. Zeidan, *Nonsmooth optimum problems with constraints*, SIAM J. Control Optim., 32 (1994), pp. 1476–1502.

[26] J. P. Penot, *Optimality Conditions for Minimax Problems, Semi-infinite Programming Problems and Their Relatives*, Report 92/16, UPRA, Laboratoire de Math. Appl., Av. de l'Université, 64000 Pau, France, 1992.

[27] J. P. Penot, *Optimality conditions in mathematical programming and composite optimization*, Math. Programming, 67 (1994), pp. 225–245.

[28] S. M. Robinson, *First order conditions for general nonlinear optimization*, SIAM J. Appl. Math., 30 (1976), pp. 597–607.

[29] S. M. Robinson, *Stability theory for systems of inequalities, Part* II: *Differentiable nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513.

[30] S. M. Robinson, *Regularity and stability for convex multivalued functions*, Math. Oper. Res.,

1 (1976), pp. 130–143.
[31] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[32] R. T. Rockafellar, *Conjugate Duality and Optimization*, CBMS-NSF Regional Conference Series in Applied Mathematics 16, SIAM, Philadelphia, PA, 1974.
[33] R. T. Rockafellar, *First and second-order epi-differentiability in nonlinear programming*, Trans. Amer. Math. Soc., 307 (1988), pp. 75–108.
[34] R. T. Rockafellar, *Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives*, Math. Oper. Res., 14 (1989), pp. 462–484.
[35] R. T. Rockafellar, *Nonsmooth analysis and parametric optimization*, in Methods of Nonconvex Analysis, Lecture Notes in Math. 1446, A. Cellina, ed., Springer-Verlag, Berlin, 1990, pp. 137–151.
[36] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1997.
[37] A. Shapiro, *Second-order derivatives of extremal-value functions and optimality conditions for semi-infinite programs*, Math. Oper. Res., 10 (1985), pp. 207–219.
[38] A. Shapiro, *Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton*, Appl. Math. Optim., 18 (1988), pp. 215–229.
[39] A. Shapiro, *On Lipschitzian stability of optimal solutions of parametrized semi-infinite programs*, Math. Oper. Res., 19 (1994), pp. 743–752.
[40] A. Shapiro and M. K. H. Fan, *On eigenvalue optimization*, SIAM J. Optim., 5 (1995), pp. 552–569.
[41] A. Shapiro, *First and second order analysis of nonlinear semidefinite programs*, Math. Programming, Series B, 77 (1997), pp. 301–320.
[42] C. Ursescu, *Multifunctions with convex closed graph*, Czechoslovak Math. J., 25 (1975), pp. 438–441.
[43] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev., 38 (1996), pp. 49–95.
[44] W. Wetterling, *Definitheitsbedingungen fürrelative Extrema bei Optimieungs- und Approximation-saufgaben*, Numer. Math., 15 (1879), pp. 122–136.
[45] J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62.