MATHEMATICS OF OPERATIONS RESEARCH
Vol. 30, No. 1, February 2005, pp. 109–126
ISSN 0364-765X | EISSN 1526-5471 | 05 | 3001 | 0109

DOI 10.1287/moor.1040.0115

Sensitivity Analysis of Parameterized Variational Inequalities

Alexander Shapiro
School of Industrial and Systems Engineering, Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA, ashapiro@isye.gatech.edu

In this paper we discuss local uniqueness, continuity, and differentiability properties of solutions of parameterized variational inequalities (generalized equations). To this end we use two types of techniques. One approach consists in formulating variational inequalities in a form of optimization problem based on regularized gap functions, and applying a general theory of perturbation analysis of parameterized optimization problems. Another approach is based on a theory of contingent (outer graphical) derivatives and some results about differentiability properties of metric projections.

Key words: variational inequalities; gap functions; sensitivity analysis; second order regularity; quadratic growth condition; locally upper Lipschitz and Hölder continuity; directional differentiability; prox-regularity; graphical derivatives
MSC2000 subject classification: Primary: 90C31
OR/MS subject classification: Programming (complementarity)
History: Received June 7, 2003; revised February 9, 2004.

1. Introduction. In this paper we discuss continuity and differentiability properties of solutions of the parameterized variational inequalities

$$F(x, u) \in N_K(x). \tag{1.1}$$

Here K is a nonempty closed subset of $\mathbb{R}^n$, F: $\mathbb{R}^n \times U \to \mathbb{R}^n$ is a mapping, U is a normed space, and $N_K(\bar{x})$ denotes a normal cone to K at $\bar{x}$. In case the set K is not convex, there are several concepts of normal cones available in the literature. To be specific we assume that $N_K(\bar{x})$ is given by the polar (negative dual) of the contingent (Bouligand) cone to K at $\bar{x} \in K$, and $N_K(\bar{x}) := \varnothing$ if $\bar{x} \notin K$. If the set K is convex, then this definition coincides with the standard notion

$$N_K(\bar{x}) := \{y\colon \langle y, x - \bar{x}\rangle \le 0,\ \forall\, x \in K\}$$

of the normal cone at the point $\bar{x} \in K$ (the notation $\langle \cdot, \cdot \rangle$ stands for the standard scalar product in $\mathbb{R}^n$). We denote variational inequality (1.1) by $\mathrm{VI}(K, F_u)$ and by $\mathrm{Sol}(K, F_u)$ its set of solutions; i.e., $\bar{x} \in \mathrm{Sol}(K, F_u)$ iff $\bar{x} \in K$ and $F(\bar{x}, u) \in N_K(\bar{x})$. For a reference value $u_0 \in U$ of the parameter vector, we often drop the corresponding subscript in the above notation. In particular, $F(\cdot) := F(\cdot, u_0)$ and $\mathrm{VI}(K, F)$ corresponds to the reference variational inequality

$$F(x) \in N_K(x). \tag{1.2}$$

It is well known that for optimization problems, $\mathrm{VI}(K, F)$ represents first order optimality conditions. That is, consider the optimization problem

$$\operatorname*{Min}_{x \in K} f(x), \tag{1.3}$$

where f: $\mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. We have that if x_0 is a locally optimal solution of (1.3), then $x_0 \in \mathrm{Sol}(K, F)$ with $F(\cdot) := -\nabla f(\cdot)$.

There is a large literature on all aspects of the theory of variational inequalities. It will be beyond the scope of this paper to give a survey of all relevant results. In that respect we may refer to the recent comprehensive monograph by Facchinei and Pang [8] and references therein. We also do not investigate existence of solutions of $\mathrm{VI}(K, F_u)$ and concentrate on continuity and differentiability properties of such solutions if they do exist. For a discussion of various conditions ensuring nonemptiness of the solution set $\mathrm{Sol}(K, F_u)$ we refer to

Facchinei and Pang [8, §2.2]. For example, a simple sufficient condition for existence of a solution of $\mathrm{VI}(K, F_u)$ is that $F(\cdot, u)$ is continuous and the set K is convex and bounded, and hence compact (see Kinderlehrer and Stampacchia [14]). In case the set K is *polyhedral*, a perturbation theory of $\mathrm{VI}(K, F_u)$ is thoroughly developed notably by Robinson [20, 21, 22], Levy and Rockafellar [17], Dontchev and Rockafellar [7], Levy [16], and Klatte and Kummer [13]. Much less is known about continuity and differentiability properties of the solution multifunction $u \mapsto \mathrm{Sol}(K, F_u)$ for a general, not necessarily polyhedral, set K. Some results of that type were obtained in Shapiro [29] by a reduction approach. It is clear from a general perturbation theory of parameterized optimization problems (see Bonnans and Shapiro [4] and references therein) and results presented in Shapiro [29] that an additional term, representing curvature of the set K, should appear in the corresponding formulas.

In contrast to variational inequalities, sensitivity analysis of locally optimal solutions of optimization problems is quite well developed. One of the reasons for this discrepancy is that powerful tools of duality theory cannot be directly applied to an analysis of variational inequalities. In this paper we investigate properties of the solution mapping $u \mapsto \mathrm{Sol}(K, F_u)$ by using two, somewhat different, techniques. One approach is based on reducing the analysis to a study of optimization problems related to the corresponding (regularized) gap functions. Another methodology approaches the problem by using tools of various derivative and coderivative mappings. For a discussion of that approach we may refer to a survey paper by Levy [16] and references therein.

For perturbation theory of optimization problems we use, as a reference, Bonnans and Shapiro [4]. In particular, we discuss the case where the set $K = K(u)$ may depend on u and is given in the form

$$K(u) := \{x\colon G(x, u) \in Q\}, \tag{1.4}$$

where $G\colon \mathbb{R}^n \times U \to \mathbb{R}^m$ is a continuously differentiable mapping and $Q \subset \mathbb{R}^m$ is a closed *convex* set. For the polyhedral set $Q := \{0\} \times \mathbb{R}^m_+ \subset \mathbb{R}^n$, sensitivity analysis of such variational inequalities (generalized equations) was developed in Levy [15], Klatte [11] and Klatte and Kummer [12, 13]. The case of a general polyhedral set Q was developed in Robinson [23] under a nondegeneracy condition. Shapiro [29] studied the case where the set Q is cone reducible and a nondegeneracy condition holds.

This paper is organized as follows. In §2 we discuss some general properties of (regularized) gap functions which allow us to reformulate variational inequalities into respective optimization problems. In §3 we study relations between optimization problems associated with gap functions and local properties of the corresponding variational inequalities (generalized equations). Results of these two sections may have an independent interest. In §§4 and 5 we investigate continuity and differentiability properties of solutions of parameterized variational inequalities and generalized equations in terms of the corresponding contingent derivative multifunctions.

We use the following notation and terminology. The space $\mathbb{R}^n$ is equipped with the Euclidean norm $\|x\| := \langle x, x\rangle^{1/2}$. It is said that $F(\cdot, \cdot)$ is directionally differentiable at a point $(x_0, u_0) \in \mathbb{R}^n \times U$, if the limit

$$F'((x_0, u_0), (h, p)) := \lim_{t \downarrow 0} \frac{F(x_0 + th, u_0 + tp) - F(x_0, u_0)}{t} \tag{1.5}$$

exists for all $(h, p) \in \mathbb{R}^n \times U$. Unless stated otherwise we assume that $F(\cdot, \cdot)$ is locally Lipschitz continuous. Then it follows from (1.5) that $F(\cdot, \cdot)$ is Hadamard directionally differentiable at (x_0, u_0); i.e.,

$$F'((x_0, u_0), (h, p)) = \lim_{k \to \infty} \frac{F(x_0 + t_k h_k, u_0 + t_k p_k) - F(x_0, u_0)}{t_k} \tag{1.6}$$

for any sequences $t_k \downarrow 0$, $h_k \to h$ and $p_k \to p$. If $F(\cdot, \cdot)$ is differentiable, we denote by $DF(x, u)$ its differential at a point (x, u); i.e., $DF(x, u)(h, p) = \nabla_x F(x, u)h + \nabla_u F(x, u)p$. By $D^2G(x)(h, h)$ we denote the quadratic form associated with the second order derivative of a mapping $G: \mathbb{R}^n \to \mathbb{R}^m$ at x. For a point $x \in \mathbb{R}^n$ we denote by $B(x, r) := \{y: \|y - x\| \le r\}$ the ball of radius r centered at x, and by $\operatorname{dist}(x, K)$ the distance from x to the set K. By $P_K(x)$ we denote the metric projection of x onto K; i.e., $P_K(x)$ is a closest point of K to x. Of course, $\operatorname{dist}(x, K) = \|x - P_K(x)\|$. Since the set K is closed, the metric projection $P_K(x)$ always exists although may not be unique if the set K is not convex. The notation $T_K(x)$ stands for the contingent (Bouligand) cone to K at x. By the definition $T_K(x) = \emptyset$ if $x \notin K$. As we mentioned earlier, the normal cone $N_K(\bar{x})$, at $\bar{x} \in K$, is defined as the negative dual of $T_K(\bar{x})$; that is, $N_K(\bar{x}) := \{y: \langle y, x\rangle \le 0, \ \forall x \in T_K(\bar{x})\}$.

For a set $S \subset \mathbb{R}^n$ we denote by $\operatorname{cl}(S)$ its topological closure, by

$$I_S(x) := \begin{cases} 0, & \text{if } x \in S, \\ +\infty, & \text{if } x \notin S, \end{cases} \tag{1.7}$$

its indicator function, and by

$$\sigma(x, S) := \sup_{y \in S} \langle x, y\rangle \tag{1.8}$$

its support function. An extended real valued function $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be proper if $f(x) > -\infty$ for all $x \in \mathbb{R}^n$ and its domain $\operatorname{dom} f := \{x: f(x) < +\infty\}$ is nonempty. If f is convex, we denote by $\partial f(x)$ its subdifferential at $x \in \operatorname{dom} f$. For a linear mapping $A: \mathbb{R}^n \to \mathbb{R}^m$ we denote by $A^*: \mathbb{R}^m \to \mathbb{R}^n$ its adjoint mapping; i.e., $\langle y, Ax\rangle = \langle A^* y, x\rangle$ for any $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. By $\mathcal{S}^n$ we denote the space of $n \times n$ symmetric matrices.

2. Preliminary results. We assume in this and the following sections that the set K and mapping $F(x)$ are independent of u, in particular, if K is defined in the form (1.4), we assume that the mapping $G(x)$ is independent of u. We introduce a gap function, associated with $\operatorname{VI}(K, F)$, and discuss its basic properties. This, in itself, may have an independent interest. Consider a (reference) point $x_0 \in K$. For a constant $\kappa > 0$ and a neighborhood V of x_0 consider the following (*regularized*) *gap* function:

$$\gamma_\kappa(x) := \inf_{y \in K \cap V} \left\{\langle F(x), x - y\rangle + \tfrac{1}{2}\kappa \|x - y\|^2\right\}. \tag{2.1}$$

For convex set K, such regularized gap functions were introduced by Auchmuty [1] and Fukushima [9], and are discussed in Facchinei and Pang [8, §10.2.1]. Let us remark that the subsequent analysis is local in nature and the above gap function also depends on the neighborhood V. For the sake of simplicity we suppress this dependence in the notation. If the set K is convex, then the optimization problem in the right-hand side of (2.1) is convex, and therefore in that case the restriction to a neighborhood of x_0 can be removed.

Clearly for $y = x$ we have that the value of the function in the right-hand side of (2.1) is zero, and hence $\gamma_\kappa(x) \le 0$ for all $x \in K \cap V$. It is known that if the set K is convex and $V = \mathbb{R}^n$, then a point $\bar{x} \in K$ is a solution of $\operatorname{VI}(K, F)$ iff $\gamma_\kappa(\bar{x}) = 0$. That is, $\bar{x} \in \operatorname{Sol}(K, F)$ iff $\gamma_\kappa(\bar{x}) = 0$ and $\bar{x}$ is an optimal solution of the optimization problem

$$\operatorname*{Max}_{x \in K \cap V} \gamma_\kappa(x). \tag{2.2}$$

We extend this result to nonconvex sets K. In order to proceed we need the following concept.

Definition 2.1. It is said that the set K is *prox-regular* at x_0 if there exists a neighborhood W of x_0 and a positive constant α such that

$$\operatorname{dist}(y - x, T_K(x)) \le \alpha \|y - x\|^2 \quad \text{for all } x, y \in K \cap W. \tag{2.3}$$

Property (2.3) was introduced in Shapiro [26] under the name "$O(2)$-convexity." The term "prox-regularity" was suggested in Poliquin and Rockafellar [19] and Rockafellar and Wets [25] where this concept was defined in a somewhat different, although equivalent, form. Any convex set K is prox-regular at its every point. If the set K is given in the form (1.4), i.e., $K := G^{-1}(Q)$, then K is prox-regular at x_0 provided that Robinson's constraint qualification holds and $DG(\cdot)$ is Lipschitz continuous in a neighborhood of x_0 (see Shapiro [26, pp. 134–135]). Prox-regular sets have the following important properties.

Proposition 2.1. *Suppose that K is prox-regular at $x_0 \in K$. Then K is Clarke regular at x_0, the cone $T_K(x_0)$ is convex, and $P_K(x)$ is unique and locally Lipschitz continuous for all x in a neighborhood of x_0.*

For a definition of Clarke regularity (of sets) see, e.g., Rockafellar and Wets [25, Definition 6.4]. The implication that Clarke regularity follows from prox-regularity is shown in Shapiro [28]. Clarke regularity, in turn, implies that $T_K(x_0)$ is convex. The implication that $P_K(x)$ is unique and locally Lipschitz continuous is shown in Shapiro [26, Theorem 2.2].

Let us discuss now some properties of the function $\gamma_\kappa(\cdot)$. We assume in what follows that the set K is prox-regular at the point x_0. It will be convenient to write the gap function in the form $\gamma_\kappa(x) = \vartheta_\kappa(x, F(x))$, where

$$\vartheta_\kappa(x, a) := \inf_{y \in K \cap V} \left\{ g_\kappa(x, a, y) := \langle a, x - y\rangle + \tfrac{1}{2}\kappa \|x - y\|^2 \right\}. \tag{2.4}$$

The function $g_\kappa(x, a, y)$ can also be written as follows:

$$g_\kappa(x, a, y) = \tfrac{1}{2}\kappa \|x + \kappa^{-1} a - y\|^2 - \tfrac{1}{2}\kappa^{-1}\|a\|^2. \tag{2.5}$$

The following properties are assumed to hold *locally*, i.e., for x in a neighborhood of x_0, bounded a and sufficiently large κ. By Proposition 2.1 we have that the minimization problem in the right-hand side of (2.4) has unique optimal solution

$$\bar{y}_\kappa(x, a) = P_K(x + \kappa^{-1} a), \tag{2.6}$$

and $\bar{y}_\kappa(\cdot,\cdot)$ is locally Lipschitz continuous. Consequently, $\gamma_\kappa(\cdot)$ is real valued and Lipschitz continuous in a neighborhood of x_0. Also it follows by Danskin's Theorem [6], that $\vartheta_\kappa(\cdot,\cdot)$ is differentiable with $D\vartheta_\kappa(x, a) = Dg_\kappa(x, a, \bar{y})$, where $\bar{y} = \bar{y}_\kappa(x, a)$. By straightforward calculations we obtain that

$$\nabla_x \vartheta_\kappa(x, a) = a + \kappa(x - \bar{y}) \quad \text{and} \quad \nabla_a \vartheta_\kappa(x, a) = x - \bar{y}, \tag{2.7}$$

and hence $D\vartheta_\kappa(\cdot,\cdot)$ is locally Lipschitz continuous. Let us also observe that if $x \in K$ and $a \in N_K(x)$, then $P_K(x + \kappa^{-1} a) = x$, and hence $\bar{y}_\kappa(x, a) = x$. In particular, if $x_0 \in \mathrm{Sol}(K, F)$, then $\bar{y}_\kappa(x_0, F(x_0)) = x_0$.

Suppose that $\bar{x} \in \mathrm{Sol}(K, F)$, and let $x = \bar{x} + h$. We have that $\bar{y}_\kappa(\bar{x}, F(\bar{x})) = \bar{x}$, and since $F(\cdot)$ is assumed to be locally Lipschitz continuous, $F(x) = F(\bar{x}) + r(h)$ with $r(h) = O(\|h\|)$. Consequently, for $\bar{a} := F(\bar{x})$,

$$\begin{aligned} \gamma_\kappa(x) - \gamma_\kappa(\bar{x}) &= \vartheta_\kappa(\bar{x} + h, \bar{a} + r(h)) - \vartheta_\kappa(\bar{x}, \bar{a}) \\ &= D\vartheta_\kappa(\bar{x}, \bar{a})(h, r(h)) + O(\|h\|^2 + \|r(h)\|^2) \\ &= \langle F(\bar{x}), h\rangle + O(\|h\|^2). \end{aligned} \tag{2.8}$$

It follows that $\gamma_\kappa(\cdot)$ is differentiable at $\bar{x}$ and $\nabla \gamma_\kappa(\bar{x}) = F(\bar{x})$. Note again that the above properties hold locally and for κ large enough.

The following proposition is an extension of the corresponding results for the case of convex set K (Auchmuty [1], Fukushima [9], Facchinei and Pang [8, Theorem 10.2.3]).

Proposition 2.2. *Suppose that the set K is prox-regular at the point x_0. Then there exist constant $\kappa > 0$ and a neighborhood V of x_0 such that for any $\bar{x} \in V$ we have that $\bar{x} \in \mathrm{Sol}(K, F)$ iff $\bar{x} \in K$ and $\gamma_\kappa(\bar{x}) = 0$.*

Proof. Consider a point $\bar{x} \in \mathrm{Sol}(K, F)$. By the definition we have that $\bar{x} \in K$ and $F(\bar{x})$ belongs to the negative dual of the cone $T_K(\bar{x})$. It follows that for any $h \in \mathbb{R}^n$,

$$\langle F(\bar{x}), h\rangle \le \|F(\bar{x})\| \,\mathrm{dist}(h, T_K(\bar{x})).$$

Together with (2.3) this implies that for $\bar{x} \in V$ and V sufficiently small,

$$\langle F(\bar{x}), \bar{x} - y \rangle \geq -\alpha \|F(\bar{x})\| \|y - \bar{x}\|^2, \quad \forall\, y \in K \cap V. \tag{2.9}$$

Moreover, we can choose the neighborhood V such that $\|F(x)\| \leq \beta$ for some $\beta > 0$ and all $x \in V$. Then for $\kappa > 2\alpha\beta$, we obtain that the function $g_\kappa(\bar{x}, F(\bar{x}), \cdot)$ attains its minimum over $K \cap V$ at $\bar{y} = \bar{x}$, and hence $\gamma_\kappa(\bar{x}) = 0$.

Conversely, suppose that $\bar{x} \in K \cap V$ and $\gamma_\kappa(\bar{x}) = 0$. Then $\bar{y} = \bar{x}$ is an optimal solution of the right-hand side of (2.1). By (2.7) and derivations similar to (2.8), we have that for κ large enough and $\bar{x}$ sufficiently close to x_0, the minimizer $\bar{y}$ is unique and $\nabla_x \gamma_\kappa(\bar{x}) = F(\bar{x})$. It follows then by first order optimality conditions that $F(\bar{x}) \in N_K(\bar{x})$, and hence $\bar{x} \in \mathrm{Sol}(K, F)$. □

Let us now discuss second order differentiability properties of $\vartheta_\kappa(\cdot, \cdot)$ and $\gamma_\kappa(\cdot)$. In general the metric projection $P_K(\cdot)$ is not differentiable everywhere, and hence $\vartheta_\kappa(\cdot, \cdot)$ is not twice differentiable, even if K is convex. Therefore we study second order directional derivatives of $\vartheta_\kappa(\cdot, \cdot)$ and $\gamma_\kappa(\cdot)$. It is said that a differentiable function $f\colon \mathbb{R}^n \to \mathbb{R}$ is *second order Hadamard directionally differentiable*, at a point x, if for any $h \in \mathbb{R}^n$ the limit

$$\lim_{\substack{t \downarrow 0 \\ h' \to h}} \frac{f(x + th') - f(x) - tDf(x)h'}{\frac{1}{2}t^2}$$

exists. In case it exists, we denote this limit by $f''(x, h)$. Note that if $f''(x, h)$ exists, it is continuous in h.

Consider the second order tangent set to K at a point $x \in K$ in a direction h:

$$T_K^2(x, h) := \left\{ w \in \mathbb{R}^n\colon\ \mathrm{dist}\left(x + th + \tfrac{1}{2}t^2 w, K \right) = o(t^2) \right\}.$$

Note that the set $T_K^2(x, h)$ can be nonempty only if $h \in T_K(x)$, and for any $t > 0$,

$$T_K^2(x, th) = t^2 T_K^2(x, h). \tag{2.10}$$

Definition 2.2. It is said that the set K is *second order regular*, at a point $\bar{x} \in K$, if for any $h \in T_K(\bar{x})$ and any sequence $x_k \in K$ of the form $x_k := \bar{x} + t_k h + \frac{1}{2}t_k^2 w_k$, where $t_k \downarrow 0$ and $t_k w_k \to 0$, the following condition holds:

$$\lim_{k \to \infty} \mathrm{dist}(w_k, T_K^2(\bar{x}, h)) = 0.$$

The above concept of second order regularity (for general, not necessarily convex, sets) was developed in Bonnans et al. [3] and Bonnans and Shapiro [4]. The class of second order regular sets is quite large; it contains polyhedral sets, cones of positive semidefinite matrices, etc. Note that second order regularity of K at $\bar{x}$ implies that $T_K^2(\bar{x}, h)$ is nonempty for any $h \in T_K(\bar{x})$.

Consider the optimization (minimization) problem in the right-hand side of (2.4). At a point $(x_0, a_0, \bar{y})$, where $x_0 \in K$, $a_0 \in N_K(x_0)$, and $\bar{y} = x_0$, the function $g_\kappa(\cdot, \cdot, \cdot)$ has the following second order Taylor expansion:

$$g_\kappa(x_0 + d, a_0 + r, \bar{y} + h) = \langle a_0, d - h \rangle + \langle r, d - h \rangle + \tfrac{1}{2}\kappa \|d - h\|^2. \tag{2.11}$$

(The above expansion is exact since $g_\kappa(\cdot, \cdot, \cdot)$ is a quadratic function.) Therefore, the corresponding so-called *critical cone*, associated with the problem of minimization of $g_\kappa(x_0, a_0, \cdot)$ over the set $K \cap V$, is defined as

$$C(x_0) := \{h\colon \langle F(x_0), h \rangle = 0,\ h \in T_K(x_0)\}. \tag{2.12}$$

Recall that if K is prox-regular at x_0, then $T_K(x_0)$ is convex, and hence $C(x_0)$ is convex.

We have the following result (Bonnans and Shapiro [4, Theorem 4.133]).

Proposition 2.3. *Let $x_0 \in K$, $a_0 \in N_K(x_0)$, and suppose that the set K is prox-regular and second order regular at x_0. Then for sufficiently small neighborhood V and κ large* *enough, $\vartheta_\kappa(\cdot,\cdot)$ is second order Hadamard directionally differentiable at (x_0, a_0) and $\vartheta''_\kappa((x_0, a_0), (d, r))$ is equal to the optimal value of the problem*:

$$\operatorname*{Min}_{h\in C(x_0)} \left\{2\langle r, d-h\rangle + \kappa\|d-h\|^2 - \sigma(a_0, T_K^2(x_0, h))\right\}. \tag{2.13}$$

From now on we assume that the mapping $F(\cdot)$ is *directionally differentiable* at x_0. Then, since it is assumed that $F(\cdot)$ is locally Lipschitz continuous, $F(\cdot)$ is directionally differentiable at x_0 in the Hadamard sense, and $F'(x_0,\cdot)$ is Lipschitz continuous.

Proposition 2.4. *Let $x_0 \in K$, $a_0 \in N_K(x_0)$, and suppose that the set K is prox-regular and second order regular at x_0. Then for sufficiently small neighborhood V and κ large enough, $\gamma_\kappa(\cdot)$ is second order Hadamard directionally differentiable at x_0 and $\gamma''_\kappa(x_0, d) = \nu(d)$, where $\nu(d)$ denotes the optimal value of the problem*:

$$\operatorname*{Min}_{h\in C(x_0)} \left\{2\langle F'(x_0, d), d-h\rangle + \kappa\|d-h\|^2 - \sigma\left(F(x_0), T_K^2(x_0, h)\right)\right\}. \tag{2.14}$$

Proof. Because of Hadamard directional differentiability of $F(\cdot)$ at x_0 and since $D_a\vartheta(x_0, F(x_0)) = 0$ and $D\vartheta(\cdot,\cdot)$ is locally Lipschitz continuous, we have that for any sequences $t_k \downarrow 0$ and $d_k \to d$,

$$\vartheta(x_0 + t_k d_k, F(x_0 + t_k d_k)) = \vartheta\,(x_0 + t_k d_k, F(x_0) + t_k F'(x_0, d_k)) + o(t_k^2),$$

and $F'(x_0, d_k) \to F'(x_0, d)$. Consequently, the limit

$$\lim_{k\to\infty} \frac{\vartheta_\kappa(x_0 + t_k d_k, F(x_0 + t_k d_k)) - \vartheta_\kappa(x_0, F(x_0)) - t_k\langle F(x_0), d_k\rangle}{\frac{1}{2}t_k^2}$$

is equal to $\vartheta''_\kappa((x_0, F(x_0)), (d, F'(x_0, d)))$. Together with Proposition 2.3, this completes the proof. □

3. Local uniqueness of solutions. In this section we discuss conditions ensuring local uniqueness of solutions of the variational inequality $\mathrm{VI}(K, F)$. We consider the case where the set K is given in the form (1.4), i.e., $K = G^{-1}(Q)$. For the reference point $x_0 \in K$ we denote by $z_0 := G(x_0)$. Unless stated otherwise, we make the following assumptions throughout this section.

(A1) The mapping $G\colon \mathbb{R}^n \to \mathbb{R}^m$ is twice continuously differentiable.

(A2) The set $Q \subset \mathbb{R}^m$ is convex, closed, and second order regular at the point z_0.

(A3) Robinson's constraint qualification holds at the point x_0.

In the considered finite dimensional case, Robinson's constraint qualification can be written in the form

$$DG(x_0)\mathbb{R}^n + T_Q(z_0) = \mathbb{R}^m. \tag{3.1}$$

Under the above assumptions (A1)–(A3), the set K is second order regular (Bonnans and Shapiro [4, Proposition 3.88]) and, as it was mentioned in the previous section, is prox-regular at the point x_0. Moreover, the tangent cone to K at x_0 can be written as follows:

$$T_K(x_0) = \{h\colon DG(x_0)h \in T_Q(z_0)\}. \tag{3.2}$$

Therefore, $F(x_0) \in N_K(x_0)$ iff the optimal value of the problem

$$\operatorname*{Max}_{h\in\mathbb{R}^n} \langle F(x_0), h\rangle \quad \text{subject to } DG(x_0)h \in T_Q(z_0), \tag{3.3}$$

is zero. By calculating the dual of the above problem (cf., Bonnans and Shapiro [4, p. 151]), we obtain that $F(x_0) \in N_K(x_0)$ iff there exists $\lambda \in \mathbb{R}^m$ satisfying the following conditions:

$$F(x_0) - DG(x_0)^*\lambda = 0 \quad \text{and} \quad \lambda \in N_Q(G(x_0)). \tag{3.4}$$

We refer to the system (3.4) as *generalized equations*. Denote by $\Lambda(x_0)$ the set of all λ satisfying (3.4). By the above discussion we have that $x_0 \in \mathrm{Sol}(K, F)$ iff the set $\Lambda(x_0)$ is nonempty. Moreover, because of Robinson's constraint qualification, the set $\Lambda(x_0)$ is bounded and hence compact. Because of (3.2) and (3.4) we have that for any $\lambda \in \Lambda(x_0)$, the critical cone $C(x_0)$ can be written in the form

$$C(x_0) = \left\{h\colon \langle \lambda, DG(x_0)h\rangle = 0,\ DG(x_0)h \in T_Q(z_0)\right\}. \tag{3.5}$$

Let us discuss properties of the function

$$\phi(h) := \begin{cases} -\sigma(F(x_0), T_K^2(x_0, h)), & \text{if } h \in C(x_0),\\ +\infty, & \text{if } h \notin C(x_0). \end{cases} \tag{3.6}$$

This function appears in formula (2.14). We refer to $\phi(\cdot)$ as the *sigma term*. By (2.10) we have that

$$\phi(th) = t^2\phi(h) \quad \text{for any } h \in \mathbb{R}^n \text{ and } t > 0. \tag{3.7}$$

LEMMA 3.1. *Let $x_0 \in \mathrm{Sol}(K, F)$ and consider $\phi(\cdot)$ defined in* (3.6)*. Suppose that the assumptions* (A1)–(A3) *hold. Then $\phi(0) = 0$, and for any $h \in C(x_0)$, $\phi(h)$ is finite valued and*

$$\phi(h) = \sup_{\lambda \in \Lambda(x_0)} \{\langle \lambda, D^2G(x_0)(h, h)\rangle - \sigma(\lambda, T_Q^2(z_0, DG(x_0)h))\}. \tag{3.8}$$

PROOF. By the second order regularity of K, the set $T_K^2(x_0, h)$ is nonempty, and hence $\phi(h) < +\infty$, for any $h \in C(x_0)$. Since $T_K^2(x_0, 0) = T_K(x_0)$ and $F(x_0)$ belongs to the negative dual of $T_K(x_0)$, it follows that $\phi(0) = 0$. Now since Q is convex, we have (cf., Cominetti [5]) that for $z_0 \in Q$ and $v \in T_Q(z_0)$,

$$T_Q^2(z_0, v) + T_{T_Q(z_0)}(v) \subset T_Q^2(z_0, v) \subset T_{T_Q(z_0)}(v), \tag{3.9}$$

$$T_{T_Q(z_0)}(v) = \mathrm{cl}\{T_Q(z_0) + \mathrm{sp}(v)\}, \tag{3.10}$$

where $\mathrm{sp}(v)$ denotes the linear space generated by vector v.

By a chain rule (Cominetti [5], Bonnans and Shapiro [4, Proposition 3.33]), we have

$$T_K^2(x_0, h) = DG(x_0)^{-1}[T_Q^2(z_0, DG(x_0)h) - D^2G(x_0)(h, h)]. \tag{3.11}$$

It follows that the term $-\sigma(F(x_0), T_K^2(x_0, h))$ is equal to the optimal value of the following problem:

$$\begin{aligned} &\operatorname*{Min}_{w \in \mathbb{R}^n}\ \langle -F(x_0), w\rangle \\ &\text{subject to}\ \ DG(x_0)w + D^2G(x_0)(h, h) \in T_Q^2(z_0, DG(x_0)h). \end{aligned} \tag{3.12}$$

The dual of this problem is

$$\operatorname*{Max}_{\lambda \in \Lambda(x_0)} \{\langle \lambda, D^2G(x_0)(h, h)\rangle - \sigma(\lambda, T_Q^2(z_0, DG(x_0)h))\}, \tag{3.13}$$

and (under Robinson's constraint qualification) the optimal values of (3.12) and (3.13) are equal to each other (cf., Bonnans and Shapiro [4, p. 175]). We obtain that for any $h \in C(x_0)$, formula (3.8) holds.

Since for any $h \in C(x_0)$ and $\lambda \in \Lambda(x_0)$, we have that $\langle \lambda, DG(x_0)h\rangle = 0$ and $\langle \lambda, v\rangle \le 0$ for $v \in T_Q(z_0)$, it follows by (3.9) and (3.10) that

$$\sigma(\lambda, T_Q^2(z_0, DG(x_0)h)) \le 0, \quad \text{for all } h \in C(x_0) \text{ and } \lambda \in \Lambda(x_0). \tag{3.14}$$

Consequently,

$$\phi(h) \ge \sup_{\lambda \in \Lambda(x_0)} \langle \lambda, D^2G(x_0)(h, h)\rangle, \quad \forall\, h \in C(x_0). \tag{3.15}$$

Recall that $\Lambda(x_0) \ne \varnothing$, since $x_0 \in \mathrm{Sol}(K, F)$. It follows that $\phi(h)$ is finite valued for any $h \in C(x_0)$. □

Remark 3.1. Suppose that, in addition to the assumptions of Lemma 3.1, the following property holds at the point z_0:

$$\operatorname{dist}(z_0 + tv, Q) = o(t^2) \quad \text{for } t > 0 \text{ and every } v \in T_Q(z_0). \tag{3.16}$$

This holds, for example, if the set Q is polyhedral. Condition (3.16) means that $0 \in T_Q^2(z_0, v)$ for any $v \in T_Q(z_0)$. It follows then by (3.9) that $T_Q^2(z_0, v) = T_{T_Q(z_0)}(v)$, and hence because of (3.10) that

$$\sigma(\lambda, T_Q^2(z_0, DG(x_0)h)) = 0 \quad \text{for all } h \in C(x_0) \text{ and } \lambda \in \Lambda(x_0)$$

(cf., Bonnans and Shapiro [4, p. 177]). In that case, for any $h \in \mathbb{R}^n$,

$$\phi(h) = \sup_{\lambda \in \Lambda(x_0)} \langle \lambda, D^2 G(x_0)(h, h)\rangle + I_{C(x_0)}(h). \tag{3.17}$$

Condition (3.16) also implies that the set Q is second order regular at z_0 (Bonnans and Shapiro [4, p. 203]).

The first-order necessary condition for x_0 to be an optimal solution of (2.2) is that

$$\nabla \gamma_\kappa(x_0) \in N_K(x_0). \tag{3.18}$$

Since $\nabla \gamma_\kappa(x_0) = F(x_0)$ if $x_0 \in \operatorname{Sol}(K, F)$, condition (3.18) follows, of course, from the assumption $x_0 \in \operatorname{Sol}(K, F)$. Also the critical cone associated with problem (2.2) at the point x_0 is the same as the one defined in (2.12).

Definition 3.1. We say that the *quadratic growth* condition holds, for the problem (2.2) at x_0, if there exist a constant $c > 0$ and a neighborhood N of x_0 such that

$$-\gamma_\kappa(x) \geq -\gamma_\kappa(x_0) + c\|x - x_0\|^2, \quad \forall x \in K \cap N. \tag{3.19}$$

Of course, if $x_0 \in \operatorname{Sol}(K, F)$, then $\gamma_\kappa(x_0) = 0$ and condition (3.19) implies that x_0 is a locally unique solution of $\operatorname{VI}(K, F)$.

Since the set Q is convex, the function $\sigma(\lambda, T_Q^2(z_0, DG(x_0)\cdot))$ is concave (Bonnans and Shapiro [4, Proposition 3.48]). Therefore, it follows from (3.8) that the function $\phi(\cdot)$ is representable, on $C(x_0)$, as the maximum of the sum of quadratic and convex functions. Since the set $\Lambda(x_0)$ is compact, the corresponding quadratic functions can be represented in the form

$$\langle \lambda, D^2 G(x_0)(h, h)\rangle = f_\lambda(h) - \gamma\|h\|^2,$$

such that $f_\lambda(\cdot)$ is convex for all $\lambda \in \Lambda(x_0)$ and γ large enough. It follows that $\phi(\cdot)$ can be represented as the difference of an extended real valued convex function and the quadratic function $\gamma\|\cdot\|^2$. The subdifferential $\partial\phi(h)$ is then defined in a natural way as the subdifferential of the corresponding convex function minus $\nabla(\gamma\|h\|^2) = 2\gamma h$. In particular, for $\phi(\cdot)$ given in (3.17) and $h \in C(x_0)$, we have

$$\partial\phi(h) = \operatorname{conv}\left\{ \bigcup_{\lambda \in \Lambda^*(x_0, h)} [2D^2 G(x_0)h]^* \lambda \right\} + N_{C(x_0)}(h), \tag{3.20}$$

where $\Lambda^*(x_0, h) := \arg\max_{\lambda \in \Lambda(x_0)} \langle \lambda, D^2 G(x_0)(h, h)\rangle$.

Consider the following conditions:

(C1) There is $\kappa > 0$ such that to every $d \in C(x_0)\backslash\{0\}$ corresponds $h \in C(x_0)$ such that

$$2\langle F'(x_0, d), d - h\rangle + \kappa\|d - h\|^2 + \phi(h) < \phi(d). \tag{3.21}$$

(C2) The system

$$0 \in -2F'(x_0, d) + \partial\phi(d) \tag{3.22}$$

has only one solution $d = 0$.

Note that, because of (3.14) we have that $0 \in \partial\phi(0)$, and since $F'(x_0, 0) = 0$, it follows that $d = 0$ is always a solution of (3.22).

We can now formulate the main result of this section.

THEOREM 3.1. *Let $x_0 \in \mathrm{Sol}(K, F)$ and consider $\phi(\cdot)$ defined in (3.6). Suppose that the assumptions* (A1)–(A3) *hold. Then for κ large enough (and V sufficiently small), conditions* (C1) *and* (C2) *are equivalent to each other and are necessary and sufficient for the quadratic growth condition* (3.19) *to hold.*

PROOF. As we mentioned earlier, since $x_0 \in \mathrm{Sol}(K, F)$ we have here that the first order necessary condition (3.18) holds. Moreover, by Proposition 2.4 we have that (for V sufficiently small and κ large enough) the function $\gamma_\kappa(\cdot)$ is second order Hadamard directionally differentiable at x_0 and $\gamma''_\kappa(d) = \nu(d)$, where $\nu(d)$ is the optimal value of problem (2.14). It follows that

$$\gamma''_\kappa(x_0; d, w) = D\gamma_\kappa(x_0)w + \nu(d), \tag{3.23}$$

where

$$\gamma''_\kappa(x_0; d, w) := \lim_{t\downarrow 0} \frac{\gamma_\kappa(x_0 + td + \frac{1}{2}t^2 w) - \gamma_\kappa(x_0) - tD\gamma_\kappa(x_0)d}{\frac{1}{2}t^2} \tag{3.24}$$

denotes the parabolic second order directional derivative. We have then the following necessary and, because of the second order Hadamard directionally differentiability, sufficient condition for the quadratic growth property (3.19):

$$\inf_{w \in T^2_K(x_0, d)} (-\gamma_\kappa)''(x_0; d, w) > 0, \quad \forall\, d \in C(x_0)\setminus\{0\} \tag{3.25}$$

(see Bonnans and Shapiro [4, Proposition 3.105]). Since $D\gamma_\kappa(x_0)w = \langle F(x_0), w\rangle$ and (3.23), we have that

$$\inf_{w \in T^2_K(x_0, d)} (-\gamma_\kappa)''(x_0; d, w) = -\nu(d) - \sigma(F(x_0), T^2_K(x_0, d)).$$

It follows that condition (3.25) can be written in the following equivalent form:

$$\nu(d) + \sigma(F(x_0), T^2_K(x_0, d)) < 0, \quad \forall\, d \in C(x_0)\setminus\{0\}. \tag{3.26}$$

By employing formula (2.14) we obtain that condition (C1) is equivalent to (3.26), and hence is necessary and sufficient for the quadratic growth (3.19).

Now for a given $d \in C(x_0)$ consider the function

$$\psi_\kappa(h) := 2\langle F'(x_0, d), d - h\rangle + \kappa\|d - h\|^2 + \phi(h). \tag{3.27}$$

Clearly, for $h = d$ we have that $\psi_\kappa(d) = \phi(d)$. Since $\psi_\kappa(h) = +\infty$ for any $h \in \mathbb{R}^n\setminus C(x_0)$, we obtain that condition (C1) means that $\inf_{h\in\mathbb{R}^n} \psi_\kappa(h) < \phi(d)$. Equivalently this can be formulated as that d is not a minimizer of $\psi_\kappa(\cdot)$ over $\mathbb{R}^n$. Since $\phi(\cdot)$ can be represented as a difference of a convex and quadratic function, the function $\psi_\kappa(\cdot)$ is convex for κ large enough. Consequently, d is a minimizer of $\psi_\kappa(\cdot)$ iff $0 \in \partial\psi_\kappa(d)$. Moreover,

$$\partial\psi_\kappa(d) = -2F'(x_0, d) + \partial\phi(d),$$

and hence (3.22) is a necessary and sufficient condition for d to be a minimizer of $\psi_\kappa(\cdot)$ for κ large enough. This shows equivalence of conditions (C1) and (C2), and hence completes the proof. □

For any $d \notin C(x_0)$ we have that $\phi(d) = +\infty$, and hence $\partial\phi(d) = \varnothing$. Therefore, the inclusion (3.22) should be verified only for $d \in C(x_0)$. Consider the following condition:

(C3) For all $d \neq 0$ it holds that $\varphi(d) \neq 0$, where

$$\varphi(d) := -\langle d, F'(x_0, d)\rangle + \phi(d). \tag{3.28}$$

Of course, since $\varphi(d) = +\infty$ for any $d \notin C(x_0)$, the above condition should be verified only for $d \in C(x_0)\setminus\{0\}$.

Proposition 3.1. *Let $x_0 \in \mathrm{Sol}(K, F)$ and suppose that the assumptions* (A1)–(A3) *hold. Then condition* (C3) *implies condition* (C2)*. Suppose, further, that $F(\cdot)$ is differentiable at x_0, the Jacobian matrix $\nabla F(x_0)$ is symmetric, and $\varphi(d) \geq 0$ for all $d \in C(x_0)$. Then conditions* (C2) *and* (C3) *are equivalent.*

Proof. Recall that by Theorem 3.1 condition (C2) is equivalent to condition (C1) for κ large enough. Suppose that condition (C3) holds. Consider $d \in C(x_0)\setminus\{0\}$. In order to show the implication (C3) $\Rightarrow$ (C2) it will suffice to verify that d is not a minimizer of the function $\psi_\kappa(\cdot)$ defined in (3.27). By (3.7) we have that for $t > -1$,

$$\psi_\kappa((1+t)d) = -2t\langle d, F'(x_0, d)\rangle + t^2\kappa\|d\|^2 + (1+t)^2\phi(d) = \psi_\kappa(d) + q(t, d),$$

where

$$q(t, d) := 2\varphi(d)t + \left(\kappa\|d\|^2 + \phi(d)\right)t^2.$$

Since $\varphi(d) \neq 0$, it follows that $q(t, d) < 0$ for all negative or positive, depending on the sign of the number $\kappa\|d\|^2 + \phi(d)$, values of t sufficiently close to zero. This implies that d is not a minimizer of $\psi_\kappa(\cdot)$.

Suppose now that $\varphi(d) \geq 0$ for all $d \in C(x_0)$, $F(\cdot)$ is differentiable at x_0, and $\nabla F(x_0)$ is symmetric. Since $\varphi(0) = 0$ and $\varphi(d) = +\infty$ for $d \notin C(x_0)$, it follows that $d = 0$ is a minimizer of $\varphi(\cdot)$. We have

$$\partial\varphi(d) = -\nabla F(x_0)d + \partial\phi(d).$$

Consequently, if $d \in C(x_0)$ is a minimizer of $\varphi(\cdot)$, then (3.22) holds by the first order necessary conditions. Therefore, condition (C2) implies that $d = 0$ is the unique minimizer of $\varphi(\cdot)$, and hence that $\varphi(d) > 0$ for all $d \in C(x_0)\setminus\{0\}$. That is, (C2) implies (C3). Since we already showed that (C3) $\Rightarrow$ (C2), it follows that (C2) and (C3) are equivalent. □

The following is a consequence of Theorem 3.1 and Proposition 3.1.

Theorem 3.2. *Let $x_0 \in \mathrm{Sol}(K, F)$. Suppose that the assumptions* (A1)–(A3) *hold, and either condition* (C2) *or* (C3) *is satisfied. Then x_0 is a locally unique solution of* $\mathrm{VI}(K, F)$.

Remark 3.2. Suppose that $G(x) \equiv x$, i.e., $G(\cdot)$ is the identity mapping. Then we can identify the set K with the set Q, and the assumptions (A1) and (A3) hold automatically and the assumption (A2) means, of course, that the set K is convex, closed and second order regular at x_0. If, moreover, the set K is *polyhedral*, then $\sigma\left(F(x_0), T_K^2(x_0, h)\right) = 0$ for any $h \in C(x_0)$ and $\phi(\cdot)$ becomes the indicator function of the set $C(x_0)$ (see Remark 3.1). Consequently, in that case the system (3.22) takes the form:

$$0 \in -F'(x_0, d) + N_{C(x_0)}(d), \tag{3.29}$$

and $\varphi(d) = -\langle d, F'(x_0, d)\rangle$ for all $d \in C(x_0)$. Therefore, for polyhedral set K condition $\varphi(d) > 0$ for all $d \in C(x_0)\setminus\{0\}$ is equivalent to the condition:

$$\langle d, F'(x_0, d)\rangle < 0, \quad \forall\, d \in C(x_0)\setminus\{0\}. \tag{3.30}$$

It is shown in Facchinei and Pang [8, Proposition 3.3.4] that for convex (not necessarily polyhedral) set K condition (3.30) implies local uniqueness of solution x_0. Also if K is polyhedral and $F(\cdot)$ is affine, then the condition: "the system (3.29) has only one solution $d = 0$," is necessary and sufficient for local uniqueness of x_0 (Facchinei and Pang [8, Proposition 3.3.7]). Since here $D^2G(x_0) = 0$, and hence $\phi(d) \geq 0$ for all $d \in C(x_0)$ (see (3.15)), condition (3.30) is stronger than condition (C3).

REMARK 3.3. Consider the optimization problem (1.3) and the associated variational inequality representing first order optimality conditions for (1.3). Let x_0 be a locally optimal solution of (1.3), and hence $x_0 \in \mathrm{Sol}(K, F)$ with $F(\cdot) := -\nabla f(\cdot)$. Suppose that the function $f(\cdot)$ is twice continuously differentiable and assumptions (A1)–(A3) hold. Then for $d \in C(x_0)$ the function $\varphi(d)$ takes the form

$$\varphi(d) = D^2 f(x_0)(d, d) - \sigma(F(x_0), T_K^2(x_0, d)). \tag{3.31}$$

By second order necessary conditions we have that $\varphi(d) \geq 0$ for all $d \in C(x_0)$. Moreover, the quadratic growth condition for the optimization problem (1.3) holds at x_0 iff $\varphi(d) > 0$ for all $d \in C(x_0) \setminus \{0\}$. It follows, by Theorem 3.1 and Proposition 3.1, that the quadratic growth condition (3.19) holds, for κ large enough, iff the corresponding quadratic growth condition for the optimization problem (1.3) is satisfied.

4. Continuity and differentiability properties of solutions. Consider the parameterized variational inequality (1.1). In this section we discuss continuity and differentiability properties of the multifunction $\mathcal{S}(u) := \mathrm{Sol}(K, F_u)$, at the (reference) point $u_0 \in U$. It will be assumed in this section that the set K is independent of u, and $x_0 \in \mathrm{Sol}(K, F)$; i.e., $x_0 \in \mathcal{S}(u_0)$, and that the mapping $F(\cdot, \cdot)$ is locally Lipschitz continuous and directionally differentiable at (x_0, u_0). We describe continuity and differentiability properties of $\mathcal{S}(u)$ in terms of the corresponding contingent derivatives. Such approach to sensitivity analysis was initiated by Rockafellar [24].

DEFINITION 4.1. For a multifunction $\mathcal{M}\colon U \rightrightarrows \mathbb{R}^n$ and a point $x_0 \in \mathcal{M}(u_0)$, the *contingent derivative* $D\mathcal{M}(u_0|x_0)$ is defined as a multifunction, from U into $\mathbb{R}^n$, with $h \in D\mathcal{M}(u_0|x_0)(w)$ iff there exist sequences $w_k \to w$, $h_k \to h$ and $t_k \downarrow 0$ such that $x_0 + t_k h_k \in \mathcal{M}(u_0 + t_k w_k)$.

The term "contingent derivative" is motivated by the fact that the graph of the multifunction $D\mathcal{M}(u_0|x_0)(\cdot)$ coincides with the contingent (Bouligand) cone to the graph of $\mathcal{M}(\cdot)$ at (u_0, x_0). The contingent derivatives are called *(outer) graphical derivatives* in Levy [15] and Rockafellar and Wets [25, p. 324].

By the definition we have that $\mathcal{S}(u)$ is the solution of the variational condition $0 \in -F(x, u) + N_K(x)$. Therefore, we can employ the following results due to Levy and Rockafellar [17, Theorem 4.1] and Levy [15, Theorem 3.1 and Corollary 3.3].

THEOREM 4.1. *Let $\mathcal{N}\colon \mathbb{R}^n \times U \rightrightarrows \mathbb{R}^n$ be a multifunction and*

$$\mathcal{S}(u) := \left\{x \in \mathbb{R}^n\colon\ 0 \in -F(x, u) + \mathcal{N}(x, u)\right\}$$

be the associated solution multifunction, and let $y_0 \in \mathcal{N}(x_0, u_0)$ with $y_0 := F(x_0, u_0)$, i.e., $x_0 \in \mathcal{S}(u_0)$. Then the following inclusion holds for every $p \in U$,

$$D\mathcal{S}(u_0|x_0)(p) \subset \left\{d\colon 0 \in -F'((x_0, u_0), (d, p)) + D\mathcal{N}((x_0, u_0)|y_0)(d, p)\right\}. \tag{4.1}$$

Moreover, the left-hand side of (4.1) is equal to the right-hand side if the parameterization is rich enough, that is, if $F(x, u)$ can be written in the form $F_1(x, u_1) + u_2$, where $u = (u_1, u_2) \in U \times \mathbb{R}^n$.

In this section we employ the above result for the normal-cone multifunction $\mathcal{N}(x) := N_K(x)$. Since this multifunction does not depend on u, formula (4.1) then takes the form

$$D\mathcal{S}(u_0|x_0)(p) \subset \{d\colon 0 \in -F'((x_0, u_0), (d, p)) + DN_K(x_0|y_0)(d)\}. \tag{4.2}$$

In order to proceed we need to calculate the contingent derivative of the multifunction $N_K(x)$ at $x_0 \in K$ for $y_0 \in N_K(x_0)$.

LEMMA 4.1. *Suppose that the set K is prox-regular at $x_0 \in K$. Then for any $y_0 \in N_K(x_0)$ sufficiently close to* 0, *the following inclusion holds*:

$$DN_K(x_0|y_0)(d) \subset \{h\colon P'_K(x_0+y_0, d+h) = d\}, \quad \forall\, d \in \mathbb{R}^n, \tag{4.3}$$

provided that P_K is directionally differentiable at $x_0 + y_0$.

PROOF. As it was mentioned earlier, it follows from prox-regularity of K at x_0 that the projection $P_K(\cdot)$ is uniquely defined and Lipschitz continuous in a neighborhood of x_0, and $P_K(x+y) = x$ iff $x \in K$ and $y \in N_K(x)$ for x in a neighborhood of x_0 and y sufficiently close to 0. We have that $h \in DN_K(x_0|y_0)(d)$ iff there exist sequences $d_k \to d$, $h_k \to h$ and $t_k \downarrow 0$ such that $y_0 + t_k h_k \in N_K(x_0 + t_k d_k)$, or equivalently

$$P_K(x_0 + y_0 + t_k(d_k + h_k)) = x_0 + t_k d_k. \tag{4.4}$$

Note that it follows from (4.4) that $x_0 + t_k d_k \in K$.

Since $P_K(\cdot)$ is Lipschitz continuous in a neighborhood of x_0 and directionally differentiable at $x_0 + y_0$, we have that

$$P_K(x_0 + y_0 + t_k(d_k + h_k)) = x_0 + t_k P'_K(x_0 + y_0, d + h) + o(t_k). \tag{4.5}$$

Consequently, it follows from (4.4) that

$$P'_K(x_0 + y_0, d + h) = d, \tag{4.6}$$

which completes the proof. □

We assume in the remainder of this section that the set K is defined in the form (1.4), i.e., $K := G^{-1}(Q)$. Suppose that the assumptions (A1)–(A3) of §3 hold. Then $P_K(\cdot)$ is directionally differentiable at the point $x_0 + y_0$, and $P'_K(x_0 + y_0, w)$ is equal to the optimal solution of the problem

$$\operatorname*{Min}_{\eta \in C(x_0)} \{\|w - \eta\|^2 - \sigma(y_0, T_K^2(x_0, \eta))\}, \tag{4.7}$$

where $C(x_0) := \{h\colon \langle y_0, h\rangle = 0,\ h \in T_K(x_0)\}$. This follows from general results of sensitivity analysis of parameterized optimization problems (see Bonnans and Shapiro [4, §4.7.3]). For convex sets K this formula was given in Bonnans et al. [2]. Note that, under assumptions (A1)–(A3), the critical cone $C(x_0)$ is convex and problem (4.7) has unique optimal solution for all y_0 in a neighborhood of 0.

PROPOSITION 4.1. *Suppose that assumptions* (A1)–(A3) *hold and let $y_0 \in N_K(x_0)$. Then the following inclusion holds for every $d \in \mathbb{R}^n$*,

$$DN_K(x_0|y_0)(d) \subset \tfrac{1}{2}\partial\phi(d), \tag{4.8}$$

where $\phi(\cdot)$ is defined in (3.6).

PROOF. As we mentioned earlier, assumptions (A1)–(A3) imply that K is prox-regular at x_0. Moreover, by rescaling $y_0 \mapsto t y_0$, if necessary with $t > 0$ small enough, we can assume that y_0 belongs to a sufficiently small neighborhood of 0. Thus P_K is directionally differentiable at the point $x_0 + y_0$ and $P'_K(x_0 + y_0, w)$ is equal to the optimal solution of the problem (4.7), and the inclusion (4.3) holds. That is, if $h \in DN_K(y_0|x_0)(d)$, then $\bar\eta = d$ is the optimal solution of (4.7) for $w := d + h$. By first order necessary conditions, this, in turn, implies that $0 \in -2h + \partial\phi(d)$. This completes the proof. □

Note that if the set K is convex, then assumptions (A1)–(A3) in the above proposition can be replaced by the assumption that K is second order regular at x_0. If the set K is convex *polyhedral*, then P_K is directionally differentiable and

$$P_K(x_0 + y_0 + tw) = x_0 + tP'_K(x_0 + y_0, w)$$

for all $t > 0$ sufficiently small. Therefore, in that case $DN_K(x_0|y_0)(d)$ is equal to the right-hand sides of (4.3) and (4.8), and hence

$$DN_K(x_0|y_0)(d) = N_{C(x_0)}(d). \tag{4.9}$$

For convex polyhedral set K, the contingent derivative $DN_K(x_0|y_0)(\cdot)$ was calculated in Levy and Rockafellar [18].

Theorem 4.1 and Proposition 4.1 imply the following theorem which is the main result of this section.

Theorem 4.2. *Suppose that assumptions* (A1)–(A3) *hold. Then the following inclusion holds for every* $p \in U$:

$$D\mathcal{S}(u_0|x_0)(p) \subset \Psi(p), \tag{4.10}$$

where $\Psi(p)$ *is the set of all vectors* $d \in \mathbb{R}^n$ *satisfying the following condition*:

$$0 \in -2F'((x_0, u_0), (d, p)) + \partial\phi(d). \tag{4.11}$$

Let us discuss now some implications of the above result. Recall that a multifunction $\mathcal{M}$: $U \rightrightarrows \mathbb{R}^n$ is said to be *locally upper Lipschitz* at $u_0 \in U$ for $x_0 \in \mathcal{M}(u_0)$ if there exist positive number ρ and neighborhoods V and W of x_0 and u_0, respectively, such that

$$\mathcal{M}(u) \cap V \subset B(x_0, \rho\|u - u_0\|), \qquad \forall\, u \in W. \tag{4.12}$$

By taking $u = u_0$ in (4.12) we obtain that $\mathcal{M}(u_0) \cap V = \{x_0\}$; i.e., $\mathcal{M}(u_0)$ restricted to a neighborhood of x_0 is single valued.

Proposition 4.2. *Multifunction* $\mathcal{M}$: $U \rightrightarrows \mathbb{R}^n$ *is locally upper Lipschitz at* u_0 *for* $x_0 \in \mathcal{M}(u_0)$ *if and only if* $D\mathcal{M}(u_0|x_0)(0) = \{0\}$.

Sufficiency of the above condition for the locally upper Lipschitz continuity of $\mathcal{M}$ is shown in King and Rockafellar [10, Proposition 2.1] and necessity in Levy [15, Proposition 4.1].

Because of the inclusion (4.10), we have that if $\Psi(0) = \{0\}$, then $D\mathcal{S}(u_0|x_0)(0) = \{0\}$. Clearly, for $p = 0$, system (4.11) coincides with system (3.22), and condition $\Psi(0) = \{0\}$, is the same as condition (C2). Therefore, we obtain the following result.

Theorem 4.3. *Let* $x_0 \in \mathrm{Sol}(K, F)$. *Suppose that assumptions* (A1)–(A3) *hold and either condition* (C2) *or* (C3) *is satisfied. Then the solution multifunction* $\mathcal{S}(u)$ *is locally upper Lipschitz at* u_0 *for* x_0.

It can be noted that the result of Theorem 3.2 (about local uniqueness of x_0) follows from Theorem 4.3. Also, as it was shown in Theorem 3.1, condition $\Psi(0) = \{0\}$ is necessary and sufficient for the quadratic growth condition (3.19).

It was shown in Remark 3.3 that, under assumptions (A1)–(A3), condition (C2) is equivalent to the corresponding quadratic growth condition for the optimization problem (1.3). It is said that a parameterization $f(x, u)$ of the objective function of problem (1.3) includes the tilt perturbation if $f(x, u) = f_1(x, u_1) + \langle u_2, x\rangle$, $u = (u_1, u_2) \in U_1 \times \mathbb{R}^n$ (see, e.g., Levy [16, p. 7]). In that case the gradient mapping $F(x, u) := \nabla_x f(x, u)$ can be written in the form $F_1(x, u_1) + u_2$, where $F_1(x, u) := \nabla_x f_1(x, u)$. It is possible to show that for a parameterization of problem (1.3) which includes the tilt perturbation, the corresponding quadratic growth condition is necessary for the locally upper Lipschitz continuity of $\mathrm{Sol}(K, F_u)$ (see the proof of Bonnans and Shapiro [4, Theorem 5.3] and the second part of Theorem 4.1). Therefore, at least for variational inequalities associated with optimization problems, the locally upper Lipschitz continuity of $\mathrm{Sol}(K, F_u)$ implies condition (C2) for a sufficiently rich parameterization.

The following result about directional differentiability of a solution mapping $\bar{x}(u) \in \mathcal{S}(u)$ is a consequence of Theorem 4.2 (cf., Levy [15, Theorem 4.6]).

COROLLARY 4.1. *For a vector $p \in U$ and a path $u(t) := u_0 + tp + o(t)$, let $\bar{x}(t) \in \mathscr{S}(u(t))$ be such that $\|\bar{x}(t) - x_0\| = O(t)$ for all $t > 0$ sufficiently small. Suppose that assumptions* (A1)–(A3) *hold and system* (4.11) *has unique solution $\bar{d} = \bar{d}(p)$. Then $(\bar{x}(t) - x_0)/t$ converges to $\bar{d}$ as $t \downarrow 0$.*

For polyhedral set K, system (4.11) takes the form (compare with (3.29)):

$$0 \in -F'((x_0, u_0), (d, p)) + N_{C(x_0)}(d). \tag{4.13}$$

For polyhedral set K and differentiable $F(\cdot,\cdot)$, the linearized system (4.13) was considered and directional differentiability of the solution mapping was established, under an assumption of strong regularity, in Robinson [21], and the corresponding locally upper Lipschitz behavior of the solution multifunction was derived in Robinson [20, Theorem 4.1]. Also for polyhedral set K and locally Lipschitz $F(\cdot,\cdot)$ the result of Theorem 4.2 follows from results presented in Klatte [11] and Klatte and Kummer [13], and an extension of the system (3.29) and upper Lipschitz continuity of $\mathscr{S}(u)$ was derived in Klatte [11, Theorem 4] and Klatte and Kummer [13, Theorem 8.30] in a framework of generalized Kojima functions. As was mentioned earlier, for a nonpolyhedral set K, condition (3.30) is stronger than condition (C3). Condition (3.30) was used in Facchinei and Pang [8, Proposition 5.1.6 and Corollary 5.1.8].

DEFINITION 4.2. It is said that the set K is *cone reducible* at $x_0 \in K$, if there exist a neighborhood V of x_0, a convex closed pointed cone $\mathscr{C} \subset \mathbb{R}^m$, and a twice continuously differentiable mapping Ξ: $V \to \mathbb{R}^m$ such that $\Xi(x_0) = 0 \in \mathbb{R}^m$, the derivative mapping $D\Xi(x_0)$: $\mathbb{R}^n \to \mathbb{R}^m$ is onto, and $K \cap V = \{x \in V\colon \Xi(x) \in \mathscr{C}\}$.

The above concept of cone reducibility is discussed in detail in Bonnans and Shapiro [4, §§3.4.4 and 4.6.1]. If the set K is cone reducible at x_0, then the function $\phi(d)$ can be represented as the restriction of a quadratic function $q(d) = \langle d, Ad\rangle$, $A \in \mathscr{S}^n$, to the set $C(x_0)$ (cf., Bonnans and Shapiro [4, p. 242]). In that case the system (4.11) takes the form:

$$0 \in -F'((x_0, u_0), (d, p)) + Ad + N_{C(x_0)}(d). \tag{4.14}$$

In some cases (e.g., for the cones of positive semidefinite matrices as discussed in Shapiro [27]) this quadratic function, and hence its gradient $\nabla q(d) = 2Ad$, can be calculated in a closed form. For cone reducible set K and differentiable $F(\cdot,\cdot)$, the system (4.14) and the result of Theorems 4.2 and 4.3 were derived in Shapiro [29] by a different method.

5. Parameterized generalized equations. In this section we discuss generalized equations of the form

$$F(x, u) - D_x G(x, u)^* \lambda = 0 \quad \text{and} \quad \lambda \in N_Q(G(x, u)). \tag{5.1}$$

Here the mapping $G_u(\cdot) = G(\cdot, u)$ and the set $K(u) := G_u^{-1}(Q)$ depend on the parameter vector $u \in U$. As before, we denote by $\mathscr{S}(u)$ the set of solutions of (5.1), i.e., $x \in \mathscr{S}(u)$ iff there exists $\lambda \in \mathbb{R}^m$ satisfying (5.1). We denote by $\Lambda(x, u)$ the set of λ satisfying generalized equations (5.1). We assume that $x_0 \in \mathscr{S}(u_0)$, i.e., that the set $\Lambda(x_0, u_0)$ is nonempty. We also make the following assumptions.

(B1) The mapping G: $\mathbb{R}^n \times U \to \mathbb{R}^m$ is twice continuously differentiable.

(B2) The set $Q \subset \mathbb{R}^m$ is convex and closed.

(B3) Robinson's constraint qualification, with respect to the mapping $G(\cdot) = G(\cdot, u_0)$, holds at x_0.

Consider function $\gamma_\kappa(x, u)$ defined in the same way as in (2.1) with K replaced by $K(u)$ and $F(x)$ replaced by $F(x, u)$. Since, by assumption (B1), $DG(x, u)$ is locally Lipschitz continuous, it follows by the Robinson-Ursescu stability theorem that property (2.3), in the definition of prox-regularity, holds for all u in a neighborhood of u_0 and $K = K(u)$, with the constant α and the corresponding neighborhood independent of u. Therefore, we have by Proposition 2.2 that there exist constant $\kappa > 0$ and neighborhoods V and W of x_0 and u_0, respectively, such that for any $u \in W$ and $\bar{x} \in V$, it holds that $\bar{x} \in \mathcal{S}(u)$ iff $\bar{x} \in K(u)$ and $\gamma_\kappa(\bar{x}, u) = 0$.

The quadratic growth condition, for $x = x_0$ and $u = u_0$, is defined here in the same way as in Definition 3.1 for the function $\gamma_\kappa(\cdot) = \gamma_\kappa(\cdot, u_0)$ and the set $K = K(u_0)$. Similar to Theorem 3.1 we have here that, under assumptions (B1)–(B3) and second order regularity of Q at $z_0 := G(x_0, u_0)$, for κ large enough the quadratic growth condition (3.19) holds iff condition (C2) is satisfied.

We say that a multifunction $\mathcal{M}\colon U \rightrightarrows \mathbb{R}^n$ is *locally upper Hölder*, of degree $\frac{1}{2}$, at u_0 for $x_0 \in \mathcal{M}(u_0)$ if there exist positive number ρ and neighborhoods V and W of x_0 and u_0, respectively, such that

$$\mathcal{M}(u) \cap V \subset B(x_0, \rho \|u - u_0\|^{1/2}), \qquad \forall\, u \in W. \tag{5.2}$$

THEOREM 5.1. *Suppose that assumptions* (B1)–(B3) *hold, $F(\cdot,\cdot)$ is continuously differentiable, and for sufficiently large κ the quadratic growth condition* (3.19) *is satisfied. Then the solution multifunction $\mathcal{S}(\cdot)$ is locally upper Hölder, of degree $\frac{1}{2}$, at u_0 for x_0.*

PROOF. Suppose that the quadratic growth condition (3.19) holds at x_0, with the corresponding constant $c > 0$ and neighborhood N. Consider a solution $\hat{x}(u) \in \mathcal{S}(u) \cap N$. We can choose the constant κ large enough and the neighborhoods V and W, such that $\hat{x}(u)$ is a maximizer of $\gamma_\kappa(\cdot, u)$ over $K(u) \cap V$ for all $u \in W$. By Bonnans and Shapiro [4, Proposition 4.37] we have then that the following estimate holds:

$$\|\hat{x}(u) - x_0\| \le c^{-1}\ell + 2\delta + c^{1/2}(\eta_1\delta + \eta_2\delta)^{1/2}. \tag{5.3}$$

Here $\ell = \ell(u)$ is a Lipschitz constant of the function

$$\chi(\cdot, u) := \gamma_\kappa(\cdot, u) - \gamma_\kappa(\cdot, u_0)$$

on a subset $\widehat{N}_u$ of N containing x_0 and $\hat{x}(u)$, η_1 and $\eta_2 = \eta_2(u)$ are Lipschitz constants of $\gamma_\kappa(\cdot, u_0)$ and $\gamma_\kappa(\cdot, u)$, respectively, on N, and $\delta = \delta(u)$ is the Hausdorff distance between $K(u_0) \cap N$ and $K(u) \cap N$. By the Robinson-Ursescu stability theorem we have that for the neighborhood N sufficiently small, $\delta(u) = O(\|u - u_0\|)$. The Lipschitz constant $\eta_2(u)$ is bounded for all u in a neighborhood of u_0. Let us estimate $\ell(u)$. We have that $\ell(u) \le \sup_{x \in \widehat{N}_u} \|D_x\chi(x, u)\|$. Moreover, since $\gamma_\kappa(x, u) = \vartheta_\kappa(x, F(x, u))$, we have by (2.7) that

$$\begin{aligned} D_x\chi(x, u) = {} & (D_xF(x, u)^* - D_xF(x, u_0)^*)(x - \bar{y}_\kappa(x, u_0)) \\ & + D_xF(x, u)^*(\bar{y}_\kappa(x, u_0) - \bar{y}_\kappa(x, u)) \\ & + [F(x, u) - F(x)] + \kappa[\bar{y}_\kappa(x, u_0) - \bar{y}_\kappa(x, u)]. \end{aligned} \tag{5.4}$$

We also have that $\bar{y}_\kappa(x_0, u_0) = x_0$ and (see Bonnans and Shapiro [4, pp. 434–435]) that the mapping $(x, u) \to P_{K(u)}(x)$, and hence the mapping $(x, u) \to P_{K(u)}(x + \kappa^{-1}F(x, u))$, are locally upper Hölder, of degree $\frac{1}{2}$, at (x_0, u_0). Therefore, by choosing $\widehat{N}_u := \{x_0, \hat{x}(u)\}$, we can bound the norm of the first term in the right-hand side of (5.4), on that subset, by $\beta\|\hat{x}(u) - x_0\|^{1/2}$ for some $\beta > 0$. Moreover, by continuity of $DF(x, u)$, the constant β can be arbitrary small for all u in a sufficiently small neighborhood of u_0. In particular, we can choose a neighborhood of u_0 such that $\beta \le c/2$. The other three terms in the right-hand side of (5.4) are of order $O(\|u - u_0\|^{1/2})$ uniformly in $x \in N$, for a sufficiently small neighborhood N. The proof can be completed then by applying the estimate (5.3). $\square$

Without additional assumptions, the power constant $\frac{1}{2}$ in the above locally upper Hölder continuity of $\mathcal{S}(\cdot)$ cannot be improved. For optimization problems this is discussed in

Bonnans and Shapiro [4, §4.5.1]. In order to apply the necessary and sufficient condition of Proposition 4.2, for the upper Lipschitz continuity of the solution mapping, together with the estimate (4.1) of Theorem 4.1, we now need to calculate the contingent derivative of the multifunction

$$\mathcal{N}(x, u) := D_x G(x, u)^*[N_Q(G(x, u))]. \tag{5.5}$$

Note that $y \in \mathcal{N}(x, u)$ iff there exists $\lambda \in N_Q(G(x, u))$ such that $y = D_x G(x, u)^* \lambda$. Therefore, generalized equations (5.1) can be written in the form

$$0 \in -F(x, u) + \mathcal{N}(x, u). \tag{5.6}$$

We say that the *strict constraint qualification* holds at $\lambda_0 \in \Lambda(x_0, u_0)$ if

$$DG(x_0)\mathbb{R}^n + T_{Q_0}(z_0) = \mathbb{R}^m, \tag{5.7}$$

where $z_0 := G(x_0)$ and

$$Q_0 := \{y \in Q: \langle \lambda_0, y - z_0 \rangle = 0\}.$$

The above condition (5.7) is just Robinson's constraint qualification with respect to the reduced set Q_0 (cf., Bonnans and Shapiro [4, Definition 4.46]). Of course, since Q_0 is a subset of Q, condition (5.7) implies Robinson's constraint qualification (3.1). It is known that if the strict constraint qualification holds, then $\Lambda(x_0, u_0) = \{\lambda_0\}$ is a singleton, and conversely if $\Lambda(x_0, u_0) = \{\lambda_0\}$ is a singleton and the set Q is polyhedral, then the strict constraint qualification follows (Bonnans and Shapiro [4, Proposition 4.47]). Note that if $DG(x_0)$ is *onto*, i.e., $DG(x_0)\mathbb{R}^n = \mathbb{R}^m$, then of course the strict constraint qualification follows.

LEMMA 5.1. *Let $y_0 := F(x_0, u_0) \in \mathcal{N}(x_0, u_0)$. Suppose that assumptions* (B1), (B2), *and the strict constraint qualification hold. Then*

$$D\mathcal{N}((x_0, u_0)|y_0)(d, p) \subset D_x G(x_0, u_0)^*[DN_Q(z_0|\lambda_0)(DG(x_0, u_0)(d, p))] \\ + [D^2_{xx} G(x_0, u_0)d + D^2_{xu} G(x_0, u_0)p]^* \lambda_0. \tag{5.8}$$

Moreover, if $DG(x_0)$ is onto, then the left- and right-hand sides of (5.8) *are equal to each other.*

PROOF. Let h be an element of $D\mathcal{N}((x_0, u_0)|y_0)(d, p)$. This means that there exist sequences $d_k \to d$, $p_k \to p$, $h_k \to h$, and $t_k \downarrow 0$ such that $y_k \in \mathcal{N}(x_k, u_k)$, where $y_k := y_0 + t_k h_k$, $x_k := x_0 + t_k d_k$, and $u_k := u_0 + t_k p_k$. The condition $y_k \in \mathcal{N}(x_k, u_k)$, in turn, means that $y_k = D_x G(x_k, u_k)^* \lambda_k$ for some sequence $\lambda_k \in N_Q(G(x_k, u_k))$. Now the strict constraint qualification implies that

$$\|\lambda_k - \lambda_0\| = O(\|G(x_k, u_k) - G(x_0, u_0)\| + \|D_x G(x_k, u_k) - D_x G(x_0, u_0)\|) \tag{5.9}$$

(see Bonnans and Shapiro [4, Proposition 4.47(ii)]). It follows that $\|\lambda_k - \lambda_0\| = O(t_k)$. Consequently, by passing to a subsequence if necessary, we can assume that $\mu_k := (\lambda_k - \lambda_0)/t_k$ converges to a vector μ. It follows that

$$\lambda_0 + t_k \mu_k \in N_Q(z_0 + t_k DG(x_0, u_0)(d, p) + o(t_k)), \tag{5.10}$$

and hence

$$\mu \in DN_Q(z_0|\lambda_0)(DG(x_0, u_0)(d, p)). \tag{5.11}$$

Moreover, since $y_k = D_x G(x_k, u_k)^* \lambda_k$ it follows that

$$y_k = y_0 + t_k\big(D_x G(x_0, u_0)^* \mu + [D^2_{xx} G(x_0, u_0)d + D^2_{xu} G(x_0, u_0)p]^* \lambda_0\big) + o(t_k). \tag{5.12}$$

It follows from (5.11) and (5.12) that h belongs to the right-hand side of (5.8), and hence the inclusion (5.8) follows.

Conversely, suppose that $DG(x_0)$ is onto. Let μ be an element of the set in the right-hand side of (5.11). This means that there exist sequences $\mu_k \to \mu$ and $w_k \to DG(x_0, u_0)(d, p)$ such that

$$\lambda_k \in N_Q(z_0 + t_k w_k), \tag{5.13}$$

where $\lambda_k := \lambda_0 + t_k \mu_k$. Moreover, since $DG(x_0)$ is onto, there exist sequences $d_k \to d$ and $p_k \to p$ such that $z_0 + t_k w_k = G(x_k, u_k)$, where $x_k := x_0 + t_k d_k$ and $u_k := u_0 + t_k p_k$. Define $y_k := D_x G(x_k, u_k)^* \lambda_k$. Then $y_k \in \mathcal{N}(x_k, u_k)$ and the equation (5.12) holds. It follows that

$$D_x G(x_0, u_0)^* \mu + [D^2_{xx} G(x_0, u_0) d + D^2_{xu} G(x_0, u_0) p]^* \lambda_0 \in D\mathcal{N}((x_0, u_0) | y_0)(d, p).$$

This shows that the opposite of the inclusion (5.8) also holds, and hence completes the proof. □

Formula (5.8), combined with (4.8), yields the following result.

Proposition 5.1. *Let $y_0 := F(x_0, u_0) \in \mathcal{N}(x_0, u_0)$. Suppose that assumptions* (B1), (B2), *and the strict constraint qualification hold, and the set Q is second order regular at the point $z_0 := G(x_0, u_0)$. Then*

$$\begin{aligned} D\mathcal{N}((x_0, u_0) | y_0)(d, p) \subset {} & \tfrac{1}{2} D_x G(x_0, u_0)^* [\partial \xi (DG(x_0, u_0)(d, p))] \\ & + [D^2_{xx} G(x_0, u_0) d + D^2_{xu} G(x_0, u_0) p]^* \lambda_0, \end{aligned} \tag{5.14}$$

where

$$\xi(v) := \begin{cases} -\sigma(\lambda_0, T^2_Q(z_0, v)), & \text{if } v \in T_Q(z_0) \text{ and } \langle \lambda_0, v \rangle = 0, \\ +\infty, & \text{otherwise.} \end{cases} \tag{5.15}$$

By Theorem 4.1 we obtain that, under the assumptions of the above proposition and Lipschitz continuity and directional differentiability of $F(x, u)$, the contingent derivative $D\mathcal{S}(u_0 | x_0)(p)$ is included in the set of vectors $d \in \mathbb{R}^n$ satisfying the following condition:

$$\begin{aligned} 0 \in {} & -2F'((x_0, u_0), (d, p)) + D_x G(x_0, u_0)^* [\partial \xi (DG(x_0, u_0)(d, p))] \\ & + 2[D^2_{xx} G(x_0, u_0) d + D^2_{xu} G(x_0, u_0) p]^* \lambda_0. \end{aligned} \tag{5.16}$$

Consequently, by Proposition 4.2 we obtain the following result.

Theorem 5.2. *Let $y_0 := F(x_0, u_0) \in \mathcal{N}(x_0, u_0)$. Suppose that assumptions* (B1), (B2), *and the strict constraint qualification hold, the set Q is second order regular at the point z_0, and $F(x, u)$ is Lipschitz continuous and directionally differentiable. Then $\mathcal{S}(u)$ is locally upper Lipschitz at u_0 for x_0 if the system*

$$0 \in -2F'(x_0, d) + DG(x_0)^* [\partial \xi (DG(x_0) d)] + 2[D^2 G(x_0) d]^* \lambda_0 \tag{5.17}$$

has only one solution $d = 0$.

Of course, for $K = Q$ and identity mapping $G(x, u) \equiv x$, system (5.16) coincides with system (4.11), and Theorem 5.2 reduces to Theorem 4.3.

Acknowledgments. The author is indebted to Diethard Klatte, Jong-Shi Pang, and Adam Levy for constructive comments and valuable discussions that helped to improve the manuscript.

References

[1] Auchmuty, G. 1989. Variational principles for variational inequalities. *Numer. Funct. Anal. Optim.* **10** 863–874.

[2] Bonnans, J. F., R. Cominetti, A. Shapiro. 1998. Sensitivity analysis of optimization problems under second order regular constraints. *Math. Oper. Res.* **23** 806–831.

[3] Bonnans, J. F., R. Cominetti, A. Shapiro. 1999. Second order optimality conditions based on parabolic second order tangent sets. *SIAM J. Optim.* **9** 466–492.

[4] Bonnans, J. F., A. Shapiro. 2000. *Perturbation Analysis of Optimization Problems*. Springer, New York.

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

[6] Danskin, J. M. 1967. *The Theory of Max-Min and Its Applications to Weapon Allocation Problems. Econometrics and Operations Research*, Vol 5. Springer-Verlag, Berlin, Germany.

[7] Dontchev, A. L., R. T. Rockafellar. 1996. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM J. Optim.* **6** 1087–1105.

[8] Facchinei, F., J.-S. Pang. 2003. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer, New York.

[9] Fukushima, M. 1992. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Math. Programming* **53** 99–110.

[10] King, A., R. T. Rockafellar. 1992. Sensitivity analysis for nonsmooth generalized equations. *Math. Programming* **55** 193–212.

[11] Klatte, D. 2000. Upper Lipschitz behavior of solutions to perturbed $C^{1,1}$ programs. *Math. Programming* **88** 285–311.

[12] Klatte, D., B. Kummer. 1999. Generalized Kojima-functions and Lipschitz stability of critical points. *Comput. Optim. Appl.* **13** 61–85.

[13] Klatte, D., B. Kummer. 2002. *Nonsmooth Equations in Optimization: Regularity Calculus, Methods and Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

[14] Kinderlehrer, D., G. Stampacchia. 1980. *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York.

[15] Levy, A. B. 1999. Implicit function theorems for sensitivity analysis of variational conditions. *Math. Programming* **74** 333–350. (Errata: *Math. Programming* **86** 439–441.)

[16] Levy, A. B. 2001. Solution sensitivity from general principles. *SIAM J. Control Optim.* **40** 1–38.

[17] Levy, A. B., R. T. Rockafellar. 1994. Sensitivity analysis for solutions to generalized equations. *Trans. Amer. Math. Soc.* **345** 661–671.

[18] Levy, A. B., R. T. Rockafellar. 1995. Sensitivity of solutions in nonlinear programming problems with nonunique multipliers. D. Z. Du, L. Qi, R. S. Womersley, eds. *Recent Advances in Nonsmooth Optimization*. Word Scientific, Singapore, 215–223.

[19] Poliquin, R. A., R. T. Rockafellar. 1996. Prox-regular functions in variational analysis. *Trans. Amer. Math. Soc.* **348** 1805–1838.

[20] Robinson, S. M. 1982. Generalized equations and their solutions, Part II: Applications to nonlinear programming. *Math. Programming Stud.* **19** 200–221.

[21] Robinson, S. M. 1985. Implicit B-differentiability in generalized equations. Technical report 2854, Mathematics Research Center, University of Wisconsin.

[22] Robinson, S. M. 1992. Normal maps induced by linear transformations. *Math. Oper. Res.* **17** 691–714.

[23] Robinson, S. M. 2003. Constraint nondegeneracy in variational analysis. *Math. Oper. Res.* **28** 201–232.

[24] Rockafellar, R. T. 1990. Nonsmooth analysis and parametric optimization. A. Cellina, ed. *Methods on Nonconvex Analysis, Lecture Notes in Mathematics*, Vol. 1446. Springer, Berlin, 137–151.

[25] Rockafellar, R. T., R. J.-B. Wets. 1998. *Variational Analysis*. Springer-Verlag, New York.

[26] Shapiro, A. 1994. Existence and differentiability of metric projections in Hilbert spaces. *SIAM J. Optim.* **4** 130–141.

[27] Shapiro, A. 1997. First and second order analysis of nonlinear semidefinite programs. *Math. Programming Ser. B* **77** 301–320.

[28] Shapiro, A. 2000. On the asymptotics of constrained local M-estimators. *Ann. Statist.* **28** 948–960.

[29] Shapiro, A. 2003. Sensitivity analysis of generalized equations. *J. Math. Sci.* **115** 2554–2565.