Tim Hoheisel; Christian Kanzow

First- and second-order optimality conditions for mathematical programs with vanishing constraints

Applications of Mathematics, Vol. 52 (2007), No. 6, 495–514

Persistent URL: `http://dml.cz/dmlcz/134692`

FIRST- AND SECOND-ORDER OPTIMALITY CONDITIONS FOR MATHEMATICAL PROGRAMS WITH VANISHING CONSTRAINTS*

Tim Hoheisel, Christian Kanzow, Würzburg

Dedicated to Jiří V. Outrata on the occasion of his 60th birthday.

Abstract. We consider a special class of optimization problems that we call *Mathematical Programs with Vanishing Constraints*, MPVC for short, which serves as a unified framework for several applications in structural and topology optimization. Since an MPVC most often violates stronger standard constraint qualification, first-order necessary optimality conditions, weaker than the standard KKT-conditions, were recently investigated in depth. This paper enlarges the set of optimality criteria by stating first-order sufficient and second-order necessary and sufficient optimality conditions for MPVCs.

Keywords: mathematical programs with vanishing constraints, mathematical programs with equilibrium constraints, first-order optimality conditions, second-order optimality conditions

MSC 2000: 90C30, 90C33

1. Introduction

We consider a constrained optimization problem of the form

$$
\begin{aligned}
(1) \qquad \min \quad & f(x) \\
\text{s.t.} \quad & g_i(x) \leqslant 0 \quad \forall\, i = 1, \ldots, m, \\
& h_j(x) = 0 \quad \forall\, j = 1, \ldots, p, \\
& H_i(x) \geqslant 0 \quad \forall\, i = 1, \ldots, l, \\
& G_i(x)H_i(x) \leqslant 0 \quad \forall\, i = 1, \ldots, l,
\end{aligned}
$$

* This research was partially supported by the DFG (Deutsche Forschungsgemeinschaft) under grant KA1296/15-1.

that we call a *Mathematical program with vanishing constraints* (MPVC for short), where all functions $f, g_i, h_j, G_i, H_i \colon \mathbb{R}^n \to \mathbb{R}$ are assumed to be at least continuously differentiable. This special class of optimization problems was first introduced in [1] and shown to serve as a unified framework for several applications in structural and topology optimization. The naming of the problem is motivated by the fact that, on the one hand, it is closely related to the class of optimization problems called *Mathematical programs with equilibrium constraints* (MPECs for short), see [7], [10] for a general treatment and [1] for the relation between MPVCs and MPECs, and, on the other hand, that, due to the characteristic constraints $H_i(x) \geqslant 0$ and $G_i(x)H_i(x) \leqslant 0$, the implicit sign restriction $G_i(x) \leqslant 0$ vanishes as soon as $H_i(x) = 0$ holds. An MPVC may also be viewed as a special case of a mathematical program with combinatorial constraints as discussed in [13].

The recent papers on MPVCs have already investigated first-order necessary optimality conditions in depth. For example, in [1] the notion of a *strongly stationary point* was introduced and it was shown that a feasible point of an MPVC is strongly stationary if and only if it satisfies the KKT-conditions from standard optimization, and herewith, strong stationarity becomes a necessary optimality criterion under the presence of certain constraint qualifications, like the *Guignard CQ*, see, in particular, [5] for a more detailed discussion.

In turn, in [6], it was pointed out that the Guignard CQ, the weakest constraint qualification to garantuee the KKT-conditions to hold at a local minimizer of a standard optimization problem, holds under reasonable assumptions at a feasible point of an MPVC, but may yet be violated in some non-pathological cases. Thus, borrowing from the MPEC-theory, a weaker stationarity condition, called *M-stationarity* and holding under weaker constraint qualifications, was introduced and investigated in [6].

The goal of this paper is to extend the set of optimality conditions that can be stated in the MPVC-context. To this end, we state a new first-order sufficient condition and present both a second-order necessary and a second-order sufficient optimality condition for MPVCs.

The first-order sufficient condition, in particular, tells us that a strongly stationary point of an MPVC is already a local minimizer provided that the constraint functions g_i, h_j, G_i, H_i have certain convexity properties. We find this result quite astonishing since the MPVC itself is still a nonconvex program even if g_i, h_j, G_i, H_i have nice convexity properties, due to the product constraint $G_i(x)H_i(x) \leqslant 0$. In that part, some ideas go back to related results for MPECs which can be found, e.g., in [14].

As to the second-order conditions, our approach is motivated by corresponding results from standard optimization theory as well as some related results in the MPEC-setting, see, in particular, [11] and [7].

The organization of the paper is as follows: We first introduce some important index sets and preliminary definitions in Section 2. In particular, we recall the above mentioned stationarity concepts: strong stationarity and M-stationarity. In Section 3, the first-order sufficient optimality condition is stated, whereas the second-order optimality conditions are presented in Section 4. We close with some final remarks in Section 5.

The notation that we use in this paper is standard, with $\|\cdot\|$ being an arbitrary norm in $\mathbb{R}^n$. The directional derivative of a mapping $f\colon \mathbb{R}^n \to \mathbb{R}$ at x in the direction d is denoted by $f'(x;d)$. Recall that we have $f'(x;d) = \nabla f(x)^T d$ whenever f is differentiable at x.

2. Preliminaries

In this section, we introduce several index sets that turned out to be vital for the analysis of MPVCs. Furthermore, we give definitions of two stationarity concepts, *strong stationarity* and *M-stationarity*, which were introduced in the context of MPVCs in [1] and [6], respectively.

For these purposes, let X denote the feasible set of (1), and let $x^* \in X$ be an arbitrary feasible point. Then we define the index sets

$$\begin{aligned} J &:= \{1,\dots,p\}, \\ I_g &:= \{i \mid g_i(x^*) = 0\}, \\ I_+ &:= \{i \mid H_i(x^*) > 0\}, \\ I_0 &:= \{i \mid H_i(x^*) = 0\}. \end{aligned} \tag{2}$$

Furthermore, we divide the index set I_+ into the following subsets:

$$\begin{aligned} I_{+0} &:= \{i \mid H_i(x^*) > 0,\ G_i(x^*) = 0\}, \\ I_{+-} &:= \{i \mid H_i(x^*) > 0,\ G_i(x^*) < 0\}. \end{aligned} \tag{3}$$

Similarly, we partition the set I_0 in the following way:

$$\begin{aligned} I_{0+} &:= \{i \mid H_i(x^*) = 0,\ G_i(x^*) > 0\}, \\ I_{00} &:= \{i \mid H_i(x^*) = 0,\ G_i(x^*) = 0\}, \\ I_{0-} &:= \{i \mid H_i(x^*) = 0,\ G_i(x^*) < 0\}. \end{aligned} \tag{4}$$

Note that the first subscript indicates the sign of $H_i(x^*)$, whereas the second subscript stands for the sign of $G_i(x^*)$.

With the above definitions, we are now in a position to define the above mentioned stationarity concepts.

Definition 2.1. Let x^* be feasible for the MPVC (1). Then x^* is called *strongly stationary* if there exist scalars $\lambda_i \in \mathbb{R}$ $(i = 1, \dots, m)$, $\mu_j \in \mathbb{R}$ $(j \in J)$, $\eta_i^H, \eta_i^G \in \mathbb{R}$ $(i = 1, \dots, l)$ such that

$$
\begin{aligned}
(5) \qquad \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j \in J} \mu_j \nabla h_j(x^*) \\
- \sum_{i=1}^{l} \eta_i^H \nabla H_i(x^*) + \sum_{i=1}^{l} \eta_i^G \nabla G_i(x^*) = 0
\end{aligned}
$$

and

$$
\begin{gathered}
(6) \qquad \lambda_i \geqslant 0, \quad g_i(x^*) \leqslant 0, \quad \lambda_i g_i(x^*) = 0 \quad \forall\, i = 1, \dots, m, \\
\eta_i^H = 0 \ \ (i \in I_+), \quad \eta_i^H \geqslant 0 \ \ (i \in I_{0-} \cup I_{00}), \quad \eta_i^H \text{ free } \ (i \in I_{0+}), \\
\eta_i^G = 0 \ \ (i \in I_{+-} \cup I_0), \quad \eta_i^G \geqslant 0 \ \ (i \in I_{+0}).
\end{gathered}
$$

From [1], we know that strong stationarity is equivalent to the usual KKT conditions of an MPVC, i.e., strong stationarity is a necessary optimality condition under the presence of, e.g., the Guignard constraint qualification. See [5] for a more detailed discussion and sufficient conditions for the Guignard constraint qualification.

It may happen that a local minimum x^* of an MPVC is not a strongly stationary point even if all the mappings g_i, h_j, G_i, H_i are linear. In this case, a weaker stationary concept was introduced in [6], with the terminology coming from a similar concept for MPECs, see [9], [12], [3].

Definition 2.2. Let x^* be feasible for the MPVC (1). Then x^* is called *M-stationary* if there exist scalars $\lambda_i \in \mathbb{R}$ $(i = 1, \dots, m)$, $\mu_j \in \mathbb{R}$ $(j \in J)$, $\eta_i^H, \eta_i^G \in \mathbb{R}$ $(i = 1, \dots, l)$ such that

$$
\begin{aligned}
(7) \qquad \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j \in J} \mu_j \nabla h_j(x^*) \\
- \sum_{i=1}^{l} \eta_i^H \nabla H_i(x^*) + \sum_{i=1}^{l} \eta_i^G \nabla G_i(x^*) = 0
\end{aligned}
$$

and

$$
\begin{gathered}
(8) \qquad \lambda_i \geqslant 0, \quad g_i(x^*) \leqslant 0, \quad \lambda_i g_i(x^*) = 0 \quad \forall\, i = 1, \dots, m, \\
\eta_i^H = 0 \ \ (i \in I_+), \quad \eta_i^H \geqslant 0 \ \ (i \in I_{0-}), \quad \eta_i^H \text{ free } \ (i \in I_{0+}), \\
\eta_i^G = 0 \ \ (i \in I_{+-} \cup I_{0-} \cup I_{0+}), \quad \eta_i^G \geqslant 0 \ \ (i \in I_{+0} \cup I_{00}), \\
\eta_i^H \eta_i^G = 0 \ \ (i \in I_{00}).
\end{gathered}
$$

Note the difference between a strongly stationary point and an M-stationary point: In the former, we have $\eta_i^H \geqslant 0$ and $\eta_i^G = 0$ for all $i \in I_{00}$, whereas in the latter case, we only have $\eta_i^G \geqslant 0$ and $\eta_i^H \eta_i^G = 0$ for all $i \in I_{00}$. In particular, differences occur only for indices from the crucial index set I_{00}. In fact, this set will play an important role also in the analysis of the subsequent sections.

From [6, Theorem 3.4], we know that M-stationarity is a necessary optimality criterion under the presence of a condition that is called MPVC-GCQ, since it is an MPVC-version of the standard Guignard constraint qualification. This MPVC-GCQ condition is satisfied under very weak assumptions, in particular, it holds when all the mappings g_i, h_j, G_i, H_i are linear, see [6] for more details.

In the analysis of optimality conditions for standard nonlinear programs, the so-called Lagrangian plays an important role. As a counterpart of this Lagrangian in our MPVC setting, we define the mapping $L\colon \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^l \times \mathbb{R}^l \to \mathbb{R}$ by

$$
\begin{aligned}
(9) \qquad L(x, \lambda, \mu, \eta^G, \eta^H) := f(x) &+ \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j \in J} \mu_j h_j(x) \\
&- \sum_{i=1}^l \eta_i^H H_i(x) + \sum_{i=1}^l \eta_i^G G_i(x)
\end{aligned}
$$

and call this function the *MPVC-Lagrangian*. For example, a feasible point x^* of (1) is strongly stationary (or M-stationary) if and only if there exist multipliers $(\lambda, \mu, \eta^G, \eta^H)$ such that

$$\nabla_x L(x^*, \lambda, \mu, \eta^G, \eta^H) = 0$$

and $(\lambda, \mu, \eta^G, \eta^H)$ satisfies (6) (or (8)).

3. A first-order sufficient optimality condition

We know from the discussion of the previous section that both strong stationarity and M-stationarity are first-order necessary optimality conditions. In the case of a standard nonlinear program, the usual KKT conditions are also known to be sufficient optimality conditions under certain convexity assumptions. In our case, however, this result cannot be applied since the product term $G_i(x)H_i(x)$ usually does not satisfy any convexity requirements. Nevertheless, we will see in this section that M- and strong stationarity are also sufficient optimality conditions for our nonconvex MPVC problem, provided that the mappings g_i, h_j, G_i, H_i satisfy some convexity assumptions (but not necessarily the products $G_i H_i$ themselves). Our analysis here is motivated by a related result from [14] in the context of MPECs.

In order to state the desired result, we first recall some well-known terms concerning certain convexity properties of real-valued functions, see, for example, [2], [8].

Definition 3.1. Let $S \subseteq \mathbb{R}^n$ be a nonempty convex set and let $f\colon S \to \mathbb{R}$. Then f is called *quasiconvex* if, for each $x, y \in S$, the following inequality holds:

$$f(\lambda x + (1-\lambda)y) \leqslant \max\{f(x), f(y)\} \quad \forall\, \lambda \in (0,1).$$

Definition 3.2. Let $S \subseteq \mathbb{R}^n$ be a nonempty open set and let $f\colon S \to \mathbb{R}$ be a differentiable function. Then f is called *pseudoconvex* if, for each $x, y \in S$, the following implication holds:

$$\nabla f(x)^T(y-x) \geqslant 0 \Longrightarrow f(y) \geqslant f(x).$$

Now, let x^* be an M-stationary point of the MPVC (1) with corresponding multipliers λ, μ, η^G, η^H. Then we define the following index sets:

$$\begin{aligned} J^+ &:= \{j \in J \mid \mu_j > 0\},\\ J^- &:= \{j \in J \mid \mu_j < 0\},\\ I_{00}^+ &:= \{i \in I_{00} \mid \eta_i^H > 0\},\\ I_{00}^- &:= \{i \in I_{00} \mid \eta_i^H < 0\},\\ I_{0-}^+ &:= \{i \in I_{0-} \mid \eta_i^H > 0\},\\ I_{0+}^+ &:= \{i \in I_{0+} \mid \eta_i^H > 0\},\\ I_{0+}^- &:= \{i \in I_{0+} \mid \eta_i^H < 0\},\\ I_{+0}^{0+} &:= \{i \in I_{+0} \mid \eta_i^H = 0,\ \eta_i^G > 0\} = \{i \in I_{+0} \mid \eta_i^G > 0\},\\ I_{00}^{0+} &:= \{i \in I_{00} \mid \eta_i^H = 0,\ \eta_i^G > 0\} = \{i \in I_{00} \mid \eta_i^G > 0\}. \end{aligned} \tag{10}$$

Note that, for a strongly stationary point, the two index sets I_{00}^- and I_{00}^{0+} are empty.

Using these index sets and definitions, we are able to state the main result of this section.

Theorem 3.3. *Let x^* be an M-stationary point of the MPVC* (1). *Suppose that f is pseudoconvex at x^* and that g_i $(i \in I_g)$, h_j $(j \in J^+)$, $-h_j$ $(j \in J^-)$, G_i $(i \in I_{+0}^{0+})$, H_i $(i \in I_{0+}^-)$, $-H_i$ $(i \in I_{0+}^+ \cup I_{00}^+ \cup I_{0-}^+)$ are quasiconvex. Then the following statements hold:*

(a) *If $I_{00}^- \cup I_{00}^{0+} = \emptyset$ then x^* is a local minimizer of* (1).

(b) *If $I_{0+}^- \cup I_{00}^- \cup I_{+0}^{0+} \cup I_{00}^{0+} = \emptyset$ then x^* is a global minimizer of* (1).

Proof. Since x^* is an M-stationary point of (1) there exist multipliers λ, μ, η^G, η^H such that

$$\begin{aligned} \nabla f(x^*) + \sum_{i \in I_g} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(x^*) - \sum_{i \in I_0} \eta_i^H \nabla H_i(x^*)&\\ + \sum_{i \in I_{+0} \cup I_{00}} \eta_i^G \nabla G_i(x^*) &= 0 \end{aligned} \tag{11}$$

with

$$\lambda_i \geqslant 0 \ \forall i \in I_g, \quad \eta_i^H \geqslant 0 \ \forall i \in I_{0-}, \tag{12}$$
$$\eta_i^G \geqslant 0 \ \forall i \in I_{00} \cup I_{+0}, \quad \eta_i^H \eta_i^G = 0 \ \forall i \in I_{00}.$$

Now let x be any feasible point of (1). For $i \in I_g$, we then have $g_i(x) \leqslant 0 = g_i(x^*)$. Thus, by the quasiconvexity of g_i $(i \in I_g)$, we obtain

$$g_i(x^* + t(x - x^*)) = g_i((1-t)x^* + tx) \leqslant \max\{g_i(x), g_i(x^*)\} = 0 = g_i(x^*)$$

for all $t \in (0,1)$, which implies

$$\begin{aligned}\nabla g_i(x^*)^T(x - x^*) &= g_i'(x^*; x - x^*) \\ &= \lim_{t \downarrow 0} \frac{g_i(x^* + t(x - x^*)) - g_i(x^*)}{t} \leqslant 0 \quad \forall i \in I_g.\end{aligned}$$

In view of (12), we therefore have

$$\lambda_i \nabla g_i(x^*)^T(x - x^*) \leqslant 0 \quad \forall i \in I_g. \tag{13}$$

By similar arguments, we also obtain

$$\nabla h_j(x^*)^T(x - x^*) \leqslant 0 \quad \forall j \in J^+, \quad \text{and} \quad -\nabla h_j(x^*)^T(x - x^*) \leqslant 0 \quad \forall j \in J^-,$$

which gives

$$\mu_j \nabla h_j(x^*)^T(x - x^*) \leqslant 0 \quad \forall j \in J, \tag{14}$$

taking the definitions of J^+ and J^- into account.

Again, since x is feasible for (1), we have in particular $-H_i(x) \leqslant 0$ for all $i = 1, \ldots, l$. Thus, by the quasiconvexity of $-H_i$ for $i \in I_{0+}^+ \cup I_{00}^+ \cup I_{0-}^+$, we obtain by the above arguments $-\nabla H_i(x^*)^T(x - x^*) \leqslant 0$ and thus, in view of the definition of the occurring index sets, we have

$$-\eta_i^H \nabla H_i(x^*)^T(x - x^*) \leqslant 0 \quad \forall i \in I_{0+}^+ \cup I_{00}^+ \cup I_{0-}^+. \tag{15}$$

We now verify the statement (b) first. To this end, let $I_{0+}^- \cup I_{00}^- \cup I_{+0}^{0+} \cup I_{00}^{0+} = \emptyset$. Then it is clear from (12), (15), and the definition of the index sets that we even have

$$-\eta_i^H \nabla H_i(x^*)^T(x - x^*) \leqslant 0 \ \forall i \in I_0, \quad \eta_i^G \nabla G_i(x^*)^T(x - x^*) \leqslant 0 \ \forall i \in I_{00} \cup I_{+0}, \tag{16}$$

where the second inequality is an equality due to the fact that $\eta_i^G = 0$ for all (remaining) indices $i \in I_{00} \cup I_{+0}$. Then (13), (14), (16) together with (11) imply

$$-\nabla f(x^*)^T(x - x^*) = \Bigg(\sum_{i \in I_g} \lambda_i \nabla g_i(x^*)^T + \sum_{j=1}^{p} \mu_j \nabla h_j(x^*) - \sum_{i \in I_0} \eta_i^H \nabla H_i(x^*) + \dots$$
$$+ \sum_{i \in I_{+0} \cup I_{00}} \eta_i^G \nabla G_i(x^*)\Bigg)^T (x - x^*) \leqslant 0.$$

Hence we have $\nabla f(x^*)^T(x - x^*) \geqslant 0$, which implies $f(x) \geqslant f(x^*)$, as f is pseudoconvex by assumption. Since x is an arbitrary feasible point of (1), x^* is a global minimizer of (1) in the case that $I_{0+}^- \cup I_{00}^- \cup I_{+0}^{0+} \cup I_{00}^{0+} = \emptyset$ holds, which proves the assertion (b).

To verify the statement (a), we only need to show, in view of the above arguments, that for any feasible x sufficiently close to x^*, we have

$$-\eta_i^H \nabla H_i(x^*)^T(x - x^*) \leqslant 0 \quad \forall\, i \in I_{0+}^- \tag{17}$$

and

$$\eta_i^G \nabla G_i(x^*)^T(x - x^*) \leqslant 0 \quad \forall\, i \in I_{+0}^{0+}, \tag{18}$$

since then we see that (13), (14) and (16) are satisfied, and thus, by analogous reasoning as above, we obtain $f(x) \geqslant f(x^*)$ for all feasible x sufficiently close to x^*.

First let $i \in I_{0+}^-$. By continuity, it follows that $G_i(x) > 0$ and thus $H_i(x) = 0$ for any $x \in X$ sufficiently close to x^*. Invoking the quasiconvexity of H_i $(i \in I_{0+}^-)$, this implies $\nabla H_i(x^*)^T(x - x^*) \leqslant 0$, and since we have $\eta_i^H < 0$ $(i \in I_{0+}^-)$, (17) follows immediately.

Second, let $i \in I_{+0}^{0+}$. By continuity, it follows that $H_i(x) > 0$ and thus $G_i(x) \leqslant 0$ for any $x \in X$ sufficiently close to x^*. Invoking the quasiconvexity of G_i $(i \in I_{+0}^{0+})$, this implies $\nabla G_i(x^*)^T(x - x^*) \leqslant 0$, which gives (18), since we have $\eta_i^G > 0$ $(i \in I_{+0}^{0+})$. □

We next state a simple consequence of Theorem 3.3 where the M-stationarity of x^* is replaced by the strong stationarity assumption.

Corollary 3.4. *Let x^* be a strongly stationary point of the MPVC* (1). *Suppose that f is pseudoconvex at x^* and that g_i $(i \in I_g)$, h_j $(j \in J^+)$, $-h_j$ $(j \in J^-)$, G_i $(i \in I_{+0}^{0+})$, H_i $(i \in I_{0+}^-)$, $-H_i$ $(i \in I_{0+}^+ \cup I_{00}^+ \cup I_{0-}^+)$ are quasiconvex. Then the following statements hold:*

(a) *x^* is a local minimizer of* (1).

(b) *If* $I_{0+}^{-} \cup I_{+0}^{0+} = \emptyset$ *then* x^* *is a global minimizer of* (1).

P r o o f. Since the assumptions of Theorem 3.3 are satisfied and strong stationarity implies that $I_{00}^{-} \cup I_{00}^{0+} = \emptyset$, (a) and (b) follow immediately from Theorem 3.3 (a) and (b), respectively. □

In nonlinear programming, the case of a convex program, where all the equality constraints are supposed to be (affine) linear and the inequality constraints are convex, is often considered. However, due to the G_iH_i-constraints, being a product of two non-constant functions, our MPVC (1) is very likely a nonconvex optimization problem. Alternatively, the concept of an *MPVC-convex program* was therefore introduced in [6], where all the functions h_j, H_i, G_i are supposed to be (affine) linear and the functions g_i are supposed to be convex. For the class of MPVC-convex programs, we now get the following first-order sufficient optimality condition as a direct consequence of our previous results.

Corollary 3.5. *Let the program* (1) *be MPVC-convex such that* f *is convex. Furthermore, let* x^* *be a strongly stationary point of* (1). *Then the following statements hold:*

(a) x^* *is a local minimizer of* (1).

(b) *If* $I_{0+}^{-} \cup I_{+0}^{0+} = \emptyset$, *then* x^* *is a global minimizer of* (1).

P r o o f follows immediately from Corollary 3.4, since convex functions are both pseudo- and quasiconvex. □

We would like to point out that we find the above result somehow remarkable: The MPVC-convex program, though being equipped with convex and linear functions g_i, h_j, H_i, G_i, must yet be assumed to be a nonconvex program, due to the G_iH_i-constraints. Nevertheless, Corollary 3.5 tells us that strong stationarity (and thus the KKT-conditions themselves) are sufficient optimality conditions. That means, we have shown the KKT-conditions to be a sufficient optimality criterion for a class of usually nonconvex programs.

At this point it might be useful to go through a simple example of an MPVC, in order to illustrate some of the above introduced concepts and results.

E x a m p l e 3.6. For $a, b \in \mathbb{R}$ consider the following two-dimensional MPVC:

$$\begin{aligned} \text{(19)} \qquad \min \quad & f(x) := (x_1 - a)^2 + (x_2 - b)^2 \\ \text{s.t.} \quad & H(x) := x_1 \geqslant 0, \\ & G(x)H(x) := x_2 x_1 \leqslant 0. \end{aligned}$$

Its feasible set and also some relevant points for the upcoming discussion are given in Fig. 1. Geometrically speaking, in (19), one is searching for the projection of (a, b) onto the feasible set.

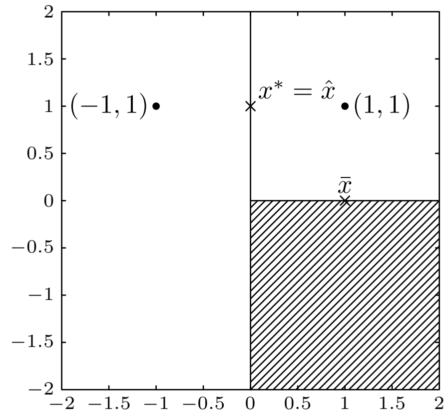

Figure 1. Feasible set of (19).

First of all, we see that the gradients $\nabla H(x) = (1,0)^T$ and $\nabla G(x) = (0,1)^T$ are linearly independent for all $x \in \mathbb{R}^2$, hence, the *MPVC-LICQ*, which will be introduced in Definition 4.1, is satisfied at any feasible point. Therefore, strong stationarity is a necessary optimality condition.

Furthermore, the function f is convex and the functions G, H are linear. Thus, the program is MPVC-convex (but still nonconvex!). By Corollary 3.5, we know then that strong stationarity is a sufficient condition for a local minimizer and, under some additional condition concerning certain index sets, even for a global minimizer. Together, the above considerations yield that a feasible point of (19) is a local minimizer if and only if it is a strongly stationary point. We will verify this by considering the above MPVC for two different choices of (a,b) and calculating the respective strongly stationary points.

For all choices (a,b), the strong stationarity conditions of (19) read

$$0 = \begin{pmatrix} 2x_1 - 2a \\ 2x_2 - 2b \end{pmatrix} - \eta^H \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \eta^G \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{20}$$

with

$$\eta^H \begin{cases} = 0, & \text{if } x_1 > 0, \\ \geqslant 0, & \text{if } x_1 = 0,\ x_2 \leqslant 0, \\ \text{free}, & \text{if } x_1 = 0,\ x_2 > 0, \end{cases} \qquad \eta^G \begin{cases} \geqslant 0, & \text{if } x_1 > 0,\ x_2 = 0, \\ = 0, & \text{else.} \end{cases} \tag{21}$$

For the choice $(a,b) := (1,1)$, it is quickly calculated that there are two strongly stationary points. The first one is $\hat{x} := (0,1)^T$ with associated multipliers $\hat{\eta}^G := 0$, $\hat{\eta}^H := -2$. The second point is $\tilde{x} := (1,0)^T$, where the corresponding multipliers are given by $\tilde{\eta}^G := 2$, $\tilde{\eta}^H := 0$. These are the only local minimizers of (19), as was argued above, for the special choice $(a,b) := (1,1)$. In fact, they are even global minimizers as can be seen easily by geometric arguments, even though the sufficient condition from Corollary 3.5 (b) is not satisfied, illustrating that this is only a *sufficient* criterion.

The next choice is $(a,b) := (-1,1)$, where we can compute only one strongly stationary point $x^* := (0,1)^T$ with multipliers given by $\eta^G := 0$, $\eta^H := 2$. In particular, we then have $I_{0+}^- \cup I_{+0}^{0+} = \emptyset$, so that, in this case, we can invoke Corollary 3.5 (b) to ensure that this is not only a local, but a global minimizer of (19).

4. Second-order optimality conditions

The goal of this section is to provide (necessary and sufficient) second-order optimality conditions for MPVCs. The analysis is motivated by general results from optimization or, more specialized, from the MPEC-field.

In order to state second-order optimality results for nonlinear programs, a suitable cone, usually a subset of the linearized cone, is needed, on which the Hessian of the Lagrangian is or is shown to be positive (semi-)definite. The cone which plays that role in our context will be defined below and is a subset of the so-called *MPVC-linearized cone* which was initially introduced in [5]. Given a feasible point x^* of (1), the MPVC-linearized cone is defined by

$$
(22)\quad \mathcal{L}_{\mathrm{MPVC}}(x^*) := \{d \in \mathbb{R}^n \mid \begin{array}{ll} \nabla g_i(x^*)^T d \leqslant 0 & (i \in I_g), \\ \nabla h_j(x^*)^T d = 0 & (j \in J), \\ \nabla H_i(x^*)^T d = 0 & (i \in I_{0+}), \\ \nabla H_i(x^*)^T d \geqslant 0 & (i \in I_{00} \cup I_{0-}), \\ \nabla G_i(x^*)^T d \leqslant 0 & (i \in I_{+0}), \\ (\nabla H_i(x^*)^T d)(\nabla G_i(x^*)^T d) \leqslant 0 & (i \in I_{00})\}. \end{array}
$$

In many situations of MPVC-analysis, the MPVC-linearized cone has been used instead of the usual linearized cone. Thus, it is not surprising that it occurs in the context of second-order optimality conditions for MPVCs, too.

For the definition of the above mentioned subset of the MPVC-linearized cone, we assume that we have a strongly stationary point $(x^*, \lambda, \mu, \eta^G, \eta^H)$ of (1). Then we define $\mathcal{C}(x^*)$ by

$$
(23)\qquad \mathcal{C}(x^*) := \{d \in \mathcal{L}_{\mathrm{MPVC}}(x^*) \mid \begin{array}{ll} \nabla g_i(x^*)^T d = 0 & (i \in I_g^+), \\ \nabla H_i(x^*)^T d = 0 & (i \in I_{00}^+ \cup I_{0-}^+), \\ \nabla G_i(x^*)^T d = 0 & (i \in I_{+0}^{0+})\}, \end{array}
$$

that is, in fact, we have (taking into account that $I_{00}^{-}=\emptyset$ at a strongly stationary point)

$$
\begin{aligned}
(24)\quad \mathcal{C}(x^*)=\{d\in\mathbb{R}^n \mid \nabla g_i(x^*)^T d &\leqslant 0 && (i\in I_g^0),\\
\nabla g_i(x^*)^T d &= 0 && (i\in I_g^+),\\
\nabla h_j(x^*)^T d &= 0 && (j\in J),\\
\nabla H_i(x^*)^T d &\geqslant 0 && (i\in I_{00}^0\cup I_{0-}^0),\\
\nabla H_i(x^*)^T d &= 0 && (i\in I_{0+}\cup I_{00}^+\cup I_{0-}^+),\\
\nabla G_i(x^*)^T d &\leqslant 0 && (i\in I_{+0}^{00}),\\
\nabla G_i(x^*)^T d &= 0 && (i\in I_{+0}^{0+}),\\
(\nabla H_i(x^*)^T d)(\nabla G_i(x^*)^T d) &\leqslant 0 && (i\in I_{00})\},
\end{aligned}
$$

where we put

$$
\begin{aligned}
(25)\qquad I_g^+ &:= \{i\in I_g \mid \lambda_i>0\},\\
I_g^0 &:= \{i\in I_g \mid \lambda_i=0\},\\
I_{00}^+ &:= \{i\in I_{00} \mid \eta_i^H>0\},\\
I_{00}^0 &:= \{i\in I_{00} \mid \eta_i^H=0\},\\
I_{0-}^+ &:= \{i\in I_{0-} \mid \eta_i^H>0\},\\
I_{0-}^0 &:= \{i\in I_{0-} \mid \eta_i^H=0\},\\
I_{+0}^{00} &:= \{i\in I_{+0} \mid \eta_i^G=0\},\\
I_{+0}^{0+} &:= \{i\in I_{+0} \mid \eta_i^G>0\}
\end{aligned}
$$

in accordance with (10).

The definition of these index sets may, again, appear a bit complicated and make the proof of our theorems somewhat technical, but on the other hand we prove pretty strong results, showing that we can use the same cone $\mathcal{C}(x^*)$ for both the necessary and the sufficient second-order condition.

The following lemma is a direct preparation for the upcoming theorem on second-order necessary optimality conditions. Its technique of proof goes back to similar considerations in the context of standard nonlinear programs, see [4], for example. Note, however, that we cannot simply apply these standard results since, e.g., the usual LICQ assumption typically does not hold for MPVCs, see [1]. Instead of this, we use the MPVC-version of LICQ which was initially introduced in [5]. We recall its definition below.

Definition 4.1. We say that MPVC-LICQ is satisfied at a feasible point x^* of (1) if the gradients

$$\begin{aligned}
\nabla h_j(x^*) &\quad (j = 1, \ldots, p),\\
\nabla g_i(x^*) &\quad (i \in I_g),\\
\nabla H_i(x^*) &\quad (i \in I_0),\\
\nabla G_i(x^*) &\quad (i \in I_{00} \cup I_{+0}),
\end{aligned}$$

are linearly independent.

Note that for the whole section, all functions occurring in (1) are assumed to be at least twice continuously differentiable.

Lemma 4.2. *Let x^* be a strongly stationary point of* (1) *such that MPVC-LICQ holds. Furthermore, let $d \in \mathcal{C}(x^*)$. Then there exists an $\varepsilon > 0$ and a twice continuously differentiable curve $x\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ such that $x(0) = x^*$, $x'(0) = d$, $x(t) \in X$ for $t \in [0, \varepsilon)$ and such that, in addition, we have*

$$\begin{aligned}
g_i(x(t)) = 0 &\quad (i \in I_g^+),\\
h_j(x(t)) = 0 &\quad (j \in J),\\
H_i(x(t)) = 0 &\quad (i \in I_{00}^+ \cup I_{0-}^+ \cup I_{0+}),\\
G_i(x(t)) = 0 &\quad (i \in I_{+0}^{0+}).
\end{aligned} \tag{26}$$

Proof. Let $d \in \mathcal{C}(x^*)$ and let $(\lambda, \mu, \eta^G, \eta^H)$ be the (unique) multipliers such that $(x^*, \lambda, \mu, \eta^G, \eta^H)$ is a strongly stationary point. We define some further subsets (depending on x^* and the particular vector d chosen from $\mathcal{C}(x^*)$) of the index sets which were defined previously:

$$\begin{aligned}
I_{g,=}^0 &:= \{i \in I_g^0 \mid \nabla g_i(x^*)^T d = 0\},\\
I_{g,<}^0 &:= \{i \in I_g^0 \mid \nabla g_i(x^*)^T d < 0\},\\
I_{00,=}^0 &:= \{i \in I_{00}^0 \mid \nabla H_i(x^*)^T d = 0\},\\
I_{00,>}^0 &:= \{i \in I_{00}^0 \mid \nabla H_i(x^*)^T d > 0\},\\
I_{0-,=}^0 &:= \{i \in I_{0-}^0 \mid \nabla H_i(x^*)^T d = 0\},\\
I_{0-,>}^0 &:= \{i \in I_{0-}^0 \mid \nabla H_i(x^*)^T d > 0\},\\
I_{+0,*=}^{00} &:= \{i \in I_{+0}^{00} \mid \nabla G_i(x^*)^T d = 0\},\\
I_{+0,*<}^{00} &:= \{i \in I_{+0}^{00} \mid \nabla G_i(x^*)^T d < 0\},\\
I_{00,>=}^0 &:= \{i \in I_{00}^0 \mid \nabla H_i(x^*)^T d > 0,\ \nabla G_i(x^*)^T d = 0\},\\
I_{00,><}^0 &:= \{i \in I_{00}^0 \mid \nabla H_i(x^*)^T d > 0,\ \nabla G_i(x^*)^T d < 0\}.
\end{aligned} \tag{27}$$

Then we define the mapping $z\colon \mathbb{R}^n \to \mathbb{R}^q$, where $q := |I_g^+ \cup I_{g,=}^0| + |J| + |I_{0+} \cup I_{00}^+ \cup I_{0-}^+ \cup I_{00,=}^0 \cup I_{0-,=}^0| + |I_{+0}^{0+} \cup I_{0+,*=}^{00} \cup I_{00,>=}^0|$, by

$$
(28) \qquad z(x) := \begin{pmatrix} g_i(x) & (i \in I_g^+ \cup I_{g,=}^0) \\ h_j(x) & (j \in J) \\ H_i(x) & (i \in I_{0+} \cup I_{00}^+ \cup I_{0-}^+ \cup I_{00,=}^0 \cup I_{0-,=}^0) \\ G_i(x) & (I_{+0}^{0+} \cup I_{+0,*=}^{00} \cup I_{00,>=}^0) \end{pmatrix},
$$

and denote the jth component function of z by z_j. Furthermore, let $\overline{H}\colon \mathbb{R}^{q+1} \to \mathbb{R}^q$ be the mapping defined by

$$
\overline{H}_j(y,t) := z_j\big(x^* + td + z'(x^*)^T y\big) \quad \forall\, j = 1, \dots, q.
$$

The system $\overline{H}(y,t) = 0$ has a solution $(y^*, t^*) := (0,0)$, and the partial Jacobian

$$
\overline{H}_y(0,0) = z'(x^*) z'(x^*)^T \in \mathbb{R}^{q \times q}
$$

is nonsingular since the matrix $z'(x^*)$ has full rank q due to the MPVC-LICQ assumption. Thus, invoking the implicit function theorem and using the twice continuous differentiability of all mappings involved in the definition of z, there exists an $\varepsilon > 0$ and a twice continuously differentiable curve $y\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^q$ such that $y(0) = 0$ and $\overline{H}(y(t), t) = 0$ for all $t \in (-\varepsilon, \varepsilon)$. Moreover, its derivative is given by

$$
y'(t) = -\big(\overline{H}_y(y(t), t)\big)^{-1} \overline{H}_t(y(t), t) \quad \forall\, t \in (-\varepsilon, \varepsilon).
$$

In particular, this implies

$$
y'(0) = -(\overline{H}_y(0,0))^{-1} \overline{H}_t(0,0) = -(\overline{H}_y(0,0))^{-1} \underbrace{z'(x^*)d}_{=0} = 0,
$$

due to the properties of d. Now define

$$
x(t) := x^* + td + z'(x^*)^T y(t).
$$

Then $x(\cdot)$ is twice continuously differentiable on $(-\varepsilon, \varepsilon)$, and we obviously have $x(0) = x^*$ and $x'(0) = d$. Hence, we still need to show that $x(t) \in X$ and that $x(\cdot)$ satisfies (26) for all t sufficiently close to 0.

For these purposes, first note that $\overline{H}_j(y(t), t) = 0$ implies $z_j(x(t)) = 0$ and thus we obtain

$$
(29) \qquad \begin{aligned} g_i(x(t)) &= 0 \quad (i \in I_g^+ \cup I_{g,=}^0), \\ h_j(x(t)) &= 0 \quad (j \in J), \\ H_i(x(t)) &= 0 \quad (i \in I_{0+} \cup I_{00}^+ \cup I_{0-}^+ \cup I_{00,=}^0 \cup I_{0-,=}^0), \\ G_i(x(t)) &= 0 \quad (i \in I_{+0}^{0+} \cup I_{+0,*=}^{00} \cup I_{00,>=}^0), \end{aligned}
$$

so that (26) and the feasibility of $x(t)$ for the above occurring index sets is guaranteed for all $t \in (-\varepsilon, \varepsilon)$.

By simple continuity arguments, one can also verify that we have $g_i(x(t)) < 0$ ($i \notin I_g$), $G_i(x(t)) < 0$ ($i \in I_{0-} \cup I_{+-}$) and $H_i(x(t)) > 0$ ($i \in I_+$) for all t sufficiently close to 0. Thus, taking the definition of $\mathcal{C}(x^*)$ into account, it remains to show that

$$\begin{aligned} g_i(x(t)) &\leqslant 0 \quad (i \in I^0_{g,<}), \\ H_i(x(t)) &\geqslant 0 \quad (i \in I^0_{00,>} \cup I^0_{0-,>}), \end{aligned} \tag{30}$$

and that

$$G_i(x(t))H_i(x(t)) \leqslant 0 \quad (i \in I^0_{00,><} \cup I^0_{0-,>} \cup I^{00}_{+0,*<}) \tag{31}$$

for $t > 0$ sufficiently small.

In order to verify (30), let $i \in I^0_{g,<}$. Then we have $\nabla g_i(x^*)^T d < 0$ by definition. This implies $\nabla g_i(x(\tau))^T x'(\tau) < 0$ for all $|\tau|$ sufficiently small. From the mean value theorem, we obtain a $\tau_t \in (0,t)$ such that $g_i(x(t)) = g_i(x(0)) + \nabla g_i(x(\tau_t))^T x'(\tau_t) \times (t-0) = t\nabla g_i(x(\tau_t))^T x'(\tau_t) < 0$ for all $t > 0$ sufficiently small, which proves the first statement of (30).

In order to prove the second statement, let $i \in I^0_{00,>} \cup I^0_{0-,>}$. Then it follows, by definition, that $\nabla H_i(x^*)^T d > 0$, and thus by continuity, it holds that $\nabla H_i((x(t))^T x'(t) > 0$ for all t sufficiently close to 0. Since we have $H_i(x(0)) = H_i(x^*) = 0$, this implies $H_i(x(t)) > 0$ for all $t > 0$ sufficiently small, using the above arguments.

To verify (31), first let $i \in I^0_{0-,>}$. Then we have $G_i(x(t)) < 0$ by continuity, and by the above reasoning we get $H_i(x(t)) > 0$ for $t > 0$ sufficiently small, so that $G_i(x(t))H_i(x(t)) \leqslant 0$ holds in this case.

Now, let $i \in I^0_{00,><}$. Then, by definition, we have $\nabla H_i(x^*)^T d > 0$ and $\nabla G_i(x^*)^T d < 0$. Then, by analogous reasoning as above, it follows that $H_i(x(t)) > 0$ and $G_i(x(t)) < 0$ for $t > 0$ sufficiently small, which gives (31) in this case.

Finally, let $i \in I^{00}_{+0,*<}$. Then we have $H_i(x(t)) > 0$ for $|t|$ sufficiently small. And since we have $\nabla G_i(x^*)^T d < 0$, we obtain $G_i(x(t)) < 0$ for all $t > 0$ sufficiently small, which eventually proves (31). $\square$

The proof of the following theorem exploits the existence of the curve x from the above lemma.

Theorem 4.3. *Let x^* be a local minimizer of* (1) *such that MPVC-LICQ holds. Then we have*

$$d^T \nabla^2_{xx} L(x^*, \lambda, \mu, \eta^G, \eta^H) d \geqslant 0 \quad \forall\, d \in \mathcal{C}(x^*),$$

where λ, μ, η^G, η^H are the (unique) multipliers corresponding to (the strongly stationary) point x^* of (1).

P r o o f. First recall from [5] that MPVC-LICQ implies that there exist (unique) multipliers such that $(x^*, \lambda, \mu, \eta^G, \eta^H)$ is a strongly stationary point.

Let $d \in \mathcal{C}(x^*)$. Using the curve $x(\cdot)$ (and $\varepsilon > 0$) from Lemma 4.2, we are in a position to define the function $\varphi\colon (-\varepsilon, \varepsilon) \to \mathbb{R}$ by

$$\varphi(t) := L(x(t), \lambda, \mu, \eta^G, \eta^H),$$

where L denotes the MPVC-Lagrangian from (9). Then φ is twice continuously differentiable with

$$\varphi'(t) = x'(t)^T \nabla_x L(x(t), \lambda, \mu, \eta^G, \eta^H)$$

and

$$\varphi''(t) = x''(t)^T \nabla_x L(x(t), \lambda, \mu, \eta^G, \eta^H) + x'(t)^T \nabla^2_{xx} L(x(t), \lambda, \mu, \eta^G, \eta^H) x'(t).$$

Using Lemma 4.2, we therefore obtain

$$\varphi'(0) = d^T \nabla_x L(x^*, \lambda, \mu, \eta^G, \eta^H) = 0$$

and

$$\varphi''(0) = d^T \nabla^2_{xx} L(x^*, \lambda, \mu, \eta^G, \eta^H) d,$$

since we have $\nabla_x L(x^*, \lambda, \mu, \eta^G, \eta^H) = 0$, as $(x^*, \lambda, \mu, \eta^G, \eta^H)$ is a strongly stationary point of (1).

Now, suppose that $\varphi''(0) = d^T \nabla^2_{xx} L(x^*, \lambda, \mu, \eta^G, \eta^H) d < 0$. By continuity, we thus have $\varphi''(t) < 0$ for t sufficiently close to 0. Invoking Taylor's formula, we obtain

$$\varphi(t) = \varphi(0) + t\varphi'(0) + \frac{t^2}{2} \varphi''(\xi_t)$$

for all $t \in (-\varepsilon, \varepsilon)$ and a suitable point ξ_t depending on t. Since we have $\varphi'(0) = 0$ and $\varphi''(\xi_t) < 0$ for t sufficiently close to 0, we thus have $\varphi(t) < \varphi(0)$ for these $t \in (-\varepsilon, \varepsilon)$. Since $(x^*, \lambda, \mu, \eta^G, \eta^H)$ is a strongly stationary point of (1), we have

$$\varphi(0) = f(x^*) + \sum_{i \in I_g} \lambda_i g_i(x^*) + \sum_{j \in J} \mu_j h_j(x^*) + \sum_{i \in I_{+0}} \eta_i^G G_i(x^*) - \sum_{i \in I_0} \eta_i^H H_i(x^*) = f(x^*)$$

and, in view of (26) and the feasibility of $x(t)$ for $t > 0$ sufficiently small, we also have

$$\begin{aligned} \varphi(t) = f(x(t)) &+ \sum_{i \in I_g} \lambda_i g_i(x(t)) + \sum_{j \in J} \mu_j h_j(x(t)) \\ &+ \sum_{i \in I_{+0}} \eta_i^G G_i(x(t)) - \sum_{i \in I_0} \eta_i^H H_i(x(t)) = f(x(t)), \end{aligned}$$

which yields $f(x(t)) < f(x^*)$ for all $t > 0$ sufficiently small, in contradiction to x^* being a local minimizer of (1). $\square$

We next state a second-order sufficiency condition. Note, again, that this result makes use of the same set $\mathcal{C}(x^*)$ as the second-order necessary condition from Theorem 4.3.

Theorem 4.4. *Let $(x^*, \lambda, \mu, \eta^G, \eta^H)$ be a strongly stationary point of the MPVC* (1) *such that*

$$d^T \nabla^2_{xx} L(x^*, \lambda, \mu, \eta^G, \eta^H) d > 0 \quad \forall\, d \in \mathcal{C}(x^*) \setminus \{0\}. \tag{32}$$

Then x^* is a strict local minimizer of (1).

Proof. Assume that x^* is not a strict local minimizer of (1). Then there exists a sequence $\{x^k\} \subseteq X$ tending to x^* with $f(x^k) \leqslant f(x^*)$ for all k. Now, put $t_k := \|x^k - x^*\|$. Then we have $t_k \downarrow 0$. Furthermore, we define the sequence $\{d^k\} \subseteq \mathbb{R}^n$ by $d^k := (x^k - x^*)/t_k$. Since we have $\|d^k\| = 1$ for all $k \in \mathbb{N}$, we can assume, without loss of generality, that $\{d^k\}$ has a limit $d \in \mathbb{R}^n \setminus \{0\}$. Furthermore, by construction, we see that d lies in the tangent cone $\mathcal{T}(x^*)$ of (1) and thus, invoking Corollary 2.5 from [5], we have in particular $d \in \mathcal{L}_{\mathrm{MPVC}}(x^*)$. Hence, we have

$$\begin{aligned}
\nabla g_i(x^*)^T d &\leqslant 0 \quad (i \in I_g),\\
\nabla h_j(x^*)^T d &= 0 \quad (j \in J),\\
\nabla H_i(x^*)^T d &= 0 \quad (i \in I_{0+}),\\
\nabla H_i(x^*)^T d &\geqslant 0 \quad (i \in I_{00} \cup I_{0-}),\\
\nabla G_i(x^*)^T d &\leqslant 0 \quad (i \in I_{+0}),
\end{aligned} \tag{33}$$

as well as

$$\bigl(\nabla G_i(x^*)^T d\bigr)\bigl(\nabla H_i(x^*)^T d\bigr) \leqslant 0 \quad (i \in I_{00}). \tag{34}$$

Furthermore, since we have $f(x^k) \leqslant f(x^*)$ for all k by assumption, the mean value theorem yields a vector ξ^k on the connecting line between x^k and x^* such that $\nabla f(\xi^k)^T (x^k - x^*) \leqslant 0$ for all k. Dividing by $\|x^k - x^*\|$ and passing to the limit thus implies

$$\nabla f(x^*)^T d \leqslant 0. \tag{35}$$

Now, we consider two different cases, which both lead to a contradiction.

First, consider the case that equality holds in (33) for all indices $i \in I_g^+ \cup I_{0-}^+ \cup I_{00}^+ \cup I_{+0}^{0+}$. Then we have $d \in \mathcal{C}(x^*)$. Since x^k is feasible for (1) for all k and we have $x^k \to x^*$, the following statements hold for all k sufficiently large:

$$
\begin{aligned}
\lambda_i \underbrace{g_i(x^k)}_{\leqslant 0} &\leqslant 0 \quad (i \in I_g), \\
\mu_j \underbrace{h_j(x^k)}_{=0} &= 0 \quad (j \in J), \\
\eta_i^H \underbrace{H_i(x^k)}_{=0} &= 0 \quad (i \in I_{0+}), \\
-\eta_i^H \underbrace{H_i(x^k)}_{\geqslant 0} &\leqslant 0 \quad (i \in I_{0-} \cup I_{00}), \\
\eta_i^G \underbrace{G_i(x^k)}_{\leqslant 0} &\leqslant 0 \quad (i \in I_{+0}),
\end{aligned}
\tag{36}
$$

where we use continuity arguments as well as the fact that we have $G_i(x^k)H_i(x^k) \leqslant 0$ for all $i = 1, \ldots, l$ and all k, for the third and fifth statement. Invoking (36) and the properties of the multipliers $(\lambda, \mu, \eta^G, \eta^H)$, we obtain

$$
\begin{aligned}
f(x^*) \geqslant f(x^k) \geqslant f(x^k) &+ \sum_{i \in I_g} \lambda_i g_i(x^k) + \sum_{j \in J} \mu_j h_j(x^k) \\
&+ \sum_{i \in I_{+0}} \eta_i^G G_i(x^k) - \sum_{i \in I_0} \eta_i^H H_i(x^k) = l(x^k),
\end{aligned}
\tag{37}
$$

where we put $l(x) := L(x, \lambda, \mu, \eta^G, \eta^H)$. Applying Taylor's formula to (37) yields a vector ξ^k on the connecting line between x^* and x^k such that

$$
\begin{aligned}
f(x^*) &\geqslant l(x^k) \\
&= \underbrace{l(x^*)}_{=f(x^*)} + \underbrace{\nabla l(x^*)^T}_{=\nabla_x L(x^*, \lambda, \mu, \eta^G, \eta^H) = 0} (x^k - x^*) + \frac{1}{2}(x^k - x^*)^T \nabla^2 l(\xi^k)(x^k - x^*) \\
&= f(x^*) + \frac{1}{2}(x^k - x^*)^T \nabla^2_{xx} L(\xi^k, \lambda, \mu, \eta^G, \eta^H)(x^k - x^*),
\end{aligned}
\tag{38}
$$

also exploiting the fact that $(x^*, \lambda, \mu, \eta^G, \eta^H)$ is a strongly stationary point of (1). Dividing by $\|x^* - x^k\|^2$ and letting $k \to \infty$ gives

$$
d^T \nabla^2_{xx} L(x^*, \lambda, \mu, \eta^G, \eta^H) d \leqslant 0,
\tag{39}
$$

which contradicts the assumption (32) of our theorem, because we have $0 \neq d \in \mathcal{C}(x^*)$.

Second, consider the opposite case, that is, assume that there is an index $i \in I_g^+ \cup I_{0-}^+ \cup I_{00}^+ \cup I_{+0}^{0+}$ such that a strict inequality holds in (33). We only consider the case that there exists an index $i \in I_g^+$ such that $\nabla g_i(x^*)^T d < 0$, since the other cases can be treated in the same way. Now, let $s \in I_g^+$ be such that $\nabla g_s(x^*)^T d < 0$. Then it follows from (33) and (35) that

$$\begin{aligned} 0 &\geqslant \nabla f(x^*)^T d \\ &= -\Bigg(\sum_{i\in I_g} \lambda_i \nabla g_i(x^*)^T d + \sum_{j\in J} \mu_j \nabla h_j(x^*)^T d + \sum_{i\in I_{+0}} \eta_i^G \nabla G_i(x^*)^T d \\ &\qquad\qquad - \sum_{i\in I_0} \eta_i^H \nabla H_i(x^*)^T d\Bigg) \\ &\geqslant -\sum_{i\in I_g^+} \lambda_i \nabla g_i(x^*)^T d \geqslant -\lambda_s \nabla g_s(x^*)^T d > 0, \end{aligned}$$

which yields the desired contradiction also in this case. □

Closing this section, we would like to point out that for Example 3.6 the conclusion of Theorem 4.3 as well as the assumptions of Theorem 4.4 are obviously satisfied, since the Hessian of the MPVC-Lagrangian is a positive multiple of the identity at any feasible point and thus in particular positive definite on the whole $\mathbb{R}^n$.

5. Final remarks

This paper contains three main results: First, it shows that the strong stationarity conditions (which are known to be equivalent to the standard KKT conditions) are sufficient optimality conditions for an interesting class of MPVCs. Second, we prove a necessary and a sufficient second-order optimality condition using the same cone in both results. It would be interesting to see whether the MPEC-counterparts of our second-order conditions are actually identical to existing second-order conditions for MPECs, cf. [7], [11], or whether we can use our technique of proof in order to obtain better results also in the context of MPECs.

Acknowledgement. We would like to thank the referee for his useful suggestions for improvement. In particular, the idea of inserting Example 3.6 for the sake of illustrating some of the presented concepts and results is due to him.

References

[1] *W. Achtziger, C. Kanzow*: Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications. Math. Program. To appear.

[2] *M. S. Bazaraa, H. D. Sherali, and C. M. Shetty*: Nonlinear Programming. Theory and Algorithms. 2nd edition. John Wiley & Sons, Hoboken, 1993. zbl

[3] *M. L. Flegel, C. Kanzow*: A direct proof for M-stationarity under MPEC-ACQ for mathematical programs with equilibrium constraints. In: Optimization with Multivalued Mappings: Theory, Applications and Algorithms (S. Dempe, V. Kalashnikov, eds.). Springer-Verlag, New York, 2006, pp. 111–122. zbl

[4] *C. Geiger, C. Kanzow*: Theorie und Numerik restringierter Optimierungsaufgaben. Springer-Verlag, Berlin, 2002. (In German.) zbl

[5] *T. Hoheisel, C. Kanzow*: On the Abadie and Guignard constraint qualification for mathematical programs with vanishing constraints. Optimization. To appear.

[6] *T. Hoheisel, C. Kanzow*: Stationary conditions for mathematical programs with vanishing constraints using weak constraint qualifications. J. Math. Anal. Appl. *337* (2008), 292–310.

[7] *Z.-Q. Luo, J.-S. Pang, and D. Ralph*: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge, 1997. zbl

[8] *O. L. Mangasarian*: Nonlinear Programming. McGraw-Hill Book Company, New York, 1969. zbl

[9] *J. V. Outrata*: Optimality conditions for a class of mathematical programs with equilibrium constraints. Math. Oper. Res. *24* (1999), 627–644. zbl

[10] *J. V. Outrata, M. Kočvara, and J. Zowe*: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Nonconvex Optimization and its Applications. Kluwer, Dordrecht, 1998. zbl

[11] *H. Scheel, S. Scholtes*: Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. Math. Oper. Res. *25* (2000), 1–22. zbl

[12] *S. Scholtes*: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM J. Optim. *11* (2001), 918–936. zbl

[13] *S. Scholtes*: Nonconvex structures in nonlinear programming. Oper. Res. *52* (2004), 368–383.

[14] *J. J. Ye*: Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. J. Math. Anal. Appl. *307* (2005), 350–369. zbl

Authors' address: *T. Hoheisel*, *C. Kanzow*, University of Würzburg, Institute of Mathematics, Am Hubland, 97074 Würzburg, Germany, e-mails: `hoheisel@mathematik.uni-wuerzburg.de`, `kanzow@mathematik.uni-wuerzburg.de`.