# THE THEORY OF 2-REGULARITY FOR MAPPINGS WITH LIPSCHITZIAN DERIVATIVES AND ITS APPLICATIONS TO OPTIMALITY CONDITIONS*

A. F. Izmailov† and M. V. Solodov‡

January 11, 2000 (Revised November 27, 2000 and May 22, 2001)


## ABSTRACT

We study local structure of a nonlinear mapping near points where standard regularity and/or smoothness assumptions need not be satisfied. We introduce a new concept of 2-regularity (a certain kind of second-order regularity) for a once differentiable mapping whose derivative is Lipschitz continuous. Under this 2-regularity condition, we obtain the representation theorem and the covering theorem (i.e., stability with respect to "right-hand side" perturbations) under assumptions which are weaker than those previously employed in the literature for results of this type. These results are further used to derive a constructive description of the tangent cone to a set defined by (2-regular) equality constraints, and optimality conditions for related optimization problems. The class of mappings introduced and studied in the paper appears to be a convenient tool for treating complementarity structures by means of an appropriate equation-based reformulation. Optimality conditions for mathematical programs with (equivalently reformulated) complementarity constraints are also discussed.


**Key words:** $C^{1,1}$-mapping, 2-regularity, covering, tangent cone, optimality conditions, complementarity, mathematical programs with equilibrium constraints.
**AMS subject classifications.** 90C30, 46T20, 47J07, 90C33.

* The first author thanks IMPA, where he was a visiting professor during the completion of this work. The second author is supported in part by CNPq Grant 300734/95-6, by PRONEX–Optimization, and by FAPERJ.

† Computing Center of the Russian Academy of Sciences, Vavilova Str. 40, Moscow, 117967, Russia. Email: `izmaf@ccas.ru`

‡ Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil.
Email : `solodov@impa.br`

# 1 Introduction

In this paper, we consider the problem of local approximation of a nonlinear mapping near a given point in the absence of standard combinations of smoothness and regularity assumptions. We further study consequences of this approximation, in particular those relevant for constructive description of tangent directions to level surfaces, and optimality conditions for problems with feasible regions defined by irregular mappings.

Let $X$ and $Y$ be Banach spaces, and consider a point $\bar{x} \in X$, its neighborhood $V$ in $X$, and a mapping $F : V \to Y$. In what follows, $\mathcal{L}(X, Y)$ stands for the space of continuous linear operators from $X$ to $Y$. For a linear operator $\Lambda \in \mathcal{L}(X, Y)$, $\ker \Lambda = \{x \in X \mid \Lambda x = 0\}$ is its null space, $\operatorname{im} \Lambda = \{y \in Y \mid y = \Lambda x \text{ for some } x \in X\}$ is its image space, and $\Lambda^* : Y^* \to X^*$ is the linear operator adjoint to $\Lambda$, where $X^*$ and $Y^*$ denote the dual spaces of $X$ and $Y$, respectively.

It is well-known that if $F$ is *(first-order) regular* at $\bar{x}$, i.e.,

$$\operatorname{im} F'(\bar{x}) = Y, \tag{1.1}$$

then for many purposes the linear approximation of $F$ is locally adequate. More precisely, we have the following fact, sometimes referred to as the Graves-Lyusternik Theorem.

**Theorem 1.1** *Let $X$ and $Y$ be Banach spaces, $V$ be a neighborhood of $\bar{x}$ in $X$. Suppose $F : V \to Y$ is Fréchet-differentiable on $V$, and the mapping $F' : V \to \mathcal{L}(X, Y)$ is continuous at $\bar{x}$. Suppose further that $F$ is regular at $\bar{x}$.*

*Then there exist a neighborhood $U$ of $0$ in $X$ and a mapping $\varphi : U \to X$ such that $\varphi$ is Fréchet-differentiable at $0$, and*

*(a)* $\varphi(U) \subset V$;

*(b)* $F(\varphi(x)) = F(\bar{x}) + F'(\bar{x})x \quad \forall\, x \in U$;

*(c)* $\varphi(0) = \bar{x}$, $\varphi'(0)$ *is the identity operator on $X$.*

This result is closely related to the implicit function theorem, and it is very useful in a variety of applications, among which is the description of tangent directions. Given a set $D$ in $X$, the *tangent cone* to $D$ at a point $\bar{x} \in D$ is the set

$$\mathcal{T}_D(\bar{x}) := \{h \in X \mid \operatorname{dist}(\bar{x} + th, D) = o(t)\,, t \in \mathbf{R}_+\},$$

where

$$\operatorname{dist}(x, D) = \inf_{\xi \in D} \|x - \xi\|, \quad x \in X.$$

Theorem 1.1 immediately implies the classical description of the tangent subspace to the set

$$F^{-1}(F(\bar{x})) = \{x \in V \mid F(x) = F(\bar{x})\}\,,$$

which is the level surface of $F$ passing through $F(\bar{x})$.

**Theorem 1.2** *Under the assumptions of Theorem 1.1, it holds that*

$$\mathcal{T}_{F^{-1}(F(\bar{x}))}(\bar{x}) = \ker F'(\bar{x}). \tag{1.2}$$

As is well known, constructive description of tangent directions is crucial for obtaining tractable optimality conditions. Let $f : V \to \mathbf{R}$ be a given (objective) function, and consider the equality-constrained problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in D\,, \end{array} \tag{1.3}$$

where

$$D = \{x \in V \mid F(x) = 0\}\,. \tag{1.4}$$

In the regular case, complete description of the tangent cone is given by Theorem 1.2, which in turn makes the primal form of classical optimality conditions for problem (1.3), (1.4) constructive. Furthermore, in that case the dual of the tangent cone is also easily computable, which allows to obtain a tractable form of primal-dual optimality conditions.

The covering property is another well-known by-product of the standard implicit function theorem for the regular case. We say that $F : V \to Y$ has *the covering property at* $\bar{x}$, if for each neighborhood $U \subset V$ of $\bar{x}$ in $X$ the set $F(U)$ is a neighborhood of $F(\bar{x})$ in $Y$.

**Theorem 1.3** *Under the assumptions of Theorem 1.1, $F$ has the covering property at $\bar{x}$. Moreover, there exist a neighborhood $W$ of $F(\bar{x})$ in $Y$ and a constant $M > 0$ such that*

$$\operatorname{dist}(\bar{x},\, F^{-1}(y)) \leq M\|F(\bar{x}) - y\| \quad \forall\, y \in W.$$

The property of covering can be interpreted as a stability result with respect to the "right-hand side" perturbations of the equation $F(x) = y$. Note that in the regular case, we could actually state the stronger property of metric regularity (e.g., see [9, Section 2.3]). However, unlike covering, this property does not carry over to the irregular case.

The focus of this paper is the irregular case when (1.1) need not be satisfied. In that setting, among other things, the representation Theorem 1.1 is not valid. As a consequence, (1.2) does not hold in general, and the classical theory does not yield a constructive description of tangent directions, and hence, also does not provide a constructive form of optimality conditions. Furthermore, other classical results describing local structure of a nonlinear mapping, such as Theorem 1.3, also are not applicable.

For the case of twice Fréchet-differentiable $F$, the questions involved have been addressed within the theory of 2-regularity [20, 1, 16, 21]; see also [27, 10] for some related issues in nonlinear analysis. Equality-constrained optimization problems from the point of view of 2-regularity have been studied in [37, 2, 6], and equality- and inequality-constrained problems in [3, 24, 18]. Even though our development is motivated by some of the cited literature, we note that this is the first paper where the problem data is not assumed to be at least twice Fréchet-differentiable. The subject of this work is specifically local structure of mappings which are once (but not necessarily twice) Fréchet-differentiable at $\bar{x}$, and whose first derivative is Lipschitz continuous on $V$. As we shall show, the difference with the twice

Fréchet-differentiable case is not merely technical. We introduce our concept of 2-regularity in Section 2. In Section 3, we obtain a result on local representation of a once differentiable nonlinear mapping using its "second-order" approximation. A covering theorem is established in Section 4. The tangent cone description is obtained as a simple consequence of the representation theorem. Having in hand the description of the tangent directions, the primal form of optimality conditions for problem (1.3), (1.4) follows immediately. However, the primal-dual optimality conditions require a nontrivial application of the covering theorem.

We emphasize that our extension from twice to once differentiable case is conceptually significant, as local structure of mappings under consideration is considerably more general than that of twice differentiable 2-regular mappings or regular mappings. Under our assumptions, the level surface $F^{-1}(F(\bar{x}))$ can have kinks, and so the tangent cone $\mathcal{T}_{F^{-1}(F(\bar{x}))}(\bar{x})$ need not contain a nontrivial linear subspace, even if the cone itself is nontrivial. In fact, it can consist of rays rather than lines, which is in contrast with the regular and higher-order differentiable 2-regular cases, where the tangent cone is always "two-sided" (i.e., if it contains some nonzero element, it always contains the entire line passing through this and zero elements). Obviously, the latter property imposes rather strong limitations on the kind of mappings that can be handled, and in particular it precludes applications to complementarity structures. One of the goals of this paper is to develop a setting which is appropriate for complementarity structures, where the difficulties arise precisely because of the feasible region having conners and kinks. This issue will be illustrated by Example 2.1. We also refer the reader to the more recent paper [19], where some new *error bounds* have been obtained using the notion of 2-regularity introduced in the present paper, including a new error bound for the nonlinear complementarity problem. Another interesting application of our theory concerns parametric complementarity constraints, which are notorious for being irregular. In Section 6, we show that such constraints can be treated via 2-regularity of some popular equation-based reformulations.

Let us introduce some notation. By $B(x,\, r)$ we denote the open ball with center $x \in X$ and radius $r > 0$. We use the notation $\mathcal{H}(D_1,\, D_2)$ for Hausdorff distance between two sets $D_1,\, D_2 \subset X$:

$$\mathcal{H}(D_1,\, D_2) = \max \left\{ \sup_{x \in D_1} \operatorname{dist}(x,\, D_2),\ \sup_{x \in D_2} \operatorname{dist}(x,\, D_1) \right\}.$$

We shall use balls and distances only in space $X$, so there would be no confusion. By $I_X$, we denote the identity operator in $X$, and similarly in $Y$. For a linear operator $\Lambda \in \mathcal{L}(X,\, Y)$, we denote

$$\|\Lambda^{-1}\| = \sup_{\substack{y \in Y, \\ \|y\|=1}} \operatorname{dist}(0,\, \Lambda^{-1}(y)) .$$

Note that when $\Lambda$ is one-to-one, then $\|\Lambda^{-1}\|$ can be considered as the usual norm of the element $\Lambda^{-1}$ in the space $\mathcal{L}(Y,\, X)$. If $K$ is a cone in $X$, then $K^* = \{x^* \in X^* \mid \langle x^*,\, x\rangle \geq 0 \ \forall\, x \in K\}$ is the (positive) dual cone of $K$. For a set $S$ in $X$, $S^\perp = \{x^* \in X^* \mid \langle x^*,\, x\rangle = 0 \ \forall\, x \in S\}$ denotes its annihilator. Sometimes it is useful to consider the *contingent cone* (Bouligand tangent), which in general can be larger than $\mathcal{T}_D(\bar{x})$, namely

$$\mathcal{B}_D(\bar{x}) := \{h \in X \mid \exists\, \{x^k\} \subset D,\, \exists\, \{\lambda_k\} \subset \mathbf{R}_+ :\, \{x^k\} \to \bar{x},\, \{\lambda_k(x^k - \bar{x})\} \to h\}.$$

Note that in the regular case, the tangent and contingent cones coincide.

We complete this section with the following fact which will be used extensively in our analysis.

**Theorem 1.4 (Set-Valued Contracting Mapping Principle)***[15]*
*Let $X$ be a Banach space, $x^0 \in X$ and $r > 0$. Suppose that $\Phi : B(x^0, r) \to 2^X$ is a (set-valued) mapping such that $\Phi(\xi)$ is nonempty and closed for every $\xi \in B(x^0, r)$. Assume further that there exists a constant $\theta \in (0, 1)$ such that*

$$\mathcal{H}(\Phi(\xi^1), \Phi(\xi^2)) \leq \theta \|\xi^1 - \xi^2\| \quad \forall \xi^1, \xi^2 \in B(x^0, r),$$

$$\operatorname{dist}(x^0, \Phi(x^0)) < (1 - \theta) r.$$

*Then there exists some $\psi \in B(x^0, r)$ such that*

$$\psi \in \Phi(\psi), \quad \|\psi - x^0\| \leq \frac{2}{1 - \theta} \operatorname{dist}(x^0, \Phi(x^0)).$$

## 2 2-Regularity and Tangent Cone

We start with specifying the *basic hypotheses* which will be assumed throughout our analysis:

(H1) $X$ and $Y$ are Banach spaces, $V$ is a neighborhood of a point $\bar{x}$ in $X$.

(H2) $F : V \to Y$ is Fréchet-differentiable on $V$, and the mapping $F' : V \to \mathcal{L}(X, Y)$ is continuous at $\bar{x}$.

(H3) The subspace $Y_1 = \operatorname{im} F'(\bar{x})$ is closed and has a closed complementary subspace $Y_2$ in $Y$.

(H4) The mapping $PF' : V \to \mathcal{L}(X, Y)$ is Lipschitz continuous on $V$ with a constant $L > 0$, where $P$ is the projector in $Y$ onto $Y_2$ parallel to $Y_1$.

Under the assumption (H3), projector $P$ is continuous, and so is $I_Y - P$, which is the projector in $Y$ onto $Y_1$ parallel to $Y_2$. Note also that (H3) holds automatically if $Y$ is finite-dimensional. Another useful observation is that if we replace (H2) by the stronger hypothesis of $F'(\cdot)$ being Lipschitz continuous on $V$, then (H4) holds automatically. However, it is clear that $P$ may have a smoothing affect on $F'(\cdot)$, and so in principle (H4) may hold also when $F'(\cdot)$ is not Lipschitz continuous on $V$. Hence, our setting is actually more general than that of $F$ Fréchet-differentiable on $V$ with Lipschitzian (on $V$) derivative.

We point out that in the regular case, hypotheses (H3) and (H4) are trivially satisfied (this is because in that case, $Y_1 = Y$, $Y_2 = \{0\}$, and $P = 0$). Hence, for the regular case, our smoothness assumptions are no stronger than classical. In this respect, it is an improvement over previous 2-regularity constructions (e.g., [20, 1, 16]), where higher-order differentiability was needed to recover standard results for the regular case.

In what follows, we shall use only the usual directional derivatives of the mapping $PF'(\cdot)$ at the point $\bar{x}$. Given some direction $h \in X$, the directional derivative of $PF'(\cdot)$ at $\bar{x}$ with respect to $h$ is

$$(PF')'(\bar{x};\, h) = \lim_{t \to 0+} \frac{PF'(\bar{x} + th) - PF'(\bar{x})}{t}.$$

Note that this derivative always exists whenever the corresponding directional derivative of the mapping $F'(\cdot)$ exists. However, $PF'(\cdot)$ can be directionally differentiable even if $F'(\cdot)$ is not. Clearly, if $(PF')'(\bar{x};\, h)$ exists then it is an element of $\mathcal{L}(X,\, Y)$, and so we can define the following linear operator whose role will be central in our analysis:

$$\Psi_2(h) \in \mathcal{L}(X,\, Y), \quad \Psi_2(h)\xi := F'(\bar{x})\xi + (PF')'(\bar{x};\, h)\xi.$$

The following definition of 2-regularity is a direct generalization of the definition for twice differentiable mappings (see, e.g., [20, 1, 16]). It would appear natural to use the term "second-order regularity" here. However, this term had already appeared in the literature in a variety of situations different from the context of the present paper. To avoid any confusion, we shall write "2-regularity" throughout.

**Definition 2.1** *The mapping $F$ is referred to as* 2-regular at the point $\bar{x}$ with respect to an element $h \in X$, *if* $\operatorname{im} \Psi_2(h) = Y$.

The following notion is well-defined provided $PF'(\cdot)$ is directionally differentiable at $\bar{x}$ with respect to every direction in $\ker F'(\bar{x})$.

**Definition 2.2** *The mapping $F$ is said to be* 2-regular at the point $\bar{x}$ *if it is 2-regular at this point with respect to every element $h \in T_2 \setminus \{0\}$, where*

$$T_2 = \{h \in \ker F'(\bar{x}) \mid (PF')'(\bar{x};\, h)h = 0\}.$$

The notion of 2-regularity introduced above is a natural extension of known regularity concepts. For example, it contains classical regularity as its special case. Indeed, if (1.1) holds then $Y_1 = Y$, $Y_2 = \{0\}$, $P = 0$, and $\Psi_2(h) = F'(\bar{x})$. Hence, in that case $F$ is 2-regular at $\bar{x}$ with respect to every $h \in X$, and in particular, it is 2-regular at the point $\bar{x}$. It is further clear that standard 2-regularity for twice differentiable $F$ is also a special case within our framework (all directional derivatives involved will be defined via ordinary derivatives; for example, $(PF')'(\bar{x};\, h)\xi = PF''(\bar{x})[h,\, \xi]$ for any $h,\, \xi \in X$).

We are now ready to state one of our principal results, from which the role of 2-regularity becomes immediately clear.

**Theorem 2.1** *Assume that the basic hypotheses (H1)-(H4) are satisfied, and that the mapping $PF'(\cdot)$ has the directional derivative at the point $\bar{x}$ with respect to a direction $h \in X$.*

*Then the following statements hold.*

*(a) If $h \in \mathcal{B}_{F^{-1}(F(\bar{x}))}(\bar{x})$, then $h \in \ker F'(\bar{x})$ and $(PF')'(\bar{x};\, h)h = 0$.*

*(b) If $h \in \ker F'(\bar{x})$ and $(PF')'(\bar{x};\, h)h = 0$, and the mapping $F$ is 2-regular at $\bar{x}$ with respect to $h$, then $h \in \mathcal{T}_{F^{-1}(F(\bar{x}))}(\bar{x})$.*

*In particular, if $PF'(\cdot)$ is directionally differentiable at $\bar{x}$ with respect to every direction in* $\ker F'(\bar{x})$*, and the mapping $F$ is 2-regular at $\bar{x}$, then*

$$\mathcal{T}_{F^{-1}(F(\bar{x}))}(\bar{x}) = \mathcal{B}_{F^{-1}(F(\bar{x}))}(\bar{x}) = T_2.$$

**Remark 2.1** The last assertion of the theorem can be established under somewhat weaker assumption than 2-regularity of $F$ at $\bar{x}$. Specifically, it is sufficient to assume that $F$ is 2-regular at $\bar{x}$ with respect to every element in some dense subset of $T_2$ (i.e., a set whose closure is $T_2$). This follows from the closedness of the tangent cone (to any set at any point) [15].

The proof of the necessary conditions of tangency (part (a) of Theorem 2.1) is quite straightforward, and we shall supply it in the end of this section. The proof of the sufficient conditions (part (b)) is considerably more involved. Proving (b) would require a generalization of the representation theorem, which is indeed of independent interest and importance, and this is the subject of the next section. First, some comments are in order regarding characterization of the tangent cone given by Theorem 2.1.

In the regular case when (1.1) holds, we have that $T_2 = \ker F'(\bar{x})$, and so Theorem 2.1 takes the form of the classical Lyusternik theorem about the tangent subspace (Theorem 1.2). Under the assumptions of twice Fréchet-differentiability of $F$ at $\bar{x}$, Theorem 2.1 was proved in [37] (for the special case when $F'(\bar{x}) = 0$), and in [2] (in complete generality, and even without the assumption that $\operatorname{im} F'(\bar{x})$ has a closed complement, which is part of our hypothesis (H3) above). Some earlier versions of this result under even stronger smoothness assumptions can be found in [27]. For other related material see also [36, 10, 3, 4, 5, 20, 1, 16, 21].

The following example (which we are going to recall in Section 6, in the context of reformulations of complementarity constraints) exhibits that the structure of a level surface described by Theorem 2.1 can be much more general than in the cases of regularity or 2-regularity and higher-order differentiability of $F$ at $\bar{x}$. This testifies to the significance of our generalization. Furthermore, this example shows that even if (at first glance) cone $T_2$ may appear to be a complicated object, it is in fact only as complex or as simple as the problem under consideration. In particular, when $F'(\cdot)$ has some special structure (for example, as in equation-based reformulations of complementarity), then all the objects involved, including $T_2$ itself, are easily explicitly computable.

**Example 2.1** Consider the point $\bar{x} = 0 \in \mathbf{R}^2$ and the mapping

$$F : \mathbf{R}^2 \to \mathbf{R}, \quad F(x) = 2x_1x_2 - (\min\{0,\, x_1 + x_2\})^2 + o(\|x\|^2)\,.$$

It is easy to see that the basic hypotheses are satisfied ($F$ has a Lipschitzian derivative and the setting is finite-dimensional). Observe further that $F'(\bar{x}) = 0$, and so $F$ is not regular at $\bar{x}$. Since $F$ is not twice differentiable at $\bar{x}$, standard theory of 2-regularity also cannot be used. Note that $\ker F'(\bar{x}) = \mathbf{R}^2$, $P$ is the identity operator in $\mathbf{R}$, and by straightforward calculations, for any $h = (h_1,\, h_2) \in \mathbf{R}^2$ we have that

$$(PF')'(\bar{x};\, h) = \lim_{t\to 0+} t^{-1}F'(th) = 2(h_2 - \min\{0,\, h_1 + h_2\},\, h_1 - \min\{0,\, h_1 + h_2\}).$$

It can be easily checked that $(PF')'(\bar{x};\, h) = (0,0)$ if, and only if, $h_1 = h_2 = 0$. It follows that $F$ is 2-regular with respect to any $h \in \mathbf{R}^2 \setminus \{0\}$, and in particular, $F$ is 2-regular at $\bar{x}$. By obvious calculations, we further have that

$$\begin{aligned} \mathcal{T}_{F^{-1}(F(\bar{x}))}(\bar{x}) &= \mathcal{B}_{F^{-1}(F(\bar{x}))}(\bar{x}) \\ &= T_2 \\ &= \{h \in \mathbf{R}^2 \mid h_1 \geq 0,\, h_2 \geq 0,\, h_1 h_2 = 0\} \\ &= \{h \in \mathbf{R}^2 \mid \min\{h_1,\, h_2\} = 0\}. \end{aligned} \tag{2.1}$$

Observe that the level surface here has a kink, and the tangent cone consists of two rays. Recall that in the regular case the tangent cone is $\ker F'(\bar{x})$, so it is a subspace. In the twice-differentiable 2-regular case, the tangent cone takes the form $\{h \in \ker F'(\bar{x}) \mid PF''(\bar{x})[h]^2 = 0\}$, a "two-sided" cone which necessarily contains entire lines passing through the zero point (assuming the cone is larger than the singleton $\{0\}$). In particular, neither of those two cases allows (one-sided) rays or kinks. On the other hand, both are special cases in our framework. An important conclusion is that 2-regularity for mappings with Lipschitzian derivatives is conceptually much more general than the regular case, or twice-differentiable 2-regular case. Furthermore, observing the structure of $F$ in this example, we are justified to state that the introduced concept of 2-regularity is adequate for handling complementarity structures, as in (2.1).

We next prove the necessity part of Theorem 2.1. The sufficiency part will be established later, as a consequence of the representation theorem in Section 3.

**Proof.** [of Theorem 2.1(a).] Take any $h \in \mathcal{B}_{F^{-1}(F(\bar{x}))}(\bar{x}) \setminus \{0\}$. The fact that $h \in \ker F'(\bar{x})$ is standard, so we shall prove only the relation $(PF')'(\bar{x};\, h)h = 0$.

By the definition of the contingent cone, there exist two sequences $\{x^k\} \subset V$ and $\{\lambda_k\} \subset \mathbf{R}_+$ such that $F(x^k) = F(\bar{x})$ for all $k$, $\{x^k\} \to \bar{x}$, $\{\lambda_k\} \to \infty$, and $\{\lambda_k(x^k - \bar{x})\} \to h$. Using the Newton–Leibniz formula for $PF(\cdot)$ (for $k$ sufficiently large), we then obtain that

$$\begin{aligned} 0 &= \lambda_k^2 \|P(F(x^k) - F(\bar{x}))\| \\ &= \lambda_k^2 \left\| \int_0^1 PF'(\bar{x} + \tau(x^k - \bar{x}))(x^k - \bar{x})\, d\tau \right\| \\ &\geq \lambda_k^2 \left\| \int_0^1 (PF')'(\bar{x};\, \tau\lambda_k^{-1}h)(x^k - \bar{x})\, d\tau \right\| \\ &\quad -\lambda_k^2 \int_0^1 \|(P(F'(\bar{x} + \tau\lambda_k^{-1}h) - F'(\bar{x})) - (PF')'(\bar{x};\, \tau\lambda_k^{-1}h))(x^k - \bar{x})\|\, d\tau \\ &\quad -\lambda_k^2 \int_0^1 \|P(F'(\bar{x} + \tau(x^k - \bar{x})) - F'(\bar{x} + \tau\lambda_k^{-1}h))(x^k - \bar{x})\|\, d\tau \\ &\geq \frac{1}{2}\|(PF')'(\bar{x};\, h)(\lambda_k(x^k - \bar{x}))\| + \lambda_k o(\lambda_k^{-1})\|\lambda_k(x^k - \bar{x})\| \\ &\quad -\frac{L}{2}\|\lambda_k(x^k - \bar{x}) - h\|\|\lambda_k(x^k - \bar{x})\|, \end{aligned}$$

where the triangle inequality and the fact that $PF'(\bar{x})\xi = 0$ for any $\xi \in X$ were used in the first inequality, and Lipschitz continuity of $PF'(\cdot)$ and the fact that the mapping $(PF')'(\bar{x};\, \cdot)$ is positively homogeneous were used in the second inequality. Passing onto the limit as $k \to \infty$, we conclude that $(PF')'(\bar{x};\, h)h = 0$. ∎

# 3 The Representation Theorem for 2-Regular Mappings

In this section, we obtain a generalization of Theorem 1.1 for the irregular case under consideration.

Recall that the mapping $PF'(\cdot)$ is referred to as *$B$-differentiable at the point $\bar{x}$ with respect to cone $K$* in $X$ if it has a directional derivative at $\bar{x}$ with respect to every direction $h \in K$, and

$$PF'(\bar{x}+h) = PF'(\bar{x}) + (PF')'(\bar{x};\, h) + o(\|h\|)\,, \quad h \in (V - \bar{x}) \cap K.$$

The notion of $B$-differentiability (with respect to $K = X$) was introduced in [32]. Note that in the finite-dimensional setting, a mapping which is Lipschitzian in a neighborhood of a point (recall hypothesis (H4)) is $B$-differentiable at this point if, and only if, it is directionally differentiable at this point with respect to every direction [34].

In the sequel, we shall need the following two technical lemmas.

**Lemma 3.1** *Assume that the basic hypotheses (H1)-(H4) are satisfied. Suppose further that the mapping $PF'(\cdot)$ is $B$-differentiable at the point $\bar{x}$ with respect to a cone $K$ in $X$.*

*Then for any $\varepsilon > 0$ there exist two constants $\delta > 0$ and $\gamma > 0$ such that $B(\bar{x},\, (1+\gamma)\delta) \subset V$, and for all $x \in B(0,\, \delta) \cap (\bar{x} + K)$ and $\xi^1,\, \xi^2 \in B(0,\, \gamma\|x\|)$, it holds that*

$$\|(I_Y - P)(F(\bar{x}+x+\xi^1) - F(\bar{x}+x+\xi^2)) - F'(\bar{x})(\xi^1 - \xi^2)\| \le \varepsilon\|\xi^1 - \xi^2\|,$$

$$\|P(F(\bar{x}+x+\xi^1) - F(\bar{x}+x+\xi^2)) - (PF')'(\bar{x};\, x)(\xi^1 - \xi^2)\| \le \varepsilon\|x\|\|\xi^1 - \xi^2\|.$$

**Proof.** The argument is completely straightforward, and similar to that in Theorem 2.1(a). One has to use the Newton–Leibniz formula for $(I_Y - P)F(\cdot)$ and $PF(\cdot)$, the equality $PF'(\bar{x}) = 0$, the triangle inequality, and $B$-differentiability of $F$ at $\bar{x}$ with respect to $K$. ∎

**Lemma 3.2** *Under the assumptions of Lemma 3.1, for any $\varepsilon > 0$ there exists a constant $\delta > 0$ such that $B(\bar{x},\, \delta) \subset V$, and for all $x \in B(0,\, \delta) \cap (\bar{x} + K)$, it holds that*

$$\|(I_Y - P)(F(\bar{x}+x) - F(\bar{x})) - F'(\bar{x})x\| \le \varepsilon\|x\|,$$

$$\left\|P(F(\bar{x}+x) - F(\bar{x})) - \frac{1}{2}(PF')'(\bar{x};\, x)x\right\| \le \varepsilon\|x\|^2.$$

**Proof.** The argument is again straightforward and similar to Lemma 3.1 and Theorem 2.1(a). One has to use additionally the fact that $(PF')'(\bar{x};\, \cdot) : K \to \mathcal{L}(X,\, Y)$ is positively homogeneous. ∎

**Remark 3.1** It is easy to see that if in Lemma 3.2 we assume that $x \neq 0$, then the constant $\delta > 0$ can be chosen in such a way that the two inequalities can be strengthened to be strict.

In the sequel, we shall use the following regularity assumption:

$$\sup\{\|(\Psi_2(h))^{-1}\| \mid h \in K,\ \|h\| = 1\} < \infty. \tag{3.1}$$

If $F$ is regular at $\bar{x}$, then (3.1) is satisfied for $K = X$. This follows from the Banach theorem, according to which a surjective continuous linear operator from one Banach space to another has a bounded (set-valued) right inverse in a sense of the "norm" defined in Section 1. Furthermore, if $F$ is 2-regular at $\bar{x}$ with respect to each $h \in K \setminus \{0\}$ and $X$ is finite-dimensional, then again (3.1) holds. To see this, one has to recall that the unit sphere in a finite-dimensional space is compact, and any small linear perturbation of a surjective linear operator is still surjective.

**Theorem 3.1** *Assume that the basic hypotheses (H1)-(H4) are satisfied. Suppose further that the mapping $PF'(\cdot)$ is B-differentiable at the point $\bar{x}$ with respect to a cone $K$ in $X$, and (3.1) is satisfied.*

*Then there exist a neighborhood $U$ of $0$ in $X$ and a mapping $\varphi : U \to X$ such that $\varphi$ is Fréchet-differentiable at $0$, and*

- *(a)* $\varphi(U) \subset V$;
- *(b)* $F(\varphi(x)) = F(\bar{x}) + F'(\bar{x})x + \frac{1}{2}(PF')'(\bar{x};\, x)x \quad \forall\, x \in U \cap K$;
- *(c)* $\varphi(0) = \bar{x}$, $\varphi'(0) = I_X$.

**Proof.** Let $\rho > 0$ be such that $B(\bar{x},\, 2\rho) \subset V$. For each $x \in B(0,\, \rho)$ define the set-valued mapping

$$\Phi_x : B(0,\, \rho) \to 2^X,$$
$$\Phi_x(\xi) = \xi - (\Psi_2(x))^{-1}(F(\bar{x} + x + \xi) - F(\bar{x}) - F'(\bar{x})x - \tfrac{1}{2}(PF')'(\bar{x};\, x)x),$$

so that

$$\eta \in \Phi_x(\xi) \quad \Leftrightarrow \quad \Psi_2(x)\eta = \Psi_2(x)\xi - F(\bar{x} + x + \xi) + F(\bar{x}) + F'(\bar{x})x + \frac{1}{2}(PF')'(\bar{x};\, x)x.$$

Note that for each fixed $\xi \in B(0,\, \rho)$, $\Phi_x(\xi)$ is either an affine set parallel to the subspace $\ker \Psi_2(x)$, or the empty set.

We shall prove our assertion by applying the set-valued contracting mapping principle (Theorem 1.4) to the mapping defined above.

It is easy to see that $\Phi_x$ is well-defined for all $x \in B(0,\, \rho) \cap K$. Moreover, by (3.1),

$$\Phi_x(\xi) \neq \emptyset \quad \forall\, x \in (B(0,\, \rho) \cap K) \setminus \{0\},\ \forall\, \xi \in B(0,\, \rho). \tag{3.2}$$

Fix an arbitrary $\theta \in (0, 1)$, and denote the constant in the left-hand side of (3.1) by $C$, $C > 0$. Applying Lemma 3.1, we conclude that there exist two constants $\delta \in (0, \rho]$ and $\gamma \in (0, \rho/\delta]$ such that for all $x \in B(0, \delta) \cap K$ and for all $\xi^1, \xi^2 \in B(0, \gamma\|x\|)$ the following relations hold:

$$\|(I_Y - P)(F(\bar{x} + x + \xi^1) - F(\bar{x} + x + \xi^2)) - F'(\bar{x})(\xi^1 - \xi^2)\| \leq \frac{\theta}{2C}\|\xi^1 - \xi^2\|, \tag{3.3}$$

$$\|P(F(\bar{x} + x + \xi^1) - F(\bar{x} + x + \xi^2)) - (PF')'(\bar{x};\, x)(\xi^1 - \xi^2)\| \leq \frac{\theta}{2C}\|x\|\|\xi^1 - \xi^2\|. \tag{3.4}$$

Next, define $r(x) = \gamma\|x\|$, $x \in B(0, \delta)$. According to Lemma 3.2 (recall Remark 3.1), there exists a neighborhood $U \subset B(0, \delta)$ of 0 in $X$ such that for all $x \in (U \cap K) \setminus \{0\}$ we have

$$\|(I_Y - P)(F(\bar{x} + x) - F(\bar{x})) - F'(\bar{x})x\| < \frac{1-\theta}{2C} r(x), \tag{3.5}$$

$$\left\| P(F(\bar{x} + x) - F(\bar{x})) - \frac{1}{2}(PF')'(\bar{x};\, x)x \right\| < \frac{1-\theta}{2C}\|x\| r(x). \tag{3.6}$$

To apply the contracting mapping principle, we have to estimate the (Hausdorff) distance between $\Phi_x(\xi^1)$ and $\Phi_x(\xi^2)$ for each $x \in (U \cap K) \setminus \{0\}$ and $\xi^1, \xi^2 \in B(0, r(x))$. Note that for any $x^i \in \Phi_x(\xi^i)$, $i = 1, 2$,

$$\Psi_2(x)(x^1 - x^2) = \Psi_2(x)(\xi^1 - \xi^2) - F(\bar{x} + x + \xi^1) + F(\bar{x} + x + \xi^2). \tag{3.7}$$

Observe further that for any $\xi \in X$ and $t > 0$,

$$\Psi_2(tx)\xi = F'(\bar{x})\xi + (PF')'(\bar{x}; tx)\xi = F'(\bar{x})\xi + t(PF')'(\bar{x}; x)\xi,$$

in particular, multiplication of $x$ does not affect the component in $Y_1 = \operatorname{im} F'(\bar{x})$, while the component in $Y_2$ is positively homogenuous. Using this observation and (3.7), we have that

$$\begin{aligned}
\Psi_2(\|x\|^{-1}x)(x^1 - x^2) \quad = \quad & \Psi_2(\|x\|^{-1}x)(\xi^1 - \xi^2) \\
& -(I_Y - P)(F(\bar{x} + x + \xi^1) - F(\bar{x} + x + \xi^2)) \\
& -\|x\|^{-1}P(F(\bar{x} + x + \xi^1) - F(\bar{x} + x + \xi^2)).
\end{aligned}$$

Since $\Phi_x(\xi^1)$ and $\Phi_x(\xi^2)$ are two parallel affine sets, we have that

$$\begin{aligned}
\mathcal{H}(\Phi_x(\xi^1),\, \Phi_x(\xi^2)) &= \inf_{\substack{x^i \in \Phi_x(\xi^i) \\ i=1,2}} \|x^1 - x^2\| \\
&= \inf\Big\{ \|\xi\| \mid \xi \in X,\, \Psi_2(\|x\|^{-1}x)\xi = \Psi_2(\|x\|^{-1}x)(\xi^1 - \xi^2) \\
&\qquad -(I_Y - P)(F(\bar{x} + x + \xi^1) - F(\bar{x} + x + \xi^2)) \\
&\qquad -\|x\|^{-1}P(F(\bar{x} + x + \xi^1) - F(\bar{x} + x + \xi^2)) \Big\} \\
&\leq C\Big( \|(I_Y - P)(F(\bar{x} + x + \xi^1) - F(\bar{x} + x + \xi^2)) - F'(\bar{x})(\xi^1 - \xi^2)\|
\end{aligned}$$

$$+\|x\|^{-1}\|P(F(\bar{x}+x+\xi^1)-F(\bar{x}+x+\xi^2))-(PF')'(\bar{x};\,x)(\xi^1-\xi^2)\|\Big)$$

$$\leq \theta\|\xi^1-\xi^2\|, \tag{3.8}$$

where the second inequality follows from (3.1), and the last inequality follows from (3.3) and (3.4).

Similarly, from (3.5) and (3.6), it follows that for all $x \in (U \cap K) \setminus \{0\}$

$$\begin{aligned}
\mathrm{dist}(0,\,\Phi_x(0)) &= \inf\Bigg\{\|\xi\| \,\Bigg|\, \xi \in X\,,\, \Psi_2(\|x\|^{-1}x)\xi = (I_Y-P)(F(\bar{x}+x)-F(\bar{x}))-F'(\bar{x})x \\
&\qquad +\|x\|^{-1}\left(P(F(\bar{x}+x)-F(\bar{x}))-\frac{1}{2}(PF')'(\bar{x};\,x)x\right)\Bigg\} \\
&\leq C\Bigg(\|(I_Y-P)(F(\bar{x}+x)-F(\bar{x}))-F'(\bar{x})x\| \\
&\qquad +\|x\|^{-1}\left\|P(F(\bar{x}+x)-F(\bar{x}))-\frac{1}{2}(PF')'(\bar{x};\,x)x\right\|\Bigg) \\
&< (1-\theta)r(x).
\end{aligned} \tag{3.9}$$

Relations (3.8) and (3.9) (also taking into account (3.2)), imply that for all $x \in (U \cap K) \setminus \{0\}$ the mapping $\Phi_x$ satisfies in $B(0,\,r(x))$ all the assumptions of Theorem 1.4. Therefore, for each $x \in (U \cap K) \setminus \{0\}$, $\Phi_x$ has a fixed point in $B(0,\,r(x))$. Specifically, for each such $x$, there exists an element $\psi(x) \in B(0,\,r(x))$ such that

$$\psi(x) \in \Phi_x(\psi(x)), \quad \|\psi(x)\| \leq \frac{2}{1-\theta}\,\mathrm{dist}(0,\,\Phi_x(0)).$$

Hence, we have that for each $x \in (U \cap K) \setminus \{0\}$,

$$F(\bar{x}+x+\psi(x)) = F(\bar{x})+F'(\bar{x})x+\frac{1}{2}(PF')'(\bar{x};\,x)x, \tag{3.10}$$

and, using the second inequality in (3.9) and Lemma 3.2,

$$\|\psi(x)\| \leq \frac{2}{1-\theta}\,\mathrm{dist}(0,\,\Phi_x(0)) = o(\|x\|). \tag{3.11}$$

Set $\psi(0) := 0$, and define the mapping

$$\varphi : U \to X, \quad \varphi(x) = \begin{cases} \bar{x}+x+\psi(x), & \text{if } x \in U \cap K, \\ \bar{x}+x, & \text{if } x \in U \setminus K. \end{cases}$$

With this definition, (3.10) and (3.11) imply that $\varphi$ possesses all the properties asserted in the theorem. ■

If the mapping $F$ is regular at $\bar{x}$, then $P = 0$, condition (3.1) is satisfied for $K = X$, and Theorem 3.1 reduces to the representation theorem given in Section 1. For the case of $F$ twice Fréchet-differentiable at $\bar{x}$, the representation theorem analogous to Theorem 3.1 was obtained in [16].

Theorem 3.1 immediately implies the sufficient conditions of tangency stated in Section 2.

**Proof.** [Theorem 2.1 (b).] Observe that by the Banach theorem, if $F$ is 2-regular at $\bar{x}$ with respect to $h$, then $\|(\Psi_2(h))^{-1}\| < \infty$. The result now follows applying Theorem 3.1 to the cone $K = \{\xi \in X \mid \xi = th,\ t \in \mathbf{R}_+\}$, where $h$ is the element under consideration. ■

# 4 Covering Property

Under the assumptions of Theorem 2.1, we immediately obtain *primal* first-order necessary optimality conditions for problem (1.3), (1.4), which are constructive in the sense that we have a constructive description of the tangent cone in terms of the problem data. These conditions will be stated in the next section. To obtain a primal-dual form of necessary optimality conditions, a covering theorem would be needed.

Covering can be used for deriving optimality conditions via the following considerations (in [12] this approach was called *simultaneous covering*). For problem (1.3), (1.4), consider the mapping

$$\Phi : V \to \mathbf{R} \times Y, \quad \Phi(x) = (f(x),\, F(x)). \tag{4.1}$$

It is clear that if $\bar{x}$ is a local minimizer for problem (1.3), (1.4), then the mapping $\Phi$ does not have the covering property at $\bar{x}$. Hence, $\Phi$ cannot fall under the assumptions of any covering theorem applied at $\bar{x}$. This fact can then be used to derive a primal-dual form of optimality conditions. We postpone further details until Section 5.

We emphasize that the property of covering is a stability result with respect to the "right-hand side" perturbations of the equation $F(x) = y$, and so it is of independent interest and importance, i.e., regardless of its applications. For a thorough study of covering theorems and their applications, we refer the reader to [12]. Below, we prove a covering theorem (with a quantitive estimate, see (4.2)!) under assumptions which are weaker than those previously used in the literature.

**Theorem 4.1** *Assume that the basic hypotheses (H1)-(H4) hold, and that the mapping $PF'(\cdot)$ has the directional derivative at the point $\bar{x}$ with respect to a direction $h$ satisfying $h \in \ker F'(\bar{x})$ and $(PF')'(\bar{x};\, h)h = 0$. Suppose further that the mapping $F$ is 2-regular at $\bar{x}$ with respect to $h$.*

*Then $F$ has the covering property at $\bar{x}$, and there exist a neighborhood $W$ of $F(\bar{x})$ in $Y$ and a constant $M > 0$ such that*

$$\mathrm{dist}(\bar{x},\, F^{-1}(y)) \le M \max\{\|(I_Y - P)(F(\bar{x}) - y)\|,\, \|P(F(\bar{x}) - y)\|^{1/2}\} \quad \forall y \in W. \tag{4.2}$$

**Proof.** The approach we use here again relies on the set-valued contracting mapping principle.

Set $C = \|(\Psi_2(h))^{-1}\|$, and fix an arbitrary constant $\theta \in (0,\, 1)$. Lemma 3.1 implies the existence of $\delta > 0$ and $\gamma > 0$ such that

$$B(\bar{x},\, (\|h\| + \gamma)\delta) \subset V,$$

and for all $t \in (0, \delta)$, we have for all $\xi^1, \xi^2 \in B(0, \gamma t)$ that

$$\|(I_Y - P)(F(\bar{x} + th + \xi^1) - F(\bar{x} + th + \xi^2)) - F'(\bar{x})(\xi^1 - \xi^2)\| \leq \frac{\theta}{2C}\|\xi^1 - \xi^2\|, \qquad (4.3)$$

$$\|P(F(\bar{x} + th + \xi^1) - F(\bar{x} + th + \xi^2)) - (PF')'(\bar{x};\, th)(\xi^1 - \xi^2)\| \leq \frac{\theta}{2C}t\|\xi^1 - \xi^2\|. \qquad (4.4)$$

Next, define the function

$$s : Y \to \mathbf{R}_+, \quad s(y) = \max\{c\|(I_Y - P)(F(\bar{x}) - y)\|,\, (c\|P(F(\bar{x}) - y)\|)^{1/2}\}, \qquad (4.5)$$

where $c$ is any constant such that

$$c > \frac{4C}{(1-\theta)\gamma}. \qquad (4.6)$$

Note that $y = F(\bar{x})$ is the unique root of $s(\cdot)$, and $s(y) \to 0$ as $y \to F(\bar{x})$.

Next, choose a neighborhood $W'$ of $F(\bar{x})$ in $Y$ such that $s(y) < \delta$ for all $y \in W'$. Then for all $y \in W'$ the set-valued mapping

$$\Phi_y : B(0, \gamma\delta) \to 2^X, \quad \Phi_y(\xi) = \xi - (\Psi_2(s(y)h))^{-1}(F(\bar{x} + s(y)h + \xi) - y),$$

is well-defined and, by 2-regularity of $F$ at $\bar{x}$ with respect to $h$,

$$\Phi_y(\xi) \neq \emptyset \quad \forall\, y \in W' \setminus \{F(\bar{x})\}, \forall\, \xi \in B(0, \gamma\delta). \qquad (4.7)$$

According to Lemma 3.2, there exists a neighborhood $W \subset W'$ of $F(\bar{x})$ in $Y$ such that for all $y \in W$ we have

$$\begin{aligned}\|(I_Y - P)(F(\bar{x} + s(y)h) - F(\bar{x}))\| &= \|(I_Y - P)(F(\bar{x} + s(y)h) - F(\bar{x})) - F'(\bar{x})(s(y)h)\| \\ &\leq \frac{1-\theta}{4C}\gamma s(y), \qquad (4.8)\end{aligned}$$

$$\begin{aligned}\|P(F(\bar{x} + s(y)h) - F(\bar{x}))\| &= \|P(F(\bar{x} + s(y)h) - F(\bar{x})) - \frac{1}{2}(PF')'(\bar{x};\, s(y)h)(s(y)h)\| \\ &\leq \frac{1-\theta}{4C}\gamma(s(y))^2. \qquad (4.9)\end{aligned}$$

Define $r(y) = \gamma s(y)$, $y \in W$. Applying (4.3) and (4.4), we obtain that for all $y \in W \setminus \{F(\bar{x})\}$ and all $\xi^1, \xi^2 \in B(0, r(y))$, it holds that

$$\begin{aligned}\mathcal{H}(\Phi_y(\xi^1), \Phi_y(\xi^2)) &= \inf_{\substack{x^i \in \Phi_x(\xi^i) \\ i=1,2}} \|x^1 - x^2\| \\ &= \inf\Big\{\|\xi\| \mid \xi \in X\,, \Psi_2(h)\xi = \Psi_2(h)(\xi^1 - \xi^2) \\ &\qquad -(I_Y - P)(F(\bar{x} + s(y)h + \xi^1) - F(\bar{x} + s(y)h + \xi^2)) \\ &\qquad -(s(y))^{-1}P(F(\bar{x} + s(y)h + \xi^1) - F(\bar{x} + s(y)h + \xi^2))\Big\}\end{aligned}$$

$$\leq C\Big(\|(I_Y - P)(F(\bar{x} + s(y)h + \xi^1) - F(\bar{x} + s(y)h + \xi^2)) - F'(\bar{x})(\xi^1 - \xi^2)\|$$
$$+(s(y))^{-1}\|P(F(\bar{x} + s(y)h + \xi^1) - F(\bar{x} + s(y)h + \xi^2)) - (PF')'(\bar{x}; s(y)h)(\xi^1 - \xi^2)\|\Big)$$
$$\leq \theta\|\xi^1 - \xi^2\| . \tag{4.10}$$

Moreover, from (4.5), (4.6), (4.8) and (4.9), it follows that for all $y \in W \setminus \{F(\bar{x})\}$

$$\begin{aligned}
\mathrm{dist}(0,\, \Phi_x(0)) &= \inf\Big\{\|\xi\| \mid \xi \in X\,, \Psi_2(h)\xi = (I_Y - P)(F(\bar{x} + s(y)h) - y) \\
&\qquad\qquad + (s(y))^{-1}P(F(\bar{x} + s(y)h) - y)\Big\} \\
&\leq C\Big(\|(I_Y - P)(F(\bar{x} + s(y)h) - y)\| + (s(y))^{-1}\|P(F(\bar{x} + s(y)h) - y)\|\Big) \\
&\leq C\Big(\|(I_Y - P)(F(\bar{x} + s(y)h) - F(\bar{x}))\| \\
&\qquad + (s(y))^{-1}\|P(F(\bar{x} + s(y)h) - F(\bar{x}))\|\Big) \\
&\qquad + C\Big(\|(I_Y - P)(F(\bar{x}) - y)\| + (s(y))^{-1}\|P(F(\bar{x}) - y)\|\Big) \\
&\leq \frac{1-\theta}{2}\gamma s(y) + \frac{2C}{c}s(y) \\
&< (1-\theta)r(y).
\end{aligned} \tag{4.11}$$

Relations (4.10) and (4.11), together with (4.7), imply that for all $y \in W \setminus \{F(\bar{x})\}$ the assumptions of Theorem 1.4 are satisfied for the mapping $\Phi_y$ in the ball $B(0,\, r(y))$. Hence, for each $y \in W \setminus \{F(\bar{x})\}$ there exists an element $\psi(y) \in B(0,\, r(y))$ such that

$$\psi(y) \in \Phi_y(\psi(y))\,, \quad \|\psi(y)\| \leq \frac{2}{1-\theta}\,\mathrm{dist}(0,\, \Phi_y(0)). \tag{4.12}$$

In particular, the inclusion above implies that

$$F(\bar{x} + s(y)h + \psi(y)) = y\,,$$

so that

$$\mathrm{dist}(\bar{x},\, F^{-1}(y)) \leq \|s(y)h + \psi(y)\|.$$

Furthermore, by (4.12) and the next to last inequality in (4.11),

$$\|\psi(y)\| \leq \left(\gamma + \frac{4C}{c(1-\theta)}\right) s(y). \tag{4.13}$$

Hence,

$$\begin{aligned}
\mathrm{dist}(\bar{x},\, F^{-1}(y)) &\leq s(y)\|h\| + \|\psi(y)\| \\
&\leq \left(\|h\| + \gamma + \frac{4C}{c(1-\theta)}\right) s(y),
\end{aligned}$$

from which the bound (4.2) follows immediately, using (4.5) and setting $M = (\|h\| + \gamma + 4C(c(1-\theta))^{-1}) \max\{c,\, c^{1/2}\}$. ∎

Theorem 4.1 establishes sufficient conditions for $F$ to have the covering property at the point $\bar{x}$ under consideration. If $F$ is regular at $\bar{x}$, then one can apply Theorem 4.1 with any $h \in \ker F'(\bar{x})$, even $h = 0$, and obtain Theorem 1.3. For the case of $F$ twice Fréchet-differentiable at $\bar{x}$, the covering property was obtained in [4]. But we emphasize that Theorem 4.1 is certainly completely meaningful beyond those two cases, see Example 4.1 below. Note also that relation (4.2) can be simplified as follows:

$$\operatorname{dist}(\bar{x}, F^{-1}(y)) \leq M\|F(\bar{x}) - y\|^{1/2} \quad \forall\, y \in W\,, \tag{4.14}$$

where the neighborhood $W$ of $F(\bar{x})$ in $Y$ and the constant $M > 0$ are possibly different from those in (4.2). In that form, Theorem 4.1 does not require any knowledge about the structure of $F'(\bar{x})$, apart from the regularity hypothesis itself (which, once again, is weaker than assumptions for comparable results in the literature). However, the simpler estimate (4.14) can be less accurate than (4.2).

**Example 4.1** Let $\bar{x}$ and $F$ be as defined in Example 2.1. As is easy to see, the assumptions of Theorem 4.1 are satisfied for any $h \in T_2 \setminus \{0\}$. Hence, for some $M > 0$ and for each $y \in \mathbf{R}$ sufficiently close to zero, there exists $x(y) \in \mathbf{R}^2$ such that

$$F(x(y)) = y, \quad \|x(y)\| \leq M|y|^{1/2}.$$

Note that this estimate cannot be improved. Indeed, let the term $o(\|x\|^2)$ appearing in the definition of $F$ be equal to zero. Then it can be shown (e.g., using Lagrange principle) that the closest to zero solution of the equation $F(x) = y$ is $x(y) = (\theta(y), \theta(y))$, where

$$\theta(y) = \frac{1}{\sqrt{2}} \begin{cases} \sqrt{y}, & \text{if } y \geq 0, \\ -\sqrt{-y}, & \text{if } y < 0. \end{cases}$$

# 5 Optimality Conditions

We now turn our attention to the optimization problem (1.3), (1.4). The primal form of optimality conditions follows immediately from Theorem 2.1 and well-known classical results.

**Theorem 5.1** *Assume that the basic hypotheses (H1)-(H4) are satisfied, and that the mapping $PF'(\cdot)$ has the directional derivative at the point $\bar{x} \in D$ with respect to a direction $h \in X$ such that $h \in \ker F'(\bar{x})$ and $(PF')'(\bar{x}; h)h = 0$. Suppose that the mapping $F$ is 2-regular at $\bar{x}$ with respect to $h$. Suppose further that the function $f : V \to \mathbf{R}$ is Fréchet-differentiable at $\bar{x}$.*

*If $\bar{x}$ is a local solution of the problem (1.3), (1.4), then*

$$\langle f'(\bar{x}),\, h\rangle \geq 0. \tag{5.1}$$

If, in addition to the assumptions of Theorem 5.1, $X$ is finite-dimensional, the mapping $PF'(\cdot)$ has the directional derivative at the point $\bar{x} \in D$ with respect to every direction in $\ker F'(\bar{x})$, and

$$\langle f'(\bar{x}),\, h\rangle > 0 \quad \forall\, h \in T_2 \setminus \{0\}\,, \tag{5.2}$$

then $\bar{x}$ is a strict local solution of (1.3), (1.4). Note that in contrast with the twice Fréchet-differentiable case, or the regular case, cone $T_2$ does not necessarily contain nontrivial linear subspaces, even if it is nontrivial itself (recall Example 2.1 above). Hence, the strict inequality in (5.2) is meaningful.

Naturally, further development of Theorem 5.1 should be the primal-dual form of necessary optimality conditions. Suppose, for the moment, that $PF'(\cdot)$ has the directional derivative at $\bar{x} \in D$ with respect to every direction in $\ker F'(\bar{x})$, and that $F$ is 2-regular at $\bar{x}$. Then, dualizing the statement of Theorem 5.1, it follows that

$$f'(\bar{x}) \in T_2^*.$$

In full generality, $T_2$ can have a rather complicated structure, since it is defined by nonlinear constraints and is nonconvex, in general. Without some further assumptions, it seems impossible to compute $T_2^*$ explicitly. However, without computing the full characterization of $T_2^*$, we are able to obtain a tractable form of primal-dual optimality conditions under the additional assumption that

$$\exists\, \bar{h} \in T_2 \text{ such that } \langle f'(\bar{x}),\, \bar{h}\rangle = 0\,, \tag{5.3}$$

where $F$ is 2-regular at $\bar{x}$ with respect to this $\bar{h}$. Assumptions similar to (5.3) are quite common in the literature, e.g., [7, 8, 31]. Note that analysis of such "critical" directions is of special importance in view of the violation of the sufficient optimality condition (5.2).

**Theorem 5.2** *Suppose that the basic hypotheses (H1)-(H4) are satisfied. Let the function $f : V \to \mathbf{R}$ be Fréchet-differentiable at $\bar{x}$. Assume that the mapping $PF'(\cdot)$ has the directional derivative at the point $\bar{x} \in D$ with respect to a direction $\bar{h} \in X$ satisfying (5.3), and that the mapping $F$ is 2-regular at $\bar{x}$ with respect to $\bar{h}$.*

*If $\bar{x}$ is a local solution of problem (1.3), (1.4), then there exists a functional $y^* = y^*(\bar{h}) \in Y^*$ such that*

$$f'(\bar{x}) = (\Psi_2(\bar{h}))^* y^*. \tag{5.4}$$

**Proof.** If $\bar{x}$ is a local solution of the problem (1.3), (1.4), then the mapping $\Phi : V \to \mathbf{R} \times Y$ defined in (4.1) does not have the covering property at $\bar{x}$. Hence, $\Phi$ at $\bar{x}$ does not fall under the assumptions of any covering theorem, and, in particular, under the assumptions of Theorem 4.1.

Consider first the case where $f'(\bar{x}) \in (\ker F'(\bar{x}))^{\perp} = \mathrm{im}(F'(\bar{x}))^*$. [In the latter relation we have used the lemma on annihilator of a null-space of a continuous linear operator between Banach spaces; e.g., see [20], where this lemma is stated for operators with closed images.] In the case under consideration, there exists a functional $y^* \in Y^*$ such that

$$f'(\bar{x}) = (F'(\bar{x}))^* y^*. \tag{5.5}$$

It is easy to see that the equality $Y = Y_1 \oplus Y_2$ implies the equality $Y^* = Y_1^{\perp} \oplus Y_2^{\perp}$. Since $\ker(F'(\bar{x}))^* = (\mathrm{im}\, F'(\bar{x}))^{\perp} = Y_1^{\perp}$, $y^*$ in (5.5) can be chosen in such a way that

$$y^* \in Y_2^{\perp} \subset (\mathrm{im}(PF')'(\bar{x};\, \bar{h}))^{\perp} = \ker((PF')'(\bar{x};\, \bar{h}))^*,$$

so that

$$(F'(\bar{x}))^* y^* = (\Psi_2(\bar{h}))^* y^*.$$

This and (5.5) imply (5.4).

Now we turn our attention to the case when $f'(\bar{x}) \notin (\ker F'(\bar{x}))^{\perp}$. By the definition of $\Phi$, we have that

$$\ker \Phi'(\bar{x}) = \ker f'(\bar{x}) \cap \ker F'(\bar{x}). \tag{5.6}$$

Furthermore, the subspace

$$Z_1 = \operatorname{im} \Phi'(\bar{x}) = \mathbf{R} \times \operatorname{im} F'(\bar{x})$$

is closed, and the subspace

$$Z_2 = \{0\} \times Y_2$$

is a closed complement of $Z_1$ in $Z = \mathbf{R} \times Y$. Note that the projector $\Pi$ in $Z$ onto $Z_2$ parallel to $Z_1$ is given by the following formula:

$$\Pi z = (0,\, Py), \quad z = (\mu,\, y) \in Z. \tag{5.7}$$

Taking into account (5.3), (5.6) and (5.7), we obtain that

$$\bar{h} \in \ker f'(\bar{x}) \cap \ker F'(\bar{x}) = \ker \Phi'(\bar{x}), \quad (\Pi\Phi')'(\bar{x};\, \bar{h})\bar{h} = (0,\, (PF')'(\bar{x};\, \bar{h})\bar{h}) = 0.$$

Hence, if $\Phi$ were 2-regular at $\bar{x}$ with respect to $\bar{h}$, then one could apply Theorem 4.1 to establish the covering property for $\Phi$ at $\bar{x}$. But this covering would contradict local optimality of $\bar{x}$. Therefore, $\Phi$ cannot be 2-regular at $\bar{x}$ with respect to $\bar{h}$. In other words,

$$\operatorname{im}(\Phi'(\bar{x}) + (\Pi\Phi')'(\bar{x};\, \bar{h})) \neq Z. \tag{5.8}$$

But (5.7) implies that

$$(\Phi'(\bar{x}) + (\Pi\Phi')'(\bar{x};\, \bar{h}))\xi = (f'(\bar{x})\xi\,,\, \Psi_2(\bar{h})\xi), \quad \xi \in X.$$

Hence, due to 2-regularity of $F$ at $\bar{x}$ with respect to $\bar{h}$, (5.8) can be satisfied only if

$$f'(\bar{x}) \in (\ker \Psi_2(\bar{h}))^{\perp} = \operatorname{im}(\Psi_2(\bar{h}))^*,$$

where again the lemma on annihilator was used. This completes the proof. ∎

It is clear that in the regular case Theorem 5.2 reduces to the classical first-order necessary optimality conditions. Indeed, suppose regularity condition (1.1) holds. As observed earlier, in that case $F$ is 2-regular at $\bar{x}$ with respect to any $h \in X$. In particular, we can choose $\bar{h} = 0$, so that (5.3) holds automatically. Furthermore, with this choice $\Psi_2(\bar{h}) = F'(\bar{x})$, and (5.4) reduces to $f'(\bar{x}) = (F'(\bar{x}))^* y^*$, i.e., we recover classical optimality conditions. Theorem 5.2 is also a generalization of similar results derived for twice Fréchet-differentiable constraints in [2, 3] (see also [20, 1, 24]). However, Theorem 5.2 is certainly meaningful beyond those two cases, as illustrated by the following example.

**Example 5.1** Let $\bar{x}$ and $F$ be as defined in Example 2.1, and the objective function be given by

$$f : \mathbf{R}^2 \to \mathbf{R}, \quad f(x) = ax_1 + bx_2 + o(\|x\|),$$

where $a, b \in \mathbf{R}$ are parameters.

Taking into account 2-regularity of $F$ at $\bar{x}$ and the representation of the cone $T_2$ in Example 2.1, primal necessary optimality conditions stated in Theorem 5.1 yield that $\bar{x}$ is not a minimizer if $a < 0$ or $b < 0$. If $a > 0$ and $b > 0$, then (5.2) holds, and hence $\bar{x}$ is a strict local minimizer. Now suppose, for instance, that $a > 0$ and $b = 0$. In this case, (5.2) is not satisfied, and (5.3) holds, e.g., with $\bar{h} = (0, 1)$. For this $\bar{h}$, (5.4) holds with $y^* = 1/2$, which illustrates Theorem 5.2. Note that similar previously known results do not apply to characterize optimality in this case.

# 6 Problems with Complementarity Constraints

To apply our results on the tangent cone and optimality conditions to a specific problem, it is clear that some knowledge of the structure of the constraints is needed (note that the stability result given by the covering theorem, on the other hand, does not require any special structure). One good example of structured problems is optimization subject to complementarity constraints, which we consider in this section. Let the feasible set $D$ be defined by

$$D = \{x = (u, v) \in \mathbf{R}^n \times \mathbf{R}^m \mid g(u, v) \geq 0, \, v \geq 0, \, \langle g(u, v), v\rangle = 0\}, \tag{6.1}$$

where $g : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^m$ is a given smooth mapping. Problem (1.3), (6.1) is an important special case of a mathematical program with equilibrium constraints (MPEC), see [25]. It is well-known [11] that for MPEC the classical Mangasarian-Fromovitz constraint qualification is violated at every feasible point, and this is the main difficulty in treating problems of this type.

In the complementarity literature, the idea of *reformulation* proved to be one of the most productive, see [29, 28, 23, 13, 35], the collection [14], and references therein. In the present context, of relevance are equation-based reformulations. It is generally considered that smooth reformulations of complementarity constraints are not adequate, because linearization of such reformulations cannot capture structural nonconvexity of the feasible set and combinatorial nature of complementarity conditions. One purpose of this section is to show that smooth reformulations can be adequate for dealing with complementarity constraints, via the theory of 2-regularity. Our other purpose is to obtain a "simplified" characterization of stationarity for MPEC. Identifying classes of problems which admit a simple description of stationarity is an important task in the MPEC literature [30, 33]. We show that 2-regularity can be used to derive simple forms of optimality conditions, which are incidentally quite different in nature from the known ones, such as in [30, 33].

To this end, define the mapping

$$\begin{aligned} F : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^m, \quad F_i(x) &= 2g_i(u, v)v_i - (\min\{0, \, g_i(u, v) + v_i\})^2, \\ & \quad i = 1, \, \ldots, \, m, \quad x = (u, v). \end{aligned}$$

It is easy to see (recall also Example 2.1) that

$$D = \{x \in \mathbf{R}^n \times \mathbf{R}^m \mid F(x) = 0\}. \tag{6.2}$$

Hence, instead of problem (1.3), (6.1) we can consider problem (1.3), (6.2) which is purely equality-constrained, and is, in a certain sense, equivalent to the original problem.

Fix a point $\bar{x} = (\bar{u}, \bar{v}) \in D$. We assume that $g$ is differentiable on some neighborhood $V$ of $\bar{x}$ in $\mathbf{R}^n \times \mathbf{R}^m$, and its derivative is Lipschitzian on $V$. Then the first derivative of $F$ is given by

$$\begin{aligned} F_i'(x) &= 2\left((v_i - \min\{0,\, g_i(u,\, v) + v_i\})\frac{\partial g_i}{\partial u}(u,\, v),\, v_i\frac{\partial g_i}{\partial v}(u,\, v)\right. \\ &\quad \left. +g_i(u,\, v)e^i - \min\{0,\, g_i(u,\, v) + v_i\}\left(\frac{\partial g_i}{\partial v}(u,\, v) + e^i\right)\right), \\ &\quad i = 1,\, \ldots,\, m, \quad x = (u,\, v) \in V, \end{aligned}$$

where $e^1, \ldots e^m$ is the standard basis in $\mathbf{R}^m$. It is clear that under our assumptions, $F'(\cdot)$ is Lipschitzian on $V$. Define the three index sets

$$\begin{aligned} I_0 &:= \{i = 1,\, \ldots,\, m \mid g_i(\bar{u},\, \bar{v}) = 0,\, \bar{v}_i = 0\}, \\ I_1 &:= \{i = 1,\, \ldots,\, m \mid g_i(\bar{u},\, \bar{v}) = 0,\, \bar{v}_i > 0\}, \\ I_2 &:= \{i = 1,\, \ldots,\, m \mid g_i(\bar{u},\, \bar{v}) > 0,\, \bar{v}_i = 0\}. \end{aligned}$$

These index sets depend on the point $\bar{x}$ under consideration, but we omit this dependency as this point is fixed throughout the analysis. With this notation, we have that

$$F_i'(\bar{x}) = 2\begin{cases} 0, & \text{if } i \in I_0, \\ \bar{v}_i g_i'(\bar{u},\, \bar{v}), & \text{if } i \in I_1, \\ g_i(\bar{u},\, \bar{v})(0,\, e^i), & \text{if } i \in I_2. \end{cases} \tag{6.3}$$

Observe that $F$ cannot be regular at $\bar{x}$ in the classical sense, except possibly under the restrictive assumption of *strict complementarity*, i.e., $I_0 = \emptyset$. Hence, classical theory is not applicable here. The theory of 2-regularity for the $C^2$ case is also not useful, because $F$ is not twice differentiable when $I_0 \neq \emptyset$.

Obviously, the null-space of $F'(\bar{x})$ is given by

$$\ker F'(\bar{x}) = \left\{ h = (p,\, q) \in \mathbf{R}^n \times \mathbf{R}^m \;\middle|\; \begin{array}{l} \langle g_i'(\bar{u},\, \bar{v}),\, h\rangle = 0,\, i \in I_1, \\ q_i = 0,\, i \in I_2 \end{array} \right\}. \tag{6.4}$$

For the sake of simplicity of demonstration, we shall assume that

$$g_i'(\bar{u},\, \bar{v})\,, i \in I_1\,, (0,\, e^i)\,, i \in I_2 \text{ are linearly independent in } \mathbf{R}^n \times \mathbf{R}^m. \tag{6.5}$$

This assumption is not necessary for the reformulation of complementarity constraints to have 2-regularity properties, but it simplifies considerably the subsequent computations. In any case, this assumption is clearly very mild. In particular, (6.5) is much weaker than the linear independence constraint qualification (LICQ) for MPEC [25, 26], which is

$$g_i'(\bar{u},\, \bar{v})\,, i \in I_0 \cup I_1\,, (0, e^i)\,, i \in I_0 \cup I_2 \quad \text{are linearly independent in } \mathbf{R}^n \times \mathbf{R}^m. \tag{6.6}$$

Under the assumption (6.5), choosing as a complementary subspace for $Y_1 = \operatorname{im} F'(\bar{x})$ its orthogonal complement in $\mathbf{R}^m$, we have that

$$Y_2 = Y_1^{\perp} = \operatorname{span}\{e^i \mid i \in I_0\}.$$

Then the associated projector $P$ is the orthogonal projector onto $Y_2$ in $\mathbf{R}^m$:

$$(Py)_i = \begin{cases} y_i, & \text{if } i \in I_0, \\ 0, & \text{if } i \notin I_0. \end{cases} \tag{6.7}$$

Let $g$ be twice differentiable at $\bar{x}$. By direct calculations, we obtain the following formula for directional derivatives of the mappings $F_i'(\cdot) : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n \times \mathbf{R}^m$, $i \in I_0$:

$$\begin{array}{rcl} (F_i')'(\bar{x};\, h) & = & 2(q_i - \min\{0,\, \langle g_i'(\bar{u},\, \bar{v}),\, h\rangle + q_i\})g_i'(\bar{u},\, \bar{v}) \\ & & +2(\langle g_i'(\bar{u},\, \bar{v}),\, h\rangle - \min\{0,\, \langle g_i'(\bar{u},\, \bar{v}),\, h\rangle + q_i\})(0,\, e^i), \\ & & i \in I_0, \quad h = (p,\, q) \in \mathbf{R}^n \times \mathbf{R}^m. \end{array} \tag{6.8}$$

Taking into account (6.4), (6.7), it is not difficult to verify that

$$T_2 = \left\{ h = (p,\, q) \in \mathbf{R}^n \times \mathbf{R}^m \;\middle|\; \begin{array}{l} \min\{\langle g_i'(\bar{u},\, \bar{v}),\, h\rangle,\, q_i\} = 0,\ i \in I_0, \\ \langle g_i'(\bar{u},\, \bar{v}),\, h\rangle = 0,\ i \in I_1, \\ q_i = 0,\ i \in I_2. \end{array} \right\}. \tag{6.9}$$

For an arbitrary element $h = (p,\, q) \in T_2$, let us define the index sets

$$I_0^0(h) = \{i \in I_0 \mid q_i = 0\}, \quad I_0^+(h) = \{i \in I_0 \mid q_i > 0\}.$$

Taking into account (6.8), (6.9), and the latter definitions, we have that

$$(F_i')'(\bar{x};\, h) = 2\begin{cases} \langle g_i'(\bar{u},\, \bar{v}),\, h\rangle(0,\, e^i), & i \in I_0^0(h), \\ q_i g_i'(\bar{u},\, \bar{v}), & i \in I_0^+(h). \end{cases}$$

Using again (6.3) and (6.7), it is easy to see that under the assumption (6.5), the mapping $F$ is 2-regular at $\bar{x}$ with respect to $h$ if, and only if,

$$\langle g_i'(\bar{u},\, \bar{v}),\, h\rangle > 0 \quad \forall\, i \in I_0^0(h), \tag{6.10}$$

and

$$g_i'(\bar{u},\, \bar{v})\,, i \in I_0^+(h) \cup I_1\,, (0,\, e^i)\,, i \in I_0^0(h) \cup I_2 \text{ are linearly independent in } \mathbf{R}^n \times \mathbf{R}^m. \tag{6.11}$$

Inequality (6.10) can be interpreted as the strict complementarity condition at the solution $h$ of the mixed complementarity problem defining $T_2$ (see (6.9)).

Next, we briefly comment on the relation between the stated 2-regularity condition and the well-known piecewise Mangasarian-Fromovitz constraint qualification for MPEC. For any pair $(A, B)$ of index sets such that $A \cup B = I_0$, $A \cap B = \emptyset$, define the *branch* of the feasible set $D$ of (1.3), (6.1) by

$$D^{(A,\,B)} = \left\{ x = (u,\,v) \in \mathbf{R}^n \times \mathbf{R}^m \;\middle|\; \begin{array}{l} v_i = 0,\ g_i(u,\,v) \geq 0,\ i \in A, \\ v_i \geq 0,\ g_i(u,\,v) = 0,\ i \in B, \\ g_i(u,\,v) = 0,\ i \in I_1, \\ v_i = 0,\ i \in I_2 \end{array} \right\}.$$

As is easy to see, there exists a neighborhood $\mathcal{O}$ of $\bar{x}$ in $\mathbf{R}^n \times \mathbf{R}^m$ such that $D^{(A,\,B)} \cap \mathcal{O} \subset D$, and if $x \in D \cap \mathcal{O}$ then there exist a branch $D^{(A,\,B)}$ such that $x \in D^{(A,\,B)}$. Hence, the feasible set $D$ can be locally decomposed into branches defined above. The piecewise Mangasarian-Fromovitz constraint qualification consists of saying that constraints defining each of the branches satisfy at $\bar{x}$ the usual MFCQ. Piecewise constraint qualifications play an important role in theoretical and numerical treatment of MPEC [22, 30].

To each branch of $D$ we can further associate the corresponding branch of cone $T_2$:

$$T_2^{(A,\,B)} = \left\{ h = (p,\,q) \in \mathbf{R}^n \times \mathbf{R}^m \;\middle|\; \begin{array}{l} q_i = 0,\ \langle g_i'(\bar{u},\,\bar{v}),\,h\rangle \geq 0,\ i \in A, \\ q_i \geq 0,\ \langle g_i'(\bar{u},\,\bar{v}),\,h\rangle = 0,\ i \in B, \\ \langle g_i'(\bar{u},\,\bar{v}),\,h\rangle = 0,\ i \in I_1, \\ q_i = 0,\ i \in I_2 \end{array} \right\}.$$

Clearly, $T_2^{(A,\,B)} \subset T_2$. On the other hand, if $h \in T_2$, then $h \in T_2^{(I_0^0(h),\,I_0^+(h))}$ with index sets $I_0^0(h)$ and $I_0^+(h)$ defined above. It is easy to see now that under the assumption (6.5), 2-regularity of $F$ at $\bar{x}$ and piecewise MFCQ are, in a certain sense, equivalent. More precisely, if $F$ is 2-regular at $\bar{x}$ with respect to $h$, i.e., (6.10) and (6.11) hold, then MFCQ holds at $\bar{x}$ for the constraints defining the branch $D^{(I_0^0(h),\,I_0^+(h))}$. And vice versa, if MFCQ holds at $\bar{x}$ for a branch $D^{(A,\,B)}$, then there exists $h \in T_2^{(A,\,B)}$ such that $F$ is 2-regular at $\bar{x}$ with respect to $h$. Moreover, such elements $h$ comprise a dense subset of $T_2^{(A,\,B)}$. In particular, if piecewise MFCQ holds at $\bar{x}$ for problem (1.3), (6.1), then $F$ is 2-regular at $\bar{x}$ with respect to every $h$ from a dense subset of $T_2$. Hence (recall Remark 2.1),

$$\mathcal{T}_D(\bar{x}) = \mathcal{B}_D(\bar{x}) = T_2.$$

Observe that in the case of (6.5), cone $T_2$ obtained via the theory of 2-regularity is the standard "linearized" cone for MPEC [25]. Hence, applying the primal optimality conditions of Theorem 5.1, we recover the concept of *B-stationarity*, as defined in [33]. The latter is generally considered to be the proper first-order primal optimality condition for MPEC. We point out that here it comes out as a direct consequence of our general theory, which confirms that 2-regularity introduced in this paper is an adequate tool for treating complementarity constraints. Note also that piecewise MFCQ is arguably the most natural assumption which ensures that the MPEC tangent cone is given by (6.9). Our 2-regularity condition is no stronger than the piecewise MFCQ (the two are equivalent under the assumption (6.5)), which shows that 2-regularity is also a natural concept for MPEC.

The practical difficulty in verifying various optimality concepts for MPEC consists in combinatorial aspect of the MPEC tangent cone, and the generally complex structure of its dual. In [30, 33], some situations have been identified when MPEC admits a simple form of primal-dual optimality conditions. The approaches of [30, 33] are essentially based on obtaining regularity conditions which ensure that the dual of the MPEC tangent cone coincides with the dual of the tangent cone of a certain standard nonlinear program related to MPEC. We next show that our approach also yields simplified optimality conditions for MPEC, but it is quite different in nature when compared to [30, 33] (this is already clear from the proof of Theorem 5.2, as it is not based on computing the dual cone).

By Theorem 5.2, if there exists some critical direction $h$ satisfying (6.10), (6.11), then optimality conditions for problem (1.3), (6.2) take the form of (5.4). It is easy to see that in the present context the latter means that

$$f'(\bar{x}) = \sum_{i \in I_0^+(h) \cup I_1} \bar{y}_i g_i'(\bar{u}, \bar{v}) + \sum_{i \in I_0^0(h) \cup I_2} \bar{y}_i (0, e^i) \tag{6.12}$$

with some $\bar{y} \in \mathbf{R}^m$. We next compare our regularity and optimality conditions with some others in the MPEC literature.

In [33], under the LICQ (6.6), the following primal-dual optimality conditions were obtained:

$$\begin{aligned} f'(\bar{x}) &= \sum_{i \in I_0 \cup I_1} \bar{\lambda}_i g_i'(\bar{u}, \bar{v}) + \sum_{i \in I_0 \cup I_2} \bar{\mu}_i (0, e^i), \\ & \quad \bar{\lambda}_i \geq 0\,, \bar{\mu}_i \geq 0\,, i \in I_0\,. \end{aligned} \tag{6.13}$$

First, it should be noted that LICQ certainly implies piecewise MFCQ, which is in a sense equivalent to 2-regularity, as discussed above. In particular, LICQ implies conditions (6.10), (6.11) for all branches, while we need 2-regularity only for one direction $h$. On the other hand, we need this $h$ to be a critical direction. It is quite clear that 2-regularity and LICQ are therefore essentially different conditions, i.e., one can hold when the other does not, and vice versa. Second, comparing formally the representations for the gradient of the objective function in (6.12) and (6.13), note that again neither is weaker or stronger than the other. On one hand, since $I_0 = I_0^0 \cup I_0^+$, (6.12) involves twice fewer gradients of constraints from the index set $I_0$ (i.e., half of the corresponding multipliers are zero), so in this respect (6.12) is sharper. But on the other hand, (6.13) comes with sign constraints for the multipliers associated with $I_0$, so it is better in this sense. The conclusion is that optimality conditions based on 2-regularity are essentially different from those in [33].

Without stating the results of [30], we point out that our regularity/optimality conditions principally differ also from that reference. Note that simplified primal-dual conditions in [30] refer to the case when the tangent cone and/or its dual can be computed as corresponding cones of certain standard nonlinear programs associated with MPEC. By contrast, our development has a certain "directional flavor", and our primal-dual optimality conditions are derived without computing full characterization of the dual of the tangent cone. Indeed, our results follow from Theorem 5.2, where the primal-dual optimality conditions are obtained by using the covering theorem, and not by computing the dual of the tangent cone.

Comparing the approach described in this section with [17], we note that the latter is based on specialized 2-regularity constructions (in the sense of twice differentiable mappings!) for a certain subset of MPEC constraints. In fact, to develop 2-regularity directly for the full set of equality and inequality constraints in (6.1) is simply impossible. This is only natural, as MPEC considered as standard nonlinear program cannot be regular in any classical sense, whether first or second order. For this reason, the approach of [17] is quite special, and may not be applicable to other related problems without nontrivial modifications. On the other hand, [17] also leads to a certain form of second-order optimality conditions. In this section, we have shown that 2-regularity proposed in this paper is applicable to (6.1) *directly*, once the feasible set is written in the form of $C^{1,1}$ equations. Since this development is based on a general theory, it should be also applicable to other problems where constraints include complementarity conditions.

Finally, we believe that our analysis can also be used with other appropriate reformulations, such as in [29, 23]. We chose to focus on just one for the clarity of presentation. Also, it is definitely possible to study 2-regularity and the resulting optimality conditions for MPEC without the simplifying assumption (6.5). However, the development in that case is technically more involved. Given that MPEC is not the primary subject of the present paper, we do not pursue here the study of this case. Nevertheless, we note that this may actually be the case where application of 2-regularity to MPEC is especially useful, because other approaches may not be applicable. This issue requires further investigation.

## References


[1] A.V. Arutyunov. *Optimality Conditions: Abnormal and Degenerate Problems.* Kluwer Academic Publishers, Dordrecht, 2000.

[2] E.R. Avakov. Extremum conditions for smooth problems with equality-type constraints. *USSR Computational Mathematics and Mathematical Physics*, 25:24–32, 1985.

[3] E.R. Avakov. Necessary extremum conditions for smooth abnormal problems with equality- and inequality constraints. *Mathematical Notes*, 45:431–437, 1989.

[4] E.R. Avakov. Theorems on estimates in the neighborhood of a singular point of a mapping. *Mathematical Notes*, 47:425–432, 1990.

[5] E.R. Avakov, A.A. Agrachyov, and A.V. Arutyunov. The level-set of a smooth mapping in a neighborhood of a singular point, and zeros of a quadratic mapping. *Mathematics USSR Sbornik*, 73:455–466, 1992.

[6] K.N. Belash and A.A. Tretyakov. Methods for solving singular problems. *USSR Computational Mathematics and Mathematical Physics*, 28:1097–1102, 1988.

[7] J.F. Ben-Tal. Second-order and related extremality conditions in nonlinear programming. *Journal of Optimization Theory and Applications*, 31:143–165, 1980.

[8] J.F. Ben-Tal and J. Zowe. A unified theory of first and second order optimality conditions for extremum problems in topological vector spaces. *Mathematical Programming Study*, 19:39–76, 1982.

[9] J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems.* Springer–Verlag, New York, NY, 2000.

[10] M. Buchner, J. Marsden, and S. Schecter. Applications of the blowing-up construction and algebraic geometry to bifurcation problems. *Journal of Differential Equations*, 48:404–433, 1983.

[11] Y. Chen and M. Florian. The nonlinear bilevel programming problem: formulations, regularity and optimality conditions. *Optimization*, 32:193–209, 1995.

[12] A.V. Dmitruk, A.A. Milyutin, and N.P. Osmolovskii. Lyusternik's theorem and the theory of extrema. *Russian Mathematical Surveys*, 35:11–51, 1980.

[13] F. Facchinei and J. Soares. A new merit function for nonlinear complementarity problems and a related algorithm. *SIAM Journal on Optimization*, 7:225–247, 1997.

[14] M. Fukushima and L. Qi (editors). *Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods.* Kluwer Academic Publishers, 1999.

[15] A.D. Ioffe and V.M. Tikhomirov. *Theory of Extremal Problems.* North–Holland, Amsterdam, Holland, 1974.

[16] A.F. Izmailov. On certain generalizations of Morse's lemma. *Proceedings of the Steklov Institute of Mathematics*, 220:138–153, 1998.

[17] A.F. Izmailov and M.V. Solodov. Complementarity constraint qualification via the theory of second-order regularity. Manuscript, 1999.

[18] A.F. Izmailov and M.V. Solodov. Optimality conditions for irregular inequality-constrained problems, 1999. *SIAM J. on Control and Optimization*, to appear.

[19] A.F. Izmailov and M.V. Solodov. Error bounds for 2-regular mappings with Lipschitzian derivatives and their applications. *Mathematical Programming*, 89:413–435, 2001.

[20] A.F. Izmailov and A.A. Tretyakov. *Factor-Analysis of Nonlinear Mappings.* Nauka, Moscow, 1994. In Russian.

[21] A.F. Izmailov and A.A. Tretyakov. 2-*Regular Solutions of Nonlinear Problems. Theory and Numerical Methods.* Fizmatlit, Moscow, 1999. In Russian.

[22] H. Jiang and D. Ralph. Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. *SIAM Journal on Optimization*, 10:779–808, 2000.

[23] C. Kanzow. Some equation-based methods for the nonlinear complementarity problem. *Optimization Methods and Software*, 3:327–340, 1994.

[24] U. Ledzewicz and H. Schättler. High-order approximations and generalized necessary conditions for optimality. *SIAM Journal on Control and Optimization*, 37:33–53, 1998.

[25] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical programs with equilibrium constraints.* Cambridge University Press, 1996.

[26] Z.-Q. Luo, J.-S. Pang, and D. Ralph. Piecewise sequential quadratic programming for mathematical programs with nonlinear complementarity constraints. In P. Pardalos A. Migdalas and P. Varbrand, editors, *Multilevel Optimization: Algorithms, Complexity and Applications*, pages 209–230. Kluwer Academic Publishers, 1998.

[27] R.J. Magnus. On the local structure of the zero-set of a Banach space valued mapping. *Journal of Functional Analysis*, 22:58–72, 1976.

[28] O. L. Mangasarian and M. V. Solodov. Nonlinear complementarity as unconstrained and constrained minimization. *Mathematical Programming*, 62:277–297, 1993.

[29] O.L. Mangasarian. Equivalence of the complementarity problem to a system of nonlinear equations. *SIAM Journal of Applied Mathematics*, 31:89–92, 1976.

[30] J.-S. Pang and M. Fukushima. Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints. *Computational Optimization and Applications*, 13:111–136, 1999.

[31] J.-P. Penot. Second-order conditions for optimization problems with constraints. *SIAM Journal on Control and Optimization*, 37:303–315, 1999.

[32] S. M. Robinson. Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity. *Mathematical Programming Study*, 30:45–66, 1987.

[33] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality and sensitivity. *Mathematics of Operations Research*, 25:1–22, 2000.

[34] A. Shapiro. On concepts of directional differentiability. *Journal of Optimization Theory and Applications*, 66:477–487, 1990.

[35] M. V. Solodov. Implicit Lagrangian. In C. Floudas and P. Pardalos, editors, *Encyclopedia of Optimization.* Kluwer Academic Publishers, 2001.

[36] A. Szulkin. Local structure of the zero-sets of differentiable mappings and applications to bifurcation theory. *Mathematica Scandnavica*, 45:232–242, 1979.

[37] A.A. Tretyakov. Necessary and sufficient conditions for optimality of $p$-th order. *USSR Computational Mathematics and Mathematical Physics*, 24:123–127, 1984.