Math. Program., Ser. A 89: 413–435 (2001)

Digital Object Identifier (DOI) 10.1007/s101070100210

A.F. Izmailov · M.V. Solodov

Error bounds for 2-regular mappings with Lipschitzian derivatives and their applications

Received: February 2000 / Accepted: November 2000

Published online January 17, 2001 – © Springer-Verlag 2001

Abstract. We obtain local estimates of the distance to a set defined by equality constraints under assumptions which are weaker than those previously used in the literature. Specifically, we assume that the constraints mapping has a Lipschitzian derivative, and satisfies a certain 2-regularity condition at the point under consideration. This setting directly subsumes the classical regular case and the twice differentiable 2-regular case, for which error bounds are known, but it is significantly richer than either of these two cases. When applied to a certain equation-based reformulation of the nonlinear complementarity problem, our results yield an error bound under an assumption more general than *b*-regularity. The latter appears to be the weakest assumption under which a local error bound for complementarity problems was previously available. We also discuss an application of our results to the convergence rate analysis of the exterior penalty method for solving irregular problems.

Key words. error bound – $C^{1,1}$-mapping – 2-regularity – nonlinear complementarity problem – exterior penalty – rate of convergence

1. Error bounds and their applications

Among the most important tools for theoretical and numerical treatment of nonlinear operator equations, optimization problems, variational inequalities, and other related problems, are the so-called *error bounds*, i.e., upper estimates of the distance to a given set in terms of some residual function. We refer the reader to [31] for a survey of error bounds and their applications. When the set is defined by functional constraints, a typical residual function is some measure of violation of constraints at the given point. When available, error bounds can often be used to obtain a *constructive* local description of the set under consideration. This description, in turn, plays a central role in the theory of optimality conditions. As another important application of error bounds, we mention development, implementation, and convergence rate analysis of numerical methods for solving optimization and related problems.

In this paper, we consider error bounds for sets given by equality constraints. However, it is worth to point out that this setting implicitly includes sets of more general structure, namely those whose constraints can be (equivalently) reformulated into equations. One such example is solution set of the nonlinear complementarity problem, which we shall study in Sect. 3.

A.F. Izmailov: Computing Center of the Russian Academy of Sciences, Vavilova Str. 40, Moscow, 117967, Russia, e-mail: izmaf@ccas.ru

M.V. Solodov: Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil, e-mail: solodov@impa.br

Mathematics Subject Classification (1991): 90C30, 49M30, 65K05, 46T20, 90C33

Let X and Y be Banach spaces, V be an open set in X, $F : V \to Y$ be a given mapping, and consider the set

$$D = \{x \in V \mid F(x) = 0\}. \tag{1}$$

We define the distance function from a point to a set in a usual way:

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

Given some point $\bar{x} \in D$, a local error bound is the following property:

$$\operatorname{dist}(x,\ D) \le M\|F(x)\|^{1/\gamma} \quad \forall\, x \in U, \tag{2}$$

where $M > 0$ and $\gamma > 0$ are two constants, and $U \subset V$ is a neighborhood of $\bar{x}$, with M, γ and U being independent of point x appearing in (2). In this paper, we shall also obtain somewhat more general forms of error bounds, but they can be all reduced to the form (2) (this reduction, however, may make them less accurate in certain situations).

In what follows, for a linear operator $\Lambda : X \to Y$, $\ker \Lambda = \{x \in X \mid \Lambda x = 0\}$ is its null space, and $\operatorname{im} \Lambda = \{y \in Y \mid y = \Lambda x$ for some $x \in X\}$ is its image space. By $F^{-1}(y)$ we shall denote the complete pre-image of an element $y \in Y$ with respect to F, namely

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

With this notation, $F^{-1}(F(\bar{x}))$ is the level surface of F passing through $F(\bar{x})$.

It is well-known that a Lipschitzian error bound (namely, (2) with $\gamma = 1$) holds under the assumption that F is smooth and *regular at* $\bar{x}$, i.e.,

$$\operatorname{im} F'(\bar{x}) = Y. \tag{3}$$

Specifically, the following result, sometimes referred to as Graves-Lyusternik Theorem, is classical (e.g., see [15, 9, 2]).

Theorem 1. *Let X and Y be Banach spaces, V be a neighborhood of a point $\bar{x}$ in X. Suppose further that the mapping $F : V \to Y$ is Fréchet-differentiable on V, its derivative is continuous at $\bar{x}$, and F is regular at $\bar{x}$.*

Then there exist a neighborhood $U \subset V$ of $\bar{x}$ in X and a constant $M > 0$ such that

$$\operatorname{dist}(x,\ F^{-1}(F(\bar{x}))) \le M\|F(x) - F(\bar{x})\| \quad \forall\, x \in U. \tag{4}$$

In fact, under the regularity assumption (3), we could state a more general property than the error bound given in Theorem 1. In [2], it is referred to as the pseudo-Lipschitzian property, and it has been studied extensively in the literature in a variety of settings, see [31] for some references.

Let us define the *tangent cone* to D at a point $\bar{x} \in D$ as

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

As an immediate application of the error bound given in Theorem 1, we mention the Graves-Lyusternik theorem [15, 9] describing the tangent subspace to a level surface of a regular mapping.

Theorem 2. *Under the assumptions of Theorem 1, the following equality holds:*

$$T_{F^{-1}(F(\bar{x}))}(\bar{x}) = \ker F'(\bar{x}).$$

The inclusion $T_{F^{-1}(F(\bar{x}))}(\bar{x}) \subset \ker F'(\bar{x})$ is standard, and holds without any regularity assumptions. To see that the error bound (4) implies the opposite inclusion, it is enough to take $x = \bar{x}+th$, $h \in \ker F'(\bar{x})$, $t \in \mathbf{R}_+$, employ the definition of the Fréchet derivative in the right-hand side of (4), take into account that $h \in \ker F'(\bar{x})$, and use the definition of the tangent cone. Theorem 2 is an important tool for deriving necessary optimality conditions for optimization problems with equality constraints.

When regularity condition (3) is not satisfied, the assertions of Theorems 1 and 2 do not hold in general. In particular, a Lipschitzian error bound is not necessarily valid. This irregular case is the main issue under consideration in the present paper.

Without the regularity assumption (3), error bound (2) is known to hold when F is an analytic function [23]. Compared to our results (see Theorem 4 and Remark 7), for error bounds based on analytic functions the constant γ is in general unknown, except in the quadratic case under a certain positivity condition, in which case $\gamma = 2$ (and, of course, in the affine case with $\gamma = 1$). In Sect. 3, we explain that our results are of a rather different nature than those for analytic functions. We also provide an example where our results imply an error bound for complementarity problems with non-analytic functions. For the case of twice Fréchet-differentiable F, certain error bounds were obtained in [5] under the so-called 2-regularity condition, which is considerably weaker than (3) (see also [18, 19]). These results are directly subsumed in our development below, so we shall not survey them here. In Sect. 2, we shall obtain error bounds in the considerably much more general setting where F need not be twice Fréchet-differentiable, but its first Fréchet-derivative is Lipschitz continuous near the point under consideration. This development will use the new notion of 2-regularity under relaxed smoothness assumptions introduced in [17]. We obtain error bounds under regularity/smoothness assumptions for which no such results were previously available. Specifically, we prove an estimate which subsumes (2) with at least $\gamma = 2$. As its corollary, we recover in the regular case Theorem 1 (i.e., (2) with $\gamma = 1$), Theorem 2, and the generalization of the Graves-Lyusternik theorem about the tangent cone for the 2-regular case. We emphasize that our extension of error bound results is significant, as local structure of mappings in our setting can be considerably more complicated than that of twice differentiable 2-regular mappings or regular mappings (in the latter two cases the tangent cone to a level surface is necessarily two-sided, but it need not be two-sided under our assumptions, see [17, Example 2.1]). In this paper, the significance of the new notion of 2-regularity is further demonstrated by an application to an equation-based reformulation of the nonlinear complementarity problem (NCP). It is well-known that useful NCP reformulations are neither twice differentiable nor regular in the classical sense (except possibly under certain restrictive assumptions), and so the theory previously available is usually not directly applicable. In Sect. 3, we obtain a new error bound for NCP under a 2-regularity condition, which we show to be weaker than b-regularity (we show that b-regularity implies our condition, and that the converse is not true). To our knowledge, b-regularity is the weakest condition under which a local error bound for NCP had been known to hold previously. As another application of our results, in Sect. 4 we discuss exterior penalty methods and their convergence rate analysis.

Let us introduce some more notation. By $B(x,\ r)$ we denote the open ball with center $x \in X$ and radius $r > 0$, and by 2^X we denote the set of all subsets of X. 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_Y, we denote the identity operator in Y. Let $\mathcal{L}(X,\ Y)$ be the space of continuous linear operators from X to Y, normed in the usual manner. For a linear operator $\Lambda \in \mathcal{L}(X,\ Y)$, we denote by Λ^{-1} its right inverse, that is $\Lambda^{-1} : Y \to 2^X$, $\Lambda^{-1}(y) = \{x \in X \mid \Lambda x = y\}$. Furthermore, we shall use the “norm”

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

Note that when Λ is one-to-one, $\|\Lambda^{-1}\|$ can be considered as the usual norm of the element Λ^{-1} in the space $\mathcal{L}(Y,\ X)$. We denote $\operatorname{span}\{x^1, \ldots, x^k\} := \{x \in X \mid x = \sum_{i=1}^{k} t_i x^i\ , t_i \in \mathbf{R}\}$ and $\operatorname{cone}\{x^1, \ldots, x^k\} := \{x \in X \mid x = \sum_{i=1}^{k} t_i x^i\ , t_i \in \mathbf{R}_+\}$. For a (finite-dimensional) vector x, x_i will denote its i-th component.

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

Theorem 3 (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

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

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

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

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

Note that in [15], Theorem 3 was established in the more general setting of a complete metric space, but we do not need this level of generality in the present paper.

2. 2-regularity. Error bounds for 2-regular mappings

Throughout this section we assume the following *basic hypotheses*.

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

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 that if Y is a Hilbert space and Y_1 is closed then Y_1 always has a closed complementary subspace in Y (see hypothesis (H3)). Hypothesis (H4) always holds when $F'(\cdot)$ is Lipschitz continuous on V. Note that in the regular case hypotheses (H3) and (H4) are trivially satisfied ($Y_1 = Y$, $Y_2 = \{0\}$, $P = 0$).

Our analysis will be based on the usual notion of 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 defined by

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

This derivative always exists whenever the corresponding directional derivative of the mapping $F'(\cdot)$ exists, but not vice versa: $PF'(\cdot)$ can be directionally differentiable even if $F'(\cdot)$ is not.

The following definitions are direct extensions of those previously introduced in the twice differentiable case, see [3–6, 18, 1, 16, 19]. If F is twice Fréchet-differentiable at $\bar{x}$, our definitions reduce to standard notions of 2-regularity theory. In particular, all directional derivatives involved will be defined via second Fréchet derivatives; for example, $(PF')'(\bar{x};\ h)\xi = PF''(\bar{x})[h,\ \xi]$, $h,\ \xi \in X$. But we stress once again that the setting (H1)–(H4) is significantly different from the case of twice differentiability. Moreover, it is also richer when it comes to applications (more on this in Sect. 3).

If $(PF')'(\bar{x};\ h)$ exists, then it is an element of $\mathcal{L}(X,\ Y)$, and so the following operator can be defined:

$$\Psi_2(h) \in \mathcal{L}(X,\ Y), \quad \Psi_2(h)\xi := F'(\bar{x})\xi + (PF')'(\bar{x};\ h)\xi. \tag{5}$$

This operator is central in the theory of 2-regularity. In this paper, its role will become clear in Theorem 4, which is our main error bound result.

Let $PF'(\cdot)$ be directionally differentiable at $\bar{x}$ with respect to an element $h \in X$.

Definition 1. *The mapping F is referred to as* 2-regular at the point $\bar{x}$ with respect to $h \in X$, *if* $\operatorname{im} \Psi_2(h) = Y$.

Remark 1. It is easy to see that if F is 2-regular at $\bar{x}$ with respect to some $h \in X$, then it is 2-regular with respect to th for all $t > 0$ (this fact is based on positive homogeneity of the directional derivative with respect to a direction).

Remark 2. Consider the case when the mapping F is regular at the point $\bar{x}$, i.e., when (3) holds. Then $Y_1 = Y$, $Y_2 = \{0\}$, $P = 0$, $\Psi_2(h) = F'(\bar{x})$, and hence, F is 2-regular with respect to *every* $h \in X$.

The next 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. *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 cone T_2 and 2-regularity play a central role in the description of tangent directions to a level surface of a nonlinear mapping (see Theorem 5).

Remark 3. If F is regular at $\bar{x}$, then $P = 0$ and so $PF'(\cdot)$ is certainly (directionally) differentiable. Taking further into account Remark 2, it follows that if F is regular at $\bar{x}$, then it is also 2-regular at $\bar{x}$. Note that in the regular case $T_2 = \ker F'(\bar{x})$.

Finally, the following definition requires directional differentiability of $PF'(\cdot)$ at $\bar{x}$ with respect to every direction in X.

Definition 3. *The mapping F is referred to as* strongly 2-regular at the point $\bar{x}$ *if there exists $\nu > 0$ such that*

$$\sup_{\substack{h \in T_2^{\nu},\\ \|h\|=1}} \left\|(\Psi_2(h))^{-1}\right\| < \infty,$$

where

$$T_2^{\nu} = \{h \in X \mid \|F'(\bar{x})h\| \leq \nu,\ \|(PF')'(\bar{x};\ h)h\| \leq \nu\}.$$

Remark 4. Clearly, if the mapping F is regular at the point $\bar{x}$, then F is strongly 2-regular at $\bar{x}$. This follows from the fact that a surjective continuous linear operator from one Banach space to another has a bounded (set-valued) right inverse in the sense of the "norm" defined in Sect. 1.

Remark 5. Note that $T_2 = T_2^0 \subset T_2^{\nu}$ for all $\nu \geq 0$. In the case of a finite-dimensional X, strong 2-regularity at $\bar{x}$ is equivalent to 2-regularity at $\bar{x}$. To see this, one has to recall the following facts.

1. Under the hypothesis (H4) the mapping $(PF')'(\bar{x};\ \cdot) : X \to \mathcal{L}(X,\ Y)$ is continuous on X (in fact, this mapping is even Lipschitzian, see [34, Lemma 2.1]).
2. A small linear perturbation of a surjective linear operator results in a small perturbation of the norm of its right inverse [18, Theorem 2, p. 26].
3. The unit sphere in a finite-dimensional space is compact.

In general Banach spaces, strong 2-regularity is somewhat stronger than 2-regularity (see [6, 18, 16, 19] for some examples and more detailed discussion of this issue).

In the sequel, the following useful notion will also be involved. The mapping $PF'(\cdot)$ is referred to as *B-differentiable at the point $\bar{x}$ with respect to a cone K* in X if it has a directional derivative at $\bar{x}$ with respect to every direction $h \in K$, and

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

We call $PF'(\cdot)$ *B-differentiable at $\bar{x}$*, if it is B-differentiable at this point with respect to the cone $K = X$. The notion of B-differentiability (with respect to $K = X$) was introduced in [35]. 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 any direction [36].

Before we proceed, two technical lemmas on the properties of mappings with B-differentiable derivatives are in order.

Lemma 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 exists a constant $\delta > 0$ such that*

$$B(\bar{x},\ \delta) \subset V,$$

and for all $x \in B(\bar{x},\ \delta) \cap (\bar{x} + K)$, it holds that

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

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

Proof. Using the Mean-Value Theorem and the equality $PF'(\bar{x}) = 0$, we have that

$$\|(I_Y - P)(F(x) - F(\bar{x})) - F'(\bar{x})(x - \bar{x})\| = o(\|x - \bar{x}\|)\,, \tag{6}$$

from which follows the first assertion of the lemma.

Using additionally the Newton–Leibniz formula, for all $x \in (\bar{x} + K)$ sufficiently close to $\bar{x}$, we have that

$$\begin{aligned}
&\left\| P(F(x) - F(\bar{x})) - \frac{1}{2}(PF')'(\bar{x};\ x - \bar{x})(x - \bar{x}) \right\| \\
&= \left\| \int_0^1 P(F'(\bar{x} + \tau(x - \bar{x})) - F'(\bar{x}))(x - \bar{x})d\tau - \frac{1}{2}(PF')'(\bar{x};\ x - \bar{x})(x - \bar{x}) \right\| \\
&= \left\| \int_0^1 (PF')'(\bar{x};\ \tau(x - \bar{x}))(x - \bar{x})d\tau - \frac{1}{2}(PF')'(\bar{x};\ x - \bar{x})(x - \bar{x}) \right\| + o(\|x - \bar{x}\|^2) \\
&= o(\|x - \bar{x}\|^2)\,,
\end{aligned} \tag{7}$$

where the second equality follows from the B-differentiability of $PF'(\cdot)$ at $\bar{x}$ with respect to K, and the last follows from the fact that the mapping $(PF')'(\bar{x};\ \cdot) : K \to \mathcal{L}(X,\ Y)$ is positively homogeneous. Relation (7) establishes the second assertion.

□

Remark 6. Obviously, for $x \neq \bar{x}$ one can guarantee strict inequalities in the assertions of Lemma 1.

Lemma 2. *Under the assumptions of Lemma 1, 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 $x \in B(\bar{x},\ \delta) \cap (\bar{x} + K)$ and $\xi^1,\ \xi^2 \in B(0,\ \gamma\|x - \bar{x}\|)$, it holds that

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

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

Proof. The argument is analogous to Lemma 1.

□

In Theorem 4, we shall need the following regularity assumption. Assuming $PF'(\cdot)$ is directionally differentiable at $\bar{x}$ with respect to every direction in some cone K in X,

$$\exists\, \nu > 0 \text{ such that } \sup_{\substack{h \in K_2^{\nu},\\ \|h\|=1}} \left\|(\Psi_2(h))^{-1}\right\| < \infty, \tag{8}$$

where

$$K_2^{\nu} = \{h \in K \mid \|F'(\bar{x})h\| \le \nu,\ \|(PF')'(\bar{x};\ h)h\| \le \nu\}.$$

Note that for $K = X$ this condition is precisely strong 2-regularity of F at $\bar{x}$ introduced in Definition 3. In the case of finite-dimensional X, this is further equivalent to 2-regularity of F at $\bar{x}$ (recall Remark 5). Another useful observation is that if F is regular at $\bar{x}$, (8) is clearly satisfied for $K = X$ with any $\nu > 0$ (recall Remark 4).

The following is our main result. After the proof, in Remark 7, we shall discuss its relation to the more standard form of error bounds, such as (2).

Theorem 4. *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 that the regularity condition* (8) *is satisfied.*

Then there exist a neighborhood $U \subset V$ of $\bar{x}$ in X and a constant $M > 0$ such that for all $x \in (U \cap (\bar{x} + K)) \setminus \{\bar{x}\}$ it holds that

$$\operatorname{dist}(x,\ F^{-1}(F(\bar{x}))) \le M\left(\|(I_Y - P)(F(x) - F(\bar{x}))\| + \frac{\|P(F(x) - F(\bar{x}))\|}{\|x - \bar{x}\|}\right). \tag{9}$$

In particular, if $K = X$ (in which case (8) *means that F is strongly 2-regular at $\bar{x}$), then* (9) *holds for all $x \in U \setminus \{\bar{x}\}$.*

Proof. Let $\rho > 0$ be such that $B(\bar{x},\ 2\rho) \subset V$. Fix an arbitrary $\theta \in (0,\ 1)$, and denote the constant in the left-hand side of (8) by C, $C > 0$. Applying Lemma 2, we conclude that there exist two constants $\delta \in (0, \rho]$ and $\gamma \in (0, \rho/\delta]$ such that for all $x \in B(\bar{x}, \delta) \cap (\bar{x} + K)$ and for all $\xi^1, \xi^2 \in B(0, \gamma\|x - \bar{x}\|)$ the following relations hold:

$$\left\|(I_Y - P)(F(x + \xi^1) - F(x + \xi^2)) - F'(\bar{x})(\xi^1 - \xi^2)\right\| \leq \frac{\theta}{2C}\|\xi^1 - \xi^2\|, \tag{10}$$

$$\|P(F(x + \xi^1) - F(x + \xi^2)) - (PF')'(\bar{x};\, x - \bar{x})(\xi^1 - \xi^2)\| \leq \frac{\theta}{2C}\|x - \bar{x}\|\|\xi^1 - \xi^2\|. \tag{11}$$

Next, choose

$$\bar{\nu} \in \left(0,\ \min\left\{\nu,\ \frac{(1-\theta)\gamma}{4C}\right\}\right). \tag{12}$$

First, we consider the case when

$$\frac{x - \bar{x}}{\|x - \bar{x}\|} \in K_2^{\bar{\nu}}. \tag{13}$$

Define $r(x) = \gamma\|x - \bar{x}\|$, $x \in B(\bar{x}, \delta)$. According to Lemma 1, Remark 6, and (12), there exists a neighborhood $U \subset B(\bar{x}, \delta)$ of $\bar{x}$ in X such that for all $x \in (U \cap (\bar{x} + K)) \setminus \{\bar{x}\}$, we have that

$$\|(I_Y - P)(F(x) - F(\bar{x}))\| < \frac{1-\theta}{2C} r(x), \tag{14}$$

$$\|P(F(x) - F(\bar{x}))\| < \frac{1-\theta}{2C}\|x - \bar{x}\| r(x). \tag{15}$$

For each $x \in U$, define the set-valued mapping

$$\Phi_x : B(0, r(x)) \to 2^X, \quad \Phi_x(\xi) = \xi - (\Psi_2(x - \bar{x}))^{-1}(F(x + \xi) - F(\bar{x})).$$

For all $x \in U \setminus \{\bar{x}\}$ satisfying (13), by (8) and (12) (recall also Remark 1), it holds that

$$\Phi_x(\xi) \neq \emptyset \quad \forall\, \xi \in B(0, r(x)). \tag{16}$$

In the case of (13), we shall prove our assertion by applying the set-valued contracting mapping principle (i.e., Theorem 3) to the mapping defined above. To apply Theorem 3, we have to estimate the (Hausdorff) distance between $\Phi_x(\xi^1)$ and $\Phi_x(\xi^2)$ for each $x \in U \setminus \{\bar{x}\}$ satisfying (13) and each $\xi^1, \xi^2 \in B(0, r(x))$. By the definition, we have that

$$\eta \in \Phi_x(\xi) \quad \Leftrightarrow \quad \Psi_2(x - \bar{x})\eta = \Psi_2(x - \bar{x})\xi - F(x + \xi) + F(\bar{x}).$$

In particular, for each fixed $\xi \in B(0, r(x))$, $\Phi_x(\xi)$ is a (nonempty, by (16)) affine set parallel to the subspace $\ker \Psi_2(x - \bar{x})$.

For any $x^i \in \Phi_x(\xi^i)$, $i = 1, 2$, we have that

$$\Psi_2(x - \bar{x})(x^1 - x^2) = \Psi_2(x - \bar{x})(\xi^1 - \xi^2) - F(x + \xi^1) + F(x + \xi^2). \tag{17}$$

Observe further that by the definition (5) of Ψ_2, for any $\xi \in X$ and $t > 0$, we have that

$$\begin{aligned}\Psi_2(t(x-\bar{x}))\xi &= F'(\bar{x})\xi + (PF')'(\bar{x};\ t(x-\bar{x}))\xi \\ &= F'(\bar{x})\xi + t(PF')'(\bar{x};\ x-\bar{x})\xi.\end{aligned}$$

In particular, multiplication of $x - \bar{x}$ 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 (17), we have that

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

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

$$\begin{aligned}&\mathcal{H}(\Phi_x(\xi^1),\ \Phi_x(\xi^2)) = \inf_{\substack{x^i \in \Phi_x(\xi^i),\\ i=1,\,2}} \|x^1-x^2\| \\ &= \inf\left\{\|\xi\| \;\middle|\; \xi \in X,\ \Psi_2\left(\frac{x-\bar{x}}{\|x-\bar{x}\|}\right)\xi = \Psi_2\left(\frac{x-\bar{x}}{\|x-\bar{x}\|}\right)(\xi^1-\xi^2) \right. \\ &\quad -(I_Y-P)(F(x+\xi^1)-F(x+\xi^2)) \\ &\quad \left. -\|x-\bar{x}\|^{-1}P(F(x+\xi^1)-F(x+\xi^2))\right\} \\ &\le C\Big(\big\|(I_Y-P)(F(x+\xi^1)-F(x+\xi^2))-F'(\bar{x})(\xi^1-\xi^2)\big\| \\ &\quad +\|x-\bar{x}\|^{-1}\|P(F(x+\xi^1)-F(x+\xi^2))-(PF')'(\bar{x};\ x-\bar{x})(\xi^1-\xi^2)\|\Big) \\ &\le \theta\|\xi^1-\xi^2\|, \qquad (18)\end{aligned}$$

where the first inequality follows from (8), and the last inequality follows from (10) and (11).

Similarly, from (14) and (15), it follows that for all $x \in U \setminus \{\bar{x}\}$ satisfying (13) the following relations hold:

$$\begin{aligned}\operatorname{dist}(0,\ \Phi_x(0)) = \inf&\left\{\|\xi\| \;\middle|\; \xi \in X,\ \Psi_2\left(\frac{x-\bar{x}}{\|x-\bar{x}\|}\right)\xi = (I_Y-P)(F(x)-F(\bar{x})) \right. \\ &\quad \left. + \|x-\bar{x}\|^{-1}(P(F(x)-F(\bar{x}))\right\} \\ &\le C(\|(I_Y-P)(F(x)-F(\bar{x}))\| + \|x-\bar{x}\|^{-1}\|P(F(x)-F(\bar{x}))\|) \\ &< (1-\theta)r(x). \qquad (19)\end{aligned}$$

Relations (18) and (19) (also taking into account (16)), imply that for all $x \in U \setminus \{\bar{x}\}$ satisfying (13) the mapping Φ_x satisfies in $B(0,\ r(x))$ all the assumptions of Theorem 3. Therefore, for each $x \in U \setminus \{\bar{x}\}$ satisfying (13), Φ_x has a fixed point in $B(0,\ r(x))$. Specifically, there exists an element $\psi(x) \in B(0,\ r(x))$ such that

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

By the definition of Φ_x, it now follows that

$$F(x + \psi(x)) = F(\bar{x}), \tag{20}$$

and, using the next to last inequality in (19),

$$\|\psi(x)\| \leq \frac{2C}{1-\theta}\left(\|(I_Y - P)(F(x) - F(\bar{x}))\| + \|x - \bar{x}\|^{-1}\|P(F(x) - F(\bar{x}))\|\right). \tag{21}$$

Obviously, (20) and (21) imply the estimate (9) with $M = 2(1-\theta)^{-1}C$ (for those $x \in (U \cap (\bar{x} + K)) \setminus \{\bar{x}\}$ that satisfy (13)).

Next we turn our attention to the second case, namely when $x \in (V \cap (\bar{x} + K)) \setminus \{\bar{x}\}$ is sufficiently close to $\bar{x}$ and

$$\frac{x - \bar{x}}{\|x - \bar{x}\|} \notin K_2^{\bar{\nu}}. \tag{22}$$

Using (6) and (7), (22) implies that one of the following two relations must hold:

$$\begin{aligned}\|(I_Y - P)(F(x) - F(\bar{x}))\| &= \|F'(\bar{x})(x - \bar{x})\| + o(\|x - \bar{x}\|) \\ &> \bar{\nu}\|x - \bar{x}\| + o(\|x - \bar{x}\|),\end{aligned}$$

or

$$\begin{aligned}\|P(F(x) - F(\bar{x}))\| &= \frac{1}{2}\|(PF')'(\bar{x};\ x - \bar{x})(x - \bar{x})\| + o(\|x - \bar{x}\|^2) \\ &> \frac{\bar{\nu}}{2}\|x - \bar{x}\|^2 + o(\|x - \bar{x}\|^2).\end{aligned}$$

Hence,

$$\|(I_Y - P)(F(x) - F(\bar{x}))\| + \|x - \bar{x}\|^{-1}\|P(F(x) - F(\bar{x}))\| > \frac{\bar{\nu}}{2}\|x - \bar{x}\| + o(\|x - \bar{x}\|),$$

and the desired estimate (9) is now obvious, as $\bar{x} \in F^{-1}(F(\bar{x}))$. In particular, we have proved that for any $M \in (2/\bar{\nu}, +\infty)$ there exists a neighborhood $U \subset V$ of $\bar{x}$ in X such that (9) holds under the additional assumption (22).

Combining the two cases (13) and (22) completes the proof.

□

Note that in the regular case, because of Remark 4 and since $P = 0$, Theorem 4 reduces to the classical error bound result given by Theorem 1.

Remark 7. The estimate (9) in Theorem 4 implies the following error bound (possibly with different M and U), which is in general somewhat weaker but may be more convenient in applications:

$$\operatorname{dist}(x,\ F^{-1}(F(\bar{x}))) \leq M\left(\|(I_Y - P)(F(x) - F(\bar{x}))\| + \|P(F(x) - F(\bar{x}))\|^{1/2}\right) \quad \forall\, x \in U \cap (\bar{x} + K)\,. \tag{23}$$

To verify that (23) follows from (9), denote $a := \operatorname{dist}(x,\ F^{-1}(F(\bar{x})))$, $b := M\|(I_Y - P)(F(x) - F(\bar{x}))\|$, and $c := M\|P(F(x) - F(\bar{x}))\|$. Since $\|x - \bar{x}\|^{-1} \leq a^{-1}$, from (9) we obtain that $a \le b + c/a$. Resolving the quadratic inequality in a, namely $a^2 - ba - c \le 0$, we obtain that $a \le b/2 + \sqrt{b^2/4 + c} \le b + \sqrt{c}$, from which the bound (23) follows, recalling definitions of the quantities involved.

To see that bound (23) can be weaker than (9), consider the case when $\bar{x}$ is not an isolated point of the set $F^{-1}(F(\bar{x}))$. Then $\mathrm{dist}(x,\ F^{-1}(F(\bar{x})))$ can tend to zero without $\|x - \bar{x}\|$ tending to zero, in which case (9) gives a linear error bound, while (23) contains also a term of order $1/2$. In the case when $\bar{x}$ is an isolated point of $F^{-1}(F(\bar{x}))$, the two error bounds can be regarded as equivalent.

Clearly, error bound (23) can be further simplified as follows:

$$\mathrm{dist}(x,\ F^{-1}(F(\bar{x}))) \le M\|F(x) - F(\bar{x})\|^{1/2} \quad \forall\, x \in U \cap (\bar{x} + K). \tag{24}$$

However, the latter bound is, in turn, weaker than (23). When $K = X$ and $F(\bar{x}) = 0$, we obtain a bound in the form (2) with $\gamma = 2$.

Applying Theorem 4 with the cone $K = \{th \mid t \in \mathbf{R}_+\}$, $h \in X$, we can establish the following generalization of the Graves-Lyusternik theorem describing the tangent directions to a level surface of a nonlinear mapping. This theorem was first obtained in [17] using another approach, not based on error bounds.

Theorem 5. *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 $h \in T_{F^{-1}(F(\bar{x}))}(\bar{x})$, 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 $h \in T_{F^{-1}(F(\bar{x}))}(\bar{x})$.

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

$$T_{F^{-1}(F(\bar{x}))}(\bar{x}) = T_2.$$

Proof. The proof of the necessary conditions of tangency (part (a)) is quite straightforward, and can be found in [17]. Sufficient conditions of tangency (part (b)) follow immediately from Theorem 4 (specifically, error bound (23)), Lemma 1, and the definition of the tangent cone.

□

By Remark 3, in the regular case we have that $T_2 = \ker F'(\bar{x})$, and so Theorem 5 takes the form of the classical Graves-Lyusternik theorem describing the tangent subspace (Theorem 2). Under the assumptions of twice Fréchet-differentiability of F at $\bar{x}$, Theorem 5 was proved in [39] (for the special case when $F'(\bar{x}) = 0$), and in [3] (without the latter assumption). Some earlier version of this result under even stronger smoothness assumptions can be found in [24]. For other related material see also [38, 8, 4–6, 18, 1, 16, 19].

Theorem 5 can be used to derive optimality conditions for 2-regular problems, including a special form of primal-dual conditions. We refer the reader to [17], where also an application to optimality conditions for mathematical programs with complementarity constraints is discussed.

3. A new error bound for the nonlinear complementarity problem

In this section, as an application of our general results, we obtain a new error bound for the classical nonlinear complementarity problem [30,11] (NCP). Given a (smooth) mapping $g : \mathbf{R}^m \to \mathbf{R}^m$, this problem consists of finding a point $x \in \mathbf{R}^m$ such that

$$g(x) \geq 0, \quad x \geq 0, \quad \langle g(x),\ x\rangle = 0. \tag{25}$$

While there exists a wide range of approaches to solving the NCP, one of the most successful and widely used is *reformulation* of the NCP as a system of equations or an optimization problem, see [27,25,20,13,10,37], the collection [14], and references therein. Among other things, such reformulations often lead to error bounds for NCP (i.e., estimates of the distance to its solution set) in terms of the objective function in optimization-based reformulations, or in terms of the equation residual in equation-based reformulations [13,31].

To our knowledge, the weakest condition under which a local error bound for NCP was previously available is *b-regularity*, introduced in [32]. Actually, there appears to be no explicit reference in the literature to an error bound for NCP under this assumption. However, the result is known to experts in the field, and it follows readily from some well-known facts, which we shall cite below.

Let $\bar{x}$ be a solution of the NCP, and define the three index sets

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

Definition 4. *A solution $\bar{x}$ of NCP is called* b-regular, *if for every $A \subset I_0$ the submatrix comprised by elements G_{ij}, $i, j \in A \cup I_1$, is nonsingular, where G is the Jacobian matrix of g at $\bar{x}$.*

By direct observation, it is easy to check that Definition 4 can be equivalently stated in the following form (which will be more convenient for our purposes): for any pair $(A,\ B)$ of index sets such that $A \cup B = I_0$, $A \cap B = \emptyset$, it holds that for $h \in \mathbf{R}^m$

$$\left.\begin{array}{ll} \langle g_i'(\bar{x}),\ h\rangle = 0, & i \in A \cup I_1 \\ h_i = 0, & i \in B \cup I_2 \end{array}\right\} \iff h = 0\,. \tag{26}$$

Or equivalently,

$$g_i'(\bar{x})\,,\ i \in A \cup I_1\,,\ e^i\,,\ i \in B \cup I_2 \text{ are linearly independent in } \mathbf{R}^m\,, \tag{27}$$

where $e^1, \dots e^m$ is the standard basis in $\mathbf{R}^m$, and the index sets A and B vary as specified above.

If $\bar{x}$ is a b-regular solution of NCP, then it is known that it is locally unique. Furthermore, the norm of the *natural residual*

$$S : \mathbf{R}^m \to \mathbf{R}^m, \quad S_i(x) = \min\{x_i,\ g_i(x)\},\ i = 1, \dots, m,$$

provides a local error bound: there exist a constant $M > 0$ and a neighborhood U of $\bar{x}$ such that

$$\|x - \bar{x}\| \leq M\|S(x)\| \quad \forall\, x \in U. \tag{28}$$

This error bound follows from the following facts.

1. Let $Q(\bar{x}) = 0$, and suppose that Q is semismooth [29,34] at $\bar{x}$, and $\bar{x}$ is BD-regular (as defined in [34]). Then the equation residual provides a local error bound [33, Proposition 3], i.e., $\|x - \bar{x}\| \leq M\|Q(x)\|$ for all x in some neighborhood of $\bar{x}$ and some $M > 0$.
2. If $\bar{x}$ is a b-regular solution of NCP, then $\bar{x}$ is a BD-regular solution of $S(x) = 0$ [21, Proposition 3.4](see also [22, Proposition 2.10]).
3. If $g(\cdot)$ is continuously differentiable then the mapping $S(\cdot)$ is semismooth (e.g., see [22, Proposition 2.7]).

In this section, we shall obtain an error bound for NCP under an assumption weaker than b-regularity. Specifically, we shall derive an error bound using the 2-regularity condition studied in Sect. 2, and show that it is implied by b-regularity, but not vice versa. Our error bound will be based on the following $C^{1,1}$-equation-based reformulation of NCP. Define the mapping

$$F : \mathbf{R}^m \to \mathbf{R}^m, \quad F_i(x) = 2g_i(x)x_i - (\min\{0,\, g_i(x) + x_i\})^2,\ i = 1, \ldots, m.$$

It is known [20], and easy to check, that solution set of NCP coincides with the set

$$D = \{x \in \mathbf{R}^m \mid F(x) = 0\}. \tag{29}$$

We believe that our analysis can also be applied to some other appropriate $C^{1,1}$ reformulations, such as in [27,20], but we shall consider only one of them for the clarity of presentation.

We assume that g is differentiable on some neighborhood V of $\bar{x}$ in $\mathbf{R}^m$, its derivative is Lipschitzian on V, and g is twice differentiable at $\bar{x}$. Then it is easy to see that F satisfies all the smoothness assumptions of Theorem 4 (with $K = \mathbf{R}^m$). Hence, an error bound for NCP in terms of the residual function $\|F(\cdot)\|$ follows immediately from Theorem 4, whenever F is 2-regular at $\bar{x}$ (recall that in the finite-dimensional setting 2-regularity is equivalent to strong 2-regularity). It remains to understand what 2-regularity means in the context of this section. Let P and T_2 be as defined in Sect. 2. Recall also Definition 2 and (5). Observe that if $h \in T_2$, then necessarily $h \in \ker \Psi_2(h)$ (this follows directly from the definitions). Since in the case under consideration the linear operator $\Psi_2(h)$ is from $\mathbf{R}^m$ to $\mathbf{R}^m$, if $h \in T_2$ and $h \neq 0$ then $\ker \Psi_2(h) \neq \{0\}$, and hence, $\operatorname{im} \Psi_2(h) \neq \mathbf{R}^m$. It follows that F cannot be 2-regular with respect to any $h \in T_2 \setminus \{0\}$. On the other hand, if

$$T_2 = \{0\}, \tag{30}$$

then Definition 2 of 2-regularity is always (trivially) satisfied. Hence, the mapping F under consideration is 2-regular at $\bar{x}$ if, and only if, (30) holds.

We are now ready to state our error bound for NCP. Afterwards, we shall compare our new regularity condition (30) with the condition of b-regularity defined above. In particular, we shall establish that the former implies the latter, and that the converse is not true. Some examples will be given to illustrate those facts.

Note that if (30) holds then $\bar{x}$ is an isolated point of the set D (this follows from Lemma 1 and compactness property of the unit sphere in $\mathbf{R}^m$). Applying now Theorem 4 (see also Remark 7), we readily obtain the following error bound.

Theorem 6. *Let V be a neighborhood of a point $\bar{x}$ in $\mathbf{R}^m$, which is a solution of NCP. Let $g : V \to \mathbf{R}^m$ be differentiable on V, its derivative be Lipschitz continuous on V, and let g be twice differentiable at $\bar{x}$. Suppose further that F is 2-regular at $\bar{x}$ (i.e., condition* (30) *is satisfied).*

Then there exist a neighborhood U of $\bar{x}$ in $\mathbf{R}^m$ and a number $M > 0$ such that

$$\|x - \bar{x}\| \leq M(\|(I - P)F(x)\| + \|PF(x)\|^{1/2}) \quad \forall\, x \in U. \tag{31}$$

Note that according to Remark 7, we can replace estimate (31) by a weaker one:

$$\|x - \bar{x}\| \leq M\|F(x)\|^{1/2} \quad \forall\, x \in U. \tag{32}$$

We proceed to establish the relationship between 2-regularity of F at $\bar{x}$ and b-regularity.

The first derivative of F is given by

$$F_i'(x) = 2\big(x_i g_i'(x) + g_i(x)e^i - \min\{0,\ g_i(x) + x_i\}\big(g_i'(x) + e^i\big)\big),$$
$$i = 1, \ldots, m, \quad x \in V,$$

where $e^1, \ldots e^m$ is again the standard basis in $\mathbf{R}^m$. As before, let $\bar{x}$ be a solution of NCP (equivalently, $\bar{x} \in D$), and let the index sets I_0 , I_1 and I_2 be as defined above. With this notation, we have that

$$F_i'(\bar{x}) = 2\begin{cases} 0, & \text{if } i \in I_0, \\ \bar{x}_i g_i'(\bar{x}), & \text{if } i \in I_1, \\ g_i(\bar{x})e^i, & \text{if } i \in I_2. \end{cases} \tag{33}$$

Observe that F cannot be regular at $\bar{x}$ when $I_0 \neq \emptyset$. Among other things, this implies that an error bound in terms of $\|F(\cdot)\|$ could not have been obtained using the classical results.

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

$$\ker F'(\bar{x}) = \left\{ h \in \mathbf{R}^m \;\middle|\; \begin{array}{l} \langle g_i'(\bar{x}),\ h\rangle = 0,\ i \in I_1, \\ h_i = 0,\ i \in I_2 \end{array} \right\}. \tag{34}$$

By (33), it is further evident that

$$\operatorname{im} F'(\bar{x}) \subset \operatorname{span}\left\{e^i \mid i \in I_1 \cup I_2\right\}. \tag{35}$$

By the directional differentiability of F', we have that for any $h \in \mathbf{R}^m$

$$(PF')'(\bar{x};\ h)h = 0 \Longleftrightarrow (F')'(\bar{x};\ h)h \in \operatorname{im} F'(\bar{x}) \Longrightarrow \left(F_i'\right)'(\bar{x};\ h)h = 0 \quad \forall\, i \in I_0,$$

where the last implication follows from (35). Hence, by the definition of T_2,

$$T_2 \subset Q := \left\{ h \in \ker F'(\bar{x}) \,\middle|\, \left(F_i'\right)'(\bar{x};\, h)h = 0\,,\, i \in I_0 \right\}. \tag{36}$$

Hence, a *sufficient* condition for 2-regularity of F at $\bar{x}$ is $Q = \{0\}$. Note that under the assumption that

$$g_i'(\bar{x})\,,\; i \in I_1\,,\; e^i\,,\; i \in I_2 \text{ are linearly independent in } \mathbf{R}^m\,, \tag{37}$$

the inclusions in (35) and (36) hold as equalities. In that case, the condition $Q = \{0\}$ is also necessary for 2-regularity of F at $\bar{x}$. However, we point out that the linear independence assumption (37) is not necessary for the reformulation of NCP to have 2-regularity properties (see Example 1 below). On the other hand, it simplifies a comparison with the b-regularity condition (27), which itself certainly subsumes (37).

By direct calculations, we obtain the following formula for the directional derivatives of $F_i'(\cdot) : V \to \mathbf{R}^m$, $i \in I_0$:

$$\begin{aligned}\left(F_i'\right)'(\bar{x};\; h) = 2\big(h_i - \min\left\{0,\; \langle g_i'(\bar{x}),\; h\rangle + h_i\right\}\big)g_i'(\bar{x})\\ +2\big(\langle g_i'(\bar{x}),\; h\rangle - \min\left\{0,\; \langle g_i'(\bar{x}),\; h\rangle + h_i\right\}\big)e^i,\\ i \in I_0, \quad h \in \mathbf{R}^m.\end{aligned}$$

Using the latter formula, and taking also into account (34), it is easy to verify that

$$Q = \left\{ h \in \mathbf{R}^m \,\middle|\, \begin{array}{l} \min\left\{\langle g_i'(\bar{x}),\; h\rangle,\; h_i\right\} = 0,\; i \in I_0, \\ \langle g_i'(\bar{x}),\; h\rangle = 0,\; i \in I_1, \\ h_i = 0,\; i \in I_2. \end{array} \right\}.$$

It is now clear that the condition $Q = \{0\}$, which is sufficient for 2-regularity of F at $\bar{x}$, can be stated as follows: for any pair of index sets (A, B) such that $A \cup B = I_0$, $A \cap B = \emptyset$, it holds that for $h \in \mathbf{R}^m$

$$\left.\begin{array}{l} \langle g_i'(\bar{x}),\; h\rangle = 0,\; h_i \geq 0,\; i \in A \\ \langle g_i'(\bar{x}),\; h\rangle \geq 0,\; h_i = 0,\; i \in B \\ \langle g_i'(\bar{x}),\; h\rangle = 0,\; i \in I_1 \\ h_i = 0,\; i \in I_2 \end{array}\right\} \quad \Longleftrightarrow \quad h = 0\,. \tag{38}$$

By direct comparison of (26) and (38), we conclude that the b-regularity of $\bar{x}$ always implies 2-regularity of F at $\bar{x}$. We have therefore established the following.

Proposition 1. *Let V be a neighborhood of a point $\bar{x}$ in $\mathbf{R}^m$, and suppose that $g : V \to \mathbf{R}^m$ is differentiable on V with its derivative being Lipschitz continuous on V, and g is twice differentiable at $\bar{x}$.*

If $\bar{x}$ is a b-regular solution of NCP, then F is 2-regular at $\bar{x}$.

Combining Proposition 1 and Theorem 6, we obtain, as a corollary, a local error bound for NCP in terms of $\|F(\cdot)\|$ under the assumption of b-regularity of $\bar{x}$. We note that this bound also follows from different considerations. Indeed, as discussed above, an error bound under the assumption of b-regularity is known to hold in terms of the natural residual $\|S(\cdot)\|$. Hence, it also holds in terms of any other residual whose local growth rate is at least of the same order. We refer the reader to [40], where some relevant comparisons can be found. In particular, it holds that $\|F(x)\|^{1/2} \geq \|S(x)\|$ for all $x \in \mathbf{R}^m$. Hence, if $\bar{x}$ is b-regular then error bound (32) follows from (28) (but note that the sharper bound (31) does not readily follow from (28)). In any case, we stress once again that our general result (Theorem 6) is applicable beyond the case of b-regularity of $\bar{x}$.

We could also state a corollary of Theorem 6, where the projector P is evaluated explicitly, at least under the assumption (37). Let $F_{I_1,\, I_2}(\cdot)$ and $F_{I_0}(\cdot)$ denote the mappings with components $F_i(\cdot)$, where $i \in I_1 \cup I_2$ and $i \in I_0$, respectively. Under the hypotheses of Theorem 6, and the additional assumption (37), error bound (31) takes the form

$$\|x - \bar{x}\| \leq M\big(\|F_{I_1,\, I_2}(x)\| + \|F_{I_0}(x)\|^{1/2}\big) \quad \forall\, x \in U. \tag{39}$$

As before, the above bound can be replaced by a weaker bound (32), which does not involve projectors.

The following two examples show that 2-regularity is indeed *weaker* than b-regularity. The first example demonstrates that 2-regularity condition (30) can hold even without the linear independence assumption (37), in which case b-regularity cannot hold. The second example illustrates that even if the linear independence assumption (37) holds, it is still possible to have 2-regularity without b-regularity. The two examples therefore indicate that 2-regularity is a significantly less restrictive assumption than b-regularity.

Example 1. Let $m = 2$ and consider NCP associated with the function

$$g(x) = \big(x_1,\ -(x_2 - a)^2\big), \quad x \in \mathbf{R}^2,$$

with $a > 0$ being any given number. Consider the point $\bar{x} = (0,\, a)$, which clearly is a solution of NCP. Obviously, $I_0 = \{1\}$, $I_1 = \{2\}$, $I_2 = \emptyset$, and $g_2'(\bar{x}) = 0$, so (37) is violated. In particular, $\bar{x}$ is not a b-regular solution of NCP. Nevertheless, it is not difficult to evaluate T_2 here. We have: $F'(\bar{x}) = 0$, hence $\ker F'(\bar{x}) = \mathbf{R}^2$, $Y_1 = \{0\}$, $Y_2 = \mathbf{R}^2$, P is the identity operator in $\mathbf{R}^2$. By direct computation, we obtain that

$$\big(F_1'\big)'(\bar{x};\ h) = 4h_1(h_1 - \min\{0,\ 2h_1\}), \quad h \in \mathbf{R}^2.$$

Furthermore, $F_2'(\cdot)$ is smooth at $\bar{x}$, and

$$F_2''(\bar{x})[h,\ h] = 8ah_2^2, \quad h \in \mathbf{R}^2.$$

It is now obvious that (30) holds, and so F is 2-regular at $\bar{x}$. It is further easy to see that the error bound (32) is valid.

Note that error bound (28) using the natural residual S does not hold in this case. Indeed, consider the sequence of points $\{x^k\}$, where $x_1^k = 0$ and $x_2^k = a - 1/k$, $k = 1, 2, \ldots$. Then $\|x^k - \bar{x}\| = 1/k$, while $\|S(x^k)\| = 1/k^2$.

Before stating our second example, let us consider yet another characterization of 2-regularity (under the additional assumption (37)). Notice that (38) is equivalent to saying that there exist no $q \in \mathbf{R}^m$ such that the following system has a solution in $h \in \mathbf{R}^m$:

$$\begin{aligned}
&\langle q, h\rangle > 0,\\
&\left\langle g_i'(\bar{x}),\ h\right\rangle = 0,\ h_i \geq 0,\ i \in A,\\
&\left\langle g_i'(\bar{x}),\ h\right\rangle \geq 0,\ h_i = 0,\ i \in B,\\
&\left\langle g_i'(\bar{x}),\ h\right\rangle = 0,\ i \in I_1,\\
&h_i = 0,\ i \in I_2.
\end{aligned}$$

By the Motzkin theorem of the alternatives [26], the latter is equivalent to the following system (in λ, μ) having a solution for all $q \in \mathbf{R}^m$:

$$\begin{aligned}
&\sum_{i\in A} \lambda_i e^i + \sum_{i\in B} \lambda_i g_i'(\bar{x}) + \sum_{i\in A\cup I_1} \mu_i g_i'(\bar{x}) + \sum_{i\in B\cup I_2} \mu_i e^i = q,\\
&\lambda_i \geq 0\,, i \in A \cup B = I_0.
\end{aligned}$$

In other words,

$$\begin{aligned}
\mathbf{R}^m = {} & \operatorname{cone}\{e^i \mid i \in A\} + \operatorname{cone}\left\{g_i'(\bar{x}) \middle| i \in B\right\}\\
& + \operatorname{span}\left\{g_i'(\bar{x}) \middle| i \in A \cup I_1\right\} + \operatorname{span}\left\{e^i \middle| i \in B \cup I_2\right\}.
\end{aligned} \tag{40}$$

Note that this is another way to see that b-regularity implies 2-regularity. Indeed, (27) clearly guarantees (40), regardless of the cones in the right-hand side. On the other hand, if b-regularity does not hold, the last two terms in (40) span a proper subspace of $\mathbf{R}^m$. But it is still possible that the sum of this subspace and the cone given by the first two terms in the right-hand side of (40), gives the whole space $\mathbf{R}^m$. The following example illustrates this case.

Example 2. Let $m = 2$ and consider NCP associated with the function

$$g(x) = (-x_2,\ -x_1), \quad x \in \mathbf{R}^2.$$

Clearly, $\bar{x} = (0, 0)$ is a solution of this NCP. We have that $I_0 = \{1, 2\}$, $g_1'(\bar{x}) = (0, -1)$, and $g_2'(\bar{x}) = (-1, 0)$. Since $I_1 = I_2 = \emptyset$, the linear independence assumption (37) is satisfied. Take $A = \{1\} \subset I_0$, and consider the submatrix G_{ij} , i, $j \in A \cup I_1$, where G is the Jacobian matrix of g at $\bar{x}$. This submatrix consists of one element, which is $(g_1'(\bar{x}))_1 = 0$. Hence, $\bar{x}$ is not a b-regular solution.

However, F is 2-regular at $\bar{x}$. Indeed, under the condition (37) 2-regularity is equivalent to (40). We have to consider the four possible cases in choosing index sets A and B.

If $B = \{1, 2\}$, or $A = \{1, 2\}$, then it is easy to see that $\operatorname{span}\{e^i \mid i \in B\} = \mathbf{R}^2$, or $\operatorname{span}\{g_i'(\bar{x}) \mid i \in A\} = \mathbf{R}^2$, respectively.

Consider now the case where $A = \{1\}$, $B = \{2\}$. In that case, we have

$$\operatorname{cone}\{(1, 0)\} + \operatorname{cone}\{(-1, 0)\} + \operatorname{span}\{(0, -1)\} + \operatorname{span}\{(0, 1)\} = \mathbf{R}^2,$$

and notice that the conic part is crucial for the equality to hold.

The case $A = \{2\}$, $B = \{1\}$ is similar:

$$\operatorname{cone}\{(0, 1)\} + \operatorname{cone}\{(0, -1)\} + \operatorname{span}\{(-1, 0)\} + \operatorname{span}\{(1, 0)\} = \mathbf{R}^2.$$

Observe that g in Examples 1 and 2 is an analytic function, and so local error bounds of some sort can also be obtained using the results of [23]. But note that in the setting of this section 2-regularity of F at $\bar{x}$ obviously is not affected if we add to g any function which is zero at $\bar{x}$, and whose first two derivatives vanish at $\bar{x}$. Now, if this function is not infinitely differentiable (say, the third derivative does not exist), then it is not analytic. If we modify g in Examples 1 and 2 in this way, all our results still apply, while the NCP function is no longer analytic.

4. Convergence rate of penalty methods

For the sake of simplicity, we shall again consider the finite-dimensional setting only. Let V be an open set in $\mathbf{R}^n$, $F : V \to \mathbf{R}^m$ be a given mapping. Consider an equality-constrained optimization problem

$$\begin{array}{l} \text{minimize } \ f(x) \\ \text{subject to } x \in D, \end{array} \tag{41}$$

where the (feasible) set D is defined in (1) and $f : V \to \mathbf{R}$ is a given (objective) function.

Similar to the standard exterior penalty framework [12, 7, 28], define the family of functions

$$\varphi_{\alpha,\,\beta} : V \to \mathbf{R}, \quad \varphi_{\alpha,\,\beta}(x) = f(x) + \alpha \|F(x)\|^{\beta}, \tag{42}$$

where $\alpha > 0$ is the penalty parameter, and $\beta > 0$ is the power of the penalty function. For simplicity, we shall assume β to be fixed. Consider the family of (unconstrained) optimization problems

$$\begin{array}{l} \text{minimize } \varphi_{\alpha,\,\beta}(x) \\ \text{subject to } x \in \mathbf{R}^n, \end{array}$$

and an associated trajectory of (possibly inexact) solutions of those problems. Specifically, we assume that trajectory $x(\cdot) : \mathbf{R}_+ \to V$ satisfies the following conditions:

$$\varphi_{\alpha,\,\beta}(\bar{x}) \geq \varphi_{\alpha,\,\beta}(x(\alpha)) \quad \forall\, \alpha > 0 \text{ sufficiently large}, \tag{43}$$

$$x(\alpha) \to \bar{x} \text{ as } \alpha \to +\infty, \tag{44}$$

where $\bar{x}$ is a local solution of problem (41), (1). We note that property (43) is standard in exterior penalty methods [28] (for example, it is trivially satisfied if $x(\alpha)$ is a global minimizer of $\varphi_{\alpha,\,\beta}(x)$ over any set containing $\bar{x}$). Condition (44) is a natural assumption in the context of estimating local *rate* of convergence of the algorithm. Some conditions under which it holds can be found in [7, Proposition 2.2] ((44) is particularly easy to guarantee if solutions of (41), (1) are isolated).

We assume $x(\cdot)$ satisfying (43), (44) to be *given*, and our goal is to estimate the rate of convergence in some way. Error bound (2) is the key for such an estimation. We start with a general rate of convergence result, which perhaps should not be considered conceptually original (although, we could not find it in this form in the literature). We include it here to make precise the subsequent application of our error bound to obtain a convergence rate result for a penalty method in the irregular case.

Theorem 7. *Let V be a neighborhood of a point $\bar{x}$ in $\mathbf{R}^n$. Let a function $f : V \to \mathbf{R}$ be Lipschitz continuous on V with a constant $l > 0$, a mapping $F : V \to \mathbf{R}^m$ be continuous on V, and let $\bar{x}$ be a local solution of problem (41), (1). Suppose further that there exist a neighborhood $U \subset V$ of $\bar{x}$ in $\mathbf{R}^n$ and numbers $M > 0$ and $\gamma > 0$ such that (2) is satisfied.*

Let $x(\cdot) : \mathbf{R}_+ \to V$ be a trajectory of the penalty method defined by (42)–(44), with $\beta > 0$ fixed. Then the following statements hold.

(a) If $\beta\gamma \leq 1$, then

$$x(\alpha) \in D, \quad f(x(\alpha)) = f(\bar{x}) \quad \forall\, \alpha > 0 \text{ sufficiently large}. \tag{45}$$

In particular, if $\bar{x}$ is an isolated local solution of problem (41), (1), then

$$x(\alpha) = \bar{x} \quad \forall\, \alpha > 0 \text{ sufficiently large}.$$

(b) If $\beta\gamma > 1$, then

$$0 \leq f(\bar{x}) - f(x(\alpha)) \leq \frac{(Ml)^{\beta\gamma/(\beta\gamma-1)}}{\alpha^{1/(\beta\gamma-1)}} \quad \forall\, \alpha > 0 \text{ sufficiently large}.$$

Proof. By the inclusion $\bar{x} \in D$, (43), (44) and (2), we obtain that for any $\alpha > 0$ sufficiently large

$$\begin{aligned} f(\bar{x}) &= f(\bar{x}) + \alpha \|F(\bar{x})\|^\beta \\ &= \varphi_{\alpha,\,\beta}(\bar{x}) \\ &\geq \varphi_{\alpha,\,\beta}(x(\alpha)) \\ &= f(x(\alpha)) + \alpha \|F(x(\alpha))\|^\beta \\ &\geq f(x(\alpha)) + \frac{\alpha}{M^{\beta\gamma}} (\mathrm{dist}(x(\alpha),\, D))^{\beta\gamma}. \end{aligned} \tag{46}$$

Since F is continuous on V and X is finite-dimensional, there exists a (possibly not unique) metric projection $\pi_D(x(\alpha))$ of $x(\alpha)$ on D. Furthermore, $\bar{x} \in D$ and (44) imply that

$$\pi_D(x(\alpha)) \to \bar{x} \text{ as } \alpha \to +\infty.$$

Hence, as $\bar{x}$ is a local solution of problem (41) and f is Lipschitz continuous on V, we conclude that

$$\begin{aligned} f(\bar{x}) - f(x(\alpha)) &\leq f(\pi_D(x(\alpha))) - f(x(\alpha)) \\ &\leq l\, \mathrm{dist}(x(\alpha),\, D). \end{aligned} \tag{47}$$

Combining (46) and (47), we further obtain

$$f(\bar{x}) - f(x(\alpha)) \geq \frac{\alpha}{(Ml)^{\beta\gamma}} (f(\bar{x}) - f(x(\alpha)))^{\beta\gamma} \quad \forall\, \alpha > 0 \text{ sufficiently large}. \tag{48}$$

Define

$$\mathcal{A} := \{\alpha > 0 \mid f(\bar{x}) \neq f(x(\alpha))\}.$$

Relation (48), together with (46), imply that

$$0 \le (f(\bar{x}) - f(x(\alpha)))^{\beta\gamma-1} \le \frac{(Ml)^{\beta\gamma}}{\alpha} \quad \forall \alpha \in \mathcal{A} \text{ sufficiently large.} \tag{49}$$

Let $\beta\gamma \le 1$. Then, by (44), the left-hand side of (49) does not tend to 0 as $\alpha \to +\infty$, while the right-hand side does. This means that $\mathcal{A}$ is necessarily bounded in that case, i.e., $f(\bar{x}) = f(x(\alpha))$ for all α large enough. Furthermore, by the third equality in (46), we have that $f(\bar{x}) - f(x(\alpha)) \ge \alpha\|F(x(\alpha))\|^{\beta}$, and it follows that $x(\alpha)$ is also feasible whenever $f(\bar{x}) = f(x(\alpha))$. Assertion (a) is established.

Assertion (b) follows immediately from (49) for $\alpha \in \mathcal{A}$, and it is trivial for $\alpha \notin \mathcal{A}$.

□

According to Theorem 7, decreasing parameter β increases the guaranteed convergence rate of the penalty method (in fact, exact solution is obtained for all values of $\beta > 0$ sufficiently small). On the other hand, decreasing β negatively affects smoothness of the penalty function, which makes the problem less tractable numerically. The estimate obtained in Theorem 7 gives an opportunity to find a compromise between theoretical rate of convergence of the method and numerical issues related to smoothness properties of subproblems, provided γ is available. This is precisely the case when there are reasons to expect F to be regular or 2-regular at $\bar{x}$.

We next discuss applications of Theorem 7 to problems with regular and 2-regular constraints.

Assume that $F : V \to \mathbf{R}^m$ is differentiable on V and regular at $\bar{x}$, and its derivative is continuous at $\bar{x}$. Then Theorem 1 guarantees that error bound (2) holds with $\gamma = 1$. Hence, by Theorem 7, in that case

(a) If $\beta \le 1$, then (45) holds.
(b) If $\beta > 1$, then

$$0 \le f(\bar{x}) - f(x(\alpha)) \le \frac{(Ml)^{\beta/(\beta-1)}}{\alpha^{1/(\beta-1)}} \quad \forall \alpha > 0 \text{ sufficiently large.}$$

Note that this result could also be obtained combining Theorem 7 with Theorem 4, as the latter subsumes Theorem 1.

Consider now the 2-regular case. Specifically, assume that $F : V \to \mathbf{R}^m$ is differentiable on V, its derivative is continuous at $\bar{x}$, and the mapping $PF'(\cdot)$ is Lipschitz continuous on V and B-differentiable at $\bar{x}$, where P is the projector in $\mathbf{R}^m$ onto complementary subspace of $\operatorname{im} F'(\bar{x})$ parallel to $\operatorname{im} F'(\bar{x})$. Assume that the mapping F is 2-regular at $\bar{x}$ (note that, by Remark 5, in the finite-dimensional case this is equivalent to strong 2-regularity). Then Theorem 4 (recall also Remark 7) guarantees that error bound (2) holds with $\gamma = 2$. Hence, by Theorem 7, we have that

(a) If $\beta \le 1/2$, then (45) holds.
(b) If $\beta > 1/2$, then

$$0 \le f(\bar{x}) - f(x(\alpha)) \le \frac{(Ml)^{2\beta/(2\beta-1)}}{\alpha^{1/(2\beta-1)}} \quad \forall \alpha > 0 \text{ sufficiently large.}$$

Note that the case of $\beta = 2$ is of special interest as this is the smallest value of β for which the penalty function is smooth, provided F is smooth. For this value of β, in the regular case we obtain the estimate with the term in the right-hand side of the order $O(\alpha^{-1})$. In the 2-regular case, the term in the right-hand side is of the order $O(\alpha^{-1/3})$, and this estimate cannot be improved, as illustrated by the following example.

Example 3. Let $n = m = 1$, $V = \mathbf{R}$,

$$f(x) = x, \quad F(x) = x^2, \quad x \in \mathbf{R},$$

and consider the point $\bar{x} = 0$. It is easy to check that F is 2-regular at $\bar{x}$. For every $\alpha > 0$, the penalty function $\varphi_{\alpha,\,2}$ has unique critical point $x(\alpha) = -(4\alpha)^{-1/3}$ (which is in fact its global minimizer), and we have that $f(\bar{x}) - f(x(\alpha)) = (4\alpha)^{-1/3}$.

Acknowledgements. 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.

References

1. Arutyunov, A.V. (1997): Extremum Conditions. Abnormal and Degenerate Cases. Factorial, Moscow, Russia. In Russian
2. Aubin, J.-P., Frankowska, H. (1990): Set-Valued Analysis. Birkhäuser, Boston
3. Avakov, E.R. (1985): Extremum conditions for smooth problems with equality-type constraints. USSR Comput. Math. Math. Phys. **25**, 24–32
4. Avakov, E.R. (1989): Necessary extremum conditions for smooth abnormal problems with equality- and inequality constraints. Math. Notes **45**, 431–437
5. Avakov, E.R. (1990): Theorems on estimates in the neighborhood of a singular point of a mapping. Math. Notes **47**, 425–432
6. Avakov, E.R., Agrachyov, A.A., Arutyunov, A.V. (1992): The level-set of a smooth mapping in a neighborhood of a singular point, and zeros of a quadratic mapping. Math. USSR Sbornik **73**, 455–466
7. Bertsekas, D.P. (1982): Constrained Optimization and Lagrange Multiplier Methods. Academic Press, New York
8. Buchner, M., Marsden, J., Schecter, S. (1983): Applications of the blowing-up construction and algebraic geometry to bifurcation problems. J. Diff. Equations **48**, 404–433
9. Dmitruk, A.V., Milyutin, A.A., Osmolovskii, N.P. (1980): Lyusternik's theorem and the theory of extrema. Russ. Math. Surveys **35**, 11–51
10. Facchinei, F., Soares, J. (1997): A new merit function for nonlinear complementarity problems and a related algorithm. SIAM J. Optim. **7**, 225–247
11. Ferris, M.C., Pang, J.-S., eds. (1997): Complementarity and variational problems: State of the Art. SIAM Publications
12. Fiacco, A.V., McCormick, G.P. (1968): Nonlinear Programming: Sequential Unconstrained Minimization Techniques. John Wiley & Sons, New York
13. Fukushima, M. (1996): Merit functions for variational inequality and complementarity problems. In: Di Pillo, G., Giannessi, F., eds., Nonlinear Optimization and Applications, pp. 155–170. Plenum Publishing Corporation, New York, New York
14. Fukushima, M., Qi, L., eds. (1999): Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods. Kluwer Academic Publishers
15. Ioffe, A.D., Tikhomirov, V.M. (1974): Theory of Extremal Problems. North–Holland, Amsterdam, Holland
16. Izmailov, A.F. (1998): On certain generalizations of Morse's lemma. Proceedings of the Steklov Institute of Mathematics **220**, 138–153
17. Izmailov, A.F., Solodov, M.V. (2000): The theory of 2-regularity for mappings with Lipschitzian derivatives and its applications to optimality conditions. Manuscript, January 2000
18. Izmailov, A.F., Tretyakov, A.A. (1994): Factor-Analysis of Nonlinear Mappings. Nauka, Moscow. In Russian

19. Izmailov, A.F., Tretyakov, A.A. (1999): 2-Regular Solutions of Nonlinear Problems. Theory and Numerical Methods. Fizmatlit, Moscow. In Russian
20. Kanzow, C. (1994): Some equation-based methods for the nonlinear complementarity problem. Optim. Methods Software **3**, 327–340
21. Kanzow, C., Fukushima, M. (1998): Solving box constrained variational inequality problems by using the natural residual with D-gap function globalization. Oper. Res. Lett. **23**, 45–51
22. De Luca, T., Facchinei, F., Kanzow, C. (2000): A theoretical and numerical comparison of some semismooth algorithms for complementarity problems. Comput. Optim. Appl. **16**, 173–205
23. Luo, Z.-Q., Pang, J.-S. (1994): Error bounds for analytic systems and their applications. Math. Program. **67**, 1–28
24. Magnus, R.J. (1976): On the local structure of the zero-set of a Banach space valued mapping. J. Funct. Anal. **22**, 58–72
25. Mangasarian, O.L., Solodov, M.V. (1993): Nonlinear complementarity as unconstrained and constrained minimization. Math. Program. **62**, 277–297
26. Mangasarian, O.L. (1969): Nonlinear Programming. McGraw–Hill, New York
27. Mangasarian, O.L. (1976): Equivalence of the complementarity problem to a system of nonlinear equations. SIAM J. Appl. Math. **31**, 89–92
28. Mangasarian, O.L. (1986): Some applications of penalty functions in mathematical programming. In: Conti, R., De Giorgi, E., Giannessi, F., eds., Optimization and Related Fields. pp. 307–329. Springer, Heidelberg, Lecture Notes in Mathematics 1190
29. Mifflin, R. (1977): Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**, 957–972
30. Pang, J.-S. (1995): Complementarity problems. In: Horst, R., Pardalos, P., eds., Handbook of Global Optimization. pp. 271–338. Kluwer Academic Publishers, Boston, Massachusetts
31. Pang, J.-S. (1997): Error bounds in mathematical programming. Math. Program. **79**, 299–332
32. Pang, J.-S., Gabriel, S.A. (1993): NE/SQP: A robust algorithm for the nonlinear complementarity problem. Math. Program. **60**, 295–337
33. Pang, J.-S., Qi, L. (1993): Nonsmooth equations: motivation and algorithms. SIAM J. Optim. **3**, 443–465
34. Qi, L. (1993): Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**, 227–244
35. Robinson, S.M. (1987): Local structure of feasible sets in nonlinear programming, part III: stability and sensitivity. Math. Program. Study **30**, 45–66
36. Shapiro, A. (1990): On concepts of directional differentiability. J. Optim. Theory Appl. **66**, 477–487
37. Solodov, M.V. (2000): Implicit Lagrangian. In: Floudas, C., Pardalos, P., eds., Encyclopedia of Optimization. Kluwer Academic Publishers
38. Szulkin, A. (1979): Local structure of the zero-sets of differentiable mappings and applications to bifurcation theory. Math. Scandinavica **45**, 232–242
39. Tretyakov, A.A. (1984): Necessary and sufficient conditions for optimality of p-th order. USSR Comput. Math. Math. Phys. **24**, 123–127
40. Tseng, P. (1996): Growth behavior of a class of merit functions for the nonlinear complementarity problem. J. Optim. Theory Appl. **89**, 17–37