SIAM J. OPTIMIZATION
Vol. xx, No. x, pp. xx–xx, 200x

000

SUPERLINEARLY CONVERGENT ALGORITHMS FOR SOLVING SINGULAR EQUATIONS AND SMOOTH REFORMULATIONS OF COMPLEMENTARITY PROBLEMS

A. F. IZMAILOV * AND M. V. SOLODOV †

Abstract. We propose a new algorithm for solving smooth nonlinear equations in the case where their solutions can be singular. Compared to other techniques for computing singular solutions, a distinctive feature of our approach is that we do not employ second derivatives of the equation mapping in the algorithm, and do not assume their existence in the convergence analysis. Important examples of once but not twice differentiable equations whose solutions are inherently singular, are smooth equation-based reformulations of the nonlinear complementarity problems. Reformulations of complementarity problems serve both as illustration and motivation for our approach, and one of them we consider in detail. We show that the proposed method possesses local superlinear/quadratic convergence under reasonable assumptions. We further demonstrate that these assumptions are in general not weaker and not stronger than regularity conditions employed in the context of other superlinearly convergent Newton-type algorithms for solving complementarity problems, which are typically based on nonsmooth reformulations. Therefore our approach appears to be an interesting complement to the existing ones.

Key words. nonlinear equations, singularity, regularity, complementarity, reformulation, superlinear convergence.

AMS subject classifications. 90C30, 46T20, 47J07, 90C33

1. Introduction.

In this paper we are interested in solving nonlinear equations in the case where their solutions can be singular, and smoothness requirements are weaker than those usually assumed in this context. Our development is partially motivated by the nonlinear complementarity problem, which we consider in detail, and for which our method takes a particularly simple and readily implementable form.

Let $F : V \to \mathbf{R}^n$ be a given mapping, where V is a neighborhood of a point $\bar{x}$ in $\mathbf{R}^n$, with $\bar{x}$ being a solution of the system of equations

$$F(x) = 0. \tag{1.1}$$

In the sequel, F is assumed to be once (but not necessarily twice) differentiable on V. In this setting $\bar{x}$ is referred to as *singular solution* if the linear operator $F'(\bar{x})$ is singular, i.e.,

$$\det F'(\bar{x}) = 0,$$

or, equivalently,

$$\operatorname{corank} F'(\bar{x}) = \dim \ker F'(\bar{x}) > 0.$$

In other cases $\bar{x}$ is referred to as *regular solution.*

*Computing Center of the Russian Academy of Sciences, Vavilova Str. 40, Moscow, GSP-1, Russia (izmaf@ccas.ru). Research of this author is supported by the Russian Foundation for Basic Research Grants 99-01-00472 and 01-01-00810. The first author also thanks IMPA, where he was a visiting professor during the completion of this work.

†Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil (solodov@impa.br). Research of this author is supported in part by CNPq Grant 300734/95-6, by PRONEX–Optimization, and by FAPERJ.

Singularity gives rise to numerous difficulties. It is well-known that for Newton-type methods at best one can guarantee linear convergence rate to a singular solution [6, 7, 9]. Moreover, it is not sufficient to choose a starting point only close enough to a solution (usually the set of appropriate starting points does not contain a full neighborhood of the solution, although this set is normally rather "dense" [18]). We refer the reader to the survey [19] and references therein. Another difficulty typical in this context is related to possible instability of a singular solution with respect to perturbations of F [27]. Certain special approaches to overcome those difficulties have been developed in the last two decades, but they employ second derivatives of F. Concerning methods for computing singular solutions, we cite [8, 20, 19, 43, 14, 1], and the more recent proposals in [26, 22, 21, 2, 27, 4] (of course, this list does not mention all contributions in this field).

One of the motivations for our new approach to solving singular equations lies in applications to the classical nonlinear complementarity problem (NCP) [37, 12, 13], which is to find an $x \in \mathbf{R}^n$ such that

$$g(x) \geq 0, \quad x \geq 0, \quad \langle g(x),\, x\rangle = 0, \tag{1.2}$$

where $g : \mathbf{R}^n \to \mathbf{R}^n$ is smooth. One of the most useful approaches to numerical and theoretical treatment of the NCP consists in reformulating it as a system of smooth or nonsmooth equations [35, 29, 46]. One possible choice of a smooth reformulation is given by the following function (for other choices, see Section 5.1):

$$F : \mathbf{R}^n \to \mathbf{R}^n, \quad F_i(x) = 2g_i(x)x_i - (\min\{0,\, g_i(x) + x_i\})^2,\ i = 1,\, \ldots,\, n. \tag{1.3}$$

It is easy to check that for this mapping the solution set of the system of equations (1.1) coincides with the solution set of the NCP (1.2) [29, 47]. If $\bar{x}$ is a solution of the NCP, by direct computations (see Section 3), we obtain that

$$F_i'(\bar{x}) = 2\begin{cases} 0, & \text{if } i \in I_0, \\ \bar{x}_i g_i'(\bar{x}), & \text{if } i \in I_1, \\ g_i(\bar{x})e^i, & \text{if } i \in I_2, \end{cases} \tag{1.4}$$

where $e^1, \ldots, e^n$ denotes the standard basis in $\mathbf{R}^n$, and the index sets I_0, I_1 and I_2 are defined by

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

It is immediately clear that $F'(\bar{x})$ cannot be nonsingular, unless the index set I_0 is empty. The latter *strict complementarity* assumption is regarded rather restrictive. Therefore, smooth NCP reformulation provided by (1.3) gives rise to *inherently* singular solutions of the corresponding system of equations. In fact, it is known that *any* other smooth NCP reformulation has the same singularity properties [31] (see also Section 5.1). Furthermore, it is clear that F is once differentiable with Lipschitz-continuous derivative (if g is twice continuously differentiable), but F is not twice differentiable when $I_0 \neq \emptyset$. This is also a common property shared by all useful smooth reformulations, e.g., see the collection [16]. Thus NCP reformulations provide an interesting example of once differentiable nonlinear equations whose solutions are inherently singular. As discussed above, application of standard numerical techniques (e.g., Newton methods) in this context is prone to difficulties (and even failure) because of singularity. On the other hand, known special approaches to computing singular solutions are inapplicable, since these require second derivatives of F. This is the apparent reason why superlinearly convergent Newton-type algorithms for solving the NCP are typically based on nonsmooth equation reformulations and nonsmooth Newton methods (see [13] for a discussion and some references). In this paper we show that it is, in fact, possible to devise superlinearly convergent algorithms based on the smooth NCP reformulations. Specifically, we propose an alternative approach based on computing singular solutions of the smooth reformulation stated above, and show that conditions needed for convergence of our method are principally different from those required for convergence of known nonsmooth algorithms. Thus the two can be considered as a complement to each other.

We complete this section with some notation, which is fairly standard. We denote by $\mathcal{L}^n$ the space of linear operators from $\mathbf{R}^n$ to $\mathbf{R}^n$. For $A \in \mathcal{L}^n$, let $\ker A = \{x \in \mathbf{R}^n \mid Ax = 0\}$ stand for its kernel (null space), and $\operatorname{im} A = \{Ax \mid x \in \mathbf{R}^n\}$ stand for its image (range space). For a bilinear mapping $B : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$ and an element $p \in \mathbf{R}^n$, we define the linear operator $B[p] \in \mathcal{L}^n$ by $B[p]\xi = B[p, \xi]$. Recall that *symmetric* bilinear mappings and linear operators of the form $p \to B[p] : \mathbf{R}^n \to \mathcal{L}^n$ are in isometrically isomorphic correspondence to each other, i.e., the correspondence is one-to-one, linear, and it preserves the norm. Therefore, in the sequel we shall not be making a formal distinction between those objects. Given a set S in a vector space, by $\operatorname{conv} S$ we denote its convex hull, and by $\operatorname{span} S$ its linear hull. Finally, by E we denote the identity operator in $\mathbf{R}^n$.

2. A General Approach to Solving Singular Equations. We start with describing an approach to computing singular solutions of twice differentiable nonlinear equations, which was developed in [26, 22, 27]. We then extend it to the setting of once differentiable mappings, and in the next section show how it applies to solving complementarity problems.

A solution $\bar{x}$ of (1.1) being regular is equivalent to saying that $\operatorname{im} F'(\bar{x}) = \mathbf{R}^n$, while singularity means that $\operatorname{im} F'(\bar{x}) \neq \mathbf{R}^n$. In this situation, one possibility to "regularize" a singular solution $\bar{x}$ is to add to the left-hand side of (1.1) another term, which vanishes at $\bar{x}$ (so that $\bar{x}$ remains a solution), and such that its Jacobian at $\bar{x}$ "compensates" for the singularity of $F'(\bar{x})$ (so that to complement $\operatorname{im} F'(\bar{x})$ in $\mathbf{R}^n$). It is natural to base this extra term on the information about the first derivative of F.

To this end, define the mappings $P : V \to \mathcal{L}^n$, $h : V \to \mathbf{R}^n$, and

$$\Phi : V \to \mathbf{R}^n, \quad \Phi(x) = F(x) + P(x)F'(x)h(x), \tag{2.1}$$

and consider the equation

$$\Phi(x) = 0. \tag{2.2}$$

Suppose that $P(\cdot)$ is defined in such a way that for $\bar{P} = P(\bar{x})$ it holds that

$$\operatorname{im} F'(\bar{x}) \subset \ker \bar{P}. \tag{2.3}$$

Then, by the structure of Φ, solution $\bar{x}$ of (1.1) is also a solution for (2.2). Furthermore, if F is sufficiently smooth (at least twice differentiable at $\bar{x}$), then under appropriate assumptions on the first two derivatives of F at $\bar{x}$, and on $P(\cdot)$ and $h(\cdot)$, it is possible to ensure that Φ is differentiable at $\bar{x}$, and $\bar{x}$ is a regular solution of (2.2). As these assumptions will not be used in this paper, we omit the details, referring the reader to [26, 27]. The regular solution $\bar{x}$ of (2.2) can be computed by means of effective special methods [26, 22, 27], or by conventional numerical techniques (the latter would typically require stronger assumptions, in order to ensure differentiability of Φ not only at $\bar{x}$ but also in its neighborhood). There exist certain general techniques to define $P(\cdot)$ and $h(\cdot)$ with necessary properties (see [26, 27]). However, when one has additional information about the structure of singularity of F at $\bar{x}$ (e.g., recall (1.4) for the NCP reformulation), it can often be used to choose $P(\cdot)$ and $h(\cdot)$ in a particularly simple and constructive way. One such application is precisely the NCP, where the subspace $\operatorname{im} F'(\bar{x})$ can be identified (locally, but without knowing $\bar{x}$), and so the two mappings can be chosen constant (see Section 3).

In this paper, we shall focus exclusively on the case where it is possible to choose $P(\cdot) \equiv \bar{P}$ on V, with some $\bar{P} \in \mathcal{L}^n$ satisfying (2.3). We emphasize that, of course, $\bar{P}$ should be determined without knowing the exact solution $\bar{x}$. The simplest case when this is possible is when we know that $\operatorname{corank} F'(\bar{x}) = n$ (i.e., $F'(\bar{x}) = 0$), or when we are interested in determining a solution specifically with this particular type of singularity. In that case, it is natural to take $\bar{P} = E$. In Section 3, we show how appropriate $\bar{P}$ for the NCP reformulation can be determined using information available at any point close enough to a solution (but without knowing the solution itself). In general, if $P(\cdot)$ is defined as a constant $\bar{P}$ satisfying (2.3), one also can usually take $h(\cdot) \equiv p$, with $p \in \mathbf{R}^n \setminus \{0\}$ being an arbitrary element. Indeed, with those choices the function defined by (2.1) takes the form

$$\Phi(x) = \Phi_p(x) := F(x) + \bar{P}F'(x)p, \quad x \in V, \tag{2.4}$$

and $\bar{x}$ is still a solution of (2.2), due to (2.3). If F is twice differentiable at $\bar{x}$, then it is clear that Φ is differentiable at this point, and

$$\Phi'(\bar{x}) = F'(\bar{x}) + \bar{P}F''(\bar{x})[p]. \tag{2.5}$$

Therefore, $\bar{x}$ is a regular solution of (2.2) if the linear operator in the right-hand side of (2.5) is nonsingular. This is possible under appropriate assumptions. Since the case of twice differentiable F is not the subject of this paper, we shall not discuss here technical details. We only note that nonsingularity of (2.5) subsumes the condition

$$\operatorname{im} F'(\bar{x}) + \operatorname{im} \bar{P} = \mathbf{R}^n.$$

Observe that the latter relation implies that (2.3) must hold as equality. Summarizing, we obtain the following assumptions on the choice of $\bar{P}$:

$$\ker \bar{P} = \operatorname{im} F'(\bar{x}), \quad \mathbf{R}^n = \operatorname{im} F'(\bar{x}) \oplus \operatorname{im} \bar{P}. \tag{2.6}$$

These assumptions clearly hold if, for example, $\bar{P}$ is the projector onto some complement of $\operatorname{im} F'(\bar{x})$ in $\mathbf{R}^n$ parallel to $\operatorname{im} F'(\bar{x})$. With this choice, nonsingularity of (2.5) formally coincides with the notion of 2-regularity of Φ at $\bar{x}$ with respect to $p \in \mathbf{R}^n$, in the sense of [23, 3, 27]. We note, however, that this connection does not seem conceptually important, and in fact, appears to be in some sense a coincidence. Indeed, in the case of once differentiable mappings considered below, nonsingularity condition that would be required no longer has any direct relation to 2-regularity for mappings with Lipschitzian derivatives, as defined in [24, 25].

As a final note, we remark that it can be shown (by simple argument, see [26, 27]) that if there exists at least one element $p \in \mathbf{R}^n$ such that the operator (2.5) is nonsingular, then it will be so for almost every $p \in \mathbf{R}^n$.

We conclude the discussion of the twice differentiable case by the following example, which is very simple but serves to illustrate the basic idea.

EXAMPLE 2.1. Let $n = 1$ and $F : V \to \mathbf{R}$ be twice continuously differentiable on V, where V is a neighbourhood of $\bar{x} \in \mathbf{R}$ which is a singular solution of (1.1). The latter means here that $F(\bar{x}) = F'(\bar{x}) = 0$. Taking $\bar{P} = E$ and any $p \in \mathbf{R} \setminus \{0\}$, we obtain the following regularized equation: $\Phi(x) = F(x) + F'(x)p = 0$. Obviously, $\Phi(\bar{x}) = 0$ and $\Phi'(\bar{x}) = F''(\bar{x})p$, which is distinct from zero for any $p \in \mathbf{R} \setminus \{0\}$, provided $F''(\bar{x}) \neq 0$. This shows that in this example, if $F''(\bar{x}) \neq 0$, singularity can be easily dealt with by using the second-order information.

In the approach outlined above, F is assumed to be twice differentiable. Suppose now that F is once (but not twice) differentiable, and its first derivative is Lipschitz-continuous on V. Then Φ defined by (2.4) is also Lipschitz-continuous on V, and it is natural to try to apply to the corresponding equation (2.2) the *generalized (non-smooth) Newton method* [32, 33, 41, 42, 40, 28]. We emphasize that we shall use the nonsmooth Newton method to solve a (nonsmooth) regularization of a smooth equation. In the context of NCP, this should be compared to the more traditional approach of solving an inherently nonsmooth reformulation by the nonsmooth Newton method. As we shall show in Section 4, the two different approaches lead to two different regularity conditions, neither of which is weaker or stronger than the other.

Let $\partial\Phi(x)$ denote the Clarke's *generalized Jacobian* [5] of Φ at $x \in V$. That is

$$\partial\Phi(x) = \operatorname{conv} \partial_B \Phi(x),$$

where $\partial_B\Phi(x)$ stands for the *B-subdifferential* [45] of Φ at x, which is the set

$$\partial_B\Phi(x) = \{H \in \mathcal{L}^n \mid \exists \{x^k\} \subset D_\Phi \ : x^k \to x \text{ and } \Phi'(x^k) \to H\},$$

with $D_\Phi \subset V$ being the set of points at which Φ is differentiable. With this notation, the nonsmooth Newton method is the following iterative procedure:

$$x^{k+1} = x^k - (H(x^k))^{-1}\Phi(x^k), \quad H(x^k) \in \partial\Phi(x^k), \quad k = 0, 1, \ldots \tag{2.7}$$

It is well-known [42, 40, 28] that if

(i) Φ is semismooth [36] at $\bar{x}$, and

(ii) all the linear operators comprising $\partial\Phi(\bar{x})$ are nonsingular,

then the process (2.7) is locally well-defined and superlinearly convergent to $\bar{x}$. Moreover, if Φ is strongly semismooth [36] then the rate of convergence is quadratic. The regularity condition (ii) can be relaxed if a more specific rule of determining $H(x^k) \in \partial\Phi(x^k)$ is employed. For example, if one chooses $H(x^k) \in \partial_B\Phi(x^k)$ then it is enough to assume *BD-regularity*, i.e., that all elements in $\partial_B\Phi(\bar{x})$ are nonsingular [41].

In applications, Φ usually has some special (tractable) structure, and at each iterate x^k we are interested in obtaining just one, preferably easily computable, $H(x^k) \in \partial\Phi(x^k)$. This would be precisely the case here. The choice of an element in $\partial\Phi(x)$ that we suggest to use in the nonsmooth Newton method for solving (2.2) with Φ given by (2.4), is the following:

$$H(x) = H_p(x) := F'(x) + (\bar{P}F')'(x; p), \quad x \in V, \tag{2.8}$$

where $(\bar{P}F')'(x; p)$ denotes the usual directional derivative of the mapping $\bar{P}F'(\cdot)$ at $x \in V$ with respect to a direction $p \in \mathbf{R}^n$. In Section 3, we show that this

$H(x)$ is explicitly and easily computable for the NCP reformulations. The validity of the choice suggested in (2.8) for an element of $\partial\Phi(x)$ is actually not so obvious. The possibility of choosing the directional derivative $(\bar{P}F')'(x;\, p)$ as an element in the generalized Jacobian of $\bar{P}F'(x)p$ is based on the following fact. At a point $x \in V$ where $\bar{P}F'(\cdot)$ is differentiable, its derivative is in fact the second derivative of $\bar{P}F(\cdot)$. Due to this, $(\bar{P}F')'(x)$ can be considered as a symmetric bilinear mapping. This symmetry will be essential in the proof of Lemma 2.1 below. For a mapping $x \to Q(x)p$, where $p \in \mathbf{R}^n$ and $x \to Q(x) : \mathbf{R}^n \to \mathcal{L}^n$ is an arbitrary Lipschitzian mapping, the inclusion $Q'(x;\, p) \in \partial(Q(x)p)$ can be in general invalid.

LEMMA 2.1. *Suppose that $F : V \to \mathbf{R}^n$ has a Lipschitzian derivative on V, where V is an open set in $\mathbf{R}^n$. Assume that for some $\bar{P} \in \mathcal{L}^n$, the mapping $\bar{P}F' : V \to \mathcal{L}^n$ is directionally differentiable at a point $x \in V$ with respect to a direction $p \in \mathbf{R}^n$.*

Then $H(x) \in \partial\Phi(x)$, where H and Φ are defined in (2.8) and (2.4), respectively.

Proof. Since $\bar{P}F'(\cdot)$ is clearly Lipschitz-continuous, using further the assumption that $\bar{P}F'$ is directionally differentiable at x with respect to p, it follows that there exists a linear operator $B \in \partial(\bar{P}F')(x)$ $(B : \mathbf{R}^n \to \mathcal{L}^n)$ such that

$$(\bar{P}F')'(x;\, p) = Bp. \tag{2.9}$$

The above conclusion can be deduced from [42, Lemma 2.2(ii)] after identifying the space $\mathcal{L}^n$ with the equivalent space $\mathbf{R}^m$, $m = n^2$, and using the equivalence of the norms in finite-dimensional spaces.

By the definition of the generalized Jacobian, $B \in \partial(\bar{P}F')(x)$ means that there exist an integer m, sequences $\{x^{i,k}\} \subset V$ and numbers λ_i, $i = 1, \ldots, m$, with the following properties: $\lambda_i \geq 0$, $\sum_{i=1}^m \lambda_i = 1$, $\bar{P}F'(\cdot)$ is differentiable at each $x^{i,k}$, and

$$x = \lim_{k\to\infty} x^{i,k}, \quad B = \sum_{i=1}^m \lambda_i \lim_{k\to\infty} (\bar{P}F')'(x^{i,k}), \tag{2.10}$$

where the limits in the right-hand side of the second equality exist for each $i = 1, \ldots, m$.

Note that differentiability of $\bar{P}F'(\cdot)$ at each $x^{i,k}$ means that the mapping $\bar{P}F : V \to \mathbf{R}^n$ is twice differentiable at these points. Taking into account the symmetry of the bilinear mapping representing the second derivative, we conclude that

$$\Phi'(x^{i,k}) = F'(x^{i,k}) + (\bar{P}F')'(x^{i,k})p.$$

Therefore,

$$\begin{aligned}
\sum_{i=1}^m \lambda_i \lim_{k\to\infty} \Phi'(x^{i,k}) &= F'(x) + \sum_{i=1}^m \lambda_i \lim_{k\to\infty} (\bar{P}F')'(x^{i,k})p \\
&= F'(x) + Bp \\
&= F'(x) + (\bar{P}F')'(x;\, p) \\
&= H(x),
\end{aligned}$$

where the second equality follows from (2.10), and the third from (2.9). Using the definition of the generalized Jacobian, we conclude that $H(x) \in \partial\Phi(x)$. $\square$

REMARK 2.1. There exists another way to construct the regularized equation $\Phi(x) = 0$, which can have advantages in certain situations over the one described above. Specifically, the mapping Φ defined by (2.4) can be modified as follows:

$$\Phi(x) := (E - \bar{P})F(x) + \bar{P}F'(x)p, \quad x \in V. \tag{2.11}$$

It is clear that with this definition, $\bar{x}$ is still a solution of $\Phi(x) = 0$. Modifying H accordingly, we have

$$H(x) := (E - \bar{P})F'(x) + (\bar{P}F')'(x; p), \quad x \in V. \tag{2.12}$$

Furthermore, it is clear that Lemma 2.1 is still valid with Φ and H defined by (2.11) and (2.12). Finally, it is easy to see that since $\bar{P}F'(\bar{x}) = 0$, the possible limits of $H(x)$ as $x \to \bar{x}$ are the same, whether H is defined by (2.8) or (2.12). Hence, the regularity condition at $\bar{x}$ that would be needed for the superlinear convergence of our method is again the same, whether the method is applied to one regularized equation or the other.

The possible advantage of the modified equation is the following. If singularity of $F'(\bar{x})$ has a certain structure, then not all the components of F may need to be computed in (2.11). Furthermore, (2.12) can also take a simpler form in that case. For example, suppose that $F'(\bar{x})$ is such that $\bar{P}$ satisfying (2.3) can be chosen as the orthogonal projector onto the subspace span$\{e^i, i \in I\}$, where $e^1, \dots, e^n$ is the standard basis in $\mathbf{R}^n$ and $I \subset \{1, \dots, n\}$. Then $E - \bar{P}$ is the orthogonal projector onto span$\{e^i, i \in \{1, \dots, n\} \setminus I\}$. It is easy to see that in this case, (2.11) would not require computing the function values $F_i(x)$, $i \in I$. Furthermore, the derivatives of F_i, $i \in I$, would not appear in (2.12), and so this part would also be simplified. This feature would be further illustrated in the context of NCP in Section 3.

In the sequel, we shall also consider the following modification of the Newton algorithm (2.7), which will be useful for solving the NCP reformulation in Section 3:

$$x^{k+1} = x^k - (\tilde{H}(x^k))^{-1}\Phi(x^k), \quad \begin{array}{l}\|\tilde{H}(x^k) - H(x^k)\| = O(\|x^k - \bar{x}\|), \ H(x^k) \in \partial\Phi(x^k),\\ k = 0, 1, \dots\end{array} \tag{2.13}$$

This modification is essentially motivated by the idea of "truncating" elements of the (generalized) Jacobian by omitting the terms which vanish at the solution $\bar{x}$. These terms typically involve some higher-order derivatives of the problem data (in the context of NCP (1.2), the second derivatives of g), and so it can be advantageous not to compute them, if possible.

Note that the regularity condition which is typically employed in nonsmooth Newton methods consists of saying that every element in the generalized Jacobian $\partial\Phi(\bar{x})$ (or the B-subdifferential $\partial_B\Phi(\bar{x})$) is nonsingular (recall condition (ii) stated above). This seems to be unnecessarily restrictive, because in most implementable algorithms some specific rule to choose $H(x^k) \in \partial\Phi(x^k)$ is used. We shall therefore replace the traditional condition by a weaker one. Specifically, we shall assume that all the possible limits of $H(x^k)$ as $x^k \to \bar{x}$ are nonsingular, where $H(x^k)$ is precisely the element given by (2.8) (or by (2.12)). To this end, we shall define the set

$$\Delta\Phi_p(\bar{x}) := \{\bar{H} \in \mathcal{L}^n \mid \exists\{x^k\} \subset V \ : x^k \to \bar{x}, \ \ H_p(x^k) \to \bar{H}\}.$$

Our regularity assumption would be that elements in $\Delta\Phi_p(\bar{x})$ are nonsingular. We remind the reader that this set is the same for both choices of H, i.e., (2.8) and (2.12). We point out that unlike in the twice differentiable case, this regularity condition cannot be related to the notion of 2-regularity [24, 25] of Φ at $\bar{x}$.

With Lemma 2.1 in hand, convergence of algorithms (2.7) and (2.13), with the data defined in (2.4) and (2.8) or (2.11) and (2.12), can be established similarly to [41], but taking into account the modified nonsingularity assumption.

THEOREM 2.2. *Suppose $F : V \to \mathbf{R}^n$ has a Lipschitz-continuous derivative on V, where V is a neighborhood of a solution $\bar{x}$ of (1.1). Let $\bar{P} \in \mathcal{L}^n$ satisfy (2.3). Assume further that the mapping $\bar{P}F' : V \to \mathcal{L}^n$ is directionally differentiable with respect to a direction $p \in \mathbf{R}^n$ at any point in V, and the mapping $\bar{P}F'(\cdot)p : V \to \mathbf{R}^n$ is semismooth at $\bar{x}$. Let Φ and H be defined by (2.4) and (2.8), or (2.11) and (2.12). Assume further that all linear operators comprising $\Delta\Phi_p(\bar{x})$ are nonsingular.*

Then the iterates given by (2.7) or (2.13) are locally well-defined, and converge to $\bar{x}$ superlinearly. If, in addition, the mapping $\bar{P}F'(\cdot)p$ is strongly semismooth at $\bar{x}$ then the rate of convergence is quadratic.

Proof. It is easy to see that under our regularity assumption, $(H(\cdot))^{-1}$ is locally uniformly bounded. Indeed, assume the contrary, i.e., that there exists a sequence $\{x^k\} \subset V$ such that $x^k \to \bar{x}$, and the sequence $\{(H(x^k))^{-1}\}$ is unbounded (this subsumes the possibility that some elements of the latter sequence are not even well-defined). Recall that the generalized Jacobian is locally bounded [5]. Since, by Lemma 2.1, $H(x^k) \in \partial\Phi(x^k)$ for every k, it follows that the sequence $\{H(x^k)\}$ is bounded. Hence, we can assume that $\{H(x^k)\}$ converges to some $\bar{H} \in \Delta\Phi_p(\bar{x})$, where the inclusion is by the very definition of the set $\Delta\Phi_p(\bar{x})$. But then $\bar{H}$ is nonsingular, which is in contradiction with the earlier assumption that $\{(H(x^k))^{-1}\}$ is unbounded.

Consider first algorithm (2.7), and suppose that the iterates are well-defined up to some index $k \geq 0$. We have that

$$\begin{aligned}\|x^{k+1} - \bar{x}\| &= \|(H(x^k))^{-1}(\Phi(x^k) - \Phi(\bar{x}) - H(x^k)(x^k - \bar{x}))\| \\ &\leq M\|\Phi(x^k) - \Phi(\bar{x}) - H(x^k)(x^k - \bar{x})\|,\end{aligned}$$

where $M > 0$. Note that when $\bar{P}F'(\cdot)p$ is (strongly) semismooth, so is $\Phi(\cdot)$. It is known [40, Proposition 1] that semismoothness of Φ at $\bar{x}$ implies that

$$\sup_{H \in \partial\Phi(\bar{x}+\xi)} \|\Phi(\bar{x} + \xi) - \Phi(\bar{x}) - H\xi\| = o(\|\xi\|)$$

(the latter property was introduced in the context of the nonsmooth Newton methods in [33]). Using Lemma 2.1 and combining the last two relations, well-definedness of the whole sequence $\{x^k\}$ and its superlinear convergence to $\bar{x}$ follow by a standard argument.

In the strongly semismooth case, one has that

$$\sup_{H \in \partial\Phi(\bar{x}+\xi)} \|\Phi(\bar{x} + \xi) - \Phi(\bar{x}) - H\xi\| = O(\|\xi\|^2),$$

and so convergence is quadratic.

Consider now the iterates $\{x^k\}$ generated by (2.13). By our regularity assumption and the classical results of linear analysis, the condition

$$\|\tilde{H}(x^k) - H(x^k)\| = O(\|x^k - \bar{x}\|)$$

implies that

$$\|(\tilde{H}(x^k))^{-1} - (H(x^k))^{-1}\| = O(\|x^k - \bar{x}\|).$$

Hence,

$$\begin{aligned}\|(\tilde{H}(x^k))^{-1}\Phi(x^k) - (H(x^k))^{-1}\Phi(x^k)\| &\leq \|(\tilde{H}(x^k))^{-1} - (H(x^k))^{-1}\|\|\Phi(x^k) - \Phi(\bar{x})\| \\ &= O(\|x^k - \bar{x}\|^2),\end{aligned}$$

where the Lipschitz-continuity of Φ was also used. It follows that the difference between the original and modified steps is of the second order. By the obvious argument, it can now be easily seen that the modified algorithm has the same convergence rate as the original one. $\square$

Note that in principle, our regularity condition depends not only on the structure of singularity of F at $\bar{x}$, but also on the choice of p. Implementation of this approach presumes that there exists at least one $p \in \mathbf{R}^n$ for which this condition is satisfied. Furthermore, a way to choose such p should be available. Fortunately, a typical situation is the following. The *existence* of one suitable p can usually be established under some reasonable regularity assumption. Then, given the existence of one such p, it can further be proven that the set of appropriate elements is, in fact, open and dense in the whole space. Hence, p can be chosen arbitrary, with the understanding that almost any is suitable. We shall come back to this issue in Section 4, where regularity conditions for NCP are discussed. In the computational experience of [22, 26], where conceptually related methods for smooth operator equations are considered, a random choice of p does the job. Even though this choice certainly affects the rate and range of convergence, the differences between different choices are usually not dramatic.

Finally, we remark that the development presented above can be extended to the case when $P(\cdot)$ is not necessarily constant, but it is a Lipschitzian mapping satisfying (2.6) with $\bar{P} = P(\bar{x})$. In that case, we would have to provide a technique to define such $P(\cdot)$ in the general setting. Such techniques are possible, but they go beyond the scope of the present paper. Here we are mainly concerned with a specific application of our approach to the nonlinear complementarity problem, which we consider next.

3. Algorithm for the Nonlinear Complementarity Problem. Consider the nonlinear complementarity problem (1.2), and its reformulation as a system of smooth equations (1.1), given by (1.3). For convenience, we re-state the associated function F, which is

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

We choose a specific reformulation for the clarity of presentation. In Section 5.1, we show that our analysis is intrinsic and extends to other smooth reformulations.

Let $\bar{x} \in \mathbf{R}^n$ be a solution of NCP. Suppose that g is twice continuously differentiable in some neighborhood V of $\bar{x}$ in $\mathbf{R}^n$. Then it is easy to see that F has a Lipschitz-continuous derivative on V, which is given by

$$\begin{aligned} F_i'(x) &= 2(x_i g_i'(x) + g_i(x)e^i - \min\{0,\, g_i(x) + x_i\}(g_i'(x) + e^i)), \\ & \qquad i = 1, \ldots, n, \quad x \in V, \end{aligned} \tag{3.1}$$

where $e^1, \ldots e^n$ is the standard basis in $\mathbf{R}^n$. Recalling the three index sets

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

from (3.1) we immediately obtain that

$$F_i'(\bar{x}) = 2 \begin{cases} 0, & \text{if } i \in I_0, \\ \bar{x}_i g_i'(\bar{x}), & \text{if } i \in I_1, \\ g_i(\bar{x})e^i, & \text{if } i \in I_2. \end{cases} \tag{3.2}$$

As already discussed in Section 1, the Jacobian $F'(\bar{x})$ is necessarily singular whenever $I_0 \neq \emptyset$, the latter being the usual situation for complementarity problems of interest. Furthermore, F is not twice differentiable. Hence, smooth NCP reformulations fall precisely within the framework of Section 2. Such equations cannot be effectively solved by previously available methods, and so our approach comes into play. We next show that in the setting of NCP, the general algorithm introduced in Section 2 takes a simple implementable form.

Given the structure of $F'(\bar{x})$, we have that

$$\operatorname{im} F'(\bar{x}) \subset \operatorname{span}\{e^i\,,\; i \in I_1 \cup I_2\}.$$

Then the natural choice of $\bar{P}$ satisfying $\operatorname{im} F'(\bar{x}) \subset \ker \bar{P}$ (recall condition (2.3)), is the operator with the matrix representation consisting of rows

$$\bar{P}_i = \begin{cases} e^i, & \text{if } i \in I_0, \\ 0, & \text{if } i \in I_1 \cup I_2. \end{cases} \tag{3.3}$$

At the end of this section, we shall show how to define $\bar{P}$ without knowing the solution $\bar{x}$ (clearly, this task reduces to identifying the index set I_0). This is possible by means of *error bound* analysis. A sufficient condition for our error bound is weaker than b-regularity [39], which is currently the weakest assumption under which Newton methods for nonsmooth NCP reformulations are known to be (superlinearly) convergent [30, 34].

Once $\bar{P}$ is defined according to (3.3), we fix $p \in \mathbf{R}^n \setminus \{0\}$ arbitrarily. Then the function Φ defined by (2.4) takes the form

$$\Phi_i(x) = \begin{cases} F_i(x) + \langle F_i'(x),\, p\rangle, & \text{if } i \in I_0, \\ F_i(x), & \text{if } i \in I_1 \cup I_2, \end{cases} \qquad x \in V. \tag{3.4}$$

According to Section 2, $\bar{x}$ is a solution of $\Phi(x) = 0$, which is our “regularized” equation. We proceed to derive explicit forms for iterations of algorithms (2.7) and (2.13), and the regularity condition needed for their convergence.

First, by (2.8) and (3.3), the matrix representation of $H(x)$, which is the element of $\partial\Phi(x)$ employed in algorithm (2.7), consists of rows

$$H_i(x) = \begin{cases} F_i'(x) + (F_i')'(x;\, p), & \text{if } i \in I_0, \\ F_i'(x), & \text{if } i \in I_1 \cup I_2, \end{cases} \qquad x \in V. \tag{3.5}$$

Furthermore, the directional derivatives employed in (3.5) exist, and can be obtained explicitly from (3.1):

$$\begin{aligned} (F_i')'(x;\, p) = {} & 2(x_i g_i''(x)p + p_i g_i'(x) + \langle g_i'(x),\, p\rangle e^i \\ & - \min\{0,\, g_i(x) + x_i\} g_i''(x)p - \gamma_i(x,\, p)(g_i'(x) + e^i)), \\ & i = 1,\, \ldots,\, n,\; x \in V, \end{aligned} \tag{3.6}$$

where

$$\begin{aligned} \gamma_i(x,\, p) = {} & \begin{cases} \langle g_i'(x),\, p\rangle + p_i, & \text{if } g_i(x) + x_i < 0, \\ \min\{0,\, \langle g_i'(x),\, p\rangle + p_i\}, & \text{if } g_i(x) + x_i = 0, \\ 0, & \text{if } g_i(x) + x_i > 0, \end{cases} \\ & i = 1,\, \ldots,\, n,\; x \in V. \end{aligned} \tag{3.7}$$

Note that according to (3.5), one has to compute $(F_i')'(\cdot;\, p)$ only for $i \in I_0$. Another useful observation which would suggest truncation of the Jacobian to be discussed later, is that for $i \in I_0$ all the terms in (3.6) involving the second derivatives of g vanish at $\bar{x}$.

Furthermore, taking into account (3.5), (3.2), (3.6) and (3.7), we conclude that the matrix representation of an arbitrary limit point $\bar{H}$ of $H(x)$ as $x \to \bar{x}$ consists of rows

$$\bar{H}_i = 2 \left\{ \begin{array}{ll} -\langle g_i'(\bar{x}),\, p\rangle g_i'(\bar{x}) - p_i e^i \;\text{ or }\; p_i g_i'(\bar{x}) + \langle g_i'(\bar{x}),\, p\rangle e^i, & \text{if } i \in I_0, \\ \bar{x}_i g_i'(\bar{x}), & \text{if } i \in I_1, \\ g_i(\bar{x}) e^i, & \text{if } i \in I_2. \end{array} \right.$$

Hence, we can state the following *sufficient* condition for nonsingularity of every linear operator in $\Delta\Phi_p(\bar{x})$. Denote by $\mathcal{J}$ the collection of pairs of index sets $(J_1,\, J_2)$ such that $J_1 \cup J_2 = I_0$, $J_1 \cap J_2 = \emptyset$. Our regularity condition consists of saying that for every pair of index sets $(J_1,\, J_2) \in \mathcal{J}$, it holds that

$$\left. \begin{array}{l} \langle g_i'(\bar{x}),\, p\rangle g_i'(\bar{x}) + p_i e^i,\; i \in J_1 \\ p_i g_i'(\bar{x}) + \langle g_i'(\bar{x}),\, p\rangle e^i,\; i \in J_2 \\ g_i'(\bar{x}),\; i \in I_1 \\ e^i,\; i \in I_2 \end{array} \right\} \quad \text{are linearly independent in } \mathbf{R}^n. \tag{3.8}$$

In Section 4, we shall discuss the relation between this condition and other regularity conditions for the NCP, as well as compare convergence results of our algorithm with convergence results of other locally superlinearly convergent equation-based methods for solving NCP.

Under our assumptions, semismoothness of $\bar{P}F'(\cdot)p$ follows readily from (3.1) and standard calculus of semismooth mappings [36, Theorem 5]. Moreover, under the additional assumption of Lipschitz-continuity of $g''(\cdot)$ on V, $\bar{P}F'(\cdot)p$ is strongly semismooth, which follows from results on the superposition of strongly semismooth mappings [15, Theorem 19]. Hence, $\Phi(\cdot)$ is (strongly) semismooth.

Note that all the elements involved in the iteration scheme (2.7) are computed in this section by explicit formulas. In principle, computing H via (3.5)-(3.7) involves second derivatives of g. However, as already noted above, the terms containing second derivatives of g tend to zero as $x \to \bar{x}$. This suggests the idea to modify the process by omitting these terms, which leads to the method represented by (2.13). We shall also take into account the structure of $\bar{P}$, and make use of Remark 2.1.

Note that for $\bar{P}$ given by (3.3), we have that $(E - \bar{P})$ is the orthogonal projector onto span $\{e^i\,,\, i \in I_1 \cup I_2\}$. According to (2.11), we can therefore re-define

$$\Phi_i(x) = \left\{ \begin{array}{ll} \langle F_i'(x),\, p\rangle, & \text{if } i \in I_0, \\ F_i(x), & \text{if } i \in I_1 \cup I_2, \end{array} \right. \qquad x \in V. \tag{3.9}$$

Taking into account (2.12) and omitting further the terms that vanish at $\bar{x}$, we can take

$$\tilde{H}_i(x) = \left\{ \begin{array}{ll} 2(p_i g_i'(x) + \langle g_i'(x),\, p\rangle e^i - \gamma_i(x,\, p)(g_i'(x) + e^i)), & \text{if } i \in I_0, \\ F_i'(x), & \text{if } i \in I_1 \cup I_2, \end{array} \right. \qquad x \in V. \tag{3.10}$$

Comparing expressions (3.9) and (3.10) with (3.4) and (3.5), one can easily observe that the former are simpler and require less computations.

Furthermore, under our smoothness assumptions, it is easy to see that

$$\|\tilde{H}(x) - H(x)\| = O(\|x - \bar{x}\|),$$

and so the modified Newton method given by (2.13) is applicable.

We next give a formal statement of the convergence result for our methods applied to NCP, which is a corollary of Theorem 2.2.

THEOREM 3.1. *Let $g : V \to \mathbf{R}^n$ be a twice continuously differentiable mapping on V, V being a neighborhood of a solution $\bar{x}$ of the NCP* (1.2). *Assume that for some $p \in \mathbf{R}^n$ condition* (3.8) *is satisfied for every pair of index sets $(J_1, J_2) \in \mathcal{J}$.*

Then the iterates given by (2.7) *or* (2.13) *(with all the objects as defined in this section) converge to $\bar{x}$ locally superlinearly. If, in addition, the second derivative of g is Lipschitz-continuous on V, then the rate of convergence is quadratic.*

We next show how to construct $\bar{P}$ without knowing the solution $\bar{x}$. Given the structure of $\bar{P}$, see (3.3), it is clear that this task reduces to correct identification of the degenerate set I_0. This can be done with the help of *error bounds*, as described next (our approach is in the spirit of the technique developed in [10] for identification of active constraints in nonlinear programming). To our knowledge, the weakest condition under which a local error bound for NCP is currently available is the 2-regularity of F given by (1.3) at the NCP solution $\bar{x}$ [25]. Specifically, if F is 2-regular at $\bar{x}$ then there exist a neighborhood U of $\bar{x}$ in $\mathbf{R}^n$ and a constant $M_1 > 0$ such that

$$\|x - \bar{x}\| \leq M_1 \|F(x)\|^{1/2} \quad \forall\, x \in U. \tag{3.11}$$

We shall not introduce the notion of 2-regularity formally here, as this would require an extensive discussion. We only emphasize that the bound (3.11) may hold when the so-called natural residual $\min\{x,\, g(x)\}$ does not provide an error bound, and always holds when it does (see [25], and in particular [25, Example 1]). Hence, 2-regularity of F is a weaker assumption than the R_0 type property or semistability, which in case of NCP are both equivalent to an error bound in terms of the natural residual [38]. And it is further weaker than b-regularity, see [25].

We note that in Lemma 3.2 below we could also use other error bounds for identifying I_0. However, they would require either stronger local assumptions, or global assumptions.

LEMMA 3.2. *Suppose that $\bar{x}$ is a solution of NCP, g is Lipschitz-continuous on V, where V is a neighborhood of $\bar{x}$. Suppose finally that the local error bound* (3.11) *holds. Then for any $\alpha \in (0,\, 1)$ there exists a neighborhood U of $\bar{x}$ such that*

$$\{i \in \{1,\, \ldots,\, n\} \mid |\max\{g_i(x),\, x_i\}| \leq \|F(x)\|^{\alpha/2}\} = I_0 \quad \forall x \in U. \tag{3.12}$$

Proof. It is easy to observe that there exist some $M_2 > 0$ and some neighborhood U of $\bar{x}$ such that

$$\text{for } i \in I_0,\ |\max\{g_i(x),\, x_i\}| = |\max\{g_i(x),\, x_i\} - \max\{g_i(\bar{x}),\, \bar{x}_i\}| \leq M_2\|x - \bar{x}\| \quad \forall x \in U,$$

where the inequality follows from the Lipschitz-continuity of the functions involved. Therefore, by (3.11) (possibly adjusting the neighborhood U), for an arbitrary fixed $\alpha \in (0,\, 1)$ we have that

$$\text{for } i \in I_0,\ |\max\{g_i(x),\, x_i\}| \leq M_1 M_2 \|F(x)\|^{1/2} \leq \|F(x)\|^{\alpha/2} \quad \forall\, x \in U.$$

In particular, the quantity in the left-hand side of the inequality above tends to zero as x tends to $\bar{x}$. On the other hand, it is clear that there exists $\varepsilon > 0$ such that

$$\text{for } i \in I_1 \cup I_2, \ |\max\{g_i(x), x_i\}| \geq \varepsilon > 0 \quad \forall\, x \in U$$

(U should be adjusted again, if necessary). Combining those facts, we obtain (3.12) for U sufficiently small. □

By Lemma 3.2, the index set I_0, and hence the mapping $\bar{P}$, are correctly identified by (3.12), provided one has a point close enough to the solution. We note that this requirement of closedness to solution is completely consistent with the setting of the paper, since the subject under consideration are superlinearly convergent Newton-like methods, which are local by nature.

Finally, we mention other considerations that can also be useful for identifying I_0. Sometimes the cardinality r of I_0 may be known from a priori analysis of the problem, or one can be interested in finding an NCP solution with a given cardinality of I_0. Then for any $x \in \mathbf{R}^n$ sufficiently close to $\bar{x}$, the set I_0 coincides with the set of indices corresponding to the r smallest values of $|\max\{g_i(x), x_i\}|$. In this case, no error bound is needed to identify I_0. We note that in the present setting, cardinality of I_0 is closely related to corank of singularity. In the literature on numerical methods for solving singular equations, the assumption that corank of singularity is known is absolutely standard [20, 19, 43, 14, 1, 2]. In the complementarity literature, on the other hand, assumptions about cardinality of I_0 are not common, except possibly for $I_0 = \emptyset$.

4. Regularity Conditions. The weakest condition under which there exists a locally superlinearly convergent Newton-type algorithm for solving a (nonsmooth) equation reformulation of the NCP, is the b-regularity assumption, which can be stated as follows: for every pair of index sets $(J_1, J_2) \in \mathcal{J}$, it holds that

$$\left.\begin{array}{ll} g_i'(\bar{x}), & i \in J_1 \cup I_1 \\ e^i, & i \in J_2 \cup I_2 \end{array}\right\} \quad \text{are linearly independent in } \mathbf{R}^n.$$

Under this assumption, the natural residual mapping $x \to \min\{x, g(x)\} : \mathbf{R}^n \to \mathbf{R}^n$, is BD-regular at $\bar{x}$. Furthermore, it is also (strongly) semismooth under standard assumptions on g. Hence, the nonsmooth Newton method (2.7) based on it converges locally superlinearly [30, 34]. Note that Newton methods applied to another popular reformulation based on the Fischer-Burmeister function [17, 11], require for convergence the stronger *R-regularity* [44] assumption, see [34].

In what follows, we compare our regularity condition (3.8) with b-regularity, and show that they are essentially different. In general, neither is weaker or stronger than the other. This implies that our approach based on the smooth NCP reformulation is a complement to nonsmooth reformulations, and vice versa, as each approach can be successful in situations when the other is not.

The next result is important to obtain an insight into the nature of our regularity condition (3.8). We start with the following definition.

DEFINITION 4.1. *A solution $\bar{x}$ of the NCP* (1.2) *is referred to as* quasi-regular*, if for every pair of index sets $(J_1, J_2) \in \mathcal{J}$, there exists an element $p = p(J_1, J_2) \in \mathbf{R}^n$ such that* (3.8) *is satisfied.*

PROPOSITION 4.2. *Suppose that the solution $\bar{x}$ of the NCP* (1.2) *is quasi-regular. Then there exists a universal $p \in \mathbf{R}^n$ which satisfies* (3.8) *for every pair $(J_1, J_2) \in \mathcal{J}$. Moreover, the set of such p is open and dense in $\mathbf{R}^n$.*

Proof. Fix a pair $(J_1, J_2) \in \mathcal{J}$, and consider the determinant of the system of vectors in (3.8) as a function of p. This function is a polynomial on $\mathbf{R}^n$, and this polynomial is not everywhere zero, since it is not zero at $p(J_1, J_2)$ (see Definition 4.1). But then the set where the polynomial is not zero is obviously open and dense in $\mathbf{R}^n$. Moreover, the intersection of such sets corresponding to pairs $(J_1, J_2) \in \mathcal{J}$ is also open and dense, since it is a finite intersection of open and dense sets. □

It follows that if $\bar{x}$ is a quasi-regular solution of NCP in the sense of Definition 4.1, then even picking a random $p \in \mathbf{R}^n$ one is extremely unlikely to pick a "wrong" p (as the set of wrong elements is "thin"). Hence, under the assumption of quasi-regularity, for the implementation of the algorithm described in Section 3 we can choose $p \in \mathbf{R}^n \setminus \{0\}$ arbitrarily, with the understanding that almost every $p \in \mathbf{R}^n$ is appropriate. In particular, for all practical purposes, we can think of quasi-regularity as the regularity condition needed for superlinear convergence of our algorithm. We next investigate the relationship between quasi-regularity and b-regularity.

First, we show that if the cardinality of I_0 is equal to one, then quasi-regularity is in fact weaker than b-regularity.

PROPOSITION 4.3. *Suppose that $\bar{x}$ is a b-regular solution of the NCP, and the cardinality of I_0 is equal to one. Then $\bar{x}$ is quasi-regular.*

Proof. Let $I_0 = \{i_0\}$ and denote $L = \operatorname{span}\{g_i'(\bar{x}),\, i \in I_1,\, e^i,\, i \in I_2\}$. In this setting, b-regularity clearly means that

$$g_{i_0}'(\bar{x}) \notin L, \quad e^{i_0} \notin L,$$

corresponding to the two possible choices of $(J_1, J_2) \in \mathcal{J}$. It follows that

$$\forall\, q \in L^{\perp} \setminus \{0\}, \quad \langle g_{i_0}'(\bar{x}),\, q\rangle \neq 0, \quad q_{i_0} \neq 0. \tag{4.1}$$

Assume for a contradiction that $\bar{x}$ is not quasi-regular. Then by Definition 4.1, there exists a pair $(J_1, J_2) \in \mathcal{J}$ such that for every $p \in \mathbf{R}^n$ condition (3.8) is violated. This means that is either

$$\langle g_{i_0}'(\bar{x}),\, p\rangle g_{i_0}'(\bar{x}) + p_{i_0} e^{i_0} \in L,$$

or

$$p_{i_0} g_{i_0}'(\bar{x}) + \langle g_{i_0}'(\bar{x}),\, p\rangle e^{i_0} \in L.$$

Taking any $q \in L^{\perp} \setminus \{0\}$, we deduce that for every $p \in \mathbf{R}^n$ either

$$\langle g_{i_0}'(\bar{x}),\, q\rangle \langle g_{i_0}'(\bar{x}),\, p\rangle + q_{i_0} p_{i_0} = 0,$$

or

$$\langle g_{i_0}'(\bar{x}),\, q\rangle p_{i_0} + q_{i_0} \langle g_{i_0}'(\bar{x}),\, p\rangle = 0.$$

Setting $p = q$, we then obtain that either

$$\langle g_{i_0}'(\bar{x}),\, q\rangle^2 + q_{i_0}^2 = 0,$$

or

$$\langle g_{i_0}'(\bar{x}),\, q\rangle q_{i_0} = 0,$$

which contradicts b-regularity, because of (4.1). $\square$

It is easy to see that in the setting of Proposition 4.3, the quasi-regularity condition can be satisfied without b-regularity. For example, let $g'_{i_0}(\bar{x}) \in L$, but $e^{i_0} \notin L$. Then b-regularity is violated. On the other hand, quasi-regularity here is equivalent to saying that there exist elements $p^1, p^2 \in \mathbf{R}^n$ (corresponding to the two possible choices of $(J_1, J_2) \in \mathcal{J}$) such that

$$p^1_{i_0} \neq 0, \quad \langle g'_{i_0}(\bar{x}), p^2 \rangle \neq 0,$$

which is satisfied for almost any p^1 and p^2, provided $g'_{i_0}(\bar{x}) \neq 0$. It is also quite clear that just choosing p^1 and p^2 randomly should do the job.

In general, i.e., in the cases of higher cardinality of I_0, b-regularity and quasi-regularity become different, not directly related conditions. In particular, neither is stronger or weaker than the other, as illustrated by the following examples.

EXAMPLE 4.1. Let $n = 2$, $I_0 = \{1, 2\}$, and

$$g'_1(\bar{x}) = (1, \sqrt{2}), \quad g'_2(\bar{x}) = (\sqrt{2}, 1).$$

Then b-regularity is obvious, but

$$\langle g'_i(\bar{x}), p \rangle g'_i(\bar{x}) + p_i e^i = (2p_1 + \sqrt{2}p_2, \sqrt{2}p_1 + 2p_2) \quad \forall\, i = 1, 2, \quad \forall\, p \in \mathbf{R}^2.$$

This means that for $J_1 = I_0$, $J_2 = \emptyset$, (3.8) does not hold for any p, and so the quasi-regularity condition is not satisfied.

EXAMPLE 4.2. Let $n = 2$, $I_0 = \{1, 2\}$, and

$$g'_1(\bar{x}) = e^2, \quad g'_2(\bar{x}) \notin \operatorname{span}\{e^1\}.$$

Then b-regularity does not hold (the linear independence condition is violated for $J_1 = \{1\}$, $J_2 = \{2\}$), but quasi-regularity is satisfied, which can be shown by straightforward computations. We omit the details, as they do not provide any further insight.

We complete our discussion with a sufficient condition for quasi-regularity of $\bar{x}$, which is meaningful when the cardinality of I_0 is not greater than $n/2$, half dimensionality of the space. Specifically, suppose that

$$g'_i(\bar{x}),\ e^i,\ i \in I_0, \text{ are linearly independent in } \mathbf{R}^n, \tag{4.2}$$

and

$$\exists\, (\bar{J}_1, \bar{J}_2) \in \mathcal{J} \text{ s.t. } \left.\begin{array}{ll} g'_i(\bar{x}), & i \in \bar{J}_1 \cup I_1 \\ e^i, & i \in \bar{J}_2 \cup I_2 \end{array}\right\} \text{ are linearly independent in } \mathbf{R}^n. \tag{4.3}$$

It is clear that (4.3) is subsumed by b-regularity (where it must hold for *all* partitions of I_0). It is also not difficult to see that (4.3) is necessary for quasi-regularity of $\bar{x}$. Hence, this assumption does not introduce any additional restrictions with respect to regularity conditions under consideration. Furthermore, for non-pathological problems the cardinality of I_0 should not be too large compared to the dimensionality of the space, and so condition (4.2) should not be difficult to satisfy. Therefore, (4.2) and (4.3) appear to be not restrictive.

PROPOSITION 4.4. *Suppose that* (4.2) *and* (4.3) *hold. Then* $\bar{x}$ *is a quasi-regular solution of NCP.*

Proof. Take any pair of index sets $(J_1, J_2) \in \mathcal{J}$, and consider the system of (twice the cardinality of I_0) linear equations

$$\begin{cases} \langle g_i'(\bar{x}), p\rangle = 1, \, p_i = 0, \, i \in J_1 \cap \bar{J}_1 \\ \langle g_i'(\bar{x}), p\rangle = 0, \, p_i = 1, \, i \in J_1 \cap \bar{J}_2 \\ \langle g_i'(\bar{x}), p\rangle = 0, \, p_i = 1, \, i \in J_2 \cap \bar{J}_1 \\ \langle g_i'(\bar{x}), p\rangle = 1, \, p_i = 0, \, i \in J_2 \cap \bar{J}_2 \end{cases}$$

in the variable $p \in \mathbf{R}^n$. Under the assumption (4.2), this system has a solution $p = p(J_1, J_2)$. Observe further that substituting this p into (3.8) reduces the system of vectors appearing in (3.8) precisely to the system of vectors appearing in (4.3), which is linearly independent by the hypothesis. □

Again, it is easy to see that the latter sufficient condition for quasi-regularity of $\bar{x}$ can hold without b-regularity. On the other hand, in general it is not implied by b-regularity. In particular, b-regularity need not imply (4.2).

Summarizing the preceding discussion, we conclude that the regularity assumption required for the algorithm proposed in Section 3 for solving the NCP is different from b-regularity, which is the typical assumption in the context of nonsmooth Newton-type methods for solving nonsmooth NCP reformulations. In fact, the two assumptions are of a rather distinct nature. This is not surprising, considering that they result from approaches which are also quite different.

5. Some Further Applications. The general approach presented in Section 2 can be also useful in other problems where complementarity is present. Below we outline applications to another class of smooth reformulations of NCP (different from (1.3)), and to the mixed complementarity problems. We limit this discussion to exhibiting the structure of singularity associated with the smooth equation reformulations of those problems. Deriving the resulting regularity conditions and comparing them to known ones requires too much space. Without going into detail, we claim that regularity assumptions needed for our approach would again be different from assumptions of Newton methods for nonsmooth equations.

5.1. Other NCP Reformulations. The analysis presented in Sections 3 and 4 for NCP is intrinsic in the sense that it is also applicable to smooth reformulations other than the one given by (1.3). Indeed, following [35], consider the family of functions

$$F : \mathbf{R}^n \to \mathbf{R}^n, \quad F_i(x) = \theta(g_i(x)) + \theta(x_i) - \theta(|g_i(x) - x_i|)\,, \; i = 1, \ldots, n,$$

where $\theta : \mathbf{R} \to \mathbf{R}$ is any strictly increasing function such that $\theta(0) = 0$. It can be checked that the NCP solution set coincides with zeroes of F. As an aside, note that reformulation (1.3) cannot be written in the form stated above, so the two are really different.

Suppose further that θ is differentiable on $\mathbf{R}$ with $\theta'(0) = 0$ and $\theta'(t) > 0$ for $t > 0$. For example, we could take

$$\theta(t) = t|t|\,.$$

Let $\bar{x}$ be some solution of NCP, and V be its neighborhood. If g is twice continuously differentiable on V and θ' is Lipschitz-continuous, then the derivative of F is Lipschitz-continuous near $\bar{x}$, and it is given by

$$F_i'(x) = \theta'(g_i(x))g_i'(x) + \theta'(x_i)e^i - \operatorname{sign}(g_i(x) - x_i)\theta'(|g_i(x) - x_i|)(g_i'(x) - e^i),$$
$$i = 1, \ldots, n, \quad x \in V.$$

As is easy to see,

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

Since $\theta'(t) > 0$ for any $t > 0$, we conclude that the structure of singularity here is absolutely identical to that for F given by (1.3) (recall (1.4)). In particular,

$$\operatorname{im} F'(\bar{x}) \subset \operatorname{span}\{e^i\,,\ i \in I_1 \cup I_2\},$$

and all the objects and the analysis in Sections 3 and 4 can be derived in a similar fashion.

5.2. Mixed Complementarity Problems. The mixed complementarity problem (MCP) is a variational inequality on a (generalized) box $B = \{x \in \mathbf{R}^n \mid l \le x \le u\}$, where $l_i \in [-\infty, +\infty)$ and $u_i \in (-\infty, +\infty]$ are such that $l_i < u_i$, $i = 1, \ldots, n$. Specifically, the problem is to find

$$x \in B \text{ such that } \langle g(x), y - x\rangle \ge 0 \text{ for all } y \in B,$$

where $g : \mathbf{R}^n \to \mathbf{R}^n$. It can be seen that this is equivalent to $x \in \mathbf{R}^n$ satisfying the following conditions: for every $i = 1, \ldots, n$

$$\begin{array}{l} \text{if } g_i(x) > 0 \text{ then } x_i = l_i, \\ \text{if } g_i(x) < 0 \text{ then } x_i = u_i, \\ \text{if } g_i(x) = 0 \text{ then } l_i \le x_i \le u_i. \end{array}$$

NCP is a special case of MCP corresponding to $l_i = 0$, $u_i = +\infty$, $i = 1, \ldots, n$. Define

$$\psi : \mathbf{R}^2 \to \mathbf{R}, \quad \psi(a, b) = 2ab - (\min\{0,\, a + b\})^2.$$

We claim that solutions of MCP coincide with zeroes of the function $F : \mathbf{R}^n \to \mathbf{R}^n$ whose components are given by

$$F_i(x) = \begin{cases} \psi(g_i(x),\, x_i - l_i), & i \in I_l := \{i \mid l_i > -\infty,\, u_i = +\infty\}, \\ \psi(-g_i(x),\, u_i - x_i), & i \in I_u := \{i \mid l_i = -\infty,\, u_i < +\infty\}, \\ g_i(x), & i \in I_g := \{i \mid l_i = -\infty,\, u_i = +\infty\}, \\ \psi(-\psi(-g_i(x),\, u_i - x_i),\, x_i - l_i), & i \in I_{lu} := \{i \mid l_i > -\infty,\, u_i < +\infty\}. \end{cases}$$

We omit the proof, which can be carried out by direct verification. Let $\bar{x}$ be some solution of MCP, and V be its neighborhood. If g is twice continuously differentiable on V then the derivative of F is Lipschitz-continuous near $\bar{x}$. Defining

$$\begin{array}{l} I_0 := \{i = 1, \ldots, n \mid g_i(\bar{x}) = 0\} \cap \{i = 1, \ldots, n \mid \bar{x}_i = l_i \text{ or } \bar{x}_i = u_i\}, \\ I_1 := \{i = 1, \ldots, n \mid g_i(\bar{x}) = 0\} \setminus I_0, \\ I_2 := \{i = 1, \ldots, n \mid g_i(\bar{x}) \ne 0\}, \end{array}$$

it can be verified that

$$F_i'(\bar{x}) = \begin{cases} 0, & \text{if } i \in I_0, \\ \rho_i g_i'(\bar{x}), & \text{if } i \in I_1, \\ \nu_i e^i, & \text{if } i \in I_2, \end{cases}$$

where

$$\rho_i = \begin{cases} 2(\bar{x}_i - l_i), & i \in \{i \in I_l \mid g_i(\bar{x}) = 0,\, \bar{x}_i > l_i\}, \\ -2(u_i - \bar{x}_i), & i \in \{i \in I_u \mid g_i(\bar{x}) = 0,\, \bar{x}_i < u_i\}, \\ 1, & i \in I_g, \\ 4(\bar{x}_i - l_i)(u_i - \bar{x}_i), & i \in \{i \in I_{lu} \mid g_i(\bar{x}) = 0,\, l_i < \bar{x}_i < u_i\}, \end{cases}$$

$$\nu_i = \begin{cases} g_i(\bar{x}), & i \in \{i \in I_l \cup I_u \mid g_i(\bar{x}) \neq 0\}, \\ -4g_i(\bar{x})(u_i - l_i), & i \in \{i \in I_{lu} \mid g_i(\bar{x}) < 0,\, \bar{x}_i = u_i\}, \\ 4g_i(\bar{x})(u_i - l_i) + 2(\min\{0,\, u_i - l_i - g_i(\bar{x})\})^2, & i \in \{i \in I_{lu} \mid g_i(\bar{x}) > 0,\, \bar{x}_i = l_i\}. \end{cases}$$

In particular, $\rho_i \neq 0$, $i \in I_1$, and $\nu_i \neq 0$, $i \in I_2$. Observing the structure of $F'(\bar{x})$, further analysis can now follow the ideas of Sections 3 and 4.

6. Concluding Remarks. We have presented a new approach to solving smooth singular equations. Unlike previously available algorithms, our method is applicable when the equation mapping is not necessarily twice differentiable. Important examples of once differentiable singular equations are reformulations of the nonlinear complementarity problems, which we have studied in detail. In particular, we have demonstrated that in the case of NCP our method takes a readily implementable simple form. Furthermore, the structure of singularity can be completely identified by means of local error bound analysis, without knowing the solution itself. It was further shown that the regularity condition required for the superlinear convergence of the presented algorithm is different from conditions needed for the nonsmooth Newton methods applied to nonsmooth NCP reformulations. Thus the two approaches should be regarded as complementing each other. Finally, it was demonstrated that the main ideas of this paper should be also applicable to other problems where complementarity structures are present.

REFERENCES

[1] E.L. Allgower and K. Böhmer. Resolving singular nonlinear equations. *Rocky Mountain Journal of Mathematics*, 18:225–268, 1988.

[2] E.L. Allgower, K. Böhmer, A. Hoy, and V. Janovský. Direct methods for solving singular nonlinear equations. *Zeitschrift für Angewandte Mathematik und Mechanik*, 79:219–231, 1999.

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

[4] O.A. Brezhneva, A.F. Izmailov, A.A. Tretyakov, and A. Khmura. An approach to finding singular solutions to a general system of nonlinear equations. *Computational Mathematics and Mathematical Physics*, 40:347–358, 2000.

[5] F.H. Clarke. *Optimization and Nonsmooth Analysis.* John Wiley & Sons, New York, 1983.

[6] D.W. Decker and C.T. Kelley. Newton's method at singular points. I. *SIAM Journal on Numerical Analysis*, 17:66–71, 1980.

[7] D.W. Decker and C.T. Kelley. Newton's method at singular points. II. *SIAM Journal on Numerical Analysis*, 17:465–471, 1980.

[8] D.W. Decker and C.T. Kelley. Convergence acceleration for Newton's method at singular points. *SIAM Journal on Numerical Analysis*, 19:219–229, 1982.

[9] D.W. Decker, H.B. Kelley, and C.T. Kelley. Convergence rates for Newton's method at singular points. *SIAM Journal on Numerical Analysis*, 20:296–314, 1983.

[10] F. Facchinei, A. Fischer, and C. Kanzow. On the accurate identification of active constraints. *SIAM Journal on Optimization*, 9:14–32, 1999.

[11] 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.

[12] M. C. Ferris and J.-S. Pang (editors). *Complementarity and variational problems: State of the Art.* SIAM Publications, 1997.

[13] M.C. Ferris and C. Kanzow. Complementarity and related problems. In P.M. Pardalos and M.G.C. Resende, editors, *Handbook of Applied Optimization*, pages 514–530. Oxford University Press, 2002.

[14] J.P. Fink and W.C. Rheinboldt. A geometric framework for the numerical study of singular points. *SIAM Journal of Numerical Analysis*, 24:618–633, 1987.

[15] A. Fischer. Solution of monotone complementarity problems with locally Lipschitzian functions. *Mathematical Programming*, 76:513–532, 1997.

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

[17] C. Geiger and C. Kanzow. On the resolution of monotone complementarity problems. *Computational Optimization and Applications*, 5:155–173, 1996.

[18] A.O. Griewank. Starlike domains of convergence for Newton's method at singularities. *Numerische Mathematik*, 35:95–111, 1980.

[19] A.O. Griewank. On solving nonlinear equations with simple singularities or nearly singular solutions. *SIAM Review*, 27:537–563, 1985.

[20] A.O. Griewank and G.W. Reddien. Characterization and computation of generalized turning points. *SIAM Journal on Numerical Analysis*, 21:176–185, 1984.

[21] M. Hermann, P. Kunkel, and W. Middelman. Augmented systems for computation of singular points in Banach space problems. *Zeitschrift für Angewandte Mathematik und Mechanik*, 78:39–50, 1998.

[22] A.F. Izmailov. Stable methods for finding 2-regular solutions of nonlinear operator equations. *Computational Mathematics and Mathematical Physics*, 36:1183–1192, 1996.

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

[24] A.F. Izmailov and M.V. Solodov. The theory of 2-regularity for mappings with Lipschitzian derivatives and its applications to optimality conditions, 2000. *Mathematics of Operations Research*, to appear.

[25] 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.

[26] A.F. Izmailov and A.A. Tretyakov. Local regularization of certain classes of nonlinear operator equations. *Computational Mathematics and Mathematical Physics*, 36:835–846, 1996.

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

[28] H. Jiang, L. Qi, X. Chen, and D. Sun. Semismoothness and superlinear convergence in nonsmooth optimization and nonsmooth equations. In G. Di Pillo and F. Giannessi, editors, *Nonlinear Optimization and Applications*, pages 197–212. Plenum Press, 1996.

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

[30] C. Kanzow and M. Fukushima. Solving box constrained variational inequality problems by using the natural residual with D-gap function globalization. *Operations Research Letters*, 23:45–51, 1998.

[31] C. Kanzow and H. Kleinmichel. A new class of semismooth Newton-type methods for nonlinear complementarity problems. *Computational Optimization and Applications*, 11:227–251, 1998.

[32] B. Kummer. Newton's method for nondifferentiable functions. In J. Guddat et al., editor, *Advances in Mathematical Optimization*, volume 45, pages 114–125. Akademie-Verlag, Berlin, 1988.

[33] B. Kummer. Newton's method based on generalized derivatives for nonsmooth functions. In W. Oettli and D. Pallaschke, editors, *Advances in Optimization*, pages 171–194. Springer–Verlag, Berlin, 1992.

[34] T. De Luca, F. Facchinei, and C. Kanzow. A theoretical and numerical comparison of some semismooth algorithms for complementarity problems. *Computational Optimization and Applications*, 16:173–205, 2000.

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

[36] R. Mifflin. Semismooth and semiconvex functions in constrained optimization. *SIAM Journal on Control and Optimization*, 15:957–972, 1977.

[37] J.-S. Pang. Complementarity problems. In R. Horst and P. Pardalos, editors, *Handbook of Global Optimization*, pages 271–338. Kluwer Academic Publishers, Boston, Massachusetts, 1995.

[38] J.-S. Pang. Private communication, January 2001.

[39] J.-S. Pang and S.A. Gabriel. NE/SQP: A robust algorithm for the nonlinear complementarity

problem. *Mathematical Programming*, 60:295–337, 1993.

[40] J.-S. Pang and L. Qi. Nonsmooth equations: Motivation and algorithms. *SIAM Journal on Optimization*, 3:443–465, 1993.

[41] L. Qi. Convergence analysis of some algorithms for solving nonsmooth equations. *Mathematics of Operations Research*, 18:227–244, 1993.

[42] L. Qi and J. Sun. A nonsmooth version of Newton's method. *Mathematical Programming*, 58:353–367, 1993.

[43] P.J. Rabier and G.W. Reddien. Characterization and computation of singular points with maximum rank deficiency. *SIAM Journal of Numerical Analysis*, 23:1040–1051, 1986.

[44] S. M. Robinson. Strongly regular generalized equations. *Mathematics of Operations Research*, 5:43–62, 1980.

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

[46] D. Sun and L. Qi. On NCP-functions. *Computational Optimization and Applications*, 13:201–220, 1999.

[47] P. Tseng. Growth behavior of a class of merit functions for the nonlinear complementarity problem. *Journal of Optimization Theory and Applications*, 89:17–37, 1996.