CONVERGENCE OF INEXACT NEWTON METHODS FOR GENERALIZED EQUATIONS[1]

Asen L. Dontchev

Mathematical Reviews, Ann Arbor, MI 48107-8604

R. T. Rockafellar

Department of Mathematics, University of Washington, Seattle, WA 98195-4350

Dedicated to Jon Borwein on the occasion of his 60th birthday

Abstract. For solving the generalized equation $f(x) + F(x) \ni 0$, where f is a smooth function and F is a set-valued mapping acting between Banach spaces, we study the inexact Newton method described by

$$(f(x_k) + Df(x_k)(x_{k+1} - x_k) + F(x_{k+1})) \cap R_k(x_k, x_{k+1}) \neq \emptyset,$$

where Df is the derivative of f and the sequence of mappings R_k represents the inexactness. We show how regularity properties of the mappings $f + F$ and R_k are able to guarantee that every sequence generated by the method is convergent either q-linearly, q-superlinearly, or q-quadratically, according to the particular assumptions. We also show there are circumstances in which at least one convergence sequence is sure to be generated. As a byproduct, we obtain convergence results about inexact Newton methods for solving equations, variational inequalities and nonlinear programming problems.

Key Words. Inexact Newton method, generalized equations, metric regularity, metric subregularity, variational inequality, nonlinear programming.

AMS Subject Classification. 49J53, 49K40, 49M37, 65J15, 90C31.

Revised Version, August 26, 2011

---

1 Introduction

In this paper we consider inclusions of the form

$$(1) \qquad\qquad f(x) + F(x) \ni 0,$$

with $f : X \to Y$ a function and $F : X \rightrightarrows Y$ a set-valued mapping. General models of such kind, commonly called "generalized equations" after Robinson[2], have been used to describe in a unified way various problems such as equations ($F \equiv 0$), inequalities ($Y = \mathbb{R}^m$ and $F \equiv \mathbb{R}^m_+$), variational inequalities (F the normal cone mapping N_C of a convex set C in Y or more broadly the subdifferential mapping ∂g of a convex function g on Y), and in particular, optimality conditions, complementarity problems and multi-agent equilibrium problems.

Throughout, X, Y and P are (real) Banach spaces, unless stated otherwise. For the generalized equation (1) we assume that the function f is continuously Fréchet differentiable everywhere with derivative mapping Df and the mapping F has closed nonempty graph[3].

A Newton-type method for solving (1) utilizes the iteration

$$(2) \qquad f(x_k) + Df(x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \quad \text{for } k = 0, 1, \dots,$$

with a given starting point x_0. When F is the zero mapping, the iteration (2) becomes the standard Newton method for solving the equation $f(x) = 0$:

$$(3) \qquad f(x_k) + Df(x_k)(x_{k+1} - x_k) = 0, \quad \text{for } k = 0, 1, \dots.$$

For $Y = \mathbb{R}^m \times \mathbb{R}^l$ and $F = \mathbb{R}^m_+ \times \{0\}_{\mathbb{R}^l}$, the inclusion (1) describes a system of equalities and inequalities and the method (2) becomes a fairly known iterative procedure for solving feasibility problems of such kind. In the case when F is the normal cone mapping appearing in the Karush-Kuhn-Tucker optimality system for a nonlinear programming problem, the method (2) is closely related to the popular sequential quadratic programming method in nonlinear optimization.

The inexact Newton method for solving equations, as introduced by Dembo, Eisenstat, and Steihaug [4], consists in approximately solving the equation $f(x) = 0$ for $X = Y = \mathbb{R}^n$ in the following way: given a sequence of positive scalars η_k and a starting point x_0, the $(k + 1)$st iterate is chosen to satisfy the condition

$$(4) \qquad \|f(x_k) + Df(x_k)(x_{k+1} - x_k)\| \le \eta_k \|f(x_k)\|.$$

---

[2]Actually, in his pioneering work [15] Robinson considered variational inequalities only.

[3]Since our analysis is local, one could localize these assumptions around a solution $\bar{x}$ of (1). Also, in some of the presented results, in particular those involving strong metric subregularity, it is sufficient to assume continuity of Df only at $\bar{x}$. Since the paper is already quite involved technically, we will not go into these refinements in order to simplify the presentation as much as possible.

Basic information about this method is given in the book of Kelley [14, Chapter 6], where convergence and numerical implementations are discussed. We will revisit the results in [4] and [14] in Section 4, below.

Note that the iteration (4) for solving equations can be also written as

$$(f(x_k) + \nabla f(x_k)(x_{k+1} - x_k)) \cap I\!B_{\eta_k \|f(x_k)\|}(0) \neq \emptyset,$$

where we denote by $I\!B_r(x)$ the closed ball centered at x with radius r. Here we extend this model to solving generalized equations, taking a much broader approach to "inexactness" and working in a Banach space setting, rather than just $\mathbb{R}^n$. Specifically, we investigate the following inexact Newton method for solving generalized equations:

$$(5) \qquad (f(x_k) + Df(x_k)(x_{k+1} - x_k) + F(x_{k+1})) \cap R_k(x_k, x_{k+1}) \neq \emptyset, \quad \text{for } k = 0, 1, \ldots,$$

where $R_k : X \times X \rightrightarrows Y$ is a sequence of set-valued mappings with closed graphs. In the case when F is the zero mapping and $R_k(x_k, x_{k+1}) = I\!B_{\eta_k \|f(x_k)\|}(0)$, the iteration (5) reduces to (4).

Two issues are essential to assessing the performance of any iterative method: convergence of a sequence it generates, but even more fundamentally, its ability to produce an infinite sequence at all. With iteration (5) in particular there is the potential difficulty that a stage might be reached in which, given x_k, there is no x_{k+1} satisfying the condition in question, and the calculations come to a halt. When that is guaranteed not to happen, we can speak of the method as being *surely executable*.

In this paper, we give conditions under which the method (5) is surely executable and every sequence generated by it converges with either q-linear, q-superlinear, or q-quadratic rate, provided that the starting point is sufficiently close to the reference solution. We recover, through specialization to (4), convergence results given in [4] and [14]. The utilization of *metric regularity* properties of set-valued mappings is the key to our being able to handle generalized equations as well as ordinary equations. Much about metric regularity is laid out in our book [9], but the definitions will be reviewed in Section 2.

The extension of the *exact* Newton iteration to generalized equations goes back to the PhD thesis of Josephy [13], who proved existence and uniqueness of a quadratically convergent sequence generated by (2) under the condition of strong metric regularity of the mapping $f + F$. We extend this here to inexact Newton methods of the form (5) and also explore the effects of weaker regularity assumptions.

An inexact Newton method of a form that fits (5) was studied recently by Izmailov and Solodov in [12] for the generalized equation (1) in finite dimensions and with a reference solution $\bar{x}$ such that the mapping $f + F$ is strongly subregular at $\bar{x}$ for 0 (in [12] this property is called semistability). It is assumed in [12, Theorem 2.1] that the mapping R_k in (5) does not depend on k and the following conditions hold:

(a) For every u near $\bar{x}$ there exists $x(u)$ solving $(f(u) + Df(u)(x - u) + F(x)) \cap R(u, x) \neq \emptyset$ such that $x(u) \to \bar{x}$ as $u \to \bar{x}$;

(b) Every $\omega \in (f(u) + Df(u)(x-u) + F(x)) \cap R(u,x)$ satisfies $\|\omega\| = o(\|x-u\| + \|u-\bar{x}\|)$ uniformly in $u \in X$ and x near $\bar{x}$.

Note that for $R(u,x) = I\!B_{\eta\|f(u)\|}(0)$ with the Jacobian $Df(\bar{x})$ being nonsingular, which is the case considered by Dembo et al. [4], the assumption (b) *never* holds. Under conditions (a) and (b) above it is demonstrated in [12, Theorem 2.1] that there exists $\delta > 0$ such that, for any starting point close enough to $\bar{x}$, there exists a sequence $\{x_k\}$ satisfying (5) and the bound $\|x_{k+1} - x_k\| \le \delta$; moreover, each such sequence is superlinearly convergent to $\bar{x}$. It is not specified however in [12] how to find a constant δ in order to identify a convergent sequence.

In contrast to Izmailov and Solodov [12], we show here that under strong subregularity only for the mapping $f + F$ plus certain conditions for the sequence of mappings R_k, *all sequences* generated by the method (5) and staying sufficiently close to a solution $\bar{x}$, converge to $\bar{x}$ at a rate determined by a bound on R_k. In particular, we recover the results in [4] and [14]. Strong subregularity of $f + F$ alone is however not sufficient to guarantee that there exist infinite sequences generated by the method (5) for any starting point close to $\bar{x}$.

To be more specific about the pattern of assumptions on which we rely, we focus on a particular solution $\bar{x}$ of the generalized equation (1), so that the graph of $f + F$ contains $(\bar{x}, 0)$, and invoke properties of *metric regularity, strong metric subregularity* and *strong metric regularity* of $f + F$ at $\bar{x}$ for 0 as quantified by a constant λ. Metric regularity of $f + F$ at $\bar{x}$ for 0 is equivalent to a property we call *Aubin continuity* of $(f + F)^{-1}$ at 0 for $\bar{x}$. However, we get involved with Aubin continuity in another way, more directly. Namely, we assume that the mapping $(u,x) \mapsto R_k(u,x)$ has the *partial* Aubin continuity property in the x argument at $\bar{x}$ for 0, uniformly in k and u near $\bar{x}$, as quantified by a constant μ such that $\lambda\mu < 1$.

In that setting in the case of (plain) metric regularity and under a bound for the inner distance $d(0, R_k(u,\bar{x}))$, we show that for any starting point close enough to $\bar{x}$ the method (5) is surely executable and moreover generates at least one sequence which is linearly convergent. In this situation however, the method might also generate, through nonuniqueness, a sequence which is not convergent at all. This kind of result for the *exact* Newton method (2) was first obtained in [5]; for extensions see e.g. [11] and [3].

We further take up the case when the mapping $f + F$ is strongly metrically subregular, making the stronger assumption on R_k that the outer distance $d^+(0, R_k(u,x))$ goes to zero as $(u,x) \to (\bar{x}, \bar{x})$ for each k, entailing $R_k(\bar{x}, \bar{x}) = \{0\}$, and also that, for a sequence of scalars γ_k and u close to $\bar{x}$, we have $d^+(0, R_k(u,\bar{x})) \le \gamma_k \|u - \bar{x}\|^p$ for $p = 1$, or instead $p = 2$. Under these conditions, we prove that every sequence generated by the iteration (5) and staying close to the solution $\bar{x}$, converges to $\bar{x}$ q-linearly (γ_k bounded and $p = 1$), q-superlinearly ($\gamma_k \to 0$ and $p = 1$) or q-quadratically (γ_k bounded and $p = 2$). The strong metric subregularity, however, does not prevent the method (5) from perhaps getting "stuck" at some iteration and thereby failing to produce an infinite sequence.

Finally, in the case of strong metric regularity, we can combine the results for metric

regularity and strong metric subregularity to conclude that there exists a neighborhood of $\bar{x}$ such that, from any starting point in this neighborhood, the method (5) is surely executable and, although the sequence it generates may be not unique, every such sequence is convergent to $\bar{x}$ either q-linearly, q-superlinearly or q-quadratically, depending on the bound for $d^+(0, R_k(u, \bar{x}))$ indicated in the preceding paragraph.

For the case of an equation $f = 0$ with a smooth $f : \mathbb{R}^n \to \mathbb{R}^n$ near a solution $\bar{x}$, each of the three metric regularity properties we employ is equivalent to the nonsingularity of the Jacobian of f at $\bar{x}$, as assumed in Dembo et. al. [4]. Even in this case, however, our convergence results extend those in [4] by passing to Banach spaces and allowing broader representations of inexactness.

In the recent paper [1], a model of an inexact Newton method was analyzed in which the sequence of mappings R_k in (5) is just a sequence of elements $r_k \in Y$ that stand for error in computations. It is shown under metric regularity of the mapping $f + F$ that if the iterations can be continued without getting stuck, and r_k converges to zero at certain rate, there exists a sequence of iterates x_k which is convergent to $\bar{x}$ with the same r-rate as r_k. This result does not follow from ours. On the other hand, the model in [1] does not cover the basic case in [4] whose extension has been the main inspiration of the current paper.

There is a vast literature on inexact Newton-type method for solving equations which employs representations of inexactness other than that in Dembo et. al. [4], see e.g. [2] and the references therein.

In the following section we present background material and some technical results used in the proofs. Section 3 is devoted to our main convergence results. In Section 4 we present applications. First, we recover there the result in [4] about linear convergence of the iteration (4). Then we deduce convergence of the exact Newton method (2), slightly improving previous results. We then discuss an inexact Newton method for a variational inequality which extends the model in [4]. Finally, we establish quadratic convergence of the sequential quadratically constrained quadratic programming method.

2   Background on metric regularity

Let us first fix the notation. We denote by $d(x, C)$ the inner distance from a point $x \in X$ to a subset $C \subset X$; that is $d(x, C) = \inf \{\|x - x'\| \mid x' \in C\}$ whenever $C \neq \emptyset$ and $d(x, \emptyset) = \infty$, while $d^+(x, C)$ is the outer distance, $d^+(x, C) = \sup \{\|x - x'\| \mid x' \in C\}$. The excess from a set C to a set D is $e(C, D) = \sup_{x \in C} d(x, D)$ under the convention $e(\emptyset, D) = 0$ for $D \neq \emptyset$ and $e(D, \emptyset) = +\infty$ for any D. A set-valued mapping F from X to Y, indicated by $F : X \rightrightarrows Y$, is identified with its graph $\operatorname{gph} F = \{(x, y) \in X \times Y \mid y \in F(x)\}$. It has effective domain $\operatorname{dom} F = \{x \in X \mid F(x) \neq \emptyset\}$ and effective range $\operatorname{rge} F = \{y \in Y \mid \exists x \text{ with } F(x) \ni y\}$. The inverse $F^{-1} : Y \rightrightarrows X$ of a mapping $F : X \rightrightarrows Y$ is obtained by reversing all pairs in the graph; then $\operatorname{dom} F^{-1} = \operatorname{rge} F$.

We start with the definitions of three regularity properties which play the main roles in this paper. The reader can find much more in the book [9], most of which is devoted to these properties.

Definition 2.1 (metric regularity). *Consider a mapping $H : X \rightrightarrows Y$ and a point $(\bar{x}, \bar{y}) \in X \times Y$. Then H is said to be metrically regular at $\bar{x}$ for $\bar{y}$ when $\bar{y} \in H(\bar{x})$ and there is a constant $\lambda > 0$ together with neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that*

$$(6) \qquad d(x, H^{-1}(y)) \leq \lambda d(y, H(x)) \quad \text{for all } (x, y) \in U \times V.$$

If $f : X \to Y$ is smooth near $\bar{x}$, then metric regularity of f at $\bar{x}$ for $f(\bar{x})$ is equivalent to the surjectivity of its derivative mapping $Df(\bar{x})$. Another popular case is when the inclusion $0 \in H(x)$ describes a system of inequalities and equalities, i.e.,

$$H(x) = h(x) + F, \quad \text{where } h = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \text{ and } F = \begin{pmatrix} \mathbb{R}^m_+ \\ 0 \end{pmatrix}$$

with smooth functions g_1 and g_2. Metric regularity of the mapping H at, say, $\bar{x}$ for 0 is equivalent to the standard Mangasarian-Fromovitz condition at $\bar{x}$, see e.g. [9, Example 4D.3].

Metric regularity of a mapping F is equivalent to *linear openness* of H and to *Aubin continuity* of the inverse H^{-1}, both with the same constant λ but perhaps with different neighborhoods U and V. Recall that a mapping $S : Y \rightrightarrows X$ is said to be *Aubin continuous* (or have the Aubin property) at $\bar{y}$ for $\bar{x}$ if $\bar{x} \in S(\bar{y})$ and there exists $\lambda > 0$ together with neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$e(S(y) \cap U, S(y')) \leq \kappa \|y - y'\| \quad \text{for all } y, y' \in V.$$

We also employ a partial version of the Aubin property for a mappings of two variables. We say that a mapping $T : P \times X \rightrightarrows Y$ is *partially Aubin continuous* at $\bar{y}$ for $\bar{x}$ uniformly in p around $\bar{p}$ if $\bar{x} \in T(\bar{y}, \bar{p})$ and there exist $\lambda > 0$ and neighborhoods U of $\bar{x}$, V of $\bar{y}$ and Q of $\bar{p}$ such that

$$e(T(p, y) \cap U, T(p, y')) \leq \kappa \|y - y'\| \quad \text{for all } y, y' \in V \text{ and all } p \in Q.$$

Definition 2.2 (strong metric regularity). *Consider a mapping $H : X \rightrightarrows Y$ and a point $(\bar{x}, \bar{y}) \in X \times Y$. Then H is said to be strongly metrically regular at $\bar{x}$ for $\bar{y}$ when $\bar{y} \in H(\bar{x})$ and there is a constant $\lambda > 0$ together with neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that (6) holds together with the property that the mapping $U \ni x \mapsto H^{-1}(x) \cap V$ is single-valued.*

When a mapping $y \mapsto S(y) \cap U'$ is single-valued and Lipschitz continuous on V', for some neighborhoods U' and V' of $\bar{x}$ and $\bar{y}$, respectively, then S is said to have a *Lipschitz localization* around $\bar{y}$ for $\bar{x}$. Strong metric regularity of a mapping H at $\bar{x}$ for $\bar{y}$ is then

equivalent to the existence of a Lipschitz localization of H^{-1} around $\bar{y}$ for $\bar{x}$. A mapping S is Aubin continuous at $\bar{y}$ for $\bar{x}$ with constant λ and has a single-valued localization around $\bar{y}$ for $\bar{x}$ if and only if S has a Lipschitz localization around $\bar{y}$ for $\bar{x}$ with Lipschitz constant λ.

Strong metric regularity is the property which appears in the classical inverse function theorem: when $f : X \to Y$ is smooth around $\bar{x}$ then f is strongly metrically regular if and only if $Df(\bar{x})$ is invertible[4]. In Section 4 we will give a sufficient condition for strong metric regularity of the variational inequality representing the first-order optimality condition for the standard nonlinear programming problem.

Our next definition is a weaker form of strong metric regularity.

Definition 2.3 (strong metric subregularity). *Consider a mapping $H : X \rightrightarrows Y$ and a point $(\bar{x}, \bar{y}) \in X \times Y$. Then H is said to be strongly metrically subregular at $\bar{x}$ for $\bar{y}$ when $\bar{y} \in H(\bar{x})$ and there is a constant $\lambda > 0$ together with neighborhoods U of $\bar{x}$ such that*

$$\|x - \bar{x}\| \leq \lambda d(\bar{y}, H(x)) \quad \text{for all } x \in U.$$

Metric subregularity of H at $\bar{x}$ for $\bar{y}$ implies that $\bar{x}$ is an isolated point in $H(\bar{y})$; moreover, it is equivalent to the so-called isolated calmness of the inverse H^{-1}, meaning that there is a neighborhood U of $\bar{x}$ such that $H^{-1}(y) \cap U \subset \bar{x} + \lambda \|y - \bar{x}\| \mathbb{B}$ for all $y \in Y$, see [9, Section 3I]. Every mapping H acting in finite dimensions, whose graph is the union of finitely many convex polyhedral sets, is strongly metrically regular at $\bar{x}$ for $\bar{y}$ if and only if $\bar{x}$ is an isolated point in $H^{-1}(\bar{y})$. As another example, consider the minimization problem

(7) $$\text{minimize } g(x) - \langle p, x \rangle \quad \text{over} \quad x \in C,$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a convex C^2 function, $p \in \mathbb{R}^n$ is a parameter, and C is a convex polyhedral set in $\mathbb{R}^n$. Then the mapping $\nabla g + N_C$ is strongly metrically subregular at $\bar{x}$ for $\bar{p}$, or equivalently, its inverse, which is the solution mapping of problem (7), has the isolated calmness property at $\bar{p}$ for $\bar{x}$, if and only if the standard second-order sufficient condition holds at $\bar{x}$ for $\bar{p}$; see [9, Theorem 4E.4].

In the proofs of convergence of the inexact Newton method (5) given in Section 3 we use some technical results. The first is the following coincidence theorem from [7]:

Theorem 2.4 (coincidence theorem). *Let X and Y be two metric spaces. Consider a set-valued mapping $\Phi : X \rightrightarrows Y$ and a set-valued mapping $\Upsilon : Y \rightrightarrows X$. Let $\bar{x} \in X$ and $\bar{y} \in Y$ and let c, κ and μ be positive scalars such that $\kappa\mu < 1$. Assume that one of the sets $\operatorname{gph} \Phi \cap (\mathbb{B}_c(\bar{x}) \times \mathbb{B}_{c/\mu}(\bar{y}))$ and $\operatorname{gph} \Upsilon \cap (\mathbb{B}_{c/\mu}(\bar{y}) \times \mathbb{B}_c(\bar{x}))$ is closed while the other is complete, or both sets $\operatorname{gph}(\Phi \circ \Upsilon) \cap (\mathbb{B}_c(\bar{x}) \times \mathbb{B}_c(\bar{x}))$ and $\operatorname{gph}(\Upsilon \circ \Phi) \cap (\mathbb{B}_{c/\mu}(\bar{y}) \times \mathbb{B}_{c/\mu}(\bar{y}))$ are complete. Also, suppose that the following conditions hold:*

---

[4]The classical inverse function theorem actually gives us more: it shows that the single-valued localization of the inverse is smooth and provides also the form of its derivative.

(a) $d(\bar{y}, \Phi(\bar{x})) < c(1 - \kappa\mu)/\mu$;

(b) $d(\bar{x}, \Upsilon(\bar{y})) < c(1 - \kappa\mu)$;

(c) $e(\Phi(u) \cap \mathbb{B}_{c/\mu}(\bar{y}), \Phi(v)) \leq \kappa\, \rho(u, v)$ for all $u, v \in \mathbb{B}_c(\bar{x})$ such that $\rho(u, v) \leq c(1 - \kappa\mu)/\mu$;

(d) $e(\Upsilon(u) \cap \mathbb{B}_c(\bar{x}), \Upsilon(v)) \leq \mu\, \rho(u, v)$ for all $u, v \in \mathbb{B}_{c/\mu}(\bar{y})$ such that $\rho(u, v) \leq c(1 - \kappa\mu)$.

Then there exist $\hat{x} \in \mathbb{B}_c(\bar{x})$ and $\hat{y} \in \mathbb{B}_{c/\mu}(\bar{y})$ such that $\hat{y} \in \Phi(\hat{x})$ and $\hat{x} \in \Upsilon(\hat{y})$. If the mappings $\mathbb{B}_c(\bar{x}) \ni x \mapsto \Phi(x) \cap \mathbb{B}_{c/\mu}(\bar{y})$ and $\mathbb{B}_{c/\mu} \ni y \mapsto \Upsilon(y) \cap \mathbb{B}_c(\bar{x})$ are single-valued, then the points $\hat{x}$ and $\hat{y}$ are unique in $\mathbb{B}_c(\bar{x})$ and $\mathbb{B}_{c/\mu}(\bar{y})$, respectively.

To prove the next technical result given below as Corollary 2.6, we apply the following extension of [1, Theorem 2.1], where the case of strong metric regularity was not included but its proof is straightforward. This is actually a "parametric" version of the Lyusternik-Graves theorem; for a basic statement see [9, Theorem 5E.1].

Theorem 2.5 (perturbed metric regularity). *Consider a mapping $H : X \rightrightarrows Y$ and any $(\bar{x}, \bar{y}) \in \operatorname{gph} H$ at which $\operatorname{gph} H$ is locally closed (which means that the intersection of $\operatorname{gph} H$ with some closed ball around $(\bar{x}, \bar{y})$ is closed). Consider also a function $g : P \times X \to Y$ with $(\bar{q}, \bar{x}) \in \operatorname{dom} g$ and positive constants λ and μ such that $\lambda\mu < 1$. Suppose that H is [resp., strongly] metrically regular at $\bar{x}$ for $\bar{y}$ with constant λ and also there exist neighborhoods Q of $\bar{q}$ and U of $\bar{x}$ such that*

$$(8) \qquad \|g(q, x) - g(q, x')\| \leq \mu\|x - x'\| \quad \text{for all} \quad q \in Q \quad \text{and} \quad x, x' \in U.$$

Then for every $\kappa > \lambda/(1 - \lambda\mu)$ there exist neighborhoods Q' of $\bar{q}$, U' of $\bar{x}$ and V' of $\bar{y}$ such that for each $q \in Q'$ the mapping $g(q, \cdot) + H(\cdot)$ is [resp., strongly] metrically regular at $\bar{x}$ for $g(q, \bar{x}) + \bar{y}$ with constant κ and neighborhoods U' of $\bar{x}$ and $g(q, \bar{x}) + V'$ of $g(q, \bar{x}) + \bar{y}$.

From this theorem we obtain the following extended version of Corollary 3.1 in [1], the main difference being that here we assume that f is merely continuously differentiable near $\bar{x}$, not necessarily with Lipschitz continuous derivative. Here we also suppress the dependence on a parameter, which is not needed, present the result in the form of Aubin continuity, and include the case of strong metric regularity; all this requires certain modifications in the proof, which is therefore presented in full.

Corollary 2.6. *Suppose that the mapping $f + F$ is metrically regular at $\bar{x}$ for 0 with constant λ. Let $u \in X$ and consider the the mapping*

$$(9) \qquad X \ni x \mapsto G_u(x) = f(u) + Df(u)(x - u) + F(x).$$

Then for every $\kappa > \lambda$ there exist positive numbers a and b such that

$$(10) \quad e(G_u^{-1}(y) \cap \mathbb{B}_a(\bar{x}), G_u^{-1}(y')) \leq \kappa\|y - y'\| \quad \text{for every} \quad u \in \mathbb{B}_a(\bar{x}) \quad \text{and} \quad y, y' \in \mathbb{B}_b(0).$$

If $f + F$ is strongly metrically regular around $\bar{x}$ for $\bar{y}$ with constant λ, then the mapping G_u is strongly metrically regular at $\bar{x}$ for 0 uniformly in u; specifically, there are positive a' and b' such that for each $u \in \mathbb{B}_{a'}(\bar{x})$ the mapping

$$y \mapsto G_u^{-1}(y) \cap \mathbb{B}_{a'}(\bar{x})$$

is a Lipschitz continuous function on $\mathbb{B}_{b'}(0)$ with Lipschitz constant κ.

Proof. First, let $\kappa > \lambda' > \lambda$. From one of the basic forms of the Lyusternik-Graves theorem, see e.g. [9, Theorem 5E.4], it follows that the mapping $G_{\bar{x}}$ is metrically regular at $\bar{x}$ for 0 with any constant $\lambda' > \lambda$ and neighborhoods $\mathbb{B}_\alpha(\bar{x})$ and $\mathbb{B}_\beta(0)$ for some positive α and β (this could be also deduced from Theorem 2.5). Next, we apply Theorem 2.5 with $H(x) = G_{\bar{x}}(x)$, $\bar{y} = 0$, $q = u$, $\bar{q} = \bar{x}$, and

$$g(q, x) = f(u) + Df(u)(x - u) - f(\bar{x}) - Df(\bar{x})(x - \bar{x}).$$

Let $\kappa > \lambda'$. Pick any $\mu > 0$ such that $\mu \kappa < 1$ and $\kappa > \lambda'/(1 - \lambda'\mu)$. Then adjust α if necessary so that, from the continuous differentiability of f around $\bar{x}$,

(11) $$\|Df(x) - Df(x')\| \le \mu \quad \text{for every } x, x' \in \mathbb{B}_\alpha(\bar{x}).$$

Then for any $x, x' \in X$ and any $u \in \mathbb{B}_\alpha(\bar{x})$ we have

$$\|g(u, x) - g(u, x')\| \le \|Df(u) - Df(\bar{x})\| \|x - x'\| \le \mu \|x - x'\|,$$

that is, condition (8) is satisfied. Thus, by Theorem 2.5 there exist positive constants $\alpha' \le \alpha$ and β' such that for any $u \in \mathbb{B}_{\alpha'}(\bar{x})$ the mapping $G_u(x) = g(u, x) + G_{\bar{x}}(x)$ is (strongly) metrically regular at $\bar{x}$ for $g(u, \bar{x}) = f(u) + Df(u)(\bar{x} - u) - f(\bar{x})$ with constant κ and neighborhoods $\mathbb{B}_{\alpha'}(\bar{x})$ and $\mathbb{B}_{\beta'}(g(q, \bar{x}))$, that is,

(12) $$d(x, G_u^{-1}(y)) \le \kappa d(y, G_u(x)) \quad \text{for every } u, x \in \mathbb{B}_{\alpha'}(\bar{x}) \text{ and } y \in \mathbb{B}_{\beta'}(g(q, \bar{x})).$$

Now choose positive scalars a and b such that

$$a \le \alpha' \quad \text{and} \quad \mu a + b \le \beta'.$$

Then, using (11), for any $u, x \in \mathbb{B}_a(\bar{x})$ we have

(13) $$\|f(x) - f(u) - Df(u)(x - u)\|$$
$$= \left\| \int_0^1 Df(u + t(u - x))(x - u)dt - Df(u)(x - u) \right\| \le \mu \|x - u\|.$$

Hence, for any $u \in \mathbb{B}_a(\bar{x})$, we obtain

$$\|f(u) + Df(u)(\bar{x} - u) - f(\bar{x})\| \le \mu \|u - \bar{x}\|,$$

9

and then, for $y \in I\!B_b(0)$,

$$
\begin{aligned}
\|g(u, \bar{x}) - y\| &\leq \|f(u) + Df(u)(\bar{x} - u) - f(\bar{x})\| + \|y\| \\
&\leq \mu\|u - \bar{x}\| + b \leq \mu a + b \leq \beta'.
\end{aligned}
$$

Thus, $I\!B_b(0) \subset I\!B_{\beta'}(g(u, \bar{x}))$. Let $y, y' \in I\!B_b(0)$ and $x \in G_u^{-1}(y) \cap I\!B_a(\bar{x})$. Then $x \in I\!B_a(\bar{x})$ and from (12) we have

$$
d(x, G_u^{-1}(y')) \leq \kappa d(y', G_u(x)) \leq \kappa\|y' - y\|.
$$

Taking the supremum on the left with respect to $x \in G_u^{-1}(y) \cap I\!B_c(\bar{x})$ we obtain (10).

If $f + F$ is strongly metrically regular, then we repeat the above argument but now by applying the strong regularity version of Theorem 2.5, obtaining constants a' and b' that might be different from a and b for metric regularity. $\qquad\square$

The following theorem is a "parametric" version of [9, Theorem 3I.6]:

Theorem 2.7 (perturbed strong subregularity). *Consider a mapping $H : X \rightrightarrows Y$ and any $(\bar{x}, \bar{y}) \in \mathrm{gph}\, H$. Consider also a function $g : P \times X \to Y$ with $(\bar{q}, \bar{x}) \in \mathrm{dom}\, g$ and let λ and μ be two positive constants such that $\lambda\mu < 1$. Suppose that H is strongly metrically subregular at $\bar{x}$ for $\bar{y}$ with constant λ and a neighborhood U of $\bar{x}$, and also there exists a neighborhood Q of $\bar{q}$ such that*

(14)
$$
\|g(q, x) - g(q, \bar{x})\| \leq \mu\|x - \bar{x}\| \quad \text{for all} \quad q \in Q \quad \text{and} \quad x \in U.
$$

Then for every $q \in Q$ the mapping $g(q, \cdot) + H(\cdot)$ is strongly metrically regular at $\bar{x}$ for $g(q, \bar{x}) + \bar{y}$ with constant $\lambda/(1 - \lambda\mu)$ and neighborhood U of $\bar{x}$.

Proof. Let $x \in U$ and $y \in H(x)$; if there is no such y the conclusion is immediate under the convention that $d(\bar{y}, \emptyset) = +\infty$. Let $q \in Q$; then, using (14),

$$
\begin{aligned}
\|x - \bar{x}\| &\leq \lambda\|\bar{y} - y\| \leq \lambda\|\bar{y} + g(q, \bar{x}) - g(q, x) - y\| + \lambda\|g(q, x) - q(q, \bar{x})\| \\
&\leq \lambda\|\bar{y} + g(q, \bar{x}) - g(q, x) - y\| + \lambda\mu\|x - \bar{x}\|,
\end{aligned}
$$

hence

$$
\|x - \bar{x}\| \leq \frac{\lambda}{1 - \lambda\mu}\|\bar{y} + g(q, \bar{x}) - g(q, x) - y\|.
$$

Since y is arbitrary in $H(x)$, we conclude that

$$
\|x - \bar{x}\| \leq \frac{\lambda}{1 - \lambda\mu}\|d(\bar{y} + g(q, \bar{x}), g(q, x) + H(x))\|
$$

and the proof is complete. $\qquad\square$

We will use the following corollary of Theorem 2.7.

Corollary 2.8. *Suppose that the mapping $f + F$ is strongly metrically subregular at $\bar{x}$ for 0 with constant λ. Let $u \in X$ and consider the mapping (9). Then for every $\kappa > \lambda$ there exists $a > 0$ such that*

$$(15) \qquad \|x - \bar{x}\| \leq \kappa d(f(u) - Df(u)(u - \bar{x}) - f(\bar{x}), G_u(x)) \quad \text{for every} \quad u, x \in \mathbb{B}_a(\bar{x}).$$

Proof. In [9, Corollary 3I.9] it is proved that if the mapping $f + F$ is strongly metrically subregular at $\bar{x}$ for 0 with constant λ then for any $\kappa > \lambda$ the mapping $G_{\bar{x}}$, as defined in (9), is strongly metrically subregular at $\bar{x}$ for 0 with constant κ. This actually follows easily from Theorem 2.7 with $H = f + F$ and

$$g(q, x) = q(x) = -f(x) + f(\bar{x}) + Df(\bar{x})(x - \bar{x}).$$

Fix $\kappa > \kappa' > \lambda$ and let $\mu' > 0$ be such that $\lambda\mu' < 1$ and $\lambda(1 - \lambda\mu') < \kappa'$. Then there exists $a' > 0$ such that (11) holds with this μ' and α replaced by a'. Utilizing (13), for any $x \in \mathbb{B}_{a'}(\bar{x})$ we obtain

$$\|g(x) - g(\bar{x})\| = \|f(x) - f(\bar{x}) - Df(\bar{x})(x - \bar{x})\| \leq \mu'\|x - x'\|,$$

that is, condition (14) is satisfied. Thus, from Theorem 2.7 the mapping $g + F = G_{\bar{x}}$ is strongly metrically subregular at $\bar{x}$ for 0 with constant κ' and neighborhood $\mathbb{B}_{a'}(\bar{x})$.

To complete the proof we apply Theorem 2.7 again but now with $H(x) = G_{\bar{x}}(x)$, $\bar{y} = 0$, $q = u$, $\bar{q} = \bar{x}$, and

$$g(u, x) = f(u) + Df(u)(x - u) - f(\bar{x}) - Df(\bar{x})(x - \bar{x}).$$

Pick any $\mu > 0$ such that $\mu \leq \mu'$, $\mu\kappa < 1$ and $\kappa > \kappa'/(1 - \kappa'\mu)$. Then there exists a positive $a \leq a'$ such that (11) and hence (13) holds with this μ. Let $u \in \mathbb{B}_a(\bar{x})$. Then for any $x \in X$ we have

$$\|g(u, x) - g(q, \bar{x})\| \leq \|Df(u) - Df(\bar{x})\|\|x - \bar{x}\| \leq \mu\|x - x'\|,$$

that is, (14) is satisfied. Thus, by Theorem 2.7 the mapping $G_u(x) = g(u, x) + G_{\bar{x}}(x)$ is strongly metrically subregular at $\bar{x}$ for $g(u, \bar{x}) = f(u) + Df(u)(\bar{x} - u) - f(\bar{x})$ with constant κ. We obtain (15). $\qquad \square$

3 Convergence of the inexact Newton method

In this section we consider the generalized equation (1) and the inexact Newton iteration (5), namely

$$(f(x_k) + Df(x_k)(x_{k+1} - x_k) + F(x_{k+1})) \cap R_k(x_k, x_{k+1}) \neq \emptyset, \quad \text{for} \quad k = 0, 1, \ldots.$$

Our first result shows that metric regularity is sufficient to make the method (5) surely executable.

Theorem 3.1 (convergence under metric regularity). *Let λ and μ be two positive constants such that $\lambda\mu < 1$. Suppose that the mapping $f + F$ is metrically regular at $\bar{x}$ for 0 with constant λ. Also, suppose that for each $k = 0, 1, \ldots$, the mapping $(u, x) \mapsto R_k(u, x)$ is partially Aubin continuous with respect to x at $\bar{x}$ for 0 uniformly in u around $\bar{x}$ with constant μ. In addition, suppose that there exist positive scalars $\gamma < (1 - \lambda\mu)/\mu$ and β such that*

$$(16) \qquad d(0, R_k(u, \bar{x})) \leq \gamma \|u - \bar{x}\| \quad \text{for all} \quad u \in {I\!\!B}_\beta(\bar{x}) \quad \text{and all} \quad k = 0, 1, \ldots.$$

Then there exists a neighborhood O of $\bar{x}$ such that for any starting point $x_0 \in O$ there exists a Newton sequence $\{x_k\}$ contained in O which is q-linearly convergent to $\bar{x}$.

Proof. Let $t \in (0, 1)$ be such that $0 < \gamma < t(1 - \lambda\mu)/\mu$. Choose a constant κ such that $\kappa > \lambda$, $\kappa\mu < 1$ and $\gamma < t(1 - \kappa\mu)/\mu$. Next we apply Corollary 2.6; let a and b be the constants entering (10) and in addition satisfying

$$(17) \qquad e(R_k(u, x) \cap {I\!\!B}_b(0), R_k(u, x')) \leq \mu\|x - x'\| \text{ for all } u, x, x' \in {I\!\!B}_a(\bar{x}).$$

Choose positive ε such that $\varepsilon < t(1 - \kappa\mu)/\kappa$ and make a even smaller if necessary so that

$$(18) \qquad \|Df(u) - Df(v)\| \leq \varepsilon \quad \text{for all} \quad u, v \in {I\!\!B}_a(\bar{x}).$$

Pick $a' > 0$ to satisfy

$$(19) \qquad a' \leq \min\{a, b/\varepsilon, \beta, b\mu\}.$$

Let $u \in {I\!\!B}_{a'}(\bar{x})$, $u \neq \bar{x}$. We apply Theorem 2.4 to the mappings $x \mapsto \Phi(x) = R_0(u, x)$ and $\Upsilon = G_u^{-1}$, with $\kappa := \kappa$, $\mu := \mu$, $\bar{x} := \bar{x}$, $\bar{y} := 0$ and $c := t\|u - \bar{x}\|$. Since $u \in {I\!\!B}_{a'}(\bar{x})$ and $a' \leq \beta$, from (16) we have

$$(20) \qquad d(0, R_0(u, \bar{x})) \leq \gamma\|u - \bar{x}\| < \frac{t(1 - \kappa\mu)}{\mu}\|u - \bar{x}\| = \frac{c(1 - \kappa\mu)}{\mu}.$$

Further, taking into account (18) in (13) and that $\varepsilon a' \leq b$, we obtain

$$\| - f(\bar{x}) + f(u) + Df(u)(\bar{x} - u)\| \leq \varepsilon a' \leq b.$$

Hence, by the assumption $0 \in f(\bar{x}) + F(\bar{x})$ and the form of G_u in (9), we have

$$-f(\bar{x}) + f(u) + Df(u)(\bar{x} - u) \in G_u(\bar{x}) \cap {I\!\!B}_b(0).$$

Then, from (10),

$$\begin{aligned} d(\bar{x}, G_u^{-1}(0)) \ &\leq \ \kappa d(0, G_u(\bar{x})) \\ &\leq \ \kappa\| - f(\bar{x}) + f(u) + Df(u)(\bar{x} - u)\| \leq \kappa\varepsilon\|u - \bar{x}\| \\ &< \ t(1 - \kappa\mu)\|u - \bar{x}\| = c(1 - \kappa\mu). \end{aligned}$$

We conclude that conditions (a) and (b) in Theorem 2.4 are satisfied. Since $u \in \mathbb{B}_{a'}(\bar{x})$, we have by (19) that $c \leq a' \leq a$ and $c/\mu \leq b$, hence (17) implies condition (c). Further, from (15) we obtain that condition (d) in Theorem 2.4 holds for the mapping $\Upsilon = G_u^{-1}$. Thus, we can apply Theorem 2.4 obtaining that there exists $x_1 \in \mathbb{B}_c(\bar{x})$ and v_1 such that $x_1 \in G_u^{-1}(v_1)$ and $v_1 \in R_0(u, x_1)$, that is, x_1 satisfies (5) with $x_0 = u$ and also $\|x_1 - \bar{x}\| \leq t\|x_0 - \bar{x}\|$. In particular, $x_1 \in \mathbb{B}_{a'}(\bar{x})$.

The induction step repeats the argument used in the first step. Having iterates $x_i \in \mathbb{B}_{a'}(\bar{x})$ from (5) for $i = 0, 1 \ldots, k-1$ with $x_0 = u$, we apply Theorem 2.4 with $c := t\|x_k - \bar{x}\|$, obtaining the existence of x_{k+1} satisfying (5) which is in $\mathbb{B}_c(\bar{x}) \subset \mathbb{B}_{a'}(\bar{x})$ and $\|x_{k+1} - \bar{x}\| \leq t\|x_k - \bar{x}\|$ for all k. □

If we assume that in addition Df is Lipschitz continuous near $\bar{x}$ and also $0 \in R_k(u, x)$ for any (u, x) near $(\bar{x}, \bar{x})$, the above theorem would follow from [9, Theorem 6C.6], where the existence of a quadratically convergent sequence is shown generated by the *exact* Newton method (2). Indeed, in this case any sequence that satisfies (2) will also satisfy (5).

Under metric regularity of the mapping $f + F$, even the exact Newton method (2) may generate a sequence which is not convergent. The simplest example of such a case is the inequality $x \leq 0$ in $\mathbb{R}$ which can be cast as the generalized equation $0 \in x + \mathbb{R}_+$ with a solution $\bar{x} = 0$. Clearly the mapping $x \mapsto x + \mathbb{R}_+$ is metrically regular at 0 for 0 but not strongly metrically subregular there. The (exact) Newton method has the form $0 \in x_{k+1} + \mathbb{R}_+$ and it generates both convergent and non-convergent sequences from any starting point.

The following result shows that strong metric subregularity of $f + F$, together with assumptions for the mappings R_k that are stronger than in Theorem 3.1, implies convergence of any sequence generated by the method (5) which starts close to $\bar{x}$, but cannot guarantee that the method is surely executable.

Theorem 3.2 (convergence under strong metric subregularity). *Let λ and μ be two positive constants such that $\lambda\mu < 1$. Suppose that the mapping $f + F$ is strongly metrically subregular at $\bar{x}$ for 0 with constant λ. Also, suppose that for each $k = 0, 1, \ldots$, the mapping $(u, x) \mapsto R_k(u, x)$ is partially Aubin continuous with respect to x at $\bar{x}$ for 0 uniformly in u around $\bar{x}$ with constant μ and also satisfies $d^+(0, R_k(u, x)) \to 0$ as $(u, x) \to (\bar{x}, \bar{x})$.*

(i) Let $t \in (0, 1)$ and let there exist positive $\gamma < t(1 - \lambda\mu)/\lambda$ and β such that

$$(21) \qquad d^+(0, R_k(u, \bar{x})) \leq \gamma\|u - \bar{x}\| \quad \text{for all} \quad u \in \mathbb{B}_\beta(\bar{x}) \; k = 0, 1, \ldots.$$

Then there exists a neighborhood $O \subset \mathbb{B}_a(\bar{x})$ of $\bar{x}$ such that for any $x_0 \in O$ every sequence $\{x_k\}$ generated by the Newton method (5) starting from x_0 and staying in O for all k satisfies

$$(22) \qquad \|x_{k+1} - \bar{x}\| \leq t\|x_k - \bar{x}\| \quad \text{for all} \quad k = 0, 1, \ldots,$$

that is, $x_k \to \bar{x}$ q-linearly;

(ii) *Let there exist a sequences of positive scalars $\gamma_k \searrow 0$, with $\gamma_0 < (1 - \lambda\mu)/\lambda$, and $\beta > 0$ such that*

$$(23) \qquad d^+(0, R_k(u, \bar{x})) \le \gamma_k \|u - \bar{x}\| \quad \text{for all} \quad u \in \mathbb{B}_\beta(\bar{x}) \ k = 0, 1, \dots .$$

Then there exists a neighborhood O of $\bar{x}$ such that for any $x_0 \in O$ every sequence $\{x_k\}$ generated by the Newton method (5) starting from x_0 and staying in O for all k, and such that $x_k \ne \bar{x}$ for all k satisfies

$$(24) \qquad \lim_{k \to \infty} \frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|} = 0,$$

that is, $x_k \to \bar{x}$ q-superlinearly;

(iii) *Suppose that the derivative mapping Df is Lipschitz continuous near $\bar{x}$ with Lipschitz constant L and let there exist positive scalars γ and β such that*

$$(25) \qquad d^+(0, R_k(u, \bar{x})) \le \gamma \|u - \bar{x}\|^2 \quad \text{for all} \quad u \in \mathbb{B}_\beta(\bar{x}) \ k = 0, 1, \dots .$$

Then for every

$$(26) \qquad C > \frac{\lambda(\gamma + L/2)}{1 - \lambda\mu}$$

there exists a neighborhood O of $\bar{x}$ such that for any $x_0 \in O$ every sequence $\{x_k\}$ generated by the Newton method (5) starting from x_0 and staying in O for all k satisfies

$$(27) \qquad \|x_{k+1} - \bar{x}\| \le C\|x_k - \bar{x}\|^2 \quad \text{for all} \quad k = 0, 1, \dots .$$

that is, $x_k \to \bar{x}$ q-quadratically.

Proof of (i). Choose t, γ and β as requested and let $\kappa > \lambda$ be such that $\kappa\mu < 1$ and $\gamma < t(1 - \kappa\mu)/\kappa$. Choose positive a and b such that (15) and (17) are satisfied. Pick $\varepsilon > 0$ such that $\gamma + \varepsilon < t(1 - \kappa\mu)$ and adjust a if necessary so that $a \le \beta$ and

$$(28) \qquad \|Df(u) - Df(\bar{x})\| \le \varepsilon \|u - \bar{x}\| \text{ for all } u \in \mathbb{B}_a(\bar{x}).$$

From (21) we have that $R_k(\bar{x}, \bar{x}) = \{0\}$ and then, by the assumptions that $d^+(0, R_k(u, x)) \to 0$ as $(u, x) \to (\bar{x}, \bar{x})$ we can make a so small that $R_k(u, x) \subset \mathbb{B}_b(0)$ whenever $u, x \in \mathbb{B}_a(\bar{x})$.

Let $x_0 \in \mathbb{B}_a(\bar{x})$ and consider any sequence $\{x_k\}$ generated by Newton method (5) starting at x_0 and staying in $\mathbb{B}_a(\bar{x})$. Then there exists $y_1 \in R_k(x_0, x_1) \cap G_{x_0}(x_1)$. From (15) and (28) via (13),

$$\|x_1 - \bar{x}\| \le \kappa\|y_1\| + \kappa\|f(x_0) - Df(x_0)(x_0 - \bar{x}) - f(\bar{x})\| \le \kappa\|y_1\| + \kappa\varepsilon\|x_0 - \bar{x}\|.$$

Since $R_0(x_0, x_1) \subset \mathbb{B}_b(0)$, from (17) there exists $y_1' \in R_0(x_0, \bar{x})$ such that

$$\|y_1 - y_1'\| \le \mu\|x_1 - \bar{x}\|$$

and moreover, utilizing (21),

$$\|y_1'\| \le \gamma\|x_0 - \bar{x}\|.$$

We obtain

$$
\begin{aligned}
\|x_1 - \bar{x}\| &\le \kappa\|y_1\| + \kappa\varepsilon\|x_0 - \bar{x}\| \\
&\le \kappa(\|y_1'\| + \|y_1 - y_1'\|) + \kappa\varepsilon\|x_0 - \bar{x}\| \le \kappa(\gamma + \varepsilon)\|x_0 - \bar{x}\| + \kappa\mu\|x_1 - \bar{x}\|.
\end{aligned}
$$

Hence,

$$\|x_1 - \bar{x}\| \le \frac{\kappa(\gamma + \varepsilon)}{1 - \kappa\mu}\|x_0 - \bar{x}\| \le t\|x_0 - \bar{x}\|,$$

Thus, (22) is established for $k = 0$. We can then repeat the above argument with x_0 replaced by x_1 and so on, obtaining by induction (22) for all k.

Proof of (ii). Choose a sequence $\gamma_k \searrow 0$ with $\gamma_0 < (1 - \lambda\mu)/\lambda$ and $\beta > 0$ such that (23) holds and then pick $\kappa > \lambda$ such that $\kappa\mu < 1$ and $\gamma_0 < (1 - \kappa\mu)/\kappa$. As in the proof of (i), choose $a \le \beta$ and b such that (15) and (17) are satisfied and since $R_k(\bar{x}, \bar{x}) = \{0\}$ from (25), adjust a so that $R_k(u, x) \subset \mathbb{B}_b(0)$ whenever $u, x \in \mathbb{B}_a(\bar{x})$.

Choose $x_0 \in \mathbb{B}_a(\bar{x})$ and consider any sequence $\{x_k\}$ generated by (5) starting from x_0 and staying in $\mathbb{B}_a(\bar{x})$. Since all assumptions in (i) are satisfied, this sequence is convergent to $\bar{x}$. Let $\varepsilon > 0$. Then there exists a natural k_0 such that

$$(29) \quad \|Df(\bar{x} + t(x_k - \bar{x})) - Df(\bar{x})\| \le \varepsilon\|x_k - \bar{x}\| \quad \text{for all} \quad t \in [0, 1] \quad \text{and all} \quad k > k_0.$$

In further lines we mimic the proof of (i). For each $k > k_0$ there exists $y_{k+1} \in R_k(x_k, x_{k+1}) \cap G_{x_k}(x_{k+1})$. From (15) and (29) via (13),

$$\|x_{k+1} - \bar{x}\| \le \kappa\|y_{k+1}\| + \kappa\|f(x_k) - Df(x_k)(x_k - \bar{x}) - f(\bar{x})\| \le \varepsilon\|x_k - \bar{x}\|.$$

By (17) there exists $y_{k+1}' \in R_k(x_k, \bar{x})$ such that

$$\|y_{k+1} - y_{k+1}'\| \le \mu\|x_{k+1} - \bar{x}\|$$

and also, from (25),

$$\|y_{k+1}'\| \le \gamma_k\|x_k - \bar{x}\|.$$

By combining the last three estimates, we obtain

$$
\begin{aligned}
\|x_{k+1} - \bar{x}\| &\le \kappa\|y_{k+1}\| + \kappa\|f(x_k) - Df(x_k)(x_k - \bar{x}) - f(\bar{x})\| \\
&\le \kappa(\|y_{k+1}'\| + \|y_{k+1} - y_{k+1}'\|) + \kappa\varepsilon\|x_k - \bar{x}\| \\
&\le \kappa\gamma_k\|x_k - \bar{x}\| + \kappa\varepsilon\|x_k - \bar{x}\| + \kappa\mu\|x_{k+1} - \bar{x}\|.
\end{aligned}
$$

Hence
$$\|x_{k+1} - \bar{x}\| \le \frac{\kappa}{1 - \kappa\mu}(\gamma_k + \varepsilon)\|x_k - \bar{x}\|.$$

Passing to the limit with $k \to \infty$ we get
$$\lim_{k\to\infty} \frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|} \le \frac{\kappa\varepsilon}{1 - \kappa\mu}.$$

Since ε can be arbitrary small and the expression on the left side does not depend on ε, we obtain (24).

Proof of (iii). Choose γ and β such that (25) holds and then pick C satisfying (26). Take $\kappa > \lambda$ such that $\kappa\mu < 1$ and $C > (\kappa + L/2)/(1 - \kappa\mu)$. Applying Corollary 2.8, choose $a \le \beta$ and b such that (15) and (17) are satisfied and $Ca < 1$. From (25) we have that $R_k(\bar{x}, \bar{x}) = \{0\}$; then adjust a so that $R_k(u, x) \subset \mathbb{B}_b(0)$ whenever $u, x \in \mathbb{B}_a(\bar{x})$. Make a smaller if necessary so that
$$\|Df(u) - Df(v)\| \le L\|u - v\| \quad \text{for all} \quad u, v \in \mathbb{B}_a(\bar{x}).$$

Then, for any $x \in \mathbb{B}_a(\bar{x})$ we have

(30)
$$\begin{aligned}
\|f(x) + Df(x)(\bar{x} - x) - f(\bar{x})\| \\
= \left\| \int_0^1 Df(\bar{x} + t(x - \bar{x}))(x - \bar{x})dt - Df(x)(x - \bar{x}) \right\| \\
\le L \int_0^1 (1 - t)dt \, \|x - \bar{x}\|^2 = \frac{L}{2}\|x - \bar{x}\|^2.
\end{aligned}$$

Let $x_0 \in \mathbb{B}_a(\bar{x})$ and consider a sequence $\{x_k\}$ generated by Newton method (5) starting at x_0 and staying in $\mathbb{B}_a(\bar{x})$ for all k. By repeating the argument of case (ii) and employing (30), we obtain
$$\begin{aligned}
\|x_{k+1} - \bar{x}\| &\le \kappa\|y_{k+1}\| + \kappa\|f(x_k) - Df(x_k)(x_k - \bar{x}) - f(\bar{x})\| \\
&\le \kappa(\|y'_{k+1}\| + \|y_{k+1} - y'_{k+1}\|) + \frac{\kappa L}{2}\|x_k - \bar{x}\|^2 \\
&\le (\kappa\gamma + \frac{\kappa L}{2})\|x_k - \bar{x}\|^2 + \kappa\mu\|x_{k+1} - \bar{x}\|.
\end{aligned}$$

Hence
$$\|x_{k+1} - \bar{x}\| \le \frac{\kappa(\gamma + L/2))}{1 - \kappa\mu}\|x_0 - \bar{x}\|^2 \le C\|x_0 - \bar{x}\|^2.$$

Thus (27) is established. $\qquad\square$

The strong metric subregularity assumed in Theorem 3.2 does not guarantee that the method (5) is surely executable. As a simple example, consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} \frac{1}{2}\sqrt{x+1} & \text{for } x \geq 0, \\ \emptyset & \text{otherwise.} \end{cases}$$

This function is strongly subregular at 0 for 0, but from any point x_0 arbitrarily close to 0 there is no Newton step x_1.

We come to the central result of this paper, whose proof is a combination of the two preceding theorems.

Theorem 3.3 (convergence under strong metric regularity). *Consider the generalized equation (1) and the inexact Newton iteration (5) and let λ and μ be two positive constants such that $\lambda\mu < 1$. Suppose that the mapping $f + F$ is strongly metrically regular at $\bar{x}$ for 0 with constant λ. Also, suppose that for each $k = 0, 1, \ldots$, the mapping $(u, x) \mapsto R_k(u, x)$ is partially Aubin continuous with respect to x at $\bar{x}$ for 0 uniformly in u around $\bar{x}$ with constant μ and satisfies $d^+(0, R_k(u, x)) \to 0$ as $(u, x) \to (\bar{x}, \bar{x})$.*

(i) Let $t \in (0, 1)$ and let there exist positive $\gamma < t(1 - \lambda\mu)\min\{1/\kappa, 1/\mu\}$ and β such that the condition (21) in Theorem 3.2 holds. Then there exists a neighborhood O of $\bar{x}$ such that for any starting point $x_0 \in O$ the inexact Newton method (5) is sure to generate a sequence which stays in O and converges to $\bar{x}$, which may be not unique, but every such sequence is convergent to $\bar{x}$ q-linearly in the way described in (22);

(ii) Let there exist sequences of positive scalars $\gamma_k \searrow 0$, with $\gamma_0 < (1 - \lambda\mu)/\lambda$, and β such that condition (23) in Theorem 3.2 is satisfied. Then there exists a neighborhood O of $\bar{x}$ such that for any starting point $x_0 \in O$ the inexact Newton method (5) is sure to generate a sequence which stays in O and converges to $\bar{x}$, which may be not unique, but every such sequence is convergent to $\bar{x}$ q-superlinearly;

(iii) Suppose that the derivative mapping Df is Lipschitz continuous near $\bar{x}$ with Lipschitz constant L and let there exist positive scalars γ and β such that (25) in Theorem 3.2 holds. Then for every constant C satisfying (26) there exists a neighborhood O of $\bar{x}$ such that for any starting point $x_0 \in O$ the inexact Newton method (5) is sure to generate a sequence which stays in O and converges to $\bar{x}$, which may be not unique, but every such sequence is convergent q-quadratically to $\bar{x}$ in the way described in (27).

If in addition the mapping R_k has a single-valued localization at $(\bar{x}, \bar{x})$ for 0, then in each of the cases (i), (ii) and (iii) there exists a neighborhood O of $\bar{x}$ such that for any starting point $x_0 \in O$ there is a unique Newton sequence $\{x_k\}$ contained in O and this sequence hence converges to $\bar{x}$ in the way described in (i), (ii) and (iii), respectively.

Proof. The statements in (i), (ii) and (iii) follow immediately by combining Theorem 3.2 and Theorem 3.1. Let R_k have a single-valued localization at $(\bar{x}, \bar{x})$ for 0. Choose a and b as above and adjust them so that $R_k(u, x) \in \mathbb{B}_b(0)$ is a singleton for all $u, x \in \mathbb{B}_a(\bar{x})$. Recall that in this case the mapping $x \mapsto R_0(u, x) \cap \mathbb{B}_b(0)$ is Lipschitz continuous on $\mathbb{B}_a((\bar{x}, \bar{x}))$

with constant μ. Then, by observing that $x_1 = G_u^{-1}(-R_0(u, x_1) \cap I\!\!B_b(0)) \cap I\!\!B_a(\bar{x})$ and the mapping $x \mapsto G_u^{-1}(-R_0(u, x) \cap I\!\!B_b(0)) \cap I\!\!B_a(\bar{x})$ is Lipschitz continuous on $I\!\!B_a(\bar{x})$ with a Lipschitz constant $\kappa\mu < 1$, hence it is a contraction, we conclude that there is only one Newton iterate x_1 from x_0 which is in $I\!\!B_a(\bar{x})$. By induction, the same argument works for each iterate x_k. $\qquad\qquad\square$

4 Applications

For the equation $f(x) = 0$ with $f : \mathbb{R}^n \to \mathbb{R}^n$ having a solution $\bar{x}$ at which $Df(\bar{x})$ is nonsingular, it is shown in Dembo et al. [4, Theorem 2.3] that when $0 < \eta_k \le \bar{\eta} < t < 1$, then any sequence $\{x_k\}$ starting close enough to $\bar{x}$ and generated by the inexact Newton method (4) is linearly convergent with

$$(31) \qquad\qquad \|x_{k+1} - x_k\| \le t\|x_k - \bar{x}\|.$$

We will now deduce this result from our Theorem 3.3(i) for X and Y Banach spaces instead of just $\mathbb{R}^n$. A constant of metric regularity of f at $\bar{x}$ could be any real number $\lambda > \|Df(\bar{x})^{-1}\|$.

Fix $\bar{\eta} < t < 1$ and choose a sequence $\eta_k \le \bar{\eta}$. Let $\nu = \max\{\|Df(\bar{x})\|, \|Df(\bar{x})^{-1}\|^{-1}\}$ and choose γ such that $\bar{\eta}\nu < \gamma < \nu$. Then pick $\beta > 0$ to satisfy $\gamma > \bar{\eta}\sup_{x\in B_\beta(x)} \|Df(x)\|$. Finally, choose $\lambda > \|Df(\bar{x})^{-1}\|$ so that $1/\lambda > \gamma$. Then, since $f(\bar{x}) = 0$, for any $u \in I\!\!B_\beta(\bar{x})$ we have

$$(32) \quad d^+(0, R_k(u, \bar{x})) = \eta_k\|f(u)\| = \eta_k\|f(u) - f(\bar{x})\| \le \eta_k \sup_{x\in B_\beta(\bar{x})} \|Df(x)\|\|u - \bar{x}\| \le \gamma\|u - \bar{x}\|.$$

Since in this case $R_k(u) = I\!\!B_{\eta_k\|f(u)\|}(0)$ doesn't depend on x, we can choose as μ any arbitrarily small positive number, in particular satisfying the bounds $\lambda\mu < 1$ and $\gamma < t(1 - \lambda\mu)/\lambda$. Then Theorem 3.3(i) applies and we recover the linear convergence (31) obtained in [4, Theorem 2.3].

For the inexact method (4) with f having Lipschitz continuous derivative near $\bar{x}$, it is proved in [14, Theorem 6.1.4] that when $\eta_k \searrow 0$ with $\eta_0 < \bar{\eta} < 1$, any sequence of iterates $\{x_k\}$ starting close enough to $\bar{x}$ is q-superlinearly convergent to $\bar{x}$. By choosing γ_0, β and λ as γ, β and λ in the preceding paragraph, and then applying (32) with γ replaced by γ_k, this now follows from Theorem 3.3(ii) without assuming Lipschitz continuity of Df.

If we take $R_k(u, x) = I\!\!B_{\eta\|f(u)\|^2}(0)$, we obtain from Theorem 3.3(iii) q-quadratic convergence, as claimed in [14, Theorem 6.1.4]. We note that Dembo et al. [4] gave results characterizing the rate of convergence in terms of the convergence of relative residuals.

When $R_k \equiv 0$ in (5), we obtain from the theorems in Section 3 convergence results for the exact Newton iteration (2) as shown in Theorem 4.1 below. The first part of this theorem is a new result which claims superlinear convergence of any sequence generated by

the method under strong metric subregularity of $f + F$. Under the additional assumption that the derivative mapping Df is Lipschitz continuous around $\bar{x}$ we obtain q-quadratic convergence; this is essentially a known result, for weaker versions see, e.g., [5], [1] and [9, Theorem 6D.1].

Theorem 4.1 (convergence of exact Newton method). *Consider the generalized equation* (1) *with a solution $\bar{x}$ and let the mapping $f + F$ be strongly metrically subregular at $\bar{x}$ for 0. Then the following statements hold for the (exact) Newton iteration* (2):

(i) *There exists a neighborhood O of $\bar{x}$ such that for any starting point $x_0 \in O$ every sequence $\{x_k\}$ generated by* (2) *starting from x_0 and staying in O is convergent q-superlinearly to $\bar{x}$.*

(ii) *Suppose that the derivative mapping Df is Lipschitz continuous near $\bar{x}$. There exists a neighborhood O of $\bar{x}$ such that for any starting point $x_0 \in O$ every sequence $\{x_k\}$ generated by* (2) *and staying in O is q-quadratically convergent to $\bar{x}$.*

If the mapping $f + F$ is not only metrically subregular but actually strongly metrically regular at $\bar{x}$ for 0, then there exists a neighborhood O of $\bar{x}$ such that in each of the cases (i) and (ii) and for any starting point $x_0 \in O$ there is a unique sequence $\{x_k\}$ generated by (2) *and staying in O, and this sequence converges to $\bar{x}$ q-superlinearly or q-quadratically, as described in (i) and (ii).*

We will next propose an inexact Newton method for the variational inequality

$$(33) \qquad \langle f(x), v - x \rangle \leq 0 \ \text{ for all } v \in C \quad \text{or, equivalently,} \quad f(x) + N_C(x) \ni 0,$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ and N_C is the normal cone mapping to the convex *polyhedral* set $C \subset \mathbb{R}^n$:

$$N_C(x) = \begin{cases} \{y \mid \langle y, v - x \rangle \leq 0 \ \text{ for all } v \in C\} \text{ for } x \in C \\ \emptyset \text{ otherwise.} \end{cases}$$

Verifiable sufficient conditions and in some cases necessary and sufficient conditions for (strong) metric (sub)regularity of the mapping $f + N_C$ are given in [9].

For the mapping $V := f + N_C$ it is proved in [8] that when V is metrically regular at $\bar{x}$ for 0, then V is strongly metrically regular there; that is, in this case metric regularity and strong metric regularity are equivalent properties. Let us assume that V is metrically regular at a solution $\bar{x}$ of (33) for 0. If we use the residual $R_k(u) = d(0, f(u) + N_C(u))$ as a measure of inexactness, we may encounter difficulties coming from the fact that the normal cone mapping may be not even continuous. A way to avoid this is to use instead the equation

$$(34) \qquad\qquad\qquad \varphi(x) = P_C(f(x) - x) - x = 0,$$

where P_C is the projection mapping into the set C. As is well known, solving (34) is equivalent to solving (33). Let us focus on the case described in Theorem 3.3(iii). If we use

$R_k(u, x) = I\!B_{\eta_k \|\varphi(u)\|^2}(0)$ we obtain an inexact Newton method for solving (33) in the form

$$(35) \qquad d(0, f(x_k) + Df(x_k)(x_{k+1} - x_k) + N_C(x_{k+1})) \leq \eta_k \|\varphi(x_k)\|^2.$$

Let $\beta > 0$ be such that f in (33) is C^1 in $I\!B_\beta(\bar{x})$. Then φ is Lipschitz continuous on $I\!B_\beta(\bar{x})$ with Lipschitz constant $L \geq 2 + \sup_{u \in B_\beta(\bar{x})} \|Df(u)\|$, and hence condition (25) holds with any $\gamma > \sup_k \eta_k L^2$. Thus, we obtain from Theorem 3.3(iii) that method (35) is sure to generate infinite sequences when starting close to $\bar{x}$ and each such sequence is quadratically convergent to $\bar{x}$. For the case of equation, that is, with $C = \mathbb{R}^n$, this result covers [14, Theorem 6.1.4]. The method (35) seems to be new and its numerical implementation is still to be explored.

As a final application, consider the standard nonlinear programming problem

$$(36) \qquad \text{minimize } g_0(x) \text{ over all } x \text{ satisfying } g_i(x) \begin{cases} = 0 & \text{for } i \in [1, r], \\ \leq 0 & \text{for } i \in [r+1, m] \end{cases}$$

with twice continuously differentiable functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \ldots, m$. Using the Lagrangian

$$L(x, y) = g_0(x) + \sum_{i=1}^{m} g_i(x) y_i$$

the associated Karush-Kuhn-Tucker (KKT) optimality system has the form

$$(37) \qquad f(x, y) + N_E(x, y) \ni (0, 0),$$

where

$$f(x, y) = \begin{pmatrix} \nabla_x L(x, y) \\ -g_1(x) \\ \vdots \\ -g_m(x) \end{pmatrix}$$

and N_E is the normal cone mapping to the set $E = \mathbb{R}^n \times [\mathbb{R}^r \times \mathbb{R}_+^{m-r}]$. It is well known that, under the Mangasarian-Fromovitz condition for the systems of constraints, for any local minimum x of (36) there exists a Lagrange multiplier y, with $y_i \geq 0$ for $i = r+1, \ldots, m$, such that (x, y) is a solution of (37).

Consider the mapping $T : \mathbb{R}^{n+m} \rightrightarrows \mathbb{R}^{n+m}$ defined as

$$(38) \qquad T : z \mapsto f(z) + N_E(z)$$

with f and E as in (37), and let $\bar{z} = (\bar{x}, \bar{y})$ solve (37), that is, $T(\bar{z}) \ni 0$. We recall a sufficient condition for strong metric regularity of the mapping T described above, which can be extracted from [9, Theorem 2G.8]. Consider the nonlinear programming problem (36)

with the associated KKT condition (37) and let $\bar{x}$ be a solution of (36) with an associated Lagrange multiplier vector $\bar{y}$. In the notation

$$I = \left\{\, i \in [1, m] \,\middle|\, g_i(\bar{x}) = 0 \,\right\} \supset \{s+1, \ldots, m\},$$
$$I_0 = \left\{\, i \in [1, s] \,\middle|\, g_i(\bar{x}) = 0 \text{ and } \bar{y}_i = 0 \,\right\} \subset I$$

and

$$M^+ = \left\{\, w \in \mathbb{R}^n \,\middle|\, w \perp \nabla_x g_i(\bar{x}) \text{ for all } i \in I \backslash I_0 \,\right\},$$
$$M^- = \left\{\, w \in \mathbb{R}^n \,\middle|\, w \perp \nabla_x g_i(\bar{x}) \text{ for all } i \in I \,\right\},$$

suppose that the following conditions are both fulfilled:

(a) the gradients $\nabla_x g_i(\bar{x})$ for $i \in I$ are linearly independent,

(b) $\langle w, \nabla_{xx}^2 L(\bar{x}, \bar{y})w \rangle > 0$ for every nonzero $w \in M^+$ with $\nabla_{xx}^2 L(\bar{x}, \bar{y})w \perp M^-$.

Then the mapping T defined in (38) is strongly metrically regular at $(\bar{x}, \bar{y})$ for 0.

The exact Newton method (2) applied to the optimality system (37) consists in generating a sequence $\{(x_k, y_k)\}$ starting from a point (x_0, y_0), close enough to $(\bar{x}, \bar{y})$, according to the iteration

$$(39) \qquad \begin{cases} \nabla_x L(x_k, y_k) + \nabla_{xx}^2 L(x_k, y_k)(x_{k+1} - x_k) + \nabla g(x_k)^\mathsf{T}(y_{k+1} - y_k) = 0, \\ g(x_k) + \nabla g(x_k)(x_{k+1} - x_k) \in N_{\mathbb{R}_+^s \times \mathbb{R}^{m-s}}(y_{k+1}). \end{cases}$$

That is, the Newton method (2) comes down to sequentially solving linear variational inequalities of the form (39) which in turn can be solved by treating them as optimality systems for associated quadratic programs. This specific application of the Newton method is therefore called the *sequential quadratic programming* (SQP) method.

Since at each iteration the method (39) solves a variational inequality, we may utilize the inexact Newton method (35) obtaining convergence in the way described above. We will not go into details here, but rather discuss an enhanced version of (39) called the *sequential quadratically constrained quadratic programming* method. This method has attracted recently the interest of people working in numerical optimization, mainly because at each iteration it solves a second-order cone programming problem to which efficient interior-point methods can be applied. The main idea of the method is to use second-order expansions for the constraint functions, thus obtaining that at each iteration one solves the following optimization problem with a quadratic objective function and quadratic constraints:

$$(40) \qquad \begin{cases} \nabla_x L(x_k, y_k) + \nabla_{xx}^2 L(x_k, y_k)(x_{k+1} - x_k) \\ \qquad + \nabla g(x_k)^\mathsf{T}(y_{k+1} - y_k) + (\nabla^2 g(x_k)(x_{k+1} - x_k))^\mathsf{T}(y_{k+1} - y_k) = 0, \\ g(x_k) + \nabla g(x_k)(x_{k+1} - x_k) \\ \qquad + (\nabla^2 g(x_k)(x_{k+1} - x_k))^\mathsf{T}(x_{k+1} - x_k) \in N_{\mathbb{R}_+^s \times \mathbb{R}^{m-s}}(y_{k+1}). \end{cases}$$

Observe that this scheme fits into the general model of the inexact Newton method (5) if $f + N_C$ is the mapping of the generalized equation, and then, denoting by $z = (x, y)$ the

variable associated with (x_{k+1}, y_{k+1}) and by $w = (u, v)$ the variable associated with (x_k, y_k), consider the "inexactness" term

$$R_k(w, z) = R(w, z) := \begin{pmatrix} (\nabla^2 g(u)(x - u))^\mathsf{T}(y - v) \\ (\nabla^2 (g(u)(x - u))^\mathsf{T}(x - u) \end{pmatrix} \quad \text{for each } k.$$

Clearly, R is Lipschitz continuous with respect to z with an arbitrarily small Lipschitz constant when z and w are close to the primal-dual pair $\bar{z} = (\bar{x}, \bar{y})$ solving the problem and $\|R(w, \bar{z})\| \leq c\|w - \bar{z}\|^2$ for some constant $c > 0$ and for w close to $\bar{z}$. Hence, from Theorem 3.3 we obtain that under the conditions (a) and (b) given above and when the starting point is sufficiently close to $\bar{z}$, the method (40) is sure to generate a unique sequence which is quadratically convergent to the reference point $(\bar{x}, \bar{y})$. This generalizes [10, Theorem 2], where the linear independence of the active constraints, the second-order sufficient condition and the strict complementarity slackness are required. It also complements the result in [12, Corollary 4.1], where the strict Mangasarian-Fromovitz condition and the second-order sufficient condition are assumed.

In this final section we have presented applications of the theoretical results developed in the preceding sections to standard, yet basic, problems of solving equations, variational inequalities and nonlinear programming problems. However, there are a number of important variational problems that go beyond these standard models, such as problems in semidefinite programming, co-positive programming, not to mention optimal control and PDE constrained optimization, for which inexact strategies might be very attractive numerically and still wait to be explored. Finally, we did not consider in this paper ways of globalization of inexact Newton methods, which is another venue for further research.

References

[1] F. J. Aragón Artacho, A. L. Dontchev, M. Gaydu, M. H. Geoffroy, V. M. Veliov, Metric regularity of Newton's iteration, SIAM J. Control Optim. **49** (2011) 339–362.

[2] I. K. Argyros, S. Hilout, Inexact Newton-type methods, J. of Complexity, [**26** (2010), 577-590.

[3] I. K. Argyros, S. Hilout, A Newton-like method for nonsmooth variational inequalities, Nonlinear Analysis **72** (2010) 3857–3864.

[4] R. S. Dembo, S. C. Eisenstat, T. Steihaug, Inexact Newton methods. SIAM J. Numer. Anal. **19** (1982), 400–408.

[5] A. L. Dontchev, Local convergence of the Newton method for generalized equation, C. R. Acad. Sci. Paris, Sér. I, **322** (1996) 327–331.

[6] A. L. Dontchev, V. M. Veliov, Metric regularity under approximation, Control & Cybern. **38** (2009) 1283–1303.

[7] A. L. Dontchev, H. Frankowska, Lyusternik-Graves theorem and fixed points II, J. Convex Analysis, accepted.

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

[9] A. L. Dontchev, R. T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer Mathematics Monographs, Springer, Dordrecht 2009.

[10] D. Fernández, M. Solodov, On local convergence of sequential quadratically-constrained quadratic-programming type methods, with an extension to variational problems, Comput. Optim. Appl. **39** (2008), 143–160.

[11] M. H. Geoffroy, A. Piétrus, Local convergence of some iterative methods for generalized equations, J. Math. Anal. Appl. **290** (2004), 497–505.

[12] A. F. Izmailov, M. V. Solodov, Inexact Josephy-Newton framework for generalized equations and its applications to local analysis of Newtonian methods for constrained optimization, Comput. Optim. Appl. **46** (2010) 347–368.

[13] N. H. Josephy, Newton's method for generalized equations, Technical Summary Report **1965**, University of Wisconsin, Madison 1979.

[14] C. T. Kelley, Solving nonlinear equations with Newton's method, Fundamentals of Algorithms, SIAM, Philadelphia, PA, 2003.

[15] S. M. Robinson, Strongly regular generalized equations. Math. Oper. Res. **5** (1980) 43–62.