TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

Operations Research and Control Systems

Metric Regularity Under Approximations

A.L. Dontchev and V.M. Veliov

Research Report 2010-01
January, 2010

Operations Research and Control Systems
Institute of Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Argentinierstraße 8/E105-4,
1040 Vienna, Austria
E-mail: orcos@eos.tuwien.ac.at

Metric Regularity Under Approximations*

A. L. Dontchev[†] V. M. Veliov[‡]

Abstract

The main goal of this paper is to show that metric regularity and strong metric regularity of a set-valued mapping imply convergence of inexact iterative methods for solving the generalized equation associated with this mapping. For that purpose we first focus on the question exactly how these properties are preserved under changes of the mapping and the reference point. As applications, we consider discrete approximations in optimal control.

Key Words. Metric regularity, inexact iterative methods, Newton's method, proximal point method, discrete approximation, optimal control.

AMS Subject Classification. 49J53, 47J07, 49J40, 49K40, 90C31.

---

*This research was partly supported by by the National Science Foundation (USA) for the first author and by the Austrian Science Foundation (FWF) under grant No P18161-N13 for the second author.

[†]Division of Mathematical Sciences, National Science Foundation, 4201 Wilson Boulevard., Arlington, VA 22230. On leave from Mathematical Reviews, AMS, Ann Arbor, MI, and from the Institute of Mathematics, Bulgarian Academy of Sciences, Sofia, Bulgaria. adontche@nsf.gov.

[‡]Institute of Mathematical Methods in Economics, Vienna University of Technology, A-1040 Vienna, Austria, e-mail: veliov@tuwien.ac.at.

1 Introduction

The Banach open mapping principle says that the surjectivity of a linear and bounded mapping A, acting from a Banach space X to another Banach space Y, is equivalent to the existence of a positive constant κ such that for any $x \in X$ and $y \in Y$ the distance from any x to the set of solutions of the equation $Az = y$ is bounded by κ times the *residual*, that is, the quantity $\|y - Ax\|$. The Banach open mapping theorem has been extended to nonlinear mappings in the works of Lyusternik and Graves. In particular, Graves' theorem asserts that when a function $f : X \to Y$ is continuously Fréchet differentiable in a neighborhood of a point $\bar{x}$, where $f(\bar{x}) = 0$, and its derivative mapping $Df(\bar{x})$ is surjective, then, for a point (x, y) close to $(\bar{x}, 0)$, the distance from x to the set of solutions $f(x) = y$ is, as in the linear case, bounded by a constant times the residual. This latter property has been recognized in the last several decades as a basic constraint qualification condition in optimization and variational analysis, under the name *metric regularity*.

Throughout X and Y are Banach spaces. The notation $g : X \to Y$ means that g is a function (a single-valued mapping) while $G : X \rightrightarrows Y$ denotes a general mapping which may be set-valued. The graph of G is the set $\operatorname{gph} G = \left\{ (x, y) \in X \times Y \,\middle|\, y \in G(x) \right\}$, and the inverse of G is the mapping $y \mapsto G^{-1}(y) = \left\{ x \,\middle|\, y \in G(x) \right\}$. All norms are denoted by $\|\cdot\|$ and the closed ball centered at x with radius r is $\mathbb{B}_r(x)$. The distance from a point x to a set C is denoted by $d(x, C)$ while the excess from a set C to a set D is $e(C, D) = \sup_{y \in C} d(x, D)$. The definition of metric regularity of a general set-valued mapping is as follows:

Definition 1. A mapping $G : X \rightrightarrows Y$ is said to be metrically regular at $\bar{x}$ for $\bar{y}$ when $\bar{y} \in G(\bar{x})$ and there is a constant $\kappa \geq 0$ together with neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$d(x, G^{-1}(y)) \leq \kappa d(y, G(x)) \quad \text{for all } (x, y) \in U \times V.$$

The infimum of κ over all such combinations of κ, U and V is called the *regularity modulus* for F at $\bar{x}$ for $\bar{y}$ and denoted by $\operatorname{reg}(G; \bar{x}|\bar{y})$.

The metric regularity property has come into play recently in various forms in the context of *generalized equations*, that are relations of the form

$$f(x) + F(x) \ni 0, \tag{1}$$

for a function f and a set-valued mapping F. The classical case of an equation corresponds to having $F(x) \equiv 0$, whereas by taking $F(x) \equiv -C$ for a fixed set $C \subset Y$ one gets various (inequality and equality) constraint systems. When Y is the dual X^* of X and F is the normal cone mapping N_C associated with a closed, convex set $C \subset X$; that is, $N_C(x)$ is empty if $x \notin C$, while

$$N_C(x) = \{y \in X^* : y(z - x) \leq 0 \text{ for all } z \in C\} \quad \text{for } x \in C,$$

then (1) becomes a variational inequality.

There are numerous results in the literature that present various forms of the Lyusternik-Graves paradigm; a comprehensive treatment of which accompanied with applications in optimization and beyond can be found in the forthcoming book [6], see also [9] and [13]. Here we recall the following version, which suits our purposes:

Theorem 1. (Lyusternik-Graves). *Let $\bar{x}$ solve the generalized equation* (1) *and suppose that f is strictly differentiable at $\bar{x}$. Then*

$$\operatorname{reg}(f+F;\bar{x}\,|\,0) = \operatorname{reg}(f(\bar{x}) + Df(\bar{x})(\cdot - \bar{x}) + F(\cdot);\bar{x}\,|\,0).$$

When a mapping $G : X \rightrightarrows Y$ is not only metrically regular at $\bar{x}$ for $\bar{y}$ but also its inverse G^{-1} localized around a point of its graph is single valued, then the mapping G is said to be *strongly metrically regular* at $\bar{x}$ for $\bar{y}$, after S. M. Robinson [11]. In this context it is useful to have the concept of a *graphical localization* of a mapping $S : Y \rightrightarrows X$ at $\bar{y}$ for $\bar{x}$, where $\bar{x} \in S(\bar{y})$. By this we mean a mapping with its graph in $Y \times X$ having the form $(V \times U) \cap \operatorname{gph} S$ for some neighborhoods U of $\bar{x}$ and V of $\bar{y}$. It is standardly known, see [6], Section 3E, that when a mapping G is metrically regular at $\bar{x}$ for $\bar{y}$ and moreover its inverse G^{-1} has localization at $\bar{y}$ for $\bar{x}$ which is not multi valued, then G is strongly regular at $\bar{x}$ for $\bar{y}$, which amounts to the existence of neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that the mapping $V \ni y \mapsto G^{-1}(y) \cap U$ is a Lipschitz continuous function with Lipschitz modulus equal to $\operatorname{reg}(G;\bar{x}\,|\,\bar{y})$. For this case we have an inverse function theorem for the generalized equation (1) which fits into the pattern of the Lyusternik-Graves theorem displayed above. Namely, on the condition of Theorem 1 the mapping $f + F$ is strongly metrically regular at $\bar{x}$ for 0 if and only of the associated linearized mapping $f(\bar{x})+Df(\bar{x})(\cdot-\bar{x})+F(\cdot)$ is strongly metrically regular at $\bar{x}$ for 0; moreover, the associated Lipschitz moduli of the localizations of these inverses are the same. This result has its origins in Robinson's implicit function theorem [11], broadly discusses in [6], Section 2B. Another regularity property which obeys the general paradigm displayed in the Lyusternik-Graves theorem is the *strong metric subregularity*, see Section 3I in [6]. In this paper we will limit our considerations to metric regularity and strong metric regularity.

In Section 2 we first focus on the question in what extend metric regularity is preserved under changes of the mapping. For a mapping $f + F$ appearing in the generalized equation (1) with a smooth f, an answer is already contained in the Lyusternik-Graves theorem (Theorem 1) above: for a mapping of the form $f + F$ metric regularity is preserved under linearization of the function f. In the following section we prove a much broader result in which the linearization is replaced by an "approximation" of f in a point near the reference point. We also show that the same type of result also holds for the property of strong metric regularity.

The central results of this paper are presented in Section 3 and are devoted to conditions for convergence of iterative methods for solving generalized equations. We focus on a general two-point iteration which covers *inexact* versions of the classical Newton's method as well as the proximal point method, but also reaches far beyond, both in general ideas and possible applications. As a sample result, we show that metric regularity of the underlaying mapping alone implies the existence of a linearly convergent sequence of iterates provided that the quantity measuring the inexactness is linearly convergent to zero. To our knowledge, inexact iteration methods have not been considered in such generality in the literature.

Section 4 gives applications of the concepts and results presented to discrete approximation in optimal control. For a standard optimal problem we show that metric regularity implies an a priori estimate for the solution of the discretized optimality system. Also, we apply a result from Section 3 to show that the inexact Newton's method associated with the discretization is linearly convergent. Finally, we pose some open problem.

2 Theorems of Lyusternik-Graves and Robinson type

Our first result is a Lyusternik-Graves type theorem that covers much more territory than the classical setting.

Theorem 2. *Consider a continuous function $f : X \to Y$ and a mapping $F : X \rightrightarrows Y$ with closed graph and suppose that $f + F$ is metrically regular at $\bar{x}$ for 0 with constant κ and neighborhoods $I\!B_a(\bar{x})$ and $I\!B_b(0)$ for some positive scalars a and b. Let $\mu > 0$ and κ' be such that $\kappa\mu < 1$ and $\kappa' > \kappa/(1-\kappa\mu)$. Then for every positive constants α and β satisfying*

$$(2) \qquad 2\alpha + 5\kappa'\beta \le a, \quad \mu\alpha + 6\beta \le b \quad \text{and} \quad \alpha \le 2\kappa'\beta,$$

every function $\tilde{f} : X \to Y$, and every $\tilde{x} \in I\!B_\alpha(\bar{x})$ and $\tilde{y} \in I\!B_\beta(0)$ with

$$(3) \qquad \tilde{y} \in \tilde{f}(\tilde{x}) + F(\tilde{x}) \quad \text{and} \quad \|\tilde{f}(\tilde{x}) - f(\tilde{x})\| \le \beta,$$

and

$$(4) \qquad \|[\tilde{f}(x') - f(x')] - [\tilde{f}(x) - f(x)]\| \le \mu\|x' - x\| \quad \text{for every } x', x \in I\!B_{\alpha+5\kappa'\beta}(\tilde{x}),$$

we have that the mapping $\tilde{f} + F$ is metrically regular at $\tilde{x}$ for $\tilde{y}$ with constant κ' and neighborhoods $I\!B_\alpha(\tilde{x})$ and $I\!B_\beta(\tilde{y})$.

The assumptions (3) and (4) describe the way the function $\tilde{f}$ approximates f so that the "approximate" mapping $\tilde{f} + F$ is metrically regular. We note that other versions of this theorem are available in the literature, see e.g. [6], Theorem 5E.1, where, however, the reference point is not perturbed. The main new element here is the specific description of the approximations involved and the constants determining the approximation error. Theorem 4 which comes below is the same type of result but for the strong metric regularity, extending Robinson's theorem [11]. Since these theorems have never been stated in the literature in the form given here, for completeness we supply them with detailed proofs.

In the proof of Theorem 2 we employ the following result:

Theorem 3. ([2]). *Let (X, ρ) be a complete metric space, and consider a set-valued mapping $\Phi : X \rightrightarrows X$, a point $\bar{x} \in X$, and positive scalars r and θ be such that $\theta < 1$, the set* $\operatorname{gph} \Phi \cap (\mathbb{B}_r(\bar{x}) \times \mathbb{B}_r(\bar{x}))$ *is closed and the following conditions hold:*
(i) $d(\bar{x}, \Phi(\bar{x})) < r(1-\theta)$;
(ii) $e(\Phi(u) \cap \mathbb{B}_r(\bar{x}), \Phi(v)) \leq \theta\rho(u, v)$ *for all* $u, v \in \mathbb{B}_r(\bar{x})$. *Then Φ there exists $x \in \mathbb{B}_r(\bar{x})$ such that $x \in \Phi(x)$.*

If Φ is assumed to be a function on X then Theorem 3 follows from the standard contraction mapping principle, see, e.g., [6], Theorem 1A.2 and around, in which case the inequality in (i) does not have to be sharp and θ in (ii) can be zero.

Proof of Theorem 2. By the definition of metric regularity, the mapping $f + F$ satisfies

$$(5) \qquad d(x, (f+F)^{-1}(y)) \leq \kappa d(y, (f+F)(x)) \text{ for every } (x, y) \in \mathbb{B}_a(\bar{x}) \times \mathbb{B}_b(0).$$

Choose $0 < \mu < 1/\kappa$ and $\kappa' > \kappa(1-\kappa\mu)$ and then the constant α and β so that the inequalities in (2) hold. Pick a function $\tilde{f} : X \to Y$ and points $\tilde{x} \in \mathbb{B}_\alpha(\bar{x})$, $\tilde{y} \in \mathbb{B}_\beta(0)$ that satisfy (3) and (4). Let $x \in \mathbb{B}_\alpha(\tilde{x})$ and $y \in \mathbb{B}_\beta(\tilde{y})$. We will first show that

$$(6) \qquad d(x, (\tilde{f}+F)^{-1}(y)) \leq \kappa' \|y - y'\| \text{ for every } y' \in (\tilde{f}(x) + F(x)) \cap \mathbb{B}_{4\beta}(\tilde{y}).$$

Choose $y' \in (\tilde{f}+F)(x) \cap \mathbb{B}_{4\beta}(\tilde{y})$. If $y' = y$ then $x \in (\tilde{f}+F)^{-1}(y)$, and hence (6) holds trivially. Suppose $y' \neq y$ and let $u \in \mathbb{B}_\alpha(\tilde{x})$. Using (3) and (4) and then the second inequality in (2), we have

$$\begin{aligned} \| -\tilde{f}(u) + f(u) + y'\| &\leq \|y' - \tilde{y}\| + \|\tilde{y}\| + \| -\tilde{f}(u) + f(u) + \tilde{f}(\tilde{x}) - f(\tilde{x})\| \\ &\quad + \|\tilde{f}(\tilde{x}) - f(\tilde{x})\| \leq 4\beta + \beta + \mu\|u - \tilde{x}\| + \beta \leq 6\beta + \mu\alpha \leq b. \end{aligned}$$

The same estimate holds of course with y' replaced by y; thus, both $-\tilde{f}(u) + f(u) + y'$ and $-\tilde{f}(u) + f(u) + y$ are in $\mathbb{B}_b(0)$ whenever $u \in \mathbb{B}_\alpha(\tilde{x})$. Consider the mapping

$$(7) \qquad \Phi : u \mapsto (f+F)^{-1}(-\tilde{f}(u) + f(u) + y) \text{ for } u \in \mathbb{B}_\alpha(\tilde{x}).$$

Denote $r := \kappa'\|y - y'\|$ and $\theta := \kappa\mu$. Then $r \leq 5\kappa'\beta$ and hence, from (2), for any $v \in \mathbb{B}_r(x)$ we have

$$\|v - \tilde{x}\| \leq \|v - x\| + \|x - \tilde{x}\| \leq 5\kappa'\beta + \alpha$$

and

$$\|v - \bar{x}\| \leq \|v - \tilde{x}\| + \|\tilde{x} - \bar{x}\| \leq 5\kappa'\beta + 2\alpha \leq a.$$

Thus, $\mathbb{B}_r(x) \subset \mathbb{B}_{5\kappa'\beta+\alpha}(\tilde{x}) \subset \mathbb{B}_a(\bar{x})$. By (4) and the assumed continuity of f, the function $\tilde{f}$ is continuous on $\mathbb{B}_r(x)$. Then, by the continuity of f, $\tilde{f}$ and the closedness of $\operatorname{gph} F$, the set $(\operatorname{gph}\Phi) \cap (\mathbb{B}_r(x) \times \mathbb{B}_r(x))$ is closed. Since $x \in (f+F)^{-1}(-\tilde{f}(x) + f(x) + y') \cap \mathbb{B}_a(\bar{x})$, utilizing (5) we obtain

$$\begin{aligned} d(x, \Phi(x)) &= d(x, (f+F)^{-1}(-\tilde{f}(x) + f(x) + y)) \le \kappa d(-\tilde{f}(x) + f(x) + y, (f+F)(x)) \\ &\le \kappa \| -\tilde{f}(x) + f(x) + y - (y' - \tilde{f}(x) + f(x)) \| = \kappa \|y - y'\| \\ &< \kappa' \|y - y'\| (1 - \kappa\mu) = r(1-\theta), \end{aligned}$$

Moreover, from (5) again we get that for any $u, v \in \mathbb{B}_r(x)$,

$$\begin{aligned} e(\Phi(u) \cap \mathbb{B}_r(x), \Phi(v)) &\le \sup_{z \in (f+F)^{-1}(-\tilde{f}(u)+f(u)+y) \cap B_a(\bar{x})} d(z, (f+F)^{-1}(-\tilde{f}(v) + f(v) + y)) \\ &\le \sup_{z \in (f+F)^{-1}(-\tilde{f}(u)+f(u)+y) \cap B_a(\bar{x})} \kappa d(-\tilde{f}(v) + f(v) + y, f(z) + F(z)) \\ &\le \kappa \| -\tilde{f}(u) + f(u) - [-\tilde{f}(v) + f(v)] \| \le \theta \|u - v\|. \end{aligned}$$

Theorem 3 then yields the existence of a point $\hat{x} \in \Phi(\hat{x}) \cap \mathbb{B}_r(x)$; that is,

$$y \in \tilde{f}(\hat{x}) + F(\hat{x}) \quad \text{and} \quad \|\hat{x} - x\| \le \kappa' \|y - y'\|.$$

Since $\hat{x} \in (\tilde{f} + F)^{-1}(y) \cap \mathbb{B}_r(x)$ we obtain (6).

Now we are ready to prove the desired inequality

$$(8) \qquad d(x, (\tilde{f} + F)^{-1}(y)) \le \kappa' d(y, \tilde{f}(x) + F(x)) \text{ for every } x \in \mathbb{B}_\alpha(\tilde{x}), \quad y \in \mathbb{B}_\beta(\tilde{y}).$$

First, note that if $\tilde{f}(x) + F(x) = \emptyset$, then (8) holds automatically since the right side is $+\infty$. Choose $\varepsilon > 0$ and any $w \in \tilde{f}(x) + F(x)$ such that

$$\|w - y\| \le d(y, \tilde{f}(x) + F(x)) + \varepsilon.$$

If $w \in \mathbb{B}_{4\beta}(\tilde{y})$, then from (6) with $y' = w$ we have that

$$d(x, (\tilde{f} + F)^{-1}(y)) \le \kappa' \|w - y\| \le \kappa' d(y, \tilde{f}(x) + F(x)) + \kappa' \varepsilon,$$

and since the left side of this inequality does not depend on ε, we obtain (8). If $w \notin \mathbb{B}_{4\beta}(\tilde{y})$, then

$$\|w - y\| \ge \|w - \tilde{y}\| - \|y - \tilde{y}\| \ge 3\beta.$$

On the other hand, from (6) applied for $x = \tilde{x}$, $y' = \tilde{y}$, and then from the last inequality in (2), we obtain

$$\begin{aligned} d(x, (\tilde{f} + F)^{-1}(y)) &\le \alpha + d(\tilde{x}, (\tilde{f} + F)^{-1}(y)) \le \alpha + \kappa' \|y - \tilde{y}\| \\ &\le \alpha + \kappa' \beta \le 3\kappa' \beta \le \kappa' \|w - y\| \\ &\le \kappa' d(y, \tilde{f}(x) + F(x)) + \kappa' \varepsilon. \end{aligned}$$

This yields (8) again and we are done. □

Theorem 1 can be deduced from Theorem 2 as follows. Let $f+F$ be metrically regular at $\bar{x}$ for 0, let $\kappa > \operatorname{reg}(f+F;\bar{x}\,|\,0)$ and choose a and b as the radia of the associated balls around $\bar{x}$ and 0. Further, choose $\mu > 0$ and κ' to satisfy $0 < \kappa/(1-\kappa\mu) < \kappa'$ and pick $\alpha > 0$ so that, from the assumed strict differentiability of f at $\bar{x}$, we have

$$\|f(x') - f(x) - Df(\bar{x})(x'-x)\| \le \mu\|x'-x\| \text{ for any } \; x', x \in I\!B_\alpha(\bar{x}).$$

Pick a positive β and adjust α, is necessary, such that the inequalities in (2) are satisfied. Then from Theorem 2 for $\tilde{x} = \bar{x}, \tilde{y} = 0$ and $\tilde{f}(x) = f(\bar{x}) + Df(\bar{x})(x-\bar{x})$ we obtain that $\operatorname{reg}(f(\bar{x}) + Df(\bar{x})(\cdot - \bar{x}) + F(\cdot); \bar{x}\,|\,0) \le \kappa'$. Since κ' can be arbitrarily close to $\operatorname{reg}(f+F;\bar{x}\,|\,0)$ we get

$$\operatorname{reg}(f(\bar{x}) + Df(\bar{x})(\cdot - \bar{x}) + F(\cdot); \bar{x}\,|\,0) \le \operatorname{reg}(f+F;\bar{x}\,|\,0).$$

By using the same argument with $f+F$ and $f(\bar{x}) + Df(\bar{x})(\cdot - \bar{x}) + F(\cdot)$ switched, we obtain the opposite inequality, hence the equality.

The kind of result stated in Theorem 2 can be extended to hold for strong metric regularity, that is, in the case when $(f+F)^{-1}$ is locally a Lipschitz continuous function around the reference point. Similar to Theorem 2, this Robinson-type result, which we present next, can be extracted from combining proofs presented in [6], where the reader can find more results centered around the implicit function theorem paradigm. The direct proof we present next echoes the proof of Theorem 2 in that it uses the standard contraction mapping principle in place of Theorem 3.

Theorem 4. *For a function $f : X \to Y$ and a mapping $F : X \rightrightarrows Y$ with $0 \in f(\bar{x}) + F(\bar{x})$, suppose that $y \mapsto (f+F)^{-1}(y) \cap I\!B_a(\bar{x})$ is a Lipschitz continuous function on $I\!B_b(0)$ with Lipschitz constant κ for positive scalars a and b. Let $\mu > 0$ and κ' be such that $\kappa\mu < 1$ and $\kappa' \ge \kappa/(1-\kappa\mu)$. Then for every positive constants α and β satisfying*

$$(9) \qquad 2\alpha \le a, \quad \mu\alpha + 3\beta \le b \quad \textit{and} \quad \kappa'\beta \le \alpha,$$

for every function $\tilde{f} : X \to Y$, and every $\tilde{x} \in I\!B_\alpha(\bar{x})$ and $\tilde{y} \in I\!B_\beta(0)$ satisfying

$$(10) \qquad \tilde{y} \in \tilde{f}(\tilde{x}) + F(\tilde{x}) \quad \textit{and} \quad \|\tilde{f}(\tilde{x}) - f(\tilde{x})\| \le \beta,$$

and

$$(11) \qquad \|[\tilde{f}(x') - f(x')] - [\tilde{f}(x) - f(x)]\| \le \mu\|x'-x\| \quad \textit{for every } x', x \in I\!B_\alpha(\tilde{x}),$$

we have that the mapping $y \mapsto (\tilde{f}+F)^{-1}(y) \cap I\!B_\alpha(\tilde{x})$ is a Lipschitz continuous function on $I\!B_\beta(\tilde{y})$ with Lipschitz constant κ', that is $\tilde{f}+F$ is strongly metrically regular at $\tilde{x}$ for $\tilde{y}$ with respective constant and neighborhoods.

Proof. Pick μ, κ' as required and then α, β to satisfy (9), then choose $\tilde{f}$ and $(\tilde{x}, \tilde{y})$ that satisfy (10) and (11). First, for any $y \in I\!B_\beta(\tilde{y})$ and any $u \in I\!B_\alpha(\tilde{x})$, noting that $I\!B_\alpha(\tilde{x}) \subset I\!B_a(\bar{x})$ by (9), we have from (10) and (11)

$$\begin{aligned} \| - \tilde{f}(u) + f(u) + y\| &\leq \|y - \tilde{y}\| + \|\tilde{y}\| + \| - \tilde{f}(u) + f(u) + \tilde{f}(\tilde{x}) - f(\tilde{x})\| \\ &\quad + \|\tilde{f}(\tilde{x}) - f(\tilde{x})\| \leq \beta + \beta + \mu\|u - \tilde{x}\| + \beta \leq \mu\alpha + 3\beta \leq b. \end{aligned}$$

By assumption, $y \mapsto s(y) := (f + F)^{-1}(y) \cap I\!B_a(\bar{x})$ is a Lipschitz continuous function on $I\!B_b(0)$ with Lipschitz constant κ. Fix $y \in I\!B_\beta(\tilde{y})$ and consider the function $\Phi(x) = s(-\tilde{f}(x) + f(x) + y)$ on $I\!B_\alpha(\tilde{x})$. Observing that $\tilde{x} = s(-\tilde{f}(\tilde{x}) + f(\tilde{x}) + \tilde{y})$, using (10) and (11), and taking into account (9), for $\theta = \kappa\mu$ we get

$$\begin{aligned} \|\tilde{x} - \Phi(\tilde{x})\| &= \|s(-\tilde{f}(\tilde{x}) + f(\tilde{x}) + \tilde{y}) - s(-\tilde{f}(\tilde{x}) + f(\tilde{x}) + y)\| \\ &\leq \kappa\|\tilde{y} - y\| \leq \kappa\beta \leq \kappa'\beta(1 - \kappa\mu) \leq \alpha(1 - \theta). \end{aligned}$$

Furthermore, for any $u, v \in I\!B_\alpha(\tilde{x})$, from (11),

$$\begin{aligned} \|\Phi(u) - \Phi(v)\| &= \|s(-\tilde{f}(u) + f(u) + y) - s(-\tilde{f}(v) + f(v) + y)\| \\ &\leq \kappa\| - \tilde{f}(u) + f(u) - [\tilde{f}(v) + f(v)]\| \leq \theta\|u - v\|. \end{aligned}$$

Hence, by the standard contraction mapping principle, there exists a unique fixed point $\hat{x} = \Phi(\hat{x})$ in $I\!B_\alpha(\tilde{x})$. Thus the mapping $y \mapsto \tilde{s}(y) := (\tilde{f} + F)^{-1}(y) \cap I\!B_\alpha(\tilde{x})$ is a function defined on $I\!B_\beta(\tilde{y})$. Let $y, y' \in I\!B_\beta(\tilde{y})$. Utilizing the equality $\tilde{s}(y) = s(-\tilde{f}(\tilde{s}(y)) + f(\tilde{s}(y)) + y)$ we obtain

$$\begin{aligned} \|\tilde{s}(y) - \tilde{s}(y')\| &= \|s(-\tilde{f}(\tilde{s}(y)) + f(\tilde{s}(y)) + y) - s(-\tilde{f}(\tilde{s}(y')) + f(\tilde{s}(y')) + y') \\ &\leq \kappa\| - \tilde{f}(\tilde{s}(y)) + f(\tilde{s}(y)) - [-\tilde{f}(\tilde{s}(y')) + f(\tilde{s}(y'))] + \kappa\|y - y'\| \\ &\leq \kappa\mu\|\tilde{s}(y) - \tilde{s}(y')\| + \kappa\|y - y'\|. \end{aligned}$$

Hence

$$\|\tilde{s}(y) - \tilde{s}(y')\| \leq \kappa'\|y - y'\|.$$

This is the desired result: the mapping $y \mapsto \tilde{s}(y) := (\tilde{f} + F)^{-1} \cap I\!B_\alpha(\tilde{x})$ is a Lipschitz continuous function on $I\!B_\beta(\tilde{y})$ with Lipschitz constant κ'. □

Note that in contrast with Theorem 2, in Theorem 4 we can choose κ' equal to $\kappa/(1-\kappa\mu)$. Also note that in the latter we do not need to assume continuity of f and closedness of the graph of F.

3 Convergence of inexact two-point iterations

In this section we consider the following general two-point iterative process for solving the generalized equation (1): Given sequences of functions $r_k : X \to Y$ and $A_k : X \times X \to Y$, and an initial point x_0, generate a sequence $\{x_k\}_{k=0}^{\infty}$ iteratively by taking x_{k+1} to be a solution to the auxiliary generalized equation

$$r_k(x_k) + A_k(x_{k+1}, x_k) + F(x_{k+1}) \ni 0 \quad \text{for } k = 0, 1, \ldots. \tag{12}$$

Here A_k is an approximation of the function f in (1) and the term r_k represents the error (inexactness) in computations. In this section we give conditions on A_k and r_k that ensure the existence of a sequence $\{x_k\}$ generated by the process (12) which converges to a solution $\bar{x}$ of the generalized equation (1), provided that the mapping $f + F$ is metrically regular at $\bar{x}$ for 0. If $f + F$ is strongly metrically regular, then, under these conditions, there is a unique such sequence $\{x_k\}$.

Specific choices of the sequence of mappings A_k lead to known computational methods for solving (1). Under the assumption that f is differentiable with derivative mapping Df, if we take $A_k(x, u) = f(u) + Df(u)(x - u)$ and $r_k = 0$ for all k, the iteration (12) becomes *Newton's method*:

$$f(x_k) + Df(x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \quad \text{for } k = 0, 1, \ldots, \tag{13}$$

If we add the term r_k to the left side of this inclusion, we obtain an inexact version of the method, see [10] for background. There are various ways to choose r_k, but we shall not go into this here. Another inexact version has $A_k(x, v) = f(v) + \Delta_k f(v)(x - v)$ where $\Delta_k f$ is an approximation of the derivative mapping Df. The iteration (13) reduces to the standard Newton's method for solving the nonlinear equation $f(x) = 0$ when F is the zero mapping. In the case when (1) represents the optimality systems for a nonlinear programming problem, the iteration (13) becomes the popular sequential quadratic programming (SQP) algorithm for optimization. See [12] for a predecessor to the general model of two-point iteration process (12).

If we choose $A_k(x, v) = \lambda_k(x - v) + f(x)$ in (12) for some sequence of positive numbers λ_k, we obtain an inexact *proximal point method*:

$$r_k(x_k) + \lambda_k(x_{k+1} - x_k) + f(x_{k+1}) + F(x_{k+1}) \ni 0, \quad \text{for } k = 0, 1, \ldots. \tag{14}$$

This method has received a lot of attention recently in particular in relation to monotone mappings and optimization problems.

Our first result establishes conditions for linear convergence of the iterative process (12).

Theorem 5. *Let the mapping $f + F$ be metrically regular at $\bar{x}$ for 0, let the non-negative numbers ε and μ satisfy*

$$\varepsilon + \mu < \frac{1}{\operatorname{reg}(f + F; \bar{x}\,|\,0)}, \tag{15}$$

and let V be a neighborhood of $\bar{x}$. Then there exists a neighborhood O of $\bar{x}$ such that for any sequences of mappings $r_k : X \to Y$ and $A_k : X \times X \to Y$ with the properties that for all $k = 0, 1, \ldots,$

$$\|f(x) - A_k(x, v) - [f(x') - A_k(x', v)]\| \leq \mu \|x - x'\| \quad \textit{for every } x, x', v \in V \tag{16}$$

and

$$\|r_k(v) + A_k(\bar{x}, v) - f(\bar{x})\| \leq \varepsilon \|v - \bar{x}\| \quad \textit{for every } v \in V, \tag{17}$$

and for any starting point $x_0 \in O$, there exists a sequence $\{x_k\}$ generated by the procedure (12) *and it converges linearly to $\bar{x}$. In addition, if $f + F$ is strongly metrically regular at $\bar{x}$ for* 0*, then the procedure* (12) *generates an unique sequence $\{x_k\}$ in O.*

Proof. Choose $\kappa > \operatorname{reg}(f + F; \bar{x}|0)$ such that, by (15),

$$(\varepsilon + \mu)\kappa < 1. \tag{18}$$

Let a and b be positive numbers such that $f + F$ is metrically regular at $\bar{x}$ for 0 with constant κ and neighborhoods $I\!B_a(\bar{x})$ and $I\!B_b(0)$. Taking a smaller a, if necessary, we may assume that $I\!B_a(\bar{x}) \subset V$. Notice that in the case of a strongly metrically regular $f + F$ (as in the last claim of the theorem) the constants a and b have to be chosen such that the mapping $y \longrightarrow (f + F)^{-1}(y) \cap I\!B_a(\bar{x})$ is single-valued and Lipschitz continuous on $I\!B_b(0)$ with Lipschitz constant κ. Then a can again be decreased, if necessary, so that $I\!B_a(\bar{x}) \subset V$, but also b has to be decreased (so that $\kappa b \leq a$ holds) in order to ensure that $(f + F)^{-1}(y) \cap I\!B_a(\bar{x})$ is still single-valued in $I\!B_b(0)$.

Let κ' satisfy

$$\varepsilon\kappa' < 1, \qquad \kappa' > \frac{\kappa}{1 - \kappa\mu}.$$

Such a κ' exists since $(\varepsilon\kappa)/(1 - \kappa\mu) < 1$ due to $\kappa\mu < 1$ and (18). Choose $\varepsilon' > \varepsilon$ such that $\varepsilon'\kappa' < 1$. Let α and β be chosen so that the conditions (2) hold. Then choose and $\delta > 0$ such that that

$$\delta \leq \alpha \quad \text{and} \quad \varepsilon\delta \leq \beta. \tag{19}$$

Finally, we set $O = I\!B_\delta(\bar{x})$.

Let r_k and A_k satisfy (16) and (17). Let x_0 be an arbitrary point in O and assume that $x_k \in O$ has been already defined for some $k \geq 0$. If $x_k = \bar{x}$ then we set $x_{k+1} = \bar{x}$ which satisfies (12) according to (17) applied for $v = \bar{x}$. Let $x_k \neq \bar{x}$. We apply Theorem 2 with $\tilde{f}(x) = r_k(x_k) + A_k(x, x_k)$, $\tilde{x} = \bar{x}$, $\tilde{y} = r_k(x_k) + A_k(\bar{x}, x_k) - f(\bar{x}) = \tilde{f}(\bar{x}) - f(\bar{x})$. According to (17) and the choice of δ in (19), we have

$$\|\tilde{y}\| = \|\tilde{f}(\tilde{x}) - f(\tilde{x})\| = \|r_k(x_k) + A_k(\bar{x}, x_k) - f(\bar{x})\| \leq \varepsilon \|x_k - \bar{x}\| \leq \varepsilon\delta \leq \beta, \tag{20}$$

and hence the condition (3) in Theorem 2 holds. Further, the condition (4) in Theorem 2 is implied by (16) because $I\!B_{\alpha+5\kappa'\beta} \subset I\!B_a(\bar{x}) \subset V$ according to the first inequality in (2).

Theorem 2 then yields that the mapping $x \mapsto r_k(x_k) + A_k(x, x_k) + F(x)$ is metrically regular at $\bar{x}$ for $\tilde{y}$ with constant κ' and neighborhoods $I\!B_\alpha(\bar{x})$ and $I\!B_\beta(\tilde{y})$. In particular, since $0 \in I\!B_\beta(\tilde{y})$ according to (20), using (17) we obtain

$$\begin{aligned} d\left(\bar{x}, (r_k(\cdot) + A_k(\cdot, x_k) + F(\cdot))^{-1}(0)\right) &\leq \kappa' d\left(0, r_k(x_k) + A_k(\bar{x}, x_k) + F(\bar{x})\right) \\ &\leq \kappa' \|r_k(x_k) + A_k(\bar{x}, x_k) - f(\bar{x})\| \\ &\leq \kappa'\varepsilon\|x_k - \bar{x}\| < \kappa'\varepsilon'\|x_k - \bar{x}\|. \end{aligned}$$

Hence there exists $x_{k+1} \in (r_k(x_k) + A_k(\cdot, x_k) + F(\cdot))^{-1}(0)$, that is, satisfying the iteration (12), which is such that

$$\|x_{k+1} - \bar{x}\| \leq \kappa'\varepsilon'\|x_k - \bar{x}\|. \tag{21}$$

In particular this implies that $x_{k+1} \in O$ due to $\kappa'\varepsilon' < 1$. Clearly, the above inequality applies also in the case $x_k = \bar{x}$. Thus the sequence $x_k \in O$ is well defined by induction and linearly convergent due to (21). If the mapping $f + F$ is strongly metrically regular, we apply Theorem 4 instead of Theorem 2 where α and β now satisfy (9), obtaining that x_{k+1} is the only point in O satisfying (12) and (21). □

Now we will consider the iteration process (12) under somewhat weaker assumptions for the error term r_k than in (17). In particular, $r_k(\bar{x})$ need not be zero, as implied by (17) provided the natural condition $A_k(\bar{x}, \bar{x}) = f(\bar{x})$.

Theorem 6. *Let the mapping $f + F$ be metrically regular at $\bar{x}$ for 0, let ε and μ be non-negative numbers satisfying (15), and let V be a neighborhood of $\bar{x}$. Then there exist $\delta > 0$, $\rho \in (0, 1)$ and $\theta > 0$, such that for any $x_k \in I\!B_\delta(\bar{x})$ and any functions $r_k : X \to Y$ and $A_k : X \times X \to Y$ that satisfy the inequalities*

$$\| [A_k(x', x_k) - f(x')] - [A_k(x, x_k) - f(x)] \| \leq \mu\|x - x'\| \quad \textit{for every } x, x' \in V, \tag{22}$$

and

$$\|A_k(\bar{x}, x_k) - f(\bar{x})\| \leq \varepsilon\|x_k - \bar{x}\|, \qquad \|r_k(x_k)\| \leq \theta, \tag{23}$$

there exists $x_{k+1} \in I\!B_\delta(\bar{x})$ solving (12) and such that

$$\|x_{k+1} - \bar{x}\| \leq \rho\|x_k - \bar{x}\| + C\|r_k(x_k)\| \quad \textit{with} \quad C = \frac{2\,\mathrm{reg}(f + F; \bar{x}\,|\,0)}{1 - \mu\,\mathrm{reg}(f + F; \bar{x}\,|\,0)}. \tag{24}$$

If $f + F$ is strongly metrically regular, then the solution x_{k+1} of (12) is unique in $I\!B_\delta(\bar{x})$.

Proof. Choose a, b, κ, κ' and ε' as in the beginning of the proof of Theorem 5. Since κ can be taken arbitrarily close to $\mathrm{reg}(f+F;\bar{x}\,|\,0)$ we may assume also that

$$(25)\qquad \kappa' < \frac{2\bar{\kappa}}{1-\mu\bar{\kappa}} = C \quad \text{with} \quad \bar{\kappa} = \mathrm{reg}(f+F;\bar{x}\,|\,0).$$

Let α and β be chosen so that the inequalities in (2) hold. Choose $\delta > 0$ so that (19) holds and moreover

$$(26)\qquad \varepsilon\delta < \beta.$$

Finally, set $\rho := \varepsilon'\kappa' < 1$ and specify $\theta > 0$ such that

$$(27)\qquad \theta \le \beta - \varepsilon\delta \quad \text{and} \quad C\theta \le \delta(1-\rho).$$

Choose $x_k \in I\!B_\delta(\bar{x})$, r_k and A_k satisfying (22) and (23). We apply Theorem 2 with $\tilde{f}(x) = r_k(x_k) + A_k(x, x_k)$, $\tilde{x} = \bar{x}$, $\tilde{y} = r_k(x_k) + A_k(\bar{x}, x_k) - f(\bar{x})$. Abbreviating $r_k(x_k) = r_k$ we obviously have

$$\tilde{y} = r_k + A_k(\bar{x}, x_k) - f(\bar{x}) = \tilde{f}(\bar{x}) - f(\bar{x}) \in \tilde{f}(\bar{x}) + F(\bar{x}),$$

and then, using (23),

$$\|\tilde{f}(\bar{x}) - f(\bar{x})\| = \|\tilde{y}\| = \|r_k + A_k(\bar{x}, x_k) - f(\bar{x})\| \le \|r_k\| + \varepsilon\|x_k - \bar{x}\| \le \theta + \varepsilon\delta \le \beta,$$

where we use (26) and the first inequality in (27). Thus (3) holds. The condition (4) follows from (22) since $I\!B_{\alpha+5\kappa'\beta}(\bar{x}) \subset I\!B_a(\bar{x}) \subset V$ due to the choice of a in the beginning of the proof of Theorem 5, and the first inequality in (2). Then, according to Theorem 2, we have

$$d(\bar{x}, (\tilde{f}+F)^{-1}(0)) \le \kappa' d(0, \tilde{f}(\bar{x}) + F(\bar{x})) \le \kappa'\|\tilde{y}\| < \kappa'\|r_k\| + \kappa'\varepsilon'\|x_k - \bar{x}\|.$$

Notice that the last inequality is strict only if $x_k \ne \bar{x}$ or $r_k \ne 0$, which we assume for the moment. Hence, there exists $x_{k+1} = (\tilde{f}+F)^{-1}(0)$, that is, satisfying (12), such that

$$(28)\qquad \|x_{k+1} - \bar{x}\| \le \kappa'\|\tilde{y}\| \le \kappa'\|r_k\| + \kappa'\varepsilon'\|x_k - \bar{x}\| \le \rho\|x_k - \bar{x}\| + C\|r_k\|.$$

In the case $x_k = \bar{x}$ and $r_k(\bar{x}) = 0$ we may choose $x_{k+1} = \bar{x}$, which solves (12) and obviously satisfies the above inequality. It remains to note that $x_{k+1} \in I\!B_\delta(\bar{x})$ due to (28) and the second inequality in (27).

In the case of strong metric regularity of $f+F$ we use Theorem 4 in place of Theorem 2, as in the end of the proof of Theorem 5, to show that x_{k+1} is unique in $\in I\!B_\delta(\bar{x})$. □

The proof of the theorem shows that one can take ρ to be any number from the non-degenerate interval $(\varepsilon\bar{\kappa}/(1-\mu\bar{\kappa}), 1)$. The number δ is independent of the choice of ρ, but θ may depend on it.

The essence of the above theorem is that if at any step k of the iterative process (12) the approximation mapping A_k is chosen in such a way that it sufficiently well approximates f (in the sense of (22) and the first inequality in (23)) and the respective error term $r_k(x)$ is sufficiently small for the current iteration x_k (i.e. $\|r_k(x_k)\| \leq \theta$), then a next iteration x_{k+1} exists (and is unique in the case of strong metric regularity), which satisfies (24). In particular, if the initial x_0 is sufficiently close to $\bar{x}$, then the iterative process can be infinitely continued, generating a sequence $\{x_k\}$. By a standard induction argument this sequence satisfies the error estimation

$$\|x_k - \bar{x}\| \leq \rho^k \|x_0 - \bar{x}\| + C \sum_{i=0}^{k-1} \rho^i \|r_{k-i}(x_{k-i})\|.$$

In particular, if $r_k(x_k)$ converges linearly to zero, then the sequence $\{x_k\}$ converges to $\bar{x}$ linearly as well. If $f + F$ is strongly metrically regular, then each x_k is unique in $I\!B_\delta(\bar{x})$. To verify the first claim we observe that if $\|r_k(x_k)\| \leq c\gamma^k$ for some constants $\gamma \in (0,1)$ and c and all k, then $\|r_k(x_k)\| \leq c'\gamma'^k/k^2$ for some $\gamma' \in (\gamma, 1)$ and c'. Hence, $\|x_k - \bar{x}\|$ can be estimated by $Cc'(\max\{\rho, \gamma'\})^k \sum_{i=0}^{\infty} 1/k^2$, which converges linearly to zero.

We will now consider the iteration (12) from a different standpoint. We will give conditions on r_k and A_k under which for any sequence generated by (12) there also exists a sequence of the exact version of (12), the one with $r_k = 0$, which starts from the same x_0 and is at distance proportional to $\{r_k\}$. Specifically, we have the following theorem:

Theorem 7. *Let the mapping $f + F$ be metrically regular at $\bar{x}$ for 0, let $\mu \geq 0$ and ρ satisfy $\mu \operatorname{reg}(f + F; \bar{x}|0) < \rho < 1$ and let V be a neighborhood of $\bar{x}$. Then there exist $\theta > 0$ and $\delta > 0$ such that for every sequences of mappings $r_k : X \to Y$ and $A_k : X \times X \to Y$ that satisfy for every $k = 0, 1, \ldots$*

$$\sup_{x \in V} \|r_k(x)\| \leq \theta \tag{29}$$

and

$$\|f(x) - A_k(x, v) - [f(x') - A_k(x', v')]\| \leq \mu(\|x - x'\| + \|v - v'\|) \text{ for all } x, x', v, v' \in V, \tag{30}$$

if a sequence $\{x_k\}$ is generated by (12) starting from a point $x_0 \in I\!B_\delta(\bar{x})$ and contained in $I\!B_\delta(\bar{x})$, there exists a sequence $\{x'_k\}$, generated again by the (12) but with $r_k = 0$ and starting from the same initial condition x_0, such that

$$\|x'_{k+1} - x_{k+1}\| \leq C \sum_{i=0}^{k} \rho^i \|r_{k-i}(x_{k-i})\| \quad \text{for all} \quad k, \tag{31}$$

where C is given in (24).

Proof. Choose $\kappa > \operatorname{reg}(f+F;\bar{x}|0)$ such that $\mu\kappa < \rho$ and let a and b be positive scalars such that $f+F$ is metrically regular at $\bar{x}$ for 0 with constant κ and neighborhoods $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(0)$. Take a smaller a if necessary so that $\mathbb{B}_a(\bar{x}) \subset V$ (see the note at the beginning of the proof of Theorem 5). Then choose κ' to satisfy

$$\mu\kappa' < \rho \quad \text{and} \quad C > \kappa' > \frac{\kappa}{1-\kappa\mu}.$$

Pick α and β so that (2) holds, then $\delta > 0$ to satisfy

$$3\mu\delta \leq \beta, \quad \alpha + \delta \leq a \quad \text{and} \quad 2\delta \leq a,$$

and finally $\theta > 0$ such that

$$\frac{C\theta}{1-\rho} \leq \delta.$$

Choose r_k and A_k that satisfy the conditions in the statement and a sequence $x_k \in \mathbb{B}_\delta(\bar{x})$ generated by (12) and starting from some $x_0 \in \mathbb{B}_\delta(\bar{x})$. By induction, let $x'_k \in \mathbb{B}_{2\delta}(\bar{x})$ be obtained by (12) but with $r_k = 0$ which has $x'_0 = x_0$ and satisfies (31) up to certain k. If $r_i(x_i) = 0$ for all $i = 0, \ldots, k$, then we take $x'_{k+1} = x_{k+1}$ and the induction step is complete. Let $r_i(x_i) \neq 0$ for some $i \in \{0, \ldots, k\}$. To prove that this holds for $k+1$, we apply Theorem 2 with

$$\tilde{x} = x_{k+1}, \quad \tilde{f}(x) = A_k(x, x'_k), \quad \tilde{y} = -r_k(x_k) + A_k(x_{k+1}, x'_k) - A_k(x_{k+1}, x_k).$$

Then of course $\tilde{y} \in \tilde{f}(\tilde{x}) + F(\tilde{x})$. Let's check the rest of the conditions in 2. Noting that from (30) $A_k(\bar{x},\bar{x}) - f(\bar{x}) = 0$, we have

$$\begin{aligned} \|A_k(x_{k+1}, x'_k) - f(x_{k+1})\| &\leq \|A_k(x_{k+1}, x'_k) - f(x_{k+1}) - [A_k(\bar{x},\bar{x}) - f(\bar{x})]\| \\ &\leq \mu\|x_{k+1} - \bar{x}\| + \mu\|x'_k - \bar{x}\| \leq 3\mu\delta \leq \beta, \end{aligned}$$

and hence the condition (3) in Theorem 2 holds. Also, from (30), for any $x, x' \in \mathbb{B}_\alpha(x_{k+1}) \subset \mathbb{B}_a(\bar{x}) \subset V$,

$$\|f(x) - A_k(x, x'_k) - [f(x') - A_k(x', x'_k)]\| \leq \mu\|x - x'\|.$$

Thus, we can apply Theorem 2 according to which

$$\begin{aligned} &d(x_{k+1}, (\tilde{f}+F)^{-1}(0)) \leq \kappa' d(0, A_k(x_{k+1}, x'_k) + F(x_{k+1})) \\ &\quad \leq \kappa'\|\tilde{y}\| = \kappa'\| - r_k(x_k) + A_k(x_{k+1}, x'_k) - A_k(x_{k+1}, x_k)\| \\ &\quad \leq \kappa'\|f(x_{k+1}) - A_k(x_{k+1}, x'_k) - [f(x_{k+1}) - A_k(x_{k+1}, x_k)]\| + \kappa'\|r_k(x_k)\| \\ &\quad \leq \kappa'\mu\|x'_k - x_k\| + \kappa'\|r_k(x_k)\| \\ &\quad < \rho C \sum_{i=1}^{k} \rho^i \|r_{k-i}(x_{k-i})\| + C\|r_k(x_k)\| \leq C \sum_{i=1}^{k} \rho^i \|r_{k-i}(x_{k-i})\|. \end{aligned}$$

The sharp inequality before the last comes from $\kappa'\mu < \rho$ if the first term (the sum) is nonzero or from $\kappa' < C$ if the sum is zero, in which case all $r_i(x_i) = 0$ for $i = 0, 1, \ldots, k-1$, but then in the second term $\|r_k(x_k)\| > 0$. Hence, there exists $x'_{k+1} \in (A_k(\cdot, x'_k) + F(\cdot))^{-1}(0)$, that is, x'_{k+1} is an exact iterate of (12), which satisfies the desired estimate (31) for $k+1$. Moreover,

$$\|x'_{k+1} - \bar{x}\| \leq \|x_{k+1} - \bar{x}\| + \|x'_{k+1} - x_{k+1}\| \leq \delta + \frac{C\theta}{1-\rho} \leq 2\delta,$$

and the proof is complete. □

The strong regularity version of Theorem 7 will have in addition that the elements of the reference sequence for the iteration with r_k and the one for $r_k = 0$ will be unique in a neighborhood of $\bar{x}$. Note that the conditions (16) in Theorem 5 as well as (22) and (23) in Theorem 6 are implied by (30) (for (23) provided that $f(\bar{x}) + A_k(\bar{x}, \bar{x}) = 0$).

We will now show what the conditions (16) and (17) mean for the Newton method and the proximal point method given in the beginning of this section. For the Newton method (13) we have $A_k(x, v) = f(v) + Df(v)(x - v)$ for all k, and then, if we assume continuous differentiability of f near $\bar{x}$, for any $\mu > 0$ there exists a neighborhood V of $\bar{x}$ such that

$$\begin{aligned}\|f(x) - f(x') - Df(v)(x - x')\| \;\leq\; & \|f(x) - f(x') - Df(\bar{x})(x - x')\| \\ & + \|Df(v) - Df(\bar{x})\|\|x - x'\| \leq \mu\|x - x'\|\end{aligned}$$

for all $x, x' \in V$. Further, the continuous differentiability of f is sufficient to have that for any $\varepsilon > 0$ there exists a neighborhood V of $\bar{x}$ such that

$$\|f(v) - Df(v)(\bar{x} - v) - f(\bar{x})\| \leq \varepsilon\|v - \bar{x}\| \quad \text{for any } v \in V.$$

If, additionally, Df is Lipschitz around $\bar{x}$, then also (30) can be easily verified for any positive μ, if the neighborhood V is taken correspondingly sufficiently small.

For the proximal point method (14) the expression on the left side of (16) is just $\lambda_k(x - x')$ and the left side of (17) is $\lambda_k(v - \bar{x})$, thus both (16) and (17) come down to the condition that each λ_k is less than the reciprocal of $2\,\mathrm{reg}(f + F; \bar{x}\,|\,0)$. Condition (30) obviously holds if $\lambda_k \leq \mu$.

4 Some results and open questions for discretization in optimal control

Consider the following optimal control problem

$$\text{(32)} \qquad \text{minimize} \int_0^1 \varphi(p(t), u(t))\,\mathrm{d}t$$

subject to

$$\dot{p}(t) = g(p(t), u(t)), \quad u(t) \in U \text{ for a.e. } t \in [0,1],$$
$$p \in W_0^{1,\infty}(I\!R^n), \quad u \in L^\infty(I\!R^m),$$

where $\varphi : I\!R^{n+m} \to I\!R$, $g : I\!R^{n+m} \to I\!R^n$, U is a convex and closed set in $I\!R^m$. Here p denotes the state trajectory of the system, u is the control function, $L^\infty(I\!R^m)$ denotes the space of essentially bounded and measurable functions with values in $I\!R^m$ and $W_0^{1,\infty}(I\!R^n)$ is the space of Lipschitz continuous functions p with values in $I\!R^n$ and such that $p(0) = 0$. We assume that problem (32) has a solution $(\bar{p}, \bar{u})$ and also that the functions φ and g are twice continuously differentiable in an open set containing all values $(\bar{p}(t), \bar{u}(t))$ for $t \in [0,1]$.

Let $W_1^{1,\infty}(I\!R^n)$ be the space of Lipschitz continuous functions q with values in $I\!R^n$ and such that $q(1) = 0$. In terms of the Hamiltonian

$$H(p, u, q) = \varphi(p, u) + q^\mathsf{T} g(p, u),$$

it is well known that the first-order necessary optimality conditions at the solution $(\bar{p}, \bar{u})$ can be expressed in the following way: there exists $\bar{q} \in W_1^{1,\infty}(I\!R^n)$, such that $\bar{x} := (\bar{p}, \bar{u}, \bar{q})$ is a solution of the following two-point boundary value problem coupled with a variational inequality

$$(33) \qquad \begin{cases} \dot{p}(t) &= g(p(t), u(t)), \quad p(0) = 0, \\ \dot{q}(t) &= -\nabla_p H(p(t), u(t), q(t)), \quad q(1) = 0, \\ 0 &\in \nabla_u H(p(t), u(t), q(t)) + N_U(u(t)), \qquad \text{for a.e. } t \in [0,1], \end{cases}$$

where $N_U(u)$ is the normal cone to the set U at the point u. Denote $X = W_0^{1,\infty}(I\!R^n) \times W_1^{1,\infty}(I\!R^n) \times L^\infty(I\!R^m)$ and $Y = L^\infty(I\!R^n) \times L^\infty(I\!R^m) \times L^\infty(I\!R^n)$.

Further, for $x = (p, q, u)$ let

$$(34) \qquad f(x) = \begin{pmatrix} \dot{p} - g(p(t), u(t)) \\ \dot{q} + \nabla_p H(p(t), u(t), q(t)) \\ \nabla_u H(p(t), u(t), q(t)) \end{pmatrix}$$

and

$$(35) \qquad F(x) = \begin{pmatrix} 0 \\ 0 \\ N_U(u) \end{pmatrix}.$$

Then the optimality system (33) (satisfied for a.e. t due to the Pontryagin maximum principle) can be written as the generalized equation (1).

In this section we will show that metric regularity of the mapping $f + F$ for the optimality systems above implies an *a priori* error estimate for a discrete approximation to the system.

A sufficient condition for *strong* metric regularity of the mapping $f + F$ for a system of the type (33), based on so-called *coercivity*, see [3]. Strong metric regularity in appropriate metric for problems which are affine with respect to the control (hence non-coercive) are given in [7, 8]. However, the known conditions for (strong) metric regularity are only sufficient and seemingly far from necessary, and also apply to limited classes of problems. Necessary and sufficient conditions for strong metric regularity plus optimality for an optimal control problem are obtained in [4]. Finding sharp conditions for metric regularity in optimal control is a challenging avenue for further research.

Suppose that the optimality system (33) is solved inexactly by means of a numerical method applied to a discrete approximation provided by Euler scheme. Specifically, let N be a natural number, let $h = 1/N$ be the mesh spacing, and let $t_i = ih$. Denote by $PL_0^N(I\!R^n)$ the space of piecewise linear and continuous functions p_N over the grid $\{t_i\}$ with values in $I\!R^n$ and such that $p_N(0) = 0$, by $PL_1^N(I\!R^n)$ the space of piecewise linear and continuous functions q_N over the grid $\{t_i\}$ with values in $I\!R^n$ and such that $q_n(0) = 0$, and by $PC^N(I\!R^m)$ the space of piecewise constant and continuous from the right functions over the grid $\{t_i\}$ with values in $I\!R^m$. Then introduce the products $X^N = PL_0^N(I\!R^n) \times PL_1^N(I\!R^n) \times PC^N(I\!R^m)$ as an approximation space for the triple (p, q, u). We identify $p \in PL_0^N(I\!R^n)$ with the vector $(p^0, \ldots, p^N)$ of its values at the mesh points (and similarly for q), and $u \in PC^N(I\!R^m)$ – with the vector $(u^0, \ldots, u^{N-1})$ of the values of u in the mesh subintervals.

Now, suppose that, as a result of the computations, for certain natural N a function $\tilde{x} = (p_N, q_N, u_N) \in X^N$ is found that satisfies the modified optimality system

$$(36) \qquad \left\{ \begin{array}{lcl} \dot{p}^i & = & g(p^i, u^i) \quad p^0 = 0, \\ \dot{q}^i & = & \nabla_p H(p^i, u^i, q^{i+1}) \quad q^N = 0, \\ 0 & \in & \nabla_u H(p^i, u^i, q^i) + N_U(u^i) \end{array} \right.$$

for $i = 0, 1, \ldots, N-1$ and consistently with the piece-wise linearity of p and q

$$\dot{p}^i = \frac{p^{i+1} - p^i}{h}.$$

The system (36) represents the Euler discretization of the optimality system (33).

Suppose that the mapping $f + F$ is metrically regular at $\bar{x}$ for 0. Then there exists positive scalars a and κ such that if $\tilde{x} \in I\!B_a(\bar{x})$, then

$$d(\tilde{x}, (f+F)^{-1}(0)) \leq \kappa d(0, f(\tilde{x}) + F(\tilde{x})),$$

where the right side of this inequality is the residual associated with the approximate solution $\tilde{x}$. In our specific case, the residual can be estimated by the norm of the function $\tilde{y}$ defined as follows:

$$\tilde{y}(t) = \left(\begin{array}{c} g(p_N(t_i), u_N(t_i)) - g(p_N(t), u_N(t_i)) \\ \nabla_x H(p_N(t_i), u_N(t_i), q_N(t_{i+1})) - \nabla_x H(p_N(t), u_N(t_i), q_N(t)) \\ \nabla_u H(p_N(t_i), u_N(t_i), q_N(t_i)) - \nabla_u H(p_N(t), u_N(t_i), q_N(t)) \end{array} \right), \quad t \in [t_i, t_{i+1}).$$

We have the estimate

$$\begin{aligned}\|\tilde{y}\| \leq & \max_{0\le i\le N-1} \sup_{t_i\le t\le t_{i+1}} [\,|g(p_N(t_i),u_N(t_i)) - g(p_N(t),u_N(t_i))| \\ & +|\nabla_x H(p_N(t_i),u_N(t_i),q_N(t_{i+1})) - \nabla_x H(p_N(t),u_N(t_i),q_N(t))| \\ & +|\nabla_u H(p_N(t_i),u_N(t_i),q_N(t_i)) - \nabla_u H(p_N(t),u_N(t_i),q_N(t))|\,].\end{aligned}$$

Observe that here p_N is a piecewise linear function across the grid $\{t_i\}$ with uniformly bounded derivative, since both p_N and u_N are in some L_∞ neighborhood of $\bar{p}$ and $\bar{u}$ respectively. Hence, taking into account that the functions g, $\nabla_x H$ and $\nabla_u H$ are continuously differentiable we obtain the following result:

Theorem 8. *Assume that the mapping of the optimality system (33) is metrically regular at $\bar{x} = (\bar{p}, \bar{q}, \bar{u})$ for 0. Then there exist constants a and c such that if the L_∞ distance from a solution $\tilde{x} = (p_N, q_N, u_N)$ to the discretized system (36) to $\bar{x}$ is not more than a, then there exists a solution $\bar{x}^N = (\bar{p}^N, \bar{q}^N, \bar{u}^N)$ of (33) such that*

$$\|\bar{p}^N - p_N\|_{W_0^{1,\infty}} + \|\bar{q}^N - q_N\|_{W_1^{1,\infty}} + \|\bar{u}^N - u_N\|_{L^\infty} \leq ch.$$

If the mapping of the optimality system (33) is strongly metrically regular at $\bar{x}$ for 0 then the above claim holds with $\bar{x}^N = \bar{x}$.

The last claim in the above result regarding the strong metric regularity case can be viewed as follows: there is a ball around $\bar{x}$ such that if $x_N = (p_N, q_N, u_N)$ is a sequence of approximate solutions to the discretized system (36) contained in this ball, then x_N converges to $\bar{x}$ with rate proportional to $1/N$.

A similar a priory error estimate is obtained in [1] under a coercivity condition acting on the discretized system (36) which implies strong metric regularity. We can obtain *a posteriori* error estimates provided that the mapping of discretized system (36) is metrically regular, say, at $\tilde{x}$ for $\tilde{y}$, uniformly in N. The system (36) fits into the approximate mapping $\tilde{f} + F$ in Section 2 but now also with approximation of the spaces X and Y with subspaces X_N and Y_N which, in the specific case considered here, are spaces of piecewise linear functions for the state and costate and piecewise constant functions for the control, and associate piecewise constant functions for Y. We were not able to find theorems of Lyusternik-Graves and Robinson type displayed in Section 2 which would also involve approximation of elements of X and Y by elements of subspaces X_N and Y_N. A main difficulty seems to be the fact that the property of metric regularity is not necessarily inherited by the restriction of the mapping on a subspace, as the following counterexample shows.

Let $X = I\!R^2$, $Y = I\!R$, $f(x_1, x_2) = x_2 - x_1^3$. Here

$$f^{-1}(y) = \{(x_1, x_2) :\ x_2 = y + x_1^3,\ x_1 \in I\!R\}.$$

The function f is metrically regular at $x = (0, 0)$ for $y = 0$ with $\kappa = 1$, since

$$d(x, f^{-1}(y)) \leq |(x_1, x_2) - (x_1, y + x_1^3)| = |y - (x_2 - x1^3)| = |y - f(x)|.$$

On the other hand, the restriction of f to $\tilde{X} = \{(x_1, x_2) : \ x_2 = 0\}$ is not metrically regular at $x_1 = 0$ for $y = 0$ because for $x \in \tilde{X}$ we have $f(x) = -x_1^3$, hence $x_1 = (-y)^{1/3}$, which is not Lipschitz at $y = 0$.

Now we turn to an application of Theorem 6 for proving convergence of a discretized (finite-dimensional) version of the Newton method for the problem (32).

The Newton mapping A_k in this case is defined for $x = (p, u, q)$, $v \in X$ as

$$A_k(x, v) = A(x, v) = \begin{pmatrix} \dot{p} - \nabla_q H(v) - \nabla^2{}_{qx} H(v)(x - v) \\ \dot{q} + \nabla_p H(v) + \nabla^2{}_{px} H(v)(x - v) \\ \nabla_u H(v) - \nabla^2{}_{ux} H(v)(x - v) \end{pmatrix}.$$

The Newton iterative process with discretization is defined as follows.

Discretized Newton Process: Let N_0 be a natural number and let $u_0 \in PC^{N_0}(I\!R^m)$ be an an initial guess for the control. Let p_0 and q_0 be the corresponding solutions of the Euler discretization of the primal and adjoint system in (36). Obviously p_0 and q_0 can be viewed as piece-wise linear functions, thus the initial approximation $x_0 = (p_0, u_0, q_0)$ belongs to the space X^{N_0}. Inductively, we assume that the k-th iteration $x_k \in X^{N_k}$ has already been defined, as well as a next mesh size $N_{k+1} = \nu_k N_k$, where ν_k is a natural number (that is, the current mesh points $\{t_i^k = i/N_k\}_{i=0,\ldots,N_k}$ are embedded in the next mesh $\{t_i^{k+1} = i/N_{k+1}\}_{i=0,\ldots,N_{k+1}}$). Then, let $x = x_{k+1} = \{x_{k+1}^i\}_i = \{(p_{k+1}^i, u_{k+1}^i, q_{k+1}^i)\}_i \in I\!R^{N_{k+1} \times n} \times I\!R^{N_{k+1} \times m} \times I\!R^{N_{k+1} \times n}$ be a solution of the discretized version of the Newton method:

$$\begin{pmatrix} \frac{p^{i+1} - p^i}{h_{k+1}} - \nabla_q H(x_k(t_{k+1}^i)) - \nabla_{qx}^2 H(x_k(t_{k+1}^i))(x^i - x_k(t_{k+1}^i)) \\ \frac{q^i - q^{i-1}}{h_{k+1}} + \nabla_p H(x_k(t_{k+1}^i)) + \nabla_{px}^2 H(x_k(t_{k+1}^i))(x^i - x_k(t_{k+1}^i)) \\ \nabla_u H(x_k(t_{k+1}^i)) - \nabla_{ux}^2 H(x_k(t_{k+1}^i))(x^i - x_k(t_{k+1}^i)) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ N_U(u^i) \end{pmatrix} \ni 0,$$

with $p_{k+1}^0 = 0$, $q_{k+1}^{N_{k+1}} = 0$, and where $h_{k+1} = 1/N_{k+1}$.[1] The sequence $\{x^i\}_{i=0,\ldots,N_{k+1}}$ is then embedded into the space $X^{N_{k+1}}$ by piecewise linear interpolation for the p and q components, and piecewise constant interpolation for the u component (so that $u_{k+1}(t) = u_{k+1}^i$ on $[t_{k+1}^i, t_{k+1}^{i+1})$). We use the same notation x_{k+1} for the so obtained next iteration belonging to the space $X^{N_{k+1}}$. In this way we obtain a sequence $x_k \in X^{N_k}$, assuming that a solution of the discretized Newton method exists at each step, although no uniqueness is a priory assumed (see the conjecture at the end of the section).

[1] We keep the argument x in the appearing derivatives of H, although in fact, $\nabla_q H$ and $\nabla_{qx}^2 H$ depend only on p and u.

The next theorem asserts that in case of strong metrical regularity of the mapping of the optimality system (33), if the discretized Newton iteration process described above starts from an initial guess $x_0 \in X^{N_0}$ which is sufficiently close to the solution $\bar{x}$ and if the sequence of discretization steps h_k converges linearly to zero, then also the sequence x_k converges linearly to $\bar{x}$ in the space $X = W_0^{1,\infty}(I\!R^n) \times W_1^{1,\infty}(I\!R^n) \times L^\infty(I\!R^m)$.

Theorem 9. *Let the mapping $f + F$ with the specifications (34), (35) (that is, the mapping of the optimality system (33) be strongly metrically regular at $\bar{x}$ for 0. Let the Hamiltonian H be twice continuously differentiable around $\bar{x}$. Then there exist constants $\delta > 0$ and $\bar{N}$ such that for every sequence $N_k = \nu^k N_0$, with $N_0 \geq \bar{N}$ and a natural number $\nu > 1$, and for every $u_0 \in PC^{N_0}(I\!R^m) \cap I\!B_\delta(\bar{x})$ any sequence x_k produced by the discretized Newton process (4) and contained in $I\!B_\delta(\bar{x})$ converges linearly to $\bar{x}$.*

Proof. We will apply Theorem 6. Let $\mu > 0$ and $\varepsilon > 0$ be chosen so small that (15) is fulfilled. According to the considerations in the end of Section 3 the Newton mapping A satisfies (22) and the first inequality in (23) with a sufficiently small neighborhood V. Let ρ, δ and θ be as in Theorem 6 in its version for the case of strong metric regularity (so that the last statement of the theorem holds true).

Let $x_{k+1} \in X^{N_{k+1}}$ be the $k+1$-st iteration of the discretized Newton process (4), $k \geq 0$. Let r_k be the residual that x_{k+1} gives when plugged into the exact Newton inclusion $A(\cdot, x_k) + F(x) \ni 0$, that is, $r_k + A(x_{k+1}, x_k) + F(x_{k+1}) \ni 0$. In order to apply Theorem 6 we have to estimate this residual r_k in the space $Y = L^\infty(I\!R^n) \times L^\infty(I\!R^m) \times L^\infty(I\!R^n)$. Since p_{k+1} and p_{k+1} are linear and u_{k+1} is constant on each subinterval $[t^i_{k+1}, t^{i+1}_{k+1})$, this amounts to estimating the expression

$$\begin{aligned} &\nabla_q H(x_k(t)) - \nabla_q H(x_k(t^i_{k+1})) \\ &\quad + \nabla^2_{qx} H(x_k(t))(x_{k+1}(t) - x_k(t)) - \nabla^2_{qx} H(x_k(t^i_{k+1}))(x_{k+1}(t^i_{k+1}) - x_k(t^i_{k+1})) \end{aligned}$$

and also the similar expressions arising from the second and the third equations in the Newton method. The iteration x_k is either the initial one ($k = 0$) in which case p_k and q_k satisfy the Euler discretization in (36), or they satisfy the first and the second equations in (4). The function u_k, being in the ball with radius δ around $\bar{u}$ in $L^\infty(I\!R^m)$, is bounded (uniformly in k). Thus, for an appropriate constant C_1 in both cases $|p^{i+1}_k - p^i_k| \leq C_1 h_k$. Hence, $|p_k(t) - p_k(t^i_{k+1})| \leq C_1 h_{k+1}$ for $t \in [t^i_{k+1}, t^{i+1}_{k+1})$. The same applies also for q. For u we have $u_k(t) - u_k(t^i_{k+1}) = 0$ due to the condition that consequent meshes are embedded. The same argument applies also to $x_{k+1}(t)) - x_k(t^i_{k+1})$. Hence, $|r_k| \leq C_2 h_{k+1}$ for an appropriate constant C_2. By choosing $\bar{N}$ sufficiently large we may ensure that $|r_k| \leq \theta$, thus Theorem 6 can be applied with the constant function r_k. We obtain that x_{k+1}, that is claimed to exists in Theorem 6, coincides with x_{k+1} obtained by the discretized Newton process, while the first claim of the same theorem implies that

$$\|x_{k+1} - \bar{x}\| \leq \rho \|x_k - \bar{x}\| + C_3 h_{k+1} \leq \rho \|x_k - \bar{x}\| + \frac{C_3}{N_0} \left(\frac{1}{\nu}\right)^k.$$

The rest of the proof only need to repeat the argument in the discussion after the proof of Theorem 6. □

In the above theorem we assume that an initial control $u_0 \in PC^{N_0}(\mathbb{R}^m) \cap \mathbb{B}_\delta(\bar{x})$ exists, which is always true if the optimal control $\bar{u}$ is integrable in Riemann sense, provided that N_0 is chosen sufficiently large.

A result related to Theorem 9 is proved in [3, Section 5], where however, Lipschitz continuity of the optimal control is a priori assumed and the strong metric regularity of the optimality system is ensured by a coercivity condition. We mention again that (local) coercivity (together with the rest of the assumptions in [3, Section 5]) is a sufficient condition, but not necessary, for strong metric regularity.

Yet another open question, an attempt for solving which was the starting point of this paper, is as follows. In [5] is was proved that for the mapping associated with a variational inequality over a convex polyhedral set, in finite dimensions, metric regularity implies strong metric regularity. Now consider the optimality system (33), which is a variational inequality, and assume that the set U is a convex polyhedron. Then, according to [5], the metric regularity will automatically implies strong metric regularity.

If we know that, for a sufficiently small discretization step the (strong) metric regularity of the discretized system (36) is equivalent to the (strong) metric regularity of the original system (33), then we would obtain that for variational system of the original optimal control problem (32) metric regularity is equivalent to strong metric regularity. We were not able to prove this statement (by now!) but we are inclined to think that it is true.

References

[1] A. L. Dontchev, An a priori estimate for discrete approximations in nonlinear optimal control, SIAM J. Control Optim. **34** (1996), 1315–1328.

[2] A. L. Dontchev, W. W. Hager, An inverse mapping theorem for set-valued maps, Proc. Amer. Math. Soc. **121** (1994) 481–489.

[3] A. L. Dontchev, W. W. Hager, V. M. Veliov, Uniform convergence and mesh independence of Newton's method for discretized variational problems, SIAM J. Control Optim. **39** (2000) 961–980.

[4] A. L. Dontchev, K. Malanowski, A characterization of Lipschitzian stability in optimal control. Calculus of variations and optimal control (Haifa, 1998), 62–76, Chapman & Hall/CRC Res. Notes Math. **411** Chapman & Hall/CRC, Boca Raton, FL, 2000.

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

[6] A. L. Dontchev, R. T. Rockafellar, Implcit Functions and Solution Mappings, Springer Mathematics Monographs, to appear in June 2009.

[7] U. Felgenhauer, The shooting approach in analyzing bang-bang extremals with simultaneous control switches. Control & Cybernetics, **37**(2) (2008) 307 - 327.

[8] U. Felgenhauer, L. Poggiolini, G. Stefani, Optimality and stability result for bang-bang optimal controls with simple and double switch behavior. This journal ???.

[9] A. D. Ioffe, Metric regularity and subdifferential calculus. Uspekhi Mat. Nauk **55** (2000), no. 3 (333),103–162; English translation Math. Surveys **55** (2000), 501–558.

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

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

[12] S. M. Robinson, Newton's method for a class of nonsmooth functions. Set-Valued Anal. **2** (1994), 291–305.

[13] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis.* Springer, Berlin, 1998.