



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

Operations
Research and
Control Systems



Metric Regularity of Newtons' Iteration

*F.J. Aragón, A.L. Dontchev,
M. Gaydu, M.H. Geoffroy and V.M. Veliov*

Research Report 2010-15

December, 2010

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

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

METRIC REGULARITY OF NEWTON'S ITERATION

F. J. Aragón Artacho¹, A. L. Dontchev²,

M. Gaydu, M. H. Geoffroy³ and V. M. Veliov⁴

Abstract. For a version of Newton's method applied to a generalized equation with a parameter, we extend the paradigm of the Lyusternik-Graves theorem to the framework of a mapping acting from the pair "parameter-starting point" to the set of corresponding convergent Newton's sequences. Under ample parameterization, metric regularity of the mapping associated with convergent Newton's sequences becomes equivalent to the metric regularity of the mapping associated with the generalized equation. We also discuss an inexact Newton's method and present an application to discretized optimal control.

Key Words. Lyusternik-Graves theorem, Newton's method, generalized equation, metric regularity, perturbations, optimal control.

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

¹Department of Statistics and Operations Research, University of Alicante, 03080 Alicante, Spain, francisco.aragon@ua.es. Supported by MICINN of Spain, grant MTM2008-06695-C03-01 and programs "José Castillejo" and "Juan de la Cierva."

²Mathematical Reviews, 416 Fourth Street, Ann Arbor, MI 48103. On leave from the Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria. ald@ams.org. Supported by the National Science Foundation Grant DMS 1008341.

³LAMIA, Dept. de Mathématiques, Université Antilles-Guyane, F-97159 Pointe-à-Pitre, Guadeloupe, michael.gaydu@univ-ag.fr, michel.geoffroy@univ-ag.fr

⁴Institute of Mathematical Methods in Economics, Vienna University of Technology, A-1040 Vienna, Austria, e-mail: veliov@tuwien.ac.at.

1 Introduction

According to the Banach open mapping principle, the surjectivity of a linear and bounded mapping A , acting from a Banach space X to another Banach space Y , not only means that for any $y \in Y$ there exists x with $Ax = y$ but is actually equivalent to the existence of a positive constant κ such that for any $(x, y) \in X \times Y$, the distance from x to the set of solutions x' of $Ax' = y$ is bounded by κ times the “residual” $\|y - Ax\|$. If $d(x, C)$ denotes the distance from a point x to a set C , the latter condition can be written as

$$(1) \quad d(x, A^{-1}(y)) \leq \kappa \|y - Ax\| \quad \text{for all } (x, y) \in X \times Y,$$

and describes a property of A known as *metric regularity*. Graves [8] extended the Banach open mapping principle to continuous functions that have surjective “approximate derivatives.” Specifically, let $f : X \rightarrow Y$ be a function that is continuous in a neighborhood of a point \bar{x} , let $A : X \rightarrow Y$ be a linear continuous mapping which is surjective and let κ be the constant in the Banach open mapping theorem associated with A . Let the difference $f - A$ be Lipschitz continuous in a neighborhood of \bar{x} with Lipschitz constant μ such that $\kappa\mu < 1$. Then a slight extension in the original proof of Graves, see e.g. [7], p. 276, gives us the same property as in (1) but now localized around the reference point:

$$d(x, f^{-1}(y)) \leq \frac{\kappa}{1 - \kappa\mu} \|y - f(x)\| \quad \text{for all } (x, y) \text{ near } (\bar{x}, f(\bar{x})).$$

Much earlier than Graves, Lyusternik [13] obtained the form of the tangent manifold to the kernel of a function, which was a stepping stone for A. A. Milyutin and his disciples [1] to develop far reaching extensions of the theorems of Lyusternik and Graves. One should note that both Lyusternik and Graves, as well as Milyutin et al., used in their proofs iterative schemes that resemble the Picard iteration or, even more directly, the Newton method. Modern versions of the theorems of Lyusternik and Graves are commonly called Lyusternik-Graves theorems. In Section 2 of this paper we present a Lyusternik-Graves theorem (Theorem 2.1) for a set-valued mapping perturbed by a function with a “sufficiently small” Lipschitz constant. Various theorems of this kind, as well as other recent developments centered around regularity properties of set-valued mappings and their role in optimization and beyond can be found in the papers [2], [9], [11], and in particular in the recent book [7].

In Section 3 we introduce Newton’s method and discuss its convergence under metric regularity, in particular establishing estimates for the convergence parameters.

The main result of this paper, presented in Section 4 as Theorem 4.2, is a Lyusternik-Graves type theorem about sequences generated by the Newton method applied to a generalized equation. Instead of considering a mapping and a perturbation of it, we study a generalized equation and its solution mapping, and also two related “approximations”: one being a linearization of the mapping associated with the equations and the other acting from

the pair “parameter – starting point” to the set of all convergent Newton’s sequences starting from that point and associated with that parameter. In our main result we show that these two mappings obey the general paradigm of the Lyusternik-Graves theorem; namely, if the linearized equation mapping is metrically regular, then the mapping associated with Newton’s sequences has the Aubin property. Under an additional condition of the so-called *ample parameterization*, the converse implication holds as well.

For illustration of our main result, consider solving a system of inequalities and equalities describing the *feasibility problem*:

$$(2) \quad \begin{aligned} g(x) &\leq p, \\ h(x) &= q, \end{aligned}$$

where $p \in \mathbb{R}^m$ and $q \in \mathbb{R}^k$ are parameters and $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $h : \mathbb{R}^n \rightarrow \mathbb{R}^k$ are continuously differentiable functions. System (2) can be put in the form of the generalized equation

$$(3) \quad y \in f(x) + F, \quad \text{where } y = \begin{pmatrix} p \\ q \end{pmatrix}, \quad f = \begin{pmatrix} g \\ h \end{pmatrix} \quad \text{and } F = \begin{pmatrix} \mathbb{R}_+^m \\ 0 \end{pmatrix}.$$

It is well known that metric regularity of the mapping $f + F$ at, say, \bar{x} for 0 is equivalent to the standard Mangasarian-Fromovitz condition at \bar{x} , see e.g. [7], Example 4D.3. Now, let us apply to (3) the Newton method described in Section 3 of this paper, namely

$$y \in f(x_k) + Df(x_k)(x_{k+1} - x_k) + F,$$

which consists of solving at each iteration a system of affine inequalities and equalities. The main result in this paper given in Theorem 4.2 and, in particular, Corollary 4.3 yields that the Mangasarian-Fromovitz condition for (2) at \bar{x} is equivalent to the following property of the set of sequences generated by the Newton method: when y is close to zero and the starting point x_0 is close to the solution \bar{x} , then the set of convergent sequences is nonempty; moreover, for every y, y' close to 0 and for every convergent sequence ξ for y there exists a convergent sequence ξ' for y' such that the l_∞ distance between ξ and ξ' is bounded by a constant times $\|y - y'\|$. This property of a set-valued mapping is known as Aubin continuity, which is a local version of the usual Lipschitz continuity with respect to the Pompeiu-Hausdorff distance. Thus, the Mangasarian-Fromovitz condition gives us not only convergence but also a kind of quantitative stability of the set of Newton sequences with respect to perturbations, and that’s all we can get from this condition. In Section 5 we present much more elaborate applications of our result to an inexact version of Newton’s method and to discretized optimal control.

This paper extends to a much broader framework the previous paper [6], see also [7], Section 6C, where Newton’s iteration for a generalized equation is considered under *strong metric regularity* and corresponding “sequential implicit function theorems” are established.

Recall that a mapping is strongly regular, a concept coined by S. M. Robinson [14], when it is metrically regular and its inverse has a Lipschitz continuous single-valued localization around the reference point. Under strong metric regularity, for each starting point close to a solution there is a unique Newton's sequence, which is then automatically convergent. This is not the case when the mapping at hand is merely metrically regular, where we have to deal with a set of sequences. In particular, the result in [6] cannot be applied to the feasibility problem (2). In order to deal with the sets of sequences, in the proof of our main result we use a technique based on "gluing" sequences each to other to construct a Newton sequence which is not only at the desired distance from the assumed one but is also convergent. At the end of the paper we derive from our Theorem 4.2 a stronger version of the main result in [6].

In the rest of this introductory section we first fix the notation and terminology. In what follows P , X and Y are Banach spaces. The notation $f : X \rightarrow Y$ means that f is a function while $F : X \rightrightarrows Y$ is a general mapping where the double arrow indicates that F may be set-valued. The *graph* of F is the set $\text{gph } F = \{(x, y) \in X \times Y \mid y \in F(x)\}$, and the *inverse* of F is the mapping $F^{-1} : Y \rightrightarrows X$ defined by $F^{-1}(y) = \{x \mid y \in F(x)\}$. All norms are denoted by $\|\cdot\|$. The closed ball centered at x with radius r is denoted by $\mathbb{B}_r(x)$, and the closed unit ball is \mathbb{B} . The *distance* from a point x to a set C is defined as $d(x, C) = \inf_{y \in C} d(x, y)$, while the *excess* from a set A to a set B is the quantity $e(A, B) = \sup_{x \in A} d(x, B)$.

Definition 1.1 (metric regularity). *A mapping $F : X \rightrightarrows Y$ is said to be metrically regular at \bar{x} for \bar{y} when $\bar{y} \in F(\bar{x})$ and there is a constant $\kappa \geq 0$ together with neighborhoods U of \bar{x} and V of \bar{y} such that*

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

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 $\text{reg}(F; \bar{x} | \bar{y})$. The absence of metric regularity is signaled by $\text{reg}(F; \bar{x} | \bar{y}) = \infty$. Thus, when we write that the modulus of metric regularity of a mapping is finite, e.g., less than a constant, we mean that the associate mapping is metrically regular.

When $A : X \rightarrow Y$ is a linear and bounded mapping, one has that A is metrically regular at every point $x \in X$ if and only if A is surjective; this is the Banach open mapping theorem. In this case the regularity modulus at any point is equal to the *inner norm* of the inverse of A , that is, $\text{reg } A = \|A^{-1}\|^- = \sup_{y \in B} d(0, A^{-1}(y))$.

Metric regularity of a mapping can be characterized by other properties; here we need the equivalence of metric regularity of a mapping with the so-called *Aubin property* of the inverse of this mapping.

Definition 1.2 (Aubin property). *A mapping $H : Y \rightrightarrows X$ is said to have the Aubin property at \bar{y} for \bar{x} if $\bar{x} \in H(\bar{y})$ and there exist a nonnegative constant κ together with neighborhoods U of \bar{x} and V of \bar{y} such that*

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

The infimum of κ over all such combinations of κ , U and V is called the Lipschitz modulus of H at \bar{y} for \bar{x} and denoted by $\text{lip}(H; \bar{y} | \bar{x})$. The absence of this property is signaled by $\text{lip}(H; \bar{y} | \bar{x}) = \infty$.

Additionally, a mapping $H : P \times Y \rightrightarrows X$ is said to have the partial Aubin property with respect to p uniformly in y at (\bar{p}, \bar{y}) for \bar{x} if $\bar{x} \in H(\bar{p}, \bar{y})$ and there is a nonnegative constant κ together with neighborhoods Q for \bar{p} , U of \bar{x} and V of \bar{y} such that

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

The infimum of κ over all such combinations of κ , Q , U and V is called the partial Lipschitz modulus of H with respect to p uniformly in y at (\bar{p}, \bar{y}) for \bar{x} and denoted by $\widehat{\text{lip}}_p(H; (\bar{p}, \bar{y}) | \bar{x})$. The absence of this property is signaled by $\widehat{\text{lip}}_p(H; (\bar{p}, \bar{y}) | \bar{x}) = \infty$.

It is now well known, see e.g. [7], Section 3E, that a mapping $F : X \rightrightarrows Y$ is metrically regular at \bar{x} for \bar{y} if and only if its inverse F^{-1} has the Aubin property at \bar{y} for \bar{x} , and moreover, $\text{lip}(F^{-1}; \bar{y} | \bar{x}) = \text{reg}(F; \bar{x} | \bar{y})$.

We recall next quantitative measures for Lipschitz continuity and partial Lipschitz continuity in a neighborhood, both of which will play an essential role in the paper. A function $f : X \rightarrow Y$ is said to be Lipschitz continuous relative to a set D , or on a set D , if $D \subset \text{dom } f$ and there exists a constant $\kappa \geq 0$ (a Lipschitz constant) such that

$$(4) \quad \|f(x') - f(x)\| \leq \kappa \|x' - x\| \quad \text{for all } x', x \in D.$$

It is said to be Lipschitz continuous around \bar{x} when this holds for some neighborhood D of \bar{x} . The Lipschitz modulus of f at \bar{x} , denoted $\text{lip}(f; \bar{x})$, is the infimum of the set of values of κ for which there exists a neighborhood D of \bar{x} such that (4) holds. Equivalently,

$$\text{lip}(f; \bar{x}) := \limsup_{\substack{x', x \rightarrow \bar{x}, \\ x \neq x'}} \frac{\|f(x') - f(x)\|}{\|x' - x\|}.$$

Further, a function $f : P \times X \rightarrow Y$ is said to be Lipschitz continuous with respect to x uniformly in p around $(\bar{p}, \bar{x}) \in \text{int dom } f$ when there are neighborhoods Q of \bar{p} and U of \bar{x} along with a constant κ and such that

$$\|f(p, x) - f(p, x')\| \leq \kappa \|x - x'\| \quad \text{for all } x, x' \in U \text{ and } p \in Q.$$

Accordingly, the partial uniform Lipschitz modulus has the form

$$\widehat{\text{lip}}_x(f; (\bar{p}, \bar{x})) := \limsup_{\substack{x, x' \rightarrow \bar{x}, p \rightarrow \bar{p}, \\ x \neq x'}} \frac{\|f(p, x') - f(p, x)\|}{\|x' - x\|}.$$

The definitions of metric regularity and the Lipschitz moduli can be extended in an obvious way to mappings acting in metric spaces.

2 Parametric Lyusternik-Graves Theorems

Our first result is a Lyusternik-Graves theorem involving a general set-valued mapping perturbed by a function which in turn depends on a parameter. It generalizes Theorem 5E.1 in [7], p. 280, in that the function now depends on a parameter and shows more transparently the interplay among constants and neighborhoods.

Theorem 2.1 (parametric Lyusternik-Graves). *Consider a mapping $F : X \rightrightarrows Y$ and any $(\bar{x}, \bar{y}) \in \text{gph } F$ at which $\text{gph } F$ is locally closed (which means that the intersection of $\text{gph } F$ with a closed ball around (\bar{x}, \bar{y}) is closed). Consider also a function $g : P \times X \rightarrow Y$ and suppose that there exist nonnegative constants κ and μ such that*

$$(5) \quad \text{reg}(F; \bar{x} | \bar{y}) \leq \kappa, \quad \widehat{\text{lip}}_x(g; (\bar{q}, \bar{x})) \leq \mu \quad \text{and} \quad \kappa\mu < 1.$$

Then for every $\kappa' > \kappa/(1 - \kappa\mu)$ 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 $g(q, \cdot) + F(\cdot)$ is metrically regular in x 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 $g(q, \bar{x}) + \bar{y}$.

An important element in this statement is that the regularity constant and neighborhoods of metric regularity of the perturbed mapping $F + g$ depend only on the regularity modulus of the underlying mapping F and the Lipschitz modulus of the perturbation function g , but not on the value of the parameter q in a neighborhood of the reference point \bar{q} . We will utilize this observation in the proof of our main result stated in Theorem 4.2. This result can be stated and proved with minor changes in notation only for the case when P is a metric space, X is a complete metric space and Y is a linear space equipped with a shift-invariant metric, as in Theorem 5E.1 in [7].

The proof of Theorem 2.1 we present below uses the contracting mapping theorem given next, similar to the second proof of Theorem 5E.1 in [7], but it is different from it in that it deals directly with metric regularity and in such a way it keeps track of the dependence of the constants and the neighborhoods.

Theorem 2.2 ([4], a contraction mapping principle for set-valued mappings). *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 a and θ be such that $\theta < 1$, the set $\text{gph } \Phi \cap (\mathbb{B}_a(\bar{x}) \times \mathbb{B}_a(\bar{x}))$ is closed and the following conditions hold:*

- (i) $d(\bar{x}, \Phi(\bar{x})) < a(1 - \theta)$;
- (ii) $e(\Phi(u) \cap \mathbb{B}_a(\bar{x}), \Phi(v)) \leq \theta\rho(u, v)$ for all $u, v \in \mathbb{B}_a(\bar{x})$.

Then there exists $x \in \mathbb{B}_a(\bar{x})$ such that $x \in \Phi(x)$.

Proof of Theorem 2.1. Pick the constants κ and μ as in (5) and then any $\kappa' > \kappa/(1 - \kappa\mu)$. Let $\lambda > \kappa$ and $\nu > \mu$ be such that $\lambda\nu < 1$ and $\lambda/(1 - \lambda\nu) < \kappa'$. Then there exist positive constants a and b such that

$$(6) \quad d(x, F^{-1}(y)) \leq \lambda d(y, F(x)) \quad \text{for all } (x, y) \in \mathbb{B}_a(\bar{x}) \times \mathbb{B}_b(\bar{y}).$$

Adjust a and b if necessary so that

$$(7) \quad \text{the set } \text{gph } F \cap (\mathbb{B}_a(\bar{x}) \times \mathbb{B}_b(\bar{y})) \text{ is closed.}$$

Then choose $c > 0$ and make a smaller if necessary such that

$$(8) \quad \|g(q, x') - g(q, x)\| \leq \nu \|x - x'\| \text{ for all } x, x' \in \mathbb{B}_a(\bar{x}) \text{ and } q \in \mathbb{B}_c(\bar{q}).$$

Choose positive constants α and β such that

$$(9) \quad \alpha + 5\kappa'\beta \leq a, \quad \alpha \leq 2\kappa'\beta, \quad \nu\alpha + 4\beta \leq b \quad \text{and} \quad \nu(\alpha + 5\kappa'\beta) + \beta \leq b.$$

Pick $q \in \mathbb{B}_c(\bar{q})$ and then $x \in \mathbb{B}_\alpha(\bar{x})$ and $y \in \mathbb{B}_\beta(g(q, \bar{x}) + \bar{y})$. We will first prove that for every $y' \in (g(q, x) + F(x)) \cap \mathbb{B}_{4\beta}(g(q, \bar{x}) + \bar{y})$

$$(10) \quad d(x, (g(q, \cdot) + F(\cdot))^{-1}(y)) \leq \kappa' \|y - y'\|.$$

Choose any $y' \in (g(q, x) + F(x)) \cap \mathbb{B}_{4\beta}(g(q, \bar{x}) + \bar{y})$. If $y = y'$ then $x \in (g(q, \cdot) + F(\cdot))^{-1}(y)$ and (10) holds since both the left side and the right side are zero. Suppose $y' \neq y$ and consider the mapping

$$\Phi : x \mapsto F^{-1}(-g(q, x) + y) \text{ for } x \in \mathbb{B}_\alpha(\bar{x}).$$

We will now prove that the mapping Φ has a fixed point in the set $\mathbb{B}_r(x)$ centered at x with radius $r := \kappa' \|y - y'\|$. Using (8) and (9), we have

$$\| -g(q, x) + y' - \bar{y} \| \leq \| -g(q, x) + g(q, \bar{x}) \| + \| y' - \bar{y} - g(q, \bar{x}) \| \leq \nu\alpha + 4\beta \leq b.$$

The same estimate holds of course with y' replaced by y because y was chosen in $\mathbb{B}_\beta(g(q, \bar{x}) + \bar{y})$. Hence, both $-g(q, x) + y'$ and $-g(q, x) + y$ are in $\mathbb{B}_b(\bar{y})$. We will now show that the set $\text{gph } \Phi \cap (\mathbb{B}_r(x) \times \mathbb{B}_r(x))$ is closed.

Let $(x_n, z_n) \in \text{gph } \Phi \cap (\mathbb{B}_r(x) \times \mathbb{B}_r(x))$ and $(x_n, z_n) \rightarrow (\tilde{x}, \tilde{z})$. Then $(z_n, -g(q, x_n) + y) \in \text{gph } F$ and also, from (9),

$$\|z_n - \bar{x}\| \leq \|z_n - x\| + \|x - \bar{x}\| \leq r + \alpha = \kappa' \|y - y'\| + \alpha \leq 5\kappa'\beta + \alpha \leq a$$

and

$$\begin{aligned} \| -g(q, x_n) + y - \bar{y} \| &\leq \| -g(q, x_n) + g(q, \bar{x}) \| + \| y - \bar{y} - g(q, \bar{x}) \| \\ &\leq \nu \|x_n - \bar{x}\| + \beta \leq \nu(\|x_n - x\| + \|x - \bar{x}\|) + \beta \\ &\leq \nu(r + \alpha) + \beta \leq \nu(5\kappa'\beta + \alpha) + \beta \leq b. \end{aligned}$$

Thus $(z_n, -g(q, x_n) + y) \in \text{gph } F \cap (\mathbb{B}_a(\bar{x}) \times \mathbb{B}_b(\bar{y}))$ which is closed by (7). Note that $r \leq \kappa'(4\beta + \beta)$ and hence, from the first relation in (9), $\mathbb{B}_r(x) \subset \mathbb{B}_a(\bar{x})$. Since $g(q, \cdot)$ is

continuous in $\mathcal{B}_a(\bar{x})$ (even Lipschitz, from (8)) and $x_n \in \mathcal{B}_r(x) \subset \mathcal{B}_a(\bar{x})$, we get that $(\tilde{z}, -g(q, \tilde{x}) + y) \in \text{gph } F \cap (\mathcal{B}_r(x) \times \mathcal{B}_b(\bar{y}))$, which in turn yields $(\tilde{x}, \tilde{z}) \in \text{gph } \Phi \cap (\mathcal{B}_r(x) \times \mathcal{B}_r(x))$. Hence the set $\text{gph } \Phi \cap (\mathcal{B}_r(x) \times \mathcal{B}_r(x))$ is closed.

Since $x \in (g(q, \cdot) + F(\cdot))^{-1}(y') \cap \mathcal{B}_a(\bar{x})$, utilizing the metric regularity of F we obtain

$$\begin{aligned} d(x, \Phi(x)) &= d(x, F^{-1}(-g(q, x) + y)) \leq \lambda d(-g(q, x) + y, F(x)) \\ &\leq \lambda \| -g(q, x) + y - (y' - g(q, x)) \| = \lambda \| y - y' \| \\ &< \kappa' \| y - y' \| (1 - \lambda\nu) = r(1 - \lambda\nu). \end{aligned}$$

Then (6), combined with (8) and the observation above that $\mathcal{B}_r(x) \subset \mathcal{B}_a(\bar{x})$, implies that for any $u, v \in \mathcal{B}_r(x)$,

$$\begin{aligned} e(\Phi(u) \cap \mathcal{B}_r(x), \Phi(v)) &\leq \sup_{z \in F^{-1}(-g(q, u) + y) \cap \mathcal{B}_a(\bar{x})} d(z, F^{-1}(-g(q, v) + y)) \\ &\leq \sup_{z \in F^{-1}(-g(q, u) + y) \cap \mathcal{B}_a(\bar{x})} \lambda d(-g(q, v) + y, F(z)) \\ &\leq \lambda \| -g(q, u) + g(q, v) \| \leq \lambda\nu \| u - v \|. \end{aligned}$$

Theorem 2.2 then yields the existence of a point $\hat{x} \in \Phi(\hat{x}) \cap \mathcal{B}_r(x)$; that is,

$$y \in g(q, \hat{x}) + F(\hat{x}) \quad \text{and} \quad \|\hat{x} - x\| \leq \kappa' \| y - y' \|.$$

Thus, since $\hat{x} \in (g(q, \cdot) + F(\cdot))^{-1}(y)$ we obtain (10).

Now we will prove the inequality

$$(11) \quad d(x, (g(q, \cdot) + F(\cdot))^{-1}(y)) \leq \kappa' d(y, g(q, x) + F(x))$$

which gives us the desired property of $g + F$. First, note that if $g(q, x) + F(x) = \emptyset$ the right side of (11) is $+\infty$ and we are done. Let $\varepsilon > 0$ and $w \in g(q, x) + F(x)$ be such that

$$\|w - y\| \leq d(y, g(q, x) + F(x)) + \varepsilon.$$

If $w \in \mathcal{B}_{4\beta}(g(q, \bar{x}) + \bar{y})$ then from (10) we have that

$$(12) \quad d(x, (g(q, \cdot) + F(\cdot))^{-1}(y)) \leq \kappa' \| y - w \| \leq \kappa' (d(y, g(q, x) + F(x)) + \varepsilon)$$

and since the left side of this inequality does not depend on ε , we obtain the desired inequality (11). If $w \notin \mathcal{B}_{4\beta}(g(q, \bar{x}) + \bar{y})$ then

$$\|w - y\| \geq \|w - g(q, \bar{x}) - \bar{y}\| - \|y - g(q, \bar{x}) - \bar{y}\| \geq 3\beta.$$

On the other hand, from (10) and then (9),

$$e(\mathcal{B}_\alpha(\bar{x}), (g(q, \cdot) + F(\cdot))^{-1}(y)) \leq \alpha + d(\bar{x}, (g(q, \cdot) + F(\cdot))^{-1}(y)) \leq \alpha + \kappa' \|\bar{y} + g(q, \bar{x}) - y\| \leq 3\kappa'\beta.$$

Since $x \in \mathcal{B}_\alpha(\bar{x})$, we obtain

$$\begin{aligned} d(x, (g(q, \cdot) + F(\cdot))^{-1}(y)) &\leq e(\mathcal{B}_\alpha(\bar{x}), (g(q, \cdot) + F(\cdot))^{-1}(y)) \\ &\leq 3\kappa'\beta \leq \kappa'\|w - y\| \leq \kappa'(d(y, g(q, x) + F(x)) + \varepsilon). \end{aligned}$$

This again implies (11) and we are done. \square

The theorem we state next concerns generalized equations of the form

$$(13) \quad f(p, x) + F(x) \ni 0,$$

for a function $f : P \times X \rightarrow Y$ and a mapping $F : X \rightrightarrows Y$, where we solve (13) with respect to the variable x for a given value of p which plays the role of a parameter. The *solution mapping* associated with the generalized equation (13) is the potentially set-valued mapping $S : P \rightrightarrows X$ defined by

$$(14) \quad S : p \mapsto \{ x \mid f(p, x) + F(x) \ni 0 \}.$$

The following result is given in [7, Theorem 3F.9] in finite dimensions but with a proof whose extension to Banach spaces needs only minor adjustments in notation, see also [7, Theorem 5E.4]. Recall that a function $f : X \rightarrow Y$ is said to be strictly differentiable at \bar{x} when there exists a linear continuous mapping $Df(\bar{x})$, the strict derivative of f at \bar{x} , such that $\text{lip}(f - Df(\bar{x}); \bar{x}) = 0$.

Theorem 2.3 (implicit mapping theorem with metric regularity). *Consider the generalized equation (13) with solution mapping S in (14) and a point (\bar{p}, \bar{x}) with $\bar{x} \in S(\bar{p})$. Suppose that f is strictly differentiable at (\bar{p}, \bar{x}) with strict partial derivatives denoted by $D_x f(\bar{p}, \bar{x})$ and $D_p f(\bar{p}, \bar{x})$ and that $\text{gph } F$ is locally closed at $(\bar{x}, -f(\bar{p}, \bar{x}))$. If the mapping*

$$(15) \quad x \mapsto G(x) := f(\bar{p}, \bar{x}) + D_x f(\bar{p}, \bar{x})(x - \bar{x}) + F(x)$$

is metrically regular at \bar{x} for 0, then S has the Aubin property at \bar{p} for \bar{x} with

$$\text{lip}(S; \bar{p} | \bar{x}) \leq \text{reg}(G; \bar{x} | 0) \cdot \|D_p f(\bar{p}, \bar{x})\|.$$

Furthermore, when f satisfies the ample parameterization condition:

$$(16) \quad \text{the mapping } D_p f(\bar{p}, \bar{x}) \text{ is surjective,}$$

then the converse implication holds as well: the mapping G is metrically regular at \bar{x} for 0 provided that S has the Aubin property at \bar{p} for \bar{x} , having

$$\text{reg}(G; \bar{x} | 0) \leq \text{lip}(S; \bar{p} | \bar{x}) \cdot \|D_p f(\bar{p}, \bar{x})^{-1}\|^{-1}.$$

The above results yield the following important corollary which is closer to the original formulations of the theorems of Lyusternik and Graves.

Corollary 2.4 (Lyusternik-Graves for linearization). *Consider the mapping $f + F$ and a point (\bar{y}, \bar{x}) with $\bar{y} \in f(\bar{x}) + F(\bar{x})$ and suppose that f is strictly differentiable at \bar{x} and that $\text{gph } F$ is locally closed at $(\bar{x}, \bar{y} - f(\bar{x}))$. Then the mapping $f + F$ is metrically regular at \bar{x} for \bar{y} if and only if the linearized mapping $x \mapsto G(x) := f(\bar{x}) + Df(\bar{x})(x - \bar{x}) + F(x)$ is metrically regular at \bar{x} for \bar{y} .*

Proof. It is enough to observe that in the case the ample parameterization condition (16) holds automatically and that $f + F$ and G can exchange places. \square

3 Newton's method under metric regularity

In this and the following sections we consider the generalized equation (13) on the following standing assumptions:

Standing assumptions: *For a given reference value \bar{p} of the parameter the generalized equation (13) has a solution \bar{x} . The function f is continuously differentiable in a neighborhood of (\bar{p}, \bar{x}) with strict partial derivatives denoted by $D_x f(\bar{p}, \bar{x})$ and $D_p f(\bar{p}, \bar{x})$ such that*

$$(17) \quad \text{lip}(D_x f; (\bar{p}, \bar{x})) < \infty,$$

and the mapping F has closed graph.

The standing assumptions are used in full strength in the main result established in Theorem 4.2 but some of them are not necessary in the preliminary results; to simplify the expositions we put aside these technical nuances.

We study the following version of Newton's method for solving (13):

$$(18) \quad f(p, x_k) + D_x f(p, x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \quad \text{for } k = 0, 1, \dots,$$

with a given starting point x_0 . If F is the zero mapping, (18) is the standard Newton's method for solving the equation $f(p, x) = 0$ with respect to x . 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 (18) becomes the popular sequential quadratic programming method.

In our further analysis we employ the following corollary of Theorem 2.1:

Corollary 3.1. *Consider the parameterized form of the mapping G given by*

$$(19) \quad X \ni x \mapsto G_{p,u}(x) = f(p, u) + D_x f(p, u)(x - u) + F(x), \quad \text{for } p \in P, u \in X.$$

and suppose that the mapping G defined in (15) is metrically regular at \bar{x} for 0. Then for every $\lambda > \text{reg}(G; \bar{x}|0)$ there exist positive numbers a , b and c such that

$$d(x, G_{p,u}^{-1}(y)) \leq \lambda d(y, G_{p,u}(x)) \quad \text{for every } u, x \in \mathcal{B}_a(\bar{x}), y \in \mathcal{B}_b(0), p \in \mathcal{B}_c(\bar{p}).$$

Proof. We apply Theorem 2.1 with the following specifications: $F(x) = G(x)$, $\bar{y} = 0$, $q = (p, u)$, $\bar{q} = (\bar{p}, \bar{x})$, and

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

Let $\lambda > \kappa \geq \text{reg}(G; \bar{x}|0)$. Pick any $\mu > 0$ such that $\mu\kappa < 1$ and $\lambda > \kappa/(1 - \kappa\mu)$. The standing assumptions yield that there exist positive constants L , α and β such that

$$(20) \quad \|f(p, x) - f(p', x)\| \leq L\|p - p'\| \quad \text{for every } p, p' \in \mathcal{B}_\beta(\bar{p}), \quad x \in \mathcal{B}_\alpha(\bar{x}),$$

$$(21) \quad \|D_x f(p, x) - D_x f(p, x')\| \leq L\|x - x'\| \quad \text{for every } x, x' \in \mathcal{B}_\alpha(\bar{x}), \quad p \in \mathcal{B}_\beta(\bar{p}),$$

and

$$(22) \quad \|D_x f(p, u) - D_x f(\bar{p}, \bar{x})\| \leq \mu \quad \text{for every } p \in \mathcal{B}_\beta(\bar{p}), \quad u \in \mathcal{B}_\alpha(\bar{x}).$$

Observe that for any $x, x' \in X$ and any $q = (p, u) \in \mathcal{B}_\beta(\bar{p}) \times \mathcal{B}_\alpha(\bar{x})$, from (22),

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

that is $\widehat{\text{lip}}_x(g; (\bar{q}, \bar{x})) \leq \mu$. Thus, the assumptions of Theorem 2.1 are satisfied and hence there exist positive constants $a' \leq \alpha$, b' and $c' \leq \beta$ such that for any $q \in \mathcal{B}_{c'}(\bar{p}) \times \mathcal{B}_{a'}(\bar{x})$ the mapping $G_{p,u}(x) = g(q, x) + G(x)$ is metrically regular at \bar{x} for $g(q, \bar{x}) = f(p, u) + D_x f(p, u)(\bar{x} - u) - f(\bar{p}, \bar{x})$ with constant λ and neighborhoods $\mathcal{B}_{a'}(\bar{x})$ and $\mathcal{B}_{b'}(g(q, \bar{x}))$. Now choose positive scalars a , b and c such that

$$(23) \quad a \leq a', \quad c \leq c', \quad \text{and} \quad La^2/2 + Lc + b \leq b'.$$

Fix any $q = (p, u) \in \mathcal{B}_c(\bar{p}) \times \mathcal{B}_a(\bar{x})$. Using (21) in the standard estimation

$$(24) \quad \begin{aligned} & \|f(p, u) + D_x f(p, u)(\bar{x} - u) - f(p, \bar{x})\| \\ &= \left\| \int_0^1 D_x f(p, \bar{x} + t(u - \bar{x}))(u - \bar{x}) dt - D_x f(p, u)(u - \bar{x}) \right\| \\ &\leq L \int_0^1 (1 - t) dt \|u - \bar{x}\|^2 = \frac{L}{2} \|u - \bar{x}\|^2, \end{aligned}$$

and applying (20) and (23), we obtain that, for $y \in \mathcal{B}_b(0)$,

$$\begin{aligned} \|g(q, \bar{x}) - y\| &\leq \|f(p, u) + D_x f(p, u)(\bar{x} - u) - f(\bar{p}, \bar{x})\| + \|y\| \\ &\leq \|f(p, u) + D_x f(p, u)(\bar{x} - u) - f(p, \bar{x})\| + \|f(p, \bar{x}) - f(\bar{p}, \bar{x})\| + \|y\| \\ &\leq \frac{L}{2} \|u - \bar{x}\|^2 + L\|p - \bar{p}\| + b \leq La^2/2 + Lc + b \leq b'. \end{aligned}$$

Thus, $\mathcal{B}_b(0) \subset \mathcal{B}_{b'}(g(q, \bar{x}))$ and the proof is complete. \square

Theorem 3.2 (convergence under metric regularity). *Suppose that the mapping G defined in (15) is metrically regular at \bar{x} for 0. Then for every*

$$(25) \quad \gamma > \frac{1}{2} \operatorname{reg}(G; \bar{x}|0) \cdot \widehat{\operatorname{lip}}_x(D_x f; (\bar{p}, \bar{x}))$$

there are positive constants \bar{a} and \bar{c} such that, for every $p \in \mathcal{B}_{\bar{c}}(\bar{p})$, $u \in \mathcal{B}_{\bar{a}}(\bar{x})$ the set $S(p) \cap \mathcal{B}_{\bar{a}/2}(\bar{x})$ is nonempty and for every $s \in S(p) \cap \mathcal{B}_{\bar{a}/2}(\bar{x})$, there exists a Newton sequence satisfying (18) for p with starting point $x_0 = u$ and components x_1, \dots, x_k, \dots all belonging to $\mathcal{B}_{\bar{a}}(\bar{x})$ and which converges quadratically to s ; moreover,

$$(26) \quad \|x_{k+1} - s\| \leq \gamma \|x_k - s\|^2 \quad \text{for all } k = 0, 1, \dots$$

Proof. Choose γ as in (25) and let $\lambda > \operatorname{reg}(G; \bar{x}|0)$ and $L > \widehat{\operatorname{lip}}_x(D_x f; (\bar{p}, \bar{x}))$ be such that

$$(27) \quad \gamma > \frac{1}{2} \lambda L.$$

According to Corollary 3.1 there exist positive a and c such that

$$d(x, G_{p,u}^{-1}(0)) \leq \lambda d(0, G_{p,u}(x)) \quad \text{for every } u, x \in \mathcal{B}_a(\bar{x}), p \in \mathcal{B}_c(\bar{p}).$$

The Aubin property of the mapping S established in Theorem 2.3 implies that for any $d > \operatorname{lip}(S; \bar{p}|\bar{x})$ there exists $c' > 0$ such that $\bar{x} \in S(p) + d\|p - \bar{p}\|\mathcal{B}$ for any $p \in \mathcal{B}_{c'}(\bar{p})$. Then

$$S(p) \cap \mathcal{B}_{d\|p-\bar{p}\|}(\bar{x}) \neq \emptyset \quad \text{for } p \in \mathcal{B}_{c'}(\bar{p}).$$

We choose next positive constants \bar{a} and \bar{c} such that the following inequalities are satisfied:

$$(28) \quad \bar{a} < a, \quad \bar{c} < \min\left\{\frac{\bar{a}}{2d}, c, c'\right\} \quad \text{and} \quad \frac{9}{2}\gamma\bar{a} \leq 1.$$

Then for every $p \in \mathcal{B}_{\bar{c}}(\bar{p})$ the set $S(p) \cap \mathcal{B}_{\frac{\bar{a}}{2}}(\bar{x})$ is nonempty. Moreover, for every $s \in S(p) \cap \mathcal{B}_{\frac{\bar{a}}{2}}(\bar{x})$ and $u \in \mathcal{B}_{\bar{a}}(\bar{x})$ we have

$$(29) \quad d(s, G_{p,u}^{-1}(0)) \leq \lambda d(0, G_{p,u}(s)).$$

Fix arbitrary $p \in \mathcal{B}_{\bar{c}}(\bar{p})$, $s \in S(p) \cap \mathcal{B}_{\frac{\bar{a}}{2}}(\bar{x})$ and $u \in \mathcal{B}_{\bar{a}}(\bar{x})$. In the next lines we will show the existence of x_1 such that

$$(30) \quad G_{p,u}(x_1) \ni 0, \quad \|x_1 - s\| \leq \gamma \|u - s\|^2 \quad \text{and} \quad x_1 \in \mathcal{B}_{\bar{a}}(\bar{x}).$$

If $d(0, G_{p,u}(s)) = 0$ we set $x_1 = s$. Since F is closed-valued, (29) implies the first relation in (30), while the second one is obvious and the third one follows from $s \in \mathcal{B}_{\frac{\bar{a}}{2}}(\bar{x})$.

If $d(0, G_{p,u}(s)) > 0$, then from (27),

$$d(s, G_{p,u}^{-1}(0)) \leq \lambda d(0, G_{p,u}(s)) < \frac{2\gamma}{L} d(0, G_{p,u}(s)).$$

and hence there exists $x_1 \in G_{p,u}^{-1}(0)$ such that

$$(31) \quad \|s - x_1\| \leq \frac{2\gamma}{L} d(0, G_{p,u}(s)).$$

Since $f(p, s) + F(s) \ni 0$ we can estimate, as in (24),

$$d(0, G_{p,u}(s)) \leq \|f(p, u) + D_x f(p, u)(s - u) - f(p, s)\| \leq \frac{L}{2} \|u - s\|^2.$$

Then (31) implies the inequality in (30). To complete the proof of (30) we estimate

$$\|x_1 - \bar{x}\| \leq \|x_1 - s\| + \|s - \bar{x}\| \leq \gamma \|u - s\|^2 + \frac{\bar{a}}{2} \leq \gamma \left(\frac{3}{2}\bar{a}\right)^2 + \frac{\bar{a}}{2} \leq \gamma \frac{9}{4} \bar{a}^2 + \frac{\bar{a}}{2} \leq \bar{a},$$

where we use (28).

Due to the inequality in (30), the same argument can be applied with $u = x_1$, to obtain the existence of x_2 such that $\|x_2 - s\| \leq \gamma \|x_1 - s\|^2$, and in the same way we get the existence of x_k satisfying (26) for all k .

Finally, noting that, from the third inequality in (28),

$$\theta := \gamma \|u - s\| \leq \gamma (\|u - \bar{x}\| + \|s - \bar{x}\|) \leq \gamma \left(\bar{a} + \frac{\bar{a}}{2}\right) < 1,$$

using (26) we obtain

$$\|x_{k+1} - s\| \leq \theta^{2^k - 1} \|u - s\|$$

and therefore the sequence $\{x_1, \dots, x_k, \dots\}$ is convergent to s with quadratic rate as in (26). This completes the proof. \square

4 A Lyusternik-Graves theorem for Newton's method

In this section we present a Lyusternik-Graves type theorem connecting the metric regularity of the linearized mapping (15) and a mapping whose values are the sets of all convergent sequences generated by Newton's method (18). This result shows that Newton's iteration is, roughly, as "stable" as the mapping of the inclusion to be solved. Such a conclusion may have important implication in the analysis of the effect of various errors, including the errors of approximations of the problem in hand, on the complexity of the method. We shall not go into this further in the current paper, only noting that the idea to consider "sequential open mapping theorems" may be applied to other classes of iterative methods.

We start with a preliminary result that extends in certain way the main step in the proof of Theorem 3.2.

Lemma 4.1. *Suppose that the mapping G defined in (15) is metrically regular at \bar{x} for 0 and let γ , γ_1 and γ_2 be positive constants such that*

$$\gamma > \frac{1}{2} \operatorname{reg}(G; \bar{x}|0) \cdot \widehat{\operatorname{lip}}_x(D_x f; (\bar{p}, \bar{x})), \quad \gamma_1 > \operatorname{reg}(G; \bar{x}|0) \cdot \|D_p f(\bar{p}, \bar{x})\|,$$

and

$$\gamma_2 > \operatorname{reg}(G; \bar{x}|0) \cdot \operatorname{lip}(D_x f; (\bar{p}, \bar{x})).$$

Then there exist positive α and ζ such that for every $p, p' \in \mathbb{B}_\zeta(\bar{p})$, $u, u' \in \mathbb{B}_\alpha(\bar{x})$ and $x \in G_{p,u}^{-1}(0) \cap \mathbb{B}_\alpha(\bar{x})$ there exists $x' \in G_{p',u'}^{-1}(0)$ satisfying

$$(32) \quad \|x - x'\| \leq \gamma \|u - u'\|^2 + \gamma_1 \|p - p'\| + \gamma_2 (\|p - p'\| + \|u - u'\|) \|x - u\|.$$

Proof. Let $\lambda' > \lambda > \operatorname{reg}(G; \bar{x}|0)$, $L > \widehat{\operatorname{lip}}_x(D_x f; (\bar{p}, \bar{x}))$, $L_1 > \|D_p f(\bar{p}, \bar{x})\| = \widehat{\operatorname{lip}}_p(f; (\bar{p}, \bar{x}))$ and $L_2 > \operatorname{lip}(D_x f; (\bar{p}, \bar{x}))$ be such that

$$(33) \quad \gamma > \frac{\lambda'}{2} L, \quad \gamma_1 > \lambda' L_1, \quad \gamma_2 > \lambda' L_2.$$

Now we chose positive α and ζ which are smaller than the numbers a and c in the claim of Corollary 3.1 corresponding to λ and such that $D_x f$ is Lipschitz with respect to $x \in \mathbb{B}_\alpha(\bar{x})$ with constant L uniformly in $p \in \mathbb{B}_\zeta(\bar{p})$, f is Lipschitz with constant L_1 with respect to $p \in \mathbb{B}_\zeta(\bar{p})$ uniformly in $x \in \mathbb{B}_\alpha(\bar{x})$, and D_x is Lipschitz with constant L_2 in $\mathbb{B}_\zeta(\bar{p}) \times \mathbb{B}_\alpha(\bar{x})$.

Let p, p', u, u', x be as in the statement of the lemma. If $d(0, G_{p',u'}(x)) = 0$ by the closedness of $G_{p',u'}^{-1}(0)$ we obtain that $x \in G_{p',u'}^{-1}(0)$ and there is nothing more to prove. If not, then, from Corollary 3.1 we get

$$d(x, G_{p',u'}^{-1}(0)) \leq \lambda d(0, G_{p',u'}(x)),$$

hence there exists $x' \in G_{p',u'}^{-1}(0)$ such that

$$(34) \quad \|x - x'\| \leq \lambda' d(0, G_{p',u'}(x)).$$

Let us estimate the right-hand side of (34). Since we have

$$\begin{aligned} 0 \in G_{p,u}(x) &= f(p, u) + D_x f(p, u)(x - u) + F(x) \\ &= G_{p',u'}(x) + f(p, u) + D_x f(p, u)(x - u) - f(p', u') - D_x f(p', u')(x - u'), \end{aligned}$$

the relation (34) implies

$$\|x - x'\| \leq \lambda' \|f(p, u) + D_x f(p, u)(x - u) - f(p', u') - D_x f(p', u')(x - u')\|.$$

By the choice of the constants γ , γ_1 and γ_2 and using an estimation analogous to (24) we obtain

$$\begin{aligned}
\|x - x'\| &\leq \lambda' [\|f(p', u) + D_x f(p', u')(x - u) - f(p', u') - D_x f(p', u')(x - u')\| \\
&\quad + L_1 \|p - p'\| + L_2 (\|p - p'\| + \|u - u'\|) \|x - u\|] \\
&\leq \lambda' \|f(p', u) + D_x f(p', u')(u' - u) - f(p', u')\| \\
&\quad + \gamma_1 \|p - p'\| + \gamma_2 (\|p - p'\| + \|u - u'\|) \|x - u\| \\
&\leq \gamma \|u - u'\|^2 + \gamma_1 \|p - p'\| + \gamma_2 (\|p - p'\| + \|u - u'\|) \|x - u\|.
\end{aligned}$$

This completes the proof. \square

We are now ready to present the main result of this paper. For that purpose we first define a mapping acting from the value of the parameter and the starting point to the set of all sequences generated by Newton's method (18).

Let $cl_\infty(X)$ be the linear space of all infinite sequences $\xi = \{x_1, x_2, \dots, x_k, \dots\}$ with elements $x_k \in X$, $k = 1, 2, \dots$, that are *convergent* to some point $x \in X$. We equip this set with the supremum norm

$$\|\xi\|_\infty = \sup_{k \geq 1} \|x_k\|$$

which makes it a linear normed space.

Define the mapping $\Xi : P \times X \rightrightarrows cl_\infty(X)$ as follows:

$$(35) \quad \Xi : (p, u) \mapsto \left\{ \xi = \{x_1, x_2, \dots\} \in cl_\infty(X) \mid \xi \text{ is such that} \right. \\
\left. \begin{aligned} &f(p, x_k) + D_x f(p, x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0 \\ &\text{for every } k = 0, 1, \dots, \text{ with } x_0 = u \end{aligned} \right\}.$$

By using the notation in (19), we can define equivalently Ξ as

$$\Xi : (p, u) \mapsto \{ \xi \in cl_\infty(X) \mid x_0 = u \text{ and } G_{p, x_k}(x_{k+1}) \ni 0 \text{ for every } k = 0, 1, \dots \}.$$

Note that if $s \in S(p)$ then the constant sequence $\{s, \dots, s, \dots\}$ belongs to $\Xi(p, s)$. Also note that if $\xi \in \Xi(p, u)$ for some (p, u) close enough to (\bar{p}, \bar{x}) , then by definition ξ is convergent and since F has closed graph its limit is a solution of (13) for p . Denote $\bar{\xi} = \{\bar{x}, \dots, \bar{x}, \dots\}$; then $\bar{\xi} \in \Xi(\bar{p}, \bar{x})$. Our main result presented next is stated in two ways: the first exhibits the qualitative side of it while the second one gives quantitative estimates.

Theorem 4.2 (Lyusternik-Graves for Newton's method). *If the mapping G defined in (15) is metrically regular at \bar{x} for 0 then the mapping Ξ defined in (35) has the partial Aubin property with respect to p uniformly in x at (\bar{p}, \bar{x}) for $\bar{\xi}$. If the function f satisfies the ample*

parameterization condition (16), then the converse implication holds as well: if the mapping Ξ has the partial Aubin property with respect to p uniformly in x at (\bar{p}, \bar{x}) for $\bar{\xi}$, then the mapping G is metrically regular at \bar{x} for 0.

In fact, we have the following stronger statement: if the mapping G defined in (15) is metrically regular at \bar{x} for 0 then the mapping Ξ has the Aubin property in both p and u at (\bar{p}, \bar{x}) for $\bar{\xi}$, with

$$(36) \quad \widehat{\text{lip}}_u(\Xi; (\bar{p}, \bar{x}) | \bar{\xi}) = 0 \quad \text{and} \quad \widehat{\text{lip}}_p(\Xi; (\bar{p}, \bar{x}) | \bar{\xi}) \leq \text{reg}(G; \bar{x} | 0) \cdot \|D_p f(\bar{p}, \bar{x})\|.$$

If the function f satisfies the ample parameterization condition (16), then

$$\text{reg}(G; \bar{x} | 0) \leq \widehat{\text{lip}}_p(\Xi; (\bar{p}, \bar{x}) | \bar{\xi}) \cdot \|D_p f(\bar{p}, \bar{x})^{-1}\|^{-1},$$

and in effect, the first relation in (36) holds as well provided that $\widehat{\text{lip}}_p(\Xi; (\bar{p}, \bar{x}) | \bar{\xi}) < \infty$.

Proof. Fix $\gamma, \gamma_1, \gamma_2$ as in Lemma 4.1 and let α, ζ be the corresponding constants from Lemma 4.1, while \bar{a} and \bar{c} are the constants from Theorem 3.2. Choose positive reals ε and d satisfying the inequalities

$$(37) \quad \varepsilon \leq \frac{\bar{a}}{2}, \quad \varepsilon \leq \alpha, \quad \tau := 2(\gamma + \gamma_2)\varepsilon < \frac{1}{8},$$

$$(38) \quad d \leq \bar{c}, \quad d \leq \zeta, \quad \frac{1}{1-\tau}(\gamma_1 + \tau)d < \frac{\varepsilon}{8},$$

$$(39) \quad \varepsilon(S(p) \cap \mathcal{B}_{\varepsilon/2}(\bar{x}), S(p')) < \gamma_1 \|p - p'\| \quad \text{for } p, p' \in \mathcal{B}_d(\bar{p}), p \neq p'.$$

The existence of ε and d such that the last relation (39) holds is implied by the Aubin property of S claimed in Theorem 2.3.

Let $p, p' \in \mathcal{B}_d(\bar{p})$, $u, u' \in \mathcal{B}_\varepsilon(\bar{x})$, and $\xi = \{x_1, x_2, \dots\} \in \Xi(p, u) \cap \mathcal{B}_{\varepsilon/2}(\bar{\xi})$. Then ξ is convergent and its limit is an element of $S(p)$. Let

$$\delta_k := \tau^k \|u - u'\| + \frac{1 - \tau^k}{1 - \tau} (\gamma_1 + \tau) \|p - p'\|, \quad k = 0, 1, \dots$$

The last inequalities in (37) and (38) imply $\delta_k < \varepsilon/2$.

First we define a sequence $\xi' = \{x'_1, x'_2, \dots\} \in \Xi(p', u')$ with the additional property that

$$(40) \quad \|x_k - x'_k\| \leq \delta_k, \quad \|x'_k - \bar{x}\| \leq \varepsilon.$$

Since $p, p' \in \mathcal{B}_d(\bar{p}) \subset \mathcal{B}_\zeta(\bar{p})$, $u, u', x_1 \in \mathcal{B}_\varepsilon(\bar{x}) \subset \mathcal{B}_\alpha(\bar{x})$ and $x_1 \in G_{p,u}^{-1}(0)$, according to Lemma 4.1 there exists $x'_1 \in G_{p',u'}^{-1}(0)$ such that

$$\|x_1 - x'_1\| \leq \gamma \|u - u'\|^2 + \gamma_1 \|p - p'\| + \gamma_2 (\|p - p'\| + \|u - u'\|) \|u - x_1\|.$$

Using (37) and (38) we obtain

$$\begin{aligned}\|x_1 - x'_1\| &\leq 2\gamma\varepsilon\|u - u'\| + \gamma_1\|p - p'\| + \gamma_2(\|p - p'\| + \|u - u'\|)2\varepsilon \\ &\leq \tau\|u - u'\| + (\gamma_1 + \tau)\|p - p'\| = \delta_1.\end{aligned}$$

In addition we have

$$(41) \quad \|x'_1 - \bar{x}\| \leq \|x'_1 - x_1\| + \|x_1 - \bar{x}\| \leq \delta_1 + \frac{\varepsilon}{2} \leq \varepsilon.$$

Now assume that x'_k is already defined so that (40) holds. Applying Lemma 4.1 for p, p', x_k, x'_k and $x_{k+1} \in G_{p, x_k}^{-1}(0) \cap \mathcal{B}_{\varepsilon/2}(\bar{x})$ (instead of (p, p', u, u', x_1)) we obtain that there exists $x'_{k+1} \in G_{p', x'_k}^{-1}(0)$ such that

$$\|x_{k+1} - x'_{k+1}\| \leq \gamma\|x_k - x'_k\|^2 + \gamma_1\|p - p'\| + \gamma_2(\|p - p'\| + \|x_k - x'_k\|)\|x_k - x_{k+1}\|.$$

In the same way as above we estimate

$$\begin{aligned}\|x_{k+1} - x'_{k+1}\| &\leq 2\gamma\varepsilon\|x_k - x'_k\| + \gamma_1\|p - p'\| + \gamma_2(\|p - p'\| + \|x_k - x'_k\|)2\varepsilon \\ &\leq 2(\gamma + \gamma_2)\varepsilon\|x_k - x'_k\| + (\gamma_1 + 2\gamma_2\varepsilon)\|p - p'\| \\ &\leq \tau\delta_k + (\gamma_1 + \tau)\|p - p'\| \\ &\leq \tau \left(\tau^k\|u - u'\| + \frac{1 - \tau^k}{1 - \tau}(\gamma_1 + \tau)\|p - p'\| \right) + (\gamma_1 + \tau)\|p - p'\| \\ &= \tau^{k+1}\|u - u'\| + \frac{1 - \tau^{k+1}}{1 - \tau}(\gamma_1 + \tau)\|p - p'\| = \delta_{k+1}.\end{aligned}$$

To complete the inductive definition of the sequence it remains to note that $\|x'_{k+1} - \bar{x}\| \leq \varepsilon$ follows from the last estimate in exactly the same way as in (41).

Since the sequence ξ is convergent to some $s \in S(p)$, there exists a natural number N such that

$$\|x_k - s\| \leq \tau(\|u - u'\| + \|p - p'\|) \quad \text{for all } k \geq N.$$

We will now take the finite sequence x'_1, \dots, x'_N and extend it to a sequence $\xi' \in \Xi(p', u')$. If $p = p'$ take $s' = s$. If not, the Aubin property of the solution map S in (39) implies that there exists $s' \in S(p')$ such that $\|s' - s\| \leq \gamma_1\|p - p'\|$. We also have

$$\|s' - \bar{x}\| \leq \|s' - s\| + \|s - \bar{x}\| \leq \gamma_1\|p - p'\| + \varepsilon/2 \leq 2d\gamma_1 + \varepsilon/2 < \bar{a}/2,$$

by (37) and (38). Thus, for $k > N$ we can define x'_k by using Theorem 3.2 as a Newton sequence for p' and initial point x'_N , quadratically convergent to s' . Observe that $p' \in$

$\mathbb{B}_d(\bar{p}) \subset \mathbb{B}_\varepsilon(\bar{p})$ and $x'_N \in \mathbb{B}_{\bar{a}}(\bar{x})$ by the second inequality in (40) and since $\varepsilon < \bar{a}/2$ as assumed in (37). Using (37) and (38) we also have

$$\begin{aligned}
\|x'_N - s'\| &\leq \|x'_N - x_N\| + \|x_N - s\| + \|s - s'\| \\
&\leq \delta_N + \tau(\|u - u'\| + \|p - p'\|) + \gamma_1 \|p - p'\| \\
(42) \quad &\leq 2\tau\|u - u'\| + \left(\frac{\gamma_1 + \tau}{1 - \tau} + \gamma_1 + \tau\right) \|p - p'\| \\
&\leq \frac{1}{4}2\varepsilon + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon.
\end{aligned}$$

According to Theorem 3.2 there is a sequence $x'_{N+1}, \dots, x'_k, \dots$ such that

$$(43) \quad \|x'_{k+1} - s'\| \leq \gamma \|x'_k - s'\|^2 \quad \text{for all } k \geq N.$$

Then, for $k > N$,

$$\|x'_k - s'\| \leq \gamma^{2^{k-N}-1} \|x'_N - s'\|^{2^{k-N}} \leq \gamma^{2^{k-N}-1} \varepsilon^{2^{k-N}} \leq \varepsilon,$$

where we apply (42) and that $\gamma\varepsilon \leq 1/4$ due to (37). Therefore, using this last estimate in (43), we get

$$(44) \quad \|x'_{k+1} - s'\| \leq \gamma\varepsilon \|x'_k - s'\| \quad \text{for all } k \geq N.$$

Recalling that $\gamma\varepsilon \leq \frac{1}{4}$ due to (37), we have

$$\|x'_k - s'\| \leq \gamma\varepsilon \|x'_N - s'\| \quad \text{for all } k > N.$$

Using (42) we obtain that for $k \geq N + 1$

$$\begin{aligned}
\|x_k - x'_k\| &\leq \|x_k - s\| + \|s - s'\| + \|s' - x'_k\| \\
&\leq \tau(\|u - u'\| + \|p - p'\|) + \gamma_1 \|p - p'\| + \varepsilon\gamma \|x'_N - s'\| \\
(45) \quad &\leq (\tau + 2\tau\varepsilon\gamma)\|u - u'\| + \left[\tau + \gamma_1 + \varepsilon\gamma \left(\frac{\gamma_1 + \tau}{1 - \tau} + \gamma_1 + \tau\right)\right] \|p - p'\|.
\end{aligned}$$

The last expression is clearly greater than δ_k for any $k \geq 1$, hence we obtain that the same estimate holds also for $k \leq N$ since for such k we have (40). Thus the distance $d(\xi, \Xi(p', u'))$ is also bounded by the expression in (45). This holds for every $p, p' \in \mathbb{B}_d(\bar{p})$, $u, u' \in \mathbb{B}_\varepsilon(\bar{x})$ and every $\xi \in \Xi(p, u) \cap \mathbb{B}_{\varepsilon/2}(\bar{\xi})$ and ε is arbitrarily small. Observe that when ε is small, then τ is also small, hence the constant multiplying $\|u - u'\|$ is arbitrarily close to zero, and that the constant multiplying $\|p - p'\|$ is arbitrarily close to γ_1 . This yields (36) and completes the proof of the first part of the theorem.

Now, let the ample parameterization condition (16) be satisfied. Let κ , c and a be positive constants such that

$$e(\Xi(p, u) \cap \Omega, \Xi(p', u)) \leq \kappa \|p - p'\| \quad \text{whenever } p, p' \in \mathcal{B}_c(\bar{p}), u \in \mathcal{B}_a(\bar{x}),$$

where Ω is a neighborhood of $\bar{\xi}$. Make a smaller if necessary so that $\mathcal{B}_a(\bar{\xi}) \subset \Omega$ and then take c smaller so that $\kappa c < a/2$. Since $\text{gph } F$ is closed, it follows that for any $p \in \mathcal{B}_c(\bar{p})$ and any sequence with components $x_k \in \mathcal{B}_a(\bar{x})$ convergent to x and satisfying

$$(46) \quad f(p, x_k) + D_x f(p, x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0 \quad \text{for all } k = 1, 2, \dots,$$

one has $f(p, x) + F(x) \ni 0$, that is, $x \in S(p)$.

We will prove that S has the Aubin property at \bar{p} for \bar{x} , and then we will apply Theorem 2.3 to show the metric regularity of G at \bar{x} for 0. Pick $p, p' \in \mathcal{B}_{c/2}(\bar{p})$ with $p \neq p'$ and $x \in S(p) \cap \mathcal{B}_{a/2}(\bar{x})$ (if there is no such x we are done). Let $\chi := \{x, x, \dots\}$. Since $\|\chi - \bar{\xi}\|_\infty = \|x - \bar{x}\| \leq a/2$, we have $\chi \in \Xi(p, x) \cap \Omega$. Hence

$$d(\chi, \Xi(p', x)) \leq \kappa \|p - p'\|.$$

Take $\varepsilon > 0$ such that $(\kappa + \varepsilon)c \leq a/2$. Then there is some $\Psi \in \Xi(p', x)$ such that

$$\|\chi - \Psi\|_\infty \leq (\kappa + \varepsilon) \|p - p'\|,$$

with $\Psi = \{x'_1, x'_2, \dots\}$ and $x'_k \rightarrow x' \in X$. For all k we have

$$\|x'_k - \bar{x}\| \leq \|x'_k - x\| + \|x - \bar{x}\| \leq \|\Psi - \chi\|_\infty + a/2 \leq (\kappa + \varepsilon)c + a/2 \leq a.$$

Hence from (46) we obtain $x' \in S(p') \cap \mathcal{B}_a(\bar{x})$. Moreover,

$$\|x - x'\| \leq \|x - x'_k\| + \|x'_k - x'\| \leq \|\chi - \Psi\|_\infty + \|x'_k - x'\| \leq (\kappa + \varepsilon) \|p - p'\| + \|x'_k - x'\|.$$

Making $k \rightarrow \infty$, we get $\|x - x'\| \leq (\kappa + \varepsilon) \|p - p'\|$. Thus,

$$d(x, S(p')) \leq \|x - x'\| \leq (\kappa + \varepsilon) \|p - p'\|.$$

Taking $\varepsilon \downarrow 0$, we have the Aubin property of S at \bar{p} for \bar{x} with constant κ , as claimed. From Theorem 2.3, G is metrically regular at \bar{x} for 0 with

$$\text{reg}(G; \bar{x} | 0) \leq \kappa \|D_p f(\bar{p}, \bar{x})^{-1}\|^-.$$

Since κ can be taken arbitrarily close to $\widehat{\text{lip}}_p(\Xi; (\bar{p}, \bar{x}) | \bar{\xi})$, we obtain the desired result. \square

To shed more light on the kind of result we just proved, consider the special case when $f(p, x)$ has the form $f(x) - p$ and take, for simplicity, $\bar{p} = 0$. Then the Newton iteration (18) becomes

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

where the ample parameterization condition (16) holds automatically. As in (19), let

$$X \ni x \mapsto G_u(x) = f(u) + Df(u)(x - u) + F(x), \quad \text{for } u \in X.$$

and define the mapping $\Gamma : cl_\infty(X) \rightrightarrows X \times P$ as

$$(47) \quad \Gamma : \xi \mapsto \left\{ \begin{pmatrix} u \\ p \end{pmatrix} \mid u = x_0 \text{ and } G_{x_k}(x_{k+1}) \ni p \text{ for every } k = 0, 1, \dots \right\}.$$

Then Theorem 4.2 becomes the following characterization result:

Corollary 4.3 (A symmetric Lyusternik-Graves for Newton's method). *The following are equivalent:*

- (i) *The mapping $G = f(\bar{x}) + Df(\bar{x})(\cdot - \bar{x}) + F$, or equivalently, the mapping $f + F$, is metrically regular at \bar{x} for 0;*
- (ii) *The mapping Γ defined in (47) is metrically regular at $\bar{\xi}$ for $(\bar{x}, 0)$.*

Next comes a statement similar to Theorem 4.2 for the case when the mapping G is *strongly metrically regular*. This case was considered in Dontchev and Rockafellar [6], see also Theorems 6D.2 and 6D.3 in [7], where a *sequential implicit function theorem* was established for Newton's method. Here, we complement these results by adding the ample parameterization case and drop one of the assumption in [7, Theorem 6D.3] that turns out to be superfluous.

To introduce the strong metric regularity property, we utilize the notion of graphical localization. A *graphical localization* of a mapping $S : X \rightrightarrows Y$ at $(\bar{x}, \bar{y}) \in \text{gph } S$ is a mapping $\tilde{S} : X \rightrightarrows Y$ such that $\text{gph } \tilde{S} = (U \times V) \cap \text{gph } S$ for some neighborhood $U \times V$ of (\bar{x}, \bar{y}) . Then we say that a mapping $S : X \rightrightarrows Y$ is *strongly metrically regular* at \bar{x} for \bar{y} if the metric regularity condition in Definition 1.1 is satisfied by some κ and neighborhoods U of \bar{x} and V of \bar{y} and, in addition, the graphical localization of S^{-1} with respect to U and V is single-valued. Equivalently, the graphical localization $V \ni y \mapsto S^{-1}(y) \cap U$ is a Lipschitz continuous function whose Lipschitz constant equals κ .

Theorem 4.4. *Suppose that the mapping G defined in (15) is strongly metrically regular at \bar{x} for 0. Then the mapping Ξ in (35) has a Lipschitz single-valued localization ξ at (\bar{p}, \bar{x}) for $\bar{\xi}$, with*

$$(48) \quad \widehat{\text{lip}}_u(\xi; (\bar{p}, \bar{x})) = 0 \quad \text{and} \quad \widehat{\text{lip}}_p(\xi; (\bar{p}, \bar{x})) \leq \text{reg}(G; \bar{x} | 0) \cdot \widehat{\text{lip}}_p(f; (\bar{p}, \bar{x})).$$

Moreover, for (p, u) close to (\bar{p}, \bar{x}) , $\xi(p, u)$ is a quadratically convergent sequence to a locally unique solution. Also, the same conclusion holds if we replace the space $cl_\infty(X)$ in the definition of Ξ by the space of all sequences with elements in X , not necessarily convergent, equipped with the $l_\infty(X)$ norm.

If the function f satisfies the ample parameterization condition (16), then the converse implication holds as well: the mapping G is strongly metrically regular at \bar{x} for 0 provided that Ξ has a Lipschitz continuous single-valued localization ξ at (\bar{p}, \bar{x}) for $\bar{\xi}$.

Proof. Assume that the mapping G is strongly metrically regular at \bar{x} for 0. Then the solution mapping S is strongly metrically regular, see e.g. [7, Theorem 5F.4]. Consider the mapping $\widehat{\Xi}$ defined in the same way as Ξ in (35) but with $cl_\infty(X)$ replaced by l_∞ ; in other words, we now do not require that the sequences in the image of the mapping be convergent. Observe that $\Xi(p, u) \subset \widehat{\Xi}(p, u)$ for all $(p, u) \in P \times X$. According to Theorem 3.1 in [6], see also [7, Theorem 6D.2], the mapping $\widehat{\Xi}$ has a single valued graphical localization ξ at (\bar{p}, \bar{x}) for $\bar{\xi}$ satisfying (48). Moreover, from Theorem 3.2, for (p, u) close to (\bar{p}, \bar{x}) , the values $\xi(p, u)$ of this localization are quadratically convergent sequences to the locally unique solution $x(p)$. Thus, any graphical localization of the mapping $\widehat{\Xi}$ with sufficiently small neighborhoods agrees with the corresponding graphical localization of the mapping Ξ defined in (35), which gives us the first claim of the theorem.

Assume that the ample parameterization condition (16) holds and let Ξ have a Lipschitz localization ξ at (\bar{p}, \bar{x}) for $\bar{\xi}$; that is, $(p, u) \mapsto \Xi(p, u) \cap \mathcal{B}_\beta(\bar{\xi})$ is a singleton $\xi(p, x)$ for any $p \in \mathcal{B}_\alpha(\bar{p})$ and $u \in \mathcal{B}_\alpha(\bar{x})$. Then in particular, Ξ has the Aubin property with respect to p uniformly in x at (\bar{p}, \bar{x}) for $\bar{\xi}$ and hence, by Theorem 4.2, G is metrically regular at \bar{x} for 0. Take \bar{a} in (28) smaller if necessary so that $\bar{a} \leq \beta$. Since by Theorem 2.3 the solution mapping S has the Aubin property at \bar{p} for \bar{x} , it remains to show that S is locally nowhere multivalued.

Take $a := \min\{\bar{a}/2, \alpha, \beta\}$ and $c := \min\{\bar{c}, \alpha\}$ and let $p \in \mathcal{B}_c(\bar{p})$ and $x, x' \in S(p) \cap \mathcal{B}_a(\bar{x})$, $x \neq x'$. Clearly, $\{x, x, \dots\} = \Xi(p, x) \cap \mathcal{B}_\beta(\bar{\xi}) = \xi(p, x)$. Further, according to Theorem 3.2 there exists a Newton sequence ξ' for p starting again from x and each element of which is in $\mathcal{B}_a(\bar{x})$, which converges to x' , thus $\xi' \in \Xi(p, x) \cap \mathcal{B}_\beta(\bar{\xi})$. But this contradicts the assumption that $\Xi(p, x) \cap \mathcal{B}_\beta(\bar{\xi})$ is a singleton. Thus, S has a single-valued graphical localization at \bar{p} for \bar{x} and by its Aubin property, this localization is Lipschitz continuous, see [7, Proposition 3G.1] whose extension to Banach spaces is straightforward. It remains to apply [7, Theorem 5F.5], which asserts that the latter property is equivalent to the strong metric regularity of G at \bar{x} for 0. \square

5 Inexact Newton method and application to optimal control

In this section we focus on the following modification of Newton's method (18):

$$(49) \quad f(x_k) + Df(x_k)(x_{k+1} - x_k) + F(x_{k+1}) \ni p_k, \quad \text{for } k = 0, 1, \dots,$$

with a given starting point x_0 , where now the parameter may change from iteration to iteration but does not appear in the function f . (A more general case where f depends on p_k could be considered but we shall not deal here with this extension.) Here the term p_k can be regarded as an error and in that case (49) can be interpreted as an inexact version of Newton's method, see [12] for background. We consider the sequence $\pi = \{p_k\}$ as an element of $l_\infty(P)$.

Theorem 5.1 (convergence of inexact Newton's method). *Suppose that the mapping $f + F$ is metrically regular at \bar{x} for 0 or, equivalently, according to Corollary 2.4, the mapping*

$$(50) \quad x \mapsto G(x) := f(\bar{x}) + Df(\bar{x})(x - \bar{x}) + F(x)$$

is metrically regular at \bar{x} for 0. Consider inexact Newton's method (49). Then there exist positive constants a and c such that:

(i) for any sequence $\pi = \{p_k\}$ which is linearly convergent to zero in the way that $\|p_k\| \leq c\theta^k$, $k = 0, 1, \dots$ for some $\theta \in (0, 1)$, and for any $x_0 \in \mathcal{B}_a(\bar{x})$ there exists a sequence $\{x_k\}$ starting from x_0 and generated by (49) for π which is linearly convergent to \bar{x} ;

(ii) for any sequence $\pi = \{p_k\}$ which is quadratically convergent to zero in a way that $\|p_k\| \leq \gamma\theta^{2^k-1}$, $k = 0, 1, \dots$ for some $\theta \in (0, 1)$ and $\gamma > 0$ and for any $x_0 \in \mathcal{B}_a(\bar{x})$ there exists a sequence $\{x_k\}$ starting from x_0 and generated by (49) for π which is quadratically convergent to \bar{x} .

Proof. We use the idea of the proof of Theorem 3.2. Choose $\lambda > \text{reg}(G; \bar{x}|0)$ and $L > \text{lip}(Df; \bar{x})$. Then according to Corollary 3.1 for the mapping $G_u(x) := f(u) + Df(u)(x - u) + F(x)$ there exist positive a and c such that

$$d(x, G_u^{-1}(p)) \leq \lambda d(p, G_u(x)) \quad \text{for every } u, x \in \mathcal{B}_a(\bar{x}), p \in \mathcal{B}_c(0).$$

Make a and c smaller if necessary so that

$$(51) \quad a \leq \frac{1}{\lambda L} \quad \text{and} \quad \lambda c \leq \frac{a}{2}.$$

Proceeding as in the proof of Theorem 3.2, we fix $u \in \mathcal{B}_a(\bar{x})$ and $p_0 \in \mathcal{B}_c(0)$ and find x_1 such that

$$(52) \quad \|x_1 - \bar{x}\| \leq \lambda d(p_0, G_u(\bar{x})) \leq \frac{\lambda L}{2} \|u - \bar{x}\|^2 + \lambda \|p_0\|.$$

From (51), $x_1 \in B_a(\bar{x})$ and also $\rho := \lambda La/2 < 1$. Hence

$$\|x_1 - \bar{x}\| \leq \rho \|u - \bar{x}\| + \lambda \|p_0\|.$$

By induction, we get

$$\|x_{k+1} - \bar{x}\| \leq \rho \|x_k - \bar{x}\| + \lambda \|p_k\| \quad \text{for all } k,$$

which gives us

$$\|x_k - \bar{x}\| \leq \rho^k \|u - \bar{x}\| + \lambda \sum_{i=1}^k \rho^{k-i} \|p_{i-1}\|.$$

We now consider separately the cases (i) and (ii).

(i) If $\|p_k\| \leq c\theta^k$ for some $\theta \in (0, 1)$ then for some c, γ and γ' with $\max\{\rho, \theta\} := \gamma' < \gamma < 1$ we have

$$\|x_k - \bar{x}\| \leq \gamma^k \|u - \bar{x}\| + \lambda c \gamma^{k-1} \sum_{i=1}^{\infty} \left(\frac{\gamma'}{\gamma}\right)^i \leq c' \gamma^k,$$

that is, $\{x_k\}$ converges linearly to \bar{x} .

(ii) Let π be quadratically convergent as described in the statement of the theorem. Take $\eta > 0$ such that

$$\frac{(\lambda + \eta)^2 \theta}{\eta^2} < 1.$$

Then decrease, if necessary, the constants a and c so that, in addition to (51), we have also

$$(53) \quad \frac{\lambda L}{2} \left(\frac{\lambda L}{2} a^2 + (\lambda + \eta)c \right) < 1 \quad \text{and} \quad (\lambda + \eta) \frac{\theta}{c\eta^2} \frac{\lambda L}{2} a^2 + \frac{(\lambda + \eta)^2 \theta}{\eta^2} < 1.$$

This can be achieved by multiplying the already defined a and c by a sufficiently small common multiplier.

Since (52) holds for any x_{k+1} in place of x_1 and x_k in place of u , we have

$$\|x_{k+1} - \bar{x}\| \leq \frac{\lambda L}{2} \|x_k - \bar{x}\|^2 + \lambda \|p_k\|.$$

Denote $\Delta_k = \|x_k - \bar{x}\|$ and $\alpha_k = c\theta^{2^k-1}$. Then

$$\Delta_{k+1} \leq \frac{\lambda L}{2} \Delta_k^2 + \lambda \alpha_k, \quad \alpha_{k+1} = \frac{\theta}{c} \alpha_k^2.$$

Thus, for $\omega_k = \Delta_k + \eta \alpha_{k-1}$, $k = 1, 2, \dots$ we obtain

$$\omega_{k+1} \leq \frac{\lambda L}{2} \Delta_k^2 + (\lambda + \eta) \alpha_k = \frac{\lambda L}{2} \Delta_k^2 + (\lambda + \eta) \frac{\theta}{c\eta^2} (\eta \alpha_{k-1})^2,$$

hence,

$$\omega_{k+1} \leq \max \left\{ \frac{\lambda L}{2}, (\lambda + \eta) \frac{\theta}{c\eta^2} \right\} \omega_k^2.$$

In order to conclude that $\{\omega_k\}$, and hence $\{\Delta_k\}$, is quadratically convergent it is enough to verify that

$$\max \left\{ \frac{\lambda L}{2}, (\lambda + \eta) \frac{\theta}{c\eta^2} \right\} \omega_1 \leq \max \left\{ \frac{\lambda L}{2}, (\lambda + \eta) \frac{\theta}{c\eta^2} \right\} \left(\frac{\lambda L}{2} a^2 + (\lambda + \eta) c \right) < 1.$$

This last inequality is implied by (53). \square

Theorem 5.2 (error in inexact Newton's method). *On the assumption of metric regularity in Theorem 5.1, there exists a constant $d > 0$ such that for every $\tau \in (0, 1)$ there are positive numbers a and c such that for any sequence $\pi = \{p_k\}$ with $\|\pi\|_\infty \leq c$, if $\{x_k\}$ is a sequence generated by (49) for π all elements of which are in $B_a(\bar{x})$, then there exists a sequence $\{\hat{x}_k\}$ generated by the (exact) Newton's method with $p_k = 0$ for all k and starting from x_0 , such that*

$$\|x_k - \hat{x}_k\| \leq d \sum_{i=0}^{k-1} \tau^{k-i-1} \|p_i\| \quad \text{for } k = 1, 2, \dots$$

In addition, if π are linearly convergent, then a sequence $\{\hat{x}_k\}$ as above exists such that $\{\hat{x}_k - x_k\}$ is linearly convergent to zero. If π are quadratically convergent, then $\{\hat{x}_k - x_k\}$ is quadratically convergent to zero.

Proof. Let $\gamma, \gamma_1, \gamma_2, \alpha$ and ζ as in Lemma 4.1. Define $d := \gamma_1 + 2\gamma_2$ and fix $\tau \in (0, 1)$. Let a and c be so small that

$$(54) \quad 2a \leq \alpha, \quad \frac{dc}{1-\tau} \leq a, \quad a \leq 1, \quad c \leq \zeta, \quad \gamma cd + 2a\gamma_2 \leq \tau(1-\tau).$$

Consider the equation for θ

$$\frac{\gamma cd}{1-\theta} + 2a\gamma_2 = \theta,$$

or

$$-\theta^2 + (1 + 2a\gamma_2)\theta - (\gamma cd + 2a\gamma_2) = 0.$$

For $\theta = 0$ the left-hand side of the above equation is clearly negative, while for $\theta = \tau$ is positive (using the last inequality in (54)). Thus there is a zero $\theta \in (0, \tau)$.

Now construct \hat{x}_k using Lemma 4.1, with $\hat{x}_0 := x_0$. Skipping some obvious details, we denote $\delta_k := \|x_k - \hat{x}_k\|$ and $\rho_k := \|p_k\|$. Then

$$(55) \quad \delta_{k+1} \leq \gamma \delta_k^2 + \gamma_1 \rho_k + \gamma_2 (\rho_k + \delta_k) \|x_{k+1} - x_k\| \leq (\gamma \delta_k + 2a\gamma_2) \delta_k + d\rho_k,$$

for $k = 0, 1, \dots$. We shall prove inductively that

$$\delta_k \leq d \sum_{i=0}^{k-1} \theta^{k-i-1} \rho_i, \quad \text{for } k = 1, 2, \dots$$

In particular,

$$\delta_k \leq d \sum_{j=0}^{\infty} \theta^j c \leq \frac{dc}{1-\theta}$$

and then

$$\|\hat{x}_k - \bar{x}\| \leq \|x_k - \bar{x}\| + \delta_k \leq a + \delta_k \leq a + \frac{dc}{1-\theta} \leq a + \frac{dc}{1-\tau} \leq 2a \leq \alpha.$$

Obviously $\delta_1 \leq d\rho_0$. Moreover,

$$\delta_{k+1} \leq \left(\frac{\gamma dc}{1-\theta} + 2a\gamma_2 \right) \delta_k + d\rho_k = \theta\delta_k + d\rho_k \leq d \sum_{i=0}^k \theta^{k-i} \rho_i,$$

which completes the proof of the first claim.

If $\rho_k \leq C\eta^k$ for some $\eta \in (0, 1)$ and $C > 0$, then taking $\tau \in (\eta, 1)$ we have

$$\|x_k - \hat{x}_k\| \leq Cd \sum_{i=0}^{k-1} \tau^{k-i-1} \eta^i \leq Cd\tau^{k-1} \sum_{i=0}^{\infty} \left(\frac{\eta}{\tau}\right)^i \leq C'\tau^k.$$

For the last part, we may assume that $\rho_k \leq C\eta^{2^k-1}$ and $\|x_k - \bar{x}\| \leq C\eta^{2^k-1}$ for all k . Then from (55) we obtain

$$\delta_{k+1} \leq \gamma\delta_k^2 + (\gamma_1 + \gamma_2)\rho_k + a\gamma_2\|x_{k+1} - x_k\| \leq \gamma\delta_k^2 + (\gamma_1 + \gamma_2 + 2a\gamma_2)C\eta^{2^k-1}.$$

By a similar argumentation as in the proof of Theorem 5.1, we conclude that δ_k converges quadratically to zero. \square

We will present an application of the last theorem to the following optimal control problem

$$(56) \quad \text{minimize} \quad \int_0^1 \varphi(\zeta(t), u(t)) dt$$

subject to

$$\begin{aligned} \dot{\zeta}(t) &= g(\zeta(t), u(t)), \quad u(t) \in U \text{ for a.e. } t \in [0, 1], \\ \zeta &\in W_0^{1,\infty}(\mathbb{R}^n), \quad u \in L^\infty(\mathbb{R}^m), \end{aligned}$$

where $\varphi : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$, $g : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$, U is a convex and closed set in \mathbb{R}^m . Here ζ denotes the state trajectory of the system, u is the control function, $L^\infty(\mathbb{R}^m)$ denotes the space of essentially bounded and measurable functions with values in \mathbb{R}^m and $W_0^{1,\infty}(\mathbb{R}^n)$ is the space of Lipschitz continuous functions ζ with values in \mathbb{R}^n and such that $\zeta(0) = 0$. We assume that problem (56) has a solution $(\bar{\zeta}, \bar{u})$ and also that there exists a closed set $\Delta \subset \mathbb{R}^n \times \mathbb{R}^m$ and a $\delta > 0$ with $B_\delta(\bar{\zeta}(t), \bar{u}(t)) \subset \Delta$ for almost every $t \in [0, 1]$ so that the functions φ and g are twice continuously differentiable in Δ .

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

$$H(\zeta, \psi, u) = \varphi(\zeta, u) + \psi^\top g(\zeta, u),$$

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

$$(57) \quad \begin{cases} \dot{\zeta}(t) &= g(\zeta(t), u(t)), \quad \zeta(0) = 0, \\ \dot{\psi}(t) &= -\nabla_\zeta H(\zeta(t), \psi(t), u(t)), \quad \psi(1) = 0, \\ 0 &\in \nabla_u H(\zeta(t), \psi(t), u(t)) + N_U(u(t)), \quad \text{for a.e. } t \in [0, 1], \end{cases}$$

where $N_U(u)$ is the normal cone to the set U at the point u . Denote $X = W_0^{1,\infty}(\mathbb{R}^n) \times W_1^{1,\infty}(\mathbb{R}^n) \times L^\infty(\mathbb{R}^m)$ and $Y = L^\infty(\mathbb{R}^n) \times L^\infty(\mathbb{R}^n) \times L^\infty(\mathbb{R}^m)$. Further, for $x = (\zeta, \psi, u)$ let

$$(58) \quad f(x) = \begin{pmatrix} \dot{\zeta} - \nabla_\psi H(\zeta(t), \psi(t), u(t)) \\ \dot{\psi} + \nabla_\zeta H(\zeta(t), \psi(t), u(t)) \\ \nabla_u H(\zeta(t), \psi(t), u(t)) \end{pmatrix}$$

and

$$(59) \quad F(x) = \begin{pmatrix} 0 \\ 0 \\ \mathcal{N}_U(u) \end{pmatrix},$$

where \mathcal{N}_U is the set of all L^∞ selections of the set-valued mapping $t \mapsto N_U(u(t))$ for $t \in [0, 1]$ (this mapping has closed graph). Thus the optimality system (57) can be written as the generalized equation $f(x) + F(x) \ni 0$. The Newton iteration applied to this system is defined for $x = (\zeta, \psi, u)$ as follows⁵:

$$(60) \quad \begin{pmatrix} \dot{\zeta}_{k+1} - \nabla_\psi H(x_k) - \nabla_{\psi x}^2 H(x_k)(x_{k+1} - x_k) \\ \dot{\psi}_{k+1} + \nabla_\zeta H(x_k) + \nabla_{\zeta x}^2 H(x_k)(x_{k+1} - x_k) \\ \nabla_u H(x_k) + \nabla_{ux}^2 H(x_k)(x_{k+1} - x_k) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \mathcal{N}_U(u_{k+1}) \end{pmatrix} \ni 0.$$

⁵ We keep the argument x in the appearing derivatives of H , although in fact, $\nabla_\psi H$ and $\nabla_{\zeta x}^2 H$ depend only on ζ and u .

We will now apply Theorem 5.2 to obtain an a priori estimate for a sequence generated by an inexact Newton iteration resulting from a discretized (finite-dimensional) version of (60) provided by the Euler scheme. (Theorem 5.1 can be also applied in this context but we shall not do this here to keep the paper in reasonable length.) For that purpose we need first to introduce the space of functions that approximate the solution of (57). Let N be a natural number, let $h = 1/N$ be the mesh spacing, and let $t^i = ih$. Denote by $PL_0^N(\mathbb{R}^n)$ the space of piecewise linear and continuous functions ζ_N over the grid $\{t^i\}$ with values in \mathbb{R}^n and such that $\zeta_N(0) = 0$, by $PL_1^N(\mathbb{R}^n)$ the space of piecewise linear and continuous functions ψ_N over the grid $\{t^i\}$ with values in \mathbb{R}^n and such that $\psi_N(1) = 0$, and by $PC^N(\mathbb{R}^m)$ the space of piecewise constant and continuous from the right functions over the grid $\{t^i\}$ with values in \mathbb{R}^m . Clearly, $PL_0^N(\mathbb{R}^n) \subset W_0^{1,\infty}(\mathbb{R}^n)$, $PL_1^N(\mathbb{R}^n) \subset W_1^{1,\infty}(\mathbb{R}^n)$ and $PC^N(\mathbb{R}^m) \subset L^\infty(\mathbb{R}^m)$. Then introduce the products $X^N = PL_0^N(\mathbb{R}^n) \times PL_1^N(\mathbb{R}^n) \times PC^N(\mathbb{R}^m)$ as an approximation space for the triple (ζ, ψ, u) . We identify $\zeta \in PL_0^N(\mathbb{R}^n)$ with the vector $(\zeta^0, \dots, \zeta^N)$ of its values at the mesh points (and similarly for ψ), and $u \in PC^N(\mathbb{R}^m)$ – with the vector (u^0, \dots, u^{N-1}) of the values of u in the mesh subintervals.

We introduce now a Newton iterative process with discretization. Let N_0 be a natural number and let $u_0 \in PC^{N_0}(\mathbb{R}^m)$ be an initial guess for the control. Let ζ_0 and ψ_0 be the corresponding solutions of the Euler discretization with uniform mesh size $h = 1/N_0$ of the primal and adjoint system in (57). Since ζ_0 and ψ_0 can be viewed as piecewise linear functions, the initial approximation $x_0 = (\zeta_0, \psi_0, u_0)$ 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 $\{t_k^i = i/N_k\}_{i=0, \dots, N_k}$ are embedded in the next mesh $\{t_{k+1}^i = i/N_{k+1}\}_{i=0, \dots, N_{k+1}}$. Then, let $x = x_{k+1} = \{x_{k+1}^i\}_i = \{(\zeta_{k+1}^i, \psi_{k+1}^i, u_{k+1}^i)\}_i \in X^N$ be a solution of the discretized Newton's method

$$(61) \quad \begin{pmatrix} \frac{\zeta^{i+1} - \zeta^i}{h_{k+1}} - \nabla_\psi H(x_k(t_{k+1}^i)) - \nabla_{\psi x}^2 H(x_k(t_{k+1}^i))(x^i - x_k(t_{k+1}^i)) \\ \frac{\psi^i - \psi^{i-1}}{h_{k+1}} + \nabla_\zeta H(x_k(t_{k+1}^i)) + \nabla_{\zeta x}^2 H(x_k(t_{k+1}^i))(x^i - x_k(t_{k+1}^i)) \\ \nabla_u H(x_k(t_{k+1}^i)) + \nabla_{ux}^2 H(x_k(t_{k+1}^i))(x^i - x_k(t_{k+1}^i)) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \mathcal{N}_U(u^i) \end{pmatrix} \ni 0,$$

with $\zeta_{k+1}^0 = 0$, $\psi_{k+1}^{N_{k+1}} = 0$, and where $h_{k+1} = 1/N_{k+1}$. The sequence of iterates $\{x^i\}_{i=0, \dots, N_{k+1}}$ is then embedded into the space $X^{N_{k+1}}$ by piecewise linear interpolation for the ζ and ψ 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 $X^{N_{k+1}}$. We note that the iteration (61) can be viewed as a sequential quadratic programming (SQP) method to the discretized optimality system.

Theorem 5.3 (a priori estimate). *Let the mapping $f + F$ with the specifications (58), (59), that is, the mapping of the optimality system (57), be metrically regular at \bar{x} for 0. Then there exist positive constants $C > 0$, $a > 0$ and a natural number \bar{N} such that for every sequence $N_k = \nu^k N_0$, with $N_0 \geq \bar{N}$, a natural number $\nu > 1$, and for every $u_0 \in PC^{N_0}(\mathbb{R}^m) \cap \mathcal{B}_a(\bar{x})$,*

if $\{x_k\}$ is a sequence generated by the discretized Newton's process (61) which is convergent and contained in $\mathbb{B}_a(\bar{x})$, then there exists a sequence $\{\hat{x}_k\}$ generated by the exact Newton's method (60) applied to the continuous optimality system (57) such that

$$\|x_k - \hat{x}_k\| \leq \frac{C}{N_0} \left(\frac{1}{\nu}\right)^k \quad \text{for } k > \frac{1}{\bar{N}}.$$

Proof. Let $x_{k+1} \in X^{N_{k+1}}$ be the $k+1$ iteration of the discretized Newton's process (61), $k \geq 0$ and denote by p_k the residual that x_{k+1} gives when plugged into the exact Newton's inclusion (60). In order to apply Theorem 5.2 we need to estimate this residual p_k in the space $Y = L^\infty(\mathbb{R}^n) \times L^\infty(\mathbb{R}^n) \times L^\infty(\mathbb{R}^m)$. Since ζ_{k+1} and ψ_{k+1} are linear and u_{k+1} is constant on each subinterval $[t_{k+1}^i, t_{k+1}^{i+1})$, this amounts to estimating the expression

$$\begin{aligned} & \nabla_\psi H(x_k(t)) - \nabla_\psi H(x_k(t_{k+1}^i)) \\ & + \nabla_{\psi x}^2 H(x_k(t))(x_{k+1}(t) - x_k(t)) \\ & - \nabla_{\psi x}^2 H(x_k(t_{k+1}^i))(x_{k+1}(t_{k+1}^i) - x_k(t_{k+1}^i)) \end{aligned}$$

and also similar expressions coming from the second and the third row of the mapping in (61). The iteration x_k is either the initial one ($k=0$) in which case ζ_k and ψ_k satisfy the Euler discretization of (57), or they satisfy the first and the second equations in (61). We have

$$\begin{aligned} & \|\nabla_\psi H(x_k(t)) - \nabla_\psi H(x_k(t_{k+1}^i))\| \\ & + \|\nabla_{\psi x}^2 H(x_k(t))(x_{k+1}(t) - x_k(t)) - \nabla_{\psi x}^2 H(x_k(t_{k+1}^i))(x_{k+1}(t_{k+1}^i) - x_k(t_{k+1}^i))\| \\ & \leq \|\nabla_\psi H(x_k(t)) - \nabla_\psi H(x_k(t_{k+1}^i))\| \\ & + \|\nabla_{\psi x}^2 H(x_k(t)) - \nabla_{\psi x}^2 H(x_k(t_{k+1}^i))\| \|x_{k+1}(t) - x_k(t)\| \\ & + \|\nabla_{\psi x}^2 H(x_k(t_{k+1}^i))\| \|x_{k+1}(t) - x_k(t) - x_{k+1}(t_{k+1}^i) + x_k(t_{k+1}^i)\| \end{aligned}$$

Noting that both $x_{k+1}(t) - x_k(t)$ and $\nabla_{\psi x}^2 H(x_k(t_{k+1}^i))$ are uniformly bounded, all jars down to estimating the expression

$$\|x_{k+1}(t) - x_{k+1}(t_{k+1}^i)\| + \|-x_k(t) + x_k(t_{k+1}^i)\|.$$

The function u_k , being in the ball with radius a around \bar{u} in $L^\infty(\mathbb{R}^m)$, is bounded (uniformly in k). Thus, for an appropriate constant C_1 in both cases $|\zeta_k^{i+1} - \zeta_k^i| \leq C_1 h_k$. Hence,

$$|\zeta_k(t) - \zeta_k(t_{k+1}^i)| \leq C_1 h_{k+1} \quad \text{for } t \in [t_{k+1}^i, t_{k+1}^{i+1}).$$

The same applies also for ψ . For u we have $u_k(t) - u_k(t_{k+1}^i) = 0$ due to the condition that consequent meshes are embedded. The same argument applies also to $x_{k+1}(t) - x_k(t_{k+1}^i)$. Hence, $\|p_k\| \leq C_2 h_{k+1}$ for an appropriate constant C_2 . By choosing \bar{N} sufficiently large we can ensure that $\|p_k\|$ is small enough for $k > 1/\bar{N}$, thus Theorem 5.2 applied with $\theta = 1/\nu$ gives us the desired result. \square

Theorem 5.3 can be interpreted as a kind of *mesh independence* result, saying roughly that the sequences of the exact and the discretized Newton's iterates behave in a similar way, independently on the discretization. For more on this topic but in a different context involving strong metric regularity, see [5]. For a recent study of discrete approximations for numerical solution of optimal control problems, see [10].

Finally, we note that we are not aware of any conditions for metric regularity of the mapping in the optimality system (57) that do not imply automatically strong metric regularity. In our opinion, finding such a condition, or showing that there are no such conditions for standard problems, is an important problem for future research.

References

- [1] A. V. DMITRUK, A. A. MILYUTIN AND N. P. OSMOLOVSKIĬ, *Lyusternik's theorem and the theory of extremum*, Uspekhi Mat. Nauk **35** no. 6 (216) (1980) 11–46 (Russian).
- [2] A. V. DMITRUK, A. Y. KRUGER, Extensions of metric regularity, Optimization **58** (2009) 561–584.
- [3] A. L. DONTCHEV, Local convergence of the Newton method for generalized equations, C. R. Acad. Sci. Paris Sér. I Math. **322** (1996) 327–331.
- [4] A. L. DONTCHEV, W. W. HAGER, An inverse mapping theorem for set-valued maps, Proc. Amer. Math. Soc. **121** (1994) 481–489.
- [5] 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.
- [6] A. L. DONTCHEV, R. T. ROCKAFELLAR, Newton's method for generalized equations: a sequential implicit function theorem, Mathematical Programming **123** (2010) 139–159.
- [7] A. L. DONTCHEV, R. T. ROCKAFELLAR, *Implicit Functions and Solution Mappings*, Springer Mathematics Monographs, Springer, Dordrecht, 2009.
- [8] L. M. GRAVES, Some mapping theorems, Duke Mathematical Journal, **17** (1950) 111–114.
- [9] A. D. IOFFE, Metric regularity and subdifferential calculus, Uspekhi Mat. Nauk **55** (2000) 103–162 (Russian).
- [10] C. Y. KAYA, Inexact restoration for Runge-Kutte discretization of optimal control problems, SIAM J. Numer. Anal. **48** (2010) 1492–1517.

- [11] D. KLATTE, B. KUMMER, Stability of inclusions: characterizations via suitable Lipschitz functions and algorithms, *Optimization* **55** (2006) 627–660.
- [12] C. T. KELLEY, Solving nonlinear equations with Newton’s method, *Fundamentals of Algorithms*, SIAM, Philadelphia, PA, 2003.
- [13] L. A. LYUSTERNIK, On the conditional extrema of functionals, *Mat. Sbornik* **41** (1934) 390–401.
- [14] S. M. ROBINSON, Strongly regular generalized equations, *Math. of Oper. Research* **5** (1980) 43–62.