

A. L. Dontchev · R. T. Rockafellar

Newton's method for generalized equations: a sequential implicit function theorem

Received: date / Accepted: date

Abstract In an extension of Newton's method to generalized equations, we carry further the implicit function theorem paradigm and place it in the framework of a mapping acting from the parameter and the starting point to the set of all associated sequences of Newton's iterates as elements of a sequence space. An inverse function version of this result shows that the strong regularity of the mapping associated with the Newton sequences is equivalent to the strong regularity of the generalized equation mapping.

Keywords Newton's method, generalized equations, variational inequalities, strong regularity, implicit function theorems, inverse function theorems, perturbations, variational analysis.

Mathematics Subject Classification (2000) 49J53, 49K40, 65J15, 90C31

This work was supported by National Science Foundation.

A. L. Dontchev

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.
E-mail: adontche@nsf.gov

R. T. Rockafellar

Dept. of Mathematics, University of Washington, Seattle, WA 98195-4350.
E-mail: rtr@math.washington.edu

1 Introduction and Background

Newton's method for solving a nonlinear equation $f(x) = 0$ for a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ focuses on successive linearization; if started near enough to a solution \bar{x} around which f is continuously differentiable with invertible derivative, it converges to \bar{x} . However, the convergence analysis depends on looking at more than just the single equation $f(x) = 0$. It is closely connected with the applicability of the inverse function theorem to f with respect to the pair $(\bar{x}, 0)$ in its graph and thus to the theoretical potential for solving $f(x) = y$ for x as a function of y in a local sense around $(\bar{x}, 0)$.

From this insight, it is easy to pass to a broader setting in which $f(p, x) = 0$ is to be solved in relation to a parameter $p \in \mathbb{R}^m$. Newton's method can be studied not only in terms of determining a solution \bar{x} for a particular instance \bar{p} of p , but also in assessing how the convergence may behave with respect to shifts in p . That puts the analysis in the context of the implicit function theorem as applied to f at a pair (\bar{p}, \bar{x}) satisfying $f(\bar{p}, \bar{x}) = 0$.

In this paper the framework is broader still. We work with Banach spaces P for p , X for x , and Y for the range space of f . We consider so-called *generalized equations* of the form

$$(1.1) \quad f(p, x) + F(x) \ni 0, \text{ or equivalently } -f(p, x) \in F(x),$$

for a function $f : P \times X \rightarrow Y$ and a mapping F from X to Y that may be *set-valued*, which we indicate by writing $F : X \rightrightarrows Y$. As is well known, the model of a generalized equation (1.1) covers a huge territory. The classical case of an equation corresponds to having $F(x) \equiv 0$, whereas by taking $F(x) \equiv -K$ for a fixed set $K \subset Y$ one gets various 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 $C \subset X$, one has a variational inequality.

We assume throughout that f is continuously Fréchet differentiable in x with derivative denoted by $D_x f(p, x)$, and that both $f(p, x)$ and $D_x f(p, x)$ depend continuously on (p, x) . 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 $F^{-1}(y) = \{x \mid y \in F(x)\}$. The norms in P , X and Y are all denoted by $\|\cdot\|$. The closed ball centered at x with radius r is symbolized by $B_r(x)$.

The extension of Newton's method to the generalized equation (1.1) operates with successive linearizations of f in the x argument while leaving F untouched. The exact formulation will be given in the next section. Our aim is to investigate its parametric properties with respect to p , and to do so moreover in terms of spaces of *sequences of iterates*. This investigation will rely heavily on extensions of the implicit function theorem that have been developed for (1.1).

The *solution mapping* associated with the generalized equation (1.1) is the potentially set-valued mapping $S : P \rightrightarrows X$ defined by

$$(1.2) \quad S : p \mapsto \{x \mid f(p, x) + F(x) \ni 0\}.$$

The implicit function paradigm asks the same questions about solutions in this case as in the classical case where $F = 0$, but for a clear expression it is

useful to have the concept of a *graphical localization* of S at \bar{p} for \bar{x} , where $\bar{x} \in S(\bar{p})$. By this we mean a set-valued mapping with its graph in $P \times X$ having the form $(Q \times U) \cap \text{gph } S$ for some neighborhoods Q of \bar{p} and V of \bar{x} . The localization is *single-valued* when this mapping reduces to a function from Q into U . If it is not only single-valued but Lipschitz continuous on Q , we speak of a *Lipschitz localization*. Beyond that, one can look for continuous differentiability, and so forth. The issue is what kind of localization properties can be deduced from assumptions on f and F in connection with a pair (\bar{p}, \bar{x}) in the graph of S , or in other words a particular solution \bar{x} to the generalized equation (1.1) corresponding to a choice of \bar{p} of p .

Robinson [12] gave an answer to this question for the solution mapping S in the case of a parameterized variational inequality, where $Y = X^*$ and F is the normal cone mapping N_C to a convex closed set C in X . His main discovery was the fact that the implicit function theorem paradigm works for this general model, albeit in a somewhat weaker sense: if the solution mapping of a “partial linearization” of (1.1) has a Lipschitz localization, then the solution mapping S in (1.2) also possesses this property. Lipschitz continuity is the most one can get in the general framework of (1.1), because elementary one-dimensional examples of variational inequalities already exhibit single-valued solution mappings that are not differentiable, but nevertheless are Lipschitz continuous. This puts Lipschitz continuity in center stage for exploring properties of solutions mappings S of the form (1.2). Much more on this circle of ideas will be presented in the forthcoming authors’ book [7].

Enabled by Robinson’s result, Josephy succeeded in his thesis [9] to be the first to show convergence of the extended Newton’s method for solving a variational inequality. In Section 2 we prove quadratic convergence of Newton’s method, for solving the parametric generalized equation (1.1), which is *locally uniform* in the parameter p . We should point out that here we deal with the weak mode of convergence of Newton’s method, where we assume the existence of a solution and then show that when the method is started close enough to this solution, we have convergence. Another way to consider convergence is in line with the classical Kantorovich theorem, where no assumptions for the existence of a solution are made but conditions are imposed on the starting point of the iteration. A Kantorovich-type theorem for generalized equations *without parameter dependence* is obtained in [2], which generalizes a previous result in [13]; extending this result to parametric generalized equations of the form (1.1) is a subject for future research. A wide overview of Newton’s method for nonsmooth equations and generalized equations is available in the book by Klatte and Kummer [10], see also the more recent survey [11].

In this paper our efforts will be concentrated on the role of the parameter p in generating sequences by Newton’s iteration that approach a solution of (1.1). In Section 2 we focus on the extent to which quadratic convergence of Newton’s method takes place which is locally uniform in the parameter. In Section 3 we go further by treating Newton’s method in terms of an implicitly defined mapping involving sequences of iterates as elements of a sequence space, and obtain a result which is strikingly similar to the corresponding implicit function theorem for generalized equations (Theorem 1.1 below). An

inverse function version of this result is also obtained. Here we build on a paper of Dontchev [3], where this idea originated. Some other results in this direction were subsequently obtained in [8]. Although oriented differently, our analysis here is related to previous studies of mesh independence of Newton's method when applied to discretized optimal control problems, see [1] and [4]. There is other work in the literature concerned with parametric versions of Newton's method; here, in contrast, we approach the dependence on a parameter in a quantitative way, as it is dealt with in the implicit function theorem.

The rest of this section introduces background facts and notation. We begin with quantitative measures for Lipschitz continuity and partial Lipschitz continuity in a neighborhood, both of which will have an essential role.

Lipschitz modulus. 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

$$(1.3) \quad \|f(x') - f(x)\| \leq \kappa \|x' - x\| \quad \text{for all } x', x \in D.$$

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 (1.3) 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 following two results can be extracted from [6] (or from [7] when this book becomes available).

Theorem 1.1 (implicit function theorem). For a generalized equation (1.1) and its solution mapping S in (1.2), let \bar{p} and \bar{x} be such that $\bar{x} \in S(\bar{p})$. Assume that f is Lipschitz continuous with respect to p uniformly in x at (\bar{p}, \bar{x}) , that is,

$$\widehat{\text{lip}}_p(f; (\bar{p}, \bar{x})) < \infty,$$

and that the inverse G^{-1} of the mapping

$$(1.4) \quad G(x) = f(\bar{p}, \bar{x}) + D_x f(\bar{p}, \bar{x})(x - \bar{x}) + F(x) \quad \text{for which } G(\bar{x}) \ni 0,$$

has a Lipschitz localization σ at 0 for \bar{x} .

Then the mapping S has a Lipschitz localization s at \bar{p} for \bar{x} with

$$(1.5) \quad \text{lip}(s; \bar{p}) \leq \text{lip}(\sigma; 0) \cdot \widehat{\text{lip}}_p(f; (\bar{p}, \bar{x})).$$

The property is assumed in Theorem 1.1 for the mapping F in the context of a variational inequality was called by Robinson [12] *strong regularity*. We give this concept a broader meaning.

Strong regularity. A mapping $T : X \rightrightarrows Y$ with $(\bar{x}, \bar{y}) \in \text{gph } T$ will be called *strongly regular at \bar{x} for \bar{y}* if its inverse T^{-1} has a Lipschitz localization at \bar{y} for \bar{x} .

In the case of the generalized equation (1.1) with $P = Y$ and $f(p, x) = g(x) - p$ for a function $g : X \rightarrow Y$, so that

$$(1.6) \quad S(p) = \{x \mid p \in g(x) + F(x)\} = (g + F)^{-1}(p),$$

the property of both the mapping G in Theorem 1.1(b) as well as the mapping $g + F$ translates as strong regularity and the inverse function version of Theorem 1.1 has the following symmetric form.

Theorem 1.2 (inverse version). *In the framework of the solution mapping (1.6), consider any pair (\bar{p}, \bar{x}) with $\bar{x} \in S(\bar{p})$. Then the mapping $g + F$ is strongly regular at \bar{x} for \bar{p} if and only if its partial linearization $x \mapsto G(x) = g(\bar{x}) + Dg(\bar{x})(x - \bar{x}) + F(x)$ is strongly regular at \bar{p} for \bar{x} . In addition, if s and σ are the associated Lipschitz localizations of $(g + F)^{-1}$ and G^{-1} respectively, then*

$$\text{lip}(s - \sigma; \bar{p}) = 0.$$

This implies in particular that $\text{lip}(s; \bar{p}) = \text{lip}(\sigma; \bar{p})$.

Throughout the paper we repeatedly use the contraction mapping principle in the following form:

Theorem 1.3 (contraction mapping principle). *Let X be a complete metric space with metric ρ . Consider a point $\bar{x} \in X$ and a function $\Phi : X \rightarrow X$ for which there exist scalars $a > 0$ and $\lambda \in [0, 1)$ such that:*

- (a) $\rho(\Phi(\bar{x}), \bar{x}) \leq a(1 - \lambda)$;
- (b) $\rho(\Phi(x'), \Phi(x)) \leq \lambda\rho(x', x)$ for every $x', x \in B_a(\bar{x})$.

Then there is a unique $x \in B_a(\bar{x})$ satisfying $x = \Phi(x)$, that is, Φ has a unique fixed point in $B_a(\bar{x})$.

Behind Theorems 1.1 and 1.2 is a more general fact about the stability of strong regularity under perturbations. Since this fact was never stated in the way we need it in the present paper, we supply it with a proof, for completeness.

Theorem 1.4 (stability of strong regularity under perturbation). *Consider a mapping $T : X \rightrightarrows Y$ and any $(\bar{x}, \bar{y}) \in \text{gph } T$ such that, for a positive constant κ and neighborhoods U of \bar{x} and V of \bar{y} , the mapping $y \mapsto T^{-1}(y) \cap U$ is a Lipschitz continuous function on V with Lipschitz constant κ . Then for every positive constant μ with $\kappa\mu < 1$ there exist neighborhoods $U' \subset U$*

of \bar{x} and $V' \subset V$ of \bar{y} such that for every function $h : X \rightarrow Y$ which is Lipschitz continuous on U with Lipschitz constant μ , the mapping $y \mapsto (h+T)^{-1}(y) \cap U'$ is a Lipschitz continuous function on $h(\bar{x})+V'$ with Lipschitz constant $\kappa/(1-\kappa\mu)$.

Proof. By assumption, for the function $s(y) = T^{-1}(y) \cap U$ for $y \in V$ we have

$$(1.7) \quad \|s(y') - s(y)\| \leq \kappa \|y' - y\| \quad \text{for all } y', y \in V.$$

Pick $\mu > 0$ such that $\kappa\mu < 1$ and then choose positive constants a and b such that

$$(1.8) \quad \mathcal{B}_a(\bar{x}) \subset U, \quad \mathcal{B}_{b+\mu a}(\bar{y}) \subset V \quad \text{and} \quad b < a(1-\kappa\mu)/\kappa.$$

Choose any function $h : X \rightarrow Y$ such that

$$(1.9) \quad \|h(x') - h(x)\| \leq \mu \|x' - x\| \quad \text{for all } x', x \in U.$$

For any $y \in \mathcal{B}_b(h(\bar{x}) + \bar{y})$ and any $x \in \mathcal{B}_a(\bar{x})$ we have

$$\| -h(x) + y - \bar{y} \| \leq \|y - h(\bar{x}) - \bar{y}\| + \|h(x) - h(\bar{x})\| \leq b + \mu a,$$

and hence, by (1.8), $-h(x) + y \in V \subset \text{dom } s$. Fix $y \in \mathcal{B}_b(h(\bar{x}) + \bar{y})$ and consider the mapping

$$\Phi_y : x \mapsto s(-h(x) + y) \quad \text{for } x \in \mathcal{B}_a(\bar{x}).$$

Then, by using (1.7), (1.8) and (1.9) we get

$$\|\bar{x} - \Phi_y(\bar{x})\| = \|s(\bar{y}) - s(y - h(\bar{x}))\| \leq \kappa \|y - \bar{y} + h(\bar{x})\| \leq \kappa b < a(1-\kappa\mu).$$

Moreover, for any $v, v' \in \mathcal{B}_a(\bar{x})$,

$$\|\Phi_y(v) - \Phi_y(v')\| = \|s(y - h(v)) - s(y - h(v'))\| \leq \kappa \|h(v) - h(v')\| \leq \kappa\mu \|v - v'\|.$$

Thus, by the contraction mapping principle, there exists a fixed point $x = \Phi_y(x)$ in $\mathcal{B}_a(\bar{x})$, and there is no more than one such fixed point in $\mathcal{B}_a(\bar{x})$. The mapping from $y \in \mathcal{B}_b(h(\bar{x}) + \bar{y})$ to the unique fixed point $x(y)$ of Φ_y in $\mathcal{B}_a(\bar{x})$ is a function which satisfies $x(y) = s(y - h(x(y)))$; therefore, for any $y, y' \in \mathcal{B}_b(h(\bar{x}) + \bar{y})$ we have

$$\begin{aligned} \|x(y) - x(y')\| &= \|s(y - h(x(y))) - s(y' - h(x(y')))\| \\ &\leq \kappa \|y - y'\| + \kappa \|h(x(y)) - h(x(y'))\| \\ &\leq \kappa \|y - y'\| + \kappa\mu \|x(y) - x(y')\|. \end{aligned}$$

Hence,

$$\|x(y) - x(y')\| \leq \frac{\kappa}{1-\mu\kappa} \|y - y'\|.$$

Choosing $U' = \mathcal{B}_a(\bar{x})$ and $V' = \mathcal{B}_b(\bar{y})$, and noting that $\mathcal{B}_b(g(\bar{x}) + \bar{y}) = g(\bar{x}) + \mathcal{B}_b(\bar{y})$, we complete the proof. \square

In the paper we utilize the following important corollary of Theorem 1.4:

Corollary 1.5. *For a mapping $F : X \rightrightarrows Y$ and a point $(\bar{x}, \bar{y}) \in \text{gph } F$, let F be strongly regular at \bar{x} for \bar{y} with associated Lipschitz localization s of F^{-1} at \bar{y} for \bar{x} . Consider also a function $r : P \times X \rightarrow Y$ such that $r(p, \bar{x})$ is continuous at \bar{p} and*

$$\text{lip}(s; \bar{y}) \cdot \widehat{\text{lip}}_x(r; (\bar{p}, \bar{x})) < 1.$$

Then for each

$$(1.10) \quad \gamma > \frac{\text{lip}(s; \bar{y})}{1 - \text{lip}(s; \bar{y}) \cdot \widehat{\text{lip}}_x(r; (\bar{p}, \bar{x}))}$$

there exist neighborhoods U of \bar{x} , V of \bar{y} and Q of \bar{p} such that for every $p \in Q$ the mapping $y \mapsto (r(p, \cdot) + F)^{-1}(y) \cap U$ is a Lipschitz continuous function on $r(\bar{p}, \bar{x}) + V$ with a Lipschitz constant γ .

Proof. Pick γ as in (1.10) and then $\kappa > \text{lip}(s; \bar{y})$ and $\mu > \widehat{\text{lip}}_x(r; (\bar{p}, \bar{x}))$ such that $\kappa\mu < 1$ and $\kappa/(1 - \kappa\mu) \leq \gamma$. Choose neighborhoods U' of \bar{x} and Q of \bar{p} such that s is Lipschitz continuous on U' with Lipschitz constant κ and for each $p \in Q$ the function $r(p, \cdot)$ is Lipschitz continuous on U' with Lipschitz constant μ . Applying Theorem 1.4, we obtain neighborhoods $U \subset U'$ of \bar{x} , and V' of \bar{y} such that for every $p \in Q$ the mapping $y \mapsto (r(p, \cdot) + F)^{-1}(y) \cap U$ is a Lipschitz continuous function on $r(p, \bar{x}) + V'$ with Lipschitz constant γ . By making Q small enough we can find a neighborhood V of \bar{y} such that $r(\bar{p}, \bar{x}) + V \subset r(p, \bar{x}) + V'$ for every $p \in Q$. \square

2 Newton's method and its convergence

As already mentioned, the version of Newton's method we consider is based on partial linearization, in which we linearize f with respect to the variable x at the current point but leave F intact.

Newton's method (in extended form). *With the aim of approximating a solution to the generalized equation (1.1) for a fixed value of the parameter p , choose a starting point x_0 and generate a sequence $\{x_k\}_{k=0}^\infty$ iteratively for $k = 0, 1, \dots$, by taking x_{k+1} to be a solution to the auxiliary generalized equation*

$$(2.1) \quad f_k(p, x_{k+1}) + F(x_{k+1}) \ni 0,$$

where

$$f_k(p, x) = f(p, x_k) + D_x f(p, x_k)(x - x_k).$$

The iteration (2.1) reduces to the standard Newton's method for solving a nonlinear equation when F is the zero mapping. If F is the normal cone mapping appearing in the first-order optimality system for a nonlinear programming problem, (2.1) becomes the popular sequential quadratic programming (SQP) method for numerical optimization.

We will study this method by reconceiving the iteration rule (2.1) as a condition that defines an element of the Banach space $l_\infty(X)$, consisting of all infinite sequences

$$\xi = \{x_1, x_2, \dots, x_k, \dots\} \text{ with elements } x_k \in X.$$

The norm on $l_\infty(X)$ is

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

Define a mapping $\Xi : X \times P \rightrightarrows l_\infty(X)$ by (2.2)

$$\Xi : (u, p) \mapsto \left\{ \xi \in l_\infty(X) \mid \bigcap_{k=0}^{\infty} (f(p, x_k) + D_x f(p, x_k)(x_{k+1} - x_k) + F(x_{k+1})) \ni 0 \text{ with } x_0 = u \right\},$$

whose value for a given (u, p) is the set of all sequences $\{x_k\}_{k=1}^{\infty}$ generated by Newton's iteration (2.1) for p that start from u . If \bar{x} is a solution of (1.1) for \bar{p} , the

$$\text{the constant sequence } \bar{\xi} = \{\bar{x}, \bar{x}, \dots, \bar{x}, \dots\} \text{ satisfies } \bar{\xi} \in \Xi(\bar{x}, \bar{p}).$$

Theorem 2.1 (uniform convergence of Newton's iteration). *In the framework of the generalized equation (1.1) with solution mapping S in (1.2), let $\bar{x} \in S(\bar{p})$. Assume that*

$$\widehat{\text{lip}}_p(f; (\bar{p}, \bar{x})) + \widehat{\text{lip}}_x(D_x f; (\bar{p}, \bar{x})) < \infty$$

and let the mapping G in (1.4) be strongly regular at \bar{x} for 0 with associated Lipschitz localization σ of the inverse G^{-1} at 0 for \bar{x} . Then for every

$$(2.3) \quad \gamma > \frac{1}{2} \text{lip}(\sigma; 0) \cdot \widehat{\text{lip}}_x(D_x f; (\bar{p}, \bar{x}))$$

there exist neighborhoods Q of \bar{p} and U of \bar{x} such that for every $p \in Q$ and $u \in U$ there is exactly one sequence $\xi(u, p)$ with components x_1, \dots, x_k, \dots , all belonging to U and generated by Newton's iteration (2.1) starting from u for the value p of the parameter. This sequence is convergent to the value $s(p)$ of the Lipschitz localization s of the solution mapping S at \bar{p} for \bar{x} whose existence is claimed in Theorem 1.1; moreover, the convergence is quadratic with constant γ , that is,

$$(2.4) \quad \|x_{k+1} - s(p)\| \leq \gamma \|x_k - s(p)\|^2 \quad \text{for all } k \geq 0.$$

In other words, the mapping Ξ in (2.2) has a single-valued graphical localization ξ at (\bar{x}, \bar{p}) for $\bar{\xi}$; moreover, for u close to \bar{x} and p close to \bar{p} the value $\xi(u, p)$ of this localization is a sequence which converges quadratically to the associated solution $s(p)$ for p in the sense of (2.4).

Proof. Choose γ as in (2.3) and then $\kappa > \text{lip}(\sigma; 0)$ and $\mu > \widehat{\text{lip}}_x(D_x f; (\bar{p}, \bar{x}))$ such that $\kappa\mu < 2\gamma$. Next, choose $\varepsilon > 0$ so that $\kappa\varepsilon < 1$ and moreover

$$(2.5) \quad \frac{\kappa\mu}{2(1 - \kappa\varepsilon)} \leq \gamma.$$

The assumed strong regularity of the mapping G in (1.4) at \bar{x} for 0 and the choice of κ mean that there exist positive constants α' and b' such that the mapping $y \mapsto \sigma(y) = G^{-1}(y) \cap \mathcal{B}_{\alpha'}(\bar{x})$ is a Lipschitz continuous function on $\mathcal{B}_{b'}(0)$ with Lipschitz constant κ . Along with the mapping G in (1.4) consider the parameterized mapping

$$x \mapsto G_{p,w}(x) = f(p, w) + D_x f(p, w)(x - w) + F(x).$$

Note that $G_{p,w}(x) = r(p, w; x) + G(x)$, where the function

$$x \mapsto r(p, w; x) = f(p, w) + D_x f(p, w)(x - w) - f(\bar{p}, \bar{x}) - D_x f(\bar{p}, \bar{x})(x - \bar{x})$$

is affine, and hence Lipschitz continuous, with Lipschitz constant

$$\eta(p, w) = \|D_x f(p, w) - D_x f(\bar{p}, \bar{x})\|.$$

Now, let κ' be such that $\kappa > \kappa' > \text{lip}(\sigma; 0)$ and let $\chi > 0$ satisfy

$$\chi\kappa' < 1 \quad \text{and} \quad \frac{\kappa'}{1 - \chi\kappa'} < \kappa.$$

Applying Corollary 1.5 to the mapping $G_{p,w}$ and noting that $r(\bar{p}, \bar{x}, \bar{x}) = 0$, we obtain that there are positive constants $\alpha \leq \alpha'$ and $b \leq b'$ such that for p and w satisfying $\eta(p, w) \leq \chi$ the mapping $y \mapsto G_{p,w}^{-1}(y) \cap \mathcal{B}_\alpha(\bar{x})$ is a Lipschitz continuous function on $\mathcal{B}_b(0)$ with Lipschitz constant κ ; we denote this function by $\Theta(p, w; \cdot)$.

Since D_x is continuous near (\bar{p}, \bar{x}) , there exist positive constants c and a such that $\eta(p, w) \leq \chi$ whenever $p \in \mathcal{B}_c(\bar{p})$ and $w \in \mathcal{B}_a(\bar{x})$. Make a and c smaller if necessary so that $a \leq \alpha$ and moreover

$$(2.6) \quad \|D_x f(p, x) - D_x f(p, x')\| \leq \mu \|x - x'\|$$

for $x, x' \in \mathcal{B}_a(\bar{x})$ and $p \in \mathcal{B}_c(\bar{p})$. The assumptions of Theorem 1.1 hold, hence, we can apply it and further adjust a and c so that the truncation $S(p) \cap \mathcal{B}_a(\bar{x})$ of the solution mapping S in (1.2) is a function s which is Lipschitz continuous on $\mathcal{B}_c(\bar{p})$ (with Lipschitz constant some $\lambda > \text{lip}(\sigma; 0) \cdot \|D_p f(\bar{p}, \bar{x})\|$). Next, take a even smaller if necessary so that

$$(2.7) \quad \frac{27}{8}a^2 < b, \quad \frac{3}{2}\mu a \leq \varepsilon, \quad \frac{1}{2}\kappa\mu a < 1 - \kappa\varepsilon \quad \text{and} \quad \frac{9}{2}\gamma a \leq 1.$$

The first and the third inequality in (2.7) allow us to choose $\delta > 0$ satisfying

$$(2.8) \quad \delta + \frac{1}{8}\mu a^2 \leq b \quad \text{and} \quad \kappa\delta + \frac{1}{2}\kappa\mu a^2 \leq a(1 - \kappa\varepsilon).$$

Then make c even smaller if necessary so that

$$(2.9) \quad \|s(p) - \bar{x}\| \leq a/2 \quad \text{and} \quad \|f(p, \bar{x}) - f(\bar{p}, \bar{x})\| \leq \delta \quad \text{for } p \in \mathcal{B}_c(\bar{p}).$$

Summarizing, we have found constants a, b and c such that for each $p \in \mathcal{B}_c(\bar{p})$ and $w \in \mathcal{B}_a(\bar{x})$ the function $\Theta(p, w, \cdot)$ is Lipschitz continuous on $\mathcal{B}_b(0)$ with constant κ , and also the conditions (2.6)–(2.9) are satisfied.

We frequently use an estimate for smooth functions obtained by simple calculus. From the standard equality

$$f(p, y) - f(p, v) = \int_0^1 D_x f(p, v + t(y - v))(y - v) dt,$$

which yields

$$\begin{aligned} & \|f(p, y) - f(p, v) - D_x f(p, v)(y - v)\| \\ &= \left\| \int_0^1 D_x f(p, v + t(y - v))(y - v) dt - D_x f(p, v)(y - v) \right\| \\ &\leq \mu \int_0^1 t dt \|y - v\|^2, \end{aligned}$$

we have from (2.6) that for all $y, v \in \mathcal{B}_a(\bar{x})$ and $p \in \mathcal{B}_c(\bar{p})$

$$(2.10) \quad \|f(p, y) - f(p, v) - D_x f(p, v)(y - v)\| \leq \frac{1}{2} \mu \|y - v\|^2.$$

Fix $p \in \mathcal{B}_c(\bar{p})$ and $w \in \mathcal{B}_a(\bar{x})$ and consider the function

$$(2.11) \quad x \mapsto g(p, w; x) := -f(p, w) - D_x f(p, w)(x - w) + f(p, s(p)) + D_x f(p, s(p))(x - s(p)).$$

Recall that here $s(p) = S(p) \cap \mathcal{B}_{a/2}(\bar{x})$ for all $p \in \mathcal{B}_c(\bar{p})$. For any $x \in \mathcal{B}_a(\bar{x})$, using (2.6) and (2.10), we have

$$\begin{aligned} \|g(p, w; x)\| &\leq \|f(p, s(p)) - f(p, w) - D_x f(p, w)(s(p) - w)\| \\ &\quad + \|(D_x f(p, w) - D_x f(p, s(p)))(x - s(p))\| \\ &\leq \frac{1}{2} \mu \|w - s(p)\|^2 + \mu \|w - s(p)\| \|x - s(p)\| \leq \frac{27}{8} \mu a^2. \end{aligned}$$

Then, from the second inequality in (2.7),

$$(2.12) \quad \|g(p, w; x)\| \leq b.$$

Using (2.9), (2.10), and the first inequality in (2.8) we come to

$$(2.13) \quad \begin{aligned} & \|f(\bar{p}, \bar{x}) - f(p, s(p)) - D_x f(p, s(p))(\bar{x} - s(p))\| \\ &\leq \|f(\bar{p}, \bar{x}) - f(p, \bar{x})\| \\ &\quad + \|f(p, \bar{x}) - f(p, s(p)) - D_x f(p, s(p))(\bar{x} - s(p))\| \\ &\leq \delta + \frac{1}{2} \mu \|s(p) - \bar{x}\|^2 \leq \delta + \frac{1}{8} \mu a^2 \leq b. \end{aligned}$$

Hence, remembering that $p \in \mathcal{B}_c(\bar{p})$ and $s(p) \in \mathcal{B}_a(\bar{x})$, both

$$g(p, w; x) \quad \text{and} \quad f(\bar{p}, \bar{x}) - f(p, s(p)) - D_x f(p, s(p))(\bar{x} - s(p))$$

are in the domain of $\Theta(p, s(p); \cdot)$ where this function is Lipschitz continuous with Lipschitz constant κ .

We now choose $p \in \mathcal{B}_c(\bar{p})$ and $u \in \mathcal{B}_a(\bar{x})$, and construct a sequence $\xi(u, p)$ generated by Newton's iteration (2.1) starting from u for the value p of the parameter, whose existence, uniqueness and quadratic convergence is claimed in the statement of the theorem.

If $u = s(p)$ there is nothing to prove, so assume $u \neq s(p)$. Our first step is to show that, for the function g defined in (2.11), the mapping

$$\Phi_0 : x \mapsto \Theta(p, s(p); g(p, u; x)),$$

has a unique fixed point in $\mathcal{B}_a(\bar{x})$. Using the equality

$$\bar{x} = \Theta(p, s(p); -f(\bar{p}, \bar{x}) + f(p, s(p)) + D_x f(p, s(p))(\bar{x} - s(p))),$$

(2.12), (2.13) and the Lipschitz continuity of $\Theta(p, s(p); \cdot)$ in $\mathcal{B}_b(0)$ with constant κ , and then the second inequality in (2.9), (2.10) and the second inequality in (2.8), we have

$$\begin{aligned} & \|\bar{x} - \Phi_0(\Theta(p, s(p); g(p, u; \bar{x})))\| \\ &= \|\Theta(p, s(p); -f(\bar{p}, \bar{x}) + f(p, s(p)) + D_x f(p, s(p))(\bar{x} - s(p))) \\ &\quad - \Theta(p, s(p); g(p, u; \bar{x}))\| \\ (2.14) \quad &\leq \kappa \left\| -f(\bar{p}, \bar{x}) + f(p, s(p)) + D_x f(p, s(p))(\bar{x} - s(p)) \right. \\ &\quad \left. - [-f(p, u) - D_x f(p, u)(\bar{x} - u) + f(p, s(p)) + D_x f(p, s(p))(\bar{x} - s(p))] \right\| \\ &= \kappa \left\| -f(\bar{p}, \bar{x}) + f(p, u) + D_x f(p, u)(\bar{x} - u) \right\| \\ &\leq \kappa \left\| -f(\bar{p}, \bar{x}) + f(p, \bar{x}) \right\| + \\ &\quad \kappa \left\| f(p, u) - f(p, \bar{x}) - D_x f(p, u)(u - \bar{x}) \right\| \\ &\leq \kappa \delta + \frac{1}{2} \kappa \mu \|u - \bar{x}\|^2 \leq \kappa \delta + \frac{1}{2} \kappa \mu a^2 \leq a(1 - \kappa \varepsilon). \end{aligned}$$

Further, for any $v, v' \in \mathcal{B}_a(\bar{x})$, by (2.12), the Lipschitz continuity of $\Theta(p, s(p), \cdot)$, (2.6), and the second inequality in (2.7), we obtain

$$\begin{aligned} & \|\Phi_0(v) - \Phi_0(v')\| = \|\Theta(p, s(p); g(p, u; v)) - \Theta(p, s(p); g(p, u; v'))\| \\ (2.15) \quad &\leq \kappa \|g(p, u; v) - g(p, u; v')\| = \\ &\quad \kappa \|(-D_x f(p, u) + D_x f(p, s(p)))(v - v')\| \\ &\leq \kappa \mu \|u - s(p)\| \|v - v'\| \leq \frac{3}{2} a \kappa \mu \|v - v'\| \leq \kappa \varepsilon \|v - v'\|. \end{aligned}$$

Hence, there exists a fixed point $x_1 \in \Phi_0(x_1) \cap \mathcal{B}_a(\bar{x})$ which translates to $g(p, u; x_1) \in G_{p, u}(x_1)$, that is, $x_1 = \Theta(p, u; 0) \in G_{p, u}^{-1}(0)$ or, equivalently,

$$0 \in f(p, u) + D_x f(p, u)(x_1 - u) + F(x_1),$$

meaning that x_1 is obtained by the Newton iteration (2.1) from u for p , and there is no more than just one such iterate in $\mathcal{B}_a(\bar{x})$.

Now we will show that x_1 satisfies a tighter estimate. Let

$$\omega_0 = \gamma \|u - s(p)\|^2.$$

Then $\omega_0 > 0$ and, by the last inequality in (2.7), $\omega_0 \leq \gamma(a + a/2)^2 \leq a/2$. We apply again the contraction mapping principle to the mapping Φ_0 but now on $\mathcal{B}_{\omega_0}(s(p))$. Noting that $s(p) = \Theta(p, s(p); 0)$ and using (2.5), (2.10) and (2.12), we have

$$\begin{aligned} & \|s(p) - \Phi_0(\Theta(p, s(p); g(p, u; s(p))))\| \\ (2.16) \quad &= \|\Theta(p, s(p); 0) - \Theta(p, s(p); g(p, u; s(p)))\| \\ &\leq \kappa \|g(p, u; s(p))\| \\ &= \kappa \left\| -f(p, u) - D_x f(p, u)(s(p) - u) + f(p, s(p)) \right\| \\ &\leq \frac{1}{2} \kappa \mu \|u - s(p)\|^2 \leq \gamma(1 - \kappa \varepsilon) \|u - s(p)\|^2 = \omega_0(1 - \kappa \varepsilon). \end{aligned}$$

Since $\mathcal{B}_{\omega_0}(s(p)) \subset \mathcal{B}_a(\bar{x})$, from (2.15) we immediately obtain

$$(2.17) \quad \|\Phi_0(v) - \Phi_0(v')\| \leq \kappa \varepsilon \|v - v'\| \quad \text{for any } v, v' \in \mathcal{B}_{\omega_0}(s(p)).$$

Hence, the contraction mapping principle applied to the function Φ_0 on the ball $\mathcal{B}_{\omega_1}(s(p))$ yields that there exists x'_1 in this ball such that $x'_1 = \Phi_0(x'_1)$. But the fixed point x'_1 of Φ_0 in $\mathcal{B}_{\omega_0}(s(p))$ must then coincide with the unique fixed point x_1 of Φ_0 in the larger set $\mathcal{B}_a(\bar{x})$. Hence, the fixed point x_1 of Φ_0 on $\mathcal{B}_a(\bar{x})$ satisfies

$$\|x_1 - s(p)\| \leq \gamma \|u - s(p)\|^2,$$

which, for $x_0 = u$, means that (2.4) holds for $k = 0$.

The induction step is now clear: if the claim holds for $k = 1, 2, \dots, n$, by defining $\Phi_n : x \mapsto \Theta(p, s(p); g(p, x_n; x))$ and replacing u by x_n in (2.14) and (2.15), we obtain that the function Φ_n has a unique fixed point x_{n+1} in $\mathcal{B}_a(\bar{x})$. This means that $g(p, x_n; x_{n+1}) \in G_{p, x_n}(x_{n+1})$ and hence x_{n+1} is the unique Newton iterate from x_n for p which is in $\mathcal{B}_a(\bar{x})$. Next, by employing again the contraction mapping principle as in (2.16) and (2.17) to Φ_n but now on the ball $\mathcal{B}_{\omega_n}(s(p))$ for $\omega_n = \gamma \|x_n - s(p)\|^2$ we obtain that x_{n+1} is at distance ω_n from $s(p)$. Using the first inequality in (2.9) and then the last one in (2.7) we have

$$\theta := \gamma \|x_0 - s(p)\| \leq \gamma (\|x_0 - \bar{x}\| + \|s(p) - \bar{x}\|) \leq \gamma \left(a + \frac{a}{2}\right) < 1.$$

Hence,

$$(2.18) \quad \|x_k - s(p)\| \leq \theta^{2^k - 1} \|x_0 - s(p)\|$$

and therefore the sequence $\{x_k\}$ is quadratically convergent to $s(p)$ as in (2.4). This completes the proof of the theorem. \square

3 An implicit function theorem for the Newton iterations

In this section we make a step further in exploring the dependence of Newton's iteration on parameters. Our main result is the following theorem which follows the general format of the implicit function theorem paradigm.

Theorem 3.1 (implicit function theorem for Newton's iteration). *In addition to the assumptions of Theorem 2.1, suppose that*

$$(3.1) \quad \text{lip}(D_x f; (\bar{p}, \bar{x})) < \infty.$$

Then the single-valued localization ξ of the mapping Ξ in (2.2) at (\bar{p}, \bar{x}) for $\bar{\xi}$ described in Theorem 2.1 is Lipschitz continuous near (\bar{x}, \bar{p}) , moreover having

$$(3.2) \quad \widehat{\text{lip}}_u(\xi; (\bar{x}, \bar{p})) = 0 \quad \text{and} \quad \widehat{\text{lip}}_p(\xi; (\bar{x}, \bar{p})) \leq \text{lip}(\sigma; 0) \cdot \widehat{\text{lip}}_p(f; (\bar{p}, \bar{x})).$$

Proof. First, recall some notation and facts established in Theorem 2.1 and its proof. We know that for any $\kappa > \text{lip}(\sigma; 0)$ there exist positive constants a, α, b and c such that $a \leq \alpha$ and for every $p \in \mathcal{B}_c(\bar{p})$ and $w \in \mathcal{B}_a(\bar{x})$ the mapping $y \mapsto G_{p, w}^{-1}(y) \cap \mathcal{B}_\alpha(\bar{x})$ is a function, with values $\Theta(p, w; y)$, which is Lipschitz continuous on $\mathcal{B}_b(0)$ with Lipschitz constant κ ; moreover, the truncation $S(p) \cap \mathcal{B}_a(\bar{x})$ of the solution mapping in (2) is a Lipschitz continuous function on $\mathcal{B}_c(\bar{p})$ and its values are in $\mathcal{B}_{a/2}(\bar{x})$; also, for any starting

point $u \in \mathcal{B}_a(\bar{x})$ and any $p \in \mathcal{B}_c(\bar{p})$ there is a unique sequence $\xi(u, p)$ starting from u and generated by Newton's method (2.1) for p whose components are contained in $\mathcal{B}_a(\bar{x})$, and this sequence is quadratically convergent to $s(p)$ as described in (2.4).

Our first observation is that for any positive $a' \leq a$, by adjusting the size of the constant c and taking as a starting point $u \in \mathcal{B}_{a'}(\bar{x})$ we can have that for any $p \in \mathcal{B}_c(\bar{p})$ all elements x_k of the sequence $\xi(u, p)$ are actually in $\mathcal{B}_{a'}(\bar{x})$. Indeed, by taking $\delta > 0$ to satisfy (2.8) with a replaced by a' and then choosing c so that (2.9) holds for the new δ and for a' , then all requirements for a will hold for a' as well and hence all Newton's iterates x_k will be at distance a' from \bar{x} .

Let us choose positive η and ν such that

$$\eta > \text{lip}(D_x f; (\bar{p}, \bar{x})) \quad \text{and} \quad \nu > \widehat{\text{lip}}_p(f; (\bar{p}, \bar{x})),$$

and then pick a positive constant $d \leq a/2$ and make c smaller if necessary so that for every $p, p' \in \mathcal{B}_c(\bar{p})$ and every $w, w' \in \mathcal{B}_d(\bar{x})$ we have

$$(3.3) \quad \|D_x f(p', w') - D_x f(p, w)\| \leq \eta(\|p' - p\| + \|w' - w\|),$$

$$(3.4) \quad \|f(p', w) - f(p, w)\| \leq \nu\|p' - p\|,$$

and, in addition, for every $x \in \mathcal{B}_d(\bar{x})$, every $p, p' \in \mathcal{B}_c(\bar{p})$ and every $w, w' \in \mathcal{B}_d(\bar{x})$

$$(3.5) \quad \|f(p', w') - D_x f(p', w')(x - w') - f(p, w) - D_x f(p, w)(x - w)\| \leq b.$$

Choose a positive τ such that $\kappa\tau < \frac{1}{3}$. Make d and c smaller if necessary so that

$$(3.6) \quad 3\eta(d + c) < \tau.$$

Since $\frac{\kappa\tau}{1 - \kappa\tau} < \frac{1}{2}$ we can take c smaller in order to have

$$(3.7) \quad \frac{\kappa\tau(2d) + 3\kappa(\tau + \nu)(2c)}{1 - \kappa\tau} \leq d.$$

Let $p, p' \in \mathcal{B}_c(\bar{p})$, $u, u' \in \mathcal{B}_d(\bar{x})$, $(p, u) \neq (p', u')$ and, according to Theorem 2.1 and the observation above, let $\xi(p, u) = (x_1, \dots, x_k, \dots)$ be the unique sequence generated by Newton's iteration (2.1) starting from u whose components x_k are all in $\mathcal{B}_d(\bar{x})$ and hence in $\mathcal{B}_{a/2}(\bar{x})$. For this sequence, having $x_0 = u$, we know that for all $k \geq 0$

$$(3.8) \quad x_{k+1} = \Theta(p, x_k; 0),$$

where

$$\Theta(p, x_k; 0) = (f(p, x_k) + D_x f(p, x_k)(\cdot - x_k) + F(\cdot))^{-1}(0) \cap \mathcal{B}_\alpha(\bar{x}).$$

Let

$$\gamma_0 = \frac{\kappa\tau\|u - u'\| + \kappa(\tau + \nu)\|p - p'\|}{1 - \kappa\tau}.$$

By using (3.7) we get that $\gamma_0 \leq d$ and then $\mathcal{B}_{\gamma_0}(x_1) \subset \mathcal{B}_a(\bar{x})$. Consider the function

$$\Phi_0 : x \mapsto \Theta(p, u; -f(p', u') - D_x f(p', u')(x - u') + f(p, u) + D_x f(p, u)(x - u)).$$

Using (3.5) and then the Lipschitz continuity of $\Theta(p, u; \cdot)$ on $\mathcal{B}_b(0)$, and applying (2.10), (3.3), (3.4) and (3.6) we obtain

$$(3.9) \quad \begin{aligned} \|x_1 - \Phi_0(x_1)\| &= \|\Theta(p, u; 0) - \\ &\quad \Theta(p, u; -f(p', u') - D_x f(p', u')(x_1 - u') \\ &\quad + f(p, u) + D_x f(p, u)(x_1 - u))\| \\ &\leq \kappa \|f(p', u) - f(p', u') - D_x f(p', u')(u - u') \\ &\quad - D_x f(p', u')(x_1 - u) - f(p', u) \\ &\quad + f(p, u) + D_x f(p, u)(x_1 - u)\| \\ &\leq \kappa \|f(p', u) - f(p', u') - D_x f(p', u')(u - u')\| \\ &\quad + \kappa \|(D_x f(p, u) - D_x f(p, u'))(x_1 - u)\| \\ &\quad + \|(D_x f(p, u') - D_x f(p', u'))(x_1 - u)\| \\ &\quad + \kappa \|-f(p', u) + f(p, u)\| \\ &\leq \frac{1}{2} \kappa \eta \|u - u'\|^2 + \kappa \eta \|u - u'\| \|x_1 - u\| \\ &\quad + \kappa \eta \|p - p'\| \|x_1 - u\| + \kappa \nu \|p - p'\| \\ &\leq 3\kappa \eta d \|u - u'\| + \kappa(2\eta d + \nu) \|p - p'\| \\ &< \kappa \tau \|u - u'\| + \kappa(\tau + \nu) \|p - p'\| = \gamma_0(1 - \kappa \tau). \end{aligned}$$

For $v, v' \in \mathcal{B}_{\gamma_0}(x_1)$, using (3.3), (3.4) and (3.6) we have

$$(3.10) \quad \begin{aligned} \|\Phi_0(v) - \Phi_0(v')\| &\leq \kappa \|(-D_x f(p', u') + D_x f(p, u))(v - v')\| \\ &\leq 2\kappa \eta(d + c) \|v - v'\| \leq \kappa \tau \|v - v'\|. \end{aligned}$$

Thus, by the contraction mapping principle there exists a unique x'_1 in $\mathcal{B}_{\gamma_0}(x_1)$ such that

$$x'_1 = \Theta(p, u; -f(p', u') - D_x f(p', u')(x - u') + f(p, u) + D_x f(p, u)(x - u)).$$

But then

$$f(p', u') + Df(p', u')(x'_1 - u') + F(x'_1) \ni 0,$$

that is, x'_1 is the unique Newton iterate from u' for p' which satisfies

$$\|x'_1 - x_1\| \leq \gamma_0.$$

Since $\gamma_0 < d$, we obtain that $x'_1 \in \mathcal{B}_a(\bar{x})$ and then x'_1 is the unique Newton's iterate from u' for p' which is in $\mathcal{B}_a(\bar{x})$.

By induction, we construct a sequence $\xi' = \{u', x'_1, x'_2, \dots, x'_k, \dots\} \in \Xi(p', u')$ such that the distance from x'_k to the corresponding components x_k of ξ satisfies the estimate

$$(3.11) \quad \|x'_k - x_k\| \leq \gamma_{k-1} := \frac{\kappa \tau \|x_{k-1} - x'_{k-1}\| + \kappa(\tau + \nu) \|p - p'\|}{1 - \kappa \tau}$$

for $k = 2, 3, \dots$. Suppose that for some $n > 1$ we have found x'_2, x'_3, \dots, x'_n with this property. First, observe that

$$\gamma_k \leq \left[\frac{\kappa \tau}{1 - \kappa \tau} \right]^{k+1} \|u - u'\| + \frac{\kappa(\tau + \nu)}{1 - \kappa \tau} \|p - p'\| \sum_{i=0}^k \left[\frac{\kappa \tau}{1 - \kappa \tau} \right]^i,$$

from where we get the estimate that for all $k = 0, 1, \dots, n-1$,

$$(3.12) \quad \gamma_k \leq \frac{\kappa\tau}{1-\kappa\tau} \|u - u'\| + \frac{\kappa(\tau + \nu)}{1-\kappa\tau} \|p - p'\|.$$

In particular, by (3.7) we obtain that $\gamma_k \leq d$ for all k and hence $x'_k \in \mathcal{B}_d(x_k) \subset \mathcal{B}_a(\bar{x})$.

To show that x'_{n+1} is a Newton iterate from x'_n for p' , we proceed in the same way as in obtaining x'_1 from u' for p' . Consider the function

$$\Phi_k : x \mapsto \Theta(p, x_k; -f(p', x'_k) - D_x f(p', x'_k)(x - x'_k) + f(p, x_k) + D_x f(p, x_k)(x - x_k)).$$

By replacing Φ_0 by Φ_k , u by x_k , u' by x'_k , and x_1 by x_{k+1} in (3.9) and (3.10) we obtain that

$$\|x_{k+1} - \Phi_k(x_{k+1})\| < \kappa\tau \|x_k - x'_k\| + \kappa(\tau + \nu) \|p - p'\| = \gamma_k(1 - \kappa\tau)$$

and

$$\|\Phi_k(v) - \Phi_k(v')\| \leq \kappa\tau \|v - v'\| \quad \text{for any } v, v' \in \mathcal{B}_{\gamma_k}(x_k).$$

Then, by the contraction mapping principle there exists a unique x'_{k+1} in $\mathcal{B}_{\gamma_k}(x_{k+1})$ with $x'_{k+1} = \Phi_k(x'_{k+1})$, which gives us

$$f(p', x'_k) + D_x f(p', x'_k)(x'_{k+1} - x'_k) + F(x'_{k+1}).$$

Moreover, since $\gamma_k \leq d$ we have that $x'_{k+1} \in \mathcal{B}_a(\bar{x})$.

We constructed a sequence x'_1, \dots, x'_k, \dots which is generated by Newton's iteration for p' starting from u' and whose components are in $\mathcal{B}_a(\bar{x})$. According to Theorem 2.1, this sequence must be the value $\xi(u', p')$ of the single-valued localization ξ whose value $\xi(u, p)$ is the sequence x_1, \dots, x_k, \dots . Taking into account (3.11) and (3.12) we obtain

$$\|\xi(u, p) - \xi(u', p')\|_\infty \leq O(\tau) \|u - u'\| + (\kappa\nu + O(\tau)) \|p - p'\|.$$

Since τ can be chosen arbitrarily small, this gives us (3.2). \square

One should note the striking similarity between the estimate (1.5) for the Lipschitz modulus of the single-valued localization of the solution mapping (1.2) and the estimate (3.2) for Newton's iteration, which indicated the sharpness of the latter result. But there is more to be said: as in the case of Theorem 1.2, the inverse function version of Theorem 3.1 becomes an "if and only if" result.

Consider the generalized equation (1.1) with $f(p, x) = g(x) - p$ whose solution mapping S is described in (1.6) and let $\bar{x} \in S(\bar{p})$. The corresponding Newton's iteration mapping in (2.2) then has the form

$$(3.13) \quad \Upsilon : (u, p) \mapsto \left\{ \xi \in l_\infty(X) \mid \bigcap_{k=0}^{\infty} (g(x_k) + Dg(x_k)(x_{k+1} - x_k) + F(x_{k+1})) \ni p \text{ with } x_0 = u \right\}.$$

Recall that a set C is locally closed at a point $x \in C$ when there exists a closed neighborhood U of \bar{x} such that $C \cap U$ is a closed set.

Theorem 3.2 (inverse function theorem for Newton's iteration). *Suppose that the mapping F has locally closed graph at $(\bar{x}, \bar{p} - g(\bar{x}))$. Then the mapping $g + F$ is strongly regular at \bar{x} for \bar{p} if and only if the mapping Υ in (3.13) has a Lipschitz localization ξ at (\bar{x}, \bar{p}) for ξ with*

$$(3.14) \quad \widehat{\text{lip}}_u(\xi; (\bar{x}, \bar{p})) < 1$$

and such that for each (u, p) close to (\bar{x}, \bar{p}) the sequence $\xi(u, p)$ is convergent. Moreover, in this case

$$(3.15) \quad \widehat{\text{lip}}_p(\xi; (\bar{x}, \bar{p})) = \text{lip}(\sigma; \bar{p}) = \text{lip}(s; \bar{p}),$$

where σ and s are the Lipschitz localizations described in Theorem 1.2.

Proof. The “only if” part follows from the combination of Theorems 2.1 and 3.1. From (3.2) we get

$$(3.16) \quad \widehat{\text{lip}}_p(\xi; (\bar{x}, \bar{p})) \leq \text{lip}(\sigma; \bar{p}).$$

To prove the “if” part, choose $\kappa > \widehat{\text{lip}}_p(\xi; (\bar{x}, \bar{p}))$, a positive $\varepsilon < 1$ and corresponding neighborhoods U of \bar{x} and Q of \bar{p} such that the sequence $\xi(u, p)$ is the only element of $\Upsilon(u, p)$ whose components x_1, \dots, x_k, \dots are in U and, moreover, the function ξ acting from $X \times Y$ to $l_\infty(X)$ is Lipschitz continuous with Lipschitz constants κ in $p \in Q$ uniformly in $u \in U$ and, from (3.14), is Lipschitz continuous with Lipschitz constants ε in $u \in U$ uniformly in $p \in Q$. From the assumed local closedness of $\text{gph } F$ we are able to make Q and U smaller is necessary so that for any $p \in Q$ and any sequence with components $v_k \in Q$ convergent to v and satisfying

$$(3.17) \quad g(v_k) + Dg(v_k)(v_{k+1} - v_k) + F(v_{k+1}) \ni p \quad \text{for all } k = 1, 2, \dots$$

one $g(v) + F(v) \ni p$.

Let $p, p' \in Q$ and let $x \in (g + F)^{-1}(p) \cap U$. The constant sequence $\chi = (x, x, \dots, x)$ is obviously convergent to the solution x of the inclusion $g(x) + F(x) \ni p$. Then $\chi \in \Upsilon^{-1}(x, p)$ and all its components are in U , hence $\chi = \xi(p, u)$. By assumption,

$$(3.18) \quad \|\chi - \xi(x, p')\|_\infty \leq \kappa \|p - p'\|,$$

and moreover $\xi(x, p') = \{x, x'_1, \dots, x'_k, \dots\}$ is convergent. Note that, by definition, $\xi(x, p')$ satisfies

$$g(x'_k) + Dg(x'_k)(x'_{k+1} - x'_k) + F(x'_{k+1}) \ni p' \quad \text{for all } k = 1, 2, \dots$$

From the property described around (3.17) we obtain that the sequence $\xi(x, p')$ is convergent to a solution $x' \in (g + F)^{-1}(p') \cap U$. Hence, using (3.18), we have

$$\begin{aligned} \|x - x'\| &\leq \|x - x'_k\| + \|x'_k - x'\| \\ &\leq \|\chi - \xi(x, p')\|_\infty + \|x'_k - x'\| \\ &\leq \kappa \|p - p'\| + \|x'_k - x'\|. \end{aligned}$$

Since $x'_k \rightarrow x'$ as $n \rightarrow \infty$, by passing to the limit in this last inequality we conclude that

$$(3.19) \quad \|x - x'\| \leq \kappa \|p - p'\|.$$

We will now show that the mapping $(g + F)^{-1}$ has a single-valued localization at \bar{p} for \bar{x} . Assume that for every neighborhoods U of \bar{x} and Q of \bar{p} there exist $p \in Q$ and $w, w' \in U$ such that $w \neq w'$ and both w and w' are in $(g + F)^{-1}(p)$. Then the constant sequences $\{w, w, \dots, w, \dots\} \in \mathcal{Y}(u, p)$ and $\{w', w', \dots, w', \dots\} \in \mathcal{Y}(w', p)$ and all their components are in U , hence $\{w, w, \dots, w, \dots\} = \xi(w, p)$ and $\{w', w', \dots, w', \dots\} = \xi(w', p)$. In the beginning of the proof we have chosen the neighborhoods U and Q such that for a fixed $p \in V$ the mapping $u \mapsto \xi(u, p)$ is a Lipschitz continuous function from X to $l_\infty(X)$ with Lipschitz constant $\varepsilon < 1$, and hence this condition holds for all of its components. This yields

$$\|w - w'\| \leq \varepsilon \|w - w'\| < \|w - w'\|,$$

which is absurd. Hence, $(g + F)^{-1}$ has a single-valued localization s at \bar{p} for \bar{x} . But then from (3.19) this localization is Lipschitz continuous with $\text{lip}(s; \bar{p}) \leq \kappa$. Theorem 1.2 says that $\text{lip}(\sigma; \bar{p}) = \text{lip}(s; \bar{p})$ and hence $\text{lip}(\sigma; \bar{p}) \leq \kappa$. Since κ could be arbitrarily close to $\widehat{\text{lip}}_p(\xi; (\bar{x}, \bar{p}))$, we get the inequality opposite to (3.16), and hence (3.15) holds. \square

Remark. In addition to the conditions in Theorem 3.1, if we assume that f is continuously differentiable in a neighborhood of (\bar{p}, \bar{x}) and the *ample parameterization condition* holds, namely the derivative $D_p f(\bar{p}, \bar{x})$ is surjective, then, by using Lemma 2.4 from [5] (stated there in finite dimensions, but whose reformulation in Banach spaces needs changes only in terminology) we can modify the proof of Theorem 3.2 to obtain the equivalence of metric regularity of the mapping G in (1.4) with the existence of a single-valued localization of the mapping Ξ in (2.2) the values of which are convergent, as in the statement of Theorem 3.2. We choose not to present this generalization here in order to simplify the already involved presentation.

As an illustration of possible applications of the results in Theorems 2.1 and 3.1 in studying complexity of Newton's iteration, we will give an estimate for the number of iterations needed to achieve certain accuracy of the method, which is the same for all values of the parameter p in some neighborhood of the reference point \bar{p} . Given an accuracy measure ρ , suppose that Newton's method (2.1) is terminated at the k -th step if

$$(3.20) \quad \text{dist}(0, f(p, x_k) + F(x_k)) \leq \rho.$$

Also suppose that the constant μ and the constants a, c and θ are chosen as in the proof of Theorem 2.1 and for $p \in \mathcal{B}_c(\bar{p})$ consider the unique sequence $\{x_k\}$ generated by (2.1) for p and starting from $x_0 \in \mathcal{B}_a(\bar{x})$ so that all elements of which are in $\mathcal{B}_a(\bar{x})$. Since x_k is a Newton iterate from x_{k-1} we have that

$$f(p, x_k) - f(p, x_{k-1}) - D_x f(p, x_{k-1})(x_k - x_{k-1}) \in f(p, x_k) + F(x_k).$$

Using (2.10), we have

$$(3.21) \quad \begin{aligned} & \text{dist}(0, f(p, x_k) + F(x_k)) \\ & \leq \|f(p, x_k) - f(p, x_{k-1}) - D_x f(p, x_{k-1})(x_k - x_{k-1})\| \leq \frac{1}{2}\mu \|x_k - x_{k-1}\|^2. \end{aligned}$$

Let k_ρ be the first iteration at which (3.20) holds. Then for $k < k_\rho$, from (3.21) we obtain

$$(3.22) \quad \rho < \frac{1}{2}\mu \|x_k - x_{k-1}\|^2.$$

Further, utilizing (2.18) we get

$$\|x_k - x_{k-1}\| \leq \|x_k - s(p)\| + \|x_{k-1} - s(p)\| \leq \theta^{2^k - 2}(1 + \theta)(\|x_0 - \bar{x}\| + \|s(p) - \bar{x}\|),$$

and from the choice of x_0 and the first inequality in (2.9) we have

$$\|x_k - x_{k-1}\| \leq \theta^{2^k - 2}(1 + \theta)\frac{3a}{2}.$$

But then, taking into account (3.22) we obtain

$$\rho < \frac{1}{2}\mu\theta^{2^{k+1}}\frac{9a^2(1 + \theta)^2}{4\theta^4}.$$

Hence, any $k < k_\rho$ satisfies

$$k \leq \log_2 \left(\log_\theta \left(\frac{8\theta^4\rho}{9a^2\mu(1 + \theta)^2} \right) \right) - 1.$$

Thus, we obtained an upper bound of the numbers of iterations needed to achieve certain accuracy, which, most importantly, is the same for all values of the parameter p in some neighborhood of the reference value \bar{p} . This tells us, for example, that, on the assumptions of Theorem 3.1, small changes of parameters in a problem do not affect the performance of Newton's method applied to this problem.

References

1. Alt W.: Discretization and mesh-independence of Newton's method for generalized equations. *Lecture Notes in Pure and Appl. Math.*, **195** 1–30. Dekker, New York (1998).
2. Dontchev A. L.: Local analysis of a Newton-type method based on partial linearization. In: *The mathematics of numerical analysis* (Park City, UT, 1995), 295–306. *Lecture Notes in Appl. Math.* **32** AMS, Providence, RI (1996).
3. Dontchev A. L.: Lipschitzian stability of Newton's method for variational inclusions. In: *System Modelling and Optimization* (Cambridge, 1999), 119–147. Kluwer Academic Publishers, Boston, MA, (2000).
4. Dontchev A. L., Hager W. W., Veliov V. M.: Uniform convergence and mesh independence of Newton's method for discretized variational problems. *SIAM J. Control Optim.* **39** 961–980 (2000).
5. Dontchev A. L., Rockafellar R. T.: Ample parameterization of variational inclusions. *SIAM J. Optim.* **12** 170–187 (2001).

-
6. Dontchev A. L., Rockafellar R. T.: Robinson's implicit function theorem and its extensions. *Math. Program.* **117** Ser. B, 129–147 (2009).
 7. Dontchev A. L., Rockafellar R. T.: *Implicit functions and solution mappings*, Springer, to appear in December 2008.
 8. Geoffroy M. H., Hilout S., Piétrus A.: Stability of a cubically convergent method for generalized equations. *Set-Valued Analysis* **14** 41–54 (2006).
 9. Josephy, N. H.: Newton's method for generalized equations and the PIES energy model. Ph.D. Dissertation, Department of Industrial Engineering, University of Wisconsin-Madison (1979).
 10. Klatte D., Kummer B.: *Nonsmooth equations in optimization. Regularity, calculus, methods and applications. Nonconvex Optimization and its Applications* **60**. Kluwer Academic Publishers, Dordrecht (2002).
 11. Klatte D., Kummer B.: Stability of inclusions: characterizations via suitable Lipschitz functions and algorithms. *Optimization* **55** 627–660 (2006).
 12. Robinson S. M.: Strongly regular generalized equations. *Math. of Oper. Research* **5** 43–62 (1980).
 13. Robinson S. M.: Newton's method for a class of nonsmooth functions. *Set-valued analysis* **2** 291–305 (1994).