A robust Kantorovich's theorem on inexact Newton method with relative residual error tolerance

O. P. Ferreira[*] B. F. Svaiter[†]

November 12, 2010

Abstract

We prove that under semi-local assumptions, the inexact Newton method with a *fixed* relative residual error tolerance converges Q-linearly to a zero of the non-linear operator under consideration. Using this result we show that Newton method for minimizing a self-concordant function or to find a zero of an analytic function can be implemented with a fixed relative residual error tolerance.

In the absence of errors, our analysis retrieve the classical Kantorovich Theorem on Newton method.
Keywords: Kantorovich's theorem, Inexact Newton method, Banach space.
MSC2010: 49M15, 90C30.

Introduction

Newton's method and its variations, including the inexact Newton methods, are the most efficient methods known for solving nonlinear equations

$$F(x) = 0,$$

where $\mathbb{X}$ and $\mathbb{Y}$ are Banach spaces, $C \subseteq \mathbb{X}$ and $F : C \to \mathbb{Y}$ is continuous and continuously differentiable on $int(C)$.

Kantorovich's Theorem on Newton's method uses semi-local assumption on F to guarantee existence of a solution of the above equation, uniqueness of this solution in a prescribed region and

[*]IME/UFG, Campus II- Caixa Postal 131, CEP 74001-970 - Goiânia, GO, Brazil (e-mail:orizon@mat.ufg.br).
The author was supported in part by FUNAPE/UFG, CNPq Grant 473756/2009-9, CNPq Grant 302024/2008-5, PRONEX–Optimization(FAPERJ/CNPq) and IMPA.

[†]IMPA, Estrada Dona Castorina, 110, Jardim Botânico, CEP 22460-320 - Rio de Janeiro, RJ, Brazil (e-mail:benar@impa.br).
The author was supported in part by CNPq Grant 301200/93-9(RN), CNPq Grant 475647/2006-8 and by PRONEX–Optimization(FAPERJ/CNPq).

also convergence of Newton's Method to such a solution, see [8, 9]. Semi-local convergence theorems for Newton method has been instrumental in the modern complexity analysis of the solution of polynomial (or analytical) equations [2, 18], linear and quadratic programming problems and linear semi-definite programming problems [15, 16]. These convergence results has also been used in the design and convergence analysis of algorithms for these problems. In all these applications, homotopy methods are combined with Newton's method, which helps the algorithm to keep track of the solution of a parametrized perturbed version of the original problem. Each Newton iteration requires the solution of a linear system, and this accounts mostly for the computational burden of these algorithms.

Since linear system are solved always inexactly in floating point computations, it is natural to investigate robustness of Kantorovich's and Kantorovich's-like theorems under errors in the computation of the Newton step. Moreover, modern implementations of the conjugate gradients, coupled with preconditioning, allows for the approximated solution of large linear systems. It would be most desirable to have an *a priori* prescribed residual error tolerance in the iterative solutions of linear system for computing the Newton steps, because this would prevent over-solving and/or under-solving the linear system in question. Although the local convergence analysis of Newton's method with relative errors in the residue [3, 4, 14] or in the steep [22] are well understood, the convergence analysis of the method under general semi-local assumptions assuming *only* bounded relative residual errors is a new contribution of this paper. Previous works on this subject include [13, 17]. The advantage of working with an error tolerance on the residual rests in the fact that the exact Newton step need not to be know for evaluating this error, which makes this criterion attractive for practical applications.

Recently, Kantorovich's theorem on Newton's Method was extended to Riemannian manifolds using a new technique which simplifies the analyses and proof of this theorem, see [5]. After that, this technique was successfully employed for proving generalized versions of Kantorovich's theorem in Riemannian Manifolds and also in the analysis of the classical version of Kantorovich's theorem in Banach spaces, see[1, 6, 10, 11, 12, 19, 20, 21]. In the present work, we will use the technique introduced in [5] to present a robust version of the Kantorovich's theorem on the inexact Newton method with residual relative error. It is worth to point out that, for null error tolerance the analysis presented merge in the usual semi-local convergence analysis on Newton's method, see [6]. The basic idea is to find good regions for the inexact Newton method. In these regions, the majorant function bounds the non-linear function which root is to be found, and the behavior of the inexact Newton iteration in these regions is estimated using iterations associated to the majorant function. Moreover, as a whole, the union of all these regions is invariant under inexact Newton's iteration.

This paper is organized as follows. In Section 1, some definitions and auxiliary results are presented. In Section 2 the main result is stated and some properties of the majorant function are established. The main relationships between the majorant function and the nonlinear operator used in the paper are presented in Section 3. In Section 4 a family of regions where the behavior of the inexact Newton iteration is estimated using the majorant function is introduced. We also show that the union of all these regions is invariant under the inexact Newton iteration with a fixed relative residual error tolerance.. In Section 5 the main result is proved. In Section 6 we show that Newton method for minimizing a self-concordant function under the usual semi-local assumption for these functions, can be implemented with a fixed residual error tolerance. Moreover, we show that Newton method for finding a zero of an analytic function, under the usual semi-local assumption of the α-theory can be also be implemented with a fixed relative residual error tolerance.

1 Basics definitions and auxiliary results

Let $\mathbb{X}$ be a Banach space. The open and closed ball at x are denoted, respectively by

$$B(x,r) = \{y \in \mathbb{X};\ \|x-y\| < r\}, \qquad B[x,r] = \{y \in \mathbb{X};\ \|x-y\| \leqslant r\}.$$

The following auxiliary results of elementary convex analysis will be needed:

Proposition 1.1. *Let $I \subset \mathbb{R}$ be an interval, and $\varphi : I \to \mathbb{R}$ be convex.*

1. *For any $u_0 \in \operatorname{int}(I)$, the application*

$$u \mapsto \frac{\varphi(u_0) - \varphi(u)}{u_0 - u}, \qquad u \in I, u \neq u_0,$$

is increasing and there exist (in $\mathbb{R}$)

$$D^-\varphi(u_0) = \lim_{u \to u_0^-} \frac{\varphi(u_0) - \varphi(u)}{u_0 - u} = \sup_{u < u_0} \frac{\varphi(u_0) - \varphi(u)}{u_0 - u}.$$

2. *If $u, v, w \in I$, $u < w$, and $u \leq v \leq w$ then*

$$\varphi(v) - \varphi(u) \leq [\varphi(w) - \varphi(u)] \frac{v-u}{w-u}.$$

Proof. See [7]. □

Proposition 1.2. *If $h : [a,b) \to \mathbb{R}$ is convex, differentiable at a, $h'(a) < 0$ and*

$$\lim_{t \to b_-} h(t) = 0,$$

then

$$a - \frac{h(a)}{h'(a)} \leq b,$$

with equality if and only if h is affine in $[a,b)$.

Proof. Since h is convex, $h(a) + h'(a)(t - a) \leq h(t)$ for any $t \in [a, b)$. Taking the limit $t \to b_-$ we obtain

$$h(a) + h'(a)(b - a) \leq 0.$$

The desired inequality now follows multiplying this inequality by the strictly positive number $-1/h'(a)$. If the above inequality holds as an equality, then

$$h'(a) = \frac{-h(a)}{b - a}.$$

Let $a \leq s < t < b$. Using again the convexity of h we have

$$h(a) + h'(a)(s - a) \leq h(s) \leq h(a)\frac{t - s}{t - a} + h(t)\frac{s - a}{t - a}.$$

Taking again the limit $t \to b_-$ in the above equation and using the previous equation we conclude that $h(s) = h(a)(b-s)/(b-a)$, i.e., h is affine. If h is affine then the the inequality of the proposition holds trivially as an equality. □

2 The inexact Newton method with relative error

Our goal is to prove the following version of Kantorovich's theorem on inexact Newton's Method with relative error.

Theorem 2.1. *Let $\mathbb{X}$ and $\mathbb{Y}$ be Banach spaces, $R \in \mathbb{R}$, $C \subseteq \mathbb{X}$ and $F : C \to \mathbb{Y}$ a continuous function, continuously differentiable on int(C). Take $x_0 \in \operatorname{int}(C)$ with $F'(x_0)$ non-singular. Suppose that $f : [0, R) \to \mathbb{R}$ is continuously differentiable, $B(x_0, R) \subseteq C$,*

$$\|F'(x_0)^{-1}\left[F'(y) - F'(x)\right]\| \leq f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|), \tag{2.1}$$

for any $x, y \in B(x_0, R)$, $\|x - x_0\| + \|y - x\| < R$,

$$\|F'(x_0)^{-1}F(x_0)\| \leq f(0)\,, \tag{2.2}$$

- **h1)** $f(0) > 0$, $f'(0) = -1$;
- **h2)** f' *is strictly increasing and convex;*
- **h3)** $f(t) < 0$ *for some* $t \in (0, R)$.

Let

$$\beta := \sup_{t\in[0,R)} -f(t), \quad t_* := \min f^{-1}(\{0\}), \quad \bar{\tau} := \sup\{t \in [0, R) \,:\, f(t) < 0\}.$$

Take $0 \le \rho < \beta/2$ and define

$$\kappa_\rho := \sup_{\rho<t<R} \frac{-(f(t)+2\rho)}{|f'(\rho)|\,(t-\rho)}, \quad \lambda_\rho := \sup\{t \in [\rho, R) : \kappa_\rho + f'(t) < 0\}, \quad \Theta_\rho := \frac{\kappa_\rho}{2-\kappa_\rho}. \tag{2.3}$$

Then for any $\theta \in [0, \Theta_\rho]$ and $z_0 \in B(x_0, \rho)$, the sequence generated by the inexact Newton method for solving $F(x) = 0$ with starting point z_0 and residual relative error tolerance θ: For $k = 0, 1, \ldots,$

$$z_{k+1} = z_k + S_k, \qquad \left\|F'(z_0)^{-1}\left[F(z_k) + F'(z_k)S_k\right]\right\| \le \theta \|F'(z_0)^{-1}F(z_k)\|,$$

is well defined (for any particular choice of each S_k),

$$\|F'(z_0)^{-1}F(z_k)\| \le \left(\frac{1+\theta^2}{2}\right)^k [f(0)+2\rho], \qquad k = 0, 1, \ldots\,, \tag{2.4}$$

the sequence $\{z_k\}$ is contained in $B(z_0, \lambda_\rho)$ and converges to a point $x_* \in B[x_0, t_*]$, which is the unique zero of F on $B(x_0, \bar\tau)$. Moreover, if

h4) $\lambda_\rho < R - \rho$,

then the sequence $\{z_k\}$ satisfies, for $k = 0, 1, \ldots,$

$$\|x_* - z_{k+1}\| \le \left[\frac{1+\theta}{2}\frac{D^- f'(\lambda_\rho)}{|f'(\lambda_\rho)|}\|x_* - z_k\| + \theta\,\frac{f'(\lambda_\rho + \rho) + 2|f'(\rho)|}{|f'(\lambda_\rho + \rho)|}\right]\|x_* - z_k\|.$$

If, additionally, $0 \le \theta < \kappa_\rho/(4+\kappa_\rho)$ then $\{z_k\}$ converges Q-linearly as follows

$$\|x_* - z_{k+1}\| \le \left[\frac{1+\theta}{2} + \frac{2\theta}{\kappa_\rho}\right]\|x_* - z_k\|, \qquad k = 0, 1, \ldots\,.$$

Remark 2.2. *In Theorem 2.1 if $\theta = 0$ we obtain the exact Newton method and its convergence properties. Now, taking $\theta = \theta_k$ in each iteration and letting θ_k goes to zero as k goes to infinity, the penultimate inequality of the Theorem 2.1 implies that the generated sequence converges to the solution with superlinear rate.*

¿From now on, we assume that the hypotheses of Theorem 2.1 hold. The scalar function f in the above theorem is called a *majorant function* for F at point x_0. Before proceeding, we will analyze some basic properties of the majorant function. Condition **h2** implies in strict convexity of f. Note that t_* is the smallest root of $f(t) = 0$ and, since f is convex, if this equation has two roots, then the second one is $\bar\tau$.

Define

$$\bar t := \sup\{t \in [0, R) : f'(t) < 0\}\,. \tag{2.5}$$

Proposition 2.3. *The following statements on the majorant function hold* *i)* $f'(t) < 0$ *for any* $t \in [0, \bar{t})$, *(and* $f'(t) \geq 0$ *for any* $t \in [0, R) \setminus [0, \bar{t})$*);*

ii) $0 < t_* < \bar{t} \leq \bar{\tau} \leq R$*;*

iii) $\beta = -\lim_{t \to \bar{t}_-} f(t), \qquad 0 < \beta < \bar{t}.$

Proof. Item *i* follows from the second part of **h1**, **h2** and the definition (2.5).

Using the first inequality in **h1**, **h3** and the continuity of f we conclude that t_* is well defined and

$$0 < t_* < R.$$

Condition **h2** implies in strict convexity of f, hence condition **h3** and the definition of t_* imply that there exists $t \in (t_*, R)$ such that

$$0 > f(t) > f(t_*) + f'(t_*)(t - t_*) = f'(t_*)(t - t_*),$$

which implies that $0 > f'(t_*)$. Therefore, using item i and the definition of $\bar{t}$ we have

$$t_* < \bar{t} \leq R.$$

Since t_* is the smallest root of $f(t) = 0$ and f is strictly decreasing in $[0, \bar{t})$ we conclude that $f < 0$ in $[t_*, \bar{t})$. So, the definition of $\bar{\tau}$ implies that

$$\bar{t} \leq \bar{\tau} \leq R,$$

and the proof of item ii is concluded.

Using **h3** and the definition of β we obtain that $0 < \beta$. Since f is convex, combining this with **h1** we have

$$f(t) \geq f(0) - t > -t, \qquad 0 \leq t < R,$$

with strict inequality for $t \neq 0$. We know that f is strictly decreasing and $f < 0$ in $[t_*, \bar{t})$. Hence, letting t goes to $\bar{t}_-$ in last inequality and using the definition of β the item iii follows. □

We will first prove Theorem 2.1 for the case $\rho = 0$ and $z_0 = x_0$. In order to simplify the notation in the case $\rho = 0$, we will use κ, λ and θ instead of κ_0, λ_0 and θ_0 respectively:

$$\kappa := \sup_{0<t<R} \frac{-f(t)}{t}, \quad \lambda := \sup\{t \in [0, R) : \kappa + f'(t) < 0\}, \quad \Theta := \frac{\kappa}{2 - \kappa}. \tag{2.6}$$

Proposition 2.4. *For* κ, λ, θ *as in* (2.6) *it holds that*

$$0 < \kappa < 1, \qquad 0 < \Theta < 1, \qquad t_* < \lambda \leq \bar{t} \leq \bar{\tau}, \tag{2.7}$$

and

$$\begin{gathered} f'(t) + \kappa < 0, \qquad \forall\, t \in [0, \lambda), \\ \inf_{0 \leq t < R} f(t) + \kappa t = \lim_{t \to \lambda_-} f(t) + \kappa t = 0, \end{gathered} \tag{2.8}$$

Proof. Since f is convex, combining this with **h1** we have

$$f(t) \geq f(0) - t > -t, \qquad 0 \leq t < R,$$

with strict inequality for $t \neq 0$. For $t \neq 0$, last inequality is equivalent to

$$\frac{-f(t)}{t} \leq 1 - \frac{f(0)}{t} < 1 - \frac{f(0)}{R} < 1, \qquad 0 < t < R,$$

and, using also **h3**, we conclude that

$$0 < \kappa < 1, \qquad 0 < \Theta < 1,$$

where the bounds on Θ follows from its definition and the bound on κ. Moreover, as f' is continuous, strictly increasing and $f'(0) = -1$, we obtain

$$0 < \lambda, \qquad f'(t) + \kappa < 0, \qquad \forall\, t \in [0, \lambda),$$
$$\inf_{0 \leq t < R} f(t) + \kappa t = \lim_{t \to \lambda_-} f(t) + \kappa t = 0,$$

where the last equalities follows from the definition of κ.

Note that $f'(t) = -\kappa < 0$ for all $t \in [0, \lambda)$. Since f' is strictly negative in $[0, \lambda)$, we conclude that $t_* < \lambda \leq \bar{t} \leq \bar{\tau}$ and the proof is concluded. □

3 Basic results

In this section we will obtain bounds on $\|F'^{-1}\|$ and on the linearization error on F using the majorant function f. This bounds will be used in the next section for analyzing the inexact Newton iterations. It is worth mentioning that in this section the inequality on **h1** and (2.2) will be used only for proving its last result and **h3** will not be used.

A Newton iteration at x requires non-singularity of $F'(x)$, which will be verified using the majorant function f.

Proposition 3.1. *If* $\|x - x_0\| \leq t < \bar{t}$ *then* $F'(x)$ *is non-singular and*

$$\|F'(x)^{-1}F'(x_0)\| \leq \frac{1}{-f'(t)}.$$

Proof. The definition (2.5) shows that $f'(t) < 0$. Direct manipulation, (2.1), **h1** and **h2** give us

$$\begin{aligned}\|F'(x_0)^{-1}F'(x) - I\| = \|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| &\leq f'(\|x - x_0\|) - f'(0)\\ &= f'(t) + 1 < 1.\end{aligned}$$

Using Banach's Lemma and the last inequality above we conclude that $F'(x_0)^{-1}F'(x)$ is non-singular and

$$\|F'(x)^{-1}F'(x_0)\| = \|(F'(x_0)^{-1}F'(x))^{-1}\| \leq \frac{1}{1-(f'(t)+1)},$$

which is the desired inequality. □

The linearization errors on F and f are, respectively

$$E_F(y,x) := F(y) - \left[F(x) + F'(x)(y-x)\right], \qquad x \in B(x_0,R), \qquad y \in C \tag{3.1}$$
$$e_f(v,t) := f(v) - [f(t) + f'(t)(v-t)], \qquad t,v \in [0,R). \tag{3.2}$$

The linearization error of the majorant function bounded the linearization error of F.

Lemma 3.2. *If $x, y \in \mathbb{X}$ and $\|x - x_0\| + \|y - x\| < R$ then*

$$\|F'(x_0)^{-1}E_F(y,x)\| \leq e_f(\|x-x_0\| + \|y-x\|, \|x-x_0\|),$$

Proof. Since

$$x + u(y-x) \in B(x_0,R), \qquad 0 \leq u \leq 1,$$

and F is continuously differentiable in $B(x_0,R)$, direct use of (3.1) gives

$$E_F(y,x) = \int_0^1 [F'(x+u(y-x)) - F'(x)](y-x)\, du.$$

Combining the above equality with (2.1) we have

$$\begin{aligned}
&\|F'(x_0)^{-1}E_F(y,x)\| \\
&\leq \int_0^1 \left\|F'(x_0)^{-1}[F'(x+u(y-x)) - F'(x)]\right\| \|y-x\| \; du \\
&\leq \int_0^1 \left[f'\left(\|x-x_0\| + u\,\|y-x\|\right) - f'\left(\|x-x_0\|\right)\right] \|y-x\|\, du
\end{aligned}$$

which after performing the integration and using the definition in (3.2) yields the desired inequality. □

Convexity of f and f' guarantee that $e_f(t+s,t)$ is increasing in s *and* t.

Lemma 3.3. *If $0 \leq b \leq t$, $0 \leq a \leq s$ and $t + s < R$ then*

$$\begin{aligned}
&e_f(a+b,b) \leq e_f(t+s,t), \\
&e_f(a+b,b) \leq \frac{1}{2}\frac{f'(t+s) - f'(t)}{s}\, a^2, \qquad s \neq 0.
\end{aligned}$$

Proof. First note that

$$e_f(a+b,b) = \int_0^a \left[f'(b+r) - f'(b)\right] dr\,.$$

Since f' is convex, for any $\tau_0 > 0$, the function $\tau \mapsto f'(\tau+\tau_0) - f'(\tau)$ is non-decreasing. So,

$$e_f(a+b,b) \le \int_0^a \left[f'(t+r) - f'(t)\right] dr \;\le \int_0^s \left[f'(t+r) - f'(t)\right] dr\,. \tag{3.3}$$

where the second inequality follows from the convexity of f, which implies positivity of the integrand. To end the proof of first inequality, note that the last term on the above inequality is $e_f(t+s,t)$.

For proving second inequality, apply Proposition 1.1 item 2 with $u = t$, $v = t+r$, $w = t+s$ and $\varphi = f'$ in first inequality in (3.3) to conclude that

$$e_f(a+b,b) \le \int_0^a [f'(t+s) - f'(t)]\,\frac{r}{s}\,dr,$$

which performing the integral gives the desired inequality. □

Now we are ready to bound the linearization error E_F using the linearization error on the majorant function.

Corollary 3.4. *If $x, y \in \mathbb{X}$, $\|x - x_0\| \le t$, $\|y - x\| \le s$ and $s + t < R$ then*

$$\|F'(x_0)^{-1}E_F(y,x)\| \le e_f(t+s,t),$$
$$\|F'(x_0)^{-1}E_F(y,x)\| \le \frac{1}{2}\frac{f'(s+t) - f'(t)}{s}\|y-x\|^2, \qquad s \neq 0.$$

Proof. The results follows by direct combination of the Lemmas 3.2 and 3.3 by taking $b = \|x - x_0\|$ and $a = \|y - x\|$. □

The first inequality in the next corollary will be useful for obtaining asymptotic bounds on the sequence generated by the inexact Newton method with relative error tolerance, while the second inequality will be used to show that this method is robust with respect to the initial iterate.

Corollary 3.5. *For any $y \in B(x_0, R)$,*

$$-f(\|y-x_0\|) \le \left\|F'(x_0)^{-1}F(y)\right\| \le f(\|y-x_0\|) + 2\|y-x_0\|\,.$$

Proof. Using Lemma 3.2 with $x = x_0$, the definition of E_F and triangle inequality we have

$$\begin{aligned} e_f(\|y-x_0\|, 0) &\ge \left\|F'(x_0)^{-1}E_F(y,x_0)\right\| \\ &\ge \left\|F'(x_0)^{-1}F(x_0) + y - x_0\right\| - \left\|F'(x_0)^{-1}F(y)\right\| \\ &\ge \|y-x_0\| - \left\|F'(x_0)^{-1}F(x_0)\right\| - \left\|F'(x_0)^{-1}F(y)\right\|. \end{aligned}$$

Combining this inequality with the definition of e_f and using the assumption **h1** and (2.2) we obtain

$$f(\|y-x_0\|) - f(0) + \|y-x_0\| \geq \|y-x_0\| - f(0) - \left\|F'(x_0)^{-1}F(y)\right\|,$$

which is equivalent to the firs inequality of the corollary.

To prove the second inequality, use again Lemma 3.2 the definition of E_F and triangle inequality to obtain

$$\begin{aligned} e_f(\|y-x_0\|, 0) &\geq \left\|F'(x_0)^{-1}E_F(y,x_0)\right\| \\ &\geq \left\|F'(x_0)^{-1}F(y)\right\| - \left\|F'(x_0)^{-1}F(x_0) + y - x_0\right\| \\ &\geq \left\|F'(x_0)^{-1}F(y)\right\| - \left\|F'(x_0)^{-1}F(x_0)\right\| - \|y-x_0\|. \end{aligned}$$

Using the above inequality, the definition of e_f, **h1** and (2.2) we have

$$f(\|y-x_0\|) - f(0) + \|y-x_0\| \geq \left\|F'(x_0)^{-1}F(y)\right\| - f(0) - \|y-x_0\|.$$

which is equivalent to the second inequality of the corollary. □

Note that the first inequality on the above corollary proves that F has no zeroes in the region $t_* < \|x-x_0\| < \bar{\tau}$.

Lemma 3.6. *If* $x \in \mathbb{X}$, $\|x-x_0\| \leq t < R$ *then*

$$\|F'(x_0)^{-1}F'(x)\| \leq 2 + f'(t).$$

Proof. Simple algebraic manipulation together with assumption (2.1) give us

$$\|F'(x_0)^{-1}F'(x)\| \leq I + F'(x_0)^{-1}[F'(x) - F'(x_0)] \leq 1 + f'(\|x-x_0\|) - f'(0).$$

Hence, **h1**, **h2** and the last inequality imply the statement of the lemma. □

Lemma 3.7. *Take* $\theta \geq 0$, $0 \leq t \leq \lambda$, $x_*, x, y \in \mathbb{X}$. *If* $\lambda < R$, $\|x-x_0\| \leq t$, $\|x_* - x\| \leq \lambda - t$, $F(x_*) = 0$ *and*

$$\left\|F'(x_0)^{-1}[F(x) + F'(x)(y-x)]\right\| \leq \theta \|F'(x_0)^{-1}F(x)\|, \tag{3.4}$$

then

$$\|x_* - y\| \leq \left[\frac{1+\theta}{2} + \frac{2\theta}{\kappa}\right] \|x_* - x\|, \tag{3.5}$$

$$\|x_* - y\| \leq \left[\frac{1+\theta}{2}\frac{D^- f'(\lambda)}{|f'(\lambda)|}\|x_* - x\| + \theta\,\frac{2+f'(\lambda)}{|f'(\lambda)|}\right] \|x_* - x\|. \tag{3.6}$$

Proof. Since $F(x_*)=0$, direct algebraic manipulation and (3.1) yield

$$y-x_* = F'(x)^{-1}\left[E_F(x_*,x)+[F(x)+F'(x)(y-x)]\right].$$

Using (3.4), properties of the norm and some simple manipulations we conclude from last equality that

$$\|x_*-y\| \le \left\|F'(x)^{-1}F'(x_0)\right\| \left[\left\|F'(x_0)^{-1}E_F(x_*,x)\right\|+\theta\left\|F(x_0)^{-1}F(x)\right\|\right].$$

On the other hand, using again $F(x_*)=0$ and the definition in (3.1) we have

$$-F'(x_0)^{-1}F(x) = F'(x_0)^{-1}\left[E_F(x_*,x)+F'(x)(x_*-x)\right],$$

which using the triangular inequality yields

$$\left\|F'(x_0)^{-1}F(x)\right\| \le \left\|F'(x_0)^{-1}E_F(x_*,x)\right\| + \left\|F'(x_0)^{-1}F'(x)\right\| \|x_*-x\|.$$

Combining two above inequalities with Proposition 3.1, Corollary 3.4 with $y=x_*$ and $s=\lambda-t$ and Lemma 3.6 we have

$$\|x_*-y\| \le \frac{1}{|f'(t)|}\left[\frac{1+\theta}{2}\frac{f'(\lambda)-f'(t)}{\lambda-t}\|x_*-x\|+\theta\,[2+f'(t)]\right]\|x_*-x\|.$$

Since $\|x_*-x\|\le\lambda-t$, $f'<-\kappa<0$ in $[0,\lambda)$ and f' is increasing the first inequality follows from last inequality.

Using Proposition 1.1 and taking in account that $f'<0$ in $[0,\lambda)$ and increasing we obtain the second inequality from above inequality. □

4 The inexact Newton iteration with relative error

In the next lemma we study a single inexact Newton iteration with relative error θ.

Lemma 4.1. *Take $t,\varepsilon,\theta\ge 0$, and $x\in C$ such that*

$$\|x-x_0\|\le t<\bar{t}, \quad \|F'(x_0)^{-1}F(x)\|\le f(t)+\varepsilon, \quad t-(1+\theta)\frac{f(t)+\varepsilon}{f'(t)}<R. \tag{4.1}$$

If $y\in X$ and

$$\|F'(x_0)^{-1}[F(x)+F'(x)(y-x)]\|\le\theta\|F'(x_0)^{-1}F(x)\|. \tag{4.2}$$

then

1. $\|y-x\|\le -(1+\theta)\dfrac{f(t)+\varepsilon}{f'(t)}$;
2. $\|y-x_0\|\le t-(1+\theta)\dfrac{f(t)+\varepsilon}{f'(t)}<R$;

3. $\|F'(x_0)^{-1}F(y)\| \leq f\left(t-(1+\theta)\dfrac{f(t)+\varepsilon}{f'(t)}\right)+\varepsilon+2\theta(f(t)+\varepsilon)$.

Proof. Using Proposition 3.1 and the first inequality in (4.1) we conclude that $F'(x)$ is non-singular and $\left\|F'(x)^{-1}F'(x_0)\right\| \leq -1/f'(t)$. Therefore, using also the identity

$$y-x = F'(x)^{-1}F'(x_0)\Big[F'(x_0)^{-1}\left[F(x)+F'(x)(y-x)\right]-F'(x_0)^{-1}F(x)\Big],$$

triangular inequality and (4.2) we conclude that

$$\|y-x\| \leq \frac{-1}{f'(t)}(1+\theta)\left\|F'(x_0)^{-1}F(x)\right\| .$$

To end the proof of item 1, use the above inequality and the second inequality on (4.1).

Item 2 follows from triangular inequality, item 1 and the first and the third inequalities in (4.1).

Using the definition of the error (3.1) we have

$$F(y) = E_F(y,x)+F'(x_0)\Big[F'(x_0)^{-1}[F(x)+F'(x)(y-x)]\Big].$$

Therefore, using the triangle inequality, (4.2) and the second inequality on (4.1) we have

$$\begin{aligned}\|F'(x_0)^{-1}F(y)\| \leq& \|F'(x_0)^{-1}E_F(y,x)\|+\theta\left\|F'(x_0)^{-1}F(x)\right\| \\ \leq& \|F'(x_0)^{-1}E_F(y,x)\|+\theta(f(t)+\varepsilon).\end{aligned}$$

Using (4.1), item 1, and Lemma 3.2 with $s=-(1+\theta)(f(t)+\varepsilon)/f'(t)$ we have

$$\begin{aligned}\|F'(x_0)^{-1}E_F(y,x)\| \leq& e_f\left(t-(1+\theta)\frac{f(t)+\varepsilon}{f'(t)},t\right) \\ =& f\left(t-(1+\theta)\frac{f(t)+\varepsilon}{f'(t)}\right)+\varepsilon+\theta(f(t)+\varepsilon).\end{aligned}$$

Direct combination of the two above equation yields the latter inequality in item 3. □

In view of Lemma 4.1 define, for $\theta \geq 0$, the auxiliary map $n_\theta : [0,\bar{t}\,)\times[0,\infty)\to\mathbb{R}\times\mathbb{R}$,

$$n_\theta(t,\varepsilon) := \left(t-(1+\theta)\frac{f(t)+\varepsilon}{f'(t)},\ \varepsilon+2\theta(f(t)+\varepsilon)\right). \tag{4.3}$$

Let

$$\Omega := \{(t,\varepsilon)\in\mathbb{R}\times\mathbb{R}\ :\ 0\leq t<\lambda,\ \ 0\leq\varepsilon\leq\kappa t,\ \ 0<f(t)+\varepsilon\}. \tag{4.4}$$

Lemma 4.2. *If $0 \leq \theta \leq \Theta$, $(t, \varepsilon) \in \Omega$ and $(t_+, \varepsilon_+) = n_\theta(t, \varepsilon)$, that is,*

$$t_+ = t - (1+\theta)\frac{f(t)+\varepsilon}{f'(t)}, \qquad \varepsilon_+ = \varepsilon + 2\theta(f(t)+\varepsilon),$$

then $n_\theta(t,\varepsilon) \in \Omega$, $t < t_+$, $\varepsilon \leq \varepsilon_+$ and

$$f(t_+) + \varepsilon_+ < \left(\frac{1+\theta^2}{2}\right)(f(t)+\varepsilon).$$

Proof. Since $0 \leq t < \lambda$, according to (2.6) we have $f'(t) < -\kappa < 0$. Therefore $t < t_+$ and $\varepsilon \leq \varepsilon_+$.

As $\varepsilon \leq \kappa t$, $f(t) + \varepsilon > 0$ and $-1 \leq f'(t) < f'(t) + \kappa < 0$,

$$\begin{aligned} -\frac{f(t)+\varepsilon}{f'(t)} &\leq -\frac{f(t)+\kappa\, t}{f'(t)} \\ &= -\frac{f(t)+\kappa\, t}{f'(t)+\kappa}\left[1 + \frac{\kappa}{f'(t)}\right] \leq -\frac{f(t)+\kappa\, t}{f'(t)+\kappa}(1-\kappa). \end{aligned} \tag{4.5}$$

The function $h(s) := f(s) + \kappa s$ is differentiable at t, $h'(t) < 0$, is strictly convex and

$$\lim_{s\to\lambda_-} h(t) = 0.$$

Therefore, using Proposition 1.2 we have $t - h(t)/h'(t) < \lambda$, which is equivalent to

$$-\frac{f(t)+\kappa\, t}{f'(t)+\kappa} < \lambda - t. \tag{4.6}$$

Combining the above inequality with (4.5) and the definition of t_+ we conclude that

$$t_+ < t + (1+\theta)(1-\kappa)(\lambda - t).$$

Using (2.6) and (2.7) we have $(1+\theta)(1-\kappa) \leq 1-\theta < 1$, which combined with the above inequality yields $t_+ < \lambda$.

Using the definition of ε_+, inequality $\varepsilon \leq \kappa\, t$ and (2.6) we obtain

$$\begin{aligned} \varepsilon_+ &\leq 2\theta(f(t)+\varepsilon) + \kappa\, t \\ &= \kappa\left(t + (1+\theta)(f(t)+\varepsilon)\right). \end{aligned}$$

Using again the inequalities $f(t)+\varepsilon > 0$ and $-1 \leq f'(t) < 0$ we have

$$f(t) + \varepsilon \leq -\frac{f(t)+\varepsilon}{f'(t)}.$$

Combining the two above inequalities with the definition of t_+ we obtain $\varepsilon_+ \leq \kappa\, t_+$.

For proving the two last inequalities, first note that from the definition of the linearization error in (3.2) we have

$$\begin{aligned}
f(t_+) + \varepsilon_+ &= f(t) + f'(t)(t_+ - t) + e_f(t_+, t) + 2\theta(f(t) + \varepsilon) + \varepsilon \\
&= \theta(f(t) + \varepsilon) + e_f(t_+, t) \\
&= \theta(f(t) + \varepsilon) + \int_t^{t_+} \big(f'(u) - f'(t)\big)\, du.
\end{aligned}$$

Since f' is strictly increasing we conclude that the integral is positive. So, last equality implies that $f(t_+) + \varepsilon_+ \geq \theta(f(t) + \varepsilon) > 0$. Taking $s \in [t_+, \lambda)$ and using the convexity of f' we have

$$\begin{aligned}
\int_t^{t_+} \big(f'(u) - f'(t)\big)\, du &\leq \int_t^{t_+} \big(f'(s) - f'(t)\big)\, \frac{u - t}{s - t}\, du \\
&= \frac{1}{2} \frac{(t_+ - t)^2}{s - t} \big(f'(s) - f'(t)\big).
\end{aligned}$$

Substituting last inequality into above equation we have

$$\begin{aligned}
f(t_+) + \varepsilon_+ &\leq \theta(f(t) + \varepsilon) + \frac{1}{2} \frac{(t_+ - t)^2}{s - t} \big(f'(s) - f'(t)\big) \\
&= \left(\theta + \frac{1}{2} \frac{(1 + \theta)^2}{(s - t)} \frac{f(t) + \varepsilon}{-f'(t)} \frac{f'(s) - f'(t)}{-f'(t)}\right) (f(t) + \varepsilon).
\end{aligned}$$

On the other hand, because $f'(s) + \kappa < 0$ and $-1 \leq f'(t)$ it easy to conclude that

$$\frac{f'(s) - f'(t)}{-f'(t)} = \frac{f'(s) + \kappa - f'(t) - \kappa}{-f'(t)} \leq 1 - \kappa.$$

Combining last two above inequalities with (4.5), (4.6) and taking in account that $(1 + \theta)(1 - \kappa) \leq 1 - \theta$ we conclude that

$$\begin{aligned}
f(t_+) + \varepsilon_+ &\leq \left(\theta + \frac{1}{2}(1 + \theta)^2 (1 - \kappa)^2 \frac{\lambda - t}{s - t}\right) (f(t) + \varepsilon) \\
&= \left(\theta + \frac{1}{2}(1 - \theta)^2 \frac{\lambda - t}{s - t}\right) (f(t) + \varepsilon),
\end{aligned}$$

and the result follows taking the limit $s \to \lambda_-$. □

The outcome of an inexact Newton iteration is any point satisfying some error tolerance. Hence, instead of a mapping for Newton iteration, we shall deal with a *family* of mappings, describing all possible inexact iterations.

Definition 4.3. *For $0 \le \theta$, $\mathcal{N}_\theta$ is the family of maps $N_\theta : B(x_0, \bar t) \to X$ such that*

$$\left\|F'(x_0)^{-1}[F(x) + F'(x)(N_\theta(x) - x)]\right\| \le \theta \left\|F'(x_0)^{-1}F(x)\right\|, \tag{4.7}$$

for each $x \in B(x_0, \bar t\,)$.

If $x \in B(x_0, \bar t)$, then $F'(x)$ is non-singular. Therefore, for $\theta = 0$, the family $\mathcal{N}_0$ has a single element, namely the exact Newton iteration map

$$N_0 : B(x_0, \bar t) \to \mathbb{X}, \quad x \mapsto N_0(x) = x - F'(x)^{-1}F(x).$$

Trivially, if $0 \le \theta \le \theta'$ then $\mathcal{N}_0 \subset \mathcal{N}_\theta \subset \mathcal{N}_{\theta'}$. Hence, $\mathcal{N}_\theta$ is non-empty for all $\theta \ge 0$.

Remark 4.4. *For any $\theta \in (0, 1)$ and $N_\theta \in \mathcal{N}_\theta$*

$$N_\theta(x) = x \iff F(x) = 0, \qquad x \in B(x_0, \bar t).$$

This means that the fixed point of the inexact Newton iteration N_θ are the same fixed points of the exact *Newton iteration, namely, the zeros of F.*

The main tool for the analysis of the inexact Newton method with a relative residual tolerance will be a family of sets described below and analyzed in the ensuing proposition, which is a combination of Lemmas 4.1 and 4.2. Define

$$K(t, \varepsilon) := \left\{x \in X \ : \ \|x - x_0\| \le t, \ \left\|F'(x_0)^{-1}F(x)\right\| \le f(t) + \varepsilon\right\}, \tag{4.8}$$

and

$$K := \bigcup_{(t,\varepsilon)\in\Omega} K(t, \varepsilon). \tag{4.9}$$

Recall that n_θ, Ω and $\mathcal{N}_\theta$ were defined in (4.3), (4.4) and Definition 4.3 respectively.

Proposition 4.5. *Take $0 \le \theta \le \Theta$ and $N_\theta \in \mathcal{N}_\theta$. Then for any $(t, \varepsilon) \in \Omega$ and $x \in K(t, \varepsilon)$*

$$N_\theta(K(t, \varepsilon)) \subset K(n_\theta(t, \varepsilon)) \subset K, \qquad \|N_\theta(x) - x\| \le t_+ - t,$$

where t_+ is the first component of $n_\theta(t, \varepsilon)$. Moreover,

$$n_\theta\left(\Omega\right) \subset \Omega, \qquad N_\theta\left(K\right) \subset K. \tag{4.10}$$

Proof. Combine definitions (4.3), (4.4), Definition 4.3, (4.8), (4.9) with Lemmas 4.1 and 4.2. □

5 Convergence analysis

Theorem 5.1. *Take $0 \le \theta \le \Theta$ and $N_\theta \in \mathcal{N}_\theta$. For any $(t_0, \varepsilon_0) \in \Omega$ and $y_0 \in K(t_0, \varepsilon_0)$ the sequences*

$$y_{k+1} = N_\theta(y_k), \qquad (t_{k+1}, \varepsilon_{k+1}) = n_\theta(t_k, \varepsilon_k), \qquad k = 0, 1, \ldots, \tag{5.1}$$

are well defined,

$$y_k \in K(t_k, \varepsilon_k), \qquad (t_k, \varepsilon_k) \in \Omega \qquad k = 0, 1, \ldots, \tag{5.2}$$

the sequence $\{t_k\}$ is strictly increasing and converges to some $\tilde{t} \in (0, \lambda]$, the sequence $\{\varepsilon_k\}$ is non-decreasing and converges to some $\tilde{\varepsilon} \in [0, \kappa\lambda]$,

$$\|F'(x_0)^{-1}F(y_k)\| \le f(t_k) + \varepsilon_k \le \left(\frac{1+\theta^2}{2}\right)^k (f(t_0) + \varepsilon_0), \quad k = 0, 1, \ldots. \tag{5.3}$$

the sequence $\{y_k\}$ is contained in $B(x_0, \lambda)$ and converges to a point $x_* \in B[x_0, t_*]$ which is the unique zero of F in $B(x_0, \bar{\tau})$ and

$$\|y_{k+1} - y_k\| \le t_{k+1} - t_k, \qquad \|x_* - y_k\| \le \tilde{t} - t_k, \qquad k = 0, 1, \ldots. \tag{5.4}$$

Moreover, if

h4') *$\lambda < R$,*

then the sequence $\{y_k\}$ satisfies, for $k = 0, 1, \ldots$

$$\|x_* - y_{k+1}\| \le \left[\frac{1+\theta}{2}\frac{D^- f'(\lambda)}{|f'(\lambda)|}\|x_* - y_k\| + \theta\,\frac{2 + f'(\lambda)}{|f'(\lambda)|}\right]\|x_* - y_k\|. \tag{5.5}$$

If, additionally, $0 \le \theta < \kappa/(4+\kappa)$ then $\{y_k\}$ converges Q-linearly as follows

$$\|x_* - y_{k+1}\| \le \left[\frac{1+\theta}{2} + \frac{2\theta}{\kappa}\right]\|x_* - y_k\|, \qquad k = 0, 1, \ldots. \tag{5.6}$$

Proof. Well definition of the sequences $\{(t_k, \varepsilon_k)\}$ and $\{y_k\}$ as defined in (5.1) follows from the assumptions on θ, (t_0, ε_0), y_0 and the two last inclusions on Proposition 4.5. Moreover, since (5.2) holds for $k = 0$, using the first inclusion in Proposition 4.5 and induction on k, we conclude that (5.2) holds for all k. The first inequality in (5.4) now follows from Proposition 4.5, (5.2) and (5.1) while the first inequality in (5.3) follows from (5.2) and the definition of $K(t, \varepsilon)$ in (4.8).

Direct inspection of the definition of Ω in (4.4) shows that

$$\Omega \subset [0, \lambda) \times [0, \kappa\lambda).$$

Therefore, using (5.2) and the definition of $K(t, \varepsilon)$ we have

$$t_k \in [0, \lambda), \quad \varepsilon_k \in [0, \kappa\lambda), \quad y_k \in B(x_0, \lambda), \qquad k = 0, 1, \ldots$$

Using (4.4) and Lemma 4.2 we conclude that $\{t_k\}$ is strictly increasing, $\{\varepsilon_k\}$ is non-decreasing and the second equality in (5.3) holds for all k. Therefore, in view of the first two above inclusions, $\{t_k\}$ and $\{\varepsilon_k\}$ converge, respectively, to some $\tilde{t} \in (0, \lambda]$ and $\tilde{\varepsilon} \in [0, \kappa\lambda]$. Convergence to $\tilde{t}$, together with the first inequality in (5.4) and the inclusion $y_k \in B(x_0, \lambda)$ implies that y_k converges to some $x_* \in B[0, \lambda]$ and that the second inequality on (5.4) holds for all k.

Using the inclusion $y_k \in B(x_0, \lambda)$, the first inequality in Corollary 3.5 and (5.3) we have

$$-f(\|y_k - x_0\|) \leq \left(\frac{1+\theta^2}{2}\right)^k (f(t_0) + \varepsilon_0), \quad k = 0, 1, \ldots.$$

According to (2.8), $f' < -\kappa$ in $[0, \lambda)$. Therefore, since $f(t_*) = 0$ and $t_* < \lambda$,

$$f(t) \leq -\kappa(t - t_*), \qquad t_* \leq t < \lambda.$$

Hence, *if* $\|y_k - x_0\| \geq t_*$, we can combine the two above inequalities, setting $t = \|y_k - x_0\|$ in the second, to obtain

$$\|y_k - x_0\| - t_* \leq \left(\frac{1+\theta^2}{2}\right)^k \frac{f(t_0) + \varepsilon_0}{\kappa}.$$

Note that the above inequality remain valid even if $\|y_k - x_0\| < t_*$. Therefore, taking the limit $k \to \infty$ in the above inequality we conclude that $\|x_* - x_0\| \leq t_*$. Moreover, now that we know that x_* is in the interior of the domain of F, we can also take the limit $k \to \infty$ in (5.3) to conclude that $F(x_*) = 0$.

The "classical" version of Kantorovich's theorem on Newton's method for a generic majorant function (see e.g. [6]) guarantee that under the assumptions of Theorem 2.1, F has a unique zero in $B(x_0, \bar{\tau})$. Hence x_* must be this zero of F.

To prove (5.5) and (5.6) , first note that from first inclusion in (5.2) we have $\|y_k - x_0\| \leq t_k$, for all $k = 0, 1, \ldots$. Now, since $\tilde{t} \in (0, \lambda]$ we obtain from second inequality in (5.4) that $\|x_* - y_k\| \leq \lambda - t_k$, for all $k = 0, 1, \ldots$. Therefore, using **h4'**, $F(x_*) = 0$ and and first equality in (5.1), the desire inequalities follows by applying Lemma 3.7. For concluding the proof, note that for $0 \leq \theta < \kappa/(4 + \kappa)$ the quantity in the bracket in (5.6) is less than one, which implies that the sequence $\{y_k\}$ converges Q-linearly. □

Proposition 5.2. *If* $0 \leq \rho < \beta/2$ *then*

$$\rho < \bar{t}/2 < \bar{t}, \qquad f'(\rho) < 0.$$

Proof. Assumption $\rho < \beta/2$ and Proposition 2.3 item iii proves the first two inequalities of the proposition. The last inequality follows from the first inequality and Proposition 2.3 item i. □

Proof of Theorem 2.1. First we will prove Theorem 2.1 with $\rho = 0$ and $z_0 = x_0$. Note that, from the definition in (2.6), we have

$$\kappa_0 = \kappa, \quad \lambda_0 = \lambda, \quad \Theta_0 = \Theta.$$

Since

$$(0,0) \in \Omega, \qquad x_0 \in K(0,0),$$

using Theorem 5.1 we conclude that Theorem 2.1 holds for $\rho = 0$.

For proving the general case, take

$$0 \le \rho < \beta/2, \qquad z_0 \in B[x_0, \rho] \ . \tag{5.7}$$

Using Proposition 5.2 and (2.5) we conclude that $\rho < \bar{t}/2$ and $f'(\rho) < 0$. Define

$$g : [0, R-\rho) \to \mathbb{R}, \quad g(t) = \frac{-1}{f'(\rho)}[f(t+\rho) + 2\rho]. \tag{5.8}$$

We claim that g is a majorant function for F at point z_0. Trivially, $B(z_0, R-\rho) \subset C$, $g'(0) = -1$, $g(0) > 0$. Moreover g' is also convex and strictly increasing. To end the proof that g satisfies **h1**, **h2** and **h3**, using Proposition 2.3 item iii and second inequality in (5.7) we have

$$\lim_{t \to \bar{t}-\rho} g(t) = \frac{-1}{f'(\rho)}(2\rho - \beta) < 0 \ .$$

Using Proposition 3.1 we have

$$\left\| F'(z_0)^{-1} F'(x_0) \right\| \le \frac{-1}{f'(\rho)} \ . \tag{5.9}$$

Therefore, using also the second inequality of Corollary 3.5 we have

$$\begin{aligned} \left\| F'(z_0)^{-1} F(z_0) \right\| &\le \left\| F'(z_0)^{-1} F'(x_0) \right\| \left\| F'(x_0)^{-1} F(z_0) \right\| \\ &\le \frac{-1}{f'(\rho)}[f(\|z_0 - x_0\|) + 2\|z_0 - x_0\|]. \end{aligned}$$

As $f' \ge -1$, the function $t \mapsto f(t) + 2t$ is (strictly) increasing. Combining this fact with the above inequality and (5.8) we conclude that

$$\left\| F'(z_0)^{-1} F'(z_0) \right\| \le g(0).$$

To end the proof that g is a majorant function for F at z_0, take $x, y \in \mathbb{X}$ such that

$$x, y \in B(z_0, R-\rho), \qquad \|x - z_0\| + \|y - x\| < R - \rho \ .$$

Hence $x, y \in B(x_0, R)$, $\|x - x_0\| + \|y - x\| < R$ and using (5.9) together with (2.1) we have

$$\begin{aligned} \|F'(z_0)^{-1}\left[F'(y) - F'(x)\right]\| &\le \|F'(z_0)^{-1}F'(x_0)\| \|F'(x_0)^{-1}\left[F'(y) - F'(x)\right]\| \\ &\le \frac{-1}{f'(\rho)}\left[f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|)\right] . \end{aligned}$$

Since f' is convex, the function $t \mapsto f'(s+t) - f'(s)$ is increasing for $s \geq 0$ and $\|x - x_0\| \leq \|x - z_0\| + \|z_0 - x_0\| \leq \|x - z_0\| + \rho$,

$$f'(\|y-x\| + \|x-x_0\|) - f'(\|x-x_0\|) \leq \\ f'(\|y-x\| + \|x-z_0\| + \rho) - f'(\|x-z_0\| + \rho).$$

Combining the two above inequalities with the definition of g we obtain

$$\|F'(z_0)^{-1}\left[F'(y) - F'(x)\right]\| \leq g'(\|y-x\| + \|x-z_0\|) - g'(\|x-z_0\|).$$

Note that for κ_ρ, λ_ρ and Θ_ρ as defined in (2.3), we have

$$\kappa_\rho = \sup_{0<t<R-\rho} \frac{-g(t)}{t}, \quad \lambda_\rho = \sup\{t \in [0, R-\rho) : \kappa_\rho + g'(t) < 0\}, \quad \Theta_\rho = \frac{\kappa_\rho}{2-\kappa_\rho},$$

which are the same as (2.3) with g instead of f. Therefore, applying Theorem 2.1 for F and the majorant function g at point z_0 and $\rho = 0$, we conclude that the sequence $\{z_k\}$ is well defined, remains in $B(z_0, \lambda_\rho)$, satisfies (2.4) and converges to some $z_* \in B[z_0, t_{*,\rho}]$ which is a zero of F, where $t_{*,\rho}$ is the smallest solution of $g(t) = 0$. Using (5.8) we conclude that $t_{*,\rho}$ is the smallest solution of

$$f(\rho + t) + 2\rho = 0.$$

Hence, in view of Proposition 2.3 item ii, we have $\rho + t_{*,\rho} < \bar{t} \leq \bar{\tau}$. and $B[(z_0, t_{*,\rho}] \subset B(x_0, \bar{\tau})$. Therefore, z_* is the unique zero of F in $B(x_0, \bar{\tau})$, which we already called x_*. Since

$$g'(t) = f'(t+\rho)/|f'(\rho)|, \;\; D^- g'(t) = D^- f'(t+\rho)/|f'(\rho)|, \qquad t \in [0, R-\rho),$$

applying again Theorem 2.1 for F and the majorant function g at point z_0 and $\rho = 0$, we conclude that item **h4** also holds. □

6 Special cases

First we use Theorem 2.1 to analyze the convergence of the inexact Newton method with a relative residual error tolerance in the setting of Smale's α-theory. Up to our knowledge, this is the first time an inexact Newton method with a relative error tolerance is analyzed in this framework.

Theorem 6.1. *Let $\mathbb{X}$ and $\mathbb{Y}$ be Banach spaces, $C \subseteq \mathbb{X}$ and $F : C \to \mathbb{Y}$ a continuous function and analytic int(C). Take $x_0 \in \mathrm{int}(C)$ with $F'(x_0)$ non-singular. Define*

$$\gamma := \sup_{n>1} \left\| \frac{F'(x_0)^{-1} F^{(n)}(x_0)}{n!} \right\|^{1/(n-1)}.$$

Suppose that $B(x_0, 1/\gamma) \subset C$, $b > 0$ and that

$$\|F'(x_0)^{-1}F(x_0)\| \leq b, \qquad b\gamma < 3 - 2\sqrt{2}, \qquad 0 \leq \theta \leq \frac{1 - 2\sqrt{\gamma b} - \gamma b}{1 + 2\sqrt{\gamma b} + \gamma b}.$$

Then, the sequence generated by the inexact Newton method for solving $F(x) = 0$ with starting point x_0 and residual relative error tolerance θ: For $k = 0, 1, \ldots,$

$$x_{k+1} = x_k + S_k, \quad \left\|F'(x_0)^{-1}\left[F(x_k) + F'(x_k)S_k\right]\right\| \leq \theta\|F'(x_0)^{-1}F(x_k)\|,$$

is well defined, the generated sequence $\{x_k\}$ converges to a point x_* which is a zero of F,

$$\|F'(x_0)^{-1}F(x_k)\| \leq \left(\frac{1+\theta^2}{2}\right)^k b, \qquad k = 0, 1, \ldots,$$

the sequence $\{x_k\}$ is contained in $B(x_0, \lambda)$, $x_* \in B[x_0, t_*]$ and x_* is the unique zero of F in $B(x_0, \bar{\tau})$, where

$$\lambda := \frac{b}{\sqrt{\gamma b} + \gamma b},$$

$$t_* = \frac{1 + \gamma b - \sqrt{1 - 6\gamma b + (\gamma b)^2}}{4}, \qquad \bar{\tau} = \frac{1 + \gamma b + \sqrt{1 - 6\gamma b + (\gamma b)^2}}{4}.$$

Moreover, the sequence $\{x_k\}$ satisfies, for $k = 0, 1, \ldots,$

$$\|x_* - x_{k+1}\| \leq \left[\frac{1+\theta}{2}\frac{D^- f'(\lambda)}{|f'(\lambda)|}\|x_* - x_k\| + \theta\,\frac{f'(\lambda) + 2}{|f'(\lambda)|}\right]\|x_* - x_k\|.$$

If, additionally, $0 \leq \theta < (1 - 2\sqrt{\gamma b} - \gamma b)/(5 - 2\sqrt{\gamma b} - \gamma b)$ then $\{x_k\}$ converges Q-linearly as follows

$$\|x_* - x_{k+1}\| \leq \left[\frac{1+\theta}{2} + \frac{2\theta}{1 - 2\sqrt{\gamma b} - \gamma b}\right]\|x_* - x_k\|, \qquad k = 0, 1, \ldots.$$

Proof. Since the function $f : [0, 1/\gamma) \to \mathbb{R}$

$$f(t) = \frac{t}{1 - \gamma t} - 2t + b,$$

is a majorant function for F in x_0, [6]. Therefore, all results follow from Theorem 2.1, applied to this particular context. □

A semi-local convergence result for Newton method is instrumental in the complexity analysis of linear and quadratic minimization problems by means of self-concordant functions [15]. Also in this setting, Theorem 2.1 provides a semi-local convergence result for Newton method with a relative error tolerance.

Theorem 6.2. *Let $C \subseteq \mathbb{R}^n$ be an open convex set and let $g : C \to \mathbb{R}$ be an a-self-concordant function with parameter $a > 0$. For $x \in C$, let*

$$\|v\|_x := \sqrt{v^T g''(x) v}, \quad v \in \mathbb{R}^n,$$

$$W_r(x) := \{z : \|z - x\|_x < r\}, \quad W_r[x] := \{z : \|z - x\|_x \le r\}.$$

Suppose that $x_0 \in C$, $g''(x_0)$ is non-singular, $b > 0$

$$\|g''(x_0)^{-1} g'(x_0)\|_{x_0} \le b < 3 - 2\sqrt{2}, \qquad 0 \le \theta \le \frac{1 - 2\sqrt{b} - b}{1 + 2\sqrt{b} + b}.$$

Then the sequence generated by the inexact Newton method for solving $g'(x) = 0$ with starting point x_0 and residual relative error tolerance θ: For $k = 0, 1, \ldots,$

$$x_{k+1} = x_k + S_k, \quad \left\| g''(x_0)^{-1} \left[g'(x_k) + g''(x_k) S_k \right] \right\|_{x_0} \le \theta \|g''(x_0)^{-1} g'(x_k)\|_{x_0},$$

is well defined, converges to a point x_* which is the (unique, global) minimizer of g,

$$\|g''(x_0)^{-1} g'(x_k)\|_{x_0} \le \left(\frac{1 + \theta^2}{2} \right)^k b, \qquad k = 0, 1, \ldots,$$

the sequence $\{x_k\}$ is contained in $W_\lambda(x_0)$ and $x_* \in W_{t_*}(x_0)$, where

$$\lambda := \frac{b}{\sqrt{b} + b}, \qquad t_* = \frac{1 + \ b - \sqrt{1 - 6\ b + b^2}}{4}.$$

Moreover, the sequence $\{x_k\}$ satisfies, for $k = 0, 1, \ldots,$

$$\|x_* - x_{k+1}\| \le \left[\frac{1 + \theta}{2} \frac{D^- f'(\lambda)}{|f'(\lambda)|} \|x_* - x_k\| + \theta \, \frac{f'(\lambda) + 2}{|f'(\lambda)|} \right] \|x_* - x_k\|.$$

If, additionally, $0 \le \theta < (1 - 2\sqrt{b} - b)/(5 - 2\sqrt{b} - b)$ then $\{x_k\}$ converges Q-linearly as follows

$$\|x_* - x_{k+1}\| \le \left[\frac{1 + \theta}{2} + \frac{2\theta}{1 - 2\sqrt{b} - b} \right] \|x_* - x_k\|, \qquad k = 0, 1, \ldots.$$

Proof. The scalar function $f : [0, 1) \to \mathbb{R}$ defined by

$$f(t) = \frac{t}{1 - t} - 2t + b,$$

is a majorant function for g' in x_0 , [6]. Therefore, the proof follows from Theorem 2.1, applied to this particular context. □

Theorem 6.3. *Let $\mathbb{X}$ and $\mathbb{Y}$ be a Banach spaces, $C \subseteq \mathbb{X}$ and $F : C \to \mathbb{Y}$ a continuous function, continuously differentiable on int(C). Take $x_0 \in \mathrm{int}(C)$ with $F'(x_0)$ non-singular. Suppose that exist constants $L > 0$ and $b > 0$ such that $bL < 1/2$, $B(x_0, 1/L) \subset C$ and*

$$\left\|F'(x_0)^{-1}\left[F'(y) - F'(x)\right]\right\| \leq L\|x - y\|, \qquad x,\, y \in B(x_0, 1/L),$$

$$\|F'(x_0)^{-1}F(x_0)\| \leq b, \qquad 0 \leq \theta \leq \frac{1-\sqrt{2bL}}{1+\sqrt{2bL}}$$

Then, the sequence generated by the inexact Newton method for solving $F(x) = 0$ with starting point x_0 and residual relative error tolerance θ: For $k = 0, 1, \ldots,$

$$x_{k+1} = x_k + S_k, \quad \left\|F'(x_0)^{-1}\left[F(x_k) + F'(x_k)S_k\right]\right\| \leq \theta\|F'(x_0)^{-1}F(x_k)\|,$$

is well defined,

$$\|F'(x_0)^{-1}F(x_k)\| \leq \left(\frac{1+\theta^2}{2}\right)^k b, \qquad k = 0, 1, \ldots .$$

the sequence $\{x_k\}$ is contained in $B(x_0, \lambda)$, converges to a point $x_* \in B[x_0, t_*]$ which is the unique zero of F in $B(x_0, 1/L)$ where

$$\lambda := \frac{\sqrt{2bL}}{L}, \quad t_* = \frac{1-\sqrt{1-2Lb}}{L}.$$

Moreover, the sequence $\{x_k\}$ satisfies, for $k = 0, 1, \ldots,$

$$\|x_* - z_{k+1}\| \leq \left[\frac{1+\theta}{2}\frac{L}{1-\sqrt{2bL}}\|x_* - x_k\| + \theta\,\frac{1+\sqrt{2bL}}{1-\sqrt{2bL}}\right]\|x_* - x_k\|.$$

If, additionally, $0 \leq \theta < (1-\sqrt{2bL})/(5-\sqrt{2bL})$ then the sequence $\{x_k\}$ converges Q-linearly as follows

$$\|x_* - x_{k+1}\| \leq \left[\frac{1+\theta}{2} + \frac{2\theta}{1-\sqrt{2bL}}\right]\|x_* - x_k\|, \qquad k = 0, 1, \ldots .$$

Proof. Since the function $f : [0, 1/L) \to \mathbb{R}$,

$$f(t) := \frac{L}{2}\,t^2 - t + b,$$

is a majorant function for F at point x_0, all result follow from Theorem 2.1, applied to this particular context. □

References

[1] F. Alvarez, J. Bolte, and J. Munier. A unifying local convergence result for Newton's method in Riemannian manifolds. *Found. Comput. Math.*, 8(2):197–226, 2008.

[2] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and real computation.* Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp.

[3] J. Chen and W. Li. Convergence behaviour of inexact Newton methods under weak Lipschitz condition. *J. Comput. Appl. Math.*, 191(1):143–164, 2006.

[4] R. S. Dembo, S. C. Eisenstat, and T. Steihaug. Inexact Newton methods. *SIAM J. Numer. Anal.*, 19(2):400–408, 1982.

[5] O. P. Ferreira and B. F. Svaiter. Kantorovich's theorem on Newton's method in Riemannian manifolds. *J. Complexity*, 18(1):304–329, 2002.

[6] O. P. Ferreira and B. F. Svaiter. Kantorovich's majorants principle for Newton's method. *Comput. Optim. Appl.*, 42(2):213–229, 2009.

[7] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms. I*, volume 305 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 1993. Fundamentals.

[8] L. V. Kantorovič. The principle of the majorant and Newton's method. *Doklady Akad. Nauk SSSR (N.S.)*, 76:17–20, 1951.

[9] L. V. Kantorovich and G. P. Akilov. *Functional analysis in normed spaces.* Translated from the Russian by D. E. Brown. Edited by A. P. Robertson. International Series of Monographs in Pure and Applied Mathematics, Vol. 46. The Macmillan Co., New York, 1964.

[10] C. Li and J. Wang. Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition. *IMA J. Numer. Anal.*, 26(2):228–251, 2006.

[11] C. Li and J. Wang. Newton's method for sections on Riemannian manifolds: generalized covariant α-theory. *J. Complexity*, 24(3):423–451, 2008.

[12] C. Li, J.-H. Wang, and J.-P. Dedieu. Smale's point estimate theory for Newton's method on Lie groups. *J. Complexity*, 25(2):128–151, 2009.

[13] I. Moret. A Kantorovich-type theorem for inexact Newton methods. *Numer. Funct. Anal. Optim.*, 10(3-4):351–365, 1989.

[14] B. Morini. Convergence behaviour of inexact Newton methods. *Math. Comp.*, 68(228):1605–1613, 1999.

[15] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*, volume 13 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[16] F. A. Potra. The Kantorovich Theorem and interior point methods. *Math. Program.*, 102(1, Ser. A):47–70, 2005.

[17] W. Shen and C. Li. Kantorovich-type convergence criterion for inexact Newton methods. *Appl. Numer. Math.*, 59(7):1599–1611, 2009.

[18] S. Smale. Newton's method estimates from data at one point. In *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*, pages 185–196. Springer, New York, 1986.

[19] J.-H. Wang, S. Huang, and C. Li. Extended Newton's method for mappings on Riemannian manifolds with values in a cone. *Taiwanese J. Math.*, 13(2B):633–656, 2009.

[20] J.-h. Wang and C. Li. Uniqueness of the singular points of vector fields on Riemannian manifolds under the γ-condition. *J. Complexity*, 22(4):533–548, 2006.

[21] J.-H. Wang and C. Li. Kantorovich's theorem for newton's method on lie groups. *Journal of Zhejiang University: Science A*, 8(6):978–986, 2007. cited By (since 1996) 0.

[22] T. J. Ypma. Local convergence of inexact Newton methods. *SIAM J. Numer. Anal.*, 21(3):583–590, 1984.