

Accepted Manuscript

Starting points for Newton's method under a center Lipschitz condition for the second derivative

J.A. Ezquerro, M.A. Hernández-Verón, Á.A. Magreñán

PII: S0377-0427(16)30623-9

DOI: <http://dx.doi.org/10.1016/j.cam.2016.12.013>

Reference: CAM 10937

To appear in: *Journal of Computational and Applied Mathematics*

Received date: 12 September 2016

Revised date: 21 November 2016

Please cite this article as: J.A. Ezquerro, M.A. Hernández-Verón, Á.A. Magreñán, Starting points for Newton's method under a center Lipschitz condition for the second derivative, *Journal of Computational and Applied Mathematics* (2016), <http://dx.doi.org/10.1016/j.cam.2016.12.013>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Starting points for Newton's method under a center Lipschitz condition for the second derivative

J. A. Ezquerro ^a, M. A. Hernández-Verón ^a and Á. A. Magreñán ^b

^a *Department of Mathematics and Computation, University of La Rioja.*

Calle Luis de Ulloa s/n, 26004 Logroño, Spain.

E-mail address: <jezquer><mahernan>@unirioja.es

^b *Escuela Superior de Ingeniería y Tecnología, Universidad Internacional de La Rioja.*

Avenida Gran Vía Rey Juan Carlos I, 41, 26002 Logroño, Spain.

E-mail address: alberto.magrenan@unir.net

Abstract

We analyse the semilocal convergence of Newton's method under a center Lipschitz condition for the second derivative of the operator involved different from that used by other authors until now. In particular, we propose to center the Lipschitz condition for the second derivative in a different point from that where Newton's method starts. This allows us to obtain different starting points for Newton's method and modify the domain of starting points.

Keywords: Newton's method, semilocal convergence, majorizing sequence, error estimates, order of convergence, region of accessibility, integral equation.

2000 Mathematics Subject Classification: 47H99, 65H10, 65J15.

This work has been partially supported by the project MTM2014-52016-C2-1-P of Spanish Ministry of Economy and Competitiveness and the International University of La Rioja, under the Plan Propio de Investigación, Desarrollo e Innovación (2015–2017, research group: Modelación matemática aplicada a la ingeniería (MOMAIN) and the grant SENECA 19374/PI/14.

1 Introduction

By using mathematical modelling, many problems from computational sciences and other disciplines can be brought in the form of the equation $F(x) = 0$, where F is a nonlinear

operator defined on a nonempty open convex subset Ω of a Banach space X with values in a Banach space Y . As the solutions of these equations can rarely be found in closed form, we usually look for numerical approximations of these solutions. That is why the solution methods for these equations are iterative. For this, starting from one initial approximation x_0 of a solution x^* of the equation $F(x) = 0$, a sequence $\{x_n\}$ of approximations is constructed such that $\|x_{n+1} - x^*\| < \|x_n - x^*\|$, $n \geq 0$, that leads to the sequence $\{x_n\}$ converges to the solution x^* .

The study about convergence matter of iterative procedures is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iterative procedure; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls.

It is well-known that Newton's method,

$$x_0 \in \Omega, \quad x_{n+1} = x_n - [F'(x_n)]^{-1}F(x_n), \quad n \geq 0, \quad (1)$$

is the one of the most used iterative methods to approximate the solution x^* of $F(x) = 0$. We analyse the semilocal convergence of the method from conditions on the starting point x_0 and the operator F , along with a condition that connects the previous conditions. The first semilocal convergence result for Newton's method in Banach spaces was given by Kantorovich [11] under the following conditions:

(A1) There exists $\Gamma_0 = [F'(x_0)]^{-1} \in \mathcal{L}(Y, X)$, for some $x_0 \in \Omega$, with $\|\Gamma_0\| \leq \beta$ and $\|\Gamma_0 F(x_0)\| \leq \eta$, where $\mathcal{L}(Y, X)$ is the set of bounded linear operators from Y to X .

(A2) $\|F''(x)\| \leq M$ for $x \in \Omega$.

Theorem 1 (The Newton-Kantorovich theorem, [11]). *Let $F : \Omega \subseteq X \rightarrow Y$ be a twice continuously differentiable operator defined on a non-empty open convex domain Ω of a Banach space X with values in a Banach space Y . Suppose that conditions (A1) and (A2) are satisfied. If $h = M\beta\eta \leq \frac{1}{2}$ and $B(x_0, s^*) \subset \Omega$, where $s^* = \frac{1-\sqrt{1-2h}}{h}\eta$, then Newton's sequence, given by (1) and starting at x_0 , converges to a solution x^* of the equation $F(x) = 0$ and $x_n, x^* \in B(x_0, s^*)$, for all $n \in \mathbb{N}$.*

A few years later, Ortega observes that condition (A2) implies that F' is Lipschitz continuous in Ω and presents in [12] a variant of the result given by Kantorovich where (A2) is replaced by condition:

$$\|F'(x) - F'(y)\| \leq K\|x - y\| \quad \text{for } x, y \in \Omega.$$

A little later, in [7], condition (A2) is relaxed by using this condition "centered" in the starting point x_0 :

$$\|F'(x) - F'(x_0)\| \leq K_0\|x - x_0\| \quad \text{for } x \in \Omega, \quad (2)$$

what is known as center Lipschitz condition for F' at x_0 .

The use of the previous condition instead of condition (A2) leads to condition $h = M\beta\eta \leq \frac{1}{2}$ is replaced by another more restrictive: $K_0\beta\eta \leq \frac{14-4\sqrt{6}}{25} = 0.1680816\dots$. This implies that the domain of starting points for Newton's method is modified, since we can find starting points x_0 such that the last condition is satisfied (observe that $K_0 \leq K$), but the second-to-last condition is not. Besides, we notice that the point x_0 satisfies the conditions of semilocal convergence or there is no possibility of choosing another starting point. As a consequence, the domain of starting points for Newton's method consists of a single point, x_0 , or is an empty set and Newton's method is never convergent. To avoid this problem that presents the last condition, we use in [5] a center condition for the first derivative on an auxiliary point and obtain a domain of starting points which is not reduced to a point or to the empty set, since a nonempty set of possible starting points is found.

On the other hand, Huang proposes in [8] an alternative, that does not consist of relaxing the condition on the operator F , and imposes a condition on F that leads to a modification, not a restriction, of the domain of starting points. In particular, Huang proposes that F'' is Lipschitz continuous in Ω . But, if we pay attention to the proof of Huang, we see that it is not necessary that F'' is Lipschitz continuous in the entire domain Ω , since it is enough that F'' is center Lipschitz continuous at x_0 (see [6]). So, the semilocal convergence of Newton's method can be proved under the following conditions:

(B1) There exists $\Gamma_0 = [F'(x_0)]^{-1} \in \mathcal{L}(Y, X)$, for some $x_0 \in \Omega$, with $\|\Gamma_0\| \leq \beta$ and $\|\Gamma_0 F(x_0)\| \leq \eta$; moreover, $\|F''(x_0)\| \leq M_0$.

(B2) $\|F''(x) - F''(x_0)\| \leq L\|x - x_0\|$ for $x \in \Omega$.

Theorem 2. *Let $F : \Omega \subseteq X \rightarrow Y$ be a twice continuously differentiable operator defined on a nonempty open convex domain Ω of a Banach space X with values in a Banach space Y . Suppose that conditions (B1) and (B2) are satisfied. If $6M_0^3\beta^3\eta + 9L^2\beta^2\eta^2 + 18LM_0\beta^2\eta - 3M_0^2\beta^2 - 8L\beta \leq 0$ and $B(x_0, r^*) \subset \Omega$, where r^* is the smallest positive zero of the polynomial $\phi(t) = \frac{L}{6}t^3 + \frac{M_0}{2}t^2 - \frac{t}{\beta} + \frac{\eta}{\beta}$, then Newton's sequence, given by (1) and starting at x_0 , converges to a solution x^* of $F(x) = 0$ and $x_n, x^* \in \overline{B(x_0, r^*)}$, for all $n \in \mathbb{N}$.*

Observe that Huang changes the Lipschitz condition on the operator F' given by (2) for a Lipschitz condition on the operator F'' given by (B2). Obviously, condition (B2) limits the number of operator equations that can be solved by applying Newton's method.

In this work, following the idea mentioned above, we propose to use condition (B2), but centered at a different point from the starting point x_0 of Newton's method, so that we modify condition (B2) by centering it in the following way:

$$\|F''(x) - F''(\tilde{x})\| \leq \tilde{L}\|x - \tilde{x}\|, \quad \text{for } x \in \Omega, \quad (3)$$

once the point $\tilde{x} \in \Omega$ is fixed. This modification leads to a modification in the domain of starting points of Newton's method.

Throughout the paper we denote $\overline{B(x, \varrho)} = \{y \in X; \|y - x\| \leq \varrho\}$ and $B(x, \varrho) = \{y \in X; \|y - x\| < \varrho\}$.

2 Motivation

If the operator F'' is Lipschitz continuous in a domain, it is obvious that F'' is Lipschitz continuous at every point of the domain. Now, we see that there are situations where F'' is only Lipschitz continuous in some points of a domain, but not in all the domain, what happens, for example, with the following nonlinear integral equation of type Hammerstein-type:

$$x(s) = u(s) + \lambda \int_a^b \mathcal{K}(s, t) x(t)^{2+\frac{1}{n}} dt, \quad s \in [a, b], \quad \lambda \in \mathbb{R}, \quad n \in \mathbb{N}, \quad (4)$$

where $-\infty < a < b < +\infty$, $u(s)$ is a given continuous function on $[a, b]$ and such that $u(s) > 0$, the kernel $\mathcal{K}(s, t)$ is continuous and positive in $[a, b] \times [a, b]$ and $x(s)$ is a function to determine. Integral equations of this kind are treated in [14].

Solving equation (4) is equivalent to solving the equation $\mathcal{F}(x) = 0$, where $\mathcal{F} : \Omega \subseteq \mathcal{C}([a, b]) \longrightarrow \mathcal{C}([a, b])$ and

$$[\mathcal{F}(x)](s) = x(s) - u(s) - \lambda \int_a^b \mathcal{K}(s, t) x(t)^{2+\frac{1}{n}} dt, \quad s \in [a, b], \quad \lambda \in \mathbb{R}, \quad n \in \mathbb{N}.$$

In this case, we consider $\Omega = \mathcal{C}([a, b])$ and have

$$[\mathcal{F}'(x)y](s) = y(s) - \lambda \left(2 + \frac{1}{n}\right) \int_a^b \mathcal{K}(s, t) x(t)^{1+\frac{1}{n}} y(t) dt,$$

$$[\mathcal{F}''(x)(yz)](s) = -\lambda \left(2 + \frac{1}{n}\right) \left(1 + \frac{1}{n}\right) \int_a^b \mathcal{K}(s, t) x(t)^{\frac{1}{n}} z(t) y(t) dt.$$

We then note that \mathcal{F}'' is not Lipschitz continuous in all Ω . For this, we consider $[a, b] = [0, 1]$, $\mathcal{K}(s, t) = 1$ and $\ell(t) = 0$. Then, $[\mathcal{F}''(\ell)(yz)](s) = 0$ and

$$\|\mathcal{F}''(x) - \mathcal{F}''(\ell)\| = |\lambda| \left(2 + \frac{1}{n}\right) \left(1 + \frac{1}{n}\right) \int_0^1 x(t)^{\frac{1}{n}} dt.$$

If \mathcal{F}'' were Lipschitz continuous in Ω , then

$$\|\mathcal{F}''(x) - \mathcal{F}''(\ell)\| \leq L_1 \|x - \ell\|,$$

or equivalently, the inequality

$$\int_0^1 x(t)^{\frac{1}{n}} dt \leq L_2 \max_{s \in [0, 1]} |x(s)|, \quad (5)$$

were satisfied for all $x \in \Omega$ and a constant L_2 . But this is not true, as we can see in the following. Indeed, if we consider the functions

$$x_i(t) = \frac{t}{i}, \quad i \geq 1, \quad t \in [0, 1],$$

and they are substituted into (5),

$$\frac{1}{i^{\frac{1}{n}}(1 + \frac{1}{n})} \leq \frac{L_2}{i} \iff i^{1-\frac{1}{n}} \leq L_2 \left(1 + \frac{1}{n}\right), \quad \text{for all } i \geq 1,$$

then inequality (5) is not true when $i \rightarrow +\infty$.

However, \mathcal{F}'' is Lipschitz continuous at a point of Ω . For this, we consider $x_0(t) = u(t)$, the max-norm and $d = \min_{s \in [a,b]} u(s)$, ($d > 0$). Then, for $y, z \in \Omega$, we have

$$\begin{aligned} \|[\mathcal{F}''(x) - \mathcal{F}''(x_0)](yz)\| &= |\lambda| \left(2 + \frac{1}{n}\right) \left(1 + \frac{1}{n}\right) \max_{s \in [a,b]} \left| \int_a^b \mathcal{K}(s, t) \left(x(t)^{\frac{1}{n}} - u(t)^{\frac{1}{n}}\right) z(t)y(t) dt \right| \\ &\leq |\lambda| \left(2 + \frac{1}{n}\right) \left(1 + \frac{1}{n}\right) \max_{s \in [a,b]} \int_a^b \frac{\mathcal{K}(s, t) \|x(t) - u(t)\|}{x(t)^{\frac{n-1}{n}} + x(t)^{\frac{n-2}{n}} u(t)^{\frac{1}{n}} + \dots + u(t)^{\frac{n-1}{n}}} dt \|z\| \|y\|, \end{aligned}$$

so that

$$\|\mathcal{F}''(x) - \mathcal{F}''(x_0)\| \leq \frac{|\lambda| \left(2 + \frac{1}{n}\right) \left(1 + \frac{1}{n}\right)}{d^{\frac{n-1}{n}}} \left(\max_{s \in [a,b]} \int_a^b \mathcal{K}(s, t) dt \right) \|x - x_0\| \leq L \|x - x_0\|,$$

where $L = \frac{|\lambda| \left(2 + \frac{1}{n}\right) \left(1 + \frac{1}{n}\right)}{d^{\frac{n-1}{n}}} \left(\max_{s \in [a,b]} \int_a^b \mathcal{K}(s, t) dt \right)$. As a consequence, \mathcal{F}'' is Lipschitz continuous at $x_0 \in \Omega$.

Finally, we note that if n is even, then we choose $\Omega = \{x(s) \in \mathcal{C}([a, b]) : x(s) \geq 0, s \in [a, b]\}$.

An important consequence of the above-mentioned is that the semilocal convergence results for Newton's method under center conditions for F'' are interesting regardless of the general situation is not centered.

3 Main results

The Newton-Kantorovich theorem, Theorem 1, guarantees the semilocal convergence of Newton's method in Banach spaces and gives a priori error estimates and information about the existence and uniqueness of solution. Kantorovich proves the theorem by using two different techniques [9, 10], although the most prominent one is the majorant principle [10], which is based on the concept of majorizing sequence. This technique has been usually used later by other authors to analyse the semilocal convergence of several iterative methods [1, 2, 3, 15]. We begin by introducing the concept of majorizing sequence and remembering how it is used to prove the convergence of sequences in Banach spaces.

Definition 3. If $\{x_n\}$ is a sequence in a Banach space X and $\{t_n\}$ is a scalar sequence, then $\{t_n\}$ is a *majorizing sequence* of $\{x_n\}$ if $\|x_n - x_{n-1}\| \leq t_n - t_{n-1}$, for all $n \in \mathbb{N}$.

From the last inequality, it follows the sequence $\{t_n\}$ is nondecreasing. The interest of the majorizing sequence is that the convergence of the sequence $\{x_n\}$ in the Banach space X is deduced from the convergence of the scalar sequence $\{t_n\}$, as we can see in the following result [11]:

Lemma 4. *Let $\{x_n\}$ be a sequence in a Banach space X and $\{t_n\}$ a majorizing sequence of $\{x_n\}$. Then, if $\{t_n\}$ converges to $t^* < +\infty$, there exists $x^* \in X$ such that $x^* = \lim_n x_n$ and $\|x^* - x_n\| \leq t^* - t_n$, for $n = 0, 1, 2, \dots$*

From the definition of majorizing sequence and Lemma 4, Kantorovich proves the Newton-Kantorovich theorem. For this, a majorizing sequence is constructed from conditions (A1) and (A2) of the Newton-Kantorovich theorem, by applying Newton's method,

$$s_0 = 0, \quad s_{n+1} = N_p(s_n) = s_n - \frac{p(s_n)}{p'(s_n)}, \quad n \geq 0,$$

to Kantorovich's polynomial

$$p(s) = \frac{M}{2}s^2 - \frac{s}{\beta} + \frac{\eta}{\beta}. \quad (6)$$

Note that (6) has two positive solutions $s^* = \frac{1-\sqrt{1-2h}}{h}\eta$ and $s^{**} = \frac{1+\sqrt{1-2h}}{h}\eta$ such that $s^* \leq s^{**}$ if $h = M\beta\eta \leq \frac{1}{2}$. Moreover, we consider $p(s)$ in some interval $[0, s']$ taking into account that $s^* \leq s^{**} < s'$.

3.1 Semilocal convergence

In this section, we establish a semilocal convergence result for Newton's method under a center condition of type (3) by using the technique of majorizing sequences. For this, we first suppose the following conditions:

- (D1) There exists $\tilde{x} \in \Omega$ such that $\|x_0 - \tilde{x}\| = \gamma$, where $x_0 \in \Omega$, and $\|F''(\tilde{x})\| \leq \delta$.
- (D2) There exists the operator $\Gamma_0 = [F'(x_0)]^{-1} \in \mathcal{L}(Y, X)$, with $\|\Gamma_0\| \leq \beta$ and $\|\Gamma_0 F(x_0)\| \leq \eta$.
- (D3) $\|F''(x) - F''(\tilde{x})\| \leq \tilde{L}\|x - \tilde{x}\|$ for $x \in \Omega$.

And second, we construct a majorizing sequence from conditions (D1), (D2) and (D3) by applying Newton's method

$$t_0 = 0, \quad t_{n+1} = N_\psi(t_n) = t_n - \frac{\psi(t_n)}{\psi'(t_n)}, \quad n \geq 0, \quad (7)$$

to polynomial

$$\psi(t) = \frac{\tilde{L}}{6}t^3 + \frac{1}{2}(\delta + \gamma\tilde{L})t^2 - \frac{t}{\beta} + \frac{\eta}{\beta}. \quad (8)$$

Note that (8) has two positive zeros t^* and t^{**} such that $t^* \leq t^{**}$ if $\psi(\alpha) \leq 0$, where α is a positive real root of $\psi'(t) = 0$. Moreover, we consider $\psi(t)$ in some interval $[0, t']$ taking into account that $t^* \leq t^{**} < t'$.

Theorem 5. Let $F : \Omega \subseteq X \longrightarrow Y$ be a twice continuously differentiable Fréchet operator defined on a nonempty open convex domain Ω of a Banach space X with values in a Banach space Y . Suppose that conditions (D1), (D2) and (D3) are satisfied. If $\psi(\alpha) \leq 0$, where α is a positive real root of $\psi'(t) = 0$ and ψ is defined in (8), and $B(x_0, t^*) \subset \Omega$, where t^* is the smallest positive zero of polynomial (8), then Newton's sequence defined in (1) and starting at x_0 converges to a solution x^* of the equation $F(x) = 0$ and $x_n, x^* \in \overline{B(x_0, t^*)}$, for all $n \in \mathbb{N}$. In addition, $\|x^* - x_n\| \leq t^* - t_n$ for $n \geq 0$, where $\{t_n\}$ is defined in (7).

Proof. First, x_1 is well-defined, since the operator $\Gamma_0 = [F'(x_0)]^{-1}$ exists by condition (D2). Moreover, taking into account that $t_0 = 0$ and (D2), we have

$$\|x_1 - x_0\| = \|\Gamma_0 F(x_0)\| \leq \eta = -\frac{\psi(0)}{\psi'(0)} = -\frac{\psi(t_0)}{\psi'(t_0)} = t_1 - t_0 < t^*$$

and $x_1 \in B(x_0, t^*)$.

Next, taking into account that

$$\begin{aligned} I - \Gamma_0 F'(x_1) &= \Gamma_0 (F'(x_0) - F'(x_1)) \\ &= -\Gamma_0 \int_{x_0}^{x_1} F''(z) dz \\ &= -\Gamma_0 \int_{x_0}^{x_1} (F''(z) - F''(\tilde{x})) dz - \Gamma_0 F''(\tilde{x})(x_1 - x_0) \\ &= -\Gamma_0 \int_0^1 (F''(x_0 + \tau(x_1 - x_0)) - F''(\tilde{x}))(x_1 - x_0) d\tau - \Gamma_0 F''(\tilde{x})(x_1 - x_0), \end{aligned}$$

it follows

$$\begin{aligned} \|I - \Gamma_0 F'(x_1)\| &\leq \|\Gamma_0\| \int_0^1 \|F''(x_0 + \tau(x_1 - x_0)) - F''(\tilde{x})\| d\tau \|x_1 - x_0\| \\ &\quad + \|\Gamma_0\| \|F''(\tilde{x})\| \|x_1 - x_0\| \\ &\leq \beta \eta \tilde{L} \int_0^1 \|x_0 + \tau(x_1 - x_0) - \tilde{x}\| d\tau + \beta \eta \delta \\ &\leq \beta \left(\tilde{L} \left(\gamma + \frac{1}{2} \eta \right) + \delta \right) \eta \\ &= -\frac{1}{\psi'(t_0)} (\psi'(t_1) - \psi'(t_0)) \\ &= 1 - \frac{\psi'(t_1)}{\psi'(t_0)} \\ &< 1, \end{aligned}$$

since $\psi'(t_1) > 0$ and $\psi'(t_1) > \psi'(t_0)$, and, by the Banach lemma on invertible operators, the operator $\Gamma_1 = [F'(x_1)]^{-1}$ exists and

$$\|\Gamma_1\| \leq -\frac{1}{\psi'(t_1)}.$$

Hence, x_2 is well-defined.

Besides, since

$$\begin{aligned}
 F(x_1) &= F(x_0) + F'(x_0)(x_1 - x_0) + \int_{x_0}^{x_1} F''(z)(x_1 - z) dz \\
 &= \int_{x_0}^{x_1} F''(z)(x_1 - z) dz \\
 &= \int_0^1 F''(x_0 + \tau(x_1 - x_0))(x_1 - x_0)^2(1 - \tau) d\tau \\
 &= \int_0^1 (F''(x_0 + \tau(x_1 - x_0)) - F''(\tilde{x}))(x_1 - x_0)^2(1 - \tau) d\tau \\
 &\quad + \frac{1}{2}F''(\tilde{x})(x_1 - x_0)^2,
 \end{aligned}$$

we have

$$\begin{aligned}
 \|F(x_1)\| &\leq \int_0^1 \|F''(x_0 + \tau(x_1 - x_0)) - F''(\tilde{x})\| (1 - \tau) d\tau \|x_1 - x_0\|^2 \\
 &\quad + \frac{1}{2} \|F''(\tilde{x})\| \|x_1 - x_0\|^2 \\
 &\leq \int_0^1 \psi''(t_0 + \tau(t_1 - t_0))(t_1 - t_0)^2(1 - \tau) d\tau \\
 &= \int_{t_0}^{t_1} \psi''(\xi)(t_1 - \xi) d\xi \\
 &= \psi(t_1).
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 \|x_2 - x_1\| &= \|\Gamma_1 F(x_1)\| \leq \|\Gamma_1\| \|F(x_1)\| \leq -\frac{\psi(t_1)}{\psi'(t_1)} = t_2 - t_1, \\
 \|x_2 - x_0\| &\leq \|x_2 - x_1\| + \|x_1 - x_0\| \leq (t_2 - t_1) + (t_1 - t_0) = t_2 - t_0 < t^*
 \end{aligned}$$

and $x_2 \in B(x_0, t^*)$.

If we assume

$$\|\Gamma_n\| \leq -\frac{1}{\psi'(t_n)}, \quad (9)$$

$$\|F(x_n)\| \leq \psi(t_n), \quad (10)$$

$$\|x_{n+1} - x_n\| \leq -\frac{\psi(t_n)}{\psi'(t_n)} = t_{n+1} - t_n, \quad (11)$$

$$\|x_{n+1} - x_0\| \leq t_{n+1} - t_0 < t^*, \quad (12)$$

where the operator $\Gamma_n = [F'(x_n)]^{-1}$ exists, it follows in the same way that the operator $\Gamma_{n+1} = [F'(x_{n+1})]^{-1}$ exists and

$$\begin{aligned}\|\Gamma_{n+1}\| &\leq -\frac{1}{\psi'(t_{n+1})}, \\ \|F(x_{n+1})\| &\leq \psi(t_{n+1}), \\ \|x_{n+2} - x_{n+1}\| &\leq -\frac{\psi(t_{n+1})}{\psi'(t_{n+1})} = t_{n+2} - t_{n+1}, \\ \|x_{n+2} - x_0\| &\leq \|x_{n+2} - x_{n+1}\| + \|x_{n+1} - x_0\| \leq t_{n+2} - t_0 < t^*,\end{aligned}$$

so that (9), (10), (11) and (12) are true for all positive integers n by mathematical induction. As a consequence, the sequence $\{x_n\}$ is well-defined and $x_n \in B(x_0, t^*)$, for $n \geq 0$.

Since $\lim_n t_n = t^*$, $\{t_n\}$ is a Cauchy sequence, so that $\{x_n\}$ is also a Cauchy sequence and convergent. So, $\lim_n x_n = x^*$ and $\|x^* - x_n\| \leq t^* - t_n$, for $n \geq 0$. Moreover, as $\|F(x_n)\| \leq \psi(t_n)$, for $n \geq 0$, then, by letting $n \rightarrow +\infty$, it follows $F(x^*) = 0$ by the continuities of F and ψ . ■

Note that Theorem 5 is reduced to Theorem 2 if $\tilde{x} = x_0$.

3.2 Uniqueness of the solution

After proving the semilocal convergence of Newton's method and locating the solution x^* , we prove the uniqueness of x^* .

Theorem 6. *Under the conditions of the last theorem, the solution x^* is unique in $B(x_0, t^{**}) \cap \Omega$ if $t^* < t^{**}$ or in $\overline{B(x_0, t^*)}$ if $t^{**} = t^*$.*

Proof. Suppose that $t^* < t^{**}$ and y^* is a solution of $F(x) = 0$ in $B(x_0, t^{**}) \cap \Omega$ different from x^* . Then, taking into account that $t_0 = 0$, we have

$$\|y^* - x_0\| \leq \rho(t^{**} - t_0) \quad \text{with} \quad \rho \in (0, 1).$$

We now suppose $\|y^* - x_i\| \leq \rho^{2i}(t^{**} - t_i)$ for $i = 0, 1, \dots, n$. In addition,

$$\begin{aligned} \|y^* - x_{n+1}\| &= \|\Gamma_n(F(y^*) - F(x_n) - F'(x_n)(y^* - x_n))\| \\ &= \left\| -\Gamma_n \int_0^1 F''(x_n + \tau(y^* - x_n))(y^* - x_n)^2(1 - \tau) d\tau \right\| \\ &= \left\| -\Gamma_n \int_0^1 (F''(x_n + \tau(y^* - x_n)) - F''(\tilde{x}))(y^* - x_n)^2(1 - \tau) d\tau \right. \\ &\quad \left. + \frac{1}{2}F''(\tilde{x})(y^* - x_n)^2 \right\| \\ &\leq \|\Gamma_n\| \int_0^1 \|F''(x_n + \tau(y^* - x_n)) - F''(\tilde{x})\|(1 - \tau) d\tau \|y^* - x_n\|^2 \\ &\quad + \frac{1}{2}\|F''(\tilde{x})\| \|y^* - x_n\|^2 \\ &\leq -\frac{\mu}{\psi'(t_n)} \|y^* - x_n\|^2, \end{aligned}$$

where $\mu = \int_0^1 \psi''(t_n + \tau(t^{**} - t_n))(1 - \tau) d\tau$.

On the other hand, we also have

$$t^{**} - t_{n+1} = -\frac{1}{\psi'(t_n)} \int_{t_n}^{t^{**}} \psi''(\xi)(t^{**} - \xi) d\xi = -\frac{\mu}{\psi'(t_n)} (t^{**} - t_n)^2.$$

Therefore,

$$\|y^* - x_{n+1}\| \leq \frac{t^{**} - t_{n+1}}{(t^{**} - t_n)^2} \|y^* - x_n\|^2 \leq \rho^{2^{n+1}}(t^{**} - t_{n+1}),$$

so that $y^* = x^*$.

If $t^{**} = t^*$ and y^* is another solution of $F(x) = 0$, different from x^* , in $\overline{B(x_0, t^{**})}$, then $\|y^* - x_0\| \leq t^* - t_0 = t^*$. Proceeding similarly to the previous case, we can prove by mathematical induction on n that $\|y^* - x_n\| \leq t^{**} - t_n$. Since $t^{**} = t^*$ and $\lim_n t_n = t^*$, the uniqueness of the solution is now easy to follow. ■

3.3 Quadratic convergence of Newton's method

Now, from the following theorem which provides some a priori error estimates for Newton's method, we deduce the quadratic convergence of the method under conditions (D1), (D2) and (D3). The proof of the theorem follows from Ostrowski's technique [13] and is analogous to that given in [4].

Notice first that if $\psi(t)$ has two real zeros t^* and t^{**} such that $0 < t^* \leq t^{**}$, we can then write

$$\psi(t) = \left(\frac{\tilde{L}}{6}t + \varepsilon \right) (t^* - t)(t^{**} - t)$$

with $\varepsilon > 0$.

Theorem 7. Suppose that conditions (D1), (D2) and (D3) are satisfied and $\psi(\alpha) \leq 0$, where α is a positive root of $\psi'(t) = 0$ and ψ is given in (8).

(a) If $t^* < t^{**}$ and $t^* > \frac{6\varepsilon}{L}$, then

$$\frac{(t^{**} - t^*)\theta^{2^n}}{P - \theta^{2^n}} \leq t^* - t_n \leq \frac{(t^{**} - t^*)\Delta^{2^n}}{Q - \Delta^{2^n}}, \quad n \geq 0,$$

where $\theta = \frac{t^*}{t^{**}}P$, $\Delta = \frac{t^*}{t^{**}}Q$, $P = \frac{\tilde{L}t^{**} - 6\varepsilon}{\tilde{L}t^* + 6\varepsilon}$, $Q = \frac{\tilde{L}(2t^* - t^{**}) + 6\varepsilon}{\tilde{L}t^* + 6\varepsilon}$ and provided that $\theta < 1$ and $\Delta < 1$.

(b) If $t^* = t^{**}$ and $t^* > \frac{12\varepsilon}{L}$, then

$$\left(\frac{\tilde{L}t^* - 6\varepsilon}{\tilde{L}t^* - 12\varepsilon} \right)^n t^* \leq t^* - t_n \leq \frac{t^*}{2^n}, \quad n \geq 0.$$

Proof. Let $t^* < t^{**}$ and denote $a_n = t^* - t_n$ and $b_n = t^{**} - t_n$ for all $n \geq 0$. Then

$$\psi(t_n) = \frac{1}{6} (\tilde{L}t_n + 6\varepsilon) a_n b_n, \quad \psi'(t_n) = \frac{\tilde{L}}{6} a_n b_n - \frac{1}{6} (\tilde{L}t_n + 6\varepsilon) (a_n + b_n)$$

and

$$a_{n+1} = t^* - t_{n+1} = t^* - t_n + \frac{\psi(t_n)}{\psi'(t_n)} = \frac{a_n^2 (\tilde{L}b_n - 6\varepsilon - \tilde{L}t_n)}{\tilde{L}a_nb_n - (\tilde{L}t_n + 6\varepsilon)(a_n + b_n)}.$$

From $\frac{a_{n+1}}{b_{n+1}} = \frac{a_n^2 (\tilde{L}b_n - (\tilde{L}t_n + 6\varepsilon))}{b_n^2 (\tilde{L}a_n - (\tilde{L}t_n + 6\varepsilon))}$ and taking into account function $d(t) = \frac{\tilde{L}t^{**} - 6\varepsilon - 2\tilde{L}t}{\tilde{L}t^* - 6\varepsilon - 2\tilde{L}t}$,

$P \leq \min\{d(t); t \in [0, t^*]\} = d(0)$ and $Q = \max\{d(t); t \in [0, t^*]\} = d(t^*)$ it follows

$$P \left(\frac{a_n}{b_n} \right)^2 \leq \frac{a_{n+1}}{b_{n+1}} \leq Q \left(\frac{a_n}{b_n} \right)^2.$$

In addition,

$$\frac{a_{n+1}}{b_{n+1}} \leq Q^{2^{n+1}-1} \left(\frac{a_0}{b_0} \right)^{2^{n+1}} = \frac{\Delta^{2^{n+1}}}{Q} \quad \text{and} \quad \frac{a_{n+1}}{b_{n+1}} \geq P^{2^{n+1}-1} \left(\frac{a_0}{b_0} \right)^{2^{n+1}} = \frac{\theta^{2^{n+1}}}{P}.$$

Taking then into account that $b_{n+1} = (t^{**} - t^*) + a_{n+1}$, it follows:

$$\frac{(t^{**} - t^*)\theta^{2^{n+1}}}{P - \theta^{2^{n+1}}} \leq t^* - t_{n+1} \leq \frac{(t^{**} - t^*)\Delta^{2^{n+1}}}{Q - \Delta^{2^{n+1}}}.$$

If $t^* = t^{**}$, then $a_n = b_n$ and

$$a_{n+1} = \frac{a_n \left(\tilde{L} a_n - (\tilde{L} t_n + 6\varepsilon) \right)}{\tilde{L} a_n - 2(\tilde{L} t_n + 6\varepsilon)}.$$

As a consequence, $\left(\frac{\tilde{L} t^* - 6\varepsilon}{\tilde{L} t^* - 12\varepsilon} \right) a_n \leq a_{n+1} \leq \frac{a_n}{2}$ and

$$\left(\frac{\tilde{L} t^* - 6\varepsilon}{\tilde{L} t^* - 12\varepsilon} \right)^{n+1} t^* \leq t^* - t_{n+1} \leq \frac{t^*}{2^{n+1}}.$$

The proof is complete. ■

From the last theorem, it follows that the convergence of Newton's method is quadratic if $t^* < t^{**}$ and linear if $t^* = t^{**}$.

4 Region of accessibility

An important aspect to consider when studying the applicability of an iterative method is the set of starting points that we can take into account, so that the iterative method converges to a solution of an equation from any point of the set, what we call region of accessibility of the iterative method.

We consider the simple academic example given by the complex equation $\omega(x) = x^{5/2} - 3/2 = 0$ and analyse the region of accessibility of the solution $x^* = 1.1760 \dots$ when it is approximated by Newton's method. For this, we consider the ball $B(0, 2)$ as the domain of the function and study two situations.

In the first one, we consider $\tilde{x} = 1$, $x_0 \in B(0, 2)$ and, to paint the region of accessibility, colour of red all the points x_0 that satisfy the main condition $(\psi(\alpha) \leq 0)$ of Theorem 5. So, once $\gamma = |x_0 - \tilde{x}| = |x_0 - 1|$ is fixed, we obtain the region of accessibility of x^* shown in Figure 1, where the white point is x^* and the black point is \tilde{x} . Observe that we can choose $\tilde{x} = 1$, the point where ω'' is center Lipschitz, as starting point for Newton's method.

In the second situation, we consider $\tilde{x} = 1/2$ and $x_0 \in B(0, 2)$, do the same as before to paint the region of accessibility and obtain the region shown in Figure 2, where the black point is now $\tilde{x} = 1/2$. Observe that, in this case, we cannot choose \tilde{x} as starting point for Newton's method.

Note that, in both situations, we can find starting points (points of the red regions) for Newton's method different from those where ω'' is center Lipschitz.

Finally, we can conjecture that as the point \tilde{x} , the point where ω'' is center Lipschitz, is further away from the solution x^* of the equation, the region of accessibility is smaller.

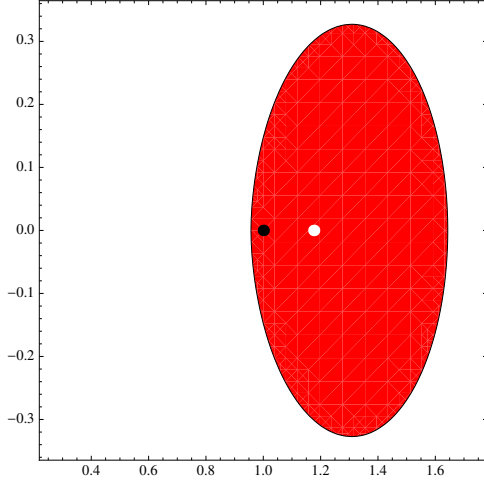


Figure 1: Region of accessibility of $x^* = 1.1760\dots$ when $\tilde{x} = 1$

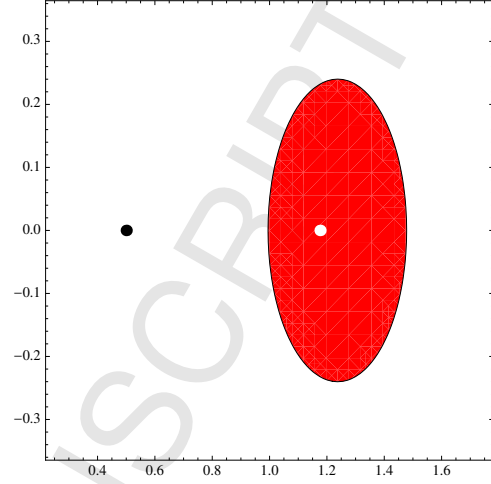


Figure 2: Region of accessibility of $x^* = 1.1760\dots$ when $\tilde{x} = 1/2$

5 Application

Now, we illustrate the above-mentioned with a system of nonlinear equations that arises from a process of discretization that transform a nonlinear integral equation of type (4) into a finite dimensional problem.

We consider the following nonlinear integral equation of mixed Hammerstein-type

$$x(s) = 1 + \frac{8}{5} \int_0^1 G(s, t) x(t)^{\frac{5}{2}} dt, \quad (13)$$

where the kernel G is the Green function in $[0, 1] \times [0, 1]$, and approximate the integral of (13) by a Gauss-Legendre quadrature formula with 8 nodes:

$$\int_a^b \varphi(t) dt \simeq \sum_{i=1}^8 w_i \varphi(t_i),$$

where the nodes t_i and the weights w_i are determined; in particular, see Table 1 for $m = 8$.

i	t_i	w_i	i	t_i	w_i
1	0.019855...	0.050614...	5	0.591717...	0.181342...
2	0.101667...	0.111191...	6	0.762766...	0.156853...
3	0.237234...	0.156853...	7	0.898333...	0.111191...
4	0.408283...	0.181342...	8	0.980145...	0.050614...

Table 1: Nodes and weights for the Gauss-Legendre formula

If we denote the approximations of $x(t_i)$ by x_i , with $i = 1, 2, \dots, m$, then equation (13) is transformed into the following nonlinear system:

$$x_i = 1 + \frac{8}{5} \sum_{j=1}^8 a_{ij} x_j^{\frac{5}{2}}, \quad j = 1, 2, \dots, 8, \quad (14)$$

where

$$a_{ij} = w_j G(t_i, t_j) = \begin{cases} w_j (1 - t_i) t_j, & j \leq i, \\ w_j (1 - t_j) t_i, & j > i. \end{cases}$$

After that, system (14) can be written as

$$\mathbb{F}(\mathbf{x}) \equiv \mathbf{x} - \mathbf{1} - \frac{8}{5} A \hat{\mathbf{x}} = 0, \quad \mathbb{F} : \Lambda \subset \mathbb{R}^8 \longrightarrow \mathbb{R}^8, \quad (15)$$

where

$$\mathbf{x} = (x_1, x_2, \dots, x_8)^T, \quad \mathbf{1} = (1, 1, \dots, 1)^T, \quad A = (a_{ij})_{i,j=1}^8, \quad \hat{\mathbf{x}} = \left(x_1^{\frac{5}{2}}, x_2^{\frac{5}{2}}, \dots, x_8^{\frac{5}{2}} \right)^T.$$

In view of what the domain Ω is for equation (13), see the introduction, we then consider $\Lambda = \{\mathbf{x} \in \mathbb{R}^8 : x_i \geq 0, \forall i = 1, 2, \dots, 8\}$. Besides, $\mathbb{F}'(\mathbf{x}) = I - 4A \operatorname{diag} \left\{ x_1^{\frac{3}{2}}, x_2^{\frac{3}{2}}, \dots, x_8^{\frac{3}{2}} \right\}$ and

$$\mathbb{F}''(\mathbf{x}) \mathbf{y} \mathbf{z} = (y_1, y_2, \dots, y_8) \mathbb{F}''(\mathbf{x}) (z_1, z_2, \dots, z_8),$$

where $\mathbf{y} = (y_1, y_2, \dots, y_8)^T$ and $\mathbf{z} = (z_1, z_2, \dots, z_8)^T$, so that

$$\mathbb{F}''(\mathbf{x}) \mathbf{y} \mathbf{z} = 6A \left(x_1^{\frac{1}{2}} y_1 z_1, x_2^{\frac{1}{2}} y_2 z_2, \dots, x_8^{\frac{1}{2}} y_8 z_8 \right)^T.$$

As $\|\mathbb{F}''(\mathbf{x})\|$ is not bounded, we cannot guarantee the convergence of Newton's method by the Newton-Kantorovich theorem.

Moreover, if we choose $\tilde{\mathbf{x}} = \mathbf{1}$ and the max-norm, we observe that $\mathbb{F}''(\mathbf{x})$ is center Lipschitz at $\tilde{\mathbf{x}} = \mathbf{1}$ with $\tilde{L} = 0.7413 \dots$, and $\delta = \tilde{L}$. If we now choose the starting point $\mathbf{x}_0 = \tilde{\mathbf{x}}$ and consider Theorem 5, which is reduced to Theorem 2 in this case, then $\gamma = 0$, $\beta = 1.8540 \dots$ and $\eta = 0.3416 \dots$,

$$\psi(t) = (0.1842 \dots) - (0.5393 \dots)t + (0.3706 \dots)t^2 + (0.1235 \dots)t^3$$

and $\psi(\alpha) = \psi(0.5668 \dots) = 0.0201 \dots > 0$. As a consequence, we cannot guarantee the convergence of Newton's method by Theorem 5. In addition, we cannot choose the point $\tilde{\mathbf{x}} = \mathbf{1}$, where $\mathbb{F}''(\mathbf{x})$ is center Lipschitz continuous, as starting point for Newton's method.

However, we can guarantee the convergence of Newton's method from starting at other points different from that, $\tilde{\mathbf{x}}$, where $\mathbb{F}''(\mathbf{x})$ is center Lipschitz continuous. For example, if we choose the starting point $\mathbf{x}_0 = (1, 1.1, 1.2, 1.3, 1.3, 1.2, 1.1, 1)^T$, different from $\tilde{\mathbf{x}} = \mathbf{1}$, then $\gamma = \|\mathbf{x}_0 - \tilde{\mathbf{x}}\| = 0.3$, $\beta = 2.7064 \dots$ and $\eta = 0.1249 \dots$,

$$\psi(t) = (0.0461 \dots) - (0.3694 \dots)t + (0.4818 \dots)t^2 + (0.1235 \dots)t^3$$

and $\psi(\alpha) = \psi(0.3391\dots) = -0.0188\dots \leq 0$. Therefore, we can now guarantee the convergence of Newton's method by Theorem 5 when Newton's method starts at $\mathbf{x}_0 \neq \tilde{\mathbf{x}}$.

Next, we use Newton's method to approximate a solution of (15). In Table 2 we can see the numerical approximation $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_8^*)^T$ of the solution after five iterations of Newton's method when the stopping criterion $\|\mathbf{x}_n - \mathbf{x}_{n-1}\| < 10^{-16}$ is used. In Table 3, we show the errors $\|\mathbf{x}_n - \mathbf{x}^*\|$ obtained with the same stopping criterion. Notice that the vector shown in Table 2 is a good approximation of a solution of system (15), since $\|F(\mathbf{x}^*)\| \leq \text{constant} \times 10^{-16}$. See the sequence $\{\|F(\mathbf{x}_n)\|\}$ in Table 3.

i	x_i^*	i	x_i^*	i	x_i^*	i	x_i^*
1	1.031098...	3	1.318208...	5	1.442167...	7	1.152085...
2	1.152085...	4	1.442167...	6	1.318208...	8	1.031098...

Table 2: Numerical solution \mathbf{x}^* of system (15)

n	$\ \mathbf{x}_n - \mathbf{x}^*\ $	$\ F(\mathbf{x}_n)\ $
0	$1.4216\dots \times 10^{-1}$	$7.3339\dots \times 10^{-2}$
1	$1.7193\dots \times 10^{-2}$	$5.5028\dots \times 10^{-3}$
2	$2.9680\dots \times 10^{-4}$	$9.4169\dots \times 10^{-5}$
3	$9.1260\dots \times 10^{-8}$	$2.8963\dots \times 10^{-8}$
4	$8.6295\dots \times 10^{-15}$	$2.7389\dots \times 10^{-15}$

Table 3: Absolute errors obtained by Newton's method and $\|F(\mathbf{x}_n)\|$ for system (15)

References

- [1] S. Amat, S. Busquier, Third-order iterative methods under Kantorovich conditions, J. Math. Anal. Appl., 336, 1 (2007) 243–261.
- [2] S. Amat, C. Bermúdez, S. Busquier and D. Mestiri, A family of Halley-Chebyshev iterative schemes for non-Fréchet differentiable operators, J. Comput. Appl. Math., 228, 1 (2009) 486–493.
- [3] I. K. Argyros, On the Newton-Kantorovich hypothesis for solving equations, J. Comput. Appl. Math., 169, 2 (2004) 315–332.
- [4] J. A. Ezquerro, D. González and M. A. Hernández, Majorizing sequences for Newton's method from initial value problems, J. Comput. Appl. Math., 236 (2012) 2246–2258.

- [5] J. A. Ezquerro and M. A. Hernández-Verón, On the domain of starting points of Newton's method under center Lipschitz conditions, *Mediterr. J. Math.*, 13 (2016) 2287–2300.
- [6] J. M. Gutiérrez, A new semilocal convergence theorem for Newton's method, *J. Comput. Appl. Math.*, 79 (1997) 131–145.
- [7] J. M. Gutiérrez and M. A. Hernández, Newton's method under weak Kantorovich conditions, *IMA J. Numer. Anal.*, 20 (2000) 521–532.
- [8] Huang Zhengda, A note on the Kantorovich theorem for Newton method, *J. Comput. Appl. Math.*, 47 (1993) 211–217.
- [9] L. V. Kantorovich, On Newton's method for functional equations, *Dokl. Akad. Nauk. SSSR*, 59 (1948) 1237–1240 (in Russian).
- [10] L. V. Kantorovich, The majorant principle and Newton's method, *Dokl. Akad. Nauk. SSSR*, 76 (1951) 17–20 (in Russian).
- [11] L. V. Kantorovich and G. P. Akilov, *Functional analysis*, Pergamon Press, Oxford, 1982.
- [12] J. M. Ortega, The Newton-Kantorovich theorem, *Amer. Math. Monthly*, 75 (1968) 658–660.
- [13] A. M. Ostrowski, *Solution of equations and systems of equations*, Academic Press, New York, 1966.
- [14] J. Rokne, Newton's method under mild differentiability conditions with error analysis, *Numer. Math.* 18 (1972) 401–412.
- [15] T. Yamamoto, Convergence theorem for Newton-like methods in Banach spaces, *Numer. Math.*, 51 (1987) 545–557.