Third-order iterative methods for operators with bounded second derivative

José M. Gutiérrez*, Miguel A. Hernández

Department of Mathematics and Computation, University of La Rioja C/Luis de Ulloa s/n, 26004, Logroño, Spain

Received 31 July 1996; revised 28 March 1997

Abstract

We analyse the classical third-order methods (Chebyshev, Halley, super-Halley) to solve a nonlinear equation \( F(x) = 0 \), where \( F \) is an operator defined between two Banach spaces. Until now the convergence of these methods is established assuming that the second derivative \( F'' \) satisfies a Lipschitz condition. In this paper we prove, by using recurrence relations, the convergence of these and other third-order methods just assuming \( F'' \) is bounded. We show examples where our conditions are fulfilled and the classical ones fail.

Keywords: Nonlinear equations in Banach spaces; Third-order method; Recurrence relations; Convergence theorem

AMS classification: 47H17; 65J15

1. Introduction

One of the most important techniques to study a nonlinear equation \( F(x) = 0 \) defined between two Banach spaces \( X \) and \( Y \), is the use of iterative processes. Certainly Newton’s method is the most famous and useful iteration for this purpose. There is a wide bibliography concerning Newton’s method and its applications (see [18, 20] for instance). But there are other methods. Third-order methods have been left aside for a long time because of their high computational cost, mainly for the evaluation of the second-order Fréchet derivative. However, in some case the rise in the velocity of convergence can justify their use. For instance, these methods have been successfully used in the solution of nonlinear integral equations [5, 6, 10]. They can also be used in problems where a quick convergence is required, such as stiff systems [19]. Besides, these methods are also interesting from the theoretical standpoint because they provide results on existence and uniqueness of solution that

* Corresponding author. E-mail: jmguti@siur.unirioja.es.

† Supported in part by a grant of the University of La Rioja (ULR-95PYB16MHV).
improve the results given by using Newton’s method. In that way, results specially important have been given for the case of quadratic equations [3, 14].

In that way, many papers on third-order methods have been published. So for Chebyshev method we have [4, 6, 17], for Halley method [1, 2, 5, 7, 8, 11, 15, 17, 23] and for super-Halley method [9, 12, 16]. In [13] a unified theory is developed for the above third-order methods. To sum up, in [13] it is studied the one-parametered family of third-order methods

$$x_{n+1} = x_n - \left( I + \frac{1}{2} L_F(x_n) [I - \alpha L_F(x_n)]^{-1} \right) F'(x_n)^{-1} F(x_n),$$  \hspace{1cm} (1)

for \( \alpha \in [0, 1] \). We have denoted \( I \) the identity operator on \( X \) and

$$L_F(x) = F'(x)^{-1} F''(x) F'(x)^{-1} F(x), \quad x \in X,$$

provided that \( F'(x)^{-1} \) exists.

This family extends the family of scalar iterative processes considered by Hernández and Salanova in [17], and includes, as particular cases, Chebyshev method (\( \alpha = 0 \)), Halley method (\( \alpha = 1/2 \)) and super-Halley method (\( \alpha = 1 \)).

Until now we know two different ways for studying third-order iterative processes. In the first one, necessary conditions for the convergence of (1) have been established assuming that the second-order Fréchet derivative of \( F \) satisfies a Lipschitz condition

$$\| F''(x) - F''(y) \| \leq k \| x - y \|$$ \hspace{1cm} (2)

for \( x \) and \( y \) in a suitable region of \( X \). See the above cites for more information. The technique developed by these authors is an extension of the technique followed by Kantorovich and other authors [18, 20] to study Newton’s method.

In the second one, Smale [21] obtained the convergence of Newton’s method for analytic maps from data at one point, instead of the region conditions (2) in the Newton-Kantorovich theorem.

Smale-like theorems for the convergence of iterative processes assume that the following inequalities are satisfied at a point \( x_0 \):

$$\frac{1}{k!} \| I_0 F^{(k)}(x_0) \| \| I_0 F(x_0) \|^{k-1} \leq h^{k-1}, \quad k \geq 2.$$ \hspace{1cm} (3)

The constant \( h \) is different for different processes [22, 24].

Our goal in this paper is to prove the convergence of (1) just assuming \( F'' \) is bounded and a punctual condition. We use recurrence relations, in a similar way that Candela and Marquina in [5, 6]. The use of these relations supposes some advantages, because we can reduce our initial problem in a Banach space to a simpler problem with real sequences and functions. Besides we obtain convergent sequences for a wider interval of values of \( \alpha \) than in [13]. We also show examples where our conditions are fulfilled and the previous ones fail.

2. Recurrence relations

Let \( X, Y \) be Banach spaces and \( F : \Omega \subseteq X \to Y \) be a nonlinear twice Fréchet differentiable operator in an open convex domain \( \Omega_0 \subseteq \Omega \). Let us assume that \( I_0 = F'(x_0)^{-1} \in \mathcal{L}(Y, X) \) exists at some \( x_0 \in \Omega_0 \), where \( \mathcal{L}(Y, X) \) is the set of bounded linear operators from \( Y \) into \( X \).
Throughout this paper we assume that
(i) \( \|F''(x)\| \leq k, x \in \Omega_0 \).
(ii) \( \|F_0\| \leq B \).
(iii) \( \|I_0 F(x_0)\| \leq \eta \).

Let us denote \( a = k \beta \eta \). Then for \( x \in \mathbb{R} \), we define the sequences

\[ a_0 = b_0 = 1; \quad c_0 = a; \quad d_0 = \frac{2 + a(1 - 2a)}{2(1 - a)}; \]
\[ a_{n+1} = \frac{a_n}{1 - aa_n d_n}; \]
\[ b_{n+1} = \frac{aa_{n+1} d_n^2}{2} \left[ 1 + \frac{4(1 - a c_n)}{(2 + (1 - 2a) c_n)^2} \right]; \]
\[ c_{n+1} = aa_{n+1} b_{n+1}; \]
\[ d_{n+1} = \frac{2 + (1 - 2a)c_{n+1}}{2(1 - ac_{n+1})} b_{n+1}. \]

Let \( \{x_n\} \) a sequence of the family (1) (for comodity, we have not written the subscript \( \alpha \), expecting the reader does not get confused). In that situation, we are going to prove that

(I) \( \|I_\alpha\| = \|F'(x_\alpha)^{-1}\| \leq a_\alpha B \).
(II) \( \|I_\alpha F(x_n)\| \leq b_\alpha \eta \).
(III) \( \|F(x_n)\| \leq c_\alpha \).
(IV) \( \|x_{n+1} - x_n\| \leq d_\alpha \eta \).

(I)–(III) follow immediately from the hypothesis. If

\[ \alpha \|F(x_0)\| \leq \alpha c_0 = \alpha a < 1, \]

then \([I - \alpha L_F(x_0)]^{-1}\) exists and

\[ \|x_1 - x_0\| \leq \left[ 1 + \frac{a}{2(1 - a \alpha)} \right] \eta = d_0 \eta, \]

and (IV) also holds.

Following an inductive procedure and assuming

\[ aa_\alpha d_\alpha < 1, \]

we have

\[ \|I - I_\alpha F'(x_{n+1})\| \leq \|I_\alpha\| \|F'(x_n) - F'(x_{n+1})\| \leq aa_\alpha d_\alpha < 1. \]
Then $I_{n+1}$ is defined and
\[ \| I_{n+1} \| \leq \frac{\| I_n \|}{1 - \| I_n \| \| F'(x_n) - F'(x_{n+1}) \|} \leq \frac{a_n B}{1 - a_n d_n} = a_{n+1} B. \]

On the other hand, we deduce from (1) that
\[ F'(x_{n+1})(x_{n+1} - x_n) = -F(x_n) - \frac{1}{2}F''(x_n)I_n F(x_n)[I - \alpha L F(x_n)]^{-1} I_n F(x_n), \]
and then
\[ F(x_{n+1}) = \int_{x_n}^{x_{n+1}} [F'(x) - F'(x_n)] \, dx - \frac{1}{2}F''(x_n)I_n F(x_n)[I - \alpha L F(x_n)]^{-1} I_n F(x_n). \]

Consequently,
\begin{align*}
\| I_{n+1} F(x_{n+1}) \| &\leq \| I_{n+1} \| \| F(x_{n+1}) \| \leq \frac{a_{n+1} B k \eta^2}{2} \left[ d_n^2 + \frac{b_n^2}{1 - \alpha c_n} \right] \\
&= \frac{a a_{n+1} d_n^2}{2} \left[ 1 + \frac{4(1 - \alpha c_n)}{(2 + (1 - 2 \alpha) c_n)^2} \right] \eta = b_{n+1} \eta. \tag{4}
\end{align*}

Finally, it can be easily deduced that
\[ \| L_F(x_{n+1}) \| \leq \| I_{n+1} \| \| F''(x_{n+1}) \| \| I_{n+1} F(x_{n+1}) \| \leq a a_{n+1} b_{n+1} = c_{n+1} \]
and, as in the case $n = 0$, if we assume $\alpha c_{n+1} < 1$,
\[ \| x_{n+2} - x_{n+1} \| \leq d_{n+1} \eta. \]

So, to study the sequence $\{x_n\}$ defined in a Banach space we must analyse the real sequences $\{a_n\}$, $\{b_n\}$, $\{c_n\}$ and $\{d_n\}$. To establish the convergence of $\{x_n\}$ we only have to prove that $\{d_n\}$ is a Cauchy sequence and the above assumptions
\[ \alpha c_n < 1, \quad n \in \mathbb{N}, \]
\[ a a_n d_n < 1, \quad n \in \mathbb{N}. \]

That is the aim of the following section.

3. Convergence study

In this section we are going to study the sequences $\{a_n\}$, $\{b_n\}$, $\{c_n\}$ and $\{d_n\}$ defined in the previous one to prove the convergence of $\{x_n\}$ defined in (1). First at all we give a technical lemma including the results concerning one and two variable functions that we are going to need. We omit the proof expecting the reader could get it patiently but without any difficulty.
Lemma 3.1. Let \( a_0 = 0.32664 \ldots \) the smallest positive root of \( 2x^4 + 7x^3 - 4x^2 - 24x + 8 = 0 \). Let us define the functions
\[
h(x) = \frac{4 - 11x + x^2 + 2x^3 - x\sqrt{1 + 8x - 4x^2}}{2x(2 - 5x + 2x^2)},
\]
\[
H(x, y) = \frac{8 + 4(1 - 3x)y + (1 - 2x)^2 y^2}{[2 - 2(1 + x)y - (1 - 2x)y^2]^2},
\]
\[
g_\alpha(x) = \frac{2 + (1 - 2\alpha)x}{2 - 2x(1 + \alpha) - x^2(1 - 2\alpha)} \quad \text{and} \quad f_\alpha(x) = \frac{2 + (1 - 2\alpha)x}{2(1 - \alpha x)}.
\]

Then \( h(x) \) is decreasing and \( h(x) < 1/x - \frac{1}{2} \) for \( x \in (0, a_0) \). \( H(x, y) \) is increasing in the variable \( y \) for \( y \in (0, a_0) \) and \( 0 \leq x \leq h(y) \). Finally, for all \( \alpha > 0 \), \( g_\alpha(x) \) and \( f_\alpha(x) \) are increasing functions.

Now we start with an easy lemma that gives us a recurrence relation for the sequence \( \{c_n\} \). The proof follows from the definition of the sequences \( \{a_n\}, \{b_n\}, \{c_n\} \) and \( \{d_n\} \).

Lemma 3.2. For the sequences \( \{c_n\} \), the following recurrence is true
\[
c_{n+1} = \frac{c_n^2}{2} \frac{8 + 4(1 - 3\alpha)c_n + (1 - 2\alpha)^2 c_n^2}{[2 - 2(1 + \alpha)c_n - (1 - 2\alpha)c_n^2]^2}.
\]

Lemma 3.3. Let \( 0 < a < a_0 \) and \( 0 \leq \alpha \leq h(a) \). Then the sequence \( \{c_n\} \) is decreasing.

Proof. By Lemma 3.2, we have to prove
\[
\frac{c_n}{2} \frac{8 + 4(1 - 3\alpha)c_n + (1 - 2\alpha)^2 c_n^2}{[2 - 2(1 + \alpha)c_n - (1 - 2\alpha)c_n^2]^2} \leq 1, \quad n \geq 0.
\]

For \( n = 0 \), we have
\[
8a + 4a^2(1 - 3\alpha) + a^3(1 - 2\alpha)^2 \leq 2[2 - 2(1 + \alpha)a - (1 - 2\alpha)a^2]^2,
\]
that is
\[
4a^2(2 - 5a + 2a^2)x^2 - 4a(4 - 11a + a^2 + 2a^3)x + (8 - 24a - 4a^2 + 7a^3 + 2a^4) \geq 0
\]
and this is true for \( 0 \leq \alpha \leq h(a) \). Then \( c_1 \leq c_0 \).

Let us assume \( c_k \leq c_{k-1} \leq \cdots \leq c_1 \leq c_0 \). As in the above situation, \( c_{k+1} \leq c_k \) if \( 0 \leq \alpha \leq h(c_k) \). Taking into account that \( h(x) \) is a decreasing function (Lemma 3.1) and the hypothesis, we have \( \alpha \leq h(a) = h(c_0) \leq h(c_k) \) and the result follows by induction. \( \Box \)

Lemma 3.4. Under the hypothesis of the Lemma 3.3, \( \alpha c_n < 1 \), \( a a_n d_n < 1 \) for \( n \geq 0 \) and \( \{a_n\} \) is an increasing sequence.

Proof. First, notice that
\[
\alpha c_n \leq \alpha c_0 = \alpha a \leq ah(a) < 1.
\]
On the other hand,

\[ a_n d_n = \frac{2 + (1 - 2x)c_n}{2(1 - xc_n)}. \]

Then \( a_n d_n < 1 \) if \( x < q(c_n) \), where \( q(x) = (2 - 2x - x^2)/(2x(1 - x)) \). As \( q(x) \) is a decreasing function and \( c_n \leq c_0, \) \( q(c_n) > q(c_0) \). Besides, \( x \leq h(a) = h(c_0) \), so the result follows if we prove \( h(a) < q(a) \) for \( a \in (0, a_0] \). Indeed,

\[ q(a) - h(a) = \frac{1 + (1 - a)\sqrt{1 + 8a - 4a^2}}{2(1 - a)(2 - 5a + 2a^2)} > 0, \quad a \in (0, a_0]. \]

Consequently, \( a_n d_n < 1, n \geq 0 \).

Finally, \( a_0 = 1, \) \( a_i = a_i/(1 - a_0 d_0) \geq a_0 = 1 \) and inductively, \( a_{n+1} = a_n/(1 - a_n d_n) \geq a_n \geq a_{n-1} \geq \cdots \geq a_1 = a_0 = 1 \). □

**Lemma 3.5.** With the previous notations, let \( 0 < a \leq a_0 \) and \( 0 < a < h(a) \)(that is, \( c_{n+1} < c_n, n \geq 0 \)). Then

\[ c_n \leq \gamma^2 \frac{c_0}{\gamma}, \quad \text{where} \quad \gamma = \frac{c_1}{c_0} < 1. \]

Consequently, \( \lim_{n \to \infty} c_n = 0. \) Furthermore,

\[ \sum_{n=0}^{\infty} c_n < \infty. \]

**Proof.** Let us write \( c_i = \gamma c_{i-1} \), with \( \gamma < 1 \). We prove that \( c_n \leq \gamma c_{n-1} \) implies \( c_{n+1} \leq \gamma^2 c_n \). With the notations of the Lemma 3.1 we have

\[ c_{n+1} = \frac{c_n^2}{\gamma} H(z,c_n) \leq \frac{\gamma^2 c_{n-1}^2}{\gamma} H(z,c_n). \]

As \( H(z,y) \) is increasing in the second variable and \( c_n < c_{n-1} \),

\[ c_{n+1} \leq \frac{\gamma^2 c_{n-1}^2}{\gamma} H(z,c_{n-1}) = \gamma^2 c_n. \]

Then we have \( c_{n+1} \leq \gamma^2 c_n \) and, using this inequality, \( c_n \leq \gamma^2 c_{0}/\gamma \). As \( \gamma < 1 \), the first part follows.

On the other hand, define \( L(x,y) = (y^2/2)H(x,y) \). So, \( c_{n+1} = L(x,c_n) \). Notice that \( (\partial L/\partial y)(x,0) = 0 \). Then, as \( c_n \to 0 \) and \( (\partial L/\partial y)(x,y) \) is a continuous function for \( 0 < y \leq a \) and \( 0 \leq x \leq h(a) \), there exists a constant \( \tau \in (0,1) \) such that \( (\partial L/\partial y)(x,y) \leq \tau \) for \( y \) near to 0. Consequently, for \( n_0 \) large enough,

\[ c_{n_0+k+1} = L(x,c_{n_0+k}) - L(x,0) \leq \tau c_{n_0+k} \]

and recurrently, \( c_{n_0+j} \leq \tau^j c_{n_0} \). Then

\[ \sum_{n=0}^{\infty} c_n = \sum_{n=0}^{n_0-1} c_n + \sum_{n=n_0}^{\infty} c_n \leq \sum_{n=0}^{n_0-1} c_n + c_{n_0} \sum_{n=n_0}^{\infty} \tau^n c_{n_0} < \infty. \]

□
Lemma 3.6. The sequence \( \{a_n\} \) is upper bounded, that is, there exists a constant \( M > 0 \) such that \( a_n \leq M \) for \( n \geq 0 \).

**Proof.** By the definition of the sequences, we get
\[
a_{n+1} = a_n \frac{2(1 - \alpha c_n)}{2 - 2(1 + \alpha)c_n - (1 - 2\alpha)c_n^2} = a_n \left[ 1 + c_n \frac{2 + (1 - 2\alpha)c_n}{2 - 2(1 + \alpha)c_n - (1 - 2\alpha)c_n^2} \right].
\]
Taking into account this equality and with the notations of the Lemma 3.1, we write
\[a_n = \prod_{k=0}^{n} \left[ 1 + c_k g_{a}(c_k) \right].\]
As, for each \( \alpha \geq 0 \), \( g_{a}(x) \) is an increasing function of \( x \) for \( 0 < x \leq a_0 \) and \( \{c_n\} \) is a decreasing sequence,
\[\ln a_n \leq \sum_{k=0}^{n} \ln \left[ 1 + g_{a}(a_0)c_k \right] = g_{a}(a_0) \sum_{k=0}^{n} c_k < \infty. \quad \square\]

Lemma 3.7. We have \( d_n \leq d_0 \gamma^{n-1} \). Consequently, \( \sum_{n=0}^{\infty} d_n < \infty \) and \( \{d_n\} \) is a Cauchy sequence.

**Proof.** Observe that
\[d_n = \frac{c_n}{a_n} \frac{2 + (1 - 2\alpha)c_n}{2(1 - \alpha c_n)} = \frac{c_n}{a_n} f_{a}(c_n) \leq \frac{c_n}{a} f_{a}(c_n) \leq d_0 \gamma^{n-1},\]
by the Lemmas 1, 4 and 6. Then, the proof is completed. \( \square \)

Now, we are ready to state the following result on the convergence of the methods defined in (1).

**Main Theorem 3.8.** Let \( X, Y \) be Banach spaces and \( F : \Omega \subseteq X \rightarrow Y \) be a nonlinear twice Fréchet differentiable operator in an open convex domain \( \Omega_0 \subseteq \Omega \). Let us assume that \( I_0 = F'(x_0)^{-1} \in \mathcal{L}(Y,X) \) exists at some \( x_0 \in \Omega_0 \) and

(i) \( \|F''(x)\| \leq k, x \in \Omega_0. \)
(ii) \( \|I_0\| \leq B. \)
(iii) \( \|F_0F(x_0)\| \leq \eta. \)

Let us denote \( a = kB\eta \). Suppose that \( 0 < a \leq a_0 = 0.32664 \ldots \), where \( a_0 \) is the smallest positive root of \( 2x^4 + 7x^3 - 4x^2 - 24x + 8 = 0 \), and \( 0 \leq \alpha \leq h(a) \) where \( h(x) \) is the function defined in the Lemma 3.1. Then, if \( B(x_0, r\eta) = \{x \in X; \|x - x_0\| \leq r\eta\} \subseteq \Omega_0 \), where \( r = \sum_{n=0}^{\infty} d_n \), the sequence \( \{x_n\} \) (depending on \( \alpha \)) defined in (1) and starting at \( x_0 \) converges to a solution \( x^* \) of the equation \( F(x) = 0 \). In that case, the solution \( x^* \) and the iterates \( x_n \) belong to \( \overline{B(x_0, r\eta)} \), and \( x^* \) is the only solution of \( F(x) = 0 \) in the open ball \( B(x_0, (2/kB) - r\eta) \cap \Omega_0 \).

Furthermore, we can give the following error estimates in terms of the real sequence \( \{d_n\} \) (or \( \{c_n\} \)):
\[
\|x^* - x_{n+1}\| \leq \sum_{k=n+1}^{\infty} d_k \eta \leq \frac{d_0}{\gamma} \sum_{k=n}^{\infty} \gamma^k, \quad \gamma = c_1/c_0.
\]
Proof. When $0 \leq \alpha < h(a)$ the convergence of the sequences $\{x_n\}$ follows immediately from the previous lemmas. When $\alpha = h(a)$ we have $c_n = c_0 = a$, for $n \geq 0$. Then, $a_n b_n = 1$, $a_{n+1} = \omega a_n$, where $\omega = g_2(a) > 1$. So $a_n = \omega^n a_0 = \omega^n$ and

$$d_n = \frac{1}{a_n} d_0 = \frac{d_0}{\omega^n}.$$

As we can see, $\{d_n\}$ is a Cauchy sequence and besides

$$\alpha c_n = h(a) a < 1, \quad n \geq 0.$$

Finally, in the same way that in the Lemma 3.4

$$a a_n d_n = a d_0 < 1$$

also holds. So the conditions required for the convergence of the sequences $\{a_n\}$, $\{b_n\}$, $\{c_n\}$ and $\{d_n\}$ are satisfied, and consequently the sequence $\{x_n\}$ is also convergent. If $x^*$ is the limit of the sequences $\{x_n\}$, then taking into account the bound (4) for $F(x_{n+1})$

$$\|F(x_{n+1})\| \leq \frac{k^2}{2} \left[ d_n^2 + \frac{b_n^2}{1 - \alpha c_n} \right],$$

that the limit of the sequences $\{b_n\}$ and $\{d_n\}$ is 0 (see Lemmas 3.6 and 3.7) and the continuity of $F$, we prove that $F(x^*) = 0$.

Besides we have $\|x_{n+1} - x_n\| \leq d_n \eta$ and therefore, for $p \geq 0$.

$$\|x_p - x_0\| \leq \|x_p - x_{p-1}\| + \cdots + \|x_1 - x_0\| \leq (d_{p-1} + \cdots + d_0) \eta.$$

By letting $p \rightarrow \infty$, we obtain the region where the solution is located, $\|x^* - x_0\| \leq r \eta$ and the error estimates (5).

Now, to show the unicity, suppose that $y^* \in B(x_0, (2/kB) - r \eta)$ is another solution of $F(x) = 0$. Then

$$0 = F(y^*) - F(x^*) = \int_0^1 F'(x^* + t(y^* - x^*)) \, dt (y^* - x^*).$$

Using, the estimate

$$\|I_0\| \int_0^1 \|F'(x^* + t(y^* - x^*)) - F'(x_0)\| \, dt \leq kB \int_0^1 \|x^* + t(y^* - x^*) - x_0\| \, dt$$

$$\leq kB \int_0^1 ((1 - t)\|x^* - x_0\| + t\|y^* - x_0\|) \, dt < \frac{kB}{2} \left( r \eta + \frac{2}{kB} - r \eta \right) = 1,$$

we have that the operator $\int_0^1 F'(x^* + t(y^* - x^*)) \, dt$ has an inverse and then $y^* = x^*$. \(\square\)

Next we give three examples where the conditions of the Main Theorem are satisfied and however conditions (2) or (3) fail.

Example 1. First, consider the function

$$f(x) = 2x^{7/3} + 23x$$

(6)
Table 1
Recurrence relations for Chebyshev method

<table>
<thead>
<tr>
<th>n</th>
<th>a_n</th>
<th>b_n</th>
<th>c_n</th>
<th>d_n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000000</td>
<td>1.000000</td>
<td>0.293097</td>
<td>1.146550</td>
</tr>
<tr>
<td>1</td>
<td>1.506139</td>
<td>0.510879</td>
<td>0.225525</td>
<td>0.568487</td>
</tr>
<tr>
<td>2</td>
<td>2.010749</td>
<td>0.172140</td>
<td>0.101450</td>
<td>0.180872</td>
</tr>
<tr>
<td>3</td>
<td>2.250662</td>
<td>0.020564</td>
<td>0.013565</td>
<td>0.020703</td>
</tr>
<tr>
<td>4</td>
<td>2.281826</td>
<td>0.000284</td>
<td>0.000190</td>
<td>0.000284</td>
</tr>
</tbody>
</table>

Table 2
Recurrence relations for Halley method

<table>
<thead>
<tr>
<th>n</th>
<th>a_n</th>
<th>b_n</th>
<th>c_n</th>
<th>d_n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000000</td>
<td>1.000000</td>
<td>0.293097</td>
<td>1.171713</td>
</tr>
<tr>
<td>1</td>
<td>1.523058</td>
<td>0.567966</td>
<td>0.253542</td>
<td>0.650421</td>
</tr>
<tr>
<td>2</td>
<td>2.146213</td>
<td>0.249250</td>
<td>0.156790</td>
<td>0.270452</td>
</tr>
<tr>
<td>3</td>
<td>2.586198</td>
<td>0.053270</td>
<td>0.040379</td>
<td>0.054368</td>
</tr>
<tr>
<td>4</td>
<td>2.697361</td>
<td>0.002313</td>
<td>0.001828</td>
<td>0.002315</td>
</tr>
</tbody>
</table>

defined in $X = [-3, 3]$ and let $x_0 = 1$. Notice that $f''(x)$ does not satisfy the Lipschitz condition (2) in any interval containing the solution $x^* = 0$. So Kantorovich theorem is not available in this case. However,

$$\|F_0\| = \left| \frac{1}{f'(x_0)} \right| = \frac{3}{83} = B, \quad \|F_0F(x_0)\| = \left| \frac{f(x_0)}{f'(x_0)} \right| = \frac{75}{83} = \eta,$$

$$\sup_{x \in X} \|F''(x)\| = f''(3) = \frac{56\sqrt{3}}{9} = k,$$

and then $a = 0.2930976 \ldots < a_0$. So to be able to apply the Main Theorem we must take $\alpha$ such that $0 \leq \alpha \leq h(a) = 0.976787 \ldots$. Let us consider with a little more detail the cases $\alpha = 0$ (Chebyshev method) and $\alpha = 1/2$ (Halley method). From the recurrence relations we obtain the following sequences that we show in Tables 1 and 2.

In both cases, we have $r\eta = \sum_{n=0}^{\infty} d_n \eta < 2$ and then $B(x_0, r \eta) \subseteq X$. Then Chebyshev and Halley sequences converge to $x^* = 0$, a solution of (6). In Table 3 we show these sequences together with Newton's method, so we can compare the velocity of convergence in that situation.

Example 2. In this example we show a function (see [24]) that does not satisfy the Smale-like conditions (3) for the Halley method. In [24] it has been proved that the constant $h$ appearing in (3) for Halley method is $3 - 2\sqrt{2}$. Let us consider the polynomial

$$p(x) = x^3 + x^2 - x + 0.2$$
and \( x_0 = 0 \). Then

\[
\frac{1}{2!} \left| \frac{f(0)}{f'(0)} \right| \frac{f''(0)}{f'(0)} = 0.2 > 3 - 2\sqrt{2}.
\]

So the starting point \( x_0 = 0 \) does not satisfy condition (3). However, our theorem can be used because \( f \) is bounded on any closed interval containing \( x_0 = 0 \).

Finally, Smale-like theorems does not work when we have nonanalytical functions. In that situation our theorem can be useful. We show this situation in the third example.

**Example 3.** Let us consider the function

\[
f(x) = 9x^{7/3} + 4x^2 - 36x + 9
\]

defined in \( X = [-1, 1] \) and let \( x_0 = 0 \). In that case the derivatives \( f^{(k)}(x) \) are not defined at \( x_0 \) for \( k \geq 3 \). Then the Smale-like conditions (3) do not work. However \( a = kB\eta = 0.25 < a_0 \) and then we can take the sequences defined in (1) for \( 0 \leq \alpha \leq h(a) = 2.12382 \ldots \). Let us consider for instance \( \alpha = 0 \) (Chebyshev method), \( \alpha = 1/2 \) (Halley method) and \( \alpha = 1 \) (super-Halley method). In all cases we have \( h\eta = \sum_{n=0}^{\infty} d_n \eta < 1 \) and then \( \overline{B}(x_0, r\eta) \subseteq X \). Then the three sequences converge to a solution of (6). In Tables 4 and 5 we compare the velocity of convergence of these sequences and Newton’s method.

### Table 3
**Newton, Chebyshev and Halley sequences**

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Newton</th>
<th>Chebyshev</th>
<th>Halley</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0000000000000000</td>
<td>1.0000000000000000</td>
<td>1.0000000000000000</td>
</tr>
<tr>
<td>1</td>
<td>0.096385542168747</td>
<td>0.004568134375485</td>
<td>-0.000346392451592</td>
</tr>
<tr>
<td>2</td>
<td>0.0004894689911883</td>
<td>-6.681578402264 ( 10^{-8} )</td>
<td>1.6283021649 ( 10^{-10} )</td>
</tr>
<tr>
<td>3</td>
<td>2.189086598062 ( 10^{-9} )</td>
<td>3.5005864407 ( 10^{-19} )</td>
<td>-2.7977225707 ( 10^{-25} )</td>
</tr>
<tr>
<td>4</td>
<td>7.214193075678 ( 10^{-22} )</td>
<td>-1.6688534227 ( 10^{-45} )</td>
<td>9.8923731005 ( 10^{-60} )</td>
</tr>
<tr>
<td>5</td>
<td>5.411849026009 ( 10^{-51} )</td>
<td>6.3835965504 ( 10^{-107} )</td>
<td>-4.0593589051 ( 10^{-140} )</td>
</tr>
<tr>
<td>6</td>
<td>5.961846623225 ( 10^{-119} )</td>
<td>-3.1470887156 ( 10^{-250} )</td>
<td>1.0943519878 ( 10^{-337} )</td>
</tr>
</tbody>
</table>

### Table 4
**Newton vs. Chebyshev**

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Newton</th>
<th>Chebyshev</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0000000000000000</td>
<td>0.0000000000000000</td>
</tr>
<tr>
<td>1</td>
<td>0.2500000000000000</td>
<td>0.2569444444444444</td>
</tr>
<tr>
<td>2</td>
<td>0.2696904362485184</td>
<td>0.2698550563853907</td>
</tr>
<tr>
<td>3</td>
<td>0.2698560959192613</td>
<td>0.2698560959192614</td>
</tr>
<tr>
<td>4</td>
<td>0.2698560959192614</td>
<td>0.2698560959192614</td>
</tr>
</tbody>
</table>
Table 5
Halley vs. super-Halley

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Halley</th>
<th>super-Halley</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0000000000000000</td>
<td>0.0000000000000000</td>
</tr>
<tr>
<td>1</td>
<td>0.2571428571428571</td>
<td>0.2573529411764706</td>
</tr>
<tr>
<td>2</td>
<td>0.2698554666753525</td>
<td>0.2698558470367463</td>
</tr>
<tr>
<td>3</td>
<td>0.2698560959192614</td>
<td>0.2698560959192614</td>
</tr>
</tbody>
</table>

4. Final remarks

The aim of this paper is to give a general vision on third order iterative methods and to soften the classical convergence conditions. In practice, the difficulties coming from the evaluation of the second Fréchet derivative are usually harder than the advantages because of the order of these methods. So, Newton’s method and its modifications are widely used. Nevertheless, third order methods are interesting in some applications. For instance, they provide better information on the existence and uniqueness of a root [13]. They are also used for solving quadratic equations [3, 14], that is, equations where the operator involved has constant second Fèchet derivative. This kind of equations appears often in the numerical treatment of nonlinear integral and differential equations (see [20] and the references given there).

In [6] we can find an application of third-order methods to hardware programming, where these methods are used to evaluate the inverse of an elementary function. In that paper, three iterations of Newton’s method are compared with two steps of the Halley and the Chebyshev ones (resulting six order in all cases). It is shown that the computational cost is very similar and, furthermore, two steps in a third-order method go faster than three steps of Newton’s method.

Finally, we show an application to matrix calculations. The following notes are merely orientative and we do not go into further details about computational questions. Let $A$ be a regular square matrix of order $n$. We can use iterative processes in order to calculate $A^{-1}$. To do that we solve the equation

$$F(X) = X^{-1} - A = 0,$$

where $F$ is an operator defined on the set of regular square matrices of order $n$. So, Newton’s method provide the iteration

$$X_{k+1} = (2I_n - X_k A)X_k, \quad k \geq 0,$$

where $I_n$ denotes the identity matrix. The Chebyshev method can be written in the form

$$Y_{k+1} = [3I_n - (3I_n - Y_k A)Y_k A] Y_k, \quad k \geq 0.$$

We establish a comparison between the above methods referring to the obtained results and the number of operations which have been done. We consider that sums and multiplications by a small integer are irrelevant in the computational cost and, so, we ignore them. As we can see, for three Newton steps we have to do six matrix multiplications, the same that for two Chebyshev steps. Although the computational effort is very similar, the results used to be better for Chebyshev method.
Let $A$ be the matrix

$$A = \begin{pmatrix} 1 & 2 & 0.3 \\ -1 & -0.1 & 12 \\ 0.01 & 1 & -0.5 \end{pmatrix}$$

and

$$X_0 = Y_0 = \begin{pmatrix} 0.5 & 0 & -1.5 \\ 0 & 0 & 1 \\ 0.1 & 0.1 & -0.1 \end{pmatrix}.$$ 

Then three iterations in Newton's method give

$$X_3 = \begin{pmatrix} 0.914675 & -0.099531 & -1.843852 \\ 0.029273 & 0.038657 & 0.945392 \\ 0.076776 & 0.075329 & -0.146034 \end{pmatrix},$$

whereas two iterations in Chebyshev method give

$$Y_2 = \begin{pmatrix} 0.916596 & -0.099726 & -1.845456 \\ 0.029241 & 0.038659 & 0.945421 \\ 0.076782 & 0.075329 & -0.146039 \end{pmatrix}.$$ 

Then, $Y_2$ is a better approximation of the exact solution

$$A^{-1} = \begin{pmatrix} 0.918545 & -0.099925 & -1.84708 \\ 0.029209 & 0.038663 & 0.945448 \\ 0.076788 & 0.075328 & -0.146045 \end{pmatrix}.$$

Third-order iterative processes can also be used in other matrix calculations, such as the square root or the $j$th root, $j \geq 2$.

Acknowledgements

The authors would like to thank the referees for his comments.

References