Available online at www.sciencedirect.com
ScienceDirect
Journal of Complexity 24 (2008) 185 – 197
Journal of COMPLEXITY
www.elsevier.com/locate/jco

A note on a modification of Moser's method

J.M. Gutiérrez, M.A. Hernández, N. Romero*

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

Received 20 February 2007; accepted 27 April 2007
Available online 8 May 2007

Abstract

We use a recurrence technique to obtain semilocal convergence results for Ulm's iterative method to approximate a solution of a nonlinear equation $F(x) = 0$

$$\begin{cases} x_{n+1} = x_n - B_n F(x_n), & n \geqslant 0, \\ B_{n+1} = 2B_n - B_n F'(x_{n+1}) B_n, & n \geqslant 0. \end{cases}$$

This method does not contain inverse operators in its expression and we prove it converges with the Newton rate. We also use this method to approximate a solution of integral equations of Fredholm-type.

MSC: 45G10; 47H17; 65J15

Keywords: Iterative processes; *R*-order of convergence; Semilocal convergence

1. Introduction

In this paper we consider an operator F defined in an open, convex and nonempty subset Ω of a Banach space X with values in another Banach space Y.

We consider the problem of approximating a solution x^* of a nonlinear equation

$$F(x) = 0. \tag{1}$$

Without any doubt Newton's method is the most used iterative process to solve this problem. It is given by the algorithm: $x_{n+1} = x_n - F'(x_n)^{-1} F(x_n)$, $n \geqslant 0$ for x_0 given. This iterative process

* Corresponding author.
E-mail addresses: jmguti@unirioja.es (J.M. Gutiérrez), mahernan@unirioja.es (M.A. Hernández), natalia.romero@unirioja.es (N. Romero).

0885-064X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jco.2007.04.003

has quadratic R-order of convergence so its speed of convergence and its operational cost is quite balanced.

Other methods, such as higher order methods also include in their expression the inverse of the operator F'. To avoid this problem, Newton-type methods: $x_{n+1} = x_n - H_n F(x_n)$, where H_n is an approximation of $F'(x_n)^{-1}$ are considered. One of these methods was proposed by Moser in [4] in this way. Given $x_0 \in \Omega$ and $A_0 \in \mathcal{L}(Y, X)$, the following sequences are defined

$$x_{n+1} = x_n - A_n F(x_n), \quad n \geqslant 0, \tag{2}$$

$$A_{n+1} = A_n - A_n(F'(x_n)A_n - I_X), \quad n \geqslant 0, \tag{3}$$

where I_X is the identity operator in X. The first equation is similar to Newton's method, but replacing the operator $F'(x_n)^{-1}$ by a linear operator A_n. The second equation is Newton's method applied to equation $g_n(A) = 0$ where $g_n : \mathcal{L}(Y, X) \to \mathcal{L}(X, Y)$ is defined by $g_n(A) = A^{-1} - F'(x_n)$. So $\{A_n\}$ gives us an approximation of $F'(x_n)^{-1}$.

In addition, it can be shown that the rate of convergence for the above scheme is $(1 + \sqrt{5})/2$, provided the root of (1) is simple [4]. However, this is unsatisfactory from a numerical point of view because the scheme uses the same amount of information per step as Newton's method, yet, it converges no faster than the secant method.

Moser's method was developed as a technical tool for investigating the stability of the N-body problem in celestial mechanics. The main difficulty in this, and similar problems involving small divisors, is the solution of a system of nonlinear partial differential equations. In essence, Moser's idea is to solve the problem by a sequence of changes of variables.

In [10] Ulm proposed the following iterative method to solve nonlinear equations. Given $x_0 \in \Omega$ where F is a Fréchet-differentiable operator and $B_0 \in \mathcal{L}(Y, X)$, Ulm defines

$$\begin{cases} x_{n+1} = x_n - B_n F(x_n), & n \geqslant 0, \\ B_{n+1} = 2B_n - B_n F'(x_{n+1})B_n, & n \geqslant 0. \end{cases} \tag{4}$$

Notice that, here $F'(x_{n+1})$ appears instead of $F'(x_n)$ in (3). This is crucial for obtaining fast convergence. Under certain assumptions, Ulm showed, that the method generates successive approximations that converge to a solution of (1), asymptotically as fast as Newton's method. Ulm applied the method to several particular classes of operator equations.

The method exhibits several attractive features. First, it converges with the Newton rate. Second, it is inverse free: you do not need to solve a linear equation at each iteration. Third, in addition to solve the nonlinear equation (1), the method produces successive approximations $\{B_n\}$ to the value of $F'(x^*)^{-1}$, being x^* a solution of (1). This property is very helpful when one investigates the sensitivity of the solution to small perturbations.

Although method (4) was firstly proposed by Ulm [10], it has been also considered by other authors. For instance, Hald [1] showed the quadratic convergence of the method. Later, Zehnder [13] or Petzeltova [5] have studied the convergence of the method under Kantorovich-type conditions.

An alternative to Kantorovich theory to study the convergence of iterative processes to solve nonlinear equations is given by the known as Smale's point estimate theory [8,9]. Roughly speaking, if x_0 is an initial value such that the sequence $\{x_n\}$ satisfies

$$\|x_n - x^*\| \leqslant \left(\frac{1}{2}\right)^{2^n-1} \|x_0 - x^*\|$$

then x_0 is said to be an approximate zero of F, (see [9]). The following conditions were introduced by Smale [8,9] in order to prove that x_0 is an approximated zero

$$\begin{aligned}
&\text{(i) } \|F'(x_0)^{-1}F(x_0)\| \leqslant \beta,\\
&\text{(ii) } \sup_{k\geqslant 2}\left(\frac{1}{k!}\|F'(x_0)^{-1}F^{(k)}(x_0)\|\right)^{1/(k-1)} \leqslant \gamma,\\
&\text{(iii) } \alpha = \beta\gamma \leqslant 3 - 2\sqrt{2}.
\end{aligned} \tag{5}$$

Wang and Zhao [11] pointed that condition (5) is too restrictive. Instead of (5) they assume

$$\begin{aligned}
&\text{(i}'\text{) } \|F'(x_0)^{-1}F(x_0)\| \leqslant \beta,\\
&\text{(ii}'\text{) } \frac{1}{k!}\|F'(x_0)^{-1}F^{(k)}(x_0)\| \leqslant \gamma_k, \quad k \geqslant 2.\\
&\text{(iii}'\text{) The equation } \phi(t) = 0 \text{ has at least a positive solution, where}\\
&\phi(t) = \beta - t + \sum_{k\geqslant 2}\gamma_k t^k.
\end{aligned} \tag{6}$$

Then they study the convergence of Newton's method by constructing a majorizing sequence in terms of the function $\phi(t)$.

In this paper, we are going to consider Wang–Zhao's type-hypothesis in order to study the convergence of Ulm's method (4). However, we are going to construct a system of recurrence relations in order to analyze the convergence of the method instead of considering the aforesaid majorizing sequence. This technique has been used successfully to prove the convergence of Newton's method and other iterative methods under different conditions, for instance see [2].

The main theorem we show in this paper (Theorem 3) provides a new result on the semilocal convergence for the iterative process given in (4) under conditions similar to the given in (6). So we consider Eq. (1), where F is a nonlinear analytic operator in an open convex nonempty subset Ω of a Banach space X in another Banach space Y. For a given $x_0 \in \Omega$ and $B_0 \in \mathcal{L}(Y, X)$ let us assume:

(C1) $\|B_0\| \leqslant \beta$.

(C2) $\|F(x_0)\| \leqslant \eta$.

(C3) $\|I - F'(x_0)B_0\| \leqslant \delta$.

(C4) $\frac{1}{k!}\|F^{(k)}(x_0)\| \leqslant \gamma_k, \quad k \geqslant 2$.

(C5) $S > 0$ is the radius of convergence of the power series $\sum_{k\geqslant 2}\gamma_k t^k$, that is,

$$S = \liminf_{k\to\infty} |\gamma_k|^{-1/k}.$$

(C6) The equation

$$t(1 - \Delta(t)) - \beta\eta = 0 \tag{7}$$

has at least a solution in $(0, S]$, where $\Delta(t) = (1 + \delta + \beta^2\eta h(t))\left(\delta + \dfrac{\beta^2\eta h(t)}{2}\right)$, being $h(t) = \sum_{k\geqslant 2} k(k-1)\gamma_k t^{k-2}$. We denote by R the smallest root of (7) in $(0, S]$.

(C7) $B(x_0, R) \subset \Omega$.

We prove that, under these conditions (C1)–(C7), the iterative process (4) is convergent to a solution of (1) with at least quadratic R-order of convergence. In addition, we find the domains where the solution is located and where it is unique, together with some estimates of the error.

In Section 2, we introduce the recurrence sequences associated to the method (4). They allow us to obtain semilocal convergence results (Section 3) and to calculate the R-order of convergence (Section 4). We prove that (4) converges to a solution of a nonlinear equation as fast as Newton's method.

In Section 5, we finish this study analyzing the application to nonlinear integral equations of Fredholm-type in this way

$$x(s) = f(s) + \lambda \int_a^b k(s,t)x(t)^p\,dt, \quad s \in [a,b].$$

To solve these integral equations it is equivalent to solve nonlinear integral equations in the form (1), where F is a nonlinear p-times Fréchet-differentiable operator.

Through this paper we denote,

$$B(x_0,R) = \{x \in X;\ \|x - x_0\| < R\} \quad \text{and} \quad \overline{B(x_0,R)} = \{x \in X;\ \|x - x_0\| \leqslant R\}.$$

2. Recurrence relations

In this section, we define the auxiliary functions and we construct the sequences that we need to establish the convergence of the method (4) under conditions (C1)–(C7). We also give some technical lemmas that we need in the proof of the main theorem (Theorem 3).

To prove that sequence (4) is well defined, we need some definitions and lemmas. First, we define the auxiliary functions

$$f(t,u) = 1 + t + uh(R), \quad g(t,u) = t + \frac{u\,h(R)}{2}, \tag{8}$$

and, for all $n \geqslant 1$, the following real sequences:

$$\begin{aligned} a_n &= (a_{n-1} + b_{n-1}h(R)d_{n-1})^2,\\ b_n &= f(a_{n-1}, b_{n-1}d_{n-1})b_{n-1},\\ d_n &= f(a_{n-1}, b_{n-1}d_{n-1})g(a_{n-1}, b_{n-1}d_{n-1})d_{n-1}, \end{aligned} \tag{9}$$

where $a_0 = \delta$, $b_0 = \beta$, $d_0 = \beta\eta$.

Observe that, the real function $h(t)$ is analytic in $(0, S]$ and the two real functions f and g given, are nondecreasing in the both arguments, for $t > 0$ and $u > 0$.

Notice that $\|x_1 - x_0\| \leqslant \beta\eta = d_0$. If we suppose that $x_1 \in B(x_0,R)$, then $x_1 \in \Omega$ and therefore

$$\begin{aligned} \|F'(x_1) - F'(x_0)\|\|B_0\| &= \left\| \sum_{k\geqslant 2} \frac{1}{(k-1)!} F^{(k)}(x_0)(x_1 - x_0)^{k-1} \right\| \|B_0\| \\ &\leqslant \left(\sum_{k\geqslant 2} k\,(k-1)\gamma_k \|x_1 - x_0\|^{k-2} \right) \|x_1 - x_0\|\|B_0\| \\ &\leqslant h(\|x_1 - x_0\|)\|x_1 - x_0\|\|B_0\| < b_0 h(R) d_0. \end{aligned}$$

Now, taking into account (C3), we obtain

$$\|I - F'(x_1)B_0\| \leqslant \|I - F'(x_0)B_0\| + \|F'(x_1) - F'(x_0)\|\|B_0\| \leqslant a_0 + b_0 h(R) d_0.$$

Therefore,

$$\|B_1 - B_0\| \leqslant \|B_0\| \|I - F'(x_1)B_0\| \leqslant b_0(a_0 + b_0 h(R) d_0).$$

On the other hand, $x_1 \in B(x_0, R)$ and from (C7), we obtain that $x_0 + \omega(x_1 - x_0) \in B(x_0, R) \subset \Omega$ for $\omega \in (0, 1)$. Taking into account that h is a nondecreasing function, it follows:

$$\begin{aligned}
&\|B_0\| \|F''(x_0 + \omega(x_1 - x_0))\| \|x_1 - x_0\| \\
&\quad = \|B_0\| \left\| \sum_{k \geqslant 2} \frac{k\,(k-1)}{k!}\, F^{(k)}(x_0)\,(x_0 + \omega(x_1 - x_0) - x_0)^{k-2} \right\| \|x_1 - x_0\| \\
&\quad \leqslant \|B_0\| h(\|x_1 - x_0\|) \|x_1 - x_0\| < b_0 h(R) d_0.
\end{aligned}$$

Now, from Taylor's formula and (4), we have for $n \geqslant 1$:

$$\begin{aligned}
F(x_n) &= F(x_{n-1}) + F'(x_{n-1})\,(-B_{n-1}F(x_{n-1})) + \int_{x_{n-1}}^{x_n} F''(x)(x_n - x)\,dx \\
&= \left(I - F'(x_{n-1})B_{n-1}\right) F(x_{n-1}) \\
&\quad + \int_0^1 F''(x_{n-1} + \omega(x_n - x_{n-1}))\,(1-\omega)\,d\omega (x_n - x_{n-1})^2. \qquad (10)
\end{aligned}$$

Taking norms for $n = 1$, it follows:

$$\begin{aligned}
\|F(x_1)\| &\leqslant \left\|I - F'(x_0)B_0\right\| \|F(x_0)\| + \int_0^1 \left\|F''(x_0 + \omega(x_1 - x_0))\right\| (1-\omega)\,d\omega\ \|x_1 - x_0\|^2 \\
&\leqslant \left(a_0 + \int_0^1 \|B_0\| \left\|F''(x_0 + \omega(x_1 - x_0))\right\| \|x_1 - x_0\| (1-\omega)\,d\omega\right) \|F(x_0)\| \\
&\leqslant \left(a_0 + \frac{b_0 h(R) d_0}{2}\right) \|F(x_0)\| = g(a_0, b_0 d_0)\, \|F(x_0)\|\,.
\end{aligned}$$

Moreover, $\|B_1\| \leqslant \|B_0\| \left\|2I - F'(x_1)B_0\right\| \leqslant b_1$, and consequently $\|x_2 - x_1\| \leqslant d_1$. Then,

$$\|x_2 - x_0\| \leqslant \|x_2 - x_1\| + \|x_1 - x_0\| \leqslant d_1 + d_0.$$

Besides, if $x_2 \in B(x_0, R) \subset \Omega$, we obtain

$$\begin{aligned}
&\|F'(x_2) - F'(x_1)\| \|B_1\| \\
&\quad = \left\| \sum_{k \geqslant 2} \frac{1}{(k-1)!} F^{(k)}(x_0) \left[(x_2 - x_0)^{k-1} - (x_1 - x_0)^{k-1}\right] \right\| \|B_1\| \\
&\quad \leqslant \left\| \sum_{k \geqslant 2} \frac{1}{(k-1)!} F^{(k)}(x_0)(x_2 - x_1) \left(\sum_{j=0}^{k-2} (x_2 - x_0)^{k-2-j} (x_1 - x_0)^j \right) \right\| \|B_1\| \\
&\quad \leqslant \sum_{k \geqslant 2} k\gamma_k \left(\sum_{j=0}^{k-2} \|x_2 - x_0\|^{k-2-j} \|x_1 - x_0\|^j \right) \|x_2 - x_1\| \|B_1\| \\
&\quad \leqslant \sum_{k \geqslant 2} k(k-1)\gamma_k R^{k-2} \|x_2 - x_1\| \|B_1\| < b_1 h(R) d_1,
\end{aligned}$$

and $\|I - F'(x_1)B_1\| = \|(I - F'(x_1)B_0)^2\| \leqslant a_1$, so that

$$\|I - F'(x_2)B_1\| \leqslant \|I - F'(x_1)B_1\| + \|F'(x_2) - F'(x_1)\| \|B_1\| \leqslant a_1 + b_1 h(R) d_1.$$

Furthermore,

$$\|B_2 - B_1\| \leqslant \|B_1\| \|I - F'(x_2)B_1\| \leqslant b_1(a_1 + b_1 h(R) d_1).$$

On the other hand, since $\omega \in (0, 1)$, it follows that $x_1 + \omega(x_2 - x_1) \in \Omega$, and moreover

$$\begin{aligned}
&\|B_1\| \|F''(x_1 + \omega(x_2 - x_1))\| \|x_2 - x_1\| \\
&\quad = \|B_1\| \left\| \sum_{k\geqslant 2} \frac{k\,(k-1)}{k!} F^{(k)}(x_0)\,(x_1 + \omega(x_2 - x_1) - x_0)^{k-2} \right\| \|x_2 - x_1\| \\
&\quad \leqslant \|B_1\| \left(\sum_{k\geqslant 2} k\,(k-1)\gamma_k\,((1-\omega)\|x_1 - x_0\| + \omega\|x_2 - x_0\|)^{k-2} \right) \|x_2 - x_1\| \\
&\quad < \|B_1\| \left(\sum_{k\geqslant 2} k\,(k-1)\gamma_k R^{k-2} \right) \|x_2 - x_1\| < b_1 h(R) d_1.
\end{aligned}$$

Now, as in the case $n = 1$, taking norms in (10) for $n = 2$, it follows:

$$\|F(x_2)\| \leqslant g(a_1, b_1 d_1) \|F(x_1)\|.$$

Moreover, $\|B_2\| \leqslant \|B_1\| \|2I - F'(x_2)B_1\| \leqslant b_2$, and consequently $\|x_3 - x_2\| \leqslant d_2$. Then

$$\|x_3 - x_0\| \leqslant \|x_3 - x_2\| + \|x_2 - x_0\| \leqslant d_2 + d_1 + d_0.$$

Now we present a system of recurrence relations in the next lemma. The proof of lemma follows from a similar way that the previous reasoning and using induction. Besides, it allows to proof the convergence of iterative process given in (4).

Lemma 1. *Let us suppose* (C1)–(C7) *and $x_n \in B(x_0, R)$, for $n \in \mathbb{N}$, then the following recurrence relations hold*:

[I$_n$] $\|I - F'(x_n)B_n\| \leqslant a_n$,
[II$_n$] $\|F'(x_{n+1}) - F'(x_n)\| \|B_n\| < b_n h(R) d_n$,
[III$_n$] $\|I - F'(x_{n+1})B_n\| \leqslant a_n + b_n h(R) d_n$,
[IV$_n$] $\|B_{n+1} - B_n\| \leqslant b_n(a_n + b_n h(R) d_n)$,
[V$_n$] If $\omega \in (0, 1)$: $\|B_n\| \|F''(x_n + \omega(x_{n+1} - x_n))\| \|x_{n+1} - x_n\| < b_n h(R) d_n$,
[VI$_n$] $\|F(x_n)\| \leqslant g(a_{n-1}, b_{n-1} d_{n-1}) \|F(x_{n-1})\|$,
[VII$_n$] $\|B_n\| \leqslant f(a_{n-1}, b_{n-1} d_{n-1}) \|B_{n-1}\| \leqslant b_n$,
[VIII$_n$] $\|x_{n+1} - x_n\| \leqslant f(a_{n-1}, b_{n-1} d_{n-1}) g(a_{n-1}, b_{n-1} d_{n-1}) \|B_{n-1}\| \|F(x_{n-1})\| \leqslant d_n$,
[IX$_n$] $\|x_{n+1} - x_0\| \leqslant \sum_{j=0}^{n} d_j$.

3. Semilocal convergence

From the previous recurrence relations, we prove that the sequence $\{x_n\}$ is well-defined. We carry out the study of the convergence of sequence $\{x_n\}$ given in (4). We see that the sequence $\{x_n\}$ converges to a solution x^* of Eq. (1). For this, we consider the following properties of the sequences $\{a_n\}$ and $\{b_n d_n\}$:

Lemma 2. *If a_0, b_0 and d_0 satisfy*

$$(a_0 + b_0 h(R) d_0)^2 \leqslant a_0 \quad \textit{and} \quad f(a_0, b_0 d_0)^2 g(a_0, b_0 d_0) \leqslant 1, \tag{11}$$

then the sequences $\{a_n\}$ and $\{b_n d_n\}$ are decreasing. In addition, $f(a_0, b_0 d_0) g(a_0, b_0 d_0) < 1$.

To establish the convergence of the sequence $\{x_n\}$ it is enough to prove that it is a Cauchy sequence, because the sequence is defined in a Banach space. We provide the following semilocal convergence theorem, which show the domain of existence of the solution.

Theorem 3. *Let X and Y be two Banach spaces and $F : \Omega \subseteq X \to Y$ a nonlinear analytic operator on a nonempty open convex domain Ω. Let $x_0 \in \Omega$ and $B_0 \in \mathcal{L}(Y, X)$. Assume all conditions* (C1)–(C7) *hold. If* (11) *is satisfied, then the sequence $\{x_n\}$ defined in* (4) *and starting from x_0, converges to a solution x^* of Eq.* (1). *The solution x^* and the iterations x_n belong to $\overline{B(x_0, R)}$.*

Proof. From Eq. (7), as $\Delta(R) = f(a_0, b_0 d_0) g(a_0, b_0 d_0) < 1$, we obtain that $\|x_1 - x_0\| \leqslant d_0 < R$. Then it follows that $x_1 \in B(x_0, R)$.

From (9) and previous Lemma 2, we have for $m \geqslant 1$:

$$\begin{aligned} d_m &= f(a_{m-1}, b_{m-1} d_{m-1}) g(a_{m-1}, b_{m-1} d_{m-1}) d_{m-1} \\ &= f(a_{m-2}, b_{m-2} d_{m-2}) g(a_{m-2}, b_{m-2} d_{m-2}) f(a_{m-1}, b_{m-1} d_{m-1}) \\ &\quad\; g(a_{m-1}, b_{m-1} d_{m-1}) d_{m-2} \\ &= \left[\prod_{k=0}^{m-1} f(a_k, b_k d_k) g(a_k, b_k d_k) \right] d_0 \\ &\leqslant f(a_0, b_0 d_0)^m g(a_0, b_0 d_0)^m d_0. \end{aligned}$$

Then, from [IX$_n$] in Lemmas 1 and 2, taking into account (7) we obtain:

$$\begin{aligned} \|x_m - x_0\| &\leqslant \sum_{k=0}^{m-1} d_k \leqslant \left(\sum_{k=0}^{m-1} f(a_0, b_0 d_0)^k g(a_0, b_0 d_0)^k \right) d_0 \\ &\leqslant \frac{1 - \Delta(R)^m}{1 - \Delta(R)} d_0 < \frac{1}{1 - \Delta(R)} d_0 = R. \end{aligned}$$

Therefore, $x_m \in B(x_0, R)$ for all $m \geqslant 1$. Now, from [VIII$_n$] in Lemmas 1 and 2, we obtain easily that

$$\|x_{n+2} - x_{n+1}\| \leqslant \left(\prod_{k=0}^{n} f(a_k, b_k d_k) g(a_k, b_k d_k) \right) \|B_0\| \|F(x_0)\| \leqslant \Delta(R)^{n+1} d_0.$$

Next, we establish the convergence of the sequence $\{x_n\}$ given by (4). For this, we consider $n, m \in \mathbb{N}$, and using the previous bound, we obtain:

$$\|x_{n+m} - x_n\| \leqslant \sum_{k=0}^{m-1} \left(\Delta(R)^{k+n} \right) d_0 = \Delta(R)^n \frac{1 - \Delta(R)^m}{1 - \Delta(R)} d_0. \tag{12}$$

Then, since $\Delta(R) < 1$, $\{x_n\}$ is a Cauchy sequence and converges to $x^* = \lim_{n\to\infty} x_n$. Besides, for $n = 0$ in (12), we have $x^* \in \overline{B(x_0, R)}$ by letting $m \to \infty$.

Now, we prove that the sequence $\{x_n\}$ converges to a solution x^* of (1). For this, we have

$$\|F(x_n)\| \leqslant \left(\prod_{k=0}^{n-1} g(a_k, b_k d_k)\right) \|F(x_0)\| \leqslant g(a_0, b_0 d_0)^n \|F(x_0)\|,$$

and since $g(a_0, b_0 d_0) < 1$ from (11), by letting $n \to \infty$ it follows $\lim_{n\to\infty} \|F(x_n)\| = 0$ and $F(x^*) = 0$. So x^* is a solution of (1) in the closed ball $\overline{B(x_0, R)}$. $\square$

Next, we prove the uniqueness of the solution. Now we introduce a new parameter $\tilde{a_0}$ such that:

$$\|I - B_0 F'(x_0)\| \leqslant \tilde{a_0}.$$

Notice that in general the operators B_0 and $F'(x_0)$ could not to commute, so the parameter $\tilde{a_0}$ could be different form $a_0 = \delta$ defined by (C3). The case $\tilde{a_0} = a_0$ happens, for instance, if we take $B_0 = F'(x_0)^{-1}$. This situation is analyzed in the first remark of Section 6.

Theorem 4. *Let $\tilde{a_0}$ be the parameter defined above and r the biggest positive root of the equation*

$$\sigma(t) + \frac{1 - \tilde{a_0}}{b_0}(R - t) = \sigma(R), \tag{13}$$

where $\sigma(t) = \sum_{k\geqslant 2} \gamma_k t^k$, then the solution x^* of (1) is unique in $B(x_0, r) \cap \Omega$.

Proof. Let us assume that $z^* \in B(x_0, r) \cap \Omega$ is a different solution of (1). Then, it follows that

$$0 = B_0(F(z^*) - F(x^*)) = \int_0^1 B_0 F'(x^* + \omega(z^* - x^*))\, d\omega\, (z^* - x^*).$$

If the operator A^{-1} exists, where

$$A = \int_0^1 B_0 F'(x^* + \omega(z^* - x^*))\, d\omega,$$

we have $z^* - x^* = 0$, and then the unicity of solution is obtained in $B(x_0, r) \cap \Omega$.

Then, from

$$\begin{aligned} &F'(x^* + \omega(z^* - x^*)) - F'(x_0) \\ &\quad = \sum_{k\geqslant 2} \frac{1}{(k-1)!} F^{(k)}(x_0)\big((1-\omega)(x^* - x_0) + \omega(z^* - x_0)\big)^{k-1}, \end{aligned}$$

and by taking norms, we have

$$\|F'(x^* + \omega(z^* - x^*)) - F'(x_0)\| \leqslant \sum_{k\geqslant 2} k\gamma_k \|(1-\omega)(x^* - x_0) + \omega(z^* - x_0)\|^{k-1}.$$

Besides, as $\|z^* - x_0\| < r$ and $\|x^* - x_0\| \leqslant R$, we have

$$\begin{aligned} \|I - A\| &\leqslant \|I - B_0 F'(x_0)\| + \|B_0 F'(x_0) - A\| < \tilde{a_0} + \frac{b_0}{r - R}\left(\sum_{k\geqslant 2} \gamma_k r^k - \sum_{k\geqslant 2} \gamma_k R^k\right) \\ &= \tilde{a_0} + \frac{b_0}{r - R}(\sigma(r) - \sigma(R)) = 1. \end{aligned}$$

Then, by Banach's lemma, A^{-1} exists and $z^* = x^*$. $\square$

Notice that R is a root of (13), therefore $r \geqslant R$.

4. The R-order of convergence

It is known [7,6] that in Banach spaces a sequence $\{x_n\}$ has R-order at least q if $\|x_n - x^*\| \leqslant C\tau^{q^n}$, for $\tau \in (0, 1)$ and $C \in \mathbb{R}^+$. In this section we show that, under conditions (C1)–(C7), the Ulm process provides a sequence which converges to a solution x^* of (1) with at least R-order two. Besides, we obtain a priori error bounds for the Ulm process in the approximation to the solution.

Lemma 5. *Let f and g be given in* (8). *Let us define $\theta_1 = b_1 d_1/b_0 d_0$, $\theta_2 = a_1/a_0$ and $\theta = \max\{\theta_1, \theta_2\}$. If* (11) *is satisfied, then*

(i) $f(\theta t, \theta u) < f(t, u)$, $g(\theta t, \theta u) = \theta g(t, u)$ *for* $\theta \in (0, 1)$,

(ii) $a_n \leqslant \theta^{2^{n-1}} a_{n-1} \leqslant \theta^{2^n - 1} a_0$,
$b_n d_n \leqslant \theta^{2^{n-1}} b_{n-1} d_{n-1} \leqslant \theta^{2^n - 1} b_0 d_0$.

Theorem 6. *Under the assumptions of Theorem* 3, *the Ulm process has R-order of convergence at least two. Moreover, the following error estimates are obtained*:

$$\|x^* - x_n\| < \frac{(\Delta(R)\theta^{-1})^n \theta^{2^n - 1}}{1 - \Delta(R)\theta^{2^n - 1}} d_0.$$

Proof. From the previous lemma and [VIII$_k$], we have

$$\|x_{k+1} - x_k\| \leqslant (\Delta(R)\theta^{-1})^k (\theta^{1/2})^{2^k - 1} \|B_0\| \|F(x_0)\|.$$

So, for $m \geqslant 1$, it follows:

$$\begin{aligned}
&\|x_{n+m} - x_n\| \\
&\quad \leqslant \sum_{k=n}^{n+m-1} \|x_{k+1} - x_k\| \leqslant \sum_{k=0}^{m-1} \left(\Delta(R)^{k+n} \theta^{-(k+n)} (\theta^{1/2})^{2^{k+n} - 1}\right) \|B_0\| \|F(x_0)\| \\
&\quad = \left(\frac{\Delta(R)}{\theta}\right)^n \theta^{(2^n - 1)/2} \left(1 + \frac{\Delta(R)}{\theta} \theta^{2^{n-1}} + \cdots + \left(\frac{\Delta(R)}{\theta}\right)^{m-1} \theta^{2^{n-1}(2^{m-1} - 1)}\right) d_0.
\end{aligned} \tag{14}$$

By Bernoulli's inequality, we obtain $2^k - 1 > k$ and therefore from (14) it follows:

$$\|x_{n+m} - x_n\| \leqslant \left(\frac{\Delta(R)}{\theta}\right)^n (\theta^{1/2})^{2^n - 1} \frac{1 - \left(\Delta(R)\theta^{2^{n-1} - 1}\right)^m}{1 - \Delta(R)\theta^{2^{n-1} - 1}} d_0. \tag{15}$$

By letting $m \to \infty$ in (15) and taking into account that $\theta < 1$ and $\Delta(R)\theta^{-1} < 1$, it follows:

$$\|x^* - x_n\| < \frac{(\Delta(R)\theta^{-1})^n (\theta^{1/2})^{2^n - 1}}{1 - \Delta(R)\theta^{2^{n-1} - 1}} d_0 < (\theta^{1/2})^{2^n} \frac{R}{\theta^{1/2}},$$

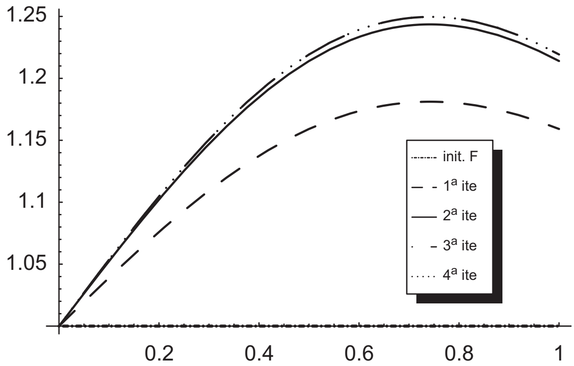

Fig. 1. Iterations $x_n(s)$ of Ulm's method.

and the Ulm process given by (4), has at least *R*-order of convergence two, since

$$\|x^* - x_n\| < C\tau^{2^n},$$

where $\tau = \theta^{1/2} < 1$ and $C = \frac{R}{\tau}$ is a positive real constant. □

5. Numerical experiment

In this section, we illustrate the theoretical results given for Ulm's method.

So we consider the following Fredholm nonlinear integral equation that appear in [3, p. 552]:

$$x(s) = 1 + \frac{1}{4}\int_0^1 \sin(s\pi t)x(t)^2\,dt. \tag{16}$$

To solve (16), it is equivalent to solve the nonlinear equation (1), where $F : \Omega \subseteq X \to Y$ is a nonlinear operator defined by

$$F(x)(s) = x(s) - 1 - \frac{1}{4}\int_0^1 \sin(s\pi t)x(t)^2 dt \tag{17}$$

and $X = Y = C[0, 1]$ is the space of continuous functions on the interval [0, 1], equipped with the max-norm, $\|x\| = \max_{s\in[a,b]} |x(s)|,\ x \in X$.

Firstly, we prove that all conditions (C1)–(C7) hold. If we choose as starting points $x_0(s) = 1$ and $B_0(y(s)) = y(s)$, we have $\|B_0\| = 1 = \beta$, moreover

$$\|F(x_0)\| = \frac{1}{4}\max_{s\in[0,1]}\left|\int_0^1 \sin(\pi st)\,dt\right| = 0.181153 = \eta,$$

$$\|I - F'(x_0)B_0\| = \frac{1}{2}\max_{s\in[0,1]}\left|\int_0^1 \sin(\pi st)\,dt\right| = 0.362306 = \delta,$$

$$\frac{1}{2!}\|F^{(2)}(x_0)\| = \frac{1}{4}\max_{s\in[0,1]}\left|\int_0^1 \sin(\pi st)\,dt\right| = 0.181153 = \gamma_2,$$

and $\frac{1}{k!}\|F^{(k)}(x_0)\| = 0 = \gamma_k$ for all $k > 2$. Hence, we check the convergence conditions of Theorem 3 and we obtain, $a_0 = 0.362306$, $b_0 = 1$, $d_0 = 0.181153$, $R = 0.415688$ and $h(R) = 0.362306$, so $(a_0 + b_0 h(R) d_0)^2 \leqslant a_0$ and $f(a_0, b_0 d_0)^2 g(a_0, b_0 d_0) \leqslant 1$. Therefore, the conditions of Theorem 3 hold and (17) has a solution x^* in $B(1, 0.415688)$. Moreover, in this case $B_0 = I$ and $\|I - B_0 F'(x_0)\| \leqslant \tilde{a_0} = a_0$, therefore it follows that the solution x^* is unique in $B(1, 3.10451)$.

On the other hand, if we demand a tolerance $\|x_{n+1} - x_n\| \leqslant C\ 10^{-5}$, where C is a positive real constant, we obtain that $\|x_4 - x_3\| \leqslant 5.1 \times 10^{-5}$, so we can consider the four iteration $x_4(s)$ as a good approximation of the solution. To finish this example, we show the first four approximations $x_n(s)$ that we calculate with starting point $x_0(s) = 1$ and six significant digits, using the *Mathematica* [12] program:

$$x_1(s) = 1 - 0.0795775 s^{-1}(-1 + \cos(3.14159 s)),$$

$$x_2(s) = 1 - 0.107004 s^{-1}(-1 + \cos(3.14159 s)),$$

$$x_3(s) = 1 - 0.109642 s^{-1}(-1 + \cos(3.14159 s)),$$

$$x_4(s) = 1 - 0.109668 s^{-1}(-1 + \cos(3.14159 s)).$$

The first four iterations of Ulm's method given in (4), to approximate a solution of (17), are shown in Fig. 1.

6. Concluding remarks

Remark 1. A special case of Ulm's method is when we consider $B_0 = F'(x_0)^{-1}$ and the first step for Ulm's method is the same as Newton's method. In this situation, we can obtain a new uniqueness result for Ulm's method. Thus, from $B_0 = F'(x_0)^{-1}$ it follows that $B_0^{-1} = F'(x_0)$ exists. Then, we consider $z^* \in B(x_0, R) \cap \Omega$ a different solution of $x^* \in B(x_0, R)$ of Eq. (1). Then, it follows that

$$0 = F(z^*) - F(x^*) = \int_0^1 F'(x^* + \omega(z^* - x^*))\, d\omega (z^* - x^*).$$

If the operator T^{-1} exists, where

$$T = \int_0^1 F'(x^* + \omega(z^* - x^*))\, d\omega,$$

then the unicity of solution is obtained in $B(x_0, r) \cap \Omega$.

Thus,

$$\|I - B_0 T\| = \|I - F'(x_0)^{-1} T\| \leqslant \|F'(x_0)^{-1}\| \|F'(x_0) - T\|$$

and

$$\left[T - F'(x_0)\right] = \int_0^1 \left[F'(x^* + \omega(z^* - x^*)) - F'(x_0)\right] d\omega.$$

Now, in the same way that in Theorem 4, it follows:

$$\|F'(x_0) - T\| < \frac{1}{r - R}(\sigma(r) - \sigma(R)).$$

Thus, if

$$\frac{b_0}{r-R}(\sigma(r)-\sigma(R))=1,$$

where $\|B_0\| \leqslant b_0$, we obtain that T^{-1} exists.

Therefore, if r is the biggest positive root of the equation

$$b_0(\sigma(t)-\sigma(R))=t-R,$$

then the solution x^* of (1) is unique in $B(x_0,r)\cap\Omega$.

Remark 2. Observe that the sequence $\{B_n\}$ converges to the bounded right inverse of $F'(x^*)$. Indeed, from [IV$_k$] it follows:

$$\|B_{k+1}-B_k\| \leqslant (a_k+b_kh(R)d_k)b_k \leqslant \theta^{2^k-1}(a_0+b_0h(R)d_0)f(a_0,b_0d_0)^k b_0,$$

since, the real function f is nondecreasing in the both arguments and $\{a_n\}$ and $\{b_nd_n\}$ are decreasing sequences. Consequently,

$$\begin{aligned}\|B_{n+m}-B_n\| &\leqslant \left(\sum_{k=0}^{m-1}\theta^{2^{n+k}-1}f(a_0,b_0d_0)^{n+k}\right)(a_0+b_0h(R)d_0)b_0\\ &\leqslant f(a_0,b_0d_0)^n\theta^{2^n-1}\left(\sum_{k=0}^{m-1}\theta^{2^n(2^k-1)}f(a_0,b_0d_0)^k\right)(a_0+b_0h(R)d_0)b_0\end{aligned}$$

and, applying the Bernoulli inequality, it follows that $\{B_n\}$ is a Cauchy sequence, since

$$\|B_{n+m}-B_n\| \leqslant f(a_0,b_0d_0)^n\theta^{2^n-1}\frac{1-(\theta^{2^n}f(a_0,b_0d_0))^m}{1-\theta^{2^n}f(a_0,b_0d_0)}(a_0+b_0h(R)d_0)b_0.$$

Thus, the sequence $\{B_n\}$ converges and we denote $B^*=\lim_{n\to\infty}B_n$. On the other hand,

$$\|I-F'(x_n)B_n\| \leqslant a_n \leqslant \theta^{2^n-1}a_0,$$

and therefore, by letting $n\to\infty$ it follows that B^* is the bounded right inverse of $F'(x^*)$.

Acknowledgments

This paper has been partly supported by the Ministry of Science and Technology (Ref. MTM2005-03091).

References

[1] O.H. Hald, On a Newton–Moser type method, Numer. Math. 23 (1975) 411–425.
[2] M.A. Hernández, The Newton method for operators with Hölder continuous first derivative, J. Optim. Theory Appl. 109 (3) (2001) 631–648.
[3] L.V. Kantorovich, G.P. Akilov, Functional Analysis, Pergamon Press, Oxford, 1982.
[4] J. Moser, Stable and random motions in dynamical systems with special emphasis on celestial mechanics. Herman Weil Lectures, Annals of Mathematics Studies, vol. 77, Princeton University Press, Princeton, NJ, 1973.
[5] H. Petzeltova, Remark on Newton–Moser type method, Commentat. Math. Univ. Carol. 21 (1980) 719–725.
[6] F.A. Potra, V. Ptak, Sharp error bounds for Newton process, Numer. Math. 34 (1980) 63–72.

[7] F.A. Potra, V. Ptak, Nondiscrete Induction and Iterative Processes, Pitman Advanced Publishing Program, London, 1984.
[8] S. Smale, Newton's method estimates from data at one point, in: R. Ewing, K. Gross, C. Martin (Eds.), The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics, Springer, New York, 1986, pp. 185–196.
[9] S. Smale, The fundamental theory for solving equations, in: Proceedings of the International Congress of Mathematicians, AMS, Providence, RI, 1987, p. 185.
[10] S. Ulm, On iterative methods with successive approximation of the inverse operator (Russian), Izv. Akad. Nauk Est. SSR 16 (1967) 403–411.
[11] D.R. Wang, F.G. Zhao, The theory of Smale's point estimation and its applications, J. Comput. Appl. Math. 60 (1995) 253–269.
[12] S. Wolfram, The Mathematica Book, fourth ed., Cambridge University Press, Cambridge, 1999 ⟨http://www.wolfram.com/⟩.
[13] E.J. Zehnder, A remark about Newton's method, Commun. Pure Appl. Math. 27 (1974) 361–366.