IMA Journal of Numerical Analysis (1997) **17**, 511–525

A modification of the Chebyshev method

JOSÉ A. EZQUERRO

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

[Received 20 January 1996 and in revised form 11 September 1996]

In this paper we use a one-parametric family of second-order iterations to solve a nonlinear operator equation in a Banach space. Two different analyses of convergence are shown. First, under standard Newton–Kantorovich conditions, we establish a Kantorovich-type convergence theorem. Second, another Kantorovich-type convergence theorem is proved, when the first Fréchet-derivative of the operator satisfies a Lipschitz condition. We also give an explicit expression for the error bound of the family of methods in terms of a real parameter $\alpha \geqslant 0$.

1. Introduction

Let us consider the problem of solving the equation

$$F(x) = 0 \tag{1.1}$$

in Banach spaces by means of iterative processes.

Let X, Y be Banach spaces and $F : \Omega \subseteq X \to Y$ a nonlinear operator on an open convex domain Ω. Let us assume that $F'(x_0)^{-1} \in \mathcal{L}(Y, X)$ exists at some $x_0 \in \Omega$, where $\mathcal{L}(Y, X)$ is the set of bounded linear operators from Y into X.

A well-known cubically convergent iterative procedure to solve (1.1) is the Chebyshev method (Argyros & Chen (1993)):

$$x_{n+1} = x_n - \left[I + \frac{1}{2}F'(x_n)^{-1}F''(x_n)F'(x_n)^{-1}F(x_n)\right]F'(x_n)^{-1}F(x_n), \quad n \geqslant 0,$$

where I is the identity operator on X. Here $F'(x_n)$ and $F''(x_n)$ denote the first and second Fréchet-derivatives of F evaluated at $x = x_n$. Note that $F'(x_n)$ is a linear operator whereas $F''(x_n)$ is a bilinear operator for all $n \geqslant 0$. For discretized versions of Chebyshev's method see Argyros (1995), Ul'm (1964).

One aim of this paper is to reduce operational costs and to ease conditions for the F operator. This is satisfied, for instance, if the second Fréchet-derivative of F is not evaluated at each x_n or this derivative does not exist. These situations are studied in Sections 2 and 3 respectively. There, the second Fréchet-derivative is replaced by a fixed bilinear operator. So we introduce a new iterative process

$$x_{n+1} = x_n - \left[I + \frac{1}{2}F'(x_n)^{-1}AF'(x_n)^{-1}F(x_n)\right]F'(x_n)^{-1}F(x_n), \quad n \geqslant 0, \tag{1.2}$$

to solve (1.1), where $A : X \times X \to Y$ is a general bounded bilinear operator which satisfies $\|A\| = \alpha$ ($\alpha \geqslant 0$). Note that for $A \equiv 0$ we obtain Newton's iteration.

Observe that the speed of convergence of (1.2) is lower than for Chebyshev's method, since the order of convergence drops from three to two. Thus we will try to obtain second-order iterative processes faster than Newton's method.

Using the method of majorizing sequences (see Kantorovich & Akilov (1982), Potra & Pták (1984), Rheinboldt (1968), Yamamoto (1988)), the following two important problems are studied: we give sufficient conditions for the convergence of (1.2) to a solution x^* of (1.1), as well as the uniqueness of x^*, and we find estimates for the distances $\|x_n - x^*\|$, $n \geqslant 0$.

We show that one of the basic assumptions in the study of the convergence of (1.2) is that F must be twice-differentiable in some ball around the initial iterate (Kantorovich & Akilov (1982), Rheinboldt (1968)), or that the linear operator F' must satisfy a Lipschitz condition (Potra & Pták (1984), Yamamoto (1988)). Notice that the latter assumption is milder than the former one.

Let us denote $\overline{B(x,r)} = \{y \in X;\ \|y - x\| \leqslant r\}$ and $B(x,r) = \{y \in X;\ \|y - x\| < r\}$.

2. First analysis: study of convergence when the nonlinear operator is twice-Fréchet differentiable

Let us assume the nonlinear operator F is twice-Fréchet differentiable on Ω. Following Argyros & Chen (1993, 1994), we write (1.2) as:

$$y_n = x_n - F'(x_n)^{-1}F(x_n), \tag{2.1}$$

$$x_{n+1} = y_n - \frac{1}{2}F'(x_n)^{-1}A(y_n - x_n)^2. \tag{2.2}$$

The following conditions are assumed:

(i) There exists a continuous linear operator $\Gamma_0 = F'(x_0)^{-1}$, $x_0 \in \Omega$. Moreover $\|\Gamma_0\| \leqslant \beta$.
(ii) $\|F''(x)\| \leqslant k$ for $x \in \Omega$.
(iii) $\|F''(x) - A\| \leqslant k - \alpha$ for $x \in \Omega$ $(\alpha \leqslant k)$.
(iv) $\|y_0 - x_0\| \leqslant \eta$.
(v) The equation

$$g(t) \equiv \frac{k}{2}t^2 - \frac{t}{\beta} + \frac{\eta}{\beta} = 0 \tag{2.3}$$

has two positive roots t^* and t^{**} ($t^* \leqslant t^{**}$). Equivalently, $k\beta\eta \leqslant \frac{1}{2}$.

Let us consider now the scalar sequences

$$s_n = t_n - \frac{g(t_n)}{g'(t_n)}, \quad t_0 = 0, \quad n \geqslant 0, \tag{2.4}$$

$$t_{n+1} = s_n - \frac{\alpha}{2}\frac{(s_n - t_n)^2}{g'(t_n)}, \quad n \geqslant 0, \tag{2.5}$$

where g is the polynomial defined in (2.3). In the next lemma we show that the sequence $\{t_n\}$ defined by (2.4) and (2.5) is increasing and quadratically convergent to t^* for all $0 \leqslant \alpha \leqslant k$.

LEMMA 2.1 Let g be the polynomial defined in (2.3). Let us write the sequence given by (2.4) and (2.5) as

$$t_0 = 0, \quad t_{n+1} = G_\alpha(t_n) = t_n - \frac{g(t_n)}{g'(t_n)}\left(1 + \frac{\alpha}{2}\frac{g(t_n)}{g'(t_n)^2}\right), \quad n \geqslant 0.$$

Then this sequence is increasing and converges quadratically to t^* for all $0 \leqslant \alpha \leqslant k$.

Proof. Note that

$$G'_\alpha(t) = \frac{g(t)}{g'(t)^2}\left[g''(t) + \alpha\left(\frac{3}{2}L_g(t) - 1\right)\right] \geqslant 0$$

in $[0, t^*]$, where

$$L_g(t) = \frac{g(t)g''(t)}{g'(t)^2}$$

(Hernández (1991)). Then by mathematical induction on n, it follows that $t_n \leqslant t^*$, $n \geqslant 0$.

On the other hand, it is easy to show that $t_n \leqslant t_{n+1}$ for all $n \in \mathbf{N}$ and consequently the proof is completed. □

We can obtain the following Ostrowski–Kantorovich representation for $F(x_{n+1})$.

LEMMA 2.2 Let F be a nonlinear operator mapping an open convex domain Ω in X to Y. Assume that F is twice-Fréchet differentiable on Ω. The following approximation is true for all $n \geqslant 0$:

$$F(x_{n+1}) = \int_{y_n}^{x_{n+1}} F''(x)(x_{n+1} - x)\,dx + \int_{x_n}^{y_n} F''(x)(x_{n+1} - y_n)\,dx$$
$$+ \int_{x_n}^{y_n} (F''(x) - A)(y_n - x)\,dx. \tag{2.6}$$

Proof. To prove the previous statement we observe that

$$F(x_{n+1}) = F(x_{n+1}) - F(y_n) - F'(y_n)(x_{n+1} - y_n) + F(y_n) + F'(y_n)(x_{n+1} - y_n).$$

We also have

$$F(x_{n+1}) - F(y_n) - F'(y_n)(x_{n+1} - y_n) = \int_{y_n}^{x_{n+1}} F''(x)(x_{n+1} - x)\,dx,$$

$$F(y_n) = \int_{x_n}^{y_n} F''(x)(y_n - x)\,dx$$

and

$$F'(y_n)(x_{n+1} - y_n) = \int_{x_n}^{y_n} F''(x)(x_{n+1} - y_n)\,dx - \frac{A}{2}(y_n - x_n)^2.$$

Substituting these in the above, we obtain (2.6). □

Next we see that the sequence $\{t_n\}$ given by (2.4) and (2.5) is the majorizing sequence of $\{x_n\}$ defined by (2.1) and (2.2).

LEMMA 2.3 Let us assume that conditions (i)–(v) are satisfied. Then

$$\|x_{n+1} - x_n\| t_{n+1} - t_n, \quad n \geqslant 0.$$

Proof. By using mathematical induction on n, it is enough to show that the following statements are true for all $n \geqslant 0$:

$[\mathrm{I}_n]$ $\|F'(x_n)^{-1}\| \leqslant \dfrac{-1}{g'(t_n)}$,
$[\mathrm{II}_n]$ $\|y_n - x_n\| \leqslant s_n - t_n$,
$[\mathrm{III}_n]$ $\|x_{n+1} - y_n\| \leqslant t_{n+1} - s_n$,
$[\mathrm{IV}_n]$ $\|F(x_{n+1})\| \leqslant g(t_{n+1})$.:

All the above statements are true for $n = 0$ with initial conditions (i)–(v). Then we assume that they are true for a fixed n and all smaller integer values. Observe that

$$I - \Gamma_0 F'(x_{n+1}) = \int_0^1 \Gamma_0 F''(x_0 + t(x_{n+1} - x_0))(x_{n+1} - x_0)\,\mathrm{d}t.$$

So

$$\|I - \Gamma_0 F'(x_{n+1})\| \leqslant \beta k \|x_{n+1} - x_0\| \leqslant \beta k t^* < 1,$$

and by the Banach lemma, $F'(x_{n+1})^{-1}$ exists and

$$\|F'(x_{n+1})^{-1}\| \leqslant \frac{\|\Gamma_0\|}{1 - \|I - \Gamma_0 F'(x_{n+1})\|} \leqslant \frac{\beta}{1 - \beta k \|x_{n+1} - x_0\|} \leqslant \frac{-1}{g'(t_{n+1})}.$$

Hence $[\mathrm{I}_{n+1}]$ is true. $[\mathrm{II}_{n+1}]$ and $[\mathrm{III}_{n+1}]$ follow easily.

Using (2.6) and taking norms we obtain

$$\|F(x_{n+1})\| = \frac{k}{2}\|x_{n+1} - y_n\|^2 + k\|x_{n+1} - y_n\|\,\|y_n - x_n\| + \frac{k-\alpha}{2}\|y_n - x_n\|^2$$

$$\leqslant \frac{k}{2}(t_{n+1} - s_n)^2 + k(t_{n+1} - s_n)(s_n - t_n) + \frac{k-\alpha}{2}(s_n - t_n)^2 = g(t_{n+1}).$$

Thus $[\mathrm{IV}_n]$ is also true.

Finally, it follows inmediately that $\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n$ for all $n \geqslant 0$. □

THEOREM 2.4 Let $x_0 \in \Omega$ be an initial value. Let us assume that conditions (i)–(v) are satisfied and $\overline{B(y_0, t^* - \eta)} \subset \Omega$. Then the procedure defined by (2.1) and (2.2) is well defined for all $n \geqslant 0$ and is convergent, and $x_n, y_n \in \overline{B(x_0, t^*)}$ for all $n \geqslant 0$. The limit x^* is the unique solution of (1.1) in $B(x_0, t^{**})$. We also have the following error bound estimates for all $n \geqslant 0$:

$$\|x^* - x_n\| \leqslant t^* - t_n \quad \text{and} \quad \|x^* - y_n\| \leqslant t^* - s_n.$$

Proof. The fact that the convergence of the sequence $\{t_n\}$ defined by (2.4) and (2.5) implies the convergence of the sequence $\{x_n\}$ given by (2.1) and (2.2) is a consequence of $\{t_n\}$ majorizing $\{x_n\}$ (see Lemma 2.3). Furthermore, making $n \to \infty$ in statement $[\mathrm{IV}_n]$ of Lemma 2.3, we deduce that $F(x^*) = 0$. We also have

$$\begin{aligned}\|x_n - y_0\| &\leqslant \|x_n - y_{n-1}\| + \|y_{n-1} - x_{n-1}\| + \cdots + \|x_1 - y_0\| \\ &\leqslant (t_n - s_{n-1}) + (s_{n-1} - t_{n-1}) + \cdots + (t_1 - s_0) \\ &= t_n - \eta \leqslant t^* - \eta\end{aligned}$$

and similarly

$$\|y_n - y_0\| \leqslant s_n - \eta \leqslant t^* - \eta.$$

For $p \geqslant 0$,

$$\|x_{n+p} - x_n\| \leqslant t_{n+p} - t_n, \quad \|x_{n+p} - y_n\| \leqslant t_{n+p} - s_n$$

and letting $p \to \infty$ we obtain

$$\|x^* - x_n\| \leqslant t^* - t_n \quad \text{and} \quad \|x^* - y_n\| \leqslant t^* - s_n, \quad n \geqslant 0.$$

Now to demonstrate uniqueness, let us assume that there exists another solution z^* of (1.1) in $B(x_0, t^{**})$. Taking into account that

$$\begin{aligned}&\|\Gamma_0\| \int_0^1 \|F'(x^* + t(z^* - x^*)) - F'(x_0)\| \, dt \\ &\leqslant \beta k \int_0^1 \|x^* + t(z^* - x^*) - x_0\| \, dt \\ &\leqslant \beta k \int_0^1 \big((1 - t)\|x^* - x_0\| + t\|z^* - x_0\|\big) \, dt \\ &< \frac{\beta k}{2}(t^* + t^{**}) = 1,\end{aligned}$$

we infer that the linear operator $\int_0^1 F'(x^* + t(z^* - x^*))\, dt$ is invertible. From the approximation

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

it follows that $x^* = z^*$. This completes the proof of the theorem. □

Now we will obtain error expressions for the sequence $\{t_n\}$ defined by (2.4) and (2.5). Following Ostrowski (1943), we derive the following error bounds.

THEOREM 2.5 Let g be the polynomial given by (2.3) and assume that g has two positive roots t^* and t^{**} ($t^* \leqslant t^{**}$). Let $\{t_n\}$ be the sequence given by (2.4) and (2.5).

(a) When $t^* < t^{**}$, let $\theta = \dfrac{t^*}{t^{**}}$ and $\Delta_\alpha = \theta\sqrt{1-\frac{\alpha}{k}}$. Then we have

$$\frac{(t^{**}-t^*)\Delta_\alpha^{2^n}}{\sqrt{1-\frac{\alpha}{k}}-\Delta_\alpha^{2^n}} < t^* - t_n < \frac{(t^{**}-t^*)\theta^{2^n}}{1-\theta^{2^n}}, \quad n \geqslant 0,$$

where $\Delta_\alpha < 1$.

(b) When $t^* = t^{**}$, we have

$$t^* - t_n = t^*\left(\frac{1}{2}-\frac{\alpha}{8k}\right)^n, \quad n \geqslant 0.$$

Proof. First we set $a_n = t^* - t_n$ and $b_n = t^{**} - t_n$. Moreover we notice that

$$g(t_n) = \frac{k}{2}a_n b_n \quad \text{and} \quad g'(t_n) = -\frac{k}{2}(a_n + b_n).$$

Now by (2.5) we have

$$a_n = a_{n-1}^2 \frac{k(a_{n-1}+b_{n-1})^2 - \alpha b_{n-1}^2}{k(a_{n-1}+b_{n-1})^3} \tag{2.7}$$

and

$$b_n = b_{n-1}^2 \frac{k(a_{n-1}+b_{n-1})^2 - \alpha a_{n-1}^2}{k(a_{n-1}+b_{n-1})^3}.$$

If $t^* < t^{**}$, denote the ratio of a_n and b_n by δ_n. So

$$\delta_n = \delta_{n-1}^2 \frac{k(1+\delta_{n-1})^2 - \alpha}{k(1+\delta_{n-1})^2 - \alpha\delta_{n-1}^2} = \delta_{n-1}^2 H(\delta_{n-1}).$$

Taking into account that the function

$$H(u) = \frac{k(1+u)^2 - \alpha}{k(1+u)^2 - \alpha u^2}$$

is nondecreasing for all $\alpha \leqslant k$, we obtain

$$\delta_n < \delta_{n-1}^2 < \cdots < \delta_0^{2^n}$$

and

$$\delta_n > \left(1-\frac{\alpha}{k}\right)\delta_{n-1}^2 > \cdots > \left(1-\frac{\alpha}{k}\right)^{\frac{2^n-1}{2}} \delta_0^{2^n}.$$

Then the first part holds.

If $t^* = t^{**}$, then $a_n = b_n$ and by (2.7) we get

$$a_n = a_{n-1}\left(\frac{1}{2}-\frac{\alpha}{8k}\right).$$

By recurrence, the second part also holds. □

To illustrate Theorem 2.4, we provide the following example.

EXAMPLE 2.6 Let us consider the system of equations $F(x, y) = 0$ where

$$F(x, y) = \left(x^2 - 2y + 1/3,\ y^2 - 4x + 2/3\right).$$

Then we have

$$F'(x, y)^{-1} = \frac{1}{2(xy - 2)} \begin{pmatrix} y & 1 \\ 2 & x \end{pmatrix}$$

if (x, y) does not belong to the hyperbola $xy = 2$. The second derivative is a bilinear operator on $\mathbb{R}^2$ given by

$$F''(x, y) = \left(\begin{array}{cc} 2 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ 0 & 2 \end{array}\right).$$

We take the max-norm in $\mathbb{R}^2$ and the norm

$$\|A\| = \max\{|a_{11}| + |a_{12}|, |a_{21}| + |a_{22}|\}$$

for

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

As in Rall (1961) we define the norm of a bilinear operator B on $\mathbb{R}^2$ by

$$\|B\| = \sup_{\|x\|=1} \max_i \sum_{j=1}^{2} \left| \sum_{k=1}^{2} b_i^{jk} x_k \right|$$

where $x = (x_1, x_2)$ and

$$B = \left(\begin{array}{cc} b_1^{11} & b_1^{12} \\ b_1^{21} & b_1^{22} \\ \hline b_2^{11} & b_2^{12} \\ b_2^{21} & b_2^{22} \end{array}\right).$$

If we choose $\mathbf{x}_0 = (\frac{1}{2}, \frac{1}{2})$ and $\Omega = B(\mathbf{x}_0, 0{\cdot}9)$, then

$$k = \|F''(x, y)\| = 2, \quad \beta = \|\Gamma_0\| = \frac{2}{7} \quad \text{and} \quad \eta = \|y_0 - x_0\| = \frac{33}{84}.$$

Now for instance if we consider the bilinear operator

$$A = \left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ 0 & 1 \end{array}\right), \tag{2.8}$$

TABLE 1
Iterative process (1.2)

n	x_n	y_n
0	0·5000000000000000	0·5000000000000000
1	0·1627287819889861	0·1570800939423388
2	0·1748627729486347	0·1819178649869130
3	0·1749448931399644	0·1819695227985917
4	0·1749448936348263	0·1819695245711170

then $\|A\| = 1 = \alpha$, and hypothesis (iii) is satisfied. Thus the polynomial (2.3) becomes

$$g(t) = t^2 - \frac{7}{2}t + \frac{231}{268}.$$

This polynomial has two positive roots $t^* = 0{\cdot}450962$ and $t^{**} = 3{\cdot}04904$. Therefore the process given by (1.2) where A is defined by (2.8) converges to

$$(x^*, y^*) = (0{\cdot}1749448936348263, 0{\cdot}181969524571117),$$

see Table 1. Moreover this solution is unique in $B(x_0, 3{\cdot}04904)$ and the error bound expressions are for all $n \geqslant 0$:

$$\frac{2{\cdot}598078\,(0{\cdot}1045831)^{2^n}}{0{\cdot}7071067 - (0{\cdot}1045831)^{2^n}} < 0{\cdot}450962 - t_n < \frac{2{\cdot}598078\,(0{\cdot}1479029)^{2^n}}{1 - (0{\cdot}1479029)^{2^n}}.$$

Observe that the process given by (1.2) converges to (x^*, y^*) faster than the Newton method (see Tables 1 and 2).

3. Second analysis: study of convergence when the first Fréchet-derivative satisfies a Lipschitz condition

Let us assume that F is a nonlinear once-Fréchet differentiable operator in an open convex domain Ω. We assume throughout this section that

(c_1) There exists a continuous linear operator $\Gamma_0 = F'(x_0)^{-1}$, $\quad x_0 \in \Omega$.
(c_2) $\|\Gamma_0(F'(x) - F'(y))\| \leqslant k\|x - y\|$, $\quad x, y \in \Omega$, $\quad k \geqslant 0$.
(c_3) $\|A\| = \alpha$, $\quad \|\Gamma_0 A\| \leqslant \alpha/b$, $\quad \|\Gamma_0 F(x_0)\| = a/b$.
(c_4) $b - 2ak \geqslant 0$.

Note that conditions (c_1)–(c_4) are milder than conditions (i)–(iv) of Section 2.

Before establishing results on existence and uniqueness of the solution of equation (1.1), we need the following two lemmas, whose proofs are trivial.

LEMMA 3.1 Let α be a fixed real number which satisfies $0 \leqslant \alpha \leqslant \dfrac{b}{4a}(b - 2ak)$. Then

(i) $\left[b + \dfrac{2\alpha}{k}, \dfrac{b^2}{2ak}\right] \neq \emptyset$.

TABLE 2
Newton's method

n	x_n	y_n
0	0·5000000000000000	0·5000000000000000
1	0·1309523809523810	0·1071428571428571
2	0·1734830412576077	0·1808104209356618
3	0·1749444542650147	0·1819683798417195
4	0·1749448936344844	0·1819695245709607

(ii) If $N \leqslant \dfrac{b^2}{2ak}$, the equation

$$p(t) \equiv \frac{kN}{2}t^2 - bt + a = 0 \tag{3.1}$$

has two positive roots r_1 and r_2 ($r_1 \leqslant r_2$). Besides $N = \dfrac{b^2}{2ak}$ iff $r_1 = r_2$.

Observe that a modification in the usual 'test' function p (see Argyros (1992, 1993b), Argyros & Chen (1993, 1994), Kantorovich & Akilov (1982), Yamamoto (1988)) has been introduced: we have inserted a parameter N in the p polynomial. Then convergence of the family (1.2) is proved under conditions (c_1)–(c_4) and the hypothesis of Lemma 3.1.

LEMMA 3.2 Let p be the polynomial defined in (3.1). Then the sequence

$$t_0 = 0, \quad t_{n+1} = t_n - \frac{p(t_n)}{p'(t_n)}\left(1 + \frac{\alpha}{2}\frac{p(t_n)}{p'(t_n)^2}\right), \quad n \geqslant 0, \tag{3.2}$$

is increasing and converges quadratically to r_1 for all $0 \leqslant \alpha \leqslant \dfrac{b}{4a}(b - 2ak)$.

Now we give a Kantorovich-type convergence theorem.

THEOREM 3.3 Let us assume that conditions (c_1)–(c_4) hold and

$$0 \leqslant \alpha \leqslant \frac{b}{4a}(b - 2ak).$$

Then the sequence $\{x_n\}$ defined by (1.2) converges to a solution x^* of (1.1) in $\overline{B(x_0, r_1)} \cap \Omega$ for

$$N \in \left[b + \frac{2\alpha}{k}, \frac{b^2}{2ak}\right].$$

The limit x^* is the unique solution of (1.1) in $B(x_0, r) \cap \Omega$ where

$$r = r_2 + \frac{2(N-b)}{kN}.$$

Moreover $\|x^* - x_n\| \leqslant r_1 - t_n$, $n \geqslant 0$.

To prove the above theorem we give the next result.

LEMMA 3.4 Under the assumptions of Theorem 3.3, the iterations (1.1) are well defined and converge to a solution x^* of (1.1). More precisely, we have:

$$\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n, \quad n \geqslant 0, \tag{3.3}$$

$$\|x^* - x_n\| \leqslant r_1 - t_n, \quad n \geqslant 0. \tag{3.4}$$

Proof. We show for $n \geqslant 0$:

$[\mathrm{I}_n]$ There exists $\Gamma_n = F'(x_n)^{-1}$,

$[\mathrm{II}_n]$ $\|\Gamma_n A\| \leqslant -\dfrac{\alpha}{p'(t_n)}$,

$[\mathrm{III}_n]$ $\|\Gamma_n F'(x_0)\| \leqslant \dfrac{p'(t_0)}{p'(t_n)}$,

$[\mathrm{IV}_n]$ $\|\Gamma_0 F(x_n)\| \leqslant -\dfrac{p(t_n)}{p'(t_0)}$.

Notice that $[\mathrm{I}_0]$–$[\mathrm{IV}_0]$ follow immediately from (c_1)–(c_4). We prove $[\mathrm{I}_{n+1}]$–$[\mathrm{IV}_{n+1}]$ by using mathematical induction. Following Altman (1961) and Yamamoto (1988), under our assumptions (c_1)–(c_4), $\Gamma_{n+1} = F'(x_{n+1})^{-1}$ exists and so $[\mathrm{II}_{n+1}]$ and $[\mathrm{III}_{n+1}]$ are true. To prove $[\mathrm{IV}_{n+1}]$, we infer by Taylor's formula and taking into account (1.2) that

$$\begin{aligned} F(x_{n+1}) &= F(x_n) + F'(x_n)(x_{n+1} - x_n) + \int_{x_n}^{x_{n+1}} (F'(x) - F'(x_n))\,\mathrm{d}x \\ &= -\frac{A}{2}(\Gamma_n F(x_n))^2 + \int_{x_n}^{x_{n+1}} (F'(x) - F'(x_n))\,\mathrm{d}x. \end{aligned}$$

Thus

$$\|\Gamma_0 F(x_{n+1})\| \leqslant \frac{\alpha}{2b}\left(\frac{p(t_n)}{p'(t_n)}\right)^2 + \frac{k}{2}(t_{n+1} - t_n)^2.$$

Repeating the same process for the p polynomial given by (3.1), we obtain

$$p(t_{n+1}) = -\frac{\alpha}{2}\left(\frac{p(t_n)}{p'(t_n)}\right)^2 + \frac{kN}{2}(t_{n+1} - t_n)^2.$$

Now to prove

$$\|\Gamma_0 F(x_{n+1})\| \leqslant -\frac{p(t_{n+1})}{p'(t_0)}, \tag{3.5}$$

it suffices to see that

$$\|\Gamma_0 F(x_{n+1})\| + \frac{p(t_{n+1})}{p'(t_0)} \leqslant \left(\frac{p(t_n)}{p'(t_n)}\right)^2 \left(\frac{\alpha}{b} + \frac{k}{2}\left(1 - \frac{N}{b}\right)\right) \leqslant 0,$$

since $N \geqslant b + \dfrac{2\alpha}{k}$. Then the induction is complete.

Furthermore

$$\|x_{n+1} - x_n\| = \left\|\left(I + \tfrac{1}{2}\Gamma_n A \Gamma_n F(x_n)\right) \Gamma_n F(x_n)\right\|$$

$$\leqslant -\left(1 + \frac{\alpha}{2} \frac{p(t_n)}{p'(t_n)^2}\right) \frac{p(t_n)}{p'(t_n)} = t_{n+1} - t_n.$$

Hence (3.3) holds. The convergence of $\{t_n\}$ implies the convergence of $\{x_n\}$ to a limit x^* (see Kantorovich & Akilov (1982)). Making $n \to \infty$ in (3.5), we infer that $F(x^*) = 0$.

Finally, for $q \geqslant 0$,

$$\|x_{n+q} - x_n\| \leqslant t_{n+q} - t_n,$$

and letting $q \to \infty$ we get (3.4). □

Proof of Theorem 3.3. The convergence of the sequence $\{x_n\}$ defined by (1.2) to a limit x^* is an immediate consequence of Lemmas 3.1, 3.2 and 3.4. To show the uniqueness of the solution x^*, let us assume that y^* is another solution of (1.1) in $B(x_0, r)$ where

$$r = r_2 + \frac{2(N-b)}{kN}.$$

Using the approximation

$$F'(x^* + t(y^* - x^*)) - F'(x_0) = \int_{x_0}^{x^*+t(y^*-x^*)} F''(z)\,\mathrm{d}z$$

and the estimate

$$\int_0^1 \|\Gamma_0\|\ \|F'(x_0) - F'(x^* + t(y^* - x^*))\|\,\mathrm{d}t$$

$$\leqslant k \int_0^1 \left(\|x_0 - x^*\|(1-t) + \|x_0 - y^*\| t\right)\,\mathrm{d}t$$

$$< k\left(\frac{r_1 + r}{2}\right) = 1,$$

we deduce that the inverse of $\int_0^1 F'(x^* + t(y^* - x^*))\,\mathrm{d}t$ exists. From the approximation

$$F(y^*) - F(x^*) = \int_0^1 F'(x^* + t(y^* - x^*))(y^* - x^*)\,\mathrm{d}t = 0,$$

we conclude that $x^* = y^*$. □

REMARK 3.5 The error estimates for the sequence $\{t_n\}$ defined by (3.2) are similar to the ones for the sequence defined by (2.4) and (2.5). In an analogous way, we get for $d = \dfrac{\alpha}{kN}$:

(a) If $r_1 < r_2$, let $\vartheta = \dfrac{r_1}{r_2}$ and $\Lambda_\alpha = \vartheta\sqrt{1-d}$. Then

$$\frac{(r_2 - r_1)\Lambda_\alpha^{2^n}}{\sqrt{1-d} - \Lambda_\alpha^{2^n}} < r_1 - t_n < \frac{(r_2 - r_1)\vartheta^{2^n}}{1 - \vartheta^{2^n}}, \qquad n \geqslant 0,$$

where $\Lambda_\alpha < 1$.

(b) If $t^* = t^{**}$, then

$$r_1 - t_n = r_1 \left(\frac{1}{2} - \frac{d}{8}\right)^n, \qquad n \geqslant 0.$$

REMARK 3.6 Notice that in practice we can always consider $N = b + 2\alpha/k$, since for this value we get the smallest error bound for (3.2). In fact, we deduce exactly as in Theorem 2.4 that

$$\mu_n = \mu_{n-1}^2 \frac{kN(1+\mu_{n-1})^2 - \alpha}{kN(1+\mu_{n-1})^2 - \alpha\mu_{n-1}^2}$$

where $\mu_n = \dfrac{r_1 - t_n}{r_2 - t_n}$. It is easy to check that the function

$$j(N) = \frac{kN(1+\mu_{n-1})^2 - \alpha}{kN(1+\mu_{n-1})^2 - \alpha\mu_{n-1}^2}$$

is nondecreasing and consequently for $N = b + 2\alpha/k$ we obtain the smallest error bound.

Now we apply our results of Theorem 3.3 by taking into account an example considered in part by Argyros (1988a, b, 1992, 1993a, b). Determining existence and uniqueness domains of solutions for a differential equation is the goal of the example.

EXAMPLE 3.7 We consider the following differential equation

$$\begin{cases} y'' + y^2 = 0, \\ y(0) = 0 = y(1). \end{cases} \tag{3.6}$$

Divide the interval $[0, 1]$ into n subintervals and set $h = 1/n$. We denote the points of subdivision by $0 = w_0 < w_1 < \cdots < w_n = 1$ with the corresponding values of the function $y_0 = y(w_0), y_1 = y(w_1), \ldots, y_n = y(w_n)$. A standard approximation for the second derivative at these points is

$$y_i'' = \frac{y_{i-1} - 2y_i + y_{i+1}}{h^2}, \qquad i = 1, 2, \ldots, n-1.$$

Noting that $y_0 = 0 = y_n$, define the operator $F : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ by

$$F(y) = Gy + h^2 g(y)$$

where

$$G=\begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2 \end{pmatrix}, \quad g(y)=\begin{pmatrix} y_1^2 \\ y_2^2 \\ \vdots \\ y_{n-1}^2 \end{pmatrix} \quad \text{and} \quad y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{pmatrix}.$$

Then

$$F'(y)=G-2h^2\begin{pmatrix} y_1 & 0 & \dots & 0 \\ 0 & y_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & y_{n-1} \end{pmatrix}.$$

Let $y \in \mathbf{R}^{n-1}$, then the norm will be $\|y\| = \max_{1\leqslant i\leqslant n-1} |y_i|$. The corresponding norm on $G \in \mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$ is

$$\|G\| = \max_{1\leqslant i\leqslant n-1} \sum_{k=1}^{n-1} |g_{ik}| = 4.$$

We note that for all $x, y \in \mathbb{R}^{n-1}$:

$$\|F'(x) - F'(y)\| = \|\mathrm{diag}\{2h^2(x_i - y_i)\}\| = 2h^2 \max_{1\leqslant i\leqslant n-1} |x_i - y_i| \leqslant 2h^2\|x - y\|.$$

To show the convergence of (1.2) to a solution y^* of (3.6), set $n = 10$ and, since the solution would vanish at the endpoints and be positive in the interior, a reasonable choice of initial approximation would seem to be $\frac{1}{4}\sin \pi x$. This gives us the following vector:

$$x_0=\begin{pmatrix} 0{\cdot}0772542 \\ 0{\cdot}1469460 \\ 0{\cdot}2022540 \\ 0{\cdot}2377640 \\ 0{\cdot}2500000 \\ 0{\cdot}2377640 \\ 0{\cdot}2022540 \\ 0{\cdot}1469460 \\ 0{\cdot}0772542 \end{pmatrix}.$$

Using the notation of Theorem 3.3 we can easily obtain the following results:

$$\frac{1}{b} = \|\Gamma_0\| = 13{\cdot}0918 \quad \text{and} \quad \frac{a}{b} = \|\Gamma_0 F(x_0)\| = 0{\cdot}255796.$$

Then $a = 0{\cdot}0195386$, $b = 0{\cdot}0763837$ and $k = 0{\cdot}261836$. For these values we have $\alpha \leqslant 0{\cdot}0646531$. Therefore we choose a constant bilinear operator A where $\|A\| = 0{\cdot}06 = \alpha$. So $N = 0{\cdot}169162$ and equation (3.1) becomes

$$0{\cdot}07t^2 - 0{\cdot}0763837t + 0{\cdot}0195386 = 0,$$

TABLE 3

n	Interval of valid α	α	r_1	r
10	[0, 0·064653]	0·06	0·409384	7·22899
12	[0, 0·044908]	0·04	0·391193	7·24823
15	[0, 0·029082]	0·02	0·334575	7·34042

whose solutions are $r_1 = 0{\cdot}409384$ and $r_2 = 0{\cdot}681812$. Hence by Theorem 3.3, iteration (1.2) converges to a solution $y^* = (y_1^*, y_2^*, \ldots, y_9^*)$ of the equation $F(y) = 0$ in $\overline{B(x_0, 0{\cdot}409384)}$. Moreover the solution y^* is unique in $B(x_0, r)$, where $r = 7{\cdot}22899$. Furthermore, we have for all $n \geqslant 0$:

$$\frac{0{\cdot}272427\,(0{\cdot}453887)^{2^n}}{0{\cdot}755929 - (0{\cdot}453887)^{2^n}} < 0{\cdot}409384 - t_n < \frac{0{\cdot}272427\,(0{\cdot}600436)^{2^n}}{1 - (0{\cdot}600436)^{2^n}}.$$

Finally note that, to solve the differential equation (3.6), the following interpolation problem is considered

$$\begin{pmatrix} 0 & \frac{1}{10} & \frac{2}{10} & \cdots & \frac{9}{10} & 1 \\ 0 & y_1^* & y_2^* & \cdots & y_9^* & 0 \end{pmatrix}$$

and its solution is an approximation to the solution of (3.6).

Notice that the larger n is, the smaller the interval for α is (see Table 3). For different values of n we get the results given in Table 3.

From Table 3, we see that for larger values of n we obtain better existence and uniqueness domains of solutions of equation (3.6).

4. Conclusions

The Chebyshev method is one of the best known iterative processes to solve a nonlinear equation $F(x) = 0$. The goal of this paper was to reduce operational costs and at the same time to relax the conditions on the second Fréchet-derivative by replacing it with a fixed bilinear operator. Moreover, results for the case when the derivative does not exist are provided. The 'penalty' for this approach is that the order of convergence drops from three to two. However, for the same computational cost, it remains faster than the Newton method. Sufficient conditions and a complete error analysis for the iterative method (1.2) are also provided.

On the other hand, two different analysis techniques are considered to study convergence of the iterative procedure (1.2). First, the F operator must be twice-differentiable in some ball around the initial iteration, then conditions on the second Fréchet-derivative of F are given. Second, the goal is to relax the conditions on F, for instance, the linear operator F' satisfies a Lipschitz condition, instead of the previous one for F''.

Unfortunately, a difficulty appears with the second technique: the decomposition obtained from Taylor's expansion is not appropriate. To solve it, we establish a new technique which consists of inserting a parameter in the 'test' function so that a suitable decomposition is obtained. Then, a scalar sequence majorizing the sequence $\{x_n\}$ given by (1.2) is obtained.

Acknowledgement

This work has been supported in part by a grant from the University of La Rioja, No 95TDA05JEF.

REFERENCES

ALTMAN, M. 1961 Concerning the method of tangent hyperbolas for operator equations. *Bull. Acad. Pol. Sci., Serie Sci. Math., Ast. et Phys.* **9**, 633–7.

ARGYROS, I. K. 1988a Newton-like methods under mild differentiability conditions with error analysis. *Bull. Aust. Math. Soc.* **37**, 131–47.

ARGYROS, I. K. 1988b On Newton's method and nondiscrete mathematical induction. *Bull. Aust. Math. Soc.* **38**, 131–40.

ARGYROS, I. K. 1992 Remarks on the convergence of Newton's method under Hölder continuity conditions. *Tamkang J. Math.* **23**, 269–77.

ARGYROS, I. K. 1993a Some methods for finding error bounds for Newton-like methods under mild differentiability conditions. *Acta Math. Hung.* **61**, 183–94.

ARGYROS, I. K. 1993b On the secant method. *Publ. Math. Debrecen* **43**, 223–38.

ARGYROS, I. K. 1995 A unified approach for constructing fast two-step Newton-like methods. *Monatshefte Math.* **119**, 1–22.

ARGYROS, I. K., & CHEN, D. 1993 Results on the Chebyshev method in Banach spaces. *Proyecciones* **12**, 119–28.

ARGYROS, I. K., & CHEN, D. 1994 Parameter-based algorithms for approximating the local solution of nonlinear complex equations. *Proyecciones* **33**, 53–61.

HERNÁNDEZ, M. A. 1991 A note on Halley's method. *Numer. Math.* **59**, 273–6.

KANTOROVICH, L. V., & AKILOV, G. P. 1982 *Functional Analysis*. Oxford: Pergamon.

OSTROWSKI, A. M. 1943 *Solution of Equations in Euclidean and Banach Spaces*. London: Academic.

POTRA, F. A., & PTÁK, V. 1984 *Nondiscrete Induction and Iterative Processes*. London: Pitman.

RALL, L. B. 1961 Quadratic equations in Banach spaces. *Rend. Circ. Mat., Palermo (2)* **10**, 314–32.

RHEINBOLDT, W. C. 1968 A unified convergence theory for a class of iterative processes. *SIAM J. Numer. Anal* **5**, 42–63.

UL'M, S. 1964 Iteration methods with divided differences of the second order. *Dokl. Akad. Nauk. SSSR* **158**, 55–8 (in Russian). (Engl. Transl. 1964 *Sov. Math. Dokl.* **5**, 1187–90.)

YAMAMOTO, T. 1988 On the method of tangent hyperbolas in Banach spaces. *J. Comput. Appl. Math.* **21**, 75–86.