JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 100, No. 2, pp. 311-326, FEBRUARY 1999

On a Convex Acceleration of Newton's Method[1]

J. A. EZQUERRO[2] AND M. A. HERNÁNDEZ[2]

Communicated by F. A. Potra

Abstract. In this study, we use a convex acceleration of Newton's method (or super-Halley method) to approximate solutions of nonlinear equations. We provide sufficient convergence conditions for this method in three space settings: real line, complex plane, and Banach space. Several applications of our results are also provided.

Key Words. Nonlinear equations, convex acceleration of Newton's method, Newton–Kantorovich assumptions, majorizing sequences.

1. Introduction

We consider the well-known convex acceleration of Newton's method (or super-Halley method), which is cubically convergent (see Refs. 1 and 2), to solve the nonlinear scalar equation

$$f(t)=0. \tag{1}$$

Using the degree of logarithmic convexity of the function f (Ref. 3),

$$L_f(t)=f(t)f''(t)/f'(t)^2$$

the convex acceleration of Newton's method is defined for some t_0 and $n\geq 0$ by

$$t_{n+1}=G(t_n)=t_n-[f(t_n)/f'(t_n)][1+L_f(t_n)/2(1-L_f(t_n))]. \tag{2}$$

[1]The research reported herein was sponsored in part by University of La Rioja Grant 96PYA17JEF. The authors are indebted to the referees for their careful reading of this paper. Their comments uncovered several weaknesses in the presentation and helped us to clarify it.
[2]Professor, Department of Mathematics and Computation, University of La Rioja, Logroño, Spain.

The convergence of this method is analyzed by means of the convexity of the functions f and f'. We prove that we can always consider it to solve any scalar equation. Next, the convergence analysis is established in the complex plane under standard Newton–Kantorovich assumptions (see Ref. 4).

We also study this method in Banach spaces. We relax the usual Lipschitz continuity of the second derivative in F'', and we need only the Hölder (k, p)-continuity, with $k \geq 0$, $p \geq 0$ (Refs. 5 and 6), in some ball around the initial iterate. The results presented here extend earlier ones (Refs. 1 and 7). In particular, the local convergence theorem appearing in Ref. 1 is shown for quadratic operators. Finally, we provide an example of two-point boundary-value problem to which our results are applied.

On the other hand, the method of tangent parabolas (Euler–Chebyshev) and the method of tangent hyperbolas (Halley) are probably the best known cubically convergent methods for solving nonlinear equations. With the exception of some special cases, these methods have little practical value because they require an evaluation of the second Fréchet derivative, which requires a number of function evaluations proportional to the cube of the dimension of the space. Discretized versions of these methods were considered by Ul'm, who used divided differences of order two, which are unattractive from a numerical point of view (Ref. 8), and by Argyros and Potra, who used divided differences of order one (Refs. 9–11).

2. Convex Acceleration Method for Scalar Nonlinear Equations

Let f: $[a, b] \subseteq \mathbb{R} \to \mathbb{R}$ be a function sufficiently differentiable and satisfying $f'(t) > 0$, $f''(t) \geq 0$ in $[a, b]$ and $f(a) < 0 < f(b)$. Therefore, there is a unique root s of Eq. (1) in $[a, b]$.

We give first a global convergence result depending on the degree of logarithmic convexity of f' in $[a, b]$.

Theorem 2.1. Assuming that $t_0 \in [a, b]$ and $L_f(t) < 1$, $L_{f'}(t) \leq 0$ in $[a, b]$, the following results hold:

(i) If $f(t_0) > 0$, the sequence $\{t_n\}$ defined by (2) is decreasing and converges to s.

(ii) If $f(t_0) < 0$, the sequence $\{t_n\}$ defined by (2) is increasing and converges to s.

Proof. If $f(t_0) > 0$, we deduce that $t_0 - s > 0$. By using the mean-value theorem, we have

$$t_1 - s = G'(z_0)(t_0 - s),$$

for some $z_0 \in (s, t_0)$ and G defined in (2). Taking into account the assumptions mentioned above and

$$G'(t) = [L_f(t)^2/2(1 - L_f(t))^2](L_f(t) - L_{f'}(t)), \tag{3}$$

we deduce that $G'(t) \geq 0$ for $t \in [s, b]$ and, consequently, $s \leq t_1$. Following an inductive process, it is easy to check that $s \leq t_n$ for $n \geq 0$.

On the other hand, we have

$$t_{n+1} - t_n = -[f(t_n)/f'(t_n)]H(L_f(t_n)),$$

where

$$H(t') = 1 + (t'/2)/(1 - t').$$

For $n \geq 0$, we derive from the hypothesis that $H(L_f(t_n)) > 0$, and, consequently, the sequence (2) is decreasing. Then, $\{t_n\}$ converges to $y \in [a, b]$. By letting $n \to \infty$ in (2) and taking into account that $H(L_f(t_n)) > 0$ for $n \geq 0$, we get $f(y) = 0$. As there is only one root of Eq. (1) in $[a, b]$, part (i) holds.

Following analogous arguments, part (ii) also holds. □

Next, we extend the previous result. The following results reduce to those obtained when $L_{f'}(t) > 0$ in $[a, b]$. We denote

$$M(h) = \max\{h(t)/t \in [a, b]\}, m(h) = \min\{h(t)/t \in [a, b]\}.$$

We need first a lemma.

Lemma 2.1. Suppose that $|L_f(t)| \leq 1/k$, with $k > 1.754877$ for $t \in [a, b]$. Then, the following conditions are satisfied:

(i) $[1/k, 2(k-1)^2 - 1/k) \neq \emptyset$,
(ii) $L_f(t)^2/2(1 - L_f(t))^2 \leq 1/2(k-1)^2$.

We can state a second result.

Theorem 2.2. Let $k > 1.754877$, let the interval $[a, b]$ satisfy $a + [(2k-1)/2(k-1)]f(b)/f'(b) \leq b$, and let $t_0 \in [a, b]$, with $f(t_0) > 0$ and $t_0 \geq a + [(2k-1)/2(k-1)]f(b)/f'(b)$. If $|L_f(t)| \leq 1/k$ and $L_{f'}(t) \in [1/k, 2(k-1)^2 - 1/k)$ in $[a, b]$, then the sequence defined by (2) converges to s and satisfies $t_{2n} \geq s$, $t_{2n+1} \leq s$ for all $n \geq 0$.

Proof. It is easy to show that $t_1 \in [a, b]$. From the assumptions on L_f and $L_{f'}$ in $[a, b]$ and (3), we derive that $|G'(t)| \leq Q < 1$ for $t \in [a, b]$ and G defined in (2). Consequently, as $f(t_0) > 0$, we have $t_1 \leq s$ and $|t_1 - s| < Q|t_0 - s|$. Now, it suffices to show that $t_{2j} \geq s$ and $t_{2j+1} \leq s$ are true for all $j \geq 0$ and $|t_n - s| < Q^n|t_0 - s|$ by mathematical induction. Therefore, the sequence (2) lie in $[a, b]$ and converges to s. □

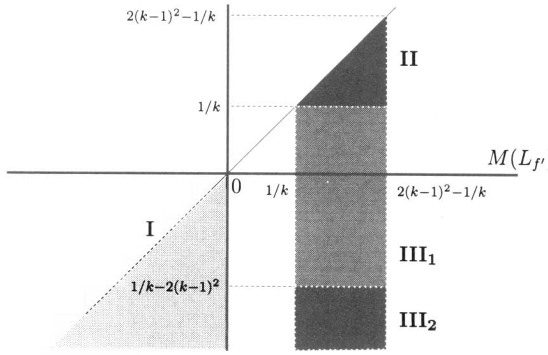

Fig. 1. Regions where $k > 1.754877$.

Notice that, for $t_0 \in [a, b]$ with $f(t_0) < 0$ and $t_0 \leq b + f(a)/f'(a)$, if the interval satisfies $a \leq b + f(a)/f'(a)$, the last result follows.

In practice, we can always apply the convex acceleration of Newton's method to solve (1). Taking into account the values of $m(L_{f'})$ and $M(L_{f'})$, we have the following cases (see Fig. 1):

(I) $M(L_{f'}) \leq 0$;
(II) $M(L_{f'}) < 2(k-1)^2 - 1/k$ and $m(L_{f'}) \geq 1/k$;
(III) $1/k < M(L_{f'}) < 2(k-1)^2 - 1/k$ and $m(L_{f'}) < 1/k$;
(III_1) $1/k \leq M(L_{f'}) < 2(k-1)^2 - 1/k$ and $1/k - 2(k-1)^2 < m(L_{f'}) < 1/k$;
(III_2) $1/k \leq M(L_{f'}) < 2(k-1)^2 - 1/k$ and $m(L_{f'}) \leq 1/k - 2(k-1)^2$.

Remark 2.1. Cases (I) and (II) follow from Theorems 2.1 and 2.2, respectively. Case (III) completes all the regions of the semiplane $M(L_{f'}) \geq m(L_{f'})$. For that, it suffices to take k going over $\mathbb{R}_+$.

We need a lemma to assure that the sequence (2) lies in $[a, b]$.

Lemma 2.2. Assuming that $|G'(t)| < 1$, where G is defined in (2), and that $|L_f(t)| \leq 1/k$ in $[a, b]$ for $k > 1.754877$, the following results hold:

(i) Let t_0, $\alpha \in [a, b]$ satisfy $f(\alpha) < 0$ and $f(t_0) > 0$. If $[(2k-1)/2(k-1)]f(b)/f'(b) \leq \alpha - a$, then the sequence (2) lies in $[a, b]$.
(ii) Let t_0, $\beta \in [a, b]$ satisfy $f(\beta) > 0$ and $f(t_0) < 0$. If $f(a)/f'(a) \geq \beta - b$, then the sequence (2) lies in $[a, b]$.

Theorem 2.3. Let t_0, α, $\beta \in [a, b]$, and let the assumptions of Lemma 2.2 be satisfied. Let $k > 1.754877$ and suppose that $|L_f(t)| \leq 1/k$ for $t \in [a, b]$.

(i) If $1/k \leq M(L_{f'}) < 2(k-1)^2 - 1/k$ and $1/k - 2(k-1)^2 < m(L_{f'}) < 1/k$ in $[a, b]$, then the sequence (2) converges to s.
(ii) If $M(|L_f|) \leq 2(k-1)^2 - |m(L_{f'})|$, $1/k \leq M(L_{f'}) < 2(k-1)^2 - 1/k$ and $m(L_{f'}) \leq 1/k - 2(k-1)^2$ in $[a, b]$, then the sequence (2) converges to s.

Proof. First, from the hypothesis it follows that $|G'(t)| \leq Q < 1$ for $t \in [a, b]$ and G defined in (2). Consequently, the sequence (2) lies in $[a, b]$. Now, it suffices to show that $|t_n - s| < Q^n |t_0 - s|$ is true for all $n \in \mathbb{N}$ by mathematical induction and that both cases hold. □

Therefore, we can apply always the convex acceleration of Newton's method to solve Eq. (1) [under certain restrictions for L_f in $[a, b]$, but $L_f(s) = 0$], and these results are applicable in a neighborhood of s.

Remark 2.2. Observe that, in practical situations, it suffices that $m(L_{f'})$ satisfies the hypotheses of Theorem 2.3 in case (i), since we can always find a finite and high enough value of k in order to take the region of plane which is necessary.

Note also that k is the least value satisfying the assumptions of Theorem 2.3 that concern us, as the interval for L_f will be bigger and then the assumption for L_f will be milder in $[a, b]$.

3. Sufficient Conditions for Convergence in the Complex Plane

In this section, under the Newton–Kantorovich conditions (Ref. 4), we establish a Kantorovich-type convergence theorem for the convex acceleration of Newton's method in the complex plane.

Here, we are concerned with the problem of approximating a locally unique solution z^* of the equation

$$f(z) = 0 \tag{4}$$

in the complex plane, where f is an holomorphic function defined on some open convex subset D of $\mathbb{C}$ with values in $\mathbb{C}$. Let $z_0 \in D$, and define the convex acceleration of Newton's method for all $n \geq 0$ by

$$z_{n+1} = H(z_n) = z_n - [f(z_n)/f'(z_n)][1 + L_f(z_n)/2(1 - L_f(z_n))]. \tag{5}$$

Let us denote

$$\overline{B(z, r)} = \{w \in \mathbb{C}; |w - z| \leq r\}.$$

The following Kantorovich-type conditions are satisfied:

(C1) $|f(z_0)/f'(z_0)| \leq a$.

(C2) $|f''(z)/f'(z_0)| \leq b$, for all $z \in D$.

(C3) The equation

$$p(t) \equiv (Mb^2/6)t^3 + (b/2)t^2 - t + a = 0, \qquad M > 0, \tag{6}$$

has a negative root and two positive roots r_1 and r_2, with $r_1 \leq r_2$. Equivalently,

$$ab \leq [1 + 4M - \sqrt{1 + 2M}]/3M(1 + \sqrt{1 + 2M}).$$

(C4) $|f(z)/f'(z_0)| \leq -p(t)/p'(t_0)$, for all $z \in D$.

(C5) $L_{f'}(z) \in \overline{B(0, -L_{p'}(r_1))}$, when $|z - z_0| \leq r_1$.

Define the scalar sequence $\{t_n\}$ given for all $n \geq 0$ by

$$t_0 = 0, \qquad t_{n+1} = P(t_n) = t_n - [p(t_n)/p'(t_n)][1 + L_p(t_n)/2(1 - L_p(t_n))], \tag{7}$$

where p is the polynomial (6). In light of Section 2, it is easy to prove that this sequence is increasing and converges to r_1.

We need first the following result for studying the convergence of (5) and the uniqueness of the solution.

Lemma 3.1. Under Conditions (C1)–(C5), $|L_f(z)| \leq L_p(t)$, provided that $|z - z_0| \leq t - t_0 < 1/b$.

Proof. Let L_f be defined by

$$L_f(z) = [f'(z_0)/f'(z)][f''(z)/f'(z_0)][f'(z_0)/f'(z)][f(z)/f'(z_0)].$$

Then, we get

$$\begin{aligned} f'(z)/f'(z_0) &\geq 1 - |z - z_0| \int_0^1 |f'(z_0)|^{-1}|f''(z_0 + t(z - z_0)| \, dt \\ &\geq 1 - b|z - z_0|, \end{aligned}$$

$$|f''(z)/f'(z_0)| \leq p''(t).$$

This completes the proof of the lemma. □

Let us denote by

$$m=(-1+\sqrt{1+2M})/bM$$

the minimum of polynomial (6).

Theorem 3.1. Let us assume that Conditions (C1)-(C5) hold. Then, sequence (5) is well defined for $n \geq 0$ and is convergent in $\overline{B(z_0, r_1)} \cap D$. If $r_1 < r_2$, the limit z^* is a unique solution of Eq. (4) in $B(z_0, m) \cap D$. If $r_1 = r_2$, z^* is unique in $\overline{B(z_0, m)} \cap D$. Furthermore, $|z^* - z_n| \leq r_1 - t_n$ for all $n \geq 0$.

Proof. Following Kantorovich (Ref. 4), the sequence (7) majorizes the sequence defined by (5) provided that

(i) $|H(z_0) - z_0| \leq P(t_0) - t_0$,
(ii) $|H'(z)| \leq P'(t)$, when $|z - z_0| \leq t - t_0$.

On the one hand, it is known that $0 \leq L_p(t) < 1/2$; then,

$$|H(z_0) - z_0| \leq (a/2)[1 + 1/(1 - |L_f(z_0)|)] \leq P(t_0) - t_0,$$

since $|L_f(z_0)| \leq L_p(t_0)$. On the other hand, we have

$$|H'(z)| \leq (1/2)[L_p(t)/(1 - L_p(t))]^2(L_p(t) + |L_{f'}(z)|)$$

and $|L_{f'}(z)| \leq -L_p(r_1)$ in $[0, r_1]$. Hence, condition (ii) holds.

To prove the uniqueness of solutions, it suffices to show that $P(t') \leq t'$ for $t' \in [r_1, m]$; see Ref. 4. □

4. Convergence in Banach Spaces

Let X, Y be Banach spaces, and consider a nonlinear operator $F: \Omega \subseteq X \to Y$ which is twice Fréchet-differentiable on an open convex set $\Omega_0 \subseteq \Omega$. Let us assume that $F'(x_0)^{-1} \in \mathscr{L}(Y, X)$ exists at some $x_0 \in \Omega_0$, where $\mathscr{L}(Y, X)$ is the set of bounded linear operators from Y into X. The convex acceleration of Newton's method for approximating a solution $x^* \in \Omega_0$ of the equation

$$F(x) = 0 \tag{8}$$

is in the form

$$x_{n+1} = x_n - [I + (1/2)L_F(x_n)(I - L_F(x_n))^{-1}]F'(x_n)^{-1}F(x_n), \qquad n \geq 0. \tag{9}$$

We denote by I the identity operator on X and by $L_F(x)$ the linear operator (Ref. 12) defined by

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

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

Following Yamamoto in Ref. 13 for the Halley method, the main assumption is that the Fréchet derivative $F''(x)$ of F satisfies a Lipschitz condition. Here, we extend this condition to operators that are Hölder (k, p)-continuous, with $k\geq 0$, $p\geq 0$; see Ref. 7.

From now on, we assume the following:

(i) there exists a continuous linear operator $\Gamma_0=F'(x_0)^{-1}$, $x_0\in\Omega_0$;
(ii) $\|\Gamma_0(F''(x)-F''(y))\|\leq k\|x-y\|^p$, $x, y\in\Omega_0$, $k\geq 0$, $p\geq 0$;
(iii) $\|\Gamma_0F(x_0)\|\leq a$, $\|\Gamma_0F''(x_0)\|\leq b$;
(iv) the equation

$$g(t)\equiv[k/(p+1)(p+2)]t^{p+2}+(b/2)t^2-t+a=0 \tag{10}$$

for $t\geq 0$, has two positive roots r_1 and r_2, with $r_1\leq r_2$;
(v) $\|\Gamma_0F'''(x)\|\leq -g'''(t)/g'(t_0)$, provided that $\|x-x_0\|\leq t-t_0$ and $p\in[0, 1)$.

Notice that the equation $g'(t)=0$ has only one positive solution u which is the minimum of g. Therefore, $g(u)\leq 0$ is a necessary and sufficient condition for the existence of positive solutions of $g(t)=0$. Let us denote these solutions by r_1 and r_2, with $r_1\leq r_2$. So, we have

$$g(t)=(r_1-t)(r_2-t)q(t),$$

with $q(r_1)\neq 0\neq q(r_2)$. Note that g is a decreasing convex function in $[0, u]$ and that $g(0)>0\geq g(u)$. These conditions are sufficient to show the convergence of the sequence

$$t_0=0,\ t_{n+1}=t_n-[g(t_n)/g'(t_n)][1+L_g(t_n)/2(1-L_g(t_n))], \qquad n\geq 0, \tag{11}$$

to r_1. Moreover, $\{t_n\}$ is an increasing sequence.

We will need the following result.

Lemma 4.1. The sequence $\{t_n\}$ given by (11) is a majorizing sequence of $\{x_n\}$ defined in (9), i.e.,

$$\|x_{n+1}-x_n\|\leq t_{n+1}-t_n, \qquad n\geq 0.$$

Proof. It is enough to show that the following statements are true, using induction on n, for all $n\geq 0$:

[I] there exists $\Gamma_n=F'(x_n)^{-1}$;

[II] $\|\Gamma_0 F''(x_n)\| \leq -g''(t_n)/g'(t_0)$;
[III] $\|\Gamma_n F'(x_0)\| \leq g'(t_0)/g'(t_n)$;
[IV] $\|\Gamma_0 F(x_n)\| \leq -g(t_n)/g'(t_0)$.

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

$$\int_{x_0}^{x_{n+1}} \Gamma_0[F''(x)-F''(x_0)]\,dx$$

$$=\Gamma_0 F'(x_{n+1})-I-\Gamma_0 F''(x_0)(x_{n+1}-x_0)$$

and (i)–(iv), we obtain

$$\|I-\Gamma_0 F'(x_{n+1})\| \leq [k/(p+1)]t_{n+1}^{p+1}+bt_{n+1}=1+g'(t_{n+1})<1.$$

Applying the Banach lemma, it follows that $(\Gamma_0 F'(x_{n+1}))^{-1}$ exists. Hence, [I] and [III] are true for $n+1$. Since we have

$$\|\Gamma_0 F''(x_{n+1})\| \leq k\|x_{n+1}-x_0\|^p+b \leq -g''(t_{n+1})/g'(t_0),$$

it follows that [II] is also true for $n+1$.

Case 1. If $p \geq 1$, we use the Altman technique (Refs. 13 and 14) and, taking into account (9), we deduce by the Taylor formula that

$$\Gamma_0 F(x_{n+1})=(1/8)\Gamma_0 F''(x_n)[L_F(x_n)(I-L_F(x_n))^{-1}\Gamma_n F(x_n)]^2$$
$$+\int_0^1 \Gamma_0[F''(x_n+t(x_{n+1}-x_n))-F''(x_n)]$$
$$\times (x_{n+1}-x_n)^2(1-t)\,dt;$$

by taking norms, we obtain

$$\|\Gamma_0 F(x_{n+1})\| \leq [g(t_n)/8]L_g(t_n)^3/(1-L_g(t_n))^2$$
$$+k(t_{n+1}-t_n)^{p+2}\int_0^1 t^p(1-t)\,dt.$$

Repeating the same process for the function g, we get

$$g(t_{n+1})=[g(t_n)/8]L_g(t_n)^3/(1-L_g(t_n))^2$$
$$+k\int_{t_n}^{t_{n+1}} (z^p-t_n^p)(t_{n+1}-z)\,dz.$$

Moreover, if $p \geq 1$, it follows immediately that

$$(t_{n+1}-t_n)^{p+2}\int_0^1 t^p(1-t)\,dt$$

$$=\int_{t_n}^{t_{n+1}} (z-t_n)^p(t_{n+1}-z)\,dz$$

$$\leq \int_{t_n}^{t_{n+1}} (z^p-t_n^p)(t_{n+1}-z)\,dz,$$

and, consequently,

$$\|\Gamma_0 F(x_{n+1})\| \leq g(t_{n+1}) = -g(t_{n+1})/g'(t_0). \tag{12}$$

Case 2. If $p \in [0, 1)$, we use the same process but taking as the remainder of the Taylor formula the approximation

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

Therefore, [IV] is true for $n+1$.

Finally, we have

$$\begin{aligned}\|x_{n+1}-x_n\| &= \|[I+(1/2)L_F(x_n)(I-L_F(x_n))^{-1}]\Gamma_n F(x_n)\| \\ &\leq -[1+L_g(t_n)/2(1-L_g(t_n))]g(t_n)/g'(t_n) \\ &= t_{n+1}-t_n. \end{aligned} \tag{13}$$

This completes the proof of Lemma 4.1. □

Theorem 4.1. Let us assume that conditions (i)–(v) hold. Let $u>0$ be the minimum of the function g defined in (10), and suppose that $g(u) \leq 0$. Assume that r_1 is the smallest positive root of Eq. (10) and $\overline{B(x_0, u)} \subset \Omega_0$. Then, the iteration $\{x_n\}$, $n \geq 0$, generated by (9) is well defined and converges to the solution x^* of the equation $F(x)=0$ in $\overline{B(x_0, r_1)}$. Moreover, we also have that $\|x^*-x_n\| \leq r_1 - t_n$, for all $n \geq 0$.

Proof. It follows from Lemma 4.1 that the sequence $\{t_n\}$ defined by (11) majorizes the sequence $\{x_n\}$ given by (9). Hence, the convergence of $\{t_n\}$ implies the convergence of $\{x_n\}$ to a limit x^*; see Refs. 4 and 15. Letting $n \to \infty$ in (12), we infer that $F(x^*)=0$.

Finally, for $v \geq 0$, it follows from (13) that

$$\|x_{n+v}-x_n\| \leq t_{n+v}-t_n;$$

and by letting $v \to \infty$, we obtain

$$\|x^* - x_n\| \leq r_1 - t_n, \qquad n \geq 0.$$

Furthermore, we also have

$$\|x^* - x_0\| \leq r_1 - t_0 = r_1. \qquad \square$$

5. Numerical Results

In the first three examples that follow, we illustrate our results by solving scalar equations. The digits shown in each example are significative.

Example 5.1. Consider the function $f(t) = t - \cos t$ defined on $[0, \pi/2]$. This function is nondecreasing and convex in $[0, \pi/2]$ with

$$L_f(t) = (t - \cos t) \cos t/(1 + \sin t)^2, \qquad L_{f'}(t) = \sin t/(\sin t - 1).$$

As $L_f(t) < 1$ and $L_{f'}(t) \leq 0$ for $t \in [0, \pi/2]$, we obtain a sequence $\{t_n\}$ decreasing and convergent to

$$s = 0.73908513321516064\,28,$$

a root of $f(t) = 0$ in $[0, \pi/2]$ by Theorem 2.1. See Table 1.

Example 5.2. To illustrate Theorem 2.2, we consider the real equation

$$f(t) = t^4 - r = 0, \text{ where } r \in (0, 1).$$

So, we get

$$L_f(t) = 3(t^4 - r)/(4t^4), \qquad L_{f'}(t) = 2/3.$$

If $k > 1.78339$, we have

$$1/k \leq L_{f'}(t) = 2/3 < 2(k-1)^2 - 1/k.$$

We choose $k = 1.79$, to obtain $|L_f(t)| \leq 1/k$ in $[0.731355, 1.18644]$ for $r > 0.499839$. Then, we take $[a, b] = [0.731355, 1.18644]$ that checks

$$1.09 = a + [(2k-1)/2(k-1)]f(b)/f'(b) \leq b = 1.18644,$$

Table 1. Results for Example 5.1.

n	t_n
0	1.0000000000000000000
1	0.7404989832636941698
2	0.7390851334050131377
3	0.7390851332151606428

Table 2. Results for Example 5.2.

n	t_n
0	1.1800000000000000000
1	0.8236846056354870351
2	0.8409040229165229872
3	0.8408964152537139204
4	0.8408964152537145430

and $r = 0.5$. Then, we take $t_0 = 1.18$ to obtain approximations of the solution

$$s = 0.8408964152537145430,$$

given in Table 2.

Example 5.3. Let us consider the nonlinear scalar equation

$$f(t) = -1 + \arcsin(1+t) = 0.$$

Then, we get

$$L_f(t) = (1+t)[-1 + \arcsin(1+t)]/\sqrt{1-(1-t)^2},$$

$$L_{f'}(t) = (2t^2 + 4t + 3)/(1+t)^2.$$

As the function $L_{f'}$ is nondecreasing in $[-0.9, -0.01]$, we have

$$M(L_{f'}) = L_{f'}(-0.9) = 102,$$

$$m(L_{f'}) = L_{f'}(-0.01) = 3.0203.$$

If $k > 8.14572$, from Theorem 2.5, we have

$$1/k \leq 102 < 2(k-1)^2 - 1/k,$$

$$1/k - 2(k-1)^2 \leq 3.0203 < 1/k.$$

Hence, we can choose $k = 8.15$ to obtain $|L_f(t)| \leq 1/k = 0.122699$ in the interval $[-0.214634, -0.123979]$. Moreover, we have

$$f(-0.214634)/f'(-0.214634) = -0.0598683 \geq \beta + 0.123979,$$

for some $\beta \in [-0.214634, -0.183847]$ with $f(\beta) > 0$. Therefore, we can take $[a, b] = [-0.214634, -0.183847]$ and choose $x_0 = -0.21$ to get the approximations to the solution

$$s = -0.1585290151921035,$$

given in Table 3.

Table 3. Results for Example 5.3.

n	t_n
0	−0.2100000000000000
1	−0.1581347287092053
2	−0.1585290154830309
3	−0.1585290151921035

An interesting application of Theorem 4.1 is given by the following example.

Example 5.4. Consider the differential equation

$$y'' + y' - y^3 = 0, \tag{14a}$$

$$y(0) = 0 = y(1). \tag{14b}$$

We divide the interval [0, 1] into n subintervals and we set $h = 1/n$. Let $\{z_k\}$ be the points of the subdivisions with

$$0 = z_0 < z_1 < \cdots < z_n = 1,$$

and corresponding values of the function

$$y_0 = y(z_0), \qquad y_1 = y(z_1), \qquad \ldots, \qquad y_n = y(z_n).$$

Standard approximations for the first and second derivatives are given respectively by

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

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

$$F(y) = G(y) + hJ(y) - 2h^2 g(y),$$

where

$$G = \begin{bmatrix} -4 & 2 & 0 & \cdots & 0 \\ 2 & -4 & 2 & \cdots & 0 \\ 0 & 2 & -4 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -4 \end{bmatrix}, \qquad J = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

$$g(y) = \begin{bmatrix} y_1^3 \\ y_2^3 \\ \vdots \\ y_{n-1}^3 \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix}.$$

Then, we get

$$F'(y)=G+hJ-6h^2\begin{bmatrix} y_1^2 & 0 & \cdots & 0 \\ 0 & y_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & y_{n-1}^2 \end{bmatrix},$$

$$F''(y)=-12h^2\begin{bmatrix} y_1 & 0 & \cdots & 0 \\ 0 & y_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & y_{n-1} \end{bmatrix}.$$

Let $x\in\mathbb{R}^{n-1}$, $A\in\mathbb{R}^{n-1}\times\mathbb{R}^{n-1}$, and define the norms of x and A by

$$\|x\|=\max_{1\leq i\leq n-1}|x_i|, \qquad \|A\|=\max_{1\leq i\leq n-1}\sum_{k=1}^{n-1}|a_{ik}|.$$

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

$$\begin{aligned}\|F''(x)-F''(y)\| &= \|\mathrm{diag}\{-12h^2(x_i-y_i)\}\| \\ &= 12h^2 \max_{1\leq i\leq n-1}|x_i-y_i| \\ &\leq 12h^2\|x-y\|.\end{aligned}$$

We choose $n=10$ and as the solution should vanish at the endpoints and be positive in the interior, a reasonable choice of initial approximation seems to be $\exp(\pi x)/100$. This gives us the following vector:

$$x_0=\begin{bmatrix} 0.0136911 \\ 0.0187446 \\ 0.0256633 \\ 0.0351359 \\ 0.0481048 \\ 0.0658606 \\ 0.0901703 \\ 0.1234530 \\ 0.1690200 \end{bmatrix}.$$

We get the following results for Theorem 4.1:

$$a=\|\Gamma_0 F(x_0)\|=0.168893,$$

$$b=\|\Gamma_0 F''(x_0)\|=0.0465696,$$

$$k = 12h^2\|\Gamma_0\| = 0.734399.$$

Therefore, Eq. (10) becomes

$$0.168893 - t + 0.0232848t^2 + 0.1224t^3 = 0.$$

This equation has two positive real solutions,

$$r_1 = 0.17017, \qquad r_2 = 2.67306.$$

Hence, by Theorem 4.1, the sequence (9) converges to a solution $y^* = (y_1^*, y_2^*, \ldots, y_0^*)$ of the equation $F(y) = 0$ in $\overline{B(x_0, r_1)}$.

Finally, note that, to solve the differential equation (14), the following interpolation problem can be considered

$$\begin{bmatrix} 0 & 1/10 & 2/10 & \cdots & 9/10 & 1 \\ 0 & y_1^* & y_2^* & \cdots & y_9^* & 0 \end{bmatrix},$$

and its solution is an approximation to the solution of Eq. (14).

References

1. CHEN, D., ARGYROS, I. K., and QIAN, Q. S., *A Local Convergence Theorem for the Super-Halley Method in a Banach Space*, Applied Mathematics Letters, Vol. 7, pp. 49–52, 1994.
2. HERNÁNDEZ, M. A., *Newton–Raphson's Method and Convexity*, Zbornik Radova Prirodno-Matematičkog Fakulteta Univerziteta u Novom Dadu, Serija Za Matematiku, Vol. 22, pp. 159–166, 1992.
3. HERNÁNDEZ, M. A., *A Note on Halley's Method*, Numerische Mathematik, Vol. 59, pp. 273–276, 1991.
4. KANTOROVICH, L. V., and AKILOV, G. P., *Functional Analysis*, Pergamon Press, Oxford, England, 1982.
5. ARGYROS, I. K., *On the Secant Method*, Publicationes Mathematicae Debrecen, Vol. 43, pp. 223–238, 1993.
6. ARGYROS, I. K., *Some Methods for Finding Error Bounds for Newton-Like Methods under Mild Differentiability Conditions*, Acta Mathematica Hungarica, Vol. 61, pp. 183–194, 1993.
7. ARGYROS, I. K., and SZIDAROVSZKY, F., *The Theory and Applications of Iteration Methods*, CRC Press, Boca Raton, Florida, 1993.
8. ULM, S., *Iterations Methods with Divided Differences of the Second Order*, Doklady Adademii Nauk SSSR, Vol. 158, pp. 55–58, 1964; Soviet Mathematics Doklady, Vol. 5, pp. 1187–1190, 1964.
9. ARGYROS, I. K., *Convergence Results for the Super-Halley Method Using Divided Differences*, Functiones et Approximatio Commentarii Mathematici, Vol. 23, pp. 109–122, 1994.

10. ARGYROS, I. K., *On the Method of Tangent Hyperbolas*, Approximation Theory and Its Applications, Vol. 12, pp. 78–96, 1996.
11. POTRA, F. A., *On an Iterative Algorithm of Order 1.839... for Solving Nonlinear Operator Equations*, Numerical Functional Analysis and Optimization, Vol. 7, pp. 75–106, 1984–85.
12. GUTIÉRREZ, J. M., HERNÁNDEZ, M. A., and SALANOVA, M. A., *Accessibility of Solutions by Newton's Method*, International Journal of Computer Mathematics, Vol. 57, pp. 239–247, 1995.
13. YAMAMOTO, T., *On the Method of Tangent Hyperbolas in Banach Spaces*, Journal of Computational and Applied Mathematics, Vol. 21, pp. 75–86, 1988.
14. ALTMAN, M., *Concerning the Method of Tangent Hyperbolas for Operator Equations*, Bulletin de L'Academie Polonaise des Sciences, Série des Sciences Mathématiques, Astronomie et Physique, Vol. 9, pp. 633–637, 1961.
15. RHEINBOLDT, W. C., *A Unified Convergence Theory for a Class of Iterative Processes*, SIAM Journal on Numerical Analysis, Vol. 5, pp. 42–63, 1968.