Applied Mathematics and Computation 117 (2001) 223–239
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

An acceleration of Newton's method: Super-Halley method ☆

J.M. Gutiérrez *, M.A. Hernández [1]

Universidad de La Rioja, Department of Matemáticas y Computación, C/ Luis de Ulloa s/n, 26004, Logroño, Spain

Abstract

From a study of the convexity we give an acceleration for Newton's method and obtain a new third order method. Then we use this method for solving non-linear equations in Banach spaces, establishing conditions on convergence, existence and uniqueness of solution, as well as error estimates © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Non-linear equation; Newton's method; Kantorovich assumptions; Iterative processes; Third order method

1. Introduction

The study of concavity and convexity of a real function is an old problem studied by the mathematicians. It is perfectly established when a function is concave or convex. However, it is not so developed how to measure this concavity or convexity. The degrees of convexity introduced by Jensen and Popoviciu [6] are interesting from the theorical standpoint, but their practical application is too difficult.

☆ Supported in part by a grant of the DGES PB-96-0120-C03-02 and a grant of the University of La Rioja API-97/B12.

* *E-mail address:* jmguti@dmc.unirioja.es (J.M. Gutiérrez).
[1] E-mail: mahernan@dmc.unirioja.es

0096-3003/01/$ - see front matter © 2001 Elsevier Science Inc. All rights reserved.
PII: S0096-3003(99)00175-7

Another measure of the convexity is suggested by Bohr–Mollerup's Theorem [2]. In this result appears the concept of log-convex function, that is, a function whose logarithm is a convex function. The degree of logarithmic convexity, introduced in Ref. [13], is a measure of this kind of convexity. Let $f : [a,b] \subseteq \mathbb{R} \to \mathbb{R}$ be a convex, twice differentiable function on an interval $[a,b]$ and $t_0 \in [a,b]$ such that $f'(t_0) \neq 0$. The degree of logarithmic convexity of f at t_0 is

$$L_f(t_0) = \frac{f(t_0)f''(t_0)}{f'(t_0)^2}.$$

One of the most interesting applications of the degree of logarithmic convexity is its relation with the velocity of convergence of Newton's method [14].

We assume for f to satisfy $f'(x) < 0$, $f''(x) > 0$, for $x \in [a,b]$ and $f(a) > 0 > f(b)$. In this situation it is well-known that the sequence defined by

$$t_{n+1} = t_n - \frac{f(t_n)}{f'(t_n)}, \quad t_0 = a \tag{1}$$

converges to t^*, the only solution of $f(t) = 0$ in $[a,b]$.

Let g be another function satisfying the same conditions as f in $[a,b]$, with $g(t^*) = 0$, and let $\{s_n\}$ be the sequence defined as follows:

$$s_{n+1} = s_n - \frac{g(s_n)}{g'(s_n)}, \quad s_0 = t_0. \tag{2}$$

If $L_g(t) < L_f(t)$ for $t \in [t_0, t^*)$, it was shown (see Ref. [14]), that the sequence $\{s_n\}$ defined by (2) converges to the root t^* faster than $\{t_n\}$ defined by (1). Moreover, $t_n \leqslant s_n \leqslant t^*$, for $n \geqslant 0$.

As an application of this result, we derive an acceleration for Newton's method and a new third order method that we call Super-Halley method. Next, we extend this method to Banach spaces and we obtain convergence results. This method has also been studied by other authors (see Ref. [5]). However, the results they have obtained are only valid for quadratic operators (see Ref. [11]). The point is that they try to majorize the sequence in the Banach space by a real sequence arising from a quadratic polynomial. This is not possible in general because, as we see, Super-Halley method is a third order method, but when it is applied to a quadratic polynomial, the order is four. In this paper it is analysed the case of taking a third degree polynomial as a majorizing function, in a similar way as Yamamoto [18] did for the method of tangent hyperbolas (see also Ref. [12]). So the convergence can be also stated for non-quadratic operators.

2. Super-Halley method for scalar equations

Let $f : [a,b] \subseteq \mathbb{R} \to \mathbb{R}$ be a function satisfying $f'(t) < 0$, $f''(t) > 0$, for $t \in [a,b]$ and $f(a) > 0 > f(b)$. Let us denote $\{t_n\}$ the Newton sequence

$$t_{n+1} = t_n - \frac{f(t_n)}{f'(t_n)}, \quad t_0 = a. \tag{3}$$

Notice that the degree of logarithmic convexity of a straight line is zero. Then, taking $g(t) = f'(t^*)(t - t^*)$ we obviously obtain a sequence that converges to t^* faster than $\{t_n\}$. The problem is that t^* is unknown. Instead of $f'(t^*)(t - t^*)$ we take its Taylor approximation

$$g(t) = f(t) - \frac{f''(t^*)}{2}(t - t^*)^2.$$

So, we deduce the following approximations:

$$g(t_n) \simeq f(t_n) - \frac{f''(t_n)}{2}(t_n - t_{n+1})^2,$$

$$g'(t_n) \simeq f'(t_n) - f''(t_n)(t_n - t_{n+1}).$$

Thus, we obtain an acceleration for Newton's method (3), which is defined by

$$s_{n+1} = t_n - \frac{g(t_n)}{g'(t_n)} = t_n - \left[1 + \frac{L_f(t_n)}{2(1 - L_f(t_n))}\right]\frac{f(t_n)}{f'(t_n)}. \tag{4}$$

Before proving that (4) is an acceleration of Newton's method, we give two elementary lemmas.

Lemma 2.1. *Let us write*

$$G(t) = t - \left[1 + \frac{L_f(t)}{2(1 - L_f(t))}\right]\frac{f(t)}{f'(t)}. \tag{5}$$

Then, we have

$$G'(t) = \frac{L_f^2(t)\Big(L_f(t) - L_{f'}(t)\Big)}{2(1 - L_f(t))^2}.$$

The following lemma was applied by Altman to prove the convergence of the method of tangent hyperbolas [1].

Lemma 2.2. *Let us assume that f satisfies $f(t) \geqslant 0$ for $a \leqslant t \leqslant t^*$, where t^* is a root of f, i.e., $f(t^*) = 0$. The second derivative f'' is non-decreasing for $a \leqslant t \leqslant t^*$. Then, $L_f(t) \leqslant 1/2$, for $a \leqslant t \leqslant t^*$.*

Theorem 2.3. *Under the previous assumptions for f, the sequence $\{s_n\}$ defined by* (4) *is an acceleration of the Newton's method* (3).

Proof. Let G the function defined by (5), $F(t) = t - (f(t)/f'(t))$, $\{t_n\}$ and $\{s_n\}$ given by (3) and (4), respectively. By Lemma 2.1 we have

$$\lim_{n\to\infty} \frac{t^* - s_{n+1}}{t^* - t_{n+1}} = \lim_{n\to\infty} \frac{t^* - G(t_n)}{t^* - F(t_n)} = \lim_{t\to t^*} \frac{t^* - G(t)}{t^* - F(t)} = 0. \qquad \square$$

From this acceleration we can define a new method as follows:

$$t_0 = a, \quad t_{n+1} = G(t_n) = t_n - \left[1 + \frac{L_f(t_n)}{2(1 - L_f(t_n))}\right] \frac{f(t_n)}{f'(t_n)}. \tag{6}$$

We call (6) Super-Halley method. In the next result we give conditions on the convergence of (6).

Theorem 2.4. *Let us assume that f satisfies $f'(t) < 0$, $f''(t) > 0$, for $t \in [a, b]$, and $f(a) > 0 > f(b)$. Let $t^* \in (a, b)$ such that $f(t^*) = 0$. Suppose*

$$L_{f'}(t) \leqslant L_f(t) < 1, \quad t \in [a, t^*). \tag{7}$$

Then the sequence $\{t_n\}$ defined by (6) *converges to t^*. In addition, $\{t_n\}$ is increasing.*

Proof. We write (6) in the form $t_{n+1} = 1/2[F(t_n) + Q(t_n)]$, where

$$F(t) = t - \frac{f(t)}{f'(t)}, \quad Q(t) = t - \frac{q(t)}{q'(t)}, \quad q(t) = \frac{f(t)}{f'(t)}. \tag{8}$$

Then, by using the Mean Value Theorem

$$t^* - t_1 = \frac{F + Q}{2}(t^*) - \frac{F + Q}{2}(t_0) \leqslant \left(\frac{F + Q}{2}\right)'(z_0)(t^* - t_0)$$

for some $z_0 \in (t_0, t^*)$. Taking into account Lemma 2.1 and (7) we deduce

$$F'(t) + Q'(t) = \frac{L_f^2(t)}{[1 - L_f(t)]^2} \left[L_f(t) - L_{f'}(t)\right] \geqslant 0, \quad t \in [a, t^*)$$

and consequently, $t_1 \leqslant t^*$. Following an inductive process is easy to check $t_n \leqslant t^*$ for $n \geqslant 0$.

On the other hand we have

$$t_{n+1} - t_n = -\frac{f(t_n)}{f'(t_n)} H\big(L_f(t_n)\big),$$

where $H(t) = 1 + (t/2(1-t))$.For $n \geqslant 0$, we derive from Lemma 2.2 that $H(L_f(t_n)) > 0$ and $t_{n+1} \geqslant t_n$. Then $\{t_n\}$ converges to $u \in [a,b]$. Making $n \to \infty$ in (6) and taking into account $H(L_f(t_n)) \geqslant 1$ for $n \geqslant 0$, we have $f(u) = 0$. As f has a unique root in $[a,b]$, we conclude $u = t^*$. □

Note. A sufficient condition for (7) to hold is

$$f'''(t) \geqslant 0, \quad t \in [t_0, t^*].$$

In fact, it is known (Lemma 2.2) that f in the previous situation satisfies $L_f(t) \leq 1/2 < 1$. The first inequality follows inmediately since $L_f(t) \geqslant 0 \geqslant L_{f'}(t)$, for $t \in [t_0, t^*]$.

If (7) is not fulfilled, the convergence of (6) is a hard problem. We cannot guarantee the sequence obtained in this case to be increasing towards t^*. In Ref. [13] more general convergence conditions for (6) have been established.

The method (6) is a third order method. This follows as a consequence of the Gander's result [8].

Theorem 2.5. *Let t^* be a simple zero of f and H any function with $H(0) = 1$, $H'(0) = 1/2$, $|H''(0)| < \infty$. The iteration*

$$t_{n+1} = t_n - H\big(L_f(t_n)\big)\frac{f(t_n)}{f'(t_n)}$$

is of third order.

However, if we use the method (6) for solving the equation $p(t) = 0$, p being a quadratic polynomial, we obtain a fourth order method, as we state at the next result.

Theorem 2.6. *Let p be a quadratic polynomial with a simple zero t^*. The Super-Halley method* (6) *for solving the equation $p(t) = 0$ is a fourth order method.*

Proof. Let G be the function defined by (5). Observe that $G(t^*) = t^*$. By Lemma 2.1 we obtain, for a quadratic polynomial p,

$$G'(t) = \frac{L_p^3(t)}{2(1 - L_p(t))^2}$$

and then $G'(t^*) = 0$. We can prove without difficulty that $G''(t^*) = 0$ and $G'''(t^*) = 0$ too. However,

$$G^{(\mathrm{iv})}(t^*) = \frac{3p''(t^*)^3}{p'(t^*)^3} \neq 0. \qquad \square$$

The previous results about the order of convergence of the sequence (6) hold if t^* is a simple zero of f. When the multiplicity of t^* is $m > 1$, we only can guarantee linear convergence [17].

In Theorem 2.6 we have established that Super-Halley method to solve a equation $p(t) = 0$, where p is a quadratic polynomial, is a fourth order method. Moreover, we can prove in this case that two iterations of Newton's method is equivalent to one iteration of Super-Halley method.

Theorem 2.7. *Let p be a quadratic polynomial with two positive roots. Let us suppose, without loss of generality, that*

$$p(t) = bt^2 - t + a, \quad ab \leqslant \frac{1}{4}.$$

Denote by $\{z_n\}$ *the Newton sequence to solve* $p(t) = 0$, *and by* $\{t_n\}$ *the iterates* (6). *Suppose* $z_0 = t_0$. *Then* $z_{2n} = t_n$ *for* $n \geqslant 0$.

Proof. Let G and F be the functions defined in (5) and (8), respectively. Then,

$$\begin{aligned} F(F(t)) &= \frac{b(bt^2 - a)^2 - a(2bt-1)^2}{(2bt-1)(2b(bt^2-a) - (2bt-1))} \\ &= \frac{b^3t^4 - 6ab^2t^2 + 4abt + a^2b - a}{(2bt-1)(2b(bt^2-a) - (2bt-1))} = G(t). \end{aligned}$$

Therefore, $z_2 = F(z_1) = F(F(z_0)) = G(z_0) = G(t_0) = t_1$, $z_4 = F(z_3) = F(F(z_2)) = G(z_2) = G(t_1) = t_2$ and so on. $\square$

3. Super-Halley method in Banach spaces

Let X and Y be Banach spaces and $F : \Omega \subseteq X \to Y$, a non-linear, twice Fréchet differentiable operator in an open convex domain $\Omega_0 \subseteq \Omega$. For solving the equation

$$F(x) = 0 \tag{9}$$

we can use the next generalization of (6):

$$x_{n+1} = x_n - \left[I + \frac{1}{2}L_F(x_n)\Delta_n^{-1}\right]\Gamma_n F(x_n), \quad n \geqslant 0, \tag{10}$$

where $x_0 \in \Omega_0$, I is the identity operator on X, $\Gamma_n = [F'(x_n)]^{-1}$, $\Delta_n = [I - L_F(x_n)]$ and $L_F(x_n)$ is the following linear operator [10]:

$$L_F(x_n)[\ ] = \Gamma_n F''(x_n)(\Gamma_n F(x_n), [\ ]),$$

provided that Γ_n and Δ_n^{-1} exist at each step.

We use the majorant principle due to Kantorovich [15] to prove the convergence of the method. Following Altman [1] we could establish convergence of (10) under majorant assumptions for F'' and F'''. In this paper, taking into account the technique employed by Yamamoto [18], we establish results on convergence of (10) under weaker conditions. From now on we assume that F given by (9) satisfies the following conditions:

(i) There exists a continous linear operator $\Gamma_0 = [F'(x_0)]^{-1}$, $x_0 \in \Omega_0$.
(ii) $\|\Gamma_0(F''(x) - F''(y))\| \leqslant k\|x-y\|$, $x, y \in \Omega_0$, $k \geqslant 0$.
(iii) $\|\Gamma_0 F(x_0)\| \leqslant a$, $\|\Gamma_0 F''(x_0)\| \leqslant b$.
(iv) The equation

$$p(t) \equiv \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a = 0 \tag{11}$$

has one negative root and two positive roots r_1 and r_2 ($r_1 \leqslant r_2$) if $k > 0$, or has two positive roots r_1 and r_2 ($r_1 \leqslant r_2$) if $k = 0$. Equivalently, [18],

$$a \leqslant \frac{b^2 + 4k - b\sqrt{b^2+2k}}{3k\left(b + \sqrt{b^2+2k}\right)} \quad \text{if } k > 0,$$

or $ab \leqslant 1/2$ if $k = 0$.

Observe that, in both cases, Super-Halley method

$$t_0 = 0, \quad t_{n+1} = t_n - \left(1 + \frac{L_p(t_n)}{2(1 - L_p(t_n))}\right)\frac{p(t_n)}{p'(t_n)}, \quad n \geqslant 0, \tag{12}$$

for solving the equation $p(t) = 0$, with p given by (11), converges to r_1, the smallest positive root of $p(t) = 0$ (see Theorem 2.4).

To establish the convergence of (10) to x^*, a solution of (9), the uniqueness of solution and the error estimates we shall need the following lemmas.

Lemma 3.1. *The iterates* (10) *are well-defined for* $n \geqslant 0$, *converge to* x^*, *a solution of* (9) *and*

$$\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n, \tag{13}$$

$$\|x^* - x_n\| \leqslant r_1 - t_n. \tag{14}$$

Proof. For $n \geqslant 0$ we prove

[I$_n$] There exists $\Gamma_n = [F'(x_n)]^{-1}$.
[II$_n$] $\|\Gamma_0 F''(x_n)\| \leqslant -p''(t_n)/p'(t_0)$.
[III$_n$] $\|\Gamma_n F'(x_0)\| \leqslant p'(t_0)/p'(t_n)$.
[IV$_n$] $\|\Gamma_0 F(x_n)\| \leqslant -p(t_n)/p'(t_0)$.

[V$_n$] There exists $\varDelta_n^{-1} = [I - L_F(x_n)]^{-1}$ and $\|\varDelta_n^{-1}\| \leqslant 1/(1 - L_p(t_n))$.

Note that **[V$_{n+1}$]** follows as a consequence of **[II$_{n+1}$]**, **[III$_{n+1}$]**, **[IV$_{n+1}$]** and Lemma 2.2 which guarantees $L_p(t) \leqslant 1/2$. Thus, we prove **[I$_{n+1}$]**–**[IV$_{n+1}$]** using induction. Applying Altman's technique, (see Ref. [1]), **[I$_{n+1}$]**, **[II$_{n+1}$]** and **[III$_{n+1}$]** follow inmediately.

To prove **[IV$_{n+1}$]**, using Taylor's formula and (10), we obtain

$$\begin{aligned}
F(x_{n+1}) = & -\frac{1}{2}F''(x_n)\varGamma_n F(x_n)\Delta_n^{-1}\varGamma_n F(x_n) + \frac{1}{2}F''(x_n)(x_{n+1} - x_n)^2 \\
& + \int_{x_n}^{x_{n+1}} \left[F''(x) - F''(x_n)\right](x_{n+1} - x)\mathrm{d}x \\
= & -\frac{1}{2}F''(x_n)\varGamma_n F(x_n)\varDelta_n^{-1}\varGamma_n F(x_n) \\
& + \frac{1}{2}F''(x_n)[\varGamma_n F(x_n)]^2 + \frac{1}{8}F''(x_n)\left[L_F(x_n)\varDelta_n^{-1}\varGamma_n F(x_n)\right]^2 \\
& + \frac{1}{2}F''(x_n)\varGamma_n F(x_n)L_F(x_n)\varDelta_n^{-1}\varGamma_n F(x_n) \\
& + \int_{x_n}^{x_{n+1}} [F''(x) - F''(x_n)](x_{n+1} - x)\mathrm{d}x.
\end{aligned}$$

Bearing in mind $\varDelta_n^{-1} = I + L_F(x_n)\varDelta_n^{-1}$ and writing $y_n = L_F(x_n)\varDelta_n^{-1}\varGamma_n F(x_n)$, we have

$$F(x_{n+1}) = \frac{1}{8}F''(x_n)y_n^2 + \int_{x_n}^{x_{n+1}} [F''(x) - F''(x_n)](x_{n+1} - x)\mathrm{d}x.$$

Denote $\tau_n = L_p(t_n)$. Notice that

$$\|y_n\| \leqslant \|L_F(x_n)\|\|\varDelta_n^{-1}\|\|\varGamma_n F(x_n)\| \leqslant -\frac{\tau_n p(t_n)}{(1 - \tau_n)p'(t_n)}$$

and therefore,

$$\|\varGamma_0 F(x_{n+1})\| \leqslant \frac{1}{8}\frac{\tau_n^3 p(t_n)}{(1 - \tau_n)^2} + \frac{k}{6}\|x_{n+1} - x_n\|^3 \leqslant \frac{1}{8}\frac{\tau_n^3 p(t_n)}{(1 - \tau_n)^2} + \frac{k}{6}(t_{n+1} - t_n)^3.$$

Consequently

$$\|\varGamma_0 F(x_{n+1})\| \leqslant p(t_{n+1}) \tag{15}$$

and we conclude the induction.

So, we have

$$\begin{aligned}
\|x_{n+1} - x_n\| = & \left\| \left[I + \frac{1}{2}L_F(x_n)\right]\varDelta_n^{-1}\varGamma_n F(x_n)\right\| \\
\leqslant & \left[1 + \frac{L_p(t_n)}{2(1 - L_p(t_n))}\right]\frac{p(t_n)}{p'(t_n)} = t_{n+1} - t_n,
\end{aligned}$$

then (13) happens and $\{t_n\}$ majorizes $\{x_n\}$. The convergence of $\{t_n\}$ (see Theorem 2.4 and its note) implies the convergence of $\{x_n\}$ to a limit x^*. Making $n \to \infty$ in (15), we deduce $F(x^*) = 0$.

Finally, for $p \geqslant 0$, $\|x_{n+p} - x_n\| \leqslant t_{n+p} - t_n$. Making $p \to \infty$ we obtain (14). □

Lemma 3.2. *Under the previous assumptions we have, for* $k > 0$,

$$\frac{\|x^* - x_{n+1}\|}{r_1 - t_{n+1}} \leqslant \left(\frac{\|x^* - x_n\|}{r_1 - t_n}\right)^3, \quad n \geqslant 0$$

and, for $k = 0$,

$$\frac{\|x^* - x_{n+1}\|}{r_1 - t_{n+1}} \leqslant \left(\frac{\|x^* - x_n\|}{r_1 - t_n}\right)^4, \quad n \geqslant 0.$$

Proof. This proof follows the technique used by Yamamoto in Ref. [18] for the method of tangent hyperbolas. Observe that $\Delta_n^{-1} = I + L_F(x_n)\Delta_n^{-1}$ and then,

$$I + \frac{1}{2}L_F(x_n)\Delta_n^{-1} = \frac{1}{2}(I + \Delta_n^{-1}) = \frac{1}{2}\Delta_n^{-1}(\Delta_n + I) = \Delta_n^{-1}\left(I - \frac{1}{2}L_F(x_n)\right).$$

Taking this into account we deduce

$$\begin{aligned}
x^* - x_{n+1} &= x^* - x_n + \Delta_n^{-1}\left(I - \frac{1}{2}L_F(x_n)\right)\Gamma_n F(x_n) \\
&= -\Delta_n^{-1}\Gamma_n\left[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n) - \frac{1}{2}F''(x_n)(x^* - x_n)^2\right] \\
&\quad + (x^* - x_n) - \frac{1}{2}\Delta_n^{-1}L_F(x_n)\Gamma_n F(x_n) - \Delta_n^{-1}(x^* - x_n) \\
&\quad - \frac{1}{2}\Delta_n^{-1}\Gamma_n F''(x_n)(x^* - x_n)^2 \\
&= -\Delta_n^{-1}\Gamma_n\int_{x_n}^{x^*}\left[F''(x) - F''(x_n)\right](x^* - x)\mathrm{d}x + \left[I - \Delta_n^{-1}\right](x^* - x_n) \\
&\quad - \frac{1}{2}\Delta_n^{-1}\Gamma_n F''(x_n)\left[(\Gamma_n F(x_n))^2 + (x^* - x_n)^2\right] \\
&= -\Delta_n^{-1}\Gamma_n\int_{x_n}^{x^*}\left[F''(x) - F''(x_n)\right](x^* - x)\mathrm{d}x \\
&\quad - \frac{1}{2}\Delta_n^{-1}\Gamma_n F''(x_n)\left[\Gamma_n\int_{x_n}^{x^*}F''(x)(x^* - x)\mathrm{d}x\right]^2
\end{aligned}$$

and therefore

$$\begin{aligned}\|x^* - x_{n+1}\| &\leqslant \|\Delta_n^{-1}\| \|\Gamma_n F'(x_0)\| \left\| \int_{x_n}^{x^*} \Gamma_0 [F''(x) - F''(x_n)](x^* - x)\mathrm{d}x \right\| \\ &\quad + \frac{1}{2} \|\Delta_n^{-1}\| \|\Gamma_n F'(x_0)\| \|\Gamma_0 F''(x_n)\| \left\| \Gamma_n \int_{x_n}^{x^*} F''(x)(x^* - x)\mathrm{d}x \right\|^2 \\ &\leqslant -\frac{k(r_1 - t_n)^3}{6p'(t_n)(1 - L_p(t_n))} \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^3 \\ &\quad - \frac{p''(t_n)}{2p'(t_n)^3(1 - L_p(t_n))} \left[\int_{t_n}^{r_1} p''(z)(r_1 - z)\mathrm{d}z \right]^2 \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^4. \end{aligned} \tag{16}$$

For $k > 0$, from (14), we have

$$\begin{aligned}\|x^* - x_{n+1}\| &\leqslant -\Bigg[\frac{k(r_1 - t_n)^3}{6p'(t_n)(1 - L_p(t_n))} + \frac{p''(t_n)}{2p'(t_n)^3(1 - L_p(t_n))} \\ &\qquad \times \left(\int_{t_n}^{r_1} p''(z)(r_1 - z)\mathrm{d}z \right)^2 \Bigg] \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^3 \\ &= (r_1 - t_{n+1}) \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^3. \end{aligned}$$

But if $k = 0$, from (16), we deduce

$$\begin{aligned}\|x^* - x_{n+1}\| &\leqslant -\frac{p''(t_n)}{2p'(t_n)^3(1 - L_p(t_n))} \left[\int_{t_n}^{r_1} p''(z)(r_1 - z)\mathrm{d}z \right]^2 \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^4 \\ &= (r_1 - t_{n+1}) \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^4. \qquad \square \end{aligned}$$

Relation (14) allows us to obtain error estimates for the sequence $\{x_n\}$ in terms of the real sequence $\{t_n\}$. When $k = 0$ we derive $\{t_n\}$ from a quadratic polynomial. Following Ostrowski [16] we obtain the next error expression when Super-Halley method is applied to a quadratic polynomial.

Lemma 3.3. *Let p be the polynomial given by* (11) *with $k = 0$, that is*

$$p(t) = \frac{b}{2}t^2 - t + a.$$

We assume p has two positive roots $r_1 \leqslant r_2$. Let $\{t_n\}$ the method defined by (12), *then*

$$r_1 - t_n = \frac{(r_2 - r_1)\theta^{4^n}}{1 - \theta^{4^n}}, \quad n \geqslant 0, \quad r_1 < r_2, \quad \theta = \frac{r_1}{r_2},$$

$$r_1 - t_n = \frac{r_1}{4^n}, \quad n \geqslant 0, \quad r_1 = r_2.$$

When $k > 0$ the real sequence $\{t_n\}$ in (14) is obtained from a cubic polynomial. In this case, it is difficult to obtain an error expression by Ostrowski method. In the next lemma and in a different way [12], we establish estimates for the error in this situation.

Lemma 3.4. *Let p be the polynomial given by* (11) *with $k > 0$, that is*

$$p(t) = \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a.$$

Let us assume that p has two positive roots $r_1 \leqslant r_2$ and a negative root, $-r_0$. Let us consider Super-Halley method $\{t_n\}$ defined by (12), *then, if $r_1 < r_2$,*

$$r_1 - t_n \sim \frac{(r_2 - r_1)\theta^{3^n}}{\sqrt{\lambda} - \theta^{3^n}}, \quad n \geqslant 0,$$

where

$$\lambda = \frac{r_2 - r_1}{r_0 + r_1} < 1, \quad \theta = \sqrt{\lambda}\frac{r_1}{r_2} < 1.$$

If $r_1 = r_2$, we have

$$\frac{r_1}{4^n} \leqslant r_1 - t_n \leqslant \frac{r_1}{3^n}.$$

Proof. The polynomial p defined above can be written in the form

$$p(t) = \frac{k}{6}(r_1 - t)(r_2 - t)(r_0 + t).$$

Notice that comparing the coefficients of t^2 we obtain $r_1 + r_2 \leqslant r_0$.

Let us write $a_n = r_1 - t_n$, $b_n = r_2 - t_n$ and $c_n = r_0 + t_n$. Then $p(t_n) = \frac{1}{6}ka_nb_nc_n$, $p'(t_n) = \frac{1}{6}k(a_nb_n - a_nc_n - b_nc_n)$ and $p''(t_n) = \frac{1}{6}k(2c_n - 2a_n - 2b_n)$. Thus, taking into account (12), we have

$$\begin{aligned} a_{n+1} = r_1 - t_{n+1} &= r_1 - t_n + \left(1 + \frac{L_p(t_n)}{2(1 - L_p(t_n))}\right)\frac{p(t_n)}{p'(t_n)} \\ &= \left(\frac{k}{6}\right)^3 \frac{a_n^3}{p'(t_n)^3(1 - L_p(t_n))}\left[a_n(b_n - c_n)(b_n^2 + c_n^2) - b_n^2c_n^2\right]. \end{aligned}$$

In a similar way, we deduce

$$b_{n+1} = \left(\frac{k}{6}\right)^3 \frac{b_n^3}{p'(t_n)^3(1 - L_p(t_n))}\left[b_n(a_n - c_n)(a_n^2 + c_n^2) - a_n^2c_n^2\right].$$

Consequently,

$$\frac{a_{n+1}}{b_{n+1}} = \frac{a_n^3}{b_n^3} H(t_n),$$

where

$$H(t_n) = \frac{b_n^2 c_n^2 + a_n(c_n - b_n)(b_n^2 + c_n^2)}{a_n^2 c_n^2 + b_n(c_n - a_n)(a_n^2 + c_n^2)}.$$

Notice that

$$H(t_n) = \frac{b_n^3 a_{n+1}}{a_n^3 b_{n+1}} = \frac{(r_1 - G(t_n))(r_2 - t_n)^3}{(r_2 - G(t_n))(r_1 - t_n)^3},$$

with G defined in (5). As $G(r_1) = r_1$, $G'(r_1) = G''(r_1) = 0$, we have for t close to r_1

$$H(t) \sim \frac{G'''(r_1)}{6}(r_2 - r_1)^2 = -\frac{f'''(r_1)}{6f'(r_1)}(r_2 - r_1)^2 = \frac{r_2 - r_1}{r_0 + r_1}.$$

Since $t_n \to r_1$ when $n \to \infty$, we obtain

$$\frac{a_n}{b_n} \sim \left(\frac{a_{n-1}}{b_{n-1}}\right)^3 \lambda \sim \cdots \sim \left(\frac{a_0}{b_0}\right)^{3^n} \lambda^{(3^n-1)/2} = \left(\sqrt{\lambda}\frac{r_1}{r_2}\right)^{3^n} \frac{1}{\sqrt{\lambda}}.$$

Then, $r_1 - t_n \sim (r_2 - r_1 + r_1 - t_n)(\theta^{3^n}/\sqrt{\lambda})$ and the first part follows.

If $r_1 = r_2$, we can write $p(t) = \frac{1}{6}k(r_1 - t)^2(r_0 + t)$, where $2r_1 \leqslant r_0$. With the same technique and notations, we obtain

$$a_{n+1} = a_n \frac{a_n^3 - c_n a_n^2 - c_n^3}{(a_n - 2c_n)(a_n^2 + 2c_n^2)} = a_n \frac{(a_n/c_n)^3 - (a_n/c_n)^2 - 1}{((a_n/c_n) - 2)\Big((a_n/c_n)^2 + 2\Big)}.$$

Taking into account that $a_n/c_n \leqslant \frac{1}{2}$ and the function

$$f(x) = \frac{1 + x^2 - x^3}{(x^2 + 2)(2 - x)}$$

increases in $[0, 1/2]$, we deduce $a_n/4 \leqslant a_{n+1} \leqslant a_n/3$ and the result follows. $\square$

Theorem 3.5. *Let us assume* (i)–(iv) *and in addition*

$$\overline{B} = \overline{B(x_1, r_1 - t_1)} = \{x \in X; \|x - x_1\| \leqslant r_1 - t_1\} \subseteq \Omega_0.$$

Then the iterative method defined by (10) *is well-defined*, $x_n \in B$ (*interior of* $\overline{B}$) *for* $n \geqslant 1$ *and the sequence* $\{x_n\}$ *is convergent to* x^*, *solution of* (9). *If* $r_1 < r_2$ *we have* *a third order method for $k > 0$ and a fouth order method for $k = 0$. If $r_1 = r_2$ we only can guarantee linear convergence. The solution, x^*, is unique in*

$$\widetilde{B} = B(x_0, r_2) \cap \Omega_0 \quad \textit{if} \ r_1 < r_2,$$
$$\widetilde{B} = \overline{B(x_0, r_2)} \cap \Omega_0 \quad \textit{if} \ r_1 = r_2.$$

Proof. The first part follows inmediately from Lemma 3.3 and because of the sequence $\{t_n\}$ defined by (12) is convergent. That (10) is a third order method for $k > 0$ follows from Lemma 3.2 and Theorem 2.5. To prove that (10) is a fourth order method for $k = 0$, we use Lemmas 3.2, 3.3 and Theorem 2.6 to obtain for $1 \leqslant \gamma < 4$

$$\frac{\|x^* - x_{n+1}\|}{\|x^* - x_n\|^\gamma} \leqslant \frac{r_1 - t_{n+1}}{(r_1 - t_n)^\gamma} = \frac{1}{(r_1 - r_2)^{\gamma-1}} \frac{(1 - \theta^{4^n})^\gamma}{1 - \theta^{4^{n+1}}} \theta^{(4-\gamma)4^n} \to 0.$$

However, taking into account Theorem 2.6,

$$0 < \lim_{n\to\infty} \frac{\|x^* - x_{n+1}\|}{\|x^* - x_n\|^4} \leqslant \frac{1}{(r_2 - r_1)^3} < \infty.$$

For the uniqueness, the proof of Yamamoto [18] also holds. $\square$

Now we center our study in error estimates for (10). As an application of Gragg and Tapia's techniques [9], Yamamoto [18] established some error estimates for Halley's method (or method of tangent hyperbolas). In the same way, the following results hold.

Theorem 3.6. *Let τ_n^* and σ_n^* be the smallest positive root and the unique positive root of the equations*

$$\phi_n(t) \equiv k_n t^3 - t + \delta_n = 0,$$
$$\psi_n(t) \equiv k_n t^3 + t - \delta_n = 0,$$

where

$$k_n = \frac{r_1 - t_{n+1}}{(r_1 - t_n)^3}, \quad \delta_n = \|x_{n+1} - x_n\| > 0.$$

Then the next estimates hold

$$\sigma_n^* \leqslant \|x^* - x_n\| \leqslant \tau_n^* \leqslant r_1 - t_n.$$

$$\|x^* - x_{n+1}\| \leqslant \tau_n^* - \delta_n.$$

Proof. By Lemma 3.2 and using the same proof as in Yamamoto [18], we can derive the result. $\square$

Corollary 3.7. *With the previous notations, the following error estimates hold*:

$$0.89\delta_n \leqslant \|x^* - x_n\| \leqslant 1.5\delta_n,$$

$$\|x^* - x_{n+1}\| \leqslant 0.5\delta_n.$$

4. Examples

We give two examples of application of Super-Halley method to a system of non-linear equations and a non-linear integral equation.

Example 1. Now, let us consider the systems of non-linear equations $F(x,y) = 0$, where $F(x,y) = (x^2 - y - 2, y^3 - x^2 + y + 1)$.

Newton sequence to solve this system can be written as follows:

$$(x_{n+1}, y_{n+1}) = (x_n, y_n) - \Gamma(x_n, y_n)F(x_n, y_n),$$

where

$$\Gamma(x_n, y_n) = [F'(x_n, y_n)]^{-1} = \frac{1}{6x_n y_n^2}\begin{pmatrix} 3y_n^2 + 1 & 1 \\ 2x_n & 2x_n \end{pmatrix}.$$

In this case, we have

$$x_{n+1} = \frac{3x_n^2 y_n^2 + 2y_n^3 + 6y_n^2 + 1}{6x_n y_n^2},$$

$$y_{n+1} = \frac{2y_n^3 + 1}{3y_n^2}.$$

The linear operator $L_F(x,y)$ defined in (10), is given by the following matrix:

$$L_F(x,y) = \frac{1}{6x^2y^3}\begin{pmatrix} y(3x^2y^2 - 2y^3 - 6y^2 - 1) & 2x(y^3 - 1) \\ 0 & 4x^2(y^3 - 1) \end{pmatrix}.$$

Super-Halley method can be written

$$(w_{n+1}, z_{n+1}) = (w_n, z_n) - \frac{1}{2}\left[I + [I - L_F(w_n, z_n)]^{-1}\right]\Gamma(w_n, z_n)F(w_n, z_n),$$

that is,

$$w_{n+1} = w_n - \frac{(9w_n^2z_n^2 + 2z_n^3 + 6z_n^2 + 1)(3w_n^2z_n^2 - 2z_n^3 - 6z_n^2 - 1)}{12w_n z_n^2(3w_n^2z_n^2 + 2z_n^3 + 6z_n^2 + 1)} - \frac{(z_n^3 - 1)^2}{6w_n z_n^5},$$

$$z_{n+1} = z_n - \frac{(2z_n^3 + 1)(z_n^3 - 1)}{3z_n^2(z_n^3 + 2)}.$$

Starting at $(x_0, y_0) = (w_0, z_0) = (6, 3)$ we obtain the iterations given in Tables 1 and 2.

In the following example we use Theorem 3.5 to show the existence and uniqueness of solution for an integral equation. The results that we obtain are compared with the ones obtained by using other third order iterative processes (see Refs. [3,4,7]).

Example 2. Let us consider the espace $X = C[0, 1]$ of continuous functions on the interval $[0, 1]$, equipped with the max-norm,

$$\|x\| = \max_{s\in[0,1]} |x(s)|.$$

Consider the equation $F(x) = 0$, where

$$F(x)(s) = x(s) - s + \frac{1}{2}\int_0^1 s\cos(x(t))\,\mathrm{d}t, \quad x \in C[0, 1],\ s \in [0, 1].$$

Using the above notation and taking as starting-point the function $x_0 = x_0(s) = s$, we obtain

Table 1
Newton's method

n	x_n	y_n
0	6.000000000000000000	3.000000000000000000
1	3.336419753086419753	2.037037037037037037
2	2.183486059447010306	1.438355269870421386
3	1.806201995430546569	1.120022383336509062
4	1.737006139001477947	1.012402260381869450
5	1.732101431071650477	1.000151311360582637
6	1.732050814916382536	1.000000022890509658
7	1.732050807568877460	1.000000000000000524
8	1.732050807568877294	1.000000000000000000
9	1.732050807568877294	1.000000000000000000

Table 2
Super-Halley method

n	w_n	z_n
0	6.000000000000000000	3.000000000000000000
1	2.195930445526441461	1.173690932311621967
2	1.726757444904059338	0.998981852656109923
3	1.732050203990691682	1.000000000352888045
4	1.732050807568877294	1.000000000000000000
5	1.732050807568877294	1.000000000000000000

$$F'(z)x(s) = x(s) - \frac{s}{2}\int_0^1 x(t)\sin(z(t))\mathrm{d}t,$$

$$[F'(z)]^{-1}y(s) = y(s) + \frac{s}{2\Phi_z}\int_0^1 y(t)\sin(z(t))\mathrm{d}t,$$

where

$$\Phi_z = 1 - \frac{1}{2}\int_0^1 t\sin(z(t))\mathrm{d}t,$$

$$F''(z)xy(s) = -\frac{s}{2}\int_0^1 x(t)y(t)\cos(z(t))\mathrm{d}t,$$

for $x, y, z \in C[0,1]$ and $s \in [0,1]$. So we deduce

$$a = b = \frac{\sin 1}{2 - \sin 1 + \cos 1}, \quad k = \frac{1}{2 - \sin 1 + \cos 1}.$$

In this case the majorizing polynomial is

$$\begin{aligned} p(t) &= \frac{k}{6}t^3 - \frac{b}{2}t^2 - t + a \\ &= \frac{1}{6(2 - \sin 1 + \cos 1)}\left[t^3 + 3(\sin 1)t^2 - 6(2 - \sin 1 + \cos 1)t + 6\sin 1\right] \end{aligned}$$

which has two positive roots

$$r_1 = 0.6095694860276291, \quad r_2 = 1.70990829134757.$$

Consequently, Theorem 3.5 guarantees that $F(x) = 0$ has a root in $\overline{B(x_0, r_1)}$ and this is the only root in $B(x_0, r_2)$ (we have written $B(x_0, r) = \{x \in X; \|x - x_0\| \leqslant r\}$ and $\overline{B(x_0, r)}$ the corresponding closed ball).

Let $\{s_n\}$, $\{t_n\}$ and $\{u_n\}$ be the sequences obtained by applying Chebyshev, Halley and Super-Halley methods to solve the equation $p(t) = 0$. Each one allows us to give an error bound for the corresponding sequence in Banach spaces. These error bounds are shown in Table 3.

Table 3
Error bounds

n	$r_1 - s_n$	$r_1 - t_n$	$r_1 - u_n$
0	0.6095694860276291	0.6095694860276291	0.6095694860276291
1	0.0534834955243040	0.0495130055348865	0.0349873303274992
2	0.0001520166774545	0.0000984825547302	0.0000560164474543
3	0.0000000000042804	0.0000000000009129	0.0000000000001218
4	0.000000000000000	0.000000000000000	0.000000000000000

Notice that the best error bounds are attained with Super-Halley method. For this equation and using Halley method, Döring [7] gave the bound

$$\|x^* - x_2\| \leqslant 0.000825.$$

Later, Candela and Marquina [3,4], gave the bounds

$$\|x^* - x_2\| \leqslant 0.00037022683427694$$

and

$$\|x^* - x_2\| \leqslant 0.00014987029635502$$

with Chebyshev and Halley method, respectively.

References

[1] M. Altman, Concerning the method of tangent hyperbolas for operator equations, Bull. Acad. Pol. Sci., Serie Sci. Math., Ast. et Phys. 9 (1961) 633–637.
[2] R. Campbell, Les Integrales Euleriennes et Leurs Aplications, Dunod, Paris, 1966.
[3] V. Candela, A. Marquina, Recurrence relations for rational cubic methods I: The Halley method, Comput. 44 (1990) 169–184.
[4] V. Candela, A. Marquina, Recurrence relations for rational cubic methods II: The Chebyshev method, Comput. 45 (1990) 355–367.
[5] D. Chen, I.K. Argyros, Q.S. Qian, A local convergence theorem for the Super-Halley method in a Banach space, App. Math. Lett. 7 (5) (1994) 49–52.
[6] Z. Ciesielski, Some propierties of convex functions of higher orders, Annales Polonici Mathematici 7 (1959) 1–7.
[7] B. Döring, Einige sätze über das verfahren der tangierenden hyperbeln in Banach-Räumen, Aplikace Mat. 15 (1970) 418–464.
[8] W. Gander, On Halley's iteration method, Amer. Math. Monthly 92 (1985) 131–134.
[9] W.B. Gragg, R.A. Tapia, Optimal error bounds for the Newton–Kantorovich theorem, SIAM J. Numer. Anal. 11 (1974) 10–13.
[10] J.M. Gutiérrez, M.A. Hernández, M.A. Salanova, Accesibility of solutions by Newton's method, Int. J. Comput. Math. 57 (1995) 239–247.
[11] J.M. Gutiérrez, M.A. Hernández, M.A. Salanova, Resolution of quadratic equations in Banach spaces, Numer. Funct. Anal. Opt. 17 (1/2) (1996) 113–121.
[12] J.M. Gutiérrez, M.A. Hernández, A family of Chebyshev–Halley type methods in Banach spaces, Bull. Austral. Math. Soc. 55 (1997) 113–130.
[13] M.A. Hernández, Newton–Raphson's method and convexity, Zb. Rad. Prirod.-Mat. Fak. Ser. Mat. 22 (1) (1993) 159–166.
[14] M.A. Hernández, M.A. Salanova, A family of Newton type iterative processes, Int. J. Comput. Math. 51 (1994) 205–214.
[15] L.V. Kantorovich, G.P. Akilov, Functional Analysis, Pergamon Press, Oxford, 1982.
[16] A.M. Ostrowski, Solution of Equations in Euclidean and Banach Spaces, Academic Press, New York, 1943.
[17] J.F. Traub, Iterative Methods for Solution of Equations, Prentice-Hall, Englewood Cliffs, NJ, 1964.
[18] T. Yamamoto, On the method of tangent hyperbolas in Banach spaces, J. Comput. Appl. Math. 21 (1988) 75–86.