ELSEVIER Applied Mathematics and Computation 103 (1999) 27–49

APPLIED MATHEMATICS AND COMPUTATION

Indices of convexity and concavity. Application to Halley method

M.A. Hernández *, M.A. Salanova

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

Abstract

We define an index to measure the convexity of a convex function f at each point. We use this index to establish conditions on the convergence of the Halley method in the complex plane and in Banach spaces. © 1999 Elsevier Science Inc. All rights reserved.

AMS Classification: 26A51; 47H10; 65J15

Keywords: Convex function; Nonlinear equations in Banach spaces; Third-order method; Halley method; Majorizing sequences

1. Introduction

The study of the concavity and convexity of a real function is an old problem studied by the mathematicians. It is perfectly known 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 [4] are interesting from the theoretical standpoint, but their practical application is too difficult.

Another measure of the convexity is suggested by Bohr–Mollerup's Theorem [3]. In this result the concept of log-convex function appears that is, a function whose logarithm is a convex function. The degree of logarithmic convexity, introduced in Ref. [7], is a measure of this kind of convexity. Let $\varphi : \Omega \subseteq \mathbb{R} \to \mathbb{R}$ be a concave, twice differentiable function on an interval Ω and $f \in C^{(m)}(V)$, the class of functions with m continuous derivatives, $m \geqslant 2$.

* Corresponding author. E-mail: mahernan@dmc.unirioja.es.

PII: S0096-3003(98)10047-4

Assume that f is a convex function on a neighbourhood V of x_0, and $y_0 = f(x_0) \in \Omega$. It is known that the curvature of a function f at x_0 [10] is given by the expression

$$K(f)(x_0) = \frac{f''(x_0)}{\left[1 + f'(x_0)^2\right]^{3/2}}.$$

If $\varphi'(y_0) = 1$ then $K(\varphi \circ f)(x_0) < K(f)(x_0)$, since

$$(\varphi \circ f)'(x_0) = \varphi'(y_0)f'(x_0), \tag{1}$$

$$(\varphi \circ f)''(x_0) = \varphi''(y_0)f'(x_0)^2 + \varphi'(y_0)f''(x_0). \tag{2}$$

So, we can say that when we compose a convex function with a concave one, we obtain a function "less convex". The index constructed in this paper measures the number of times that f must be composed with the operator φ to get a concave function.

Next, we make a similar study considering the action of a convex operator on the set of concave functions in $C^{(m)}(V)$, $m \geqslant 2$.

As an application of this index, we obtain convergence results of the Halley method in the complex plane and Banach spaces.

2. φ-convex functions

Before defining φ-convex functions we give the following definitions.

Definition 2.1. Let $f \in C^{(m)}(V)$, $m \geqslant 2$. We say that

(A) f is a strictly convex function at x_0 if $f''(x_0) > 0$.

(B) f is a non-strictly convex function at x_0 if there exists an even number $k \in \mathbb{N}$ such that $f''(x_0) = \cdots = f^{(k-1)}(x_0) = 0$ and $f^{(k)}(x_0) > 0$.

(C) f has a strong minimum at x_0 if $f'(x_0) = 0$ and $f''(x_0) > 0$.

(D) f has a non-strong minimum at x_0 if $f'(x_0) = f''(x_0) = \cdots = f^{(k-1)}(x_0) = 0$ and $f^{(k)}(x_0) > 0$ where k is an even number.

Analogously to Definition 2.1 we can give the concepts of strictly or non-strictly concave function and a strong or non-strong maximum.

Definition 2.2. Let φ be a concave function. With the above notation, f is a φ-convex function at x_0 if $\varphi \circ f$ is a convex function at x_0.

Next we give an analytic characterization of the φ-convex functions. Notice that if φ is a decreasing function in a neighbourhood of y_0, $\varphi'(y_0) \leqslant 0$, then $\varphi \circ f$ will be always a concave function. The same result is obtained if φ is a non-strictly concave function at $y_0 = f(x_0)$, because $\varphi''(y_0) = 0$. Therefore, consider φ to be an increasing and strictly concave function in a neighbourhood W of y_0 ($\varphi'(y) > 0$ and $\varphi''(y) < 0$ in W).

Taking into account (A), (B) and the different types of convexity, we get the following result.

Theorem 2.3. *Let x_0 be such that $f'(x_0) \neq 0$.*

(i) *If f is a strictly convex function at x_0, then we have*:

(a) *f is a strictly φ-*convex function at x_0 *if and only if*

$$-\frac{\varphi'(y_0)f''(x_0)}{\varphi''(y_0)f'(x_0)^2} > 1.$$

(b) *If f is a non-strictly φ-*convex function at x_0, *then*

$$-\frac{\varphi'(y_0)f''(x_0)}{\varphi''(y_0)f'(x_0)^2} = 1. \tag{3}$$

(ii) *If f is a non-strictly convex function at x_0, then f is not φ-convex.*

Proof. (a) and (b) follow from (A) and (B). As $f''(x_0) = 0$, we obtain $(\varphi \circ f)''(x_0) = \varphi''(y_0)f'(x_0)^2 < 0$, thus (ii) holds. □

As we can see in the following example, the condition (3) is not sufficient for the function f to be non-strictly φ-convex.

Example 1. Consider the functions

$$f_1(x) = \exp\left(\frac{x^4}{4} + x\right), \quad f_2(x) = \exp\left(x - \frac{x^4}{4}\right),$$

$$f_3(x) = \exp\left(\frac{x^3}{3} + x\right) \quad \text{and} \quad \varphi(x) = \log x.$$

These functions satisfy the condition (3) at x_0. Denote $F_i = \varphi \circ f_i$, for $i = 1, 2, 3$, respectively. Fig. 1 shows that $F_1''(0) > 0$, $F_2''(0) < 0$ and F_3 has an inflexion point at x_0.

If x_0 is a minimum for f then

$$\lim_{x \to x_0} -\frac{\varphi'(f(x))f''(x)}{\varphi''(f(x))f'(x)^2} = +\infty.$$

f has a strong minimum at x_0 if and only if $\varphi \circ f$ has a strong minimum at x_0, and f has a non-strong minimum at x_0 if and only if $\varphi \circ f$ has a non-strong minimum at x_0.

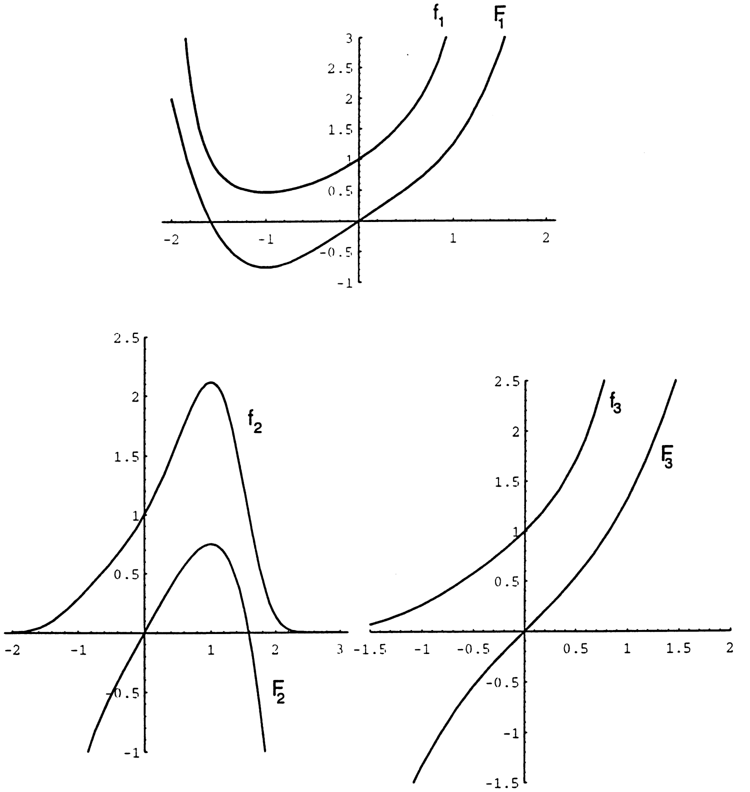

Fig. 1. Example 1.

Now we define an index to measure of convexity of a function at each point. It is defined as the number of times that we need to apply the concave operator φ to obtain a concave function. For this, denote $N[F](x) = F(x) - F(x_0) + f(x_0)$ and define the following sequences:

$$\begin{array}{ll} F_1(x) = \varphi \circ f(x), & G_1(x) = N[\varphi \circ f](x), \\ F_2(x) = \varphi \circ G_1(x), & G_2(x) = N[F_2](x), \\ \vdots & \vdots \\ F_n(x) = \varphi \circ G_{n-1}(x), & G_n(x) = N[F_n](x). \end{array}$$

Notice that $F_k(x_0) = \varphi \circ f(x_0)$ and $G_k(x_0) = f(x_0)$ for each $n \in \mathbb{N}$. To measure the resistance of f to be "concaved" by the operator φ, we study the convexity of the functions F_n at x_0. So we introduce the concept of $n - \varphi$-convex function, which is a generalization of the φ-convex function that we have analysed before.

Definition 2.4. f is an $n-\varphi$-convex function at x_0 if F_n is convex at x_0. If f is $n-\varphi$-convex function at x_0 for each $n \in \mathbb{N}$, will say that f is indefinitely φ-convex.

Next, we give an analytic characterization of these concepts and obtain an expression for the successive derivatives of the function F_n, for each $n \in \mathbb{N}$. The proofs follow inductively.

Lemma 2.5. *For each $n \in \mathbb{N}$, we have*

(i) $F_n'(x_0) = \varphi'(y_0)^n f'(x_0)$.

(ii) $F_n''(x_0) = \varphi''(y_0)\varphi'(y_0)^{n-1}[\sum_{k=0}^{n-1}\varphi'(y_0)^k]f'(x_0)^2 + \varphi'(y_0)^n f''(x_0)$.

Let us observe that if f is $n-\varphi$-convex function at x_0, then f is $k-\varphi$-convex at x_0 for each $1 \leqslant k \leqslant n$.

If F_n is non-strictly convex, we can compute the derivatives of F_n from (A) and (B) by recurrence. We have

$$F_n^{(k+1)}(x) = \sum_{t=0}^{k}\binom{k}{t}[\varphi'(G_{n-1}(x))]^{(t)}F_{n-1}^{(k-t+1)}(x). \tag{4}$$

Now we characterize the concept of $n-\varphi$-convex function at x_0. We start this study for x_0 being a minimum of f. When x_0 is not a critical point of f, there are two situations: f is strictly convex at x_0 or f is non-strictly convex.

Theorem 2.6. *With the above notation, we have*

(I) *Let x_0 be a minimum for f.*

(i) *If x_0 is a minimum for f, then x_0 is a minimum of the same type for F_n, for each $n \in \mathbb{N}$.*

(ii) *x_0 is a minimum for f if and only if f is indefinitely φ-convex at x_0.*

(II) *Let f be a strictly convex function at x_0, with $f'(x_0) \neq 0$, then*

(i) *If*

$$\sum_{k=0}^{n}\varphi'(y_0)^k \underset{(a)}{>} -\left(\varphi'(y_0)f''(x_0)/\varphi''(y_0)f'(x_0)^2\right) \underset{(b)}{>} \sum_{k=0}^{n-1}\varphi'(y_0)^k,$$

*then F_n is strictly convex at x_0 and F_{n+1} is strictly concave at x_0. Therefore, f is $n-\varphi$-*convex *and is not $(n+1)-\varphi$-convex.*

(ii) *If $-(\varphi'(y_0)f''(x_0)/\varphi''(y_0)f'(x_0)^2) = \sum_{k=0}^{n-1}\varphi'(y_0)^k$, then f is $(n-1)-\varphi$-convex at x_0.*

Moreover, F_{n-1} is strictly convex at x_0, and F_{n+1} is strictly concave at x_0.

Proof. The results of I follow using induction and taking into account (4). To prove II, by Lemma 2.5, (ii), we have

$$\frac{F_n''(x_0)}{\varphi''(y_0)\varphi'(y_0)^{n-1}f'(x_0)^2} = \sum_{k=0}^{n-1}\varphi'(y_0)^k + \frac{\varphi'(y_0)f''(x_0)}{\varphi''(y_0)f'(x_0)^2} < 0$$

from (b), then it follows $F_n''(x_0) > 0$ so, F_n is strictly convex.

From (a) and in a similar way we can derive that F_{n+1} is a strictly concave function. Then (i) holds.

To prove (ii), from the hypothesis we have that

$$F_{n-1}(x_0) = -\varphi''(y_0)\varphi'(y_0)^{2n-3}f'(x_0)^2 > 0$$

and as $\varphi''(y_0) < 0$ we obtain that F_{n-1} is a strictly concave function. □

From this last result, it is clear that the quantity $-(\varphi'(y_0)f''(x_0)/\varphi''(y_0)f'(x_0)^2)$ and its relation with $s = \sum_{k=0}^{n-1}\varphi'(y_0)^k$, defines the φ-convexity degree of f at x_0 as the natural number n such that F_n is a convex function and F_{n+1} is not a convex function. Taking into account the differents values of s it seems natural to normalize the operator φ in the form $\varphi'(y_0) = 1$. This normalization does not alter the geometric properties of φ, that is, φ is increasing and a concave function in a neighbourhood W of y_0.

The geometric interpretation of the φ-convexity degree in relation to the curvature becomes clear. Thus, given the function φ, we consider $\Psi(y) = \varphi(y)/\varphi'(y_0)$ and we repeat the previous process with Ψ. Computing, as before, the sequences

$$H_n(x) = \Psi \circ K_{n-1}(x), \quad \text{where } K_n(x) = N[H_n](x),$$

$$H_1(x) = \Psi \circ f(x) \quad \text{and} \quad K_1(x) = N[H_1](x),$$

we derive the following result, similar to the last theorem.

Corollary 2.7. *Consider f to be a strictly convex function at x_0, and $f'(x_0) \neq 0$, then*

(i) *If it is verified that*

$$n+1 > \left[\frac{-1}{\Psi''(y_0)}\right]\frac{f''(x_0)}{f'(x_0)^2} > n,$$

then, H_n is strictly convex at x_0 and H_{n+1} is strictly concave at x_0. Therefore, f is $n-\Psi$-convex and is not $(n+1)-\Psi$-convex.

(ii) *If $[-1/\Psi''(y_0)](f''(x_0)/f'(x_0)^2) = n$, then f is $(n-1)-\Psi$-convex at x_0 and is not $(n+1)-\Psi$-convex, being H_{n-1} strictly convex and H_{n+1} strictly concave at x_0.* □

It is immediate to prove the next relation between the curvatures of the functions G_i: $K(f)(x_0) > K(G_1)(x_0) > K(G_2)(x_0) > \cdots > K(G_n)(x_0) \geqslant 0 > K(G_{n+1})(x_0) > \cdots$

Example 2. Consider the family of convex functions $f_\alpha(x) = \frac{1}{2}\exp(x^\alpha - 1)$ for $\alpha \in \mathbb{R} - \{0, 1\}$ and $x_0 = 1$. If we take $\varphi(y) = \log y$ then $y_0 = f(x_0) = \frac{1}{2}$ and $\Psi(y) = \log y/2$. In this situation we obtain that

$$\left[\frac{-1}{\Psi''(y_0)}\right] \frac{f''_\alpha(x_0)}{f'_\alpha(x_0)^2} = 2 - \frac{1}{\alpha}.$$

We obtain different values for this quantity according to the values of α. From these quantities we know which in terms of the sequence $\{H_n\}$ are convex functions. For example, if we consider $\alpha = -2$, then we obtain that H_1 and H_2 are convex and H_3 is a concave function (see Fig. 2),

$$H_1(x) = \frac{1}{2}\left[\frac{1}{x^2} - 1 - \log 2\right], \quad K_1(x) = \frac{1}{2x^2},$$

$$H_2(x) = -\frac{\log 2}{2} - \log x, \quad K_2(x) = \frac{1}{2} - \log x$$

$$H_3(x) = \frac{1}{2}\log\left[\frac{1}{2} - \log x\right].$$

Moreover, it is easy to prove that (see Fig. 3)

$$K(f)(1) > K(H_1)(1) > K(H_2)(1) > 0 > K(H_3)(1).$$

Thus, we can define the following.

Definition 2.8. We call φ-convexity degree of f at x_0 to the real number given by the expression

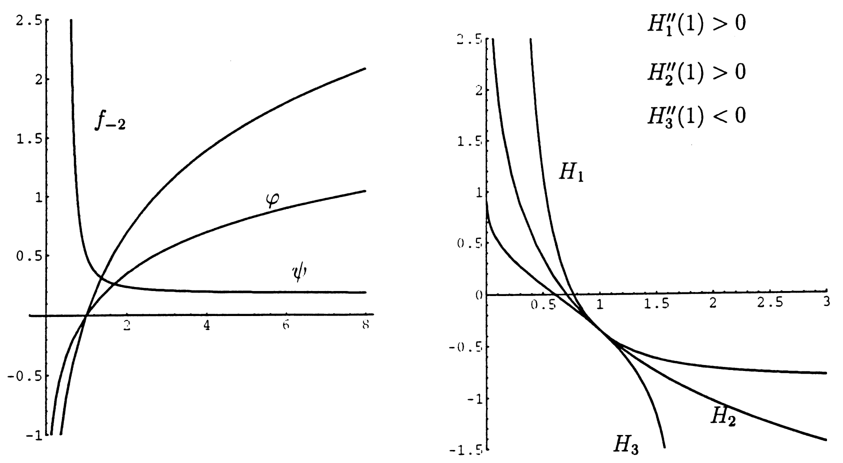

Fig. 2. Example 2.

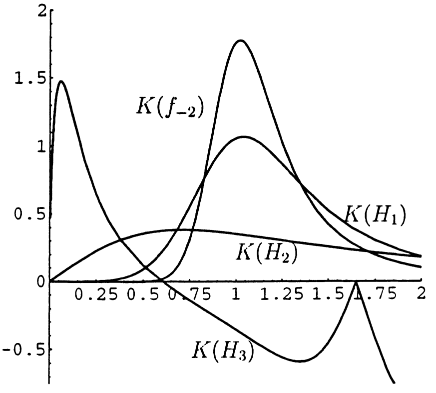

Fig. 3. Example 2.

$$U_\varphi(f)(x_0) = \frac{-1}{\Psi''(y_0)} \frac{f''(x_0)}{f'(x_0)^2}, \tag{5}$$

if x_0 is not a minimum for f. In this case $U_\varphi(f)(x_0) = +\infty$.

Making some calculations in Eq. (5), we have

$$U_\varphi(f)(x_0) = -\frac{\varphi'(y_0)}{\varphi''(y_0)} \frac{f''(x_0)}{f'(x_0)^2}. \tag{6}$$

Remarks. This quantity give us the measure of the convexity before normalization. The above normalization does not affect our study, just simplifies it.

The next example shows that $U_\varphi(f)(x_0) \in [0, +\infty)$ and it is not bounded.

Example 3. Let the family of functions $f_\alpha(x) = -(1/\alpha)\log \alpha x$ in [0,1], with α a positive real number, and we consider the family of functions $\varphi_\beta(x) = 2 + x - \beta x^2/2$, where β is a positive real number. For $x_0 = 1/\alpha$ we obtain $y_0 = f(x_0) = 0$ and $U_{\varphi\beta}(f)(1/\alpha) = \alpha/\beta$.

From Definition 2.7 we can derive the following properties of the φ-convexity degree.

Corollary 2.9. *With the above notation we have*

(i) *For each real number* r,

(a) *If we denote* $k = \Psi''(y_0)/\Psi''(r + y_0)$, *then* $U_\varphi(r + f)(x_0) = kU_\varphi(f)(x_0)$.

(b) *If* $r > 0$ *and we denote* $k = \Psi''(y_0)/\Psi''(ry_0)$, *then* $U_\varphi(rf)(x_0) = (k/r)U_\varphi(f)(x_0)$.

(ii) *If* $1/f$ *is convex in* (*a*), (*b*), *then*

$$U_\varphi\left(\frac{1}{f}\right)(x_0) = -\frac{y_0}{\Psi''\left(\frac{1}{y_0}\right)}[2 + \Psi''(y_0)y_0U_\varphi(f)(x_0)].$$

(iii) $U_\varphi(f)(x_0) \geqslant 0$. *Besides,* $U_\varphi(f)(x_0) = 0$ *in W if and only if f is affine in W, that is* $F(x) = cx + d$ *with,* $c, d \in \mathbb{R}$.
(iv) *Let g be a function satisfying the same conditions than f and such that* $g \circ f$ *is defined, then*

$$U_\varphi(g \circ f)(x_0) = U_\varphi(g)(f(x_0)) + \frac{1}{g'(f(x_0))}\frac{\Psi''(f(x_0))}{g(\Psi''(f(x_0))}U_\varphi(f)(x_0). \qquad \square$$

Now, we obtain an index of convexity concerning all the concave operators. For this, we consider a convex operator φ, as before, and normalize φ by the transformation

$$\varphi^*(y) = \frac{\varphi(y) - [\varphi'(y_0) + \varphi''(y_0)]y}{-\varphi''(y_0)}$$

so, we obtain a new operator φ^* with the same geometric characteristics than φ satisfying $(\varphi^*)'(y_0) = 1$ and $(\varphi^*)''(y_0) = -1$. Therefore, from Eq. (6) we obtain

$$U_{\varphi^*}(f)(x_0) = \frac{f''(x_0)}{f'(x_0)^2}.$$

So we give the following definition.

Definition 2.10. The convexity degree of f *at* x_0 is the real positive number given by the expression

$$U(f)(x_0) = \frac{f''(x_0)}{f'(x_0)^2}, \tag{7}$$

if x_0 is not a minimum for f. In this case $U(f)(x_0) = +\infty$.

It is easy to prove the following properties of this convexity degree.

Corollary 2.11. *It is satisfied*:
(i) *For each real number r we have*
(a) $U(r + f)(x_0) = U(f)(x_0)$, *that is,* $U(f)$ *is invariant by translations.*
(b) *If* $r > 0$, *then* $U(rf)(x_0) = U(f)(x_0)/r$.
(ii) *If* $1/f$ *is convex in* (*a*), (*b*), *then*

$$U\left(\frac{1}{f}\right)(x_0) = f(x_0)[2 - f(x_0)U(f)(x_0)].$$

(iii) $U(f)(x_0) \geqslant 0$. *Furthermore* $U(f)(x_0) = 0$ *in* (*a*), (*b*) *if and only if* f *is affine, that is,* $f(x) = cx + d$, *with* $c, d \in \mathbb{R}$.

(iv) *If* f *and* g *are two increasing functions, we have*

$$U(f+g)(x_0) \leqslant U(f)(x_0) + U(g)(x_0).$$

(v) *If* f *and* g *are two increasing functions such that* $g \circ f$ *is defined, then*

$$U(g \circ f)(x_0) = U(g)(f(x_0)) + \frac{1}{g'(f(x_0))} U(f)(x_0). \qquad \square$$

This results tell us that $U(f)$ is a good measure of convexity at each point [12].

If we consider a concave function f in a neighbourhood of the point x_0, and a convex operator φ in a neighbourhood of $y_0 = f(x_0)$, we can make a similar process to obtain the φ-concavity degree of f at x_0. It is given by the expression

$$\cap_\varphi (f)(x_0) = -\frac{\varphi'(y_0)}{\varphi''(y_0)} \frac{f''(x_0)}{f'(x_0)^2}. \tag{8}$$

As in the case of the φ-convexity degree if we normalize the operator φ, making some calculations in Eq. (8) we obtain an index to measure the concavity of f at x_0. This one is given by the real positive number

$$\cap (f)(x_0) = -\frac{f''(x_0)}{f'(x_0)^2}, \tag{9}$$

if x_0 is not a maximum. In this case $\cap(f)(x_0) = +\infty$.

Geometrically, the concavity degree of f at x_0 is the convexity degree of $-f$ at x_0. Indeed from Eqs. (7) and (9) we obtain

$$\cap (f)(x_0) = U(-f)(x_0). \tag{10}$$

From Eq. (10), it is clear that the concavity degree of f at x_0 is going to have the same properties as the convexity degree of f at x_0, as we have seen in the previous corollary.

3. Convexity and the Halley method

As a particular case of these measures of convexity, we consider the degree of logarithmic convexity of f, $L_f(x)$. It is obtained taking $\varphi(x) = \log x$ in Eq. (6) and is given by the expression

$$L_f(x) = \frac{f(x) f''(x)}{f'(x)^2}. \tag{11}$$

In the resolution of nonlinear scalar equations

$$f(x) = 0, \tag{12}$$

by iterative process, there exist methods, as the Whittaker and Newton methods, in which the velocity of convergence depends on the convexity or concavity of the function f [11,8]. Besides in Ref. [6], it is proved that all iterative processes with cubical convergence have an expression of the form

$$F(x) = x - \frac{f(x)}{f'(x)} G(x), \quad \text{with} \quad G(x) = H(L_f(x)) + f(x)^2 b(x)$$

where $b(x)$ is a bounded function. In particular, for the Halley method $H(L_f(x)) = 2/(2 - L_f(x))$ and $b(x) = 0$.

Now, we are going to give conditions for the convergence of the Halley method for the solution of nonlinear scalar and complex equations by means of the degree of logarithmic convexity of f and f'.

Moreover we make a practical study of the Halley method where it is proved that we can always apply this method to solve a nonlinear equation according to the convexity of f and f'. Next, we derive a type of Kantorovich conditions [9] for the convergence of the Halley method for a nonlinear complex equation by using majorizing sequences [11]. To finish we realize a study of this method in Banach spaces.

From now on, we consider a real function $f \in C^{(m)}([a,b]), m \geqslant 2$, satisfying the Fourier conditions, i.e., $f(a)f(b) < 0$, $f' \neq 0$ and the sign of f is constant in $[a,b]$. Under these conditions there is only one root t^* of Eq. (12) in $[a,b]$. We can assume, without loss of generality, that f is a convex and strictly increasing function in $[a,b]$, with $f(a)f(b) < 0$, because otherwise it is sufficient to change $f(t)$ by $f(-t), -f(t)$ or $-f(-t)$.

Besides, for each function h, denote

$$M(h) = \max\{h(x) | x \in [a,b]\} \quad \text{and} \quad m(h) = \min\{h(x) | x \in [a,b]\}.$$

Next, we will study the convergence of the sequence $\{t_n\}$ obtained by the Halley method which, starting at t_0, is given by the expression

$$t_n = P(t_{n-1}) = t_{n-1} - \frac{f(t_{n-1})}{f'(t_{n-1})} \frac{2}{2 - L_f(t_{n-1})}. \tag{13}$$

Theorem 3.1. *Suppose that* $L_{f'}(t) \leqslant \frac{3}{2}, L_f(t) < 2$ *in* $[a,b]$ *and* $t_0 \in [a,b]$.

(i) *If* $f(t_0) > 0$, *then the sequence* $\{t_n\}$ *is decreasing to* t^*.

(ii) *If* $f(t_0) < 0$, *then the sequence* $\{t_n\}$ *is increasing to* t^*.

Proof. The result follows by induction, taking into account the mean value theorem and the expression of the derivative of the function P

$$P'(t) = \frac{L_f(t)^2}{[2 - L_f(t)]^2} [3 - 2L_{f'}(t)]. \tag{14}$$

In both cases the sequence $\{t_n\}$ obtained converges to a limit u, and making $n \to \infty$ in Eq. (13), we deduce $f(u) = 0$, then $u = t^*$. □

Notice that the condition $L_f(t) < 2$ in $[a, b]$ is not very restrictive because $L_f(t^*) = 0$, see Eq. (11), and $L_f(t) \geqslant 0$ in $[t^*, b]$.

Definition 3.2. *Let $\{t_n\}$ be a sequence that converges to t^*. We say that $\{t_n\}$ oscillates about t^*,* if $t_{2(j-1)} \geqslant t_{2j} \geqslant t^* \geqslant t_{2j+1} \geqslant t_{2j-1}$ *or* $t_{2j-1} \geqslant t_{2j+1} \geqslant t^* \geqslant t_{2j} \geqslant t_{2(j-1)}$, *for $j \geqslant 0$.*

In the following theorems the sequence $\{t_n\}$ oscillates about t^*. Then we need a condition to assure $\{t_n\} \subseteq [a, b]$.

Lemma 3.3. *Assume that $|P'(t)| < 1$ and $|L_f(t)| < 1$ in $[a, b]$.*

(i) *Let $t_0, \alpha \in (a, b)$ be such that $f(\alpha) < 0$ and $f(t_0) > 0$. If $2f(t_0)/f'(t_0) \leqslant \alpha - a$, then $\{t_n\} \subseteq [a, b]$.*

(ii) *Let $t_0, \beta \in (a, b)$ be such that $f(\beta) > 0$ and $f(t_0) < 0$. If $2f(t_0)/f'(t_0) \geqslant b - \beta$, then $\{t_n\} \subseteq [a, b]$.*

Proof. (i) As $t_1 = P(t_0) \geqslant t_0 - 2f(t_0)/f'(t_0) \geqslant t_0 - \alpha + a > a$, it is clear that $t_1 \in (a, t_0) \subseteq [a, b]$. Now we distinguish two situations, $t_1 < t^*$ or $t_1 > t^*$.

(a) $t_1 < t^*$ implies $f(t_1) < 0$, then $-1 < L_f(t_1) < 0$, so $t_1 < t_2$. Moreover, we have $|t_2 - t^*| < |t_1 - t^*| < |t_0 - t^*|$ since $|P'(t)| < 1$. Therefore, $t_2 \in (t_1, t_0) \subseteq [a, b]$.

(b) If $t_1 > t^*$ then $t^* < t_1 < t_0$, then $t_1 \in [a, b]$ and

$$t_2 = P(t_1) \geqslant t_1 - \frac{2f(t_1)}{f'(t_1)} \geqslant t_1 - \frac{2f(t_0)}{f'(t_0)} \geqslant t_0 - \alpha + a > a,$$

since f/f' is an increasing function in (t^*, b). Thus we have, $t_2 \in (a, t_1) \subseteq (a, t_0) \subseteq [a, b]$. At this moment there are two cases:

(b.1) $t_k \in (t^*, b)$ for all $k = 1, \ldots, n-1$. By means of an analogous procedure it is proved that $t_n \in (a, t_0)$ and $\{t_n\} \subseteq [a, b]$.

(b.2) If there exists $t_k < t^*$ with $t_1, \ldots, t_{k-1} \in (t^*, b)$, a similar reasoning to (a) proves that $t_{k+1} \in (t_k, t_0) \subseteq [a, b]$.

(ii) follows in the same way. □

From now on, we choose $t_0, \alpha, \beta \in (a, b)$ as in the Lemma 3.3.

Theorem 3.4. *With the above notation, we have*

(i) *If $L_{f'}(t) \in (3/2, 2]$ and $|L_f(t)| < 1$ in $[a, b]$, then the sequence $\{t_n\}$ converges to t^*. Besides this sequence oscillates about t^*.*

(ii) *If* $L_{f'}(t) > 2$ *and* $|L_f(t)| < w$ *in* $[a,b]$, *where*

$$w = \frac{-1 + \sqrt{2M(L_{f'}) - 3}}{M(L_{f'}) - 2},$$

then $\{t_n\}$ *converges to* t^*. *Moreover, this sequence oscillates about* t^*.

Proof. From Eq. (14) as $L_{f'}(t) > 3/2$, it follows that $P'(t) < 0$. So, under the previous hypothesis, we have $|P'(t)| < |3 - 2L_{f'}(t)| < 1$ in $[a,b]$. From Lemma 3.3 we obtain that $\{t_n\} \subseteq [a,b]$.

Now we prove that $\{t_n\}$ is bounded by t^* and oscillates about it. For that, if we consider $t_0 > t^*$, then $t_1 - t^* = P'(\theta_0)(t_0 - t^*)$ with $\theta_0 \in (t^*, t_0)$, so $t_1 \leqslant t^*$ and $|t_1 - t^*| < M(|P'|)|t_0 - t^*|$. Then, by induction, it is easy to prove that $t_{2j} \geqslant t^*, t_{2j+1} \leqslant t^*$ and $|t_{j+1} - t^*| < M(|P'|)^{j+1}|t_0 - t^*|$ for all $j \geqslant 0$. Therefore, as $M(|P'|) < 1$, we obtain that $\{|t_n - t^*|\}$ is a decreasing sequence and converges to zero. So, the result holds.

We can repeat the above proof when $f(t_0) < 0$.

To show (ii), we know that $P'(t) < 0$ and $|P'(t)| < 1$ in $[a,b]$. Besides $|L_f(t)| < w < 1$ in $[a,b]$. Then, by Lemma 3.3, we obtain that $\{t_n\} \subseteq [a,b]$. And by a similar way as in Theorem 3.1 the result holds. □

In practice, we can always apply the Halley method to solve Eq. (12) from the previous theorem. Depending on the values of $M(L_{f'})$ and $m(L_{f'})$, we have the situations of Fig. 4.

(A) $M(L_{f'}) \leqslant \frac{3}{2}$.

(B) $M(L_{f'}) \leqslant 2$ and $m(L_{f'}) > \frac{3}{2}$.

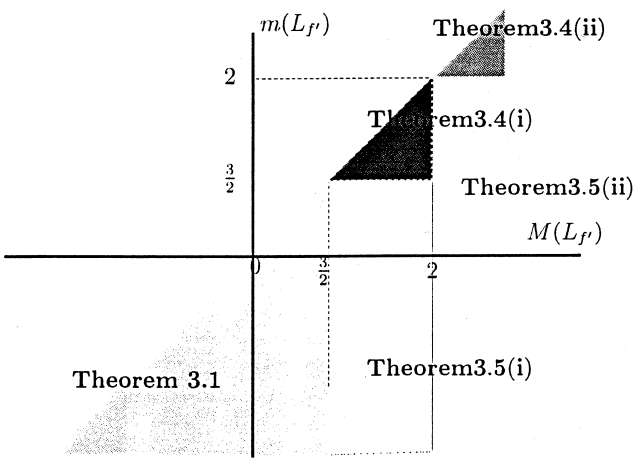

Fig. 4. Convergence results.

(C) $\frac{3}{2} < M(L_{f'}) \leqslant 2$ and $m(L_{f'}) \leqslant \frac{3}{2}$.
(D) $M(L_{f'}) > 2$ and $m(L_{f'}) \leqslant 2$.
(E) $m(L_{f'}) > 2$.

Notice that, in the situations (A), (B), and (E), the convergence of $\{t_n\}$ follows from Theorems 3.1 and 3.3. The cases (C) and (D) are studied in the next result with some conditions for L_f.

Theorem 3.5. *Consider $\alpha, t_0 \in [a,b]$ as in the Lemma 3.3, $K = \max\{|3-2M(L_{f'})|, |3-2m(L_{f'})|\}$, and assume that $M(|L_f)|) < 2/(1+\sqrt{K})$.*

(i) *If $\frac{3}{2} < M(L_{f'}) \leqslant 2, m(L_{f'}) \leqslant \frac{3}{2}$ and $|L_f(t)| < 1$ in $[a,b]$, then $\{t_n\}$ converges to t^*.*

(ii) *If $M(L_{f'}) > 2$ and $m(L_{f'}) \leqslant 2$, then $\{t_n\}$ converges to t^*.*

Proof. In the case (i), from Eq. (14) and $|L_f(t)| < 1$ we derive

$$|P'(t)| < \frac{M(|L_f)|)^2 K}{[2 - M(|L_f)|)]^2},$$

so, $|P'(t)| < 1$ in $[a,b]$ and, by Lemma 3.3, $\{t_n\} \subseteq [a,b]$.

In order to prove (ii), in that situation we have $K > 1$, therefore $M(|L_f|) < 1$. Thus $|L_f(t)| < 1$ in $[a,b]$. Now we have to prove that $|P'(t)| < 1$ in $[a,b]$ i.e. $(K-1)M(|L_f|)^2 + 4M(|L_f|) - 4 < 0$. That follows by the hypothesis that $0 < M(|L_f|) < 2/(1+\sqrt{K})$.

In both cases, (i) and (ii), we can finish the proof in a way similar to Theorem 3.4. □

Notice that if $f(t_0) < 0$, the previous result is also true, with analogous conditions.

So, we can always apply the Halley method to solve Eq. (12), with some restrictions according to the values of $L_f(t)$ in $[a,b]$. But, since $L_f(t^*) = 0$, see Eq. (11), these results can always be applied in a neighbourhood of the root t^*.

Numerical example: To illustrate this convergence study for the Halley method, we consider the equation $f(t) = -8 + e^{t^2-4} = 0$ with $[a,b] = [1,5]$. We have

$$L_f(t) = \frac{(1+2t^2)(-8+e^{t^2-4})}{2t^2\, e^{t^2-4}} \quad \text{and} \quad L_{f'}(t) = \frac{2t^2(3+2t^2)}{(1+2t^2)^2}.$$

As we can show $L_{f'}(t) \leqslant \frac{3}{2}$ and $L_f(t) < 2$ in $[0,+\infty)$. So, we can take as t_0 any point in [1,5] such that $f(t_0) > 0$. Taking $t_0 = 4$, by Theorem 3.1 we obtain that the sequence $\{t_n\}$ given by Eq. (13) is decreasing and converges to the root $t^* = 2.465652356209171$ of $f(t) = 0$ (see Table 1).

Table 1
Halley method

Iteration	t_n
0	4.000000000000000
1	3.741961670527600
2	3.465034280857143
3	3.165567038560662
4	2.846385484142443
5	2.563813967367478
6	2.467671172927391
7	2.465652373241162
8	2.465652356209171
9	2.465652356209171

4. Kantorovich type conditions for the Halley method in the complex plane

Denote by D a non-empty convex open subset of $\mathbb{C}$ and f an holomorphic complex function in D. We will solve

$$f(z) = 0. \tag{15}$$

In this section we give a result on convergence and uniqueness of the solution for Eq. (15). To do that, we use the Halley method in the complex plane. We assume that f satisfies some Kantorovich type conditions [2], we use the majorizing theory [11] and the results obtained in Section 3 for scalar equations.

Starting from $z_0 \in D$, the Halley iteration $\{z_n\}$ is given by

$$z_n = F(z_{n-1}) = z_{n-1} - \frac{f(z_{n-1})}{f'(z_{n-1})} \frac{2}{2 - L_f(z_{n-1})}. \tag{16}$$

We try to prove the convergence of the sequence $\{z_n\}$ to a root z^* of Eq. (15). First, assume that $|f(z_0)/f'(z_0)| \leqslant \eta$, with $\eta \leqslant \frac{1}{2}$, and denote by $t^* = 1 - \sqrt{1 - 2\eta}$ the smallest root of $p(t) = 0$, for $p(t) = t^2 - 2t + 2\eta$.

Lemma 4.1. *If* $t_0 = 0$, *the sequence* $\{t_n\}$ *given by*

$$t_n = P(t_{n-1}) = t_{n-1} - \frac{p(t_{n-1})}{p'(t_{n-1})} \frac{2}{2 - L_p(t_{n-1})} \tag{17}$$

is increasing and converges to t^*.

Proof. Consider $q(t) = p(-t) = t^2 + 2t + 2\eta$. This is an increasing convex function in $[-1,0]$. Then if $s_0 = 0$, the sequence

$$s_n = s_{n-1} - \frac{q(s_{n-1})}{q'(s_{n-1})} \frac{2}{2 - L_q(s_{n-1})}$$

is decreasing to $1 - \sqrt{1 - 2\eta}$ from Theorem 3.1, since $L_{q'}(t) = 0 \leqslant \frac{3}{2}$ and $L_q(t) < 2$ in $[-1,0]$.

On the other hand, as $t_n = -s_n$ for each $n \in \mathbb{N}$, the result holds. $\square$

Denote $B(a,r) = \{z \in \mathbb{C} : |z-a| < r\}$ and $\overline{B(a,r)} = \{z \in \mathbb{C} : |z-a| \leqslant r\}$ for $a \in \mathbb{C}$ and $r \in (0, +\infty)$.

Theorem 4.2. *Under the above conditions assume that there exists* $\alpha \geqslant 1$ *such that*

(i) $|L_f(z)| \leqslant L_p(t)/\alpha$ *when* $|z - z_0| \leqslant t - t_0$,

(ii) $|L_{f'}(z)| \in \overline{B(3/2, 3\alpha^2/2)}$ *for all* $z \in \overline{B(z_0, t^*)}$.

Then, the sequence $\{z_n\}$ given by Eq. (16) *is well defined, lies in* $\overline{B(z_0, t^*)}$ *and converges to the only root* z^* of Eq. (15) *in* $B(z_0, t^*)$. *Besides* $|z^* - z_n| \leqslant t^* - t_n$.

Proof. Let $\{t_n\}$ be the real sequence defined in Lemma 4.1. Following Ref. [9] the sequence $\{t_n\}$ majorizes $\{z_n\}$ provided that

(a) $|F(z_0) - z_0| \leqslant P(t_0) - t_0$,

(b) $|F'(z)| \leqslant P'(t)$ when $|z - z_0| \leqslant t - t_0$.

Taking into account (i),

$$|F(z_0) - z_0| = \left| \frac{f(z_0)}{f'(z_0)} \left[1 + \frac{L_f(z_0)}{2} \right] \right|$$

and $P(0) - 0 = 2\eta/2 - \eta$, then (a) holds.

On the other hand, by Eq. (14), (i) and (ii), we can derive (b). The uniqueness is clear since $P(t^*) = t^*$ and t^* is the only root of $p(t) = 0$ in $[0, t^*]$. $\square$

Notice that the convergence of this iterative process does not depend on the initial point $z_0 \in D$. Indeed, if $\eta > \frac{1}{2}$, we can consider $p(t) = t^2 - 2\eta t + \eta^2 - \frac{1}{4}$, and obtain, like in Lemma 4.1, that the sequence given by Eq. (17) is increasing to $\eta - \frac{1}{2}$, for $t_0 = 0$. So we can apply Theorem 4.2, and thus the Halley method converges for all $z_0 \in D$.

To finish, notice that we can obtain a convergence result for the Halley method in the complex plane from Theorem 3.3. Results on uniqueness are not given in this case.

5. Halley method in Banach spaces

Let X, Y be Banach spaces and $F : \Omega \subseteq X \to Y$ be a nonlinear twice Fréchet differentiable operator in an open convex domain $\Omega_0 \subseteq \Omega$. We denote by I the identity operator on X and $L_F(x)$ the linear operator defined by

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

provided that $\Gamma(x) = F'(x)^{-1}$ exists. This operator and its connection with Newton's method were studied in Ref. [5].

The Halley method for solving the equation

$$F(x) = 0 \tag{18}$$

is defined, starting from x_0, by

$$x_{n+1} = x_n - \left[I - \frac{1}{2}L_F(x_n)\right]^{-1} \Gamma_n F(x_n). \tag{19}$$

Using majorizing sequences a convergence result for the Halley method is given by Yamamoto in Ref. [14]. To prove this result he assumes that

(I) There exists $\Gamma_0 = [F'(x_0)]^{-1}, x_0 \in \Omega_0$.

(II) $\|\Gamma_0 F(x_0)\| \leqslant a$ and $\|\Gamma_0 F''(x_0)\| \leqslant b$.

(III*) $\|\Gamma_0(F''(x) - F''(y))\| \leqslant k\|x - y\|$ for $x, y \in \Omega_0$ and $k \geqslant 0$.

The condition (III*) reduces the Kantorovich–Altman condition [1] because he relaxes the requirement that there exist F''' and $\sup\{\|F'''(x)\|; x \in \Omega_0\} \leqslant M$.

In this paper we extend the condition (III*) and we assume that

(III) $\|\Gamma_0(F''(x) - F''(y))\| \leqslant k\|x - y\|^p \quad$ for $\quad x, y \in \Omega_0,\ k \geqslant 0$ and $p > 0$,

and the equation

$$f(t) = a - t + \frac{b}{2}t^2 + \frac{k}{(p+1)(p+2)}t^{p+2} = 0. \tag{20}$$

has two positive roots r_1 and r_2 $(r_1 \leqslant r_2)$.

Newton's method under this kind of conditions for F' and $p \in [0, 1]$ has been studied by different authors [13–15].

(IV) Besides, if $p \in [0, 1)$, we assume that there exist F''' and

$$\|\Gamma_0 F'''(x)\| \leqslant -\frac{f'''(t)}{f'(t_0)} \quad \text{when} \quad \|x - x_0\| \leqslant t - t_0.$$

With the previous notation, we have the following result.

Theorem 5.1. *Let $x_0 \in \Omega_0$ satisfy (I), (II), (III),* and *(IV). And let $m > 0$ be the minimum of the function f given by* Eq. (20). *We denote by r_1 the smallest positive root of* Eq. (20) *and assume that $f(m) \leqslant 0$ and $B(x_0, m) \subseteq \Omega_0$. Then, the Halley sequence $\{x_n\}$ given by* Eq. (19) *is well defined and converges to a solution x^* of* Eq. (18) *in $\overline{B(x_0, r_1)}$. Moreover we also have*

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

Notice that the equation $f'(t) = 0$ has only one positive solution m which is a minimum of f. Therefore $f(m) \leqslant 0$ is a necessary and sufficient condition for the existence of positive solutions of $f(t) = 0$. So

$$f(t) = (r_1 - t)(r_2 - t)q(t)$$

with $q(r_1) \neq 0 \neq q(r_2)$. Note that f, given by Eq. (20), is a decreasing convex function in $[0, m], f(0) > 0 \geqslant f(m), L_{f'}(t) \leqslant 0$ in $[0, m]$ and $L_f(t) \leqslant 1/2$ in $[0, r_1]$

as it is well known [1]. These conditions are sufficient to show the convergence to r_1 of the sequence

$$t_0 = 0, \quad t_{n+1} = t_n - \left[1 + \frac{L_f(t_n)}{2 - L_f(t_n)}\right] \frac{f(t_n)}{f'(t_n)}, \;\; n \geqslant 0, \tag{21}$$

by Theorem 3.1. To prove the Theorem 5.1, we need the following previous result.

Lemma 5.2. *The sequence* $\{t_n\}$ *defined by* Eq. (21) *is a majorizing sequence of* $\{x_n\}$ *given by* Eq. (19), *i.e.* $\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n$ for $n \geqslant 0$.

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

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

[II_n] $\|\Gamma_0 F''(x_n)\| \leqslant -\dfrac{f''(t_n)}{f'(t_0)}$,

[III_n] $\|\Gamma_n F'(x_0)\| \leqslant \dfrac{f'(t_0)}{f'(t_n)}$,

[IV_n] $\|\Gamma_0 F(x_n)\| \leqslant -\dfrac{f(t_n)}{f'(t_0)}$,

[V_n] $\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n$.

All the above statements [I_0]–[IV_0] are true for $n = 0$ by initial conditions (I)–(IV). [V_0] follows taking into account that under the previous assumptions for f we have

$$\|L_F(x_0)\| \leqslant L_f(t_0) < \frac{1}{2}, \quad \text{for } t \in [0, r_1].$$

Therefore

$$\|x_1 - x_0\| \leqslant t_1 - t_0 = t_1 \leqslant r_1 < m.$$

Now, assuming [I_n]–[V_n] prove [I_{n+1}]–[V_{n+1}]. As

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

then

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

Therefore

$$\|I - \Gamma_0 F'(x_{n+1})\| \leqslant \frac{k}{p+1}\|x_{n+1} - x_0\|^{p+1} + b\|x_{n+1} - x_0\| \leqslant 1 + f'(t_{n+1}) < 1.$$

Applying the Banach lemma, $(\Gamma_0 F'(x_{n+1}))^{-1}$ exists. Hence $[\mathrm{I}_{n+1}]$ and $[\mathrm{III}_{n+1}]$ are true. Since

$$\|\Gamma_0 F''(x_{n+1})\| \leqslant k\|x_{n+1} - x_0\|^p + b \leqslant f''(t_{n+1}),$$

$[\mathrm{II}_{n+1}]$ is also true.

Firstly, if $p \geqslant 1$, using the Altman technique [1,14], and taking into account Eq. (19), we deduce by Taylor's formula that

$$F(x_{n+1}) = \frac{1}{2}F''(x_n)\frac{1}{2}\Gamma_n F'(x_n)L_F(x_n)\left[\left[I - \frac{1}{2}L_F(x_n)\right]^{-1}\Gamma_n F(x_n)\right]^2 + \int_0^1 [F''(x_n + t(x_{n+1} - x_n)) - F''(x_n)](x_{n+1} - x_n)^2(1-t)\,\mathrm{d}t,$$

and taking norms we obtain

$$\|\Gamma_0 F(x_{n+1})\| \leqslant f(t_n)\frac{L_f(t_n)^2}{(2 - L_f(t_n))^2} + k(t_{n+1} - t_n)^{p+2}\int_0^1 t^p(1-t)\,\mathrm{d}t.$$

Repeating the same process for the function f, we get

$$f(t_{n+1}) = f(t_n)\frac{L_f(t_n)^2}{(2 - L_f(t_n))^2} + k\int_{t_n}^{t_{n+1}} (z^p - t_n^p)(t_{n+1} - z)\,\mathrm{d}z.$$

Moreover it follows immediately if $p \geqslant 1$ that

$$(t_{n+1} - t_n)^{p+2}\int_0^1 t^p(1-t)\,\mathrm{d}t = \int_{t_n}^{t_{n+1}} (z - t_n)^p(t_{n+1} - z)\,\mathrm{d}z \leqslant \int_{t_n}^{t_{n+1}} (z^p - t_n^p)(t_{n+1} - z)\,\mathrm{d}z$$

and consequently, for $p \geqslant 1$, we have

$$\|\Gamma_0 F(x_{n+1})\| \leqslant f(t_{n+1}) = -\frac{f(t_{n+1})}{f'(t_0)}.$$

Secondly, if $p \in [0,1)$, we use the same process but taking as remainder of Taylor's formula

$$\frac{1}{2}\int_{r_n}^{r_{n+1}} F'''(x)(x_{n+1} - x_n)^2\,\mathrm{d}x.$$

Therefore $[\mathrm{IV}_{n+1}]$ is true.

Finally, we have

$$\|x_{n+2}-x_{n+1}\| = \left\| -\left[I-\frac{1}{2}L_F(x_{n+1})\right]^{-1}\Gamma_{n+1}F(x_{n+1})\right\|$$
$$\leqslant -\left[I-\frac{1}{2}L_f(t_{n+1})\right]^{-1}\frac{f(t_{n+1})}{f'(t_{n+1})} = t_{n+2}-t_{n+1}.$$

That completes the proof of Lemma 5.2. □

Proof of Theorem 5.1. The convergence of $\{t_n\}$ implies the convergence of $\{x_n\}$ to x^*, and making $n\to\infty$ in [IV$_n$] we obtain $F(x^*)=0$.

Besides, from [V$_n$] we have

$$\|x_{n+s}-x_n\| \leqslant t_{n+s}-t_n, \quad \text{for } s\geqslant 0,$$

and making $s\to\infty$, we obtain $\|x^*-x_n\|\leqslant r_1-t_n$. Therefore, it follows that

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

Numerical example: To illustrate Theorem 5.1, we consider the following differential equation:

$$y''+4y'+y^3=12, \quad y(0)=0=y(1). \tag{22}$$

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 the corresponding values of the function

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

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

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

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

$$F(y)=G(y)+hJ(y)+h^2g(y)-12h^2M,$$

where

$$G=\begin{pmatrix} -2 & 1 & 0 & 0 & \ldots & 0 & 0\\ 1 & -1 & 1 & 0 & \ldots & 0 & 0\\ 0 & 1 & -2 & 1 & \ldots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \vdots & \ldots & 1 & -2 \end{pmatrix},$$

$$J = \begin{pmatrix} 0 & 2 & 0 & 0 & \dots & 0 & 0 \\ -2 & 0 & 2 & 0 & \dots & 0 & 0 \\ 0 & -2 & 0 & 2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \dots & -2 & 0 \end{pmatrix},$$

$$g(y) = \begin{pmatrix} y_1^3 \\ y_2^3 \\ \vdots \\ y_{n-1}^3 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Then

$$F'(y) = G + hJ + 3h^2 \begin{pmatrix} y_1^2 & 0 & \dots & 0 \\ 0 & y_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & y_{n-1}^2 \end{pmatrix}$$

and

$$F''(y) = 6h^2 \begin{pmatrix} A_1 \\ \vdots \\ A_{n-1} \end{pmatrix} \quad \text{with } A_j = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & y_j & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \; j = 1, 2, \dots, n-1.$$

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

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

Note that for each $x, y \in \mathbb{R}^{n-1}$ we get

$$\|F''(x) - F''(y)\| = \|\operatorname{diag} 6h^2 (x_i - y_i)\| = 6h^2 \max_{1 \leqslant i \leqslant n-1} |x_i - y_i| \leqslant 6h^2 \|x - y\|.$$

We choose $n = 10$. As a solution would vanish at the endpoints and be positive in the interior, a reasonable choice of initial approximation seems to be $\sin \pi x/10$. This gives us the following vector,

$$x_0 = \begin{pmatrix} 0.0309017 \\ 0.0587785 \\ 0.0809017 \\ 0.0951057 \\ 0.1 \\ 0.0951057 \\ 0.0809017 \\ 0.0587785 \\ 0.0309017 \end{pmatrix}.$$

We get the following results for $p = 1$ in Theorem 5.1,

$$a = \|\Gamma_0 F(x_0)\| = 0.806994, \quad b = \|\Gamma_0 F''(x_0)\| = 0.0288905,$$
$$k = 6h^2\|\Gamma_0\| = 0.355981.$$

Therefore, Eq. (22) becomes

$$0.806994 - t + 0.0144452t^2 + 0.0593302t^3 = 0.$$

This equation has two positive real solutions: $r_1 = 0.854571$ and $r_2 = 3.47812$. Hence by Theorem 5.1 the sequence (21) converges to a solution $y^* = (y_1^*, y_1^*, \ldots, y_9^*)$ of equation $f(y) = 0$ in $\overline{B(x_0, r_1)}$.

Finally note that so as to solve differential equation (22), 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 Eq. (22).

Acknowledgements

This work was supported in part by a grant of the University of La Rioja (grants: API-98/A2S and API-98/B12) and DGES (grant: PB96-0120-C03-02).

References

[1] D. Altman, Concerning the method of tangent hyperbolas for operator equations, Bull. Polish Acad. Sci., Ser. Sci., Math., Astronom. Phys. 9 (1961) 633–637.
[2] I.K. Argyros, D. Chen, Results on the Chebyshev Method in Banach Spaces, Proyecciones 12 (2) (1993) 119–128.
[3] R. Campbell, Les Intégrales Euleriennes et Leurs Aplications, Dunod, 1966.

[4] Z. Ciesielski, Some properties of convex functions on higher orders, Ann. Polon. Math. 7 (1959) 1–7.
[5] J.M. Gutiérrez, M.A. Hernández, M.A. Salanova, Accessibility of solutions by Newton's method, Int. J. Comput. Math. 57 (1995) 239–247.
[6] M.A. Hernández, A note on Halley's method, Numer. Math. 59 (3) (1991) 273–276.
[7] M.A. Hernández, Newton–Raphson's method and convexity, Zb. Rad. Prirod.-Mat. Fak. Ser. Mat. 22-1 (1992) 159–166.
[8] A.S. Householder, The Numerical Treatment of a Single Nonlinear Equation, McGraw-Hill, New York, 1970.
[9] L.V. Kantorovich, G.P. Akilov, Functional Analysis, Pergamon Press, Oxford, 1982.
[10] B. O'Neill, Elementary Differential Geometry, Academic Press, New York, 1966.
[11] W.C. Rheinboldt, A unified convergence theory for a class of iterative process, SIAM J. Numer. Anal. 5 (1968) 42–63.
[12] A.W. Roberts, D.E. Varberg, Convex Functions, Academic Press, New York, 1973.
[13] J. Rokne, Newton's method under mild differentiability conditions with error analysis, Numer. Math. 18 (1972) 401–412.
[14] T. Yamamoto, On the method of tangent hyperbolas in Banach spaces, J. Comput. Appl. Math. 21 (1988) 75–86.
[15] H. Zhengda, A note on the Kantorovich theorem for Newton iteration, J. Comput. Appl. Math. 47 (1993) 211–217.