Available online at www.sciencedirect.com

SCIENCE DIRECT®

J. Math. Anal. Appl. 303 (2005) 591–601

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.elsevier.com/locate/jmaa

On the R-order of the Halley method ☆

J.A. Ezquerro *, M.A. Hernández

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

Received 15 March 2004

Available online 23 November 2004

Submitted by M. Iannelli

Abstract

An R-order bound for the Halley method is obtained in this work, where an analysis of the convergence of the method is also presented under mild differentiability conditions. To do this, a new technique is developed, where the involved operator must satisfy some recurrence relations.

Keywords: Nonlinear equations in Banach spaces; Halley's method; Semilocal convergence theorem; Recurrence relations; A priori error estimates; R-order of convergence

1. Introduction

The problem of approximating a solution of a nonlinear equation

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

is very interesting, since we can then solve a large number of different types of problems. So, if F is a nonlinear operator defined on a nonempty open convex subset Ω of a Banach space X with values in a Banach space Y, Eq. (1) can represent a differential equation,

☆ Preparation of this paper was partly supported by the Ministry of Science and Technology (BFM 2002-00222), the University of La Rioja (API-03/15) and the Government of La Rioja (ACPI 2003/2004).

* Corresponding author.

E-mail addresses: jezquer@dmc.unirioja.es (J.A. Ezquerro), mahernan@dmc.unirioja.es (M.A. Hernández).

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2004.08.057

a boundary value problem, an integral equation, etc. The normal way to approximate a solution of (1) is by means of iterative processes. An iterative process is defined by an algorithm such that, from an initial approximation x_0, it is constructed a sequence $\{x_n\}$ that satisfies $\lim_n x_n = x^*$ and $F(x^*) = 0$.

In the study of iterative methods there are two especially important sides: the convergence of the sequence $\{x_n\}$ to a solution x^* of (1) and the speed of this convergence. Moreover, different types of convergence analysis can be done. The semilocal convergence analysis takes into account some conditions for the operator F and the initial approximation x_0 of the iteration considered for approximating the solutions of (1). If we only require conditions for the operator F, the convergence result is global, and if the solution x^* must satisfy some conditions, the convergence result is local.

This paper considers Halley's method (see [4])

$$\begin{aligned} y_n &= x_n - \left[F'(x_n)\right]^{-1} F(x_n), \\ H(x_n, y_n) &= F(x_n)^{-1} F''(x_n)(y_n - x_n), \\ x_{n+1} &= y_n - \frac{1}{2} H(x_n, y_n)\left[I + \frac{1}{2} H(x_n, y_n)\right]^{-1}(y_n - x_n), \quad n \geqslant 0, \end{aligned} \tag{2}$$

for solution of (1). The Halley method is one of the well-known numerical processes for solving (1) (see [10], where an extensive reference list can be found). Basic results concerning the convergence of the process, existence and uniqueness regions of solutions are given by other authors (see [3,12] for the references appearing there). The results concerning convergence have been published under assumptions of Newton–Kantorovich type. In [4,5,11], an abundant list of references can be found, where several techniques for finding sufficient conditions for the convergence of Halley's iteration appear.

In this paper, we pay attention to the semilocal convergence analysis. Initially, see [9], the required assumptions to study the convergence of Halley's method were:

(A1) suppose that $\Gamma_0 = F'(x_0)^{-1} \in \mathcal{L}(Y, X)$ exists at some $x_0 \in \Omega$, where $\mathcal{L}(Y, X)$ is the set of bounded linear operators from Y into X and $\|\Gamma_0\| \leqslant \beta$,
(A2) $\|y_0 - x_0\| = \|\Gamma_0 F(x_0)\| \leqslant \eta$,
(A3) $\|F''(x)\| \leqslant M$, $x \in \Omega$,
(A4) $\|F'''(x)\| \leqslant N$, $x \in \Omega$.

Under assumptions (A1)–(A4) a semilocal convergence result is obtained. Next, this study can be modified by replacing condition (A4) for

$$\left\|F''(x) - F''(y)\right\| \leqslant K\|x - y\|, \quad K \geqslant 0,\ x, y \in \Omega \tag{3}$$

(see [1,4,11]), which is milder, and keeping (A1), (A2) and (A3). The next step is to relax condition (3) by the following:

$$\left\|F''(x) - F''(y)\right\| \leqslant L\|x - y\|^p, \quad L \geqslant 0,\ p \in [0, 1],\ x, y \in \Omega \tag{4}$$

(see [7]). Notice that conditions (3) and (4) mean that F'' is Lipschitz continuous in Ω and F'' is (L, p)-Hölder continuous in Ω, respectively.

Under conditions (3) and (4) the number of equations that can be solved by Halley's iteration is limited. For instance, we cannot analyze the convergence of the Halley method to a solution of equations where additions of operators, which satisfy (3) or (4), appear. We then consider the following nonlinear integral equation of mixed Hammerstein type [6]:

$$x(s)+\sum_{i=1}^{m}\int_a^b k_i(s,t)\ell_i\big(x(t)\big)\,dt=u(s),\quad s\in[a,b],$$

where $-\infty<a<b<\infty$, u, ℓ_i, and k_i, for $i=1,2,\dots,m$, are known functions and x is a solution to be determined. If $\ell_i''(x(t))$ is (L_i,p_i)-Hölder continuous in Ω, for $i=1,2,\dots,m$, the corresponding operator $F:\Omega\subseteq C[0,1]\to C[0,1]$,

$$\big[F(x)\big](s)=x(s)+\sum_{i=1}^{m}\int_a^b k_i(s,t)\ell_i\big(x(t)\big)\,dt-u(s),\quad s\in[a,b], \tag{5}$$

does not satisfy (3) neither (4) when, for instance, the max-norm is considered. In this case,

$$\big\|F''(x)-F''(y)\big\|\leqslant\sum_{i=1}^{m}L_i\|x-y\|^{p_i},\quad L_i\geqslant 0,\ p_i\in[0,1],\ x,y\in\Omega.$$

To solve this type of equations and to relax conditions (3) and (4) we can consider

$$\big\|F''(x)-F''(y)\big\|\leqslant\omega\big(\|x-y\|\big),\quad x,y\in\Omega, \tag{6}$$

where $\omega(z)=\sum_{i=1}^{m}L_iz^{p_i}$. We then require that $\omega(z)$ is a nondecreasing continuous real function for $z>0$, such that $\omega(0)\geqslant 0$.

Obviously conditions (A4), (3) and (4) are relaxed by condition (6). Besides the former ones are generalized by the latter one if $\omega(z)=N$, $\omega(z)=Kz$ and $\omega(z)=Lz^p$, respectively.

On the other hand, the convergence properties depends on the choice of the distance $\|\cdot\|$, but for a given distance the speed of convergence of the sequence $\{x_n\}$ is characterized by the speed of convergence of the sequence of nonnegative numbers $\|x^*-x_n\|$. An important measure of the speed of convergence is the R-order of convergence (see [8]). It is known that a sequence $\{z_n\}$ converges with R-order at least $\tau>1$ if there are constants $C\in(0,\infty)$ and $\gamma\in(0,1)$ such that $z_n\leqslant C\gamma^{\tau^n}$, $n=0,1,\dots$.

Under conditions (A1)–(A4) or (A1)–(A3) and (3), the Halley method is of R-order at least three (see [2]) and the R-order of Halley's iteration has not been studied under conditions (A1)–(A3) and (4). Here we present a new technique consisting of a system of recurrence relations for analyzing the semilocal convergence of the Halley method and prove that, under the mildest conditions (A1)–(A3) and (6), the Halley process is of R-order at least two, but if $\omega(tz)\leqslant t^q\omega(z)$, for $z>0$, $t\in[0,1]$, $q\in[0,1]$, the R-order of convergence is at least $2+q$. From this, the R-order at least $2+p$ is deduced if (A1)–(A3) and (4) are satisfied.

Moreover, to find a priori estimates for the distances $\|x^*-x_n\|$, $n=1,2,\dots$, we look for a function $\alpha:\mathbb{N}\to\mathbb{R}_+$ such that $\|x^*-x_n\|\leqslant\alpha(n)$, $n=1,2,\dots$. On the basis of the new technique developed here, a priori error bounds are derived for the Halley sequence.

Throughout the paper we denote $\overline{B(x,r)} = \{y \in X;\ \|y-x\| \leqslant r\}$ and $B(x,r) = \{y \in X;\ \|y-x\| < r\}$.

2. Semilocal convergence of Halley's method

To establish a semilocal convergence result for the Halley method, certain conditions for the operator F and the initial approximation x_0 are required. Conclusions about the existence and uniqueness of solutions of (1) are also obtained. We provide the regions of existence and uniqueness of solutions from the theoretical significance of the Halley method, without finding the solutions themselves. This is sometimes more important than the actual knowledge of a solution.

A new technique is developed to prove the semilocal convergence of sequence (2), where we construct, from some scalar parameters, a system of recurrence relations where two real sequences of positive real numbers are involved. The convergence of iteration (2) is then guaranteed from the fact that (2) is a Cauchy sequence.

2.1. Recurrence relations

Let us suppose that $\Gamma_0 = F'(x_0)^{-1} \in \mathcal{L}(Y,X)$ exists at some $x_0 \in \Omega$, where $\mathcal{L}(Y,X)$ is the set of bounded linear operators from Y into X. Moreover, we assume the following assumptions:

(C1) $\|\Gamma_0\| \leqslant \beta$,
(C2) $\|y_0 - x_0\| = \|\Gamma_0 F(x_0)\| \leqslant \eta$,
(C3) $\|F''(x)\| \leqslant M$, $x \in \Omega$,
(C4) $\|F''(x) - F''(y)\| \leqslant \omega(\|x-y\|)$, $x, y \in \Omega$, where $\omega(z)$ is a nondecreasing continuous real function for $z > 0$, such that $\omega(0) \geqslant 0$,
(C5) there exists a positive real function $\varphi \in C[0,1]$, with $\varphi(t) \leqslant 1$, such that $\omega(tz) \leqslant \varphi(t)\omega(z)$, for $t \in [0,1]$ and $z \in (0,+\infty)$.

Note that condition (C5) is not restrictive, since we can always consider $\varphi(t) = 1$, as a consequence of ω is a nondecreasing function, but its interest is to sharp the a priori error bounds. We denote $A = \int_0^1 (1-t)\varphi(t)\,dt$.

From (C1)–(C5), we consider $a_0 = M\beta\eta$ and $b_0 = \beta\eta\omega(\eta)$. Observe that if $x_1 \in \Omega$ and $a_0 < 2$, we have $\|H(x_0,y_0)\| \leqslant a_0$ and, by the Banach lemma, $[I + \frac{1}{2}H(x_0,y_0)]^{-1} = P(x_0,y_0)$ exists and $\|P(x_0,y_0)\| \leqslant \frac{2}{2-a_0}$, since $a_0 < 2$. Moreover, if $b_0 < \frac{a_0^2 - 4a_0 + 2}{A(2-a_0)}$,

$$\|x_1 - y_0\| \leqslant \frac{1}{2}\big\|H(x_0,y_0)\big\|\,\big\|P(x_0,y_0)\big\|\,\|y_0 - x_0\| \leqslant \frac{a_0}{2-a_0}\|y_0 - x_0\|$$

and

$$\|x_1 - x_0\| \leqslant \|x_1 - y_0\| + \|y_0 - x_0\| \leqslant \frac{2}{2-a_0}\|y_0 - x_0\| < R\eta,$$

where

$$R = \frac{2}{(2-a_0)(1-\frac{2}{2-3a_0}(\frac{a_0^2}{2(2-a_0)}+Ab_0))}.$$

This value of R is deduced later. Consequently, $y_0, x_1 \in B(x_0, R\eta)$. Furthermore, from

$$\|I - \Gamma_0 F'(x_1)\| \leqslant \|\Gamma_0\| \|F'(x_0) - F'(x_1)\| \leqslant M\|\Gamma_0\|\|x_1 - x_0\| \leqslant \frac{2a_0}{2-a_0}$$

and $a_0 < 2$, it follows $\|I - \Gamma_0 F'(x_1)\| < 1$ and, by the Banach lemma, Γ_1 exists and $\|\Gamma_1\| \leqslant \frac{2-a_0}{2-3a_0}\|\Gamma_0\|$. Therefore, x_1 is well defined.

To prove the sequence $\{x_n\}$, defined by (2), is well defined, we first define the following real functions:

$$f(x) = \frac{2}{2-x}, \qquad g(x) = \frac{2-x}{2-3x} \quad \text{and} \quad h(x,y) = \frac{x^2}{2(2-x)} + Ay \tag{7}$$

that satisfy the properties appearing in the following lemma.

Lemma 2.1. *Let f, g and h be the three scalar functions given in* (7). *Then*

(a) *$f(x)$ and $g(x)$ are increasing in $x \in (0, \frac{3-\sqrt{5}}{2})$,*
(b) *$h(x,y)$ is increasing in its first and second arguments for $x \in (0, \frac{3-\sqrt{5}}{2})$ and $y > 0$.*

Now, we introduce an approximation of F in Lemma 2.2, where the approximations introduced in (2) are used. From a similar approximation presented in [4], the proof of Lemma 2.2 follows immediately.

Lemma 2.2. *Let F be a nonlinear operator defined on an open convex subset Ω of a Banach space X with values in a Banach space Y. Suppose that the operator F has continuous second-order Fréchet-derivatives on Ω. Then, the following approximations is true for all $n \geqslant 0$:*

$$\begin{aligned} F(x_{n+1}) = {} & \int_0^1 F''\big(y_n + t(x_{n+1}-y_n)\big)(1-t)\,dt\,(x_{n+1}-y_n)^2 \\ & - \frac{1}{2}\int_0^1 F''\big(x_n + t(y_n - x_n)\big)t\,dt\,(y_n - x_n)P(x_n,y_n)H(x_n,y_n)(y_n - x_n) \\ & + \int_0^1 \big[F''\big(x_n + t(y_n-x_n)\big) - F''(x_n)\big](1-t)\,dt \\ & \times (y_n - x_n)P(x_n,y_n)(y_n - x_n), \end{aligned}$$

where $P(x_n,y_n) = [I + \frac{1}{2}H(x_n,y_n)]^{-1}$.

Taking into account now the approximation of Lemma 2.2 for $n = 0$, we obtain the next bound

$$\|F(x_1)\| \leqslant \left(\frac{M\eta}{8}a_0^2 f(a_0)^2 + \frac{M\eta}{4}a_0 f(a_0) + A\eta\omega(\eta) f(a_0)\right)\|y_0 - x_0\|.$$

Next, we suppose

$$a_0 \in \left(0, \frac{3-\sqrt{5}}{2}\right) \quad \text{and} \quad b_0 < \frac{2(a_0^2 - 3a_0 + 1)}{A(2 - a_0)}. \tag{8}$$

Notice that the bounds for the parameters a_0 and b_0 have been restricted as a consequence of the following required development. Then $f(a_0)g(a_0)h(a_0, b_0) = c_0 < 1$ and

$$\begin{aligned}\|y_1 - x_1\| &= \|\Gamma_1 F(x_1)\| \leqslant \|\Gamma_1\|\|F(x_1)\| \leqslant f(a_0)g(a_0)h(a_0, b_0)\|y_0 - x_0\| \\ &= c_0\|y_0 - x_0\| < \eta,\end{aligned}$$

so that

$$M\|\Gamma_1\|\|y_1 - x_1\| \leqslant M\|\Gamma_0\|g(a_0)c_0\|y_0 - x_0\| \leqslant a_0 g(a_0)c_0$$

and

$$\begin{aligned}\|\Gamma_1\|\omega(\|y_1 - x_1\|)\|y_1 - x_1\| &\leqslant \|\Gamma_0\|g(a_0)\omega(c_0\|y_0 - x_0\|)c_0\|y_0 - x_0\| \\ &\leqslant b_0 g(a_0)c_0\varphi(c_0).\end{aligned}$$

Now, from the Banach lemma, $P(x_1, y_1) = [I + \frac{1}{2}H(x_1, y_1)]^{-1}$ exists, since $\|H(x_1, y_1)\| \leqslant a_0 g(a_0)c_0$, and $\|P(x_1, y_1)\| \leqslant f(a_0 g(a_0)c_0)$. Thus

$$\begin{aligned}&\|x_2 - y_1\| \leqslant \frac{1}{2}a_0 g(a_0)c_0 f(a_0 g(a_0)c_0)\|y_1 - x_1\|, \\ &\|x_2 - x_1\| \leqslant \|x_2 - y_1\| + \|y_1 - x_1\| \leqslant f(a_0 g(a_0)c_0)\|y_1 - x_1\|\end{aligned}$$

and, as f is increasing in $(0, \frac{3-\sqrt{5}}{2})$,

$$\|x_2 - x_0\| \leqslant \|x_2 - x_1\| + \|x_1 - x_0\| \leqslant (1 + c_0) f(a_0)\|y_0 - x_0\| < R\eta,$$

since $a_0 g(a_0)c_0 < a_0 < \frac{3-\sqrt{5}}{2}$.

Finally, from

$$\begin{aligned}\|I - \Gamma_1 F'(x_2)\| &\leqslant \|\Gamma_1\|\|F'(x_1) - F'(x_2)\| \leqslant M\|\Gamma_1\|\|x_2 - x_1\| \\ &\leqslant a_0 g(a_0)c_0 f(a_0 g(a_0)c_0)\end{aligned}$$

and $a_0 g(a_0)c_0 < a_0 < \frac{3-\sqrt{5}}{2}$, it follows $\|I - \Gamma_1 F'(x_2)\| < 1$ and, by the Banach lemma, Γ_2 exists and $\|\Gamma_2\| \leqslant g(a_0 g(a_0)c_0)\|\Gamma_1\|$. Consequently, x_2 is also well defined.

Note that we can do then $a_0 g(a_0)c_0 = a_1$ and $b_0 g(a_0)c_0\varphi(c_0) = b_1$ to define the following real sequences:

$$\begin{aligned}&a_n = a_{n-1}g(a_{n-1})c_{n-1}, \quad n \geqslant 1, \\ &b_n = b_{n-1}g(a_{n-1})c_{n-1}\varphi(c_{n-1}), \quad n \geqslant 1, \\ &c_n = f(a_n)g(a_n)h(a_n, b_n), \quad n \geqslant 1,\end{aligned}$$

that satisfy the properties of Lemma 2.3.

Lemma 2.3. *Let f, g and h be the three scalar functions given in* (7)*. If a_0 and b_0 satisfy* (8)*, then*

(a) $c_0 < 1$ *and* $g(a_0)c_0 < 1$,
(b) *the sequences* $\{a_n\}$, $\{b_n\}$ *and* $\{c_n\}$ *are decreasing,*
(c) $a_n < \frac{3-\sqrt{5}}{2}$, *for all* $n \geqslant 0$.

Proof. Item (a) is trivial from the hypotheses. Next, we invoke the induction hypotheses and use Lemma 2.1 to prove item (b). Finally, as the sequence $\{a_n\}$ is decreasing and $a_0 \in (0, \frac{3-\sqrt{5}}{2})$, it follows that $a_n < a_0 < \frac{3-\sqrt{5}}{2}$, for all $n \geqslant 0$. □

Since our goal is to show the sequence $\{x_n\}$, given by (2), is well defined, we present in Lemma 2.4 a system of recurrence relations from which we obtain the last. From a similar way that the mentioned above and using induction the proof of Lemma 2.4 follows.

Lemma 2.4. *Under the hypotheses of Lemma* 2.3*, the following items are true for all* $n \geqslant 1$:

(I) Γ_n *exists and* $\|\Gamma_n\| = \|F'(x_n)^{-1}\| \leqslant g(a_{n-1})\|\Gamma_{n-1}\|$,
(II) $\|y_n - x_n\| \leqslant c_{n-1}\|y_{n-1} - x_{n-1}\| \leqslant c_0^n\|y_0 - x_0\| < \eta$,
(III) $M\|\Gamma_n\|\|y_n - x_n\| \leqslant a_n$,
(IV) $\|\Gamma_n\|\omega(\|y_n - x_n\|)\|y_n - x_n\| \leqslant b_n$,
(V) $P(x_n, y_n) = [I + \frac{1}{2}H(x_n, y_n)]^{-1}$ *exists and* $\|P(x_n, y_n)\| \leqslant f(a_n)$,
(VI) $\|x_{n+1} - y_n\| \leqslant \frac{a_n}{2} f(a_n)\|y_n - x_n\|$,
(VII) $\|x_{n+1} - x_n\| \leqslant f(a_n)\|y_n - x_n\|$,
(VIII) $\|x_{n+1} - x_0\| \leqslant f(a_0)\frac{1-c_0^{n+1}}{1-c_0}\|y_0 - x_0\| < R\eta$, *where* $R = \frac{f(a_0)}{1-c_0}$.

2.2. A semilocal convergence result and R-order of convergence two

Once the sequence $\{x_n\}$ is well defined, the next goal is to prove that $\{x_n\}$ is a Cauchy sequence and it is consequently convergent. To do this, we see that (2) is a Cauchy sequence and the condition $a_n < 2$ is satisfied, for all $n \geqslant 1$.

We first provide some properties that satisfy the sequences $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$.

Lemma 2.5. *Let f, g and h be the three scalar functions given in* (7) *respectively and define $\gamma = a_1/a_0$. If* (8) *is satisfied, then*

(i) $f(\gamma x) < f(x)$, $g(\gamma x) < g(x)$ *and* $h(\gamma x, \gamma y) < \gamma h(x, y)$, *for* $\gamma \in (0, 1)$, *with* $x \in (0, \frac{3-\sqrt{5}}{2})$ *and* $y \in (0, \frac{2(x^2-3x+1)}{A(2-x)})$,
(ii) $a_n < \gamma^{2^{n-1}} a_{n-1} < \gamma^{2^n-1} a_0$, $b_n < \gamma^{2^{n-1}} b_{n-1} < \gamma^{2^n-1} b_0$, *for all* $n \geqslant 2$ *and* $c_n < \gamma^{2^{n-1}} c_{n-1} < \gamma^{2^n-1} c_0 = \gamma^{2^n}/g(a_0)$, *for all* $n \geqslant 1$.

Proof. Item (i) is obvious, since f and g are increasing in $(0, \frac{3-\sqrt{5}}{2})$ and h is increasing in its first and second arguments for $(0, \frac{3-\sqrt{5}}{2})$ and $(0, \frac{2(x^2-3x+1)}{A(2-x)})$, respectively.

To prove (ii), invoke the induction hypothesis and use Lemma 2.1. As $a_1 = \gamma a_0$, we have $b_1 = \gamma b_0 \varphi(c_0) \leqslant \gamma b_0$, since $\varphi(c_0) \leqslant 1$, and $c_1 \leqslant \gamma c_0$. If we suppose that (ii) is true for $n = k$, then

$$\begin{aligned} a_{k+1} &= a_k g(a_k) c_k < \gamma^{2^{k-1}} a_{k-1} g\big(\gamma^{2^{k-1}} a_{k-1}\big) \gamma^{2^{k-1}} c_{k-1} \\ &< \gamma^{2^k} a_{k-1} g(a_{k-1}) c_{k-1} = \gamma^{2^k} a_k, \\ b_{k+1} &= b_k g(a_k) c_k \varphi(c_k) < b_k g(a_0) \gamma^{2^k - 1} c_0 \varphi(c_k) \leqslant \gamma^{2^k} b_k, \end{aligned}$$

since $\varphi(c_k) \leqslant 1$, and

$$c_{k+1} = f(a_{k+1}) g(a_{k+1}) h(a_{k+1}, b_{k+1}) < \gamma^{2^k} f(a_k) g(a_k) h(a_k, b_k) = \gamma^{2^k} c_k.$$

Moreover,

$$\begin{aligned} a_n &< \gamma^{2^{n-1}} a_{n-1} < \gamma^{2^{n-1}} \gamma^{2^{n-2}} a_{n-2} < \cdots < \gamma^{2^n - 1} a_0, \\ b_n &< \gamma^{2^{n-1}} b_{n-1} < \gamma^{2^{n-1}} \gamma^{2^{n-2}} b_{n-2} < \cdots < \gamma^{2^n - 1} b_0, \\ c_n &< \gamma^{2^{n-1}} c_{n-1} < \gamma^{2^{n-1}} \gamma^{2^{n-2}} c_{n-2} < \cdots < \gamma^{2^n - 1} c_0. \end{aligned}$$

The lemma is proved. □

We then provide the following semilocal convergence result, which is also used to draw conclusions about the existence of a solution and the domain in which it is located, along with some error estimates that lead to Halley's method converges with R-order of convergence at least two under conditions (C1)–(C5).

Theorem 2.6. *Let X and Y be two Banach spaces and $F : \Omega \subseteq X \to Y$ a twice Fréchet differentiable operator on a nonempty open convex domain Ω. We suppose that $\Gamma_0 = F'(x_0)^{-1} \in \mathcal{L}(Y, X)$ exists for some $x_0 \in \Omega$ and conditions* (C1)–(C5) *hold. Denote $a_0 = M\beta\eta$ and $b_0 = \beta\eta\omega(\eta)$, and suppose* (8)*. If $\overline{B(x_0, R\eta)} \subseteq \Omega$, where $R = \frac{f(a_0)}{1-c_0}$ and $c_0 = f(a_0) g(a_0) h(a_0, b_0)$, then the sequence $\{x_n\}$, defined in* (2) *and starting from x_0, converges to a solution x^* of Eq.* (1)*, the solution x^* and the iterates x_n, y_n belong to $\overline{B(x_0, R\eta)}$. Furthermore, the following error bounds are obtained*:

$$\|x^* - x_n\| \leqslant f(a_0) \eta \gamma^{2^n - 1} \frac{\Delta^n}{1 - \gamma^{2^n} \Delta}, \quad n \geqslant 0, \tag{9}$$

where $\gamma = a_1/a_0$ and $\Delta = 1/g(a_0)$.

Proof. Firstly, we prove that sequence (2) is a Cauchy one. From (II), we have

$$\|y_n - x_n\| \leqslant c_{n-1} \|y_{n-1} - x_{n-1}\| \leqslant \cdots \leqslant \left(\prod_{i=0}^{n-1} c_i \right) \|y_0 - x_0\|$$

and, by Lemma 2.5, it follows that

$$\prod_{i=0}^{n-1} c_i < \prod_{i=0}^{n-1} \gamma^{2^i} \Delta = \gamma^{2^n - 1} \Delta^n,$$

where $\gamma = a_1/a_0 < 1$ and $\Delta = 1/g(a_0) < 1$. In consequence, from $m \geqslant 1$,

$$\begin{aligned}\|x_{n+m} - x_n\| &\leqslant \|x_{n+m} - x_{n+m-1}\| + \|x_{n+m-1} - x_{n+m-2}\| + \cdots + \|x_{n+1} - x_n\| \\ &\leqslant f(a_{n+m-1})\|y_{n+m-1} - x_{n+m-1}\| + f(a_{n+m-2})\|y_{n+m-2} - x_{n+m-2}\| + \cdots \\ &\quad + f(a_n)\|y_n - x_n\| \\ &\leqslant f(a_n) \sum_{i=n}^{n+m-1} \left(\prod_{j=0}^{i-1} c_j \right) \|y_0 - x_0\| \\ &\leqslant f(a_0)\eta\gamma^{2^n-1}\Delta^n \frac{1-\gamma^{2^n(2^m-1)}\Delta^m}{1-\gamma^{2^n}\Delta},\end{aligned} \tag{10}$$

since $\gamma^{2^i+2^n} \geqslant \gamma^{2^{i+1}}$, for $i = n, n+1, \ldots, n+m-1$. In addition, $\{x_n\}$ converges to $x^* = \lim_n x_n$.

Obviously, $x_m \in B(x_0, R\eta)$, for all $m \geqslant 1$, as if $n = 0$ in (10), we obtain

$$\|x_m - x_0\| \leqslant f(a_0)\eta \frac{1-\gamma^{2^m-1}\Delta^m}{1-\gamma\Delta} < R\eta.$$

Following a similar procedure, we have $y_n \in B(x_0, R\eta)$, for all $n \geqslant 0$.

By letting now $n \to \infty$ in (II), it follows that $\|\Gamma_n F(x_n)\| \to 0$. Besides $\|F(x_n)\| \to 0$, since $\|F(x_n)\| \leqslant \|F'(x_n)\|\|\Gamma_n F(x_n)\|$ and $\{\|F'(x_n)\|\}$ is a bounded sequence. Therefore $F(x^*) = 0$ by the continuity of F in $\overline{B(x_0, R\eta)}$.

Finally, by letting $m \to \infty$ in (10), we obtain (9). □

Note that the following result on the R-order of Halley's method is clear from (9).

Corollary 2.7. *Under the conditions of Theorem* 2.6*, the Halley method is of R-order at least two.*

2.3. Uniqueness of the solution

Now we establish the uniqueness of the solution x^* of Eq. (1) by the next theorem.

Theorem 2.8. *Let us suppose conditions* (C1)–(C4) *hold. The solution x^* of Eq.* (1) *is unique in the region $B(x_0, \frac{2}{M\beta} - R\eta) \cap \Omega$.*

Proof. We assume z^* is another solution of (1) in $B(x_0, \frac{2}{M\beta} - R\eta) \cap \Omega$. Then, from

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

we have to prove that the operator $T = \int_0^1 F'(x^* + t(z^* - x^*))\, dt$ is invertible to obtain $x^* = z^*$. By the Banach lemma, we have to prove $\|I - T\| < 1$. Indeed,

$$\begin{aligned}\|I-T\| &\leqslant \|\Gamma_0\| \int_0^1 \big\|F'\big(x^*+t(z^*-x^*)\big)-F'(x_0)\big\|\,dt\\ &\leqslant M\beta \int_0^1 \big\|x^*+t(z^*-x^*)-x_0\big\|\,dt\\ &\leqslant M\beta \int_0^1 \big((1-t)\|x^*-x_0\|+t\|z^*-x_0\|\big)\,dt\\ &< \frac{M\beta}{2}\left(R\eta+\frac{2}{M\beta}-R\eta\right)=1.\end{aligned}$$

This completes the proof. □

3. On the *R*-order of convergence

Observe that for the operator (5) we have $\omega(z)=\sum_{i=1}^m L_i z^{p_i}$. In consequence, $\omega(tz)=\sum_{i=1}^m (L_i t^{p_i} z^{p_i})$, and then, $\varphi(t)=t^q$, where $q=\min\{p_1,p_2,\dots,p_m\}$, since $t\in[0,1]$ and $p_i\in[0,1]$, for all $i=1,2,\dots,m$. In this situation, $A=\frac{1}{(1+q)(2+q)}$ and the sequence $\{b_n\}$ is reduced to

$$b_n=b_{n-1}g(a_{n-1})c_{n-1}^{1+q},\quad n\geqslant 1.$$

Besides,

$$h(\gamma x,\gamma^{1+p}y)<\gamma^{1+p}h(x,y),\quad \text{for } \gamma\in(0,1),\ p\in[0,1],$$

with $x\in(0,\frac{3-\sqrt{5}}{2})$ and $y\in(0,2(1+q)(2+q)\frac{x^2-3x+1}{2-x})$. Hence, for all $n\geqslant 2$,

$$\begin{aligned}a_n &< \gamma^{(2+q)^{n-1}}a_{n-1}<\gamma^{\frac{(2+q)^n-1}{1+q}}a_0,\\ b_n &< \big(\gamma^{(2+q)^{n-1}}\big)^{1+q}b_{n-1}<\gamma^{(2+q)^n-1}b_0\end{aligned}$$

and, for all $n\geqslant 1$,

$$c_n<\gamma^{(2+q)^n-1}c_0=\gamma^{(2+q)^n}/g(a_0).$$

Therefore, we obtain new error bounds for the Halley's method

$$\|x^*-x_n\|\leqslant f(a_0)\eta\gamma^{\frac{(2+q)^n-1}{1+q}}\frac{\Delta^n}{1-\gamma^{(2+q)^n}\Delta},\quad n\geqslant 0,$$

from which we derive that the Halley sequence converges with R-order at least $2+q$, since

$$\|x^*-x_n\|\leqslant \frac{f(a_0)\eta}{\gamma^{\frac{1}{1+q}}(1-\Delta)}\big(\gamma^{\frac{1}{1+q}}\big)^{(2+q)^n},\quad n\geqslant 0.$$

Remark. Observe that if F'' is Lipschitz continuous in Ω, then F'' satisfies (3) and $\omega(z) = Kz$, $K \geqslant 0$, so that Halley's method is of R-order at least three. If F'' is (L, p)-Hölder continuous in Ω, then F'' satisfies (4), $\omega(z) = Lz^p$, $L \geqslant 0$, and the Halley process is of R-order at least $2 + p$.

References

[1] I.K. Argyros, The Halley method in Banach spaces and the Ptak error estimates, Rev. Acad. Cienc. Zaragoza (2) 52 (1997) 31–41.

[2] V. Candela, A. Marquina, Recurrence relations for rational cubic methods I: the Halley method, Computing 44 (1990) 169–184.

[3] D. Chen, On the convergence of the Halley method for nonlinear equation of one variable, Tamkang J. Math. 24 (1993) 461–467.

[4] D. Chen, I.K. Argyros, Q.S. Qian, A note on the Halley method in Banach spaces, Appl. Math. Comput. 58 (1993) 215–224.

[5] B. Döring, Einige Sätze über das Verfahren der tangierenden Hyperbeln in Banach–Räumen, Aplikace Mat. 15 (1975) 418–464.

[6] M. Ganesh, M.C. Joshi, Numerical solvability of Hammerstein integral equations of mixed type, IMA J. Numer. Anal. 11 (1991) 21–31.

[7] M.A. Hernández, M.A. Salanova, Indices of convexity and concavity: application to Halley method, Appl. Math. Comput. 103 (1999) 27–49.

[8] F.A. Potra, V. Pták, Nondiscrete Induction and Iterative Processes, Pitman, New York, 1984.

[9] R.A. Safiev, On some iterative processes, Ž. Vyčcisl. Mat. Fiz. 4 (1964) 139–143, translated into English by L.B. Rall as MRC Technical Summary Report, No. 649, Univ. Wisconsin-Madison, 1966.

[10] T.R. Scavo, J.B. Thoo, On the geometry of Halley's method, Amer. Math. Monthly 102 (1995) 417–426.

[11] T. Yamamoto, On the method of tangent hyperbolas in Banach spaces, J. Comput. Appl. Math. 21 (1988) 75–86.

[12] S. Zheng, D. Robbie, A note on the convergence of Halley's method for solving operator equations, J. Austral. Math. Soc. Ser. B 37 (1995) 16–25.