APPLICATIONES MATHEMATICAE
39,2 (2012), pp. 243–255

Ioannis K. Argyros (Lawton, OK)
Hongmin Ren (Hangzhou)

ON THE HALLEY METHOD IN BANACH SPACES

Abstract. We provide a semilocal convergence analysis for Halley's method using convex majorants in order to approximate a locally unique solution of a nonlinear operator equation in a Banach space setting. Our results reduce and improve earlier ones in special cases.

1. Introduction. In this study, we are concerned with the problem of approximating a locally unique solution $x^\star$ of the equation

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

where F is a twice Fréchet-differentiable operator defined on a nonempty open and convex subset Ω of a Banach space X with values in a Banach space Y.

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations [4, 11–14]. For example, dynamic systems are mathematically modeled by difference or differential equations and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time-invariant system is driven by the equation $\dot{x} = T(x)$, for some suitable operator T, where x is the state. Then the equilibrium states are determined by solving equation (1.1). Similar equations are used in the case of discrete systems. The unknowns of engineering equations can be functions (difference, differential and integral equations), vectors (systems of linear of nonlinear algebraic equations), or real or complex numbers (single algebraic equations with single unknowns). Except in special cases, the most commonly used solution methods are iterative—when starting from one or several initial approximations a sequence is constructed that converges to a solution of the equation.

2010 *Mathematics Subject Classification*: 65H10, 65J15, 65G99, 47H17, 47J05, 49M15.
Key words and phrases: Halley's method, Banach space, majorant functions, semilocal convergence.

DOI: 10.4064/am39-2-9

Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequences converge to an optimal solution of the problem at hand. Since all of these methods have the same recursive structure, they can be introduced and discussed in a general framework.

Newton's method

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \quad (n \geq 0,\, x_0 \in \Omega), \tag{1.2}$$

is undoubtedly the most famous quadratically convergent method for approximating $x^\star$. A survey on recent results for Newton-type methods

$$x_{n+1} = x_n - A(x_n)^{-1}F(x_n) \quad (n \geq 0,\, x_0 \in \Omega) \tag{1.3}$$

under very general Lipschitz conditions can be found in [4], and in the references there (see also [2], [3], [12], [15], [16]). Here, $A(x) \in L(X, Y)$, the space of bounded linear operators from X into Y. Note that if $A(x) = F'(x)$ $(x \in \Omega)$, then we obtain Newton's method (1.2). If, $A(x) = F'(x)(I - L_F(x))$ $(x \in \Omega)$, where $L_F(x) = \frac{1}{2}F'(x)^{-1}F''(x)F'(x)^{-1}F(x)$, then we obtain *Halley's method* [1, 4, 6, 10, 12]

$$x_{n+1} = x_n - [I - L_F(x_n)]^{-1}F'(x_n)^{-1}F(x_n) \quad (n \geq 0,\, x_0 \in \Omega). \tag{1.4}$$

Although sufficient convergence conditions and error estimates on the distances $\|x_{n+1} - x_n\|$, $\|x_n - x^\star\|$ for Newton's method or Halley's method have been given as special cases of Newton-type methods, a direct approach may yield a finer error analysis.

Ferreira and Svaiter [8] used the special majorant conditions

$$\|F'(x_0)^{-1}[F'(y) - F'(x)]\| \leq f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|) \tag{1.5}$$

$x, y \in U(x_0, R) = \{x \in X \mid \|x - x_0\| < R\}$ for $R > 0$, and

$$\|y - x\| + \|x - x_0\| < R$$

to provide an elegant Kantorovich-type semilocal convergence analysis for Newton's method (1.2). The corresponding local convergence analysis was given in [7]. The function $f : [0, R) \to (-\infty, \infty)$ is continuously differentiable satisfying $f(0) > 0$, having zeros in $(0, R)$, whereas f' is convex, strictly increasing with $f'(0) = -1$. The analysis provides a relationship between the majorizing function f and the nonlinear operator F. Moreover, the assumptions for the Q-quadratic convergence are relaxed. However, if we look at the sufficient convergence conditions in the special case when

$$f(t) = \frac{L}{2}t^2 - t + \eta, \qquad L > 0, \tag{1.6}$$

where

$$\|F'(x_0)^{-1}F(x_0)\| \leq f(0) = \eta, \tag{1.7}$$

then we get

$$h_K = L\eta \leq 1/2. \tag{1.8}$$

Condition (1.8) is the Newton–Kantorovich hypothesis for solving nonlinear equations [4, 12], famous for its simplicity and clarity. The corresponding majorizing sequence is given by

$$s_{n+1} = s_n + \frac{L(s_n - s_{n-1})^2}{2(1 - Ls_n)}, \qquad s_0 = 0,\ s_1 = \eta. \tag{1.9}$$

That is Ferreira–Svaiter's approach cannot extend the applicability of Newton's method even in the simplest possible case, where f is given by (1.6).

In our studies [2–5] we have shown that under the same hypotheses and computational cost, weaker sufficient conditions and tighter error bounds can be obtained for Newton's method (1.2). Indeed, in view of (1.5) there exists a function f_0 such that

$$\|F'(x_0)^{-1}[F'(y) - F'(x_0)]\| \leq f_0'(\|y - x_0\|) - f_0'(0) \tag{1.10}$$

for all $y \in U(x_0, R)$, where f_0 has the same properties as f but

$$f_0(t) \leq f(t), \qquad t \in (0, R], \tag{1.11}$$

and $f(t)/f_0(t)$ can be arbitrarily large [2–5]. Note that if

$$f_0(t) = L_0 t^2 - t + \eta, \qquad L_0 > 0, \tag{1.12}$$

then

$$L_0 \leq L.$$

The sufficient convergence condition corresponding to (1.8) is given by

$$h_A = \bar{L}\eta \leq 1/2, \tag{1.13}$$

where

$$\bar{L} = \frac{L + 4L_0 + \sqrt{L^2 + 8L_0 L}}{8}. \tag{1.14}$$

We have

$$h_K \leq 1/2 \ \Rightarrow\ h_A \leq 1/2 \tag{1.15}$$

but not necessarily vice versa unless $L_0 = L$. By dividing (1.13) by (1.8), we get

$$\frac{h_A}{h_K} = \frac{L + 4L_0 + \sqrt{L^2 + 8L_0 L}}{8L} \to \frac{1}{4} \qquad \text{as } L_0/L \to 0. \tag{1.16}$$

Hence, we deduce that our condition (1.13) can always replace the Kantorovich condition (1.8) and the applicability of Newton's method is expanded by at most 4 times. Our majorizing sequence for Newton's method

(1.2) is given by

$$t_{n+1} = t_n + \frac{L(t_n - t_{n-1})^2}{2(1 - L_0 t_n)}, \qquad t_0 = 0,\ t_1 = \eta. \tag{1.17}$$

Under condition (1.8), we have

$$t_n \le s_n, \tag{1.18}$$

$$0 \le t_{n+1} - t_n \le s_{n+1} - s_n, \tag{1.19}$$

$$t^\star = \lim_{n\to\infty} t_n \le s^\star = \lim_{n\to\infty} s_n. \tag{1.20}$$

Moreover, if $L_0 < L$, then strict inequality holds in (1.18) and (1.19) for $n \ge 1$. That is our upper bounds on the distances $\|x_{n+1} - x_n\|$, $\|x_n - x^\star\|$ $(n \ge 0)$ are tighter than the ones in [8], [7], [12], [14–[17]], [19].

Motivated by the above advantages of our technique over earlier ones for Newton's method, we extend our approach to Halley's method (1.4). In a way analogous to (1.5) and (1.10), we consider the majorant conditions

$$\|F'(x_0)^{-1}[F''(y) - F''(x)]\| \le f''(\|y-x\| + \|x-x_0\|) - f''(\|x-x_0\|) \tag{1.21}$$

for all $x, y \in U(x_0, R)$, where $\|y - x\| + \|x - x_0\| < R$ and $f : [0, R) \to (-\infty, \infty)$ is a twice continuously differentiable function.

Here, $F''(z)$ denotes a bilinear operator $X \times X \to Y$ for each fixed $z \in X$. For simplicity we use $F''(z)z_1z_2$, $F''(z)z_1^2$ instead of the notations $F''(z)(z_1, z_2)$, $F''(z)(z_1, z_1)$, respectively for all z_1 and $z_2 \in X$. Moreover, if $B : X \times X \to Y$ is a bilinear operator and $L : X \to X$ is a linear operator, then for $x \in X$, by BLx^2 we mean $B(L(x), x)$ (or $BLxx$). Furthermore, LBx^2 is used to denote $LB(x)(x)$.

We assume the following conditions hold:

- (H_1) $f(0) > 0$, $f''(0) > 0$, $f'(0) = -1$,
- (H_2) f'' is convex and strictly increasing in $[0, R)$,
- (H_2) f has zeros in $(0, R)$. Denote by $t^\star$ the minimal zero and assume $f'(t^\star) < 0$.

In view of (1.21) there exists a twice continuously differentiable function f_0 satisfying (H_1)–(H_3) (replacing f by f_0 and $t^\star$ by $t_0^\star$) with $t_0^\star \le t^\star$ such that

$$f_0(t) \le f(t), \tag{1.22}$$

$$\|F'(x_0)^{-1}[F''(y) - F''(x_0)]\| \le f_0''(\|y - x_0\|) - f_0''(0) \quad \text{for all } y \in U(x_0, R). \tag{1.23}$$

We shall also assume:

- (H_4) $-1/f_0'(t) \le -1/f'(t)$ for $t \in (0, t^\star)$,
- (H_5) $\|F'(x_0)^{-1}F(x_0)\| \le f(0)$, $\|F'(x_0)^{-1}F''(x_0)\| \le f''(0)$,
- (H_6) $U(x_0, R) \subseteq \Omega$.

The paper is organized as follows: Section 2 contains the semilocal convergence analysis of Halley's method (1.4) under conditions (H_1)–(H_6). Special cases and applications are given in Section 3.

2. Semilocal convergence analysis of Halley's method (1.4). It is convenient to define

$$H_F(x) = x - (I - L_F(x))^{-1}F'(x)^{-1}F(x), \qquad x \in \Omega, \tag{2.1}$$

$$H_f(t) = t - \frac{1}{1 - L_f(t)}\frac{f(t)}{f'(t)}, \qquad t \in [0, R), \tag{2.2}$$

$$H_{f_0}(t) = t - \frac{1}{1 - L_{f_0}(t)}\frac{f(t)}{f_0{}'(t)}, \qquad t \in [0, R), \tag{2.3}$$

where

$$L_f(t) = \frac{f(t)f''(t)}{2(f'(t))^2}, \tag{2.4}$$

$$L_{f_0}(t) = \frac{f(t)f''(t)}{2(f_0{}'(t))^2}, \tag{2.5}$$

$$s_0 = 0, \qquad s_{n+1} = H_f(s_n) \quad (n \geq 0), \tag{2.6}$$

$$t_0 = 0, \qquad t_{n+1} = H_{f_0}(t_n) \quad (n \geq 0). \tag{2.7}$$

Using (H_4), as in (1.18)–(1.20), a simple inductive argument shows

$$t_n \leq s_n, \tag{2.8}$$

$$0 \leq t_{n+1} - t_n \leq s_{n+1} - s_n, \tag{2.9}$$

$$t_0^\star = \lim_{n\to\infty} t_n \leq t^\star = \lim_{n\to\infty} s_n. \tag{2.10}$$

Inequalities (2.8) and (2.9) are strict if $-1/f_0'(t) < -1/f'(t)$.

LEMMA 2.1 ([11]). *If $p : [0, R) \to (-\infty, \infty)$ is continuously differentiable and convex, then*

(1) $(1 - \theta)p'(\theta t) \leq \frac{p(t)-p(\theta t)}{t} \leq (1 - \theta)p'(t)$ *for all* $t \in (0, R)$, $\theta \in [0, 1]$;
(2) $\frac{p(x)-p(\theta x)}{x} \leq \frac{p(y)-p(\theta y)}{y}$ *for all* $x, y \in [0, R)$, $x < y$, $\theta \in [0, 1]$.

LEMMA 2.2 ([11]). *Let $p : I \subseteq (-\infty, \infty) \to (-\infty, \infty)$ be convex. Then*

(1) *for any $z_0 \in \operatorname{int}(I)$ there exists the left derivative given by*

$$D^- p(z_0) := \lim_{z\to z_0^-} \frac{p(z_0) - p(z)}{z_0 - z} = \sup_{z<z_0} \frac{p(z_0) - p(z)}{z_0 - z};$$

(2) *if $x, y, z \in I$ and $x \leq y \leq z$, then*

$$p(y) - p(x) \leq [p(z) - p(x)]\frac{y - x}{z - x}.$$

Using the above standard results from convex analysis [4], [11], Halley's method [4] and majorant properties [4], [8], [7] we arrive at the following results on majorizing sequences for Halley's method:

Lemma 2.3. *Let f_0, $f : [0,R) \to (-\infty,\infty)$ be twice continuously differentiable functions satisfying (H_1)–(H_4). Then the following hold:*

1. (a) *f_0', f' are strictly convex and strictly increasing on $[0,R)$;*
   (b) *f_0, f are strictly convex on $[0,R)$, $f_0(t) > 0$ for $t \in [0,t_0^\star)$, $f(t) > 0$ for $t \in (0,t^\star)$, $f_0(t)$, $f(t)$ have at most one zero on $[t_0^\star,R)$, $[t^\star,R)$, respectively;*
   (c) *$-1 < f_0'(t) < 0$, $t \in (0,t_0^\star)$, $-1 < f'(t) < 0$, $t \in (0,t^\star)$.*
2. *$0 \le L_{f_0}(t) \le 1/4$ for all $t \in [0,t_0^\star]$, $0 \le L_f(t) \le 1/4$ for all $t \in [0,t^\star]$, $L_{f_0}(t) \le L_f(t)$ for all $t \in [0,t_0^\star]$.*
3. *$t < H_{f_0}(t) < t_0^\star$ for all $t \in [0,t_0^\star)$,*
   $f_0'(t_0^\star) < 0 \Leftrightarrow \exists t \in (t_0^\star,R)$ such that $f_0(t) < 0$,
   $t < H_f(t) < t^\star$ for all $t \in [0,t^\star)$,
   $f'(t^\star) < 0 \Leftrightarrow \exists t \in (t^\star,R)$ such that $f(t) < 0$.
4. *$t^\star - H_{f_0}(t) \le g_0(t^\star,t_0^\star)(t^\star - t)^3$, $\quad t \in [0,t_0^\star)$,*
   $t^\star - H_f(t) \le g(t^\star,t^\star)(t^\star - t)^3$, $\quad t \in [0,t^\star)$,
   $g_0(t^\star,t_0^\star) \le g(t^\star,t^\star)$,

where

$$g_0(u,v) = \frac{1}{3}\,\frac{f''(u)}{f_0'(v)^2} + \frac{2}{9}\,\frac{D^- f''(u)}{-f_0'(v)}, \qquad g(u,v) = \frac{1}{3}\,\frac{f''(u)}{f'(v)^2} + \frac{2}{9}\,\frac{D^- f''(u)}{-f'(v)}.$$

Moreover, the sequences $\{s_n\}$, $\{t_n\}$ converge to $t^\star$, $t_0^\star$ respectively with Q-cubic order.

Proof. We show the results for f. The results for f_0 follow in the same way.

1. Item (a) follows from (H_2) and $f''(0) > 0$ in (H_1). (a) implies that f is strictly convex. By (H_1), (a) and $f(t^\star) = 0$, we infer that $f(t) = 0$ has at most one zero in $(t^\star,R)$. In view of $f(t^\star) = 0$ and $f(0) > 0$, we obtain $f(t) > 0$ for $t \in [0,t^\star)$. To show (b) we have, by Lemma 2.1,

$$f'(t) < \frac{f(t^\star) - f(t)}{t^\star - t}, \qquad t \in [0,t^\star).$$

Hence, $0 = f(t^\star) > f(t) + f'(t)(t^\star - t)$. But $f(t) > 0$ on $[0,t^\star)$, so $f'(t) < 0$. Moreover, f' strictly increasing and $f'(0) = -1$ implies $f'(t) > -1$ for $t \in (0,t^\star)$. That completes the proof for (c).

2. Define

$$q(s) = f(t) + f'(t)(s-t) + \tfrac{1}{2}f''(t)(s-t)^2, \qquad s \in [t,t^\star].$$

By 1(b), we get $q(t) = f(t) > 0$ and

$$q(t^\star) = f(t) + f'(t)(t^\star - t) + \tfrac{1}{2} f''(t)(t^\star - t)^2. \tag{2.11}$$

We also have, by Taylor's formula,

$$\begin{aligned} f(t^\star) = f(t) &+ f'(t)(t^\star - t) + \tfrac{1}{2} f''(t)(t^\star - t)^2 \\ &+ \int_0^1 (1-\theta)[f''(t + \theta(t^\star - t)) - f''(t)](t^\star - t)^2 \, d\theta. \end{aligned} \tag{2.12}$$

Since $f(t^\star) = 0$ and f'' is increasing, by (2.11), (2.12) we have $q(t^\star) \le 0$. Hence there is a zero of q in $[t, t^\star]$. That is, the discriminant of q is nonnegative, so $f'(t)^2 - 2f''(t)f(t) \ge 0$, and hence $0 \le f''(t)f(t)/(f'(t))^2 \le 1/2$. That completes the proof.

3. For $t \in [0, t^\star)$, by the above we have $f(t) > 0$, $-1 < f'(t) < 0$ and $0 \le L_f(t) \le 1/4$, which imply $t < H_f(t)$. Using Lemma 2.2(1) and (H_2) we have $D^- f''(t) > 0$. That is,

$$D^- H_f(t) = \frac{f(t)^2[3f''(t)^2 - 2f'(t)D^- f''(t)]}{(f(t)f''(t) - 2f'(t))^2} > 0, \qquad t \in (0, t^\star].$$

Hence $H_f(t) < H_f(t^\star) = t^\star$ for $t \in [0, t^\star)$. Thus, the first part is shown.

For the second part we see that there exists $t \in (t^\star, R)$ such so $f(t) < 0$. Conversely, since $f(t^\star) = 0$, by Lemma 2.1 we get $f(t) > f(t^\star) + f'(t^\star)(t - t^\star)$ for $t \in [t^\star, R)$, so $f'(t^\star) < 0$.

By (H_3), $f'(t^\star) < 0$, so $f(t^{\star\star}) = 0$ for some $t^{\star\star} \in (t^\star, R)$. Moreover $f(t) < 0$ for some $t \in [t^\star, R)$.

A similar result follows for f_0. Hence, we have:

$$f_0'(t_0^\star) < 0 \;\Rightarrow\; \begin{cases} f_0(t_0^{\star\star}) = 0, & t_0^{\star\star} \in (t_0^\star, R), \\ f_0(t) < 0, & t \in (t_0^{\star\star}, R), \end{cases}$$

$$f'(t^\star) < 0 \;\Rightarrow\; \begin{cases} f(t^{\star\star}) = 0, & t^{\star\star} \in (t^\star, R), \\ f(t) < 0, & t \in (t^{\star\star}, R). \end{cases}$$

4. From the definition of H_f we have in turn:

$$\begin{aligned} t^\star - H_f(t) &= \frac{1}{1 - L_f(t)} \left[(1 - L_f(t))(t^\star - t) + \frac{f(t)}{f'(t)} \right] \\ &= -\frac{1}{f'(t)(1 - L_f(t))} \int_0^1 [f''(t + \theta(t^\star - t)) - f''(t)](t^\star - t)^2 (1 - t) \, d\theta \\ &\quad + \frac{t^\star - t}{2(1 - L_f(t))} \frac{f''(t)}{f'(t)^2} \int_0^1 f''(t + \theta(t^\star - t))(t^\star - t)^2 (1 - \theta) \, d\theta. \end{aligned}$$

By the convexity of f'' for $t < t^\star$ and Lemma 2.2(2) we get

$$f''(t+\theta(t^\star - t)) - f''(t) \le [f''(t^\star) - f''(t)]\frac{\theta(t^\star - t)}{t^\star - t}.$$

In view of the fact that f'' is strictly increasing we get

$$\begin{aligned} (2.13)\quad & t^\star - H_f(t) \\ &\le -\frac{f''(t^\star) - f''(t)}{6f'(t)(1-L_f(t))}(t^\star - t)^2 + \frac{f''(t^\star)f''(t)}{4(f'(t))^2(1-L_f(t))}(t^\star - t)^3 \\ &\le \frac{2}{9}\frac{f''(t^\star) - f''(t)}{-f'(t)}(t^\star - t)^2 + \frac{1}{3}\frac{f''(t^\star)^2}{f'(t^\star)^2}(t^\star - t)^3. \end{aligned}$$

Since $f'(t) < 0$, $f''(0) > 0$, the functions f', f'' are increasing on $[0, t^\star)$ and $0 \le L_f(t) \le 1/4$, we also have

$$\begin{aligned} (2.14)\qquad \frac{f''(t^\star) - f''(t)}{-f'(t)} &\le \frac{f''(t^\star) - f''(t)}{-f'(t^\star)} \\ &= \frac{f''(t^\star) - f''(t)}{-f'(t^\star)(t^\star - t)}(t^\star - t) \\ &\le \frac{D^- f''(t^\star)}{-f'(t^\star)}(t^\star - t) \end{aligned}$$

by Lemma 2.2(1).

The proof is finished by combining (2.13) and (2.14).

We also need some lemmas relating F to f_0 and f.

Lemma 2.4. *Assume:*

- *there exists $x_0 \in \Omega$ such that $F'(x_0)^{-1} \in L(Y, X)$;*
- *$f_0'(0) = -1$, $-1 < f_0'(t) < 0$ for $t \in (0, t_0^\star)$;*
- *$\|x - x_0\| \le t < t_0^\star$,*

where $f_0 : [0, t_0^\star] \to (-\infty, \infty)$ is twice continuously differentiable and satisfies (1.23). *Then $F'(x)^{-1} \in L(Y, X)$ and*

$$(2.15)\qquad \|F'(x)^{-1}F'(x_0)\| \le -\frac{1}{f_0'(\|x - x_0\|)} \le -\frac{1}{f_0'(t)}.$$

Proof. Let $x \in \overline{U}(x_0, t)$. Then we have by Taylor's formula

$$\begin{aligned} (2.16)\qquad F'(x) = F'(x_0) &+ \int_0^1 [F''(x_0 + \theta(x - x_0)) - F''(x_0)](x - x_0)\, d\theta \\ &+ F''(x_0)(x - x_0), \end{aligned}$$

so, in view of (1.23) and (2.16), we obtain in turn

$$\begin{aligned}(2.17)\quad &\|F'(x_0)^{-1}[F'(x)-F'(x_0)]\|\\ &\le \int_0^1 \|F'(x_0)^{-1}[F''(x_0+\theta(x-x_0))-F''(x_0)]\|\,\|x-x_0\|\,d\theta\\ &\quad+\|F'(x_0)^{-1}F''(x_0)\|\,\|x-x_0\|\\ &\le \int_0^1 [f_0''(\theta\|x-x_0\|)-f_0''(0)]\|x-x_0\|\,d\theta+f_0''(0)\|x-x_0\|.\end{aligned}$$

But $f_0'(0)=-1$ and $-1<f_0'(t)<0$ by hypotheses. Hence, (2.17) yields

$$\|F'(x_0)^{-1}[F'(x)-F'(x_0)]\|\le f_0'(t)-f_0'(0). \tag{2.18}$$

It follows from (2.18) and the Banach lemma on invertible operators [4], [12] that $F'(x)^{-1}\in L(Y,X)$ so that (2.15) is satisfied.

That completes the proof of the lemma.

REMARK 2.5. Clearly f can replace f_0 in Lemma 2.4. However, in view of (H_4), (2.15) is a tighter upper bound for $\|F'(x)^{-1}F'(x_0)\|$ than

$$\|F'(x)^{-1}F'(x_0)\|\le -\frac{1}{f'(\|x-x_0\|)}\le -\frac{1}{f'(t)},\qquad t\in(0,t^\star). \tag{2.19}$$

This observation leads to a more precise majorizing sequence $\{t_n\}$ (see (2.7)) than $\{s_n\}$ (see (2.6)) for Halley's method (1.4). The same observation leads to the advantages already stated in the introduction for Newton's method (1.2).

With the exception of the uniqueness part the following semilocal result for Halley's method (1.4) uses the standard proofs for this method [1, 4, 5, 9, 10], Lemmas 2.3, 2.4, and the formulae

$$\begin{aligned}(2.20)\quad F(x_{n+1})&=\frac{1}{2}F''(x_n)L_F(x_n)(x_{n+1}-x_n)^2\\ &\quad+\int_0^1[F''(x_n+\theta(x_{n+1}-x_n))-F''(x_n)](x_{n+1}-x_n)^2\,d\theta,\end{aligned}$$

$$\begin{aligned}(2.21)\quad x^\star-x_{n+1}&=-\mathit{\Gamma}_F(x_n)F'(x_n)^{-1}\int_0^1[F''(x_n^\theta)-F''(x_n)](x^\star-x_n)^2(1-\theta)\,d\theta\\ &+\frac{1}{2}\mathit{\Gamma}_F(x_n)F'(x_n)^{-1}F''(x_n)F'(x_n)^{-1}\int_0^1 F''(x_n^\theta)(x^\star-x_n)^2(1-\theta)\,d\theta(x^\star-x_n),\end{aligned}$$

where

$$\mathit{\Gamma}_F=(I-L_F(x))^{-1}, \tag{2.22}$$

$$x_n^\theta=x_n+\theta(x^\star-x_n). \tag{2.23}$$

Theorem 2.6. *Under conditions* (H_1)–(H_6), *the sequence* $\{x_n\}$ *generated by Halley's method* (1.4) *is well defined, remains in* $U(x_0, t^\star)$, *and converges to a solution* $x^\star \in \overline{U}(x_0, t^\star)$ *of* $F(x) = 0$. *Moreover, the following estimates hold for all* $n \geq 0$*:*

$$F'(x_n)^{-1} \in L(Y, X),$$

$$\|F'(x_n)^{-1}F'(x_0)\| \leq -\frac{1}{f_0'(\|x_n - x_0\|)} \leq -\frac{1}{f_0'(t_n)} \leq -\frac{1}{f'(s_n)},$$

$$\|F'(x_0)^{-1}F''(x_n)\| \leq f''(s_n),$$

$$\|F'(x_0)^{-1}F(x_n)\| \leq f(s_n),$$

$$(I - L_F(x_n))^{-1} \in L(Y, X),$$

$$\|(I - L_F(x_n))^{-1}\| \leq \frac{1}{1 - L_{f_0}(t_n)} \leq \frac{1}{1 - L_f(s_n)},$$

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n \leq s_{n+1} - s_n,$$

$$\|x_n - x^\star\| \leq t^\star - s_n,$$

$$\|x_{n+1} - x^\star\| \leq (t^\star - s_{n+1})\left(\frac{\|x_n - x^\star\|}{t^\star - s_n}\right)^3,$$

$$\|x_{n+1} - x^\star\| \leq g_0(t^\star, t_0^\star)\|x_n - x^\star\|^3 \leq g(t^\star, t^\star)\|x_n - x^\star\|^3.$$

Furthermore, $x^\star$ *is the unique solution of* $F(x) = 0$ *in* $\overline{U}(x_0, t^\star)$. *Finally,* $x^\star$ *is the unique solution of* $F(x) = 0$ *in* $U(x_0, \alpha)$, *where*

$$\alpha = \sup\{t \in [t_0^\star, R) : f_0(t) \leq 0\},$$

provided that $x^\star \in \overline{U}(x_0, t_0^\star)$, *and* (H_5) *holds with* f_0 *replacing* f.

Proof. We shall show $x^\star$ is the unique solution of $F(x) = 0$ in $\overline{U}(x_0, t^\star)$.

Let $y^\star$ be a solution of $F(x) = 0$ in $\overline{U}(x_0, t^\star)$. A simple induction (since $\|y^\star - x_0\| \leq t^\star$) shows

$$\|y^\star - x_n\| \leq t^\star - s_n \quad (n \geq 0). \tag{2.24}$$

Indeed, (2.24) holds for $n = 0$, since $s_0 = 0$. Assume (2.24) is true for all $n \leq k$. Then, from Lemma 2.3(4), (2.13), (2.15) and (2.21)–(2.23) we have

$$\|y^\star - x_{n+1}\| \leq (t^\star - s_{n+1})\left(\frac{\|y^\star - x_n\|}{t^\star - s_n}\right)^3, \tag{2.25}$$

which shows (2.24) for all $n \geq 0$. We see by (2.25) that $\lim_{n\to\infty} x_n = y^\star$ since $\lim_{n\to\infty} s_n = t^\star$. But $\lim_{n\to\infty} x_n = x^\star$. Hence, we deduce $x^\star = y^\star$.

Finally, we must show F has no zeros in $U(x_0, \alpha) - \overline{U}(x_0, t_0^\star)$. Let $z^\star$ be a solution of $F(x) = 0$ such that $t_0^\star < \|z^\star - x_0\| < \alpha$. We have the identity

$$F(z^\star) = F(x_0) + F'(x_0)(z^\star - x_0) + \frac{1}{2}F''(x_0)(z^\star - x_0)^2 + \int_0^1 [F''(x_0 + \theta(z^\star - x_0)) - F''(x_0)](z^\star - x_0)^2(1-\theta)\,d\theta.$$

Using (1.23) we obtain the estimates

$$\begin{aligned}\Big\|\int_0^1 & F'(x_0)^{-1}[F''(x_0 + \theta(z^\star - x_0)) - F''(x_0)](z^\star - x_0)^2(1-\theta)\,d\theta\Big\| \\ &\le \int_0^1 [f_0''(\theta\|z^\star - x_0\|) - f_0''(0)]\|z^\star - x_0\|^2(1-\theta)\,d\theta \\ &= f_0(\|z^\star - x_0\|) - f_0(0) - f_0'(0)\|z^\star - x_0\| - \tfrac{1}{2}f_0''(0)\|z^\star - x_0\|^2\end{aligned}$$

and

$$\begin{aligned}\|F'(x_0)^{-1}&[F(x_0) + F'(x_0)(z^\star - x_0) + \tfrac{1}{2}F''(x_0)(z^\star - x_0)^2]\| \\ &\ge \|z^\star - x_0\| - \|F'(x_0)^{-1}F(x_0)\| - \tfrac{1}{2}\|F'(x_0)^{-1}F''(x_0)\|\,\|z^\star - x_0\|^2 \\ &\ge \|z^\star - x_0\| - f_0(0) - \tfrac{1}{2}f_0''(0)\|z^\star - x_0\|^2.\end{aligned}$$

It then follows from the estimate

$$\begin{aligned}f_0(\|z^\star - x_0\|) - f_0(0) + \|z^\star - x_0\| - \tfrac{1}{2}f_0''(0)\|z^\star - x_0\|^2 \\ \ge \|z^\star - x_0\| - f_0(0) - \tfrac{1}{2}f_0''(0)\|z^\star - x_0\|^2,\end{aligned}$$

that $f_0(\|z^\star - x_0\|) \ge 0$. But f_0 is strictly positive on $(\|z^\star - x_0\|, R)$ (since f_0 is strictly convex). So, we have $\alpha \le \|z^\star - x_0\|$, which is a contradiction.

That completes the proof of the theorem.

3. Special cases and applications

Remark 3.1. Let $f_0 = f$. Moreover, if we assume the Lipschitz condition

$$\|F'(x_0)^{-1}[F''(y) - F''(x)]\| \le L\|y - x\| \quad \text{for all } x, y \in \Omega,$$

define

$$f(t) = \eta - t + \frac{M}{2}t^2 + \frac{L}{6}t^3,$$

where $\|F'(x_0)^{-1}F(x_0)\| \le \eta = f(0)$ and $\|F'(x_0)^{-1}F''(x)\| \le M = f''(0)$. Then it follows from Theorem 2.6 that the sufficient convergence condition reduces to the Kantorovich-type condition for Halley's method [1, 4, 6, 10, 12]

$$\eta \le \eta_0, \tag{3.1}$$

where

$$\eta_0 = \frac{2(M + 2\sqrt{M^2 + 2L})}{3(M + \sqrt{M^2 + 2L})^2}$$

That is, in this case we do not improve (3.1). However, if $f_0 < f$, then again under (3.1) we have convergence for Halley's method but the errors are tighter, and the information on the location of the solution is at least as precise.

Note that a direct study of the majorizing iteration $\{t_n\}$ in [4], [5] (see (2.7)) has led to sufficient convergence conditions which can be weaker than (3.1).

The results obtained here for special cases of f_0 and f can immediately produce Smale-type [18], [20] and Nemirovskii-type [13] theorems. We leave the details to the motivated reader.

Acknowledgments. This work was supported by National Natural Science Foundation of China (Grant No. 10871178), Natural Science Foundation of Zhejiang Province of China (Grant No. Y606154), and Scientific Research Fund of Zhejiang Provincial Education Department of China (Grant No. 20071362).

References

[1] I. K. Argyros, *The convergence of a Halley–Chebysheff-type method under Newton–Kantorovich hypotheses*, Appl. Math. Lett. 6 (1993), 71–74.

[2] —, *A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space*, J. Math. Anal. Appl. 298 (2004), 374–397.

[3] —, *On the Newton–Kantorovich hypothesis for solving equations*, J. Comput. Appl. Math. 169 (2004), 315–332.

[4] —, *Convergence and Application of Newton-Type Iterations*, Springer, New York, 2008.

[5] I. K. Argyros and S. Hilout, *On the semilocal convergence of the Halley method using recurrent functions*, J. Appl. Math. Comput. 37 (2011), 221–246.

[6] J. A. Ezquerro and M. A. Hernández, *New Kantorovich-type conditions for Halley's method*, Appl. Numer. Anal. Comput. Math. 2 (2005), 70–77.

[7] O. P. Ferreira, *Local convergence of Newton's method in Banach space from the viewpoint of the majorant principle*, IMA J. Numer. Anal. 29 (2009), 746–759.

[8] O. P. Ferreira and B. F. Svaiter, *Kantorovich's majorants principle for Newton's method*, Comput. Optim. Appl. 42 (2009), 213–229.

[9] J. M. Gutiérrez and M. A. Hernández, *A family of Chebyshev–Halley type methods in Banach spaces*, Bull. Austral. Math. Soc. 55 (1997), 113–130.

[10] M. A. Hernández, *A note on the Halley method*, Numer. Math. 59 (1991), 273–276.

[11] J.-B. Hilliart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, Part 1, Springer, Berlin, 1993.

[12] L. V. Kantorovich and G. P. Akilov, *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1982.

[13] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Stud. Appl. Math. 13, SIAM. Philadelphia, 1994.

[14] J. M. Ortega and W. C. Rheinbolt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[15] P. D. Proinov, *General local convergence theory for a class of iterative processes and its applications to Newton's method*, J. Complexity 25 (2009), 38–62.

[16] —, *New general convergence theory for iterative processes and its applications to Newton–Kantorovich type theorems*, ibid. 26 (2010), 3–42.

[17] W. C. Rheinboldt, *An adaptive continuation process for solving systems of nonlinear equations*, in: Banach Center Publ. 3, PWN–Polish Sci. Publ., 1975, 129–142.

[18] S. Smale, *Newton's method estimates from data at one point*, in: The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics, R. Ewing et al. (eds.), Springer, New York, 1986, 185–196.

[19] J. F. Traub and H. Woźniakowski, *Convergence and complexity of Newton iteration for operator equations*, J. Assoc. Comput. Mach. 26 (1979), 250–258.

[20] X. Wang, *Convergence of Newton's method and inverse function theorem in Banach space*, Math. Comp. 68 (1999), 169–186.

Ioannis K. Argyros
Department of Mathematics Sciences
Cameron University
Lawton, OK 73505, U.S.A.
E-mail: iargyros@cameron.edu

Hongmin Ren
College of Information and Electronics
Hangzhou Polytechnic
Hangzhou 31140
Zhejiang, P.R. China
E-mail: rhm65@126.com

Received on 26.4.2011;
revised version on 16.9.2011 (2085)