J. Austral. Math. Soc. Ser. B 37(1995), 16–25

A NOTE ON THE CONVERGENCE OF HALLEY'S METHOD FOR SOLVING OPERATOR EQUATIONS

SHIMING ZHENG[1] and DESMOND ROBBIE[2]

(Received 9 October 1992; revised 28 January 1994)

Abstract

Halley's method is a famous iteration for solving nonlinear equations. Some Kantorovich-like theorems have been given. The purpose of this note is to relax the region conditions and give another Kantorovich-like theorem for operator equations.

1. Introduction

Three hundred years ago Halley [6] presented a famous iteration method of order three for solving nonlinear equations. For real-valued functions, the method is usually written as

$$x_{k+1} = x_k - \frac{f(x_k)/f'(x_k)}{1 - \dfrac{1}{2}\dfrac{f(x_k) f''(x_k)}{f'(x_k)^2}}, \quad k = 0, 1, \dots .$$

This method is also called the method of tangent hyperbolas, as in Salehov and Mertvetsova [7], because x_{k+1} given above is the intercept with the x-axis of a hyperbola which is osculatory to the curve $y = f(x)$ at $x = x_k$. A number of papers have been written about Halley's method (for example, [1]–[7]). Davies and Dawson [5] showed that the convergence of Halley's method is monotonic global when applied to entire functions of genus 0 or 1, real for real arguments and having only real zeros. G. Alefeld has given the following theorem.

THEOREM A [1]. *Let $f(x)$ be a real-valued function of a real variable x, and let $f(x_0)f'(x_0) \neq 0$ for some x_0. Furthermore, let*

$$f'(x_0) - \frac{1}{2} f''(x_0) \frac{f(x_0)}{f'(x_0)} \neq 0.$$

[1]Department of Mathematics, Hangzhou University, Hangzhou 310028, China.
[2]Department of Mathematics, University of Melbourne, Parkville 3052, Australia.

Define

$$h_0 = -\frac{f(x_0)/f'(x_0)}{1 - \frac{1}{2}\frac{f(x_0)f''(x_0)}{f'(x_0)^2}},$$

and set

$$J_0 = \begin{cases} [x_0, x_0 + 2h_0], & h_0 > 0, \\ [x_0 + 2h_0, x_0], & h_0 < 0. \end{cases}$$

For $x \in J_0$, let f have a continuous third derivative. Suppose that f' does not change sign in J_0 and that with $g(x) = f(x)/\sqrt{f'(x)}$, we have

$$|g''(x)| \leq M_0$$

and

$$2|h_0|M_0 \leq |g'(x_0)|.$$

Then starting with x_0 the feasibility of Halley's method is guaranteed. All x_k are contained in J_0, and the sequence $\{x_k\}$ converges to a zero x^* of f (which is unique in J_0).

Cuyt and Rall [4] discussed the computational implementation of the multivariate Halley's method for solving nonlinear systems of equations. More generally, some authors have considered Halley's method for solving operator equations in a Banach space.

Let $f : D \subset X \to X$ be a twice Frechet differentiable map, where X is a Banach space and D is an open and convex set in X. In this case Halley's method for solving the operator equation

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

is of the form

$$\begin{cases} x_{k+1} &= x_k - Q_k^{-1} f'(x_k)^{-1} f(x_k), \\ Q_k &= I - \frac{1}{2} f'(x_k)^{-1} f''(x_k) f'(x_k)^{-1} f(x_k), \ k = 0, 1, \ldots \end{cases} \tag{2}$$

Let $S(x_0, r) = \{x : \|x - x_0\| < r\}$, $\overline{S}(x_0, r) = \{x : \|x - x_0\| \leq r\}$. For the case when X is a real or complex space and f is triple differentiable and satisfies

$$\|f''(x)\| \leq M_1, \ \ \|f'''(x)\| \leq N_1, \ \forall x \in D, \tag{3}$$

Salehov and Mertvetsova [7] and one of the authors [9] have given some Kantorovich-like convergence theorems.

THEOREM B [7]. *Suppose that* $x_0 \in D$ *satisfies the following conditions:*

(1) $\|f'(x_0)^{-1}\| \le B_0$, $\|x_1 - x_0\| \le \eta_0'$;
(2) $h_0' = M_1 B_0 \eta_0' \le \frac{1}{2}$;
(3) $\overline{S}(x_0, 2\eta_0') \subset D$;
(4) $R_0 = [\frac{N_1}{M_1^2 B_0}(2 + h_0') + 3](1 + h_0') \le 9$.

Then there is a solution x^* *of (1) in* $\overline{S}(x_0, 2\eta_0')$ *and the Halley iteration converges to* x^* *and satisfies the estimation*

$$\|x^* - x_k\| \le (2h_0')^{3^k - 1}\eta_0'/2^{k-1}, \quad k = 1, 2, \dots .$$

THEOREM C [9]. *Suppose that* $x_0 \in D$ *satisfies the following conditions:*

(1) $\|f'(x_0)^{-1}\| \le B_0$, $\|f'(x_0)^{-1} f(x_0)\| \le \eta_0$;
(2) $h_0 = K B_0 \eta_0 \le \frac{1}{2}$ *with*

$$K = \sqrt{M_1^2 + \frac{2N_1}{3B_0(1 - \frac{1}{2}M_1 B_0 \eta_0)}};$$

(3) $S(x_0, (1 + \theta_0)\eta_0) \subset D$, *where*

$$\theta_0 = (1 - \sqrt{1 - 2h_0})/(1 + \sqrt{1 + 2h_0}).$$

Then we conclude that

(i) $x_k \in S(x_0, (1 + \theta_0)\eta_0) (k = 0, 1, \dots)$;
(ii) $\lim_{k\to\infty} x_k = x^* \in \overline{S}(x_0, (1 + \theta_0)\eta_0)$, $f(x^*) = 0$;
(iii) $\|x^* - x_k\| \le (1 + \theta_0)\eta_0 \theta_0^{3^k - 1} / \sum_{i=0}^{3^k - 1} \theta_0^i \quad k = 1, 2, \dots .$

Recently, Chen, Argyros and Qian [3] also gave a result on the convergence of Halley's method, in which the conditions and the conclusions are almost all the same as those of Theorem C except for the constant

$$K = \sqrt{M_1^2 + \frac{2N_1}{3B_0}}$$

instead of

$$K = \sqrt{M_1^2 + \frac{2N_1}{3B_0(1 - \frac{1}{2}M_1 B_0 \eta_0)}}$$

in the condition (2) of Theorem C.

The conditions of Theorems B and C are not comparable. For example, the function

$$f_1(x) = x^2 + 4x - 5, \quad x_0 = 0$$

satisfies the conditions of Theorem B but not Theorem C, and for the function

$$f_2(x) = x^2 - 3x + 2, \quad x_0 = 0$$

the opposite is true.

Several years ago, Smale presented first a new concept of point estimation theory. Instead of the region conditions in the Newton-Kantorovich theorem, he obtained the convergence of Newton's method for analytic maps from the data at one point. A Smale-like theorem for Halley's method has been given as follows:

THEOREM D [10]. *Suppose that $f : X \to X$ is an analytic map, where X is a real or complex Banach space. Let*

$$\alpha = \alpha(x, f) = \beta\gamma,$$

where $x \in X$,

$$\beta = \beta(x, f) = ||f'(x)^{-1} f(x)||$$

and

$$\gamma = \gamma(x, f) = \sup_{k \geq 2} \left\| f'(x)^{-1} \frac{f^{(k)}(x)}{k!} \right\|^{\frac{1}{k-1}}.$$

If $\alpha(x, f) \leq 3 - 2\sqrt{2}$, then Halley's iteration (2) with $x_0 = x$ is well defined and there is a limit $\lim_{k\to\infty} x_k = x^*$ such that $f(x^*) = 0$. The constant $3 - 2\sqrt{2}$ is the best possible.

Moreover, the optimal error estimation has been given as well in Zheng [10].

In this note we give another Kantorovich-like theorem for operator equations. We prove the following theorem.

THEOREM. *Suppose that $f : D \subset X \to X$ satisfies*

$$\|f'(x_0)^{-1}[f''(x) - f''(y)]\| \leq N\|x - y\|, \ \forall x, y \in D \tag{4}$$

and that $x_0 \in D$ satisfies the following conditions:

$$\|f'(x_0)^{-1} f(x_0)\| \leq \eta, \ \|f'(x_0)^{-1} f''(x_0)\| \leq M, \tag{5}$$

$$9N^2\eta^2 + 18NM\eta + 6M^3\eta \leq 3M^2 + 8N, \tag{6}$$

$$S = \{x : \|x - x_0\| \leq t_1^*\} \subset D,$$

where t_1^* is the smaller positive zero of the real function

$$\varphi(t) = \frac{N}{6}t^3 + \frac{M}{2}t^2 - t + \eta. \tag{7}$$

https://doi.org/10.1017/S0334270000007542 Published online by Cambridge University Press

Then the sequence $\{x_k\}$ generated by Halley's iteration is well defined, remains in S and converges to a solution x^* of (1) which satisfies the error estimation

$$\|x^* - x_k\| \le t_1^* - t_k,\ k = 0, 1, \dots,$$

where $\{t_k\}$ is the sequence produced by Halley's iteration for $\varphi(t)$ with $t_0 = 0$.

We see that in Theorems A, B and C the region conditions are concerned with the second and third derivatives and the conditions about the initial point only concern $f(x)$ and its first derivative. However, in Theorem D, there is no region condition but the data of all higher derivatives at the initial point are required. The conditions of the theorem of this note are between these two sets of conditions. Our theorem has a region condition concerning only the third but not the second derivative, and the initial point conditions concern $f(x_0)$ and its first and second derivatives but not any higher ones.

There are some maps and initial points such that the conditions of our theorem are satisfied but those of Theorems B or C are not.

EXAMPLE 1. Let $X = R$, $D = (-1, 1)$, $x_0 = 0$, $f_3(x) = \frac{x^3}{5} + \frac{x^2}{5} - x + \frac{2}{5}$.

If we consider Theorem B of Salehov and Mertvetsova with regard to example 1, then

$$M_1 = \frac{8}{5}, N_1 = \frac{6}{5}, \eta_0' = \frac{10}{23}, B_0 = 1, h_0' = \frac{16}{23} > \frac{1}{2}.$$

This means that the conditions of Theorem B are not satisfied. If we consider instead Theorem C of Zheng [9], or the result of Chen *et al.* [3], with regard to Example 1, then

$$\eta_0 = \frac{2}{5}, \quad K \ge M_1 = \frac{8}{5}, \quad h_0 \ge \frac{16}{25} > \frac{1}{2}.$$

Thus the conditions there are not satisfied either. However, considering our theorem in regard to Example 1, we have

$$\eta = M = \frac{2}{5}, \quad N = \frac{6}{5},$$

so that

$$9N^2\eta^2 + 18NM\eta + 6M^2\eta = 9.8304 < 3M^2 + 8N = 12.48.$$

This shows that the conditions of our theorem are satisfied. Thus the theorem of this note applies and the earlier results quoted do not.

Example 2 shows that the result of our theorem is not contained in Theorem D above.

EXAMPLE 2. Let $X = R$, $D = (-1, 1)$, $x_0 = 0$, $f_4(x) = x^3 + x^2 - x + 0.2$.

Considering Theorem D in regard to Example 2, we have that

$$\gamma(0, f_4) = 1, \quad \alpha(0, f_4) = \beta(0, f_4) = 0.2 > 3 - 2\sqrt{2} = 0.17157.$$

Thus the conditions of Theorem D are not satisfied. However, our theorem does apply to Example 2, because we have

$$M = 2, \quad N = 1, \quad \eta = 0.2,$$

giving

$$9N^2\eta^2 + 18NM\eta + 6M^3\eta = 17.16 < 3M^2 + 8N = 20.$$

2. Some lemmas

We give some lemmas to prove the theorem.

LEMMA 1. *Under the condition (5), the function $\varphi(t)$ defined by (6) has three real zeros.*

PROOF. From (6) we see that

$$\varphi'(t) = \frac{N}{2}t^2 + Mt - 1,$$

has two real zeros

$$t_+ = \frac{-M + \sqrt{M^2 + 2N}}{N} \geq 0 \quad \text{and} \quad t_- = \frac{-M - \sqrt{M^2 + 2N}}{N} \leq 0.$$

Since $\varphi'(t_+) = 0$, $t_+^2 = \frac{2(1-Mt_+)}{N}$. Also

$$\begin{aligned}\varphi(t) &= \frac{t}{3}[\frac{N}{2}t^2 + \frac{3}{2}Mt - 3] + \eta \\ &= \frac{t}{3}[\varphi'(t) + \frac{M}{2}t - 2] + \eta,\end{aligned}$$

so

$$\begin{aligned}\varphi(t_+) &= \frac{M}{6}t_+^2 - \frac{2}{3}t_+ + \eta = \frac{M}{3N}(1 - Mt_+) - \frac{2}{3}t_+ + \eta \\ &= \frac{M}{3N} - \frac{2N + M^2}{3N}t_+ + \eta \\ &= \frac{1}{3N^2}[3NM + M^3 + 3N^2\eta - (2N + M^2)^{3/2}].\end{aligned}$$

https://doi.org/10.1017/S0334270000007542 Published online by Cambridge University Press

Under the condition (5), it is easy to prove that

$$(3NM + M^3 + 3N^2\eta)^2 \le (2N + M^2)^3.$$

Therefore $\varphi(t_+) \le 0$. Finally, it is clear that $\varphi(-\infty) = -\infty$, $\varphi(0) > 0$, $\varphi(\infty) = \infty$, and so the lemma is proved.

The following lemma 2 can be proved using the result of Davies and Dawson [5].

LEMMA 2. *Under the condition [6], the sequence $\{t_k\}$ produced by Halley's iteration for $\varphi(t)$,*

$$\begin{cases} t_0 &= 0, \\ t_{k+1} &= t_k - \dfrac{\varphi(t_k)/\varphi'(t_k)}{1 - \dfrac{1}{2}\dfrac{\varphi(t_k)\varphi''(t_k)}{\varphi'(t_k)^2}} \end{cases}$$

is monotonic increasing and convergent to the smaller positive zero t_1^* of $\varphi(t)$.

LEMMA 3. *Under the conditions of the theorem, if $\|x - x_0\| < t_+$, where $t_+ > 0$ is the positive zero of $\varphi'(t)$, then the inverse $f'(x)^{-1}$ exists and*

$$\|f'(x)^{-1} f'(x_0)\| \le -1/\varphi'(\|x - x_0\|),$$

$$\|f'(x_0)^{-1} f''(x)\| \le \varphi''(\|x - x_0\|).$$

PROOF. From the proof of Lemma 1 we see that, when $\|x - x_0\| < t_+$,

$$\varphi'(\|x - x_0\|) = \frac{N}{2}\|x - x_0\|^2 + M\|x - x_0\| - 1 < 0.$$

Thus, under the conditions of the theorem, we have

$$\|f'(x_0)^{-1} f''(x_0)(x - x_0) + \int_0^1 f'(x_0)^{-1}[f''(x_0 + t(x - x_0)) - f''(x_0)]\,dt(x - x_0)\|$$

$$\le M\|x - x_0\| + \frac{N}{2}\|x - x_0\|^2 < 1.$$

From the Mean Value Theorem we obtain

$$\begin{aligned} f'(x) &= f'(x_0) + \int_0^1 f''(x_0 + t(x - x_0))dt(x - x_0), \\ f'(x_0)^{-1} f'(x) &= I + f'(x_0)^{-1} f''(x_0)(x - x_0) \\ &\quad + \int_0^1 f'(x_0)^{-1}[f''(x_0 + t(x - x_0)) - f''(x_0)]dt(x - x_0). \end{aligned}$$

By the Neumann Lemma, the inverse $[f'(x_0)^{-1}f'(x)]^{-1} = f'(x)^{-1}f'(x_0)$ exists and

$$\|f'(x)^{-1}f'(x_0)\| \le \frac{1}{1 - M\|x - x_0\| - \frac{N}{2}\|x - x_0\|^2} = -\frac{1}{\varphi'(\|x - x_0\|)}.$$

It is clear that

$$\begin{aligned}\|f'(x_0)^{-1}f''(x)\| &\le \|f'(x_0)^{-1}f''(x_0)\| + \|f'(x_0)^{-1}[f''(x_0) - f''(x)]\| \\ &\le M + N\|x - x_0\| = \varphi''(\|x - x_0\|).\end{aligned}$$

And so the lemma is proved.

3. The proof of the theorem

We want to prove that for all nonnegative integers k, $x_k \in S = \{x : \|x - x_0\| \le t_1^*\}$, $f'(x_k)^{-1}$ exists and

$$\|f'(x_0)^{-1}f(x_k)\| \le \varphi(t_k), \tag{11}$$

$$\|f'(x_k)^{-1}f'(x_0)\| \le \frac{-1}{\varphi'(t_k)}, \tag{12}$$

$$\|f'(x_0)^{-1}f''(x_k)\| \le \varphi''(t_k), \tag{13}$$

$$\|x_{k+1} - x_k\| \le t_{k+1} - t_k. \tag{14}$$

First we show that (14) must hold when (11)–(13) hold. In fact, from (11)–(13) we obtain

$$\|f'(x_k)^{-1}f(x_k)\| \le \|f'(x_k)^{-1}f'(x_0)\|\|f'(x_0)^{-1}f(x_k)\| \le -\frac{\varphi(t_k)}{\varphi'(t_k)} \tag{15}$$

and

$$\|f'(x_k)^{-1}f''(x_k)\| \le \|f'(x_k)^{-1}f'(x_0)\|\|f'(x_0)^{-1}f''(x_k)\| \le -\frac{\varphi''(t_k)}{\varphi'(t_k)}. \tag{16}$$

Hence, using the monotonicity of $\{t_k\}$ from Lemma 2, we have

$$\frac{1}{2}\|f'(x_k)^{-1}f''(x_k)f'(x_k)^{-1}f(x_k)\| \le \frac{1}{2}\frac{\varphi(t_k)\varphi''(t_k)}{\varphi'(t_k)^2} < 1.$$

By the Neumann Lemma, the inverse of

$$Q_k = I - \frac{1}{2}f'(x_k)^{-1}f''(x_k)f'(x_k)^{-1}f(x_k)$$

https://doi.org/10.1017/S0334270000007542 Published online by Cambridge University Press

exists and

$$\|Q_k^{-1}\| \le \frac{1}{1 - \dfrac{1}{2}\dfrac{\varphi(t_k)\varphi''(t_k)}{\varphi'(t_k)^2}}.$$

Then (14) follows from (2) and (15).

Now we prove (11)–(13). It is clear that they hold for $k = 0$. And then as we have seen, (14) must hold. Suppose that (11)–(13) are true for $k \le n$. Then

$$\|x_{n+1} - x_0\| \le t_{n+1} - t_0 = t_{n+1} < t_1^* \le t_+,$$

that is, $x_{n+1} \in S$. From Lemma 3 we see that $f'(x_{n+1})^{-1}$ exists and (12), (13) hold for $k = n + 1$. From (2) we have

$$\begin{aligned} 0 &= f(x_n) + f'(x_n)(x_{n+1} - x_n) - \frac{1}{2} f''(x_n) f'(x_n)^{-1} f(x_n)(x_{n+1} - x_n) \\ &= f(x_n) + f'(x_n)(x_{n+1} - x_n) + \frac{1}{2} f''(x_n)(x_{n+1} - x_n)^2 \\ &\quad - \frac{1}{2} f''(x_n) f'(x_n)^{-1}[f(x_n) + f'(x_n)(x_{n+1} - x_n)](x_{n+1} - x_n) \\ &= f(x_n) + f'(x_n)(x_{n+1} - x_n) + \frac{1}{2} f''(x_n)(x_{n+1} - x_n)^2 \\ &\quad - \frac{1}{4} f''(x_n) f'(x_n)^{-1} f''(x_n) f'(x_n)^{-1} f(x_n)(x_{n+1} - x_n)^2. \end{aligned}$$

Therefore

$$f(x_{n+1}) = A_n + B_n,$$

where

$$\begin{aligned} A_n &= f(x_{n+1}) - f(x_n) - f'(x_n)(x_{n+1} - x_n) - \frac{1}{2} f''(x_n)(x_{n+1} - x_n)^2 \\ &= \int_0^1 (1 - u)[f''(x_n + u(x_{n+1} - x_n)) - f''(x_n)]\, du (x_{n+1} - x_n)^2, \\ B_n &= \frac{1}{4} f''(x_n) f'(x_n)^{-1} f''(x_n) f'(x_n)^{-1} f(x_n)(x_{n+1} - x_n)^2. \end{aligned}$$

It can be shown in a similar way that

$$\varphi(t_{n+1}) = \frac{N}{6}(t_{n+1} - t_n)^3 - \frac{1}{4}\frac{\varphi''(t_n)^2 \varphi(t_n)}{\varphi'(t_n)^3}(t_{n+1} - t_n).$$

When $u \in [0, 1]$,

$$\|x_n + u(x_{n+1} - x_n) - x_0\| \le t_n + u(t_{n+1} - t_n) \le t_{n+1} < t_1^* \le t_+.$$

From Lemma 3 we have

$$\|f'(x_0)^{-1}A_n\| \le N\int_0^1 (1-u)u\,du(t_{n+1}-t_n)^3 = \frac{N}{6}(t_{n+1}-t_n)^3, \tag{17}$$

$$\|f'(x_0)^{-1}B_n\| \le -\frac{1}{4}\frac{\varphi''(t_n)^2\varphi(t_n)}{\varphi'(t_n)^3}(t_{n+1}-t_n). \tag{18}$$

Hence

$$\|f'(x_0)^{-1}f(x_{n+1})\| \le \varphi(t_{n+1}),$$

that is, (11) is true for $k = n+1$. Thus (11)-(13) hold for all $k = 0, 1, \ldots$, and so (14) also holds. Therefore, the limit, $\lim_{k\to\infty} x_k = x^*$, exists. Letting $k \to \infty$ in (11), we obtain $f(x^*) = 0$. The error estimation follows from (14) and the proof of the theorem is completed.

Acknowledgements

The authors wish to thank the referee of the paper for many very helpful comments and are also most grateful to Professor Sloan, Associate Editor, for his kind suggestions and for forwarding the reference [3].

The project was Supported by National Natural Science Foundation of China and by Natural Science Foundation of Zhejiang Province.

References

[1] G. Alefeld, "On the convergence of Halley's method", *Amer. Math. Monthly* **88** (1981) 530–536.

[2] G. N. Brown, "On Halley's variation of Newton's method", *Amer. Math. Monthly* **84** (1977) 726–728.

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

[4] A. Cuyt and L. B. Rall, "Computational implementation of the multivariate Halley method for solving nonlinear systems of equations", *ACM Trans. on Math. Softw.* **11** (1985) 20–36.

[5] M. Davies and B. Dawson, "On the global convergence of Halley's iteration formula", *Numer. Math.* **24** (1975) 133–135.

[6] E. Halley, "Methodus nova, accurata & facilis inveniendi radices aequationum quarumcumque generaliter, sine praevia reductione", *Philos. Trans. Roy. Soc. London* **18** (1694) 136–148.

[7] G. S. Salehov and M. A. Mertvetsova, "On the convergence of some iterative processes", *Izv. Kazansk. Phil. ANSSSR, Ser. Phiz-Matem. i Techn.* **5** (1954) 77–108, (in Russian).

[8] S. Smale, "Newton's method estimates from data at one point", in *Proc. of a conf. in Honor of Gail Young*, (Laramie, New York, 1986), 1–16.

[9] S. Zheng, "The convergence of Halley's method and its optimal error bounds", *J. Hangzhou University* **9** (1982) 285–289, (in Chinese).

[10] S. Zheng, "Point estimates for Halley's iteration", *Acta Math. Appl. Sinica* **14** (1991) 376–383, (in Chinese).