ON THE CONVERGENCE OF HALLEY'S METHOD

G. ALEFELD
Fachbereich Mathematik, Technische Universität Berlin
Strasse des 17. Juni 135, 1 Berlin 12, West Germany

1. Introduction. A number of papers have been written about Halley's method, a third-order method for the solution of a nonlinear equation. (See, for example, [8].) For real-valued functions, this method is usually written as

\[ x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k) - \frac{1}{2}f''(x_k)\frac{f(x_k)}{f'(x_k)}}, \quad k \geq 0. \] (0)

This method is also called the method of tangent hyperbolas, as in [3], because \( x_{k+1} \) as given by (0) is the intercept with the x-axis of a hyperbola that is osculatory to the curve \( y = f(x) \) at \( x = x_k \). Construction of the appropriate hyperbola, given \( f(x_k), f'(x_k), \) and \( f''(x_k) \), is an interesting exercise.

Many of the authors writing on Halley's method have, in particular, been concerned with developing a convergence theorem similar to the so-called Newton-Kantorovich theorem. (See, for example, [5, p. 421 ff.]) The most far-reaching results can undoubtedly be found in [3], where also a comprehensive list of references is given. Error bounds, also, are given in [3]. These reflect for the first time the correct order of the error.

In this note we add some new results on Halley's method for real functions. From a remark by G. H. Brown, Jr. [2], Halley's method can be derived by applying Newton's method to the function \( g(x) = f(x)/\sqrt{f'(x)} \). The adaptation of Theorem 7.1 of Ostrowski [7] to Halley's method gives us results on the existence and uniqueness of a zero and on the convergence to this
zero. In particular, we get error bounds that reflect correctly the order of the error. This method is demonstrated by a simple well-tried numerical example.

2. Theoretical Results.

**Theorem.** Let \( f(x) \) be a real-valued function of the real variable \( x \), and let \( f(x_0)f'(x_0) \neq 0 \) for some \( x_0 \). Furthermore let

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

Define

\[
h_0 = -\frac{f(x_0)}{f'(x_0) - \frac{1}{2} f''(x_0) \frac{f(x_0)}{f'(x_0)}}, \quad x_1 = x_0 + h_0,
\]

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 \( f' \) doesn't change sign in \( J_0 \) and that with

\[
g(x) = \frac{f(x)}{\sqrt{f'(x)}}
\]

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 \( J_0 \)).

Defining

\[
h_k = -\frac{f(x_k)}{f'(x_k) - \frac{1}{2} f''(x_k) \frac{f(x_k)}{f'(x_k)}}, \quad k \geq 0,
\]

\[
J_k = \begin{cases} [x_k, x_k + 2h_k], & h_k > 0 \\ [x_k + 2h_k, x_k], & h_k < 0\end{cases}, \quad k \geq 0,
\]

we have the error estimates

\[
|x_{k+1} - x_k| \leq \frac{1}{2} \frac{M_k^{-1}}{|g'(x_k)|} |x_k - x_{k-1}|^2, \quad k \geq 1,
\]

\[
|x^* - x_{k+1}| \leq \frac{1}{2} \frac{M_k^{-1}}{|g'(x_k)|} |x_k - x_{k-1}|^2, \quad k \geq 1,
\]

\[
|x^* - x_k| \leq \frac{M_k^{-1}}{|g'(x_k)|} |x_k - x_{k-1}|^2, \quad k \geq 1.
\]
If
\[ |f'(x)| \leq N, \quad x \in J_0, \]
and if
\[ \frac{1}{2} \frac{1}{f'(x)} \left| \frac{f''(x)}{f'(x)} + \frac{3}{2} \left( \frac{f''(x)}{f'(x)} \right)^2 \right| \leq M, \quad x \in J_0, \]
then we have the coarser estimates
\[ |x_{k+1} - x_k| \leq \frac{NM}{|g'(x_k)|} |x_k - x_{k-1}|^3, \quad k \geq 1, \quad (3') \]
\[ |x^* - x_{k+1}| \leq \frac{NM}{|g'(x_k)|} |x_k - x_{k-1}|^3, \quad k \geq 1, \quad (4') \]
\[ |x^* - x_k| \leq \frac{2NM}{|g'(x_k)|} |x_k - x_{k-1}|^3, \quad k \geq 1. \quad (5') \]

Proof. In some essential parts the proof is similar to that of Theorem 7.1 in [7]. Because of \( f'(x_0) \neq 0 \), we can assume that \( f'(x_0) > 0 \). We then consider the function
\[ g(x) = -\frac{f(x)}{\sqrt{f'(x)}}, \quad x \in J_0. \]
For \( x \in J_0 \), \( g \) has a continuous second derivative and we have
\[ g'(x) = \sqrt{f'(x)} - \frac{1}{2} \frac{f''(x)f(x)}{f'(x)^{3/2}}, \]
from which
\[ g''(x) = \frac{1}{2} g(x) \left[ -\frac{f''(x)}{f'(x)} + \frac{3}{2} \left( \frac{f''(x)}{f'(x)} \right)^2 \right]. \quad (8) \]
Furthermore
\[ |g'(x) - g'(x_0)| = \left| \int_{x_0}^{x} g''(t) \, dt \right| \leq |x - x_0| M_0, \]
and, using (2), we have
\[ |g'(x_1) - g'(x_0)| \leq |x_1 - x_0| M_0 = |h_0| M_0 \leq \frac{|g'(x_0)|}{2}, \quad (9) \]
from which we may estimate
\[ |g'(x_1)| \geq |g'(x_0)| - |g'(x_0) - g'(x_1)| \geq \frac{|g'(x_0)|}{2}. \quad (10) \]
Integrating by parts and using
\[ h_0 = -\frac{g(x_0)}{g'(x_0)}, \]
we get
\[ \int_{x_0}^{x_1} (x_1 - x) g''(x) \, dx = g(x_1) \]
and therefore

$$|g(x_1)| \leq \frac{1}{2} |h_0|^2 M_0.$$  \hspace{1cm} (11)

Because of (10), the feasibility of Halley's method (0) is guaranteed for \( k = 1 \) and we have

$$x_2 = x_1 + h_1.$$ 

Since

$$h_1 = - \frac{g(x_1)}{g'(x_1)}$$

we have from (10) and (11)

$$|h_1| \leq \frac{|h_0|^2 M_0}{|g'(x_0)|}.$$ 

Finally in a way similar to that in [7]

$$2|h_1|M_0 \leq |g'(x_1)|,$$ 

and

$$|h_1| \leq \frac{1}{2} |h_0|.$$ 

From these inequalities it follows that \( x_2 \) lies in \( J_0 \) and \( J_1 \) is contained in \( J_0 \); that is, we have

$$M_1 \leq M_0.$$ 

Therefore (12) can be replaced by

$$2|h_1|M_1 \leq |g'(x_1)|.$$ 

(13)

For the sequence

$$x_{k+1} = x_k + h_k$$ 

as computed by Halley's method it holds in general that

$$h_k = - \frac{g(x_k)}{g'(x_k)}.$$ 

(15)

Therefore (13) shows that our assumptions remain true if we replace \( x_0 \) by \( x_1 \) and \( h_0 \) by \( h_1 \), and by \( x_k \) and \( h_k \), respectively, in general. The convergence to a zero \( x^* \) that is unique in \( J_0 \) can now be proved as in [7]. To prove the error estimates we start with (11): We have

$$|g(x_1)| \leq \frac{1}{2} |h_0|^2 M_0$$

and, therefore, using (15),

$$|x_2 - x_1| = |h_1| = \left| \frac{g(x_1)}{g'(x_1)} \right| \leq \frac{M_0}{2|g'(x_1)|} |x_1 - x_0|^2.$$ 

This is (3) for \( k = 1 \). For \( k > 1 \) the assertion is proved by mathematical induction. We omit the details. Since \( |x^* - x_{k+1}| \leq h_k \) we immediately get (4) from (3).

Finally we have

$$|x^* - x_k| \leq |x_{k+1} - x_k| + |x_{k+1} - x^*|,$$

and (5) is therefore proved by using (3) and (4). Applying the mean-value theorem we have, for \( x \in J_{k-1} \),

$$|g''(x)| = \frac{1}{2} \left[ \frac{f(x)}{f'(x)} \left[ - \frac{f''(x)}{f'(x)} + \frac{3}{2} \left( \frac{f'''(x)}{f'(x)} \right)^2 \right] \right] \leq NM|x - x^*| \leq 2NM|x_k - x_{k-1}|.$$
Replacing, therefore, in (3), (4), and (5) $M_{k-1}$ by the upper bound $2NM_k |x_k - x_{k-1}|$, we establish (3'), (4') and (5'). □

Without going into details of the proof we remark that the estimation (5) can further be improved. If we define

$$t_k = \frac{1}{|g'(x_k)|} |h_k| M_k, \quad k \geq 0,$$

then

$$|x^* - x_k| \leq \frac{1}{1 + \sqrt{1 - 2t_k}} M_{k-1} |x_k - x_{k-1}|^2, \quad k \geq 1.$$  (5'')

This last can be proved in exactly the same manner as the corresponding inequality for Newton's method (see [4, p. 34 ff]).

Since one can easily show that

$$(5'')$$

(5') can be replaced by

$$|x^* - x_k| \leq \frac{1}{1 + \sqrt{1 - \frac{t_{k-1}}{t_k}} \frac{M_{k-1}}{|g'(x_k)|}} |x_k - x_{k-1}|^2.$$  (5''')

Since $t_k \to 0$, (5'') and (5''') are asymptotically better than (5) by a factor of $\frac{1}{2}$.

3. Numerical Example. In order to compare our theorem with other results, we consider the simple well-tried example (see, for example, [3, p. 453, Tabelle 2])

$$f(x) = x^3 - 10.$$  

We have

$$f'(x) = 3x^2, \quad f''(x) = 6x, \quad f'''(x) = 6.$$  

As in [3], we choose $x_0 = 2$ and get

$$h_0 = -\frac{g(x_0)}{g'(x_0)} = 2.153 \, 846 \, 154, \quad x_1 = x_0 + h_0 = 2.153 \, 846 \, 154, \quad \text{and} \quad J_0 = [2, 2.307 \, 692 \, 308].$$

For $x \in J_0$ we have

$$g''(x) \leq M_0 = 0.288 \, 675 \, 135.$$  

Furthermore

$$|g'(x_0)| = 3.752 \, 776 \, 750;$$

hence the main inequality (2) holds.

We have

$$|g'(x_1)| = 3.731 \, 590 \, 624.$$  

Using this value in (5) for $k = 1$ we have

$$|x^* - x_1| \leq \frac{M_0}{|g'(x_1)|} |x_1 - x_0|^2 = 0.001 \, 831 \, 001.$$  

All numerical values have been computed using a HP21 pocket calculator.
The actual error is
\[ |x^* - x_1| \approx 0.000 588 556. \]

The error estimation reflects the order of the actual error and is only three times as large as the actual error.

For \( h_1 \) we get
\[ h_1 = 0.000 588 556. \]

Therefore we have
\[ J_1 = [x_1, x_1 + 2h_1] \]
\[ = [2.153 846 154, 2.155 023 227], \]
and, for \( x \in J_1 \),
\[ g''(x) \in [-0.000 946 821, 0.000 945 788], \]

hence
\[ M_1 = 0.000 946 821. \]

For \( x_2 \) we get the value
\[ x_2 = 2.154 434 690, \]
and therefore
\[ |g'(x_2)| = 3.731 530 346. \]

Using (5) for \( k = 2 \), we finally have
\[ |x^* - x_2| \leq \frac{M_1}{|g'(x_2)|} |x_2 - x_1|^2 \approx 8.78 \times 10^{-11}. \]

The actual error for \( x_2 \) is
\[ |x^* - x_2| \approx 2.93 \times 10^{-11}. \]

Therefore the actual error is overestimated only by a factor of 3. (5') gives us the estimate
\[ |x^* - x_2| \approx 4.39 \times 10^{-11}. \]

None of the error estimates that are to be found in Table 2 in [3, p. 453] give a smaller bound for the error than (5').

4. Conclusion. If we compare the assumptions of our theorem with other results (especially with those given in [3]), then it seems that the most drastic assumption is the requirement that \( f' \) doesn't change sign in \( J_0 \). We need this assumption, however, in order to define \( g(x) \) in the whole interval \( J_0 \).

If one wants to use the more precise error estimate (5), then one has to compute the bound \( M_{k-1} \) for the second derivative of \( g \). There is no essential difficulty in doing this since one can get these bounds very simply by using interval arithmetic in (8). See, for example, [1, p. 28 ff], or [6, p. 161 ff], or [9]. The same is true for the inequality (5') and the bounds \( M \) and \( N \) appearing there.

\[ ^t \text{The "\( \approx \)" sign means here and in the sequel that the number is rounded upwards in the usual manner.} \]
In conclusion we remark that the most important applications of Halley’s method are to nonlinear equations in Banach spaces. See, for instance, the discussion and examples in [3].

The generalization of the results of this paper to this general case and some numerical examples will be discussed in another paper.

5. Acknowledgment. The author wants to express his thanks to Professor Jon Rokne for reading the manuscript. Thanks also to the Department of Mathematics and to the Department of Computer Science of the University of Calgary, Alberta, for the hospitality during the academic year 1979/80 when this paper was written.

References

2. G. H. Brown, Jr., On Halley’s variation of Newton’s method, this MONTHLY, 84 (1977) 726–727.