Rate of convergence of higher order methods

Trond Steihaug*, Sara Suleiman

University of Bergen, Department of Informatics, P.O. Box 7803, N-5020 Bergen, Norway

A R T I C L E   I N F O

Article history:
Available online 30 July 2011

Keywords:
Nonlinear system of equations
Newton’s methods
Schröder’s method
Chebyshev’s method
Halley’s method

A B S T R A C T

Methods like the Chebyshev and the Halley method are well known methods for solving nonlinear systems of equations. They are members in the Halley class of methods and all members in this class have local and third order rate of convergence. They are single point iterative methods using the first and second derivatives. Schröder’s method is another single point method using the first and second derivatives. However, this method is only quadratically convergent. In this paper we derive a unified framework for these methods and show their local convergence and rate of convergence. We also use the same approach to derive inexact methods. The methods in the Halley class require solution of two linear systems of equations for each iteration. However, in the Chebyshev method the coefficient matrices will be the same. Using the unified framework we show how to extend this to all methods in the class. We will illustrate these results with some numerical experiments.

© 2011 IMACS. Published by Elsevier B.V. All rights reserved.

1. Introduction

Consider the system of nonlinear equations

$$ F(x) = 0, $$

where the function $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is sufficiently smooth. The iterative methods we consider only use function, first and second derivatives of $F$ evaluated at the single point $x_k$ where $k$ is the iteration index, to generate the next iterate $x_{k+1}$. Hernández and Salanova [16] made popular a class of methods for scalar functions which later was extended by Gutiérrez and Hernández [11,12] to systems of equations. This class is referred to as the Halley class and is parameterized with a parameter $\alpha$. It is convenient to introduce the $n \times n$ matrix $L$:

$$ L(x) = F'(x)^{-1} F''(x) F'(x)^{-1} F(x). $$

For a given initial guess $x_0$, the Halley class is

$$ x_{k+1} = x_k - \left[ I + \frac{1}{2} L(x_k) \right]^{-1} F'(x_k)^{-1} F(x_k), \quad k = 0, 1, \ldots, $$

where $I$ is the identity matrix in $\mathbb{R}^{n \times n}$. The Halley class has been derived independently by several authors. The class was suggested by Kaazik [17] and the class is a subset of a family of methods proposed by Schwetlick [27]. The Halley class includes well known methods like the Chebyshev method $\alpha = 0$ [21], the Halley method $\alpha = \frac{1}{2}$ [19] and the super-Halley method $\alpha = 1$ [14,18]. All methods in the Halley class have local and third order rate of convergence under suitable assumptions.

* Corresponding author.
E-mail addresses: Trond.Steihaug@ii.uib.no (T. Steihaug), Sara.Suleiman@ii.uib.no (S. Suleiman).

0168-9274/$30.00 © 2011 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.apnum.2011.06.016
By rewriting the iteration (3) Gundersen and Steihaug [10] showed that an iteration of the Halley class can be written as a two-stage method

\[
\begin{align*}
&\text{Solve for } s_k^{(1)}: \quad F'(x_k) s_k^{(1)} = -F(x_k), \\
&\text{Solve for } s_k^{(2)}: \quad \left( F'(x_k) + \alpha F''(x_k) s_k^{(1)} \right) s_k^{(2)} = -\frac{1}{2} F''(x_k) s_k^{(1)} s_k^{(1)}. \\
&\text{Update the iterate: } \quad x_{k+1} = x_k + s_k^{(1)} + s_k^{(2)}.
\end{align*}
\]

Schröder [26] suggested to consider the nonlinear equation

\[ g(x) = f(x) / f'(x) = 0, \]

where \( f : \mathbb{R} \to \mathbb{R} \) is two times continuously differentiable. The functions \( g(x) \) and \( f(x) \) will have the same roots but the roots of \( g(x) \) are simple. Schröder’s method can easily be extended to system of nonlinear equations (1). We follow the derivation in the scalar case and apply Newton’s method on \( G(x) = 0 \) where \( G(x) \equiv F'(x)^{-1} F(x) = 0. \) Using the definition of \( G(x) \) and substituting the value of \( G'(x) \) in Newton’s method, the new iteration is \( x_{k+1} = x_k - G'(x_k)^{-1} G(x_k). \) Let \( L \) be the \( n \times n \) matrix given in (2), Schröder’s method is

\[ x_{k+1} = x_k - \left( I - L(x_k) \right)^{-1} F'(x_k)^{-1} F(x_k), \quad k = 0, 1, 2, 3, \ldots. \]

For a given starting point \( x_0 \) sufficient close to the solution, the method will have second order rate of convergence under suitable assumptions on \( F. \)

Using the same rewriting technique as for the Halley class, the Schröder method can be written as a two-stage iteration:

\[
\begin{align*}
&\text{Solve for } s_k^{(1)}: \quad F'(x_k) s_k^{(1)} = -F(x_k), \\
&\text{Solve for } s_k^{(2)}: \quad \left( F'(x_k) + F''(x_k) s_k^{(1)} \right) s_k^{(2)} = -F''(x_k) s_k^{(1)} s_k^{(1)}. \\
&\text{Update the iterate: } \quad x_{k+1} = x_k + s_k^{(1)} + s_k^{(2)}.
\end{align*}
\]

The paper is organized as follows. In Section 2 we give an algorithmic framework based on the observation in [10] that the super-Halley method is two steps of Newton on the quadratic Taylor approximation of \( F \) at iterate \( x_k. \) It is shown that all methods in the algorithmic framework are locally convergent and we give sufficient conditions for rate of convergence of the methods. The higher order methods we consider all require solutions of two linear systems of equations for each iteration. This may require factorizations of two matrices per iteration. However, in the Chebyshev method we only need one. In Section 3 we give an efficient way to use the same factorization also in the second stage of the methods and still achieve the good convergence properties of the super-Halley method. In Section 4 some numerical experiments are given to illustrate the behavior of the methods.

2. Framework for higher order methods

It is well known that we can derive Newton’s method by considering a linear approximation \( S_k(s) = F(x_k) + F'(x_k) s \) of \( F(x_k + s) \) in (1) and then find \( s \) so that \( S_k(s) = 0. \) In this section we consider a quadratic approximation \( T_k(s) \) of \( F(x_k + s) \) and instead of finding a solution of \( T_k(s) = 0 \) we will find an approximate solution.

Define the quadratic model

\[ T_k(s) = F(x_k) + F'(x_k) s + \frac{1}{2} F''(x_k) s s. \tag{4} \]

Since \( T_k(0) = F(x_k), \) the relative error must satisfy \( \| T_k(s_k) \| / \| F(x_k) \| < 1. \)

To find a solution of the problem (1) we consider methods in the Algorithmic Framework 2.1.

\begin{algorithm}[h]
\textbf{Algorithm 2.1: Framework}
\begin{algorithmic}
\State Given \( x_0; \)
\For {\( k = 0, 1, 2, \ldots \) until convergence}
\State Find approximate solution \( s_k \) of \( T_k(s_k) = 0 \) such that for \( \eta_k \leq \eta < 1, \| T_k(s_k) \| \leq \eta_k \| F(x_k) \|; \)
\State Update \( x_{k+1} = x_k + s_k; \)
\EndFor
\end{algorithmic}
\end{algorithm}

We will show that if the approximate solution \( s_k \) in addition satisfies \( \| s_k \| = O(\| F(x_k) \|) \) then any method in the Algorithmic Framework 2.1 converges.

\textbf{Standard Assumptions 1.} Assume that \( F : \mathbb{R}^n \to \mathbb{R}^n \) is three times continuously differentiable in a neighborhood \( N \) of a point \( x^* \) where \( F(x^*) = 0 \) and \( F'(x^*) \) is nonsingular.
We first state a lemma using the smoothness of $F$ and nonsingularity of $F'(x)$ in a neighborhood of the solution $x^*$.

**Lemma 1.** Assume that the Standard Assumptions 1 hold. There exist positive $\varepsilon$ and $K_1$ so that
\[
\|F'(x)^{-1}\| \leq K_1,
\|F''(x)\| \leq K_2,
\|F'(x) - \alpha F''(x) F'(x)^{-1} F(x)\|^{-1} \leq K_3,
\|F(x)\| \leq 1,
\]
for all $\|x - x^*\| \leq \varepsilon$.

The proof of the lemma can be derived from basic properties of $F$ and linear algebra. An observation that is repeatedly utilized is based on the Taylor expansion of a quadratic function. Let $s^{(1)}$ and $s^{(2)}$ be any vectors in $\mathbb{R}^n$, then
\[
T_k(s^{(1)} + s^{(2)}) = T_k(s^{(1)}) + T'_k(s^{(1)}) s^{(2)} + \frac{1}{2} T''_k(s^{(1)}) s^{(2)} s^{(2)}.
\]

Define $M = K + K^4$ and $K = \max\{K_1, K_2, K_3\}$.

**Lemma 2.** Assume that the Standard Assumptions 1 hold. For a given $\alpha$ there exists $\varepsilon > 0$ so that for all $x_k$, $\|x_k - x^*\| \leq \varepsilon$ the Newton step $s_k = s^{(1)}_k$ or $s_k = s^{(1)}_k + s^{(2)}_k$, where $s^{(1)}_k$ is defined either by a method in the Halley class or Schröder’s method satisfies
\[
\|s_k\| \leq M\|F(x_k)\|.
\]

Furthermore, the quadratic model (4) at the Newton step satisfies
\[
\|T_k(s^{(1)}_k)\| \leq \frac{K^3}{2}\|F(x_k)\|^2.
\]

Further, a step $s_k$ of a method in the Halley class, satisfies
\[
\|T_k(s_k)\| \leq \frac{1}{2}|1 - \alpha|K^6\|F(x_k)\|^3 + \frac{1}{8}K^9\|F(x_k)\|^4
\]
and for Schröder’s method
\[
\|T_k(s_k)\| \leq \frac{1}{2}\left(K^3 + \frac{1}{4}K^9\right)\|F(x_k)\|^2.
\]

**Proof.** Choose $\varepsilon > 0$ so that the results in Lemma 1 hold. In Newton’s method $s^{(1)}_k = -F'(x_k)^{-1}F(x_k)$, thus from Lemma 1
\[
\|s^{(1)}_k\| \leq K_1\|F(x_k)\|.
\]

Since $F(x_k) + F'(x_k)s^{(1)}_k$ vanishes, the model $T_k$ at the Newton step $s^{(1)}_k$ is
\[
T_k(s^{(1)}_k) = \frac{1}{2} F''(x_k)s^{(1)}_k s^{(1)}_k.
\]

Thus, the norm of the model $T_k$ at the Newton step is obtained by using Lemma 1 and the inequality (10)
\[
\|T_k(s^{(1)}_k)\| \leq \frac{1}{2}\|F''(x_k)\|\|s^{(1)}_k\|^2 \leq \frac{1}{2}K_2 K^2\|F(x_k)\|^2.
\]

In the Halley class $s_k = s^{(1)}_k + s^{(2)}_k$, where $s^{(1)}_k$ is the Newton step and $s^{(2)}_k$ defined by
\[
s^{(2)}_k = \frac{1}{2} (F'(x_k) - \alpha F''(x_k) F'(x_k)^{-1} F(x_k))^{-1} F''(x_k) s^{(1)}_k s^{(1)}_k.
\]

Taking the norm and using Lemma 1 and (10)
\[
\|s^{(2)}_k\| \leq \frac{1}{2} K_3 K_2 K^2\|F(x_k)\|^2.
\]

Combining (10) and (11), the norm on $s_k$ is obtained by
\[
\|s_k\| \leq \|s^{(1)}_k\| + \|s^{(2)}_k\| \leq K_1\|F(x_k)\| + \frac{1}{2}K_3 K_2 K^2\|F(x_k)\|^2.
\]
Lemma 3. Let \( \| \cdot \| \) be a norm and condition (13) follows by observing that if \( x \) and \( x^* \) are in this neighborhood, so that

\[
\| F(x + s) - F(x) - F'(x)s - \frac{1}{2} F''(x)ss \| \leq \varepsilon \quad \text{and} \quad \| x + s - x^* \| \leq \varepsilon.
\]

Consider

\[
F(x) = (F(x) - F(x^*) - F'(x^*)(x - x^*)) + F'(x^*)(x - x^*),
\]

and condition (13) follows by observing that if \( F \) is two times continuously differentiable, then [22, 3.3.6]

\[
\| F(x) - F'(x^*)(x - x^*) \| \leq \frac{1}{2} \sup_{0 \leq t \leq 1} \| F''(x^* + t(x - x^*)) \| \| x - x^* \|^2.
\]

Let \( \| \cdot \| \) be a norm and \( a \) be any vector in \( \mathbb{R}^n \). The weighted norm \( \| \cdot \|_* \) is defined by

\[
\| a \|_* = \| F'(x^*)a \|.
\]
Theorem 1. Let $\eta_k < \eta < 1$; $k \geq 0$. There exists $\epsilon > 0$ so that for all $x_0$, $\|x_0 - x^*\| \leq \epsilon$, the sequence of iterates $\{x_k\}_{k \geq 0}$ given by Algorithmic Framework 2.1 satisfies $\|x_k - x^*\| \leq \epsilon$ and converges Q-linearly to $x^*$ in the sense

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

for some $1 > \rho > \eta$. Let $x_k \to x^*$, then the rate of convergence is at least

1. Q-super-linear if $\eta_k \to 0$.
2. Q-quadratic if $\eta_k = O(\|F(x_k)\|)$ (Newton, Schröder).
3. Q-cubic if $\eta_k = O(\|F(x_k)\|^2)$ (Halley class).
4. Q-order min{3, 4} if $\eta_k = O(\|F(x_k)\|^{\rho - 1})$, $1 < \rho$.

Proof. Choose $\epsilon > 0$ so that Lemmas 1, 2 and 3, for $1 > \delta > 0$, $\frac{1 + \delta}{1 - \delta} \tilde{\eta} < 1$, and

$$\eta + K_4 M^3 \|F(x)\|^2 \leq \tilde{\eta} < 1$$

hold for some $\eta < \tilde{\eta} < 1$ for $\|x - x^*\| \leq \epsilon$. Consider

$$F(x_k + s_k) = T_k(s_k) + F(x_k + s_k) - T_k(s_k).$$

Taking the norm in using (12) in Lemma 3, we get

$$\|F(x_k + s_k)\| \leq \|T_k(s_k)\| + \|F(x_k + s_k) - T_k(s_k)\| \leq \|T_k(s_k)\| + K_4 \|s_k\|^3 \leq \|T_k(s_k)\| + K_4 M^3 \|F(x_k)\|^3,$$

using Lemma 2. Since $\|T_k(s_k)\| \leq \eta_k \|F(x_k)\|$, from (15)

$$\|F(x_k + s_k)\| \leq (\eta_k + K_4 M^3 \|F(x_k)\|^2) \|F(x_k)\| \leq \tilde{\eta} \|F(x_k)\|.$$

Using Lemma 3, we have

$$(1 - \delta) \|F'(x^*)(x_k + s_k - x^*)\| \leq \|F(x_k + s_k)\| \leq \tilde{\eta} \|F(x_k)\| \leq (1 + \delta) \tilde{\eta} \|F'(x^*)(x_k - x^*)\|.$$

Using the weighted norm (14)

$$\|x_k + s_k - x^*\|_* \leq \rho \|x_k - x^*\|_*,$$

where $\rho = \frac{1 + \delta}{1 - \delta} \tilde{\eta} < 1$. We have thus shown that $x_{k+1} = x_k + s_k$ will be in the neighborhood $\|x_{k+1} - x^*\|_* \leq \epsilon$ and the iterates are converging linearly.

Assume that $x_k \to x^*$. To prove Q-super-linear rate of convergence, assume that $\eta_k \to 0$. Since $\|T_k(s_k)\| \leq \eta_k \|F(x_k)\|$, then the inequality (16) becomes

$$\|F(x_k + s_k)\| \leq (\eta_k + K_4 M^3 \|F(x_k)\|^2) \|F(x_k)\|.$$

Then for $\rho_k = (1 + \delta)(\eta_k + K_4 M^3 \|F(x_k)\|^2)/(1 - \delta) \to 0$ and

$$\|x_{k+1} - x^*\|_* \leq \rho_k \|x_k - x^*\|_*.$$

For Newton’s and Schröder’s methods $\eta_k = O(\|F(x_k)\|)$ using (7) and (9) in Lemma 2. Let $C_1 = \max\{\frac{k_1^1}{3} + \frac{k_2^9}{3}, K_4 M^3\}$ then from Eq. (16)

$$\|F(x_k + s_k)\| \leq C_1 \|F(x_k)\|^2$$

which implies quadratic rate of convergence using Lemma 3

$$\|x_{k+1} - x^*\|_* \leq \frac{C_1(1 + \delta)^2}{1 - \delta} \|x_k - x^*\|^2_*.$$

To prove cubic rate of convergence for a given method in the Halley class let $\eta_k = O(\|F(x_k)\|^2)$. From (8) and (16), it follows
\[ \| F(x_k + s_k) \| \leq |1 - \alpha| \frac{K^6}{2} \| F(x_k) \|^3 + \frac{K^9}{8} \| F(x_k) \|^4 + K_4 M^3 \| F(x_k) \|^3 \]

\[ \leq \left( |1 - \alpha| \frac{K^6}{2} + \frac{K^9}{8} + K_4 M^3 \right) \| F(x_k) \|^3. \]

Thus by Lemma 3 all methods in the Halley class have a cubic rate of convergence.

To prove the final result let \( p > 1 \) be any number and assume that \( \eta_k = O(\| F(x_k) \|^{p-1}) \). Then there exist \( C_2 > 0 \) so that \( \eta_k \leq C_2 \| F(x_k) \|^p \). Since \( \| T_k(s_k) \| \leq \eta_k \| F(x_k) \|^{p-2} \), then Eq. (16) becomes

\[ \| F(x_k + s_k) \| \leq C \| F(x_k) \|^{\hat{p}} + K_4 M^3 \| F(x_k) \|^3 \]

\[ \leq C_3 \| F(x_k) \|^{|\min(3, \hat{p})|}, \]

where \( C_3 = \max(C_1, K_4 M^3) \). From Lemma 3, we get

\[ \| x_{k+1} - x^* \|_e \leq \frac{C_3 (1 + \delta)^{|\min(3, \hat{p})|}}{1 - \delta} \| x_k - x^* \|_e^{|\min(3, \hat{p})|}. \quad \square \]

3. An inexact approach

Consider the quadratic model (4) and the system \( T_k(s) = 0 \). For a given starting point \( s^{(0)} \) an inexact Newton method [7]

\[ \text{For } l = 0, 1, \ldots \text{ until termination do} \]

\[ \text{Solve } T'(s^{(l)}) t^{(l)} = -T(s^{(l)}) + r^{(l+1)} \]

\[ \text{Update } s^{(l+1)} = s^{(l)} + t^{(l)} \]

where we have eliminated the subscript \( k \). For each iterate \( l \), linear system is solved approximately

\[ r^{(l+1)} \equiv T'(s^{(l)}) t^{(l)} + T(s^{(l)}), \quad \| r^{(l+1)} \| \leq \hat{\eta}^{(l)} \| T(s^{(l)}) \|. \]

Assume that the iteration is terminated after two iterations of an inexact Newton method. For a given initial \( s^{(0)} \) two iterations are given by

\[ T_k'(s^{(0)}) s^{(1)} = -T_k(s^{(0)}) + r_k^{(1)}, \quad (19) \]

\[ T_k'(s^{(1)}) s^{(2)} = -T_k(s^{(1)}) + r_k^{(2)}. \quad (20) \]

and the approximate solution is \( s = s^{(0)} + s^{(1)} + s^{(2)}. \) A suitable initial guess is \( s^{(0)} = 0 \) which is used in the following. Then \( T_k(s^{(0)}) = F(x_k), \ T_k'(s^{(0)}) = F'(x_k), \)

\[ T_k(s^{(1)}) = r^{(1)} + \frac{1}{2} F''(x_k) s^{(1)} s^{(1)}, \quad \text{and} \quad T_k'(s^{(1)}) = F'(x_k) + F''(x_k) s^{(1)}. \]

We can rewrite Eqs. (19) and (20)

\[ F'(x_k) s_k^{(1)} = -F(x_k) + r_k^{(1)} \]

\[ \left(F'(x_k) + \alpha F''(x_k) s_k^{(1)}\right) s_k^{(2)} = -r_k^{(1)} - \frac{1}{2} F''(x_k) s_k^{(1)} s_k^{(1)} + r_k^{(2)}. \quad (22) \]

We have extended the inexact Newton approach to include \( \alpha \neq 1 \). This will be referred to as one step of an inexact Halley class method. It can be shown that if \( \alpha = 1 \) and \( r_k^{(1)} = r_k^{(2)} = 0 \) then this is two steps of Newton’s method on the quadratic system \( T_k(s) = 0 \) [10].

A modified method is derived from an inexact Halley class method where we assume that the first linear system of equations (21) is solved exactly and we have an error in the second equation (22)

\[ F'(x_k) s_k^{(1)} = -F(x_k), \quad (23) \]

\[ (F'(x_k) + \alpha F''(x_k) s_k^{(1)}) s_k^{(2)} = \frac{1}{2} F''(x_k) s_k^{(1)} s_k^{(1)} + r_k^{(2)}. \quad (24) \]

This would be the case when we have an LU-factorization of \( F'(x_k) \) and reuse this factorization in the next equation (24) and improve the solution by doing a few linear fixed-point iterations given in Algorithm 3.1.
Algorithm 3.1: Computing $s_k^{(2)}$ and $t_k^{(2)}$

Define $B = F'(x_k)$, $A = F'(x_k) + \alpha F''(x_k)x_k^{(1)}$;
Define $b = -\frac{1}{2} F''(x_k)x_k^{(1)}x_k^{(1)}$;
Define $b_0 = 0, r_0 = b$;

for $l = 1, 2, \ldots$ until termination do

1. Solve for $z_{l-1}$: $Bz_{l-1} = n_{l-1}$;
2. Update $w_l = w_{l-1} + z_{l-1}$;
3. Update $n_l = b - Aw_l$;

end

$s_k^{(2)} = w_l, t_k^{(2)} = n_l, j = l$;

If the iterative method terminates in $j$ iterations, then the residual will satisfy

$$r_k^{(2)} = r_j = (I - AB^{-1})r_{j-1} = (I - AB^{-1})^j r_0.$$  \hfill (25)

From the definitions in Algorithm 3.1

$$I - AB^{-1} = \alpha F''(x_k)F'(x_k)^{-1}F(x_k)F'(x_k)^{-1}$$

$$= \alpha F'(x_k)L(x_k)F'(x_k)^{-1},$$  \hfill (26)

where $L$ is given by (2). From (26) and Lemma 1, we have $\|I - AB^{-1}\| \leq |\alpha| K_1^2 K_2 \|F(x_k)\| < 1$, provided $\|F(x_k)\|$ is sufficiently small. Thus, the residual $r_k^{(2)}$ in (24) will satisfy

$$\|r_k^{(2)}\| \leq \|I - AB^{-1}\|^j \|b\| \leq (|\alpha| K_1^2 K_2)^j \left(\frac{1}{2} K_2 K_1^2 \|F(x_k)\|\right)^{j+2}.  \hfill (28)$$

Further, $s_k^{(2)}$ is given by

$$s_k^{(2)} = w_j = \sum_{m=0}^{j-1} B^{-1}(I - AB^{-1})^m b$$

$$= -\frac{1}{2} \left[ \sum_{m=1}^{j} \alpha^{m-1} L(x_k)^m \right] F'(x_k)^{-1}F(x_k),$$

using that $(F'(x_k)L F'(x_k)^{-1})^m = F'(x_k)L^m F'(x_k)^{-1}$.

Thus the modified Halley iterate using $j$ inner iterations is given by

$$x_{k+1} = x_k + s_k^{(1)} + s_k^{(2)}$$

$$= x_k - \left( I + \frac{1}{2} \sum_{m=1}^{j} \alpha^{m-1} L(x_k)^m \right) F'(x_k)^{-1}F(x_k).$$  \hfill (29)

Chebyshev's method is given by modified Halley (29) for $j = 1$. The C-methods [3,4] are defined

$$x_{k+1} = x_k - \left( I + \frac{1}{2} L(x_k) + CL(x_k)^2 \right) F'(x_k)^{-1}F(x_k)$$  \hfill (30)

where $L$ is given by (2). C-methods with $C = \frac{\alpha}{2}$ is the modified Halley class with $j = 2$ inner iterations. In addition, when $\alpha = 1$, the modified super-Halley with two inner iterations, reduces to a method proposed by Popovski [23] in the scalar case. Further, the Halley class (3) can be reduced to the modified Halley class (29) with $j$ inner iterations, when

$$(I - \alpha L(x_k))^{-1} \approx \sum_{m=0}^{j-1} (\alpha L(x_k))^m.$$  \hfill (31)

**Theorem 2.** Assume that the Standard Assumptions 1 hold. Let $s_k^{(1)}$ and $s_k^{(2)}$ satisfy equations (23) and (24) and the linear fixed-point method is terminated in $j$ iterations. Then methods in the modified Halley class are locally convergent and the rate of convergence is at least $O$-order $\min \{3, j + 2\}$ for any $\alpha$. 
Proof. Choose \( \varepsilon > 0 \) so that Lemmas 1 and 3 and
\[
|\alpha| K_1^2 K_2 \| F(x) \| < 1
\]  
(32)
hold for all \( \| x - x^* \| \leq \varepsilon \). Let \( s_k^{(1)} \) be the solution of (23). From Lemma 1
\[
\| s_k^{(1)} \| \leq K_1 \| F(x_k) \|.
\]  
(33)
Let \( s_k^{(2)} \) be the solution of (24). Since \( T_k(s_k^{(1)}) = \frac{1}{2} F''(x_k) s_k^{(1)} s_k^{(1)} \), then by Lemma 2, we have
\[
\| s_k^{(2)} \| \leq \frac{1}{2} K_2 \| s_k^{(1)} \| ^2 + K_3 \| r_k^{(2)} \|.
\]
Using (33) and (28) in the previous inequality, we get
\[
\| s_k^{(2)} \| \leq \frac{1}{2} K_3 K_2 K_1^2\left\| F(x_k) \right\|^2 + K_3 K_5 \| F(x_k) \| ^{2+j}.
\]
where \( K_5 = \frac{|\alpha|}{2} K_1^2 K_1^2 K_2^2 + K_2^2 + K_3 K_3 \). Therefore,
\[
\| s_k \| = \| s_k^{(1)} \| + \| s_k^{(2)} \| \leq K_1 \| F(x_k) \| + \left( \frac{1}{2} K_2 K_2 K_2 + K_3 K_3 \right) \| F(x_k) \|^2
\]
\[
\leq \left( K_1 + \left( \frac{1}{2} K_2 K_2 K_2 + K_3 K_3 \right) \right) \| F(x_k) \|.
\]
From observation (5), and (24)
\[
T_k(s_k) = (1 - \alpha) F''(x_k) s_k^{(1)} s_k^{(2)} + \frac{1}{2} F''(x_k) s_k^{(2)} s_k^{(2)} + r_k^{(2)},
\]
then
\[
\| T_k(s_k) \| \leq |1 - \alpha| K_2 \| s_k^{(1)} \| \| s_k^{(2)} \| + \frac{1}{2} K_2 \| s_k^{(2)} \| ^2 + \| r_k^{(2)} \|
\]
\[
\leq \frac{1}{2} \left| 1 - \alpha \right| K_2 K_1 C_1 \| F(x_k) \|^3 + \frac{1}{2} K_2 C_2 K_2 \| F(x_k) \|^4 + K_3 \| F(x_k) \| ^{2+j},
\]
where \( C_7 = K_1 + (\frac{1}{2} K_2 K_2 K_2 + K_3 K_3) \). It follows that for \( x_k \) sufficiently close to \( x^* \)
\[
\| s_k \| \leq M \| F(x_k) \| \quad \text{and} \quad \| T_k(s_k) \| \leq \eta \| F(x_k) \|,
\]
for some \( M \) and \( 0 < \eta < 1 \) and methods in the modified Halley class are locally convergent.

In Theorem 1 choose \( \eta_k = O(\| F(x_k) \| ^{\min\{2, j+1\}}) \) then methods in the modified Halley class have at least \( \min\{3, 2+j\} \) Q-order rate of convergence. \( \Box \)

In the case of quadratic functions, it has been shown by [3,5,15] that the super-Halley method has fourth order rate of convergence. Similarly, the \( C \)-methods with \( C = \frac{1}{2} \) has fourth order rate of convergence [4]. In the following corollary it is shown that modified Halley has fourth order rate of convergence when \( \alpha = 1 \) for any \( j \geq 2 \).

**Corollary 1.** Assume that \( F \) is a quadratic function. The modified super-Halley using \( j \geq 2 \) number of inner iterations has fourth order rate of convergence. For \( \alpha \neq 1 \) the rate of convergence is three for \( j \geq 1 \).

**Proof.** Let \( T_k(s) \) be the quadratic model defined in (4) then \( F(x_k + s) = T_k(s) \) since \( F \) is a quadratic function. Choose \( \alpha = 1 \). Using Lemmas 1 and 2 and (28), it follows that
\[
\| T_k(s_k) \| \leq C_8 \| F(x_k) \| ^{\min\{4,2+j\}},
\]
where \( C_8 \) is constant. Therefore, using Lemma 3 and the weighted norm (14), we obtain
\[
\| x_{k+1} - x^* \| _a \leq C_9 \| x_k - x^* \| _a ^{\min\{4,2+j\}},
\]
where \( C_9 \) is a nonnegative constant. For \( \alpha \neq 1 \) the result follows from the observation \( \| T_k(s_k) \| \leq C_10 \| F(x_k) \| ^{\min\{3,2+j\}} \). \( \Box \)
In the general case both linear systems of equations (19) and (20) will have errors
\[
\| r_k^{(1)} \| \leq \eta_k^{(1)} \| T_k(s_k^{(0)}) \|, \\
\| r_k^{(2)} \| \leq \eta_k^{(2)} \| T_k(s_k^{(1)}) \|.
\] (34) (35)

Since \( s_k^{(0)} = 0 \) then \( T_k(s_k^{(0)}) = F(x_k) \) and
\[
T_k(s_k^{(1)}) = x_k^{(1)} + \frac{1}{2} F''(x_k) s_k^{(1)} s_k^{(1)}. 
\] (36)

Let
\[
\eta_k^{(1)} \leq \eta < 1 \quad \text{and} \quad \eta_k^{(2)} \leq \eta < 1,
\] (37)
and assume that the residuals \( r_k^{(1)} \) and \( r_k^{(2)} \) for \( 0 < p, q \) satisfy the following
\[
\| r_k^{(1)} \| = O (\| F(x_k) \|^{1+p}), \\
\| r_k^{(2)} \| = O (\| T_k(s_k^{(1)}) \|^{1+q})
\] (38) (39)
when \( x_k \to x^* \).

**Theorem 3.** Assume that the Standard Assumptions 1 hold. Let \( s_k^{(1)} \) and \( s_k^{(2)} \) be solutions of Eqs. (21) and (22) that satisfy (34), (35) and (37). Then an inexact Halley method is locally convergent.

If the sequence of iterates \( \{x_k\} \) converges to the solution \( x^* \) and \( 0 < p, q \leq 1 \) in (38) and (39) then the Q-order rate of convergence is at least
\[
\min \{(1+p)(1+q), 3\}, \quad \text{for} \ \alpha = 1,
\] (40)
and
\[
\min \{(1+p)(1+q), 2+p, 3\}, \quad \text{for} \ \alpha \neq 1.
\] (41)

**Proof.** We will first prove (40) and (41). Assume that the sequence of iterates \( \{x_k\} \) converges to the solution \( x^* \). Let \( s_k^{(1)} \) be defined by Eq. (21), then using Lemma 1, (34) and (37)
\[
\| s_k^{(1)} \| \leq \| F(x_k)^{-1} \| (\| F(x_k) \| + \| r_k^{(1)} \|) \leq 2K_1 \| F(x_k) \|. 
\] (42)

From Eqs. (38) and (39) there exists \( C_3 > 0 \) so that
\[
\| r_k^{(1)} \| \leq C_3 \| F(x_k) \|^{1+p} \quad \text{and} \quad \| r_k^{(2)} \| \leq C_3 \| T_k(s_k^{(1)}) \|^{1+q}. 
\] (43)

From (36), taking the norm, and using the inequality (42) we have
\[
\| T_k(s_k^{(1)}) \| \leq \| r_k^{(1)} \| + \frac{1}{2} \| F''(x_k) \| \| s_k^{(1)} \| \| s_k^{(1)} \| \\
\leq C_3 \| F(x_k) \|^{1+p} + 2K_2 K_1^2 \| F(x_k) \|^2 \\
\leq (C_3 + 2K_2 K_1^2) \| F(x_k) \|^{\min(1+p,2)}. 
\] (44)

Let \( s_k^{(2)} \) be defined by Eq. (22), and from Lemma 1, (35) and (37)
\[
\| s_k^{(2)} \| \leq \| (F'(x_k) + \alpha F''(x_k) s_k^{(1)})^{-1} \| (\| T_k(s_k^{(1)}) \| + \| r_k^{(2)} \|) \\
\leq 2K_3 \| T_k(s_k^{(1)}) \| \\
\leq 2K_3 (C_3 + 2K_2 K_1^2) \| F(x_k) \|^{\min(1+p,2)}
\] (45)

using (44). From (43) and (44)
\[
\| r_k^{(2)} \| \leq C_3 \| T_k(s_k^{(1)}) \|^{1+q} \leq C_3 (C_3 + 2K_2 K_1^2)^{1+q} \| F(x_k) \|^{\min((1+p)(1+q), 2(1+q))}. 
\] (46)

Let \( s_k = s_k^{(1)} + s_k^{(2)} \), then by using observation (5), the model \( T_k(s) \) at \( s_k \) is
\[
T_k(s_k) = r_k^{(2)} + (1 - \alpha) T''(s_k^{(1)}) s_k^{(1)} s_k^{(2)} + \frac{1}{2} T''(s_k^{(1)}) s_k^{(2)} s_k^{(2)}.
\] (47)
Then from Lemma 1,
\[
\|T_k(s_k)\| \leq \|r_k^{(1)}\| + |1 - \alpha|K_2\|s_k^{(1)}\| \|s_k^{(2)}\| + \frac{1}{2}K_2\|s_k^{(2)}\|^2
\leq C_4\|F(x_k)\|^{\min\{(1+p)(1+q),2(1+p),3\}} + |1 - \alpha|C_5\|F(x_k)\|^{\min\{2+p,3\}} + C_6\|F(x_k)\|^{\min\{2(1+p),4\}},
\]
by choosing appropriate constants $C_4, C_5$ and $C_6$.

For $\alpha = 1$ and from Theorem 1 the sequence of iterates has $Q$-order rate of convergence at least $\min\{(1+p)(1+q),2(1+q),2(1+p),3\}$. Using that $p \leq 1$ and $q \leq 1$ we have the desired result (40). For $\alpha \neq 1$, the $Q$-order rate of convergence will be at least $\min\{2+p,3,(1+p)(1+q),2(1+q)\}$. □

From Theorems 1 and 3, we conclude that a method in the inexact Halley class of methods has $Q$-order 3 rate of convergence.

For unconstrained optimization, an inexact super-Halley using the conjugate gradient method as a solver for the linear systems is presented in [9]. Gundersen and Steihaug [9] used (34) and (35) as a termination with
\[
\eta_k^{(i)} \leq \min\{\eta, \|F(x_k)\|\}, \quad i = 1, 2.
\]

Using (44) with $p = 1$, we have $\|T_k(s_k^{(1)})\| = O(\|F(x_k)\|)$. Then
\[
\|r_k^{(1)}\| = O(\|F(x_k)\|^2) \quad \text{and} \quad \|r_k^{(2)}\| = O(\|F(x_k)\|^3).
\]

From (45), we have $\|s_k^{(2)}\| = O(\|F(x_k)\|)$. Using (47), with $\alpha = 1$, we get
\[
\|T_k(s_k)\| = O(\|F(x_k)\|^3).
\]
Then from Theorem 1, the inexact super-Halley will have local convergence with $Q$-order 3 rate of convergence as proved in [9].

It is well known that the Chebyshev method can be written as a two-stage method with the same matrix in the two-stage [24,28]. The Halley method can be written as well as two-stage method [28]. Consider the inexact Halley class method based on the Chebyshev method for $\alpha = 0$. Versions of an inexact Chebyshev method for unconstrained optimization have been introduced [8,29,30] on the form
\[
\begin{align*}
F'(x_k)s_k^{(1)} & = -F(x_k) + \tilde{r}_k^{(1)}, \\
F'(x_k)s_k^{(2)} & = -\frac{1}{2}F''(x_k)s_k^{(1)}s_k^{(1)} + \tilde{r}_k^{(2)},
\end{align*}
\]
where the function $F$ is the gradient. The new iterate is $x_{k+1} = x_k + s_k^{(1)} + s_k^{(2)}$.

Deng and Zhang [8] proposed a new improved method of tangent hyperbolas (NIMTH). The two linear systems (48) and (49) are solved using either a Cholesky factorization or a preconditioned conjugate gradient method (PCG). After solving the systems using a Cholesky factorization, the linear systems are solved using PCG with the factorization as a preconditioner. The number of iterations between each factorization is user defined and kept fixed. Deng and Zhang [8] used the stopping conditions in the PCG
\[
\begin{align*}
\|z_k^{(1)}\| & \leq \|F(x_k)\|^3, \\
\|z_k^{(2)}\| & \leq \frac{1}{2}F''(x_k)s_k^{(1)}s_k^{(1)}\|^{-\frac{1}{2}},
\end{align*}
\]
respectively. Moreover, the derivatives of the function are calculated by AD.

Yan and Tian [29] used the same approach as Deng and Zhang [8] to solve the linear systems (48) and (49), but the stopping conditions for the PCG iterations are
\[
\begin{align*}
\|z_k^{(1)}\| & \leq \|F(x_k)\|^3 + \epsilon, \quad i = 1, 2, \\
\|z_k^{(2)}\| & \leq \frac{1}{2}F''(x_k)s_k^{(1)}s_k^{(1)}\|^{-\frac{1}{2}},
\end{align*}
\]
respectively, where $0 < \epsilon < \frac{1}{4\pi}$ and $p$ is the number of iterations between each Cholesky factorization.

An inexact approach combined with automatic differentiation has been applied to the Halley method (\(\alpha = 1/2\)) [31]. It follows from the definitions of the residuals that $r_k^{(1)} = \tilde{r}_k^{(1)}$ and $r_k^{(2)} = -\tilde{r}_k^{(1)} + \tilde{r}_k^{(2)}$.

If $p = 1$ in (38) and $q = 1/2$ in (39), then (39) becomes
\[
\|r_k^{(2)}\| = O(\|F(x_k)\|^3),
\]
and we have an inexact Chebyshev method of $Q$-order 3.
Table 1
The number of iterations sum over all test problems with different values of \( \tau \).

<table>
<thead>
<tr>
<th>Value of ( \tau )</th>
<th>Modified SH with num. of iters ( j )</th>
<th>Other methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( j = 1 )</td>
<td>( j = 2 )</td>
</tr>
<tr>
<td>( 10^{-6} )</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>( 10^{-8} )</td>
<td>20</td>
<td>18</td>
</tr>
</tbody>
</table>

Fig. 1. The stopping criteria is \( \|F(x_k)\|/\|F(x_0)\| \leq 10^{-10} \).

The first linear system in the inexact Chebyshev method is terminated when

\[
\|r^{(1)}_k\| = O(\|F(x_k)\|^2),
\]

which can be achieved earlier than the conditions of Deng and Zhang [8] and Yan and Tian [29]. For the second linear system, the stopping condition has order \( O(\|F(x_k)\|^3) \) in [8,29] and (53).

4. Numerical experiments

We compare the performance of the different methods with third order rate of convergence with Newton’s and Schröder’s methods on a set of large-scale test problems. The third order methods used are super-Halley, Halley, Chebyshev and modified super-Halley (\( \alpha = 1 \)) using \( j = 1, 2, 3 \) inner iterations. In all our experiments we use MATLAB 7.11 release R21010b on a personal computer with processor Intel Core 2 Duo 2.4 GHz and 3.8 GB memory.

The test functions are from the test sets [20,2,6,1,25]. The number of unknowns are chosen to be \( n = 200 \) and we consider local convergence and use standard starting points.

1. Broyden banded function [20,25].
2. Discrete boundary value function [20,25].
3. Discrete integral equation function [20,25].
4. Broyden tridiagonal function [20,25].
5. Exponential function 2 [6].
6. Test 205 [25].
7. Discretized nonlinear boundary value function [1,25].
8. Variably dimensioned function [20,25,2].

The starting point for problem 7 is specified in Test 213 [25], and for problem 8 the starting point is defined by \( x_0 = x^* + \varepsilon (x^0_0 - x^*)/\|x^0_0 - x^*\| \) where \( x^* \) is the solution, \( x^0_0 \) is the standard starting point and \( \varepsilon = 5.86 \times 10^{-5} \).

In Table 1, for each method, we compute the number of iterations for all problems to reach the tolerance

\[
\|F(x_k)\| \leq \tau \|F(x_0)\|, \tag{54}
\]
where the norm is the standard $\ell_2$-norm. We use two values of $\tau$, $\tau = 10^{-6}$ and $\tau = 10^{-8}$. The columns marked `Modified SH with num. of iters $j$’ are the modified super-Halley ($\alpha = 1$) using $j = 1, 2, 3$ number of inner iterations in Algorithm 3.1. Columns marked ‘SH’, ‘H’, ‘N’ and ‘S’ are the number of iterations summed over all problems for super-Halley, Halley, Newton’s and Schröder’s methods respectively. This table shows little difference between the third order methods, and between Newton’s and Schröder’s methods in terms of number of iterations. However, as expected there is a significant difference between methods with second and third order rate of convergence. We also note that super-Halley need fewest iterations among the third order methods to reach the tolerances and there are only minor differences between three inner iterations (column marked $j = 3$) and the super-Halley method.

In Fig. 1 $\|F(x_k)\|$ is shown as a function of the iteration index $k$ for the different methods using the Broyden banded function with $n = 200$. The methods can be divided in three groups. The fastest methods are super-Halley, modified super-Halley $j = 3$ and $j = 2$, while the other third order methods Chebyshev and Halley methods are faster than Newton’s and Schröder’s methods.

In order to see the difference in contraction on the last step between the different methods we consider Broyden tridiagonal function. This is a quadratic function and the methods, modified super-Halley $j = 2$ and $j = 3$ and super-Halley, will all have fourth order rate of convergence. For a given $\varepsilon$, 100 random starting points are chosen so that $\|x_0 - x^*\| = \varepsilon$. For
each starting point we compute $\|x_1 - x^*\|$ and the average is taken over all starting points. In Fig. 2 the vertical axis is the average $|x_1 - x^*|$ and the horizontal axis is $\varepsilon$. From the figure we see that to reach an absolute error of $10^{-13}$ Newton’s method need to be closer than $10^{-6}$ to the solution while the modified super-Halley with $j = 2$ only need to be closer than $10^{-3}$. The third order methods (the Halley and the Chebyshev methods) are between the second and fourth order methods. It has been shown experimentally that the error bound obtained by super-Halley is much better than the one obtained by Chebyshev and Halley methods [13]. It is shown in this experiment that super-Halley has the smallest error. For the experiment presented in Fig. 2 the number of unknowns in Broyden tridiagonal is chosen to be $n = 10$.

To see the effect on the number of iterations to reach a certain accuracy $\|x_k - x^*\| \leq \tau_1$ as a function on the distance to the solution a similar experiment presented in Fig. 2 is done. For each $\varepsilon$, 100 starting points are chosen so that $\|x_0 - x^*\| = \varepsilon$ and the average number of iterations to reach the accuracy $\tau_1 = 10^{-8}$ is computed. Again we very little difference between modified super-Halley $j = 3$, $j = 2$ and super-Halley. Similarly, only a minor difference between modified super-Halley $j = 1$ and Halley method. Fig. 3 shows only modified super-Halley $j = 2$ and $j = 1$. Further there are basically no differences between Schröder’s and Newton’s method. So Schröder method is not included in the figure. However, the number of iterations of the higher order methods are generally strictly less than for Newton’s method. The largest $\varepsilon$ corresponds to an average of 4 iterations and all methods converging for all starting points.

Acknowledgement

The authors would like to thank the reviewers for their valuable comments and suggestions during the preparation of this paper.

References