

A unified local convergence for Chebyshev-Halley-type methods in Banach space under weak conditions

Ioannis K. Argyros and Santhosh George

Abstract. We present a unified local convergence analysis for Chebyshev-Halley-type methods in order to approximate a solution of a nonlinear equation in a Banach space setting. Our methods include the Chebyshev; Halley; super-Halley and other high order methods. The convergence ball and error estimates are given for these methods under the same conditions. Numerical examples are also provided in this study.

Mathematics Subject Classification (2010): 65D10, 65D99, 65G99, 47H17, 49M15.

Keywords: Chebyshev-Halley-type methods, Banach space, convergence ball, local convergence.

1. Introduction

In this study we are concerned with the problem of approximating a solution x^* of the equation

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

where F is a Fréchet-differentiable operator defined on a convex subset D of a Banach space X with values in a Banach space Y .

Many problems in computational sciences and other disciplines can be brought in a form like (1.1) using mathematical modeling [2, 3, 4, 5, 11, 14, 15]. The solutions of these equations can rarely be found in closed form. That is why most solution methods for these equations are iterative. The study about convergence matter of iterative procedures is usually based on two types: semi-local and local convergence analysis. The semi-local convergence matter is, based on the information around an initial point, to give conditions ensuring the convergence of the iterative procedure; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. In particular, the practice of Numerical Functional Analysis for finding solution x^* of equation (1.1) is essentially connected to variants

of Newton’s method. This method converges quadratically to x^* if the initial guess is close enough to the solution. Iterative methods of convergence order higher than two such as Chebyshev-Halley-type methods [1, 3, 5, 7]–[16] require the evaluation of the second Fréchet-derivative, which is very expensive in general. However, there are integral equations, where the second Fréchet-derivative is diagonal by blocks and inexpensive [10]–[13] or for quadratic equations the second Fréchet-derivative is constant [4, 12]. Moreover, in some applications involving stiff systems [2], [5], [9], high order methods are useful. That is why in a unified way we study the local convergence of Chebyshev-Halley-type methods(CHTM) defined for each $n = 0, 1, 2, \dots$ by

$$x_{n+1} = x_n - \left[I + \frac{1}{2}L_n(I - \theta T_n)^{-1} \right] \Gamma_n F(x_n), \tag{1.2}$$

where x_0 is an initial point, I is the identity operator, $\Gamma_n = F'(x_n)^{-1}$, $T_n = \Gamma_n B(x_n)\Gamma_n F(x_n)$, B a bilinear operator and θ a real parameter. If $B(x_n) = F''(x_n)$, then: for $\theta = 0$ we obtain the Chebyshev method; for $\theta = \frac{1}{2}$ we obtain the Halley method, for $\theta = 1$ we obtain the super-Halley method [3], [5], [7]–[16] and for $\theta \in [0, 1]$ we obtain the method studied by Gutierrez and Hernandez [10], [11]. Other choices of operator B and parameter θ are possible [3]–[5]. The usual conditions for the semi-local convergence of these methods are (\mathcal{C}) :

- (\mathcal{C}_1) There exists $\Gamma_0 = F'(x_0)^{-1}$ and $\|\Gamma_0\| \leq \beta$;
- (\mathcal{C}_2) $\|\Gamma_0 F(x_0)\| \leq \eta$;
- (\mathcal{C}_3) $\|F''(x)\| \leq \beta_1$ for each $x \in D$;
- (\mathcal{C}_4) $\|F'''(x)\| \leq \beta_2$ for each $x \in D$;
- (\mathcal{C}_5) $\|F'''(x) - F'''(y)\| \leq \beta_3\|x - y\|$ for each $x, y \in D$.

The local convergence conditions are similar but x_0 is x^* in (\mathcal{C}_1) and (\mathcal{C}_2) . There is a plethora of local and semi-local convergence results under the (\mathcal{C}) conditions [1]–[16]. The conditions (\mathcal{C}_4) and (\mathcal{C}_5) restrict the applicability of these methods. That is why, in our study we assume conditions (\mathcal{A}) :

- (\mathcal{A}_1) $F : D \rightarrow Y$ is Fréchet-differentiable and there exists $x^* \in D$ such that

$$F(x^*) = 0 \text{ and } F'(x^*)^{-1} \in L(Y, X);$$

- (\mathcal{A}_2) $\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leq L_0\|x - x^*\|$ for each $x \in D$;
- (\mathcal{A}_3) $\|F'(x^*)^{-1}(F'(x) - F'(y))\| \leq L\|x - y\|$ for each $x, y \in D$;
- (\mathcal{A}_4) $\|F'(x^*)^{-1}F'(x)\| \leq N$ for each $x \in D$
- (\mathcal{A}_5) $\|F'(x^*)^{-1}B(x)\| \leq M$ for each $x \in D$.

Notice that the (\mathcal{A}) conditions are weaker than the (\mathcal{C}) conditions.

In the rest of this study, $U(w, q)$ and $\bar{U}(w, q)$ stand, respectively, for the open and closed ball in X with center $w \in X$ and of radius $q > 0$.

As a motivational example, let us define function f on $D = [-\frac{1}{2}, \frac{5}{2}]$ by

$$f(x) = \begin{cases} x^3 \ln x^2 + x^5 - x^4, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Choose $x^* = 1$. We have that

$$\begin{aligned} f'(x) &= 3x^2 \ln x^2 + 5x^4 - 4x^3 + 2x^2, \\ f''(x) &= 6x \ln x^2 + 20x^3 - 12x^2 + 10x \\ f'''(x) &= 6 \ln x^2 + 60x^2 - 24x + 22. \end{aligned}$$

Notice that $f'''(x)$ is unbounded on D . That is condition (C_4) is not satisfied. Hence, the results depending on (C_4) cannot apply in this case. However, we have $f'(x^*) = 3$ and $f(x^*) = 0$. That is, conditions (A_1) , (A_2) , (A_3) , (A_4) are satisfied for $L_0 = L = 146.6629073$, $N = 101.5578008$ and if e.g, we choose $B = 0$, then we can set $M = 0$. Then condition (A_5) is also satisfied. Hence, the results of our Theorem 2.1 that follows can apply to solve equation $f(x) = 0$ using CHTM. Hence, the applicability of CHTM is expanded under the conditions (A) .

The paper is organized as follows: In Section 2 we present the local convergence of these methods. The numerical examples are given in the concluding Section 3.

2. Local convergence

We present the local convergence of method CHTM in this section. It is convenient for the local convergence of CHTM to introduce some functions and parameters. Define parameter r_s by

$$r_s = \frac{2}{2L_0 + |\theta|MN + \sqrt{(2L_0 + |\theta|MN)^2 - 4L_0^2}}. \tag{2.1}$$

Let

$$\varphi(t) = L_0^2 t^2 - (2L_0 + |\theta|MN)t + 1. \tag{2.2}$$

Notice that r_s is the smallest positive root of polynomial φ . Note also that

$$r_s \leq \frac{1}{L_0} \tag{2.3}$$

and $\varphi(t)$ is decreasing for all $t \in [0, \frac{2L_0 + |\theta|MN}{2L_0^2}]$.

Let us define function f on $[0, r_s)$ by

$$f(t) = \frac{1}{2} \left[L + \frac{MN^2}{(1 - L_0 t)^2 - |\theta|MNt} \right] \frac{t}{1 - L_0 t}. \tag{2.4}$$

Then $f(t)$ is increasing for all $t \in [0, r_s)$. This can be seen as follows:

$$f(t) = f_1(t)f_2(t)$$

where $f_1(t) = \frac{1}{2} \left[L + \frac{MN^2}{\varphi(t)} \right]$ and $f_2(t) = \frac{t}{1 - L_0 t}$ are increasing for all $t \in [0, r_s)$.

Define polynomial g by

$$\begin{aligned} g(t) &= L_0^2(L + 2L_0)t^3 - [(2L_0 + |\theta|MN)L + 2L_0(2L_0 + |\theta|MN) + 2L_0^2]t^2 \\ &\quad + [L + MN^2 + 2(2L_0 + |\theta|MN) + 2L_0]t - 2. \end{aligned} \tag{2.5}$$

It follows from the definition of r_s and f that function f is well defined on $[0, r_s)$. We have that $g(0) = -2$ and $g(t) \rightarrow \infty$ as $t \rightarrow \infty$. Hence, polynomial g has roots in $(0, \infty)$. Denote by r_m the smallest such root. Set

$$r^* = \min\{r_s, r_m\}. \tag{2.6}$$

Then for

$$r \in [0, r^*), \tag{2.7}$$

we have that

$$g(r) < 0 \tag{2.8}$$

and

$$f(r) < 1. \tag{2.9}$$

Then, we can show the following local convergence result for method (1.2) under (\mathcal{A}) conditions

Theorem 2.1. *Suppose that the (\mathcal{A}) conditions and $\bar{U}(x^*, r) \subseteq D$, hold, where r is given by (2.1). Then, sequence $\{x_n\}$ generated by CHTM method (1.2) for some $x_0 \in U(x^*, r)$ is well defined, remains in $U(x^*, r)$ for each $n = 0, 1, 2, \dots$ and converges to x^* . Moreover, the following estimates hold for each $n = 0, 1, 2, \dots$.*

$$\|x_{n+1} - x^*\| \leq f(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\|. \tag{2.10}$$

Proof. We shall use induction to show that estimates (2.10) hold and $x_{n+1} \in U(x^*, r)$ for each $n = 0, 1, 2, \dots$. Using (\mathcal{A}_2) and the hypothesis $x_0 \in U(x^*, r)$ we have that

$$\|F'(x^*)^{-1}(F'(x_0) - F'(x^*))\| \leq L_0\|x_0 - x^*\| < L_0r < 1. \tag{2.11}$$

It follows from (2.11) and the Banach Lemma on invertible operators [2, 5, 14] that $F'(x_0)^{-1} \in L(Y, X)$ and

$$\|F'(x_0)^{-1}F'(x^*)\| \leq \frac{1}{1 - L_0\|x_0 - x^*\|} < \frac{1}{1 - L_0r}. \tag{2.12}$$

We also have by (\mathcal{A}_4) , (\mathcal{A}_5) and (2.12) that, since

$$\begin{aligned} T_0 &= [F'(x_0)^{-1}F'(x^*)][F'(x^*)^{-1}F''(x_0)][F'(x_0)^{-1}F'(x^*)] \\ &\quad \times [F'(x^*)^{-1} \int_0^1 F'(x^* + \tau(x_0 - x^*))(x_0 - x^*)d\tau] \\ \|\theta T_0\| &\leq |\theta| \|F'(x_0)^{-1}F'(x^*)\|^2 \|F'(x^*)^{-1}F''(x_0)\| \\ &\quad \times \|F'(x^*)^{-1} \int_0^1 F'(x^* + \tau(x_0 - x^*))d\tau\| \|x_0 - x^*\| \\ &\leq \frac{|\theta|MN\|x_0 - x^*\|}{(1 - L_0\|x_0 - x^*\|)^2} \\ &\leq \frac{|\theta|MNr}{(1 - L_0r)^2} < 1 \end{aligned} \tag{2.13}$$

by the choice of r and r_s .

It follows from (2.13) and the Banach lemma that $(I - \theta T_0)^{-1}$ exists and

$$\begin{aligned} \|(I - \theta T_0)^{-1}\| &\leq \frac{1}{1 - \frac{|\theta|MN\|x_0 - x^*\|}{(1 - L_0\|x_0 - x^*\|)^2}} \\ &\leq \frac{1}{1 - \frac{|\theta|MNr}{(1 - L_0r)^2}}. \end{aligned} \tag{2.14}$$

It follows from CHTM for $n = 0$ that x_1 is well defined. We shall show (2.10) holds for $n = 0$ and $x_1 \in U(x^*, r)$. Using CHTM for $n = 0$, we get the identity

$$\begin{aligned} x_1 - x^* &= x_0 - x^* - F'(x_0)^{-1}F'(x_0) - \frac{1}{2}T_0(1 - \theta T_0)^{-1}\Gamma_0F(x_0) \\ &= -[F'(x_0)^{-1}F'(x^*)][F'(x^*)^{-1} \int_0^1 (F'(x^* + \tau(x_0 - x^*))) \\ &\quad - F'(x_0)]d\tau(x_0 - x^*)] - \frac{1}{2}[T_n(1 - \theta T_n)^{-1}][F'(x_0)^{-1}F'(x^*)] \\ &\quad \times [F'(x^*)^{-1} \int_0^1 F'(x^* + \tau(x_0 - x^*))d\tau(x_0 - x^*)]. \end{aligned} \tag{2.15}$$

Using (\mathcal{A}_3) , (2.4), (2.9), (2.12)- (2.15) we get in turn

$$\begin{aligned} \|x_1 - x^*\| &\leq \frac{L\|x_0 - x^*\|^2}{2(1 - L_0\|x_0 - x^*\|)} \\ &\quad + \frac{1}{2} \frac{\frac{MN\|x_0 - x^*\|}{(1 - L_0\|x_0 - x^*\|)^2}}{1 - \frac{|\theta|MN\|x_0 - x^*\|}{(1 - L_0\|x_0 - x^*\|)^2}} \frac{N\|x_0 - x^*\|}{1 - L_0\|x_0 - x^*\|} \\ &\leq f(\|x_0 - x^*\|)\|x_0 - x^*\| \\ &\leq f(r)\|x_0 - x^*\| < \|x_0 - x^*\|, \end{aligned} \tag{2.16}$$

which shows (2.10) for $n = 0$ and $x_1 \in U(x^*, r)$. The induction is completed, if we simply replace x_0 by x_k in the preceding estimates to obtain that

$$\begin{aligned} \|x_{k+1} - x^*\| &\leq f(\|x_k - x^*\|)\|x_k - x^*\| \\ &\leq f(r)\|x_k - x^*\| < \|x_k - x^*\|, \end{aligned} \tag{2.17}$$

which implies that (2.10) holds for each $k = 0, 1, 2, \dots$, $x_{k+1} \in U(x^*, r)$ for each $k = 0, 1, 2, \dots$, and from $\|x_{k+1} - x^*\| < \|x_k - x^*\|$ we deduce that $\lim_{k \rightarrow \infty} x_k = x^*$. \square

Remark 2.2. (a) Condition (\mathcal{A}_2) can be dropped, since this condition follows from (\mathcal{A}_3) . Notice, however that

$$L_0 \leq L \tag{2.18}$$

holds in general and $\frac{L}{L_0}$ can be arbitrarily large [2]-[6].

(b) In view of condition (\mathcal{A}_2) and the estimate

$$\begin{aligned} \|F'(x^*)^{-1}F'(x)\| &= \|F'(x^*)^{-1}[F'(x) - F'(x^*)] + I\| \\ &\leq 1 + \|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \\ &\leq 1 + L_0\|x - x^*\|^p, \end{aligned}$$

condition (\mathcal{A}_4) can be dropped and N can be replaced by

$$N(r) = 1 + L_0 r^p. \tag{2.19}$$

(c) It is worth noticing that it follows from the first term in (2.16) that r is such that

$$r < r_A = \frac{2}{2L_0 + L}. \tag{2.20}$$

The convergence ball of radius r_A was given by us in [2, 3, 5] for Newton’s method under conditions (\mathcal{A}_1) - (\mathcal{A}_3) . Estimate (2.20) shows that the convergence ball of higher than two CHTM methods are smaller than the convergence ball of the quadratically convergent Newton’s method. The convergence ball given by Rheinboldt [15] for Newton’s method is

$$r_R = \frac{2}{3L} < r_A \tag{2.21}$$

if $L_0 < L$ and $\frac{r_R}{r_A} \rightarrow \frac{1}{3}$ as $\frac{L_0}{L} \rightarrow 0$. Hence, we do not expect r to be larger than r_A no matter how we choose θ, L_0, L, M and N .

(d) The local results can be used for projection methods such as Arnoldi’s method, the generalized minimum residual method (GMREM), the generalized conjugate method(GCM) for combined Newton/finite projection methods and in connection to the mesh independence principle in order to develop the cheapest and most efficient mesh refinement strategy [2]- [5], [14, 15].

(e) The results can also be used to solve equations where the operator F' satisfies the autonomous differential equation [2]- [5], [14, 15]:

$$F'(x) = T(F(x)),$$

where T is a known continuous operator. Since

$$F'(x^*) = T(F(x^*)) = T(0), \quad F''(x^*) = F'(x^*)T'(F(x^*)) = T(0)T'(0),$$

we can apply the results without actually knowing the solution x^* . Let as an example $F(x) = e^x - 1$. Then, we can choose $T(x) = x + 1$ and $x^* = 0$.

3. Numerical Examples

We present numerical examples where we compute the radii of the convergence balls.

Example 3.1. Let $X = Y = \mathbb{R}^3$, $D = \bar{U}(0, 1)$ and $B(x) = F''(x)$ for each $x \in D$. Define F on D for $v = (x, y, z)^T$ by

$$F(v) = (e^x - 1, \frac{e - 1}{2}y^2 + y, z)^T. \tag{3.1}$$

Then, the Fréchet-derivative is given by

$$F'(v) = \begin{bmatrix} e^x & 0 & 0 \\ 0 & (e - 1)y + 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Notice that $x^* = (0, 0, 0)^T$, $F'(x^*) = F'(x^*)^{-1} = \text{diag}\{1, 1, 1\}$, $L_0 = e - 1 < L = M = N = e$. The values of r_s, r_m, r^*, r_A and r_R are given in Table 1.

Table 1

θ	r_s	r_m	r^*	r_R	r_A
0	0.5820	0.0636	0.0636	0.2453	0.3249
0.5	0.1495	0.0527	0.0527	0.2453	0.3249
1	0.0948	0.0448	0.0448	0.2453	0.3248

Example 3.2. Let $X = Y = C[0, 1]$, the space of continuous functions defined on $[0, 1]$ and be equipped with the max norm. Let $D = \overline{U}(0, 1)$ and $B(x) = F''(x)$ for each $x \in D$. Define function F on D by

$$F(\varphi)(x) = \varphi(x) - 5 \int_0^1 x\theta\varphi(\theta)^3 d\theta. \tag{3.2}$$

We have that

$$F'(\varphi(\xi))(x) = \xi(x) - 15 \int_0^1 x\theta\varphi(\theta)^2\xi(\theta)d\theta, \text{ for each } \xi \in D.$$

Then, we get that $x^* = 0$, $L_0 = 7.5$, $L = M = 15$ and $N = N(t) = 1 + 7.5t$. The values of r_s, r_m, r^*, r_A and r_R are given in Table 2.

Table 2

θ	r_s	r_m	r^*	r_R	r_A
0	0.1333	0.0018	0.0018	0.0444	0.0667
0.5	0.0128	0.0016	0.0016	0.0444	0.0667
1	0.0070	0.0014	0.0014	0.0444	0.0667

Example 3.3. Returning back to the motivational example at the introduction of this study, we have

Table 3

θ	r_s	r_m	r^*	r_R	r_A
0	0.0068	0.0045	0.0045	0.0045	0.0045

Note that, since $M = 0$ the value of r_s, r_m, r^*, r_A and r_R will not change for different values of θ .

References

- [1] Amat, S., Busquier, S., Gutiérrez, J.M., *Geometric constructions of iterative functions to solve nonlinear equations*, J. Comput. Appl. Math., **157**(2003), 197-205.
- [2] Argyros, I.K., *A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach spaces*, J. Math. Anal. Appl., **20**(2004), 373-397.
- [3] I Argyros, I.K., *Computational theory of iterative methods. Series: Studies in Computational Mathematics*, 15, Editors: C.K. Chui and L. Wuytack, Elsevier Publ. Co. New York, U.S.A, 2007.
- [4] Argyros, I.K., *Quadratic equations and applications to Chandrasekhar's and related equations*, Bull. Austral. Math. Soc., **32**(1985), 275-292.
- [5] Argyros, I.K., Hilout, S., *Numerical methods in Nonlinear Analysis*, World Scientific Publ. Comp., New Jersey, 2013.

- [6] Argyros, I.K., Hilout, S., *Weaker conditions for the convergence of Newton's method*, J. Complexity, **28**(2012), 364–387.
- [7] Candela, V., Marquina, A., *Recurrence relations for rational cubic methods I: The Halley method*, Computing, **44**(1990), 169-184.
- [8] Candela, V., Marquina, A., *Recurrence relations for rational cubic methods II: The Chebyshev method*, Computing, **45**(1990), 355-367.
- [9] Chun, C., Stanica, P., Neta, B., *Third order family of methods in Banach spaces*, Computer Math. Appl., **61**(2011), 1665-1675.
- [10] Gutiérrez, J.M., Hernández, M.A., *Recurrence relations for the super-Halley method*, Computer Math. Appl., **36**(1998), 1-8.
- [11] Gutiérrez, J.M., Hernández, M.A., *Third-order iterative methods for operators with bounded second derivative*, Journal of Computational and Applied Mathematics, **82**(1997), 171-183.
- [12] Hernández, M.A., Salanova, M.J.A., *Modification of the Kantorovich assumptions for semilocal convergence of the Chebyshev method*, Journal of Computational and Applied Mathematics, **126**(2000), 131-143.
- [13] Hernández, M.A., *Chebyshev's approximation algorithms and applications*, Computer Math. Appl., **41**(2001), 433-455.
- [14] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.
- [15] Ortega, J.M., Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
- [16] Parida, P.K., Gupta, D.K., *Recurrence relations for semi-local convergence of a Newton-like method in Banach spaces*, J. Math. Anal. Appl., **345**(2008), 350-361.

Ioannis K. Argyros
Department of Mathematical Sciences
Cameron University
Lawton, OK 73505, USA
e-mail: iargyros@cameron.edu

Santhosh George
Department of Mathematical and Computational Sciences
National Institute of Technology Karnataka
India-757 025
e-mail: sgeorge@nitk.ac.in