

# Applied Mathematics and Nonlinear Sciences 

# After notes on Chebyshev's iterative method 

S. Amat ${ }^{\dagger}$ and S. Busquier ${ }^{\ddagger}$

Department of Applied Mathematics and Statistics, Technical University of Cartagena, Cartagena, SPAIN

Submission Info
Communicated by Juan L.G. Guirao
Received 11th October 2016
Accepted 1st January 2017
Available online 1st January 2017


#### Abstract

This paper is a small review of Chebyshev's method. The geometric interpretation as a generalization of Newton's method is derived. Using this interpretation its global convergence is proved. Some dynamical properties are studied. As a higher order method, they are more complicated than in Newton's method. Finally, some applications are revisited pointing out the advantages of Chebyshev's method with respect Newton's method.


Keywords: Nonlinear equations; Chebyshev's iterative method; geometric interpretation; global convergence; dynamics; applications AMS 2010 codes: 65J15

## 1 Introduction

One of the most classical problems in Numerical Analysis is the approximation of the zeros of a given function $f$, that is, finding the values $x^{*}$ for which $f\left(x^{*}\right)=0$.

In order to approximate these equations we can use iterative methods. The most used scheme is the second order Newton method,

$$
\begin{equation*}
x_{n+1}=x_{n}-\frac{f\left(x_{n}\right)}{f^{\prime}\left(x_{n}\right)} \tag{1}
\end{equation*}
$$

This paper is devoted to the analysis of Chebyshev's method that is a third order extension of Newton's method. We present the geometric interpretation of the method and its global convergence. We introduce its extension to Banach spaces and review some applications where this high order method is a good alternative to Newton's method.

[^0]Chebyshev is one of the famous mathematicians of the nineteenth century, creator of several mathematical schools in Russia: number theory, probability theory, function approximation theory, theory of mechanisms and machines, etc. He received primary education at home. His mother, taught him to read and write, while his cousin taught arithmetic and the French language, which will be very useful in his relation with Europe. He followed also completing his secondary education at home, but having as tutor in mathematics Prof. Pogorelsky, known in his day as the best teacher of elementary mathematics in Moscow. Prof. Brashman was who practically ran the university studies of Chebyshev ended in 1841. The department of physics and mathematics in which Chebyshev studied convened an award in the 1840-41 course. Chebyshev submitted a paper on the calculation of roots of equations, using the series expansion of the inverse function. The work, unpublished at the time, was rewarded with the silver medal. Chebyshev worked as a professor at the University of St. Petersburg for 35 years. He is recognized as the founder of the mathematical school in St. Petersburg whose echo and influence has reached our time in many branches of mathematics. This school was distinguished by the tendency to relate the theoretical problems of mathematics with problems in the art and nature.

## 2 Geometry, dynamics and convergence of Chebyshev's method in the scalar case

The geometric interpretation of Newton's method is well known, given an iterate $x_{n}$, the next iterate is the zero of the tangent line

$$
y(x)-f\left(x_{n}\right)=f^{\prime}\left(x_{n}\right)\left(x-x_{n}\right),
$$

to the graph of $f$ at $\left(x_{n}, f\left(x_{n}\right)\right)$.
The following well-known theorem, giving enough conditions for the global convergence of Newton's method, follows easily from its geometric interpretation.

Theorem 1. Let $f^{\prime \prime}$ be continuous on an interval J, containing a root $x^{*}$ of $f ;$ let $f^{\prime} \neq 0$ and $f^{\prime \prime} \geq 0$ or $f^{\prime \prime} \leq 0$ on $J$. Then Newton's method converges monotonically to $x^{*}$ from any point $x_{0} \in J$ such that $f\left(x_{0}\right) f^{\prime \prime}\left(x_{0}\right) \geq 0$.

Chebyshev's method is obtained by quadratic interpolation of the inverse function of $f$, in order to approximate $f^{-1}(0)$, [22]. But it also admits a geometric derivation, from a parabola in the form

$$
\begin{equation*}
a y(x)^{2}+y(x)+b x+c=0, \tag{2}
\end{equation*}
$$

that after the imposition of the super tangency conditions, $y\left(x_{n}\right)=f\left(x_{n}\right), y^{\prime}\left(x_{n}\right)=f^{\prime}\left(x_{n}\right)$ and $y^{\prime \prime}\left(x_{n}\right)=f^{\prime \prime}\left(x_{n}\right)$, can be written as

$$
-\frac{f^{\prime \prime}\left(x_{n}\right)}{2 f^{\prime}\left(x_{n}\right)^{2}}\left(y(x)-f\left(x_{n}\right)\right)^{2}+y(x)-f\left(x_{n}\right)-f^{\prime}\left(x_{n}\right)\left(x-x_{n}\right)=0 .
$$

By calculating the intersection of this parabola with the $O X$-axis we obtain the next step of Chebyshev's method:

$$
x_{n+1}=x_{n}-\left(1+\frac{1}{2} L_{f}\left(x_{n}\right)\right) \frac{f\left(x_{n}\right)}{f^{\prime}\left(x_{n}\right)},
$$

where $L_{f}\left(x_{n}\right)=\frac{f\left(x_{n}\right) f^{\prime \prime}\left(x_{n}\right)}{\left.f^{\prime}\left(x_{n}\right)\right)^{2}}$.
We refer to [1] for the geometric interpretation of other third order methods.
By using the geometric interpretation of Chebyshev's method, we can obtain the following global convergence theorem.

Theorem 2. Let $f^{\prime \prime \prime}$ be continuous on an interval $J$ containing a root $x^{*}$ of $f$, let $f^{\prime} \neq 0, L_{f}(x)>-2$ and $\left(\left(\frac{\eta}{f^{\prime}(x)}\right)^{2}\right)^{\prime \prime} \geq 0$ in $J$, with $\eta=\operatorname{sgn}\left(f^{\prime}\right)$. Then Chebyshev's method converges monotonically to $x^{*}$ from any point of the interval.

Proof. We suppose $f^{\prime}>0$ (for $f^{\prime}<0$ the proof is similar).
First, we begin from a point on the left of $x^{*}, \bar{x} \leq x^{*}$.
We would like to show that the intersection $\widehat{x}$ of the parabola $y(x)$ given in (2) with the $O X$-axis will be in $\left[\bar{x}, x^{*}\right]$. By hypothesis $L_{f}(\bar{x})>-2$, in particular, $\bar{x} \leq \widehat{x}$.

Thus, it will be enough, if for $x \geq \bar{x}$, we can prove that

$$
\begin{equation*}
y(x)=\frac{-1+\sqrt{1-4 a(b x+c)}}{2 a} \geq f(x) . \tag{3}
\end{equation*}
$$

In this case, we will obtain a monotonic increasing sequence, bounded from above by $x^{*}$, then it converges at the limit $\gamma \leq x^{*}$. So, because the construction of the method and the continuity of $f$ the convergence is obtained, $\gamma=x^{*}$.

Inequality (3) is equivalent to

$$
\frac{-1+\sqrt{1-4 a(b x+c)}}{2 a}-\frac{-1+\sqrt{1-4 a(b \bar{x}+c)}}{2 a} \geq f(x)-f(\bar{x}),
$$

or

$$
\begin{equation*}
\int_{\bar{x}}^{x} \frac{-b}{\sqrt{1-4 a(b t+c)}} d t \geq \int_{\bar{x}}^{x} f^{\prime}(t) d t . \tag{4}
\end{equation*}
$$

As $f^{\prime}>0$ then for hypothesis $\left(\left(\frac{1}{f^{\prime}}\right)^{2}\right)^{\prime \prime} \geq 0$ in $J$, i.e., $\left(\frac{1}{f^{\prime}}\right)^{2}$ is convex, and therefore

$$
\left(\frac{1}{f^{\prime}(x)}\right)^{2} \geq \frac{1-4 a(b x+c)}{(-b)^{2}}
$$

because $\frac{1-4 a(b x+c)}{(-b)^{2}}$ approximates $\left(\frac{1}{f^{\prime}(x)}\right)^{2}$ up to second order.
Thus,

$$
\frac{-b}{\sqrt{1-4 a(b x+c)}} \geq f^{\prime}(x)>0,
$$

and consequently the relation (4) holds.
Finally, if we begin from a point on the right of the root, we will obtain $\frac{-1+\sqrt{1-4 a(b x+c)}}{2 a} \leq f(x)$ and the convergence will be monotonic from the right.

We refer to [3] for the global convergence of other third order schemes and some comparisons.
In general, the method has not global convergence. We only are able in these cases to ask for local or semilocal convergence. Around the solutions we looking for regions of convergence.

Consider for instance the problem to find the zeros of a polynomial, $p(z)=0$. Let $R(z)=\frac{P(z)}{Q(z)}$, where $P(z)$ and $Q(z)$ are complex polynomials with no common factors, be a rational map on the Riemann sphere. We say that $z_{0}$ is a fixed point of $R(z)$ if $R\left(z_{0}\right)=z_{0}$. For $z \in \overline{\mathbb{C}}$ we define its orbit as the set $\operatorname{orb}(z)=$ $\left\{z, R(z), R^{2}(z), \ldots, R^{k}(z), \ldots\right\}$, where $R^{k}$ means the $k$-fold iterate of $R$. A periodic point of period $n$ is a point $z_{0}$ such that $R^{n}\left(z_{0}\right)=z_{0}$ and $R^{j}\left(z_{0}\right) \neq z_{0}$ for $0<j<n$. Observe that if $z_{0} \in \overline{\mathbb{C}}$ is a periodic point of period $n \geq 1$, then $z_{0}$ is a fixed point of $R^{n}$. Also, recall that a fixed point $z_{0}$ is respectively attracting, repelling or indifferent in case $\left|R^{\prime}\left(z_{0}\right)\right|$ is less than, greater than or equal to 1 . A periodic point of period $n$ is said to be attracting, repelling or indifferent if as a fixed point of $R^{n}(z)$ is respectively attracting, repelling or indifferent. A superattracting fixed point of $R(z)$ is a fixed point which is also a zero of the derivative $R^{\prime}(z)$. A periodic point of period $n$ is said to be a superattracting periodic point of $R(z)$ if, as a fixed point of $R^{n}(z)$, is superattracting.

Let $\zeta$ be an attracting fixed point of $R(z)$. The basin of attraction of $\zeta$ is the set $B(\zeta)=\left\{z \in \mathbb{C}: R^{n}(z) \rightarrow\right.$ $\zeta$ as $n \rightarrow \infty\}$. The immediate basin of attraction of an attracting fixed point $\zeta$ of $R(z)$, denoted by $B^{*}(\zeta)$, is the connected component of $B(\zeta)$ containing $\zeta$. Finally, if $z_{0}$ is an attracting periodic point of period $n$ of $R(z)$, the basin of attraction of the orbit $\operatorname{orb}\left(z_{0}\right)$ is the set $B\left(\operatorname{orb}\left(z_{0}\right)\right)=\cup_{j=0}^{n-1} R^{j}\left(B\left(z_{0}\right)\right)$, where $B\left(z_{0}\right)$ is the
attraction basin of $z_{0}$ as a fixed point of $R^{n}$. The Julia set of a rational map $R(z)$, denoted by $\mathscr{J}(R)$, is the closure of the set of repelling periodic points. Its complement is the Fatou set $\mathscr{F}(R)$. If $R(z)$ has an attracting fixed point $z_{0}$, then the basin of attraction $B\left(z_{0}\right)$ is contained in the Fatou set and $\mathscr{J}(R)=\partial B\left(z_{0}\right)$. Therefore, the chaotic dynamics of $R(z)$ is contained in its Julia set.

The iterative rational function for Chebyshev's method is given by

$$
\begin{equation*}
C h_{p}(z)=z-\left(1+\frac{1}{2} L_{p}(z)\right) \frac{p(z)}{p^{\prime}(z)} . \tag{5}
\end{equation*}
$$

We recall the definition of conjugacy.
Definition 1. Let $R_{1}, R_{2}: \overline{\mathbb{C}} \rightarrow \overline{\mathbb{C}}$ be two rational maps. We say that $R_{1}$ and $R_{2}$ are conjugated if there exists a Möbius transformation $\phi: \overline{\mathbb{C}} \rightarrow \overline{\mathbb{C}}$ such that $R_{2} \circ \phi(z)=\phi \circ R_{1}(z)$ for all $z$.

An important feature of conjugation of rational maps is given by the following classical result.
Theorem 3. Let $R_{1}$ and $R_{2}$ be two rational maps and let $\phi$ be a Möbius transformation conjugating $R_{1}$ and $R_{2}$, that is, $R_{2}=\phi \circ R_{1} \circ \phi^{-1}$. Then $\mathscr{F}\left(R_{2}\right)=\phi\left(\mathscr{F}\left(R_{1}\right)\right)$ and $\mathscr{J}\left(R_{2}\right)=\phi\left(\mathscr{J}\left(R_{1}\right)\right)$.

Conjugacy plays a central rôle in understanding the behavior of classes of maps from the point of view of dynamical systems in the following sense. Suppose that one wishes to describe the quantitative as well as the qualitative behavior of the map $z \rightarrow \Phi_{f}(z)$, where $\Phi_{f}(z)$ is some iterative function. Since conjugacy preserves fixed and periodic points and their type as well as attraction basins, the dynamical data concerning $f$ is carried by the fixed points of $\Phi_{f}(z)$ as well as by the nature of such fixed points which may be (super)attracting, repelling or indifferent. Therefore, it is worthwhile to build up, for polynomials of degree two and three, a parametrized family consisting of polynomials as simple as possible such that a conjugacy exists between the corresponding iterative functions.

The result that follows, due to A. Cayley [8,21], has great historical importance. In an attempt to understand the dynamics of Newton's method in the complex plane, Cayley investigated the dynamics of Newton's method applied to polynomials of a particularly simple form. He realized that major difficulties would arise when attempting to extend the following result for quadratics to cubics and beyond. It is believed that this circumstance motivated further work of P. Fatou and G. Julia along these lines.

Theorem 4. Let $N_{p}(z)=\frac{z^{2}-a b}{2 z-(b+a)}$ be the rational map obtained from Newton's method applied to a generic quadratic polynomial $p(z)=(z-a)(z-b)$. Then $N_{p}$ is conjugated to the map $z \rightarrow z^{2}$ by the Möbius transformation $M(z)=\frac{z-a}{z-b}$ and $\mathscr{J}\left(N_{p}\right)$ is the straight line in the complex plane corresponding to the locus of points equidistant from $a$ and $b$.

For Chebyshev's method, also known as super-Newton method, the following result holds.
Theorem 5. ([18]) Let

$$
C h_{p}(z)=\frac{3 z^{4}-2(a+b) z^{3}-6 a b z^{2}+6 a b(a+b) z-a b\left(a^{2}+3 a b+b^{2}\right)}{(2 z-a-b)^{3}}
$$

be the rational map obtained from Chebyshev's method applied to a generic quadratic polynomial $p(z)=$ $(z-a)(z-b)$. Then $C h_{p}(z)$ is conjugated to the map $S_{3}(z)=\frac{z^{+}+2 z^{3}}{2 z+1}$ via the Möbius transformation $M(z)=\frac{z-a}{z-b}$.

For cubic polynomials we have the following result.
Theorem 6. Let $p(z)=\left(z-z_{0}\right)\left(z-z_{1}\right)\left(z-z_{2}\right)$ be a generic cubic polynomial with roots ordered as follows: $0 \leq\left|z_{0}\right| \leq\left|z_{1}\right| \leq\left|z_{2}\right|$. Let $T(z)=\left(z_{2}-z_{0}\right) z+z_{0}$. Then

1. $p(z)$ reduces to a polynomial belonging to the parametrized family $q_{\lambda, \rho}(z)=p \circ T(z)=\lambda^{3} z(z-1)(z-$ $\rho)$, where $\lambda=z_{2}-z_{0}$ and $\rho=\frac{z_{1}-z_{0}}{z_{2}-z_{0}}$,
2. $T$ is a conjugacy between $C h_{p}$ and $C h_{q_{\lambda, \rho}}$, that is, $T^{-1} \circ C h_{q_{\lambda, p}} \circ T=C h_{p}$, and
3. if $\tilde{q}_{\rho}(z)=z(z-1)(z-\rho)$, then $C h_{q_{\lambda, \rho}}$ is conjugated to $C h_{\tilde{q}_{\rho}}$. Consequently, $C h_{p}$ is conjugated to $C h_{\tilde{q}_{\rho}}$.

It is possible to see that the one-parameter family $\tilde{q}_{\rho}(z)=z(z-\rho)(z-1)$ reduces to the well-known oneparameter family $p_{A}(z)=z^{3}+(A-1) z-A$ (see [10]).

In the two pictures that follow we show the attraction basins of Newton's and Chebyshev's iterations applied to $z^{3}-1=0$. we can see a more rich dynamics in Chebyshev's method.


Fig. 1 Basins of attraction for the polynomial $p(z)=z^{3}-1$. Newton's method.


Fig. 2 Basins of attraction for the polynomial $p(z)=z^{3}-1$. Chebyshev's method.

In [4], we study the dynamics of a family of third-order iterative algorithms which includes Chebyshev's iteration function, Halley's iterative method, super-Halley's iterative method and the c-iterative methods. Using results of conjugation of rational maps and the definition of the universal Julia set we find the conjugacy classes of these iterative methods explicitly.

## 3 Extension to Banach spaces and some applications

To approximate a solution of the nonlinear equation

$$
\begin{equation*}
F(x)=0 \tag{6}
\end{equation*}
$$

$F: X \rightarrow Y, X, Y$ Banach spaces, Chebyshev's method can be written as

$$
\begin{equation*}
x_{n+1}=x_{n}-\left(I+\frac{1}{2} L_{F}\left(x_{n}\right)\right) F^{\prime}\left(x_{n}\right)^{-1} F\left(x_{n}\right), \tag{7}
\end{equation*}
$$

where

$$
\begin{aligned}
L_{F}\left(x_{n}\right) & =F^{\prime}\left(x_{n}\right)^{-1} F^{\prime \prime}\left(x_{n}\right) F^{\prime}\left(x_{n}\right)^{-1} F\left(x_{n}\right), \\
\beta & \in[0,1] .
\end{aligned}
$$

Next, we review some applications where Chebyshev's method can be considered a good alternative to Newton's method.

### 3.1 Quadratic equations

In this case, $F^{\prime \prime}(x)$ is a constant bilinear operator that we denote by $B$.
Using Taylor expansions,

$$
\begin{equation*}
F\left(y_{n}\right)=F\left(x_{n}\right)+F^{\prime}\left(x_{n}\right)\left(y_{n}-x_{n}\right)+\frac{1}{2} F^{\prime \prime}\left(x_{n}\right)\left(y_{n}-x_{n}\right)^{2}=\frac{1}{2} B\left(y_{n}-x_{n}\right)^{2} . \tag{8}
\end{equation*}
$$

Thus, $F^{\prime}\left(x_{n}\right)^{-1} F\left(y_{n}\right)=-\frac{1}{2} L_{F}\left(x_{n}\right)\left(y_{n}-x_{n}\right)$, and the method becomes

$$
\begin{aligned}
y_{n} & =x_{n}-F^{\prime}\left(x_{n}\right)^{-1} F\left(x_{n}\right), \\
x_{n+1} & =y_{n}-F^{\prime}\left(x_{n}\right)^{-1} F\left(y_{n}\right),
\end{aligned}
$$

equivalently,

$$
\begin{align*}
F^{\prime}\left(x_{n}\right)\left(y_{n}-x_{n}\right) & =-F\left(x_{n}\right),  \tag{9}\\
F^{\prime}\left(x_{n}\right)\left(x_{n+1}-y_{n}\right) & =-F\left(y_{n}\right) .
\end{align*}
$$

Notice that we only need a $L U$ decomposition and three evaluations $\left(F\left(x_{n}\right), F\left(y_{n}\right)\right.$ and $\left.F^{\prime}\left(x_{n}\right)\right)$ in each iteration. In particular, for this problem Chebyshev's method is more efficient than Newton's method.

Moreover, we can obtain a very simple semilocal convergence:
Theorem 7. Given $x_{0}$ such that there exists $F^{\prime}\left(x_{0}\right)^{-1}$ and condition

$$
\begin{equation*}
\left\|F^{\prime}\left(x_{0}\right)^{-1} B\right\|\left\|F^{\prime}\left(x_{0}\right)^{-1} F\left(x_{0}\right)\right\| \leq \frac{1}{2} \tag{10}
\end{equation*}
$$

holds, then Chebyshev's method is well defined and converges to $x^{*}$, solution of $F(x)=0$.
We refer [12] and its references for a general convergence analysis of this type of methods.
Let us consider the equation

$$
\begin{equation*}
F(x)=x^{T} B x+C x+D=0, \tag{11}
\end{equation*}
$$

where

$$
\begin{aligned}
& B=\operatorname{rand}(N * N * N), \\
& C=\operatorname{rand}(N * N),
\end{aligned}
$$

and $D$ is compute in order to obtain $x_{i}^{*}=1, i=1, \ldots, N$, as solution.
In Table1 we consider $N=30$. The Chebyshev's method has third order of convergence.
For quadratic equations we can find efficient higher order methods [2].

| Iteration | Cheby. |
| :---: | :---: |
| 1 | $8.41 e-03$ |
| 2 | $1.78 e-09$ |
| 3 | $2.90 e-13$ |
| 4 | $0.00 e+00$ |

Table 1 Error max-norm, $x_{0}$ such that $\left\|x^{*}-x_{0}\right\|=0.1$

### 3.2 The inverse of a matrix

Approximating inverse operators is a very common task in several areas of interest, such as physics, chemistry, engineering, etc. In a general context, we can formulate the following problem: given $g$, we are interested in calculating $f \in \Omega$ such that $H(f)=g$, where $H: \Omega \rightarrow Y$ is an operator defined in a domain $\Omega$ of a Banach space $X$ with values in a Banach space $Y$. It is clear that we have to calculate or approximate the inverse operator $H^{-1}$ for solving the previous equation. In this case, if $g$ is in the domain of $H^{-1}$, there is a solution $f=H^{-1}(g)$.

To approximate the inverse operator $H^{-1}$, we use Newton-type methods, so that better successive approximations to $H^{-1}$ are constructed from an initial approximation. The methods considered in this paper, when they are applied to compute an inverse operator, are not be based on solving linear systems, if not on the product of matrices. Notice that the formulation of the problem in this manner is very interesting.

Let $X$ and $Y$ be two Banach spaces and $G L(X, Y)=\left\{H \in \mathscr{L}(X, Y): H^{-1}\right.$ exists $\}$, where $\mathscr{L}(X, Y)$ is the set of bounded linear operators from the Banach space $X$ into the Banach space $Y$. The problem that we are thinking about is the following: given an operator $H \in G L(X, Y)$, we approximate $H^{-1}$. To do this, we first consider

$$
\mathscr{F}: G L(Y, X) \rightarrow \mathscr{L}(X, Y) \quad \text { and } \quad \mathscr{F}(G)=G^{-1}-H,
$$

so that $H^{-1}$ is the solution of the equation $\mathscr{F}(G)=0$.
If we observe Newton's method,

$$
\left\{\begin{array}{l}
G_{0} \text { given, } \\
G_{n+1}=G_{n}-\left[\mathscr{F}^{\prime}\left(G_{n}\right)\right]^{-1} \mathscr{F}\left(G_{n}\right), \quad n \geq 0,
\end{array}\right.
$$

it is clear that inverse operators are used, but if we take into account the definition of $\mathscr{F}$, calculate $\mathscr{F}^{\prime}\left(G_{n}\right)$ and do

$$
\begin{equation*}
\mathscr{F}^{\prime}\left(G_{n}\right)\left(G_{n+1}-G_{n}\right)=-\mathscr{F}\left(G_{n}\right), \quad n \geq 0, \tag{12}
\end{equation*}
$$

then we can avoid the use of inverse operators for approximating $G_{n+1}$.
Indeed, to obtain the corresponding algorithm, we only need to compute $\mathscr{F}^{\prime}\left(G_{n}\right)$. So, given $G \in G L(Y, X)$, as $G^{-1}$ exists, if

$$
0<\varepsilon<\frac{1}{\|\alpha\|\left\|G^{-1}\right\|}
$$

we have $\|\varepsilon \alpha\|<\frac{1}{\left\|G^{-1}\right\|}$ for $\alpha \in G L(Y, X)$. Therefore, it is known that $G+\varepsilon \alpha \in G L(Y, X)$ and then

$$
\mathscr{F}^{\prime}(G) \alpha=\lim _{\varepsilon \rightarrow 0} \frac{1}{\varepsilon}[\mathscr{F}(G+\varepsilon \alpha)-\mathscr{F}(G)]=-G^{-1} \alpha G^{-1} .
$$

In consequence, Newton's method is now given by the following algorithm:

$$
\left\{\begin{array}{l}
G_{0} \text { given, }  \tag{13}\\
G_{n+1}=2 G_{n}-G_{n} H G_{n}, \quad n \geq 0
\end{array}\right.
$$

In addition, Newton's method does not use inverse operators for approximating an inverse operator and the order of convergence is two.

If we now consider Chebyshev's method,

$$
\left\{\begin{array}{l}
G_{0} \text { given, } \\
G_{n+1}=G_{n}-\left[I+\frac{1}{2} L_{\mathscr{F}}\left(G_{n}\right)\right]\left[\mathscr{F}^{\prime}\left(G_{n}\right)\right]^{-1} \mathscr{F}\left(G_{n}\right), \quad n \geq 0,
\end{array}\right.
$$

where $L_{\mathscr{F}}\left(G_{n}\right)=\left[\mathscr{F}^{\prime}\left(G_{n}\right)\right]^{-1} \mathscr{F}^{\prime \prime}\left(G_{n}\right)\left[\mathscr{F}^{\prime}\left(G_{n}\right)\right]^{-1} \mathscr{F}\left(G_{n}\right)$, we can think that inverse operators are used, but we can do the same as in Newton's method to see that Chebyshev's method does not use them:

$$
\left\{\begin{array}{l}
\mathscr{F}^{\prime}\left(G_{n}\right)\left(P_{n}-G_{n}\right)=-\mathscr{F}\left(G_{n}\right), \quad n \geq 0,  \tag{14}\\
\mathscr{F}^{\prime}\left(G_{n}\right)\left(G_{n+1}-P_{n}\right)=-\frac{1}{2} \mathscr{F}^{\prime \prime}\left(G_{n}\right)\left(P_{n}-G_{n}\right)^{2},
\end{array}\right.
$$

so that we can also avoid the use of inverse operators for approximating $G_{n+1}$.
Let $\alpha, \beta \in G L(Y, X)$ and then

$$
0<\varepsilon<\frac{1}{\|\beta\|\left\|G^{-1}\right\|}
$$

so that $G+\varepsilon \beta \in G L(Y, X)$ and

$$
\mathscr{F}^{\prime \prime}(G) \alpha \beta=\lim _{\varepsilon \rightarrow 0} \frac{1}{\varepsilon}\left[\mathscr{F}^{\prime}(G+\varepsilon \beta) \alpha-\mathscr{F}^{\prime}(G) \alpha\right]=G^{-1} \alpha G^{-1} \beta G^{-1}+G^{-1} \beta G^{-1} \alpha G^{-1} .
$$

In consequence, we write Chebyshev's method as

$$
\left\{\begin{array}{l}
G_{0} \text { given, }  \tag{15}\\
G_{n+1}=3 G_{n}-3 G_{n} H G_{n}+G_{n} H G_{n} H G_{n}, \quad n \geq 0
\end{array}\right.
$$

Observe that Chebyshev's method does not use inverse operators for approximating an inverse operator and the order of convergence is three.
Theorem 8. If $\left\|I-H G_{0}\right\|<1$, Newton and Chebyshev iterative methods are convergent. Moreover, if $H G_{0}=$ $G_{0} H$, then $\lim _{n \rightarrow \infty} G_{n}=H^{-1}$.

On the other hand, if we observe Newton's and Chebyshev's methods, two well-known iterative methods, we conclude that they can approximate inverse operators without using any inverse operator in their application. From this feature of both methods, we pay attention in [5] to the construction of iterative methods of any prefixed order of convergence. Moreover, the final formulation of the methods uses only potencies of matrices that are close to the identity. This fact is crucial to avoid any stability problem in the implementation of the methods. The best method of the family will depend on the particular problem to solve. In any case, we find iterative methods with better behavior than Newton's method. Finally, we would like to emphasize that in the applications, where we are interested in the computation of an inverse operator, our methods use matrix-matrix multiplications with a computational cost similar to that of Gauss-type methods. But, if we are interested only in the application of the inverse operator, we will be able to implement our methods using only matrix-vector multiplications, so that we reduce considerably the computational cost.

### 3.3 The $p$ th root of a matrix

If we consider the complex equation $f(x)=x^{p}-a$ with $a \in \mathbb{C}$, then $L_{f}(x)=\left(\frac{p-1}{p}\right) \frac{x^{p}-a}{x^{p}}$ and Chebyshev's method is reduced to

$$
x_{0} \in D, \quad x_{n+1}=\frac{2 p^{2}-3 p+1}{2 p^{2}} x_{n}+\frac{2 p-1}{p^{2}} a x_{n}^{1-p}-\frac{p-1}{2 p^{2}} a^{2} x_{n}^{1-2 p}, \quad n \geq 0 .
$$

If we extend Chebyshev's method for the computation of the $p$ th root of a matrix, we then consider the space

$$
\Theta=\left\{A \in \mathbb{C}^{r \times r} \text { s.t. } A \text { has no nonpositive real eigenvalues }\right\}
$$

and approximate $A^{\frac{1}{p}}$ for a given matrix $X \in \Theta$. For this, we consider

$$
\mathscr{F}: \Theta \rightarrow \mathbb{C}^{r \times r} \quad \text { and } \quad \mathscr{F}(X)=X^{p}-A
$$

so that $A^{\frac{1}{p}}$ is the solution of the equation $\mathscr{F}(X)=0$. So, Chebyshev's method is reduced to

$$
\begin{equation*}
X_{0} \in \Theta, \quad X_{n+1}=\frac{2 p^{2}-3 p+1}{2 p^{2}} X_{n}+\frac{2 p-1}{p^{2}} A X_{n}^{1-p}-\frac{p-1}{2 p^{2}} A^{2} X_{n}^{1-2 p}, \quad n \geq 0 \tag{16}
\end{equation*}
$$

Chebyshev's method, as Newton's method, cannot be used directly to approximate the principal $p$ th root. Using the same idea as in [15], one can prove that it is not stable in a neighborhood of $A^{\frac{1}{p}}$. A small perturbation on the value of $X_{n}$ is amplified in the following steps and in a finite arithmetic computation the algorithm diverges. This problem can be overcome using another algorithm which provides the same sequence but it is stable in a neighborhood of $A^{\frac{1}{p}}$.

The following iteration, given by Denman and Beavers in [9],

$$
\left\{\begin{array}{l}
X_{0}=A, \quad Y_{0}=I \\
X_{n+1}=\frac{1}{2}\left(X_{n}+Y_{n}^{-1}\right) \\
N_{n+1}=\frac{1}{2}\left(Y_{n}+X_{n}^{-1}\right), \quad n \geq 0
\end{array}\right.
$$

is a stable variant of the Newton method for the matrix square root.
The instability of the simplified Newton iterations $X_{n+1}=\left(X_{n}+A X_{n}^{-1}\right) / 2$ and $X_{n+1}=\left(X_{n}+X_{n}^{-1} A\right) / 2$, shown by Higham [15] for the matrix square root, is mainly due to the one-sided multiplication by $A$. On the other hand, since $X_{n}$ and $A$ commute if $A X_{0}=X_{0} A$, the iteration can be rewritten as

$$
X_{n+1}=\frac{X_{n}+A^{\frac{1}{2}} X_{n}^{-1} A^{\frac{1}{2}}}{2}, \quad n \geq 0
$$

and the iteration becomes stable in this form. However, it is useless since it involves the square root of $A$, but it helps us to stabilize the iteration by introducing the variable $Y_{n}=A^{\frac{1}{2}} X_{n}^{-1} A^{\frac{1}{2}}=A^{-1} X_{n}=X_{n} A^{-1}$. The resulting iteration is that of Denman and Beavers, see [16].

Following these ideas, Iannazo [17] proposes the following two stable versions of the Newton method for the matrix $p$ th root:

$$
X_{n+1}=\frac{(p-1) X_{n}+\left(A^{\frac{1}{p}} X_{n}^{-1}\right)^{p-1} A^{\frac{1}{p}}}{p}, \quad n \geq 0
$$

and

$$
\left\{\begin{array}{l}
X_{0}=I, \quad N_{0}=A \\
X_{n+1}=X_{n}\left(\frac{(p-1) I+N_{n}}{p}\right) \\
N_{n+1}=\left(\frac{(p-1) I+N_{n}}{p}\right)^{-p} N_{n}, \quad n \geq 0
\end{array}\right.
$$

Observe that in the second case, the matrix $A$ does not explicitly appear in the iteration.
Similarly, for Chebyshev's method, we propose

$$
\begin{equation*}
X_{n+1}=\frac{\left(p-\frac{3}{2}+\frac{1}{2 p}\right) X_{n}+\left(2-\frac{1}{p}\right)\left(A^{\frac{1}{p}} X_{n}^{-1}\right)^{p-1} A^{\frac{1}{p}}}{p}-\frac{(p-1)\left(A^{\frac{1}{p}} X_{n}^{-1}\right)^{2 p-1} A^{\frac{1}{p}}}{2 p^{2}}, \quad n \geq 0 \tag{17}
\end{equation*}
$$

and

$$
\left\{\begin{array}{l}
X_{0}=I, \quad N_{0}=A,  \tag{18}\\
X_{n+1}=X_{n}\left(\frac{\left(p-\frac{3}{2}+\frac{1}{2 p}\right) I+\left(2-\frac{1}{p}\right) N_{n}}{p}-\frac{(p-1) N_{n}^{2}}{2 p^{2}}\right), \quad n \geq 0, \\
N_{n+1}=\left(\frac{\left(p-\frac{3}{2}+\frac{1}{2 p}\right) I+\left(2-\frac{1}{p}\right) N_{n}}{p}-\frac{(p-1) N_{n}^{2}}{2 p^{2}}\right)^{-p} N_{n} .
\end{array}\right.
$$

Note that (17) is useless, since it involves the matrix $p$ th root of $A$.
Observe that we need to calculate $X_{n}^{-1}$ at each step of Chebychev's method given by (16). We are interesting in obtaining other expression of Chebyshev's method which avoids the computation of these inverses. So, if we consider the complex function $f(x)=\frac{1}{x^{p}}-\frac{1}{a}=0$, where $f: D \subseteq \mathbb{C} \rightarrow \mathbb{C}$, then algorithm (17) is reduced to

$$
\begin{equation*}
x_{0} \in D, \quad x_{n+1}=x_{n}\left(1+\frac{1}{p}\left(1-\frac{x_{n}^{p}}{a}\right)+\frac{p+1}{2 p^{2}}\left(1-\frac{x_{n}^{p}}{a}\right)^{2}\right), \quad n \geq 0 \tag{19}
\end{equation*}
$$

As a consequence, the algorithm of Chebyshev's method for solving $\mathscr{F}(X)=0$, where $\mathscr{F}: \Theta \rightarrow \mathbb{C}^{r \times r}$, is then

$$
\begin{equation*}
X_{0} \in \Theta, \quad X_{n+1}=X_{n}\left(I+\frac{1}{p}\left(I-A^{-1} X_{n}^{p}\right)+\frac{p+1}{2 p^{2}}\left(I-A^{-1} X_{n}^{p}\right)^{2}\right), \quad n \geq 0 \tag{20}
\end{equation*}
$$

Observe that we only need to calculate $A^{-1}$ and not $X_{n}^{-1}$ at each step, so that algorithm (20) is more efficient than algorithm (16).

For Halley's method, a similar strategy in order to avoid the use of inverses is not possible. Remember that the algorithm of Halley's method,

$$
x_{0} \in D, \quad x_{n+1}=x_{n}-\left(\frac{1}{1+\frac{1}{2} L_{f}\left(x_{n}\right)}\right) \frac{f\left(x_{n}\right)}{f^{\prime}\left(x_{n}\right)}, \quad n \geq 0
$$

always involves inverses (the quotient includes $L_{f}\left(x_{n}\right)$ ). This is the main advantage of Chebyshev's method with respect to Halley's method.

After that, we analyze the convergence of Chebyshev's method. For this, we suppose $A X_{0}=X_{0} A$ and consider the following residual of algorithm (20):

$$
R\left(X_{n}\right)=I-A^{-1} X_{n}^{p}, \quad n \geq 0
$$

To prove the convergence of algorithm (20), we consider a submultiplicative matrix norm $\|\cdot\|$ defined in $\mathbb{C}^{r \times r}$ and prove that $\left\{\left\|R\left(X_{n}\right)\right\|\right\}$ is a scalar decreasing sequence convergent to zero. First of all, from $A X_{0}=X_{0} A$, it follows $A X_{n}=X_{n} A, A X_{n}^{p}=X_{n}^{p} A$ and $A^{-1} X_{n}^{p}=X_{n}^{p} A^{-1}$. As a consequence, we have

$$
R\left(X_{n}\right)=I-\left(I-R\left(X_{n-1}\right)\right)\left(I+\frac{1}{p} R\left(X_{n-1}\right)+\frac{p+1}{2 p^{2}} R\left(X_{n-1}\right)^{2}\right)^{p}
$$

Next, taking into account Villareal's formula,

$$
P(y)^{p}=\left(\sum_{i=0}^{m} a_{i} y^{i}\right)^{p}=\sum_{i=0}^{m p} P_{i} y^{i}
$$

where $P_{0}=a_{0}^{p}, P_{i}=\sum_{j=0}^{i-1} P_{j} \frac{a_{i-j}}{a_{0}} \frac{(i-j)(p+1)-i}{i}$, for $i=1,2, \ldots, m p$, and $a_{i}=0$, for all $i \geq m+1$, it follows that

$$
\begin{equation*}
R\left(X_{n}\right)=I-\left(I-R\left(X_{n-1}\right)\right) \sum_{i=0}^{2 p} P_{i} R\left(X_{n-1}\right)^{i} \tag{21}
\end{equation*}
$$

where $a_{0}=1, a_{1}=\frac{1}{p}, a_{2}=\frac{p+1}{2 p^{2}}, P_{0}=1$ and $P_{i}=\sum_{j=0}^{i-1} P_{j} \frac{a_{i-j}}{a_{0}} \frac{(i-j)(p+1)-i}{i}$, for $i=1,2, \ldots, 2 p$, and $a_{i}=0$, for all $i \geq 3$. In addition, we have $P_{1}=P_{2}=1$ and $P_{3}<1$.

Note that we can always obtain $X_{n} A=A X_{n}$. For this, it is enough to choose $X_{0}=I_{r \times r}$, as Guo does in [11]. Now, we establish the semilocal convergence of Chebyshev's method in the following theorem.

Theorem 9. Let $A \in \mathbb{C}^{r \times r}$ and $X_{0} \in \Theta$ such that $A X_{0}=X_{0} A$ and $\left\|R\left(X_{0}\right)\right\|=\left\|I-A^{-1} X_{0}^{p}\right\|<1$. Suppose that $\left\{P_{i}\right\}_{i=3}^{2 p}$ is a nonincreasing sequence such that $P_{i}$ are nonnegative for all $i=2,3, \ldots, 2 p$. Then, Chebyshev's method defined in (20) converges to $A^{\frac{1}{p}}$. Moreover, $\left\|R\left(X_{n}\right)\right\| \leq\left\|R\left(X_{0}\right)\right\|^{3^{n}}, n \in \mathbb{N}$.

In addition, we observe in Table 2 some representative values of $p$ that help to see that the conditions of Theorem 9 are satisfied.

| $p$ | $\left\{P_{i}\right\}_{i=3}^{2 p}$ |
| :---: | :--- |
| 2 | $0.375,0.1406 \ldots$ |
| 3 | $0.4814 \ldots, 0.2222 \ldots, 0.0493 \ldots, 0.0109 \ldots$ |
| 5 | $0.56,0.296,0.1059 \ldots, 0.0355 \ldots, 0.0080 \ldots$, |
|  | $0.0017 \ldots, 2.0736 \ldots \times 10^{-4}, 2.4883 \ldots \times 10^{-5}$ |
| 7 | $0.5918 \ldots, 0.3294 \ldots, 0.1345 \ldots, 0.0512 \ldots, 0.0151 \ldots, 0.0041 \ldots, 0.0008 \ldots$, |
|  | $0.0001 \ldots, 2.6281 \ldots \times 10^{-5}, 3.6251 \ldots \times 10^{-6}, 2.9592 \ldots \times 10^{-7}, 2.4157 \ldots \times 10^{-8}$ |

Table 2 Sequence $\left\{P_{i}\right\}_{i=3}^{2 p}$ for $p=2,3,5,7$

From Theorem 9, we also deduce that Chebyshev's method has $R$-order of convergence at least three.
In [6], we present a family of high-order iterative methods including both Newton and Chebyshev methods. We find algorithms of the family with better numerical behavior than Newton and Halley methods. These two algorithms are basically the iterative methods proposed in the literature to solve this problem.

## 4 Conclusions

In this paper, we have reviewed some important properties of the classical Chebyshev method. We have point out the possibility to prove its global convergence from its geometric interpretation. We have mentioned the richer dynamics of this scheme in comparison with Newton's method. Finally, we have presented several applications where this high order method is a good alternative to Newton's method. These applications include the approximation of quadratic equations, the approximation of the inverse of a matrix or the $p$ th root of a matrix.

## Acknowledgements

Research supported in part by Programa de Apoyo a la investigación de la Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia 19374/PI/14 and MTM2015-64382-P (MINECO/FEDER).

## References

[1] S. Amat, S. Busquier and J.M. Gutiérrez, (2003), Geometric constructions of iterative functions to solve nonlinear equations, Journal of Computational and Applied Mathematics 157, No 1, 197-205. doi 10.1016/S0377-0427(03)00420-5
[2] S. Amat, S. Busquier and J.M. Gutiérrez, (2006), An adaptive version of a fourth-order iterative method for quadratic equations, Journal of Computational and Applied Mathematics 191, No 2, 259-268. doi 10.1016/j.cam.2005.06.042
[3] S. Amat, S. Busquier, J.M. Gutiérrez and M.A. Hernández, (2008), On the global convergence of Chebyshev's iterative method, Journal of Computational and Applied Mathematics 220, No 1-2, 17-21. doi 10.1016/j.cam.2007.07.022
[4] S. Amat, S. Busquier and S. Plaza, (2007), On the dynamics of a family of third-order iterative functions, The ANZIAM Journal 48, No 3, 343-359. doi 10.1017/S1446181100003539
[5] S. Amat, J. A. Ezquerro and M. A. Hernández-Verón, (2014), Approximation of inverse operators by a new family of high-order iterative methods, Numerical Linear Algebra with Applications 21, No 5, 629-644. doi 10.1002/nla. 1917
[6] S. Amat, J. A. Ezquerro and M. A. Hernández-Verón, (2015), On a new family of high-order iterative methods for the matrix $p$ th root. Numerical Linear Algebra with Applications 22, No 4, 585-595. doi 10.1002/nla. 1974
[7] S. Amat, J. A. Ezquerro and M. A. Hernández-Verón, (2015), Iterative methods for computing the matrix square root, S $\vec{e}$ MA Journal 70, No 1, 11-21. doi 10.1007/s40324-015-0038-9
[8] Professor Cayley, (1879), Desiderata and Suggestions: No. 3. The Newton-Fourier Imaginary Problem, American Journal of Mathematics 2, No 1, 97-97. doi 10.2307/2369201
[9] E. D. Denman and A. N. Beavers Jr., (1976), The matrix sign function and computations in systems, Applied Mathematics and Computation 2, No 1, 63-94. doi 10.1016/0096-3003(76)90020-5
[10] J. H. Curry , L. Garnett and D. Sullivan, (1983), On the iteration of a rational function: Computer experiments with Newton's method, Communications in Mathematical Physics 91, No 2, 267-277. doi 10.1007/BF01211162
[11] C.-H. Guo, (2010), On Newton's method and Halley's method for the principal pth root of a matrix, Linear Algebra and its Applications 432, No 8, 1905-1922. doi 10.1016/j.laa.2009.02.030
[12] J.M. Gutiérrez and M.A. Hernández, (1997), A family of Chebyshev-Halley type methods in Banach spaces, Bulletin of the Australian Mathematical Society 55, No 1, 113-130. doi 10.1017/S0004972700030586
[13] M.A. Hernández and M. A. Salanova, (1993), A family of Chebyshev-Halley type methods, International Journal of Computer Mathematics 47, No 1-2, 59-63. doi 10.1080/00207169308804162
[14] M.A. Hernández and M.A. Salanova, (1998), Chebyshev method and convexity, Applied Mathematics and Computation 95, No 1, 51-62. doi 10.1016/S0096-3003(97)10091-1
[15] N. J. Higham, (1986), Newton's method for the matrix square root, Mathematics of Computation 46 537-549. doi 10.1090/S0025-5718-1986-0829624-5
[16] B. Iannazzo, (2003), A note on computing the matrix square root. Calcolo 40, No 4, 273-283. doi 10.1007/s10092-003-0079-9
[17] B. Iannazzo, (2006), On the Newton Method for the Matrix pth Root, SIAM Journal on Matrix Analysis and Applications 28, No 2, 503-523. doi 10.1137/050624790
[18] K. Kneisl, (2001), Julia sets for the super-Newton method, Cauchy's method, and Halley's method, Chaos: An Interdisciplinary Journal of Nonlinear Science 11, No 2, 359-370. doi 10.1063/1.1368137
[19] A. Melman, (1997), Geometry and Convergence of Euler's and Halley's Methods, SIAM Review 39, No 4, 728-735. doi 10.1137/S0036144595301140
[20] T. R. Scavo and J. B. Thoo, (1995), On the Geometry of Halley's Method, The American Mathematical Monthly 102, No 5, 417-426. doi 10.2307/2975033
[21] E. Schröder, (1870), Ueber iterirte Functionen, Mathematische Annalen 3, No 2, 296-322. doi 10.1007/BF01443992
[22] J.F. Traub, (1964), Iterative Methods for Solution of Equations, Prentice-Hall, Englewood Cliffs, New Jersey.


[^0]:    ${ }^{\dagger}$ Corresponding author.
    Email address: sergio.amat@upct.es
    ${ }^{\ddagger}$ Email address: sonia.busquier@upct.es
    

