Geometric constructions of iterative functions to solve nonlinear equations

S. Amat[a,*,1], S. Busquier[a], J.M. Gutiérrez[b,2]

[a]*Dep. Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena, Paseo de Alfonso XIII, 52, Cartagena (Murcia) 30203, Spain*
[b]*Dep. Matemáticas y Computación, Universidad de La Rioja, Spain*

Abstract

In this paper we present the geometrical interpretation of several iterative methods to solve a nonlinear scalar equation. In addition, we also review the extension to general Banach spaces and some computational aspects of these methods.
© 2003 Elsevier B.V. All rights reserved.

1. Introduction

One of the most important problems in numerical analysis is the solution of nonlinear equations

$$F(x) = 0.$$

To solve these equations we can use iterative methods. Roughly speaking, an iterative method starts from an initial guess x_0 (called *pivot*), which is improved by means of an iteration, $x_{n+1} = \Phi(x_n)$.

In general, an iterative method $x_{n+1} = \Phi(x_n)$ is of order q if the error $|x^* - x_{n+1}|$ is proportional to $|x^* - x_n|^q$ as $n \to \infty$, where x^* is the solution of $F(x) = 0$. Consequently, as higher the order

---

is, higher the velocity of convergence is. However, the operational cost of a method also increases with the order. This fact leads to find an equilibrium between the high velocity and the operational cost. This equilibrium used to be attained with the well-known Newton's method, that has quadratic convergence (order two). In fact, Newton's method and similar second-order methods are the most used to solve this problem and probably this situation has led to the wrong idea that higher-order methods have no more than theoretical interest. Of course, third-order methods require more computational cost than other simpler methods, which makes them disadvantageous to be used in general. But, in some cases they can be considered in practice, as we show later.

In this paper we deal with third-order methods. First, we present the geometrical interpretation of several of them, by extending the geometrical construction of a few well-known third-order methods such as Halley's or Euler's [14,17,18]. Next, we consider their extension to Banach spaces and analyze some computational aspects. Finally, in Section 4 a numerical experiment is presented.

2. The real case

In this section, we analyze the geometric construction of different iterative processes to solve a scalar nonlinear equation

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

The geometric interpretation of Newton's method is well-known: given an iterate x_n, we calculate the tangent line

$$y(x) - f(x_n) = f'(x_n)(x - x_n),$$

to the graph of f at $(x_n, f(x_n))$, and the next iterate is the zero of this tangent line:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \geqslant 0. \tag{2}$$

There exists a huge number of works studying the convergence of (2) to a root of (1) and other characteristics of this sequence. But this is not the aim of this paper. Here we only analyze the geometric derivation of other methods if, instead a straight line, we consider other tangent curves to the graph of f at $(x_n, f(x_n))$.

The tangent line in Newton's method can be seen as the first-degree Taylor polynomial of f at x_n. Another well-known construction consists in considering the second-degree polynomial, that is, the parabola

$$y(x) - f(x_n) = f'(x_n)(x - x_n) + \frac{f'(x_n)}{2}(x - x_n)^2. \tag{3}$$

The point x_{n+1} where the graph of y intersects the x-axis gives us the following sequence:

$$x_{n+1} = x_n - \frac{2}{1 + \sqrt{1 - 2L_f(x_n)}} \frac{f(x_n)}{f'(x_n)}, \quad n \geqslant 0,$$

where

$$L_f(x) = \frac{f(x)f''(x)}{f'(x)^2}.$$

This method is called *Euler's method* or *irrational Halley's method* and it has been studied, for instance, in [19] or [14]. Euler's method has cubic convergence (order three).

If we consider instead of the parabola (3) an hyperbola in the following way:

$$axy + y + bx + c = 0,$$

and we impose the tangency conditions

$$y(x_n) = f(x_n), \quad y'(x_n) = f'(x_n) \quad \text{and} \quad y''(x_n) = f''(x_n), \tag{4}$$

we have

$$y - f(x_n) - f'(x_n)(x - x_n) - \frac{f''(x_n)}{2f'(x_n)}(x - x_n)(y - f(x_n)) = 0.$$

The corresponding iterative process (again a third-order one) is

$$x_{n+1} = x_n - \frac{2}{2 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)}, \quad n \geqslant 0,$$

that is the well-known *Halley's method*. About this method Traub says in [19, pp. 91]: "Halley's method must share with the secant method the distinction of being the most frequently rediscovered methods in the literature". This assertion is confirmed by the big number of publications about it, as for example, [14,18] and the references therein. The geometric derivation of Halley's method was suggested in [17].

These two methods are, probably the two best-known third-order methods, but they are not the only ones. There are others whose geometric interpretation is not as well known, if known at all. For instance, *Chebyshev's method* is obtained by quadratic interpolation of the inverse function of f, in order to approximate $f^{-1}(0)$ [19]. But it also admits a geometric derivation, from a parabola in the form

$$ay^2 + y + bx + c = 0,$$

that after the imposition of tangency conditions (4) can be written

$$-\frac{f''(x_n)}{2f'(x_n)^2}(y - f(x_n))^2 + y - f(x_n) - f'(x_n)(x - x_n) = 0.$$

The expression of Chebyshev's method is then

$$x_{n+1} = x_n - \left(1 + \frac{1}{2}L_f(x_n)\right)\frac{f(x_n)}{f'(x_n)}, \quad n \geqslant 0.$$

Now, we consider hyperbolas in the form

$$ay^2 + bxy + y + cx + d = 0,$$

or, equivalently,

$$x = -\frac{y + ay^2 + d}{by + c}.$$

For convenience and taking into account that the hyperbola pass through the point $(x_n, f(x_n))$, it can be written in the equivalent form

$$x - x_n = -(y - f(x_n))\frac{1 + a_n(y - f(x_n))}{b_n(y - f(x_n)) + c_n}.$$

The rest of tangency conditions (4) allow us to fix the value of $c_n = -f'(x_n)$ and to settle the following relation between a_n and b_n,

$$a_n = -\frac{f''(x_n)}{2f'(x_n)^2} - \frac{b_n}{f'(x_n)}.$$

The next iterate x_{n+1} is the intersection of these hyperbolas with the x-axis:

$$x_{n+1} = x_n - \left(1 + \frac{1}{2}\frac{L_f(x_n)}{1 + b_n(f(x_n)/f'(x_n))}\right)\frac{f(x_n)}{f'(x_n)}, \quad n \geqslant 0, \tag{5}$$

where b_n is a parameter depending on n.

We have obtained a family of third-order methods that includes, as particular cases, the following ones:

1. The famous methods of Chebyshev and Halley are obtained for $b_n = 0$ and $b_n = -f''(x_n)/(2f'(x_n))$, respectively.
2. Another third-order method, not as famous as the previous ones, is *super-Halley method* [10]:

$$x_{n+1} = x_n - \left(1 + \frac{1}{2}\frac{L_f(x_n)}{1 - L_f(x_n)}\right)\frac{f(x_n)}{f'(x_n)}, \quad n \geqslant 0.$$

   It also appears in our family for $b_n = -f''(x_n)/f'(x_n)$.
3. Even more, for $b_n = -\alpha f''(x_n)/f'(x_n)$, $\alpha \in \mathbb{R}$, we obtain the family of third-order methods studied in [9]:

$$x_{n+1} = x_n - \left(1 + \frac{1}{2}\frac{L_f(x_n)}{1 - \alpha L_f(x_n)}\right)\frac{f(x_n)}{f'(x_n)}, \quad n \geqslant 0.$$

4. Finally, as a limit case, when $b_n = \pm\infty$, Newton's method is obtained.

By using the previous ideas, we can obtain other iterative processes by considering different osculating curves. For instance, if we take now the cubic function

$$a_n(y - f(x_n))^3 + b_n(y - f(x_n))^2 + (y - f(x_n)) + d_n(x - x_n) = 0,$$

and we impose the tangency conditions (4), we have

$$a_n = C\frac{f''(x_n)^2}{(f'(x_n))^2},$$

$$b_n = -\frac{f''(x_n)}{2f'(x_n)^2},$$

$$d_n = -f'(x_n).$$

So we have obtained another family of third-order methods:

$$x_{n+1} = x_n - \left(1 + \frac{1}{2}L_f(x_n) + CL_f(x_n)^2\right)\frac{f(x_n)}{f'(x_n)}, \quad n \geqslant 0.$$

They are called C-methods and they have been studied in [2].

As a final remark in this section, notice that all methods obtained here can be written in the form

$$x_{n+1} = x_n - \left(1 + \frac{1}{2}L_f(x_n) + O(L_f(x_n)^2)\right)\frac{f(x_n)}{f'(x_n)}, \quad n \geqslant 0,$$

as Altman pointed out in [1].

3. On the generalization to Banach spaces

In a formal way, most of the methods studied in the previous section can be extended to Banach spaces, by writing inverse operators instead of quotients. For instance, the expression of methods (5) can be written in terms of linear operators as

$$x_{n+1} = x_n - (I + \tfrac{1}{2}(I + b_n F'(x_n)^{-1}F(x_n))^{-1}L_F(x_n))F'(x_n)^{-1}F(x_n),$$

where I is the identity operator on X and for each $x \in X$, $L_F(x)$ is a linear operator on X defined by

$$L_F(x) = F'(x)^{-1}F''(x)F'(x)^{-1}F(x),$$

assuming that $F'(x)^{-1}$ exists. Taking this into account, the above sequence defined in Banach spaces is well defined if $b_n F'(x_n)^{-1}F(x_n)$ are linear operators on X. A reasonable choice for the parameters b_n is then $b_n = -\alpha F'(x_n)^{-1}F''(x_n)$, with α being a real number, although b_n could be any bilinear operator from $X \times X$ to X.

The main practical difficulty related to the class of methods we analyze is the evaluation of the second-order Fréchet derivative. For a nonlinear system of N equations and N unknowns, the first Fréchet derivative is a matrix with N^2 values, while the second Fréchet derivative has N^3 values. This implies a huge amount of operations in order to evaluate every iteration. To overcome these difficulties many authors have considered methods that do not use the second derivative. An alternative is given by the quasi-Newton methods or by methods that evaluate several times the function and its first derivative. A complete survey about these methods appears in [13]. Nevertheless, there are also quite articles that studies third-order methods in Banach spaces, mainly Halley or Chebyshev methods. As a sample, we can cite [3–6,8,12] or [21].

In [16], this (two-step) third-order recurrence is proposed:

$$y_{n+1} = x_n - F'(x_n)^{-1}F(x_n),$$
$$x_{n+1} = y_{n+1} - F'(x_n)^{-1}F(y_{n+1}), \quad n \geqslant 0. \tag{6}$$

This method is, in general, cheaper than any third-order methods requiring the evaluation of the second derivative and then very interesting from the practical point of view. We consider it in this paper for numerical comparisons.

Notice that the two-step method (6) can also be obtained from the generalization of (5) if we consider b_n the bilinear operator from $X \times X$ into X such that

$$b_n F(y_{n+1})F'(x_n)^{-1}F(x_n) = \tfrac{1}{2}L_F(x_n)F(x_n) - F(y_{n+1}).$$

Then it is easy to check that the different iterations of the two-step method can be interpreted also as the root of the hyperbolas

$$x - x_n = -(y - F(x_n)) \frac{1 + a_n(y - F(x_n))}{b_n(y - F(x_n)) + c_n}.$$

Of course, in practice it has used version (6) of the scheme.

But in some cases the second derivative is easy to evaluate and the methods given in the previous section can be also used in practice. For instance:

Example 1 (Quadratic equations). For these equations the second derivative is constant. Therefore, the family of methods (5) for $b_n = -\alpha F'(x_n)^{-1} F''(x_n)$ can be written [7] as

$$y_n = x_n - F'(x_n)^{-1} F(x_n),$$

$$z_n = x_n + \theta(y_n - x_n), \quad 0 \leqslant \theta \leqslant 1,$$

$$x_{n+1} = y_n - F'(z_n)^{-1} F(y_n).$$

The value $\theta = 0$ corresponds with Chebyshev method, $\theta = \frac{1}{2}$ with Halley method and $\theta = 1$ with super-Halley method. Notice that for quadratic equations, the two-step method agrees with Chebyshev method.

In addition, the fact of having a family of methods allows us to use one or another depending of our interest. For instance, if we are interested in calculating as less inverses as possible, in order to simplify the calculations, we can use Chebyshev method.

Example 2 (Calculus of the inverse). Chebyshev method allows us to obtain an algorithm for calculating the inverse of a regular matrix A. Indeed, if we apply the method to the equation $F(x) = x^{-1} - A$ we have

$$x_{n+1} = (3I_n - (3I_n - x_n A) x_n A) x_n,$$

and we only need to compute several matrix multiplications.

If we are interested in the velocity of convergence, we can choose the value of b_n in (5) in order to increase the order the convergence. So, if $b_n = b(x_n)$ with

$$b(x) = -\alpha(x) \frac{f''(x)}{f'(x)}, \quad \alpha(x) = 1 - \frac{L_{f'}(x)}{3},$$

we obtain a fourth-order method. This value is, in general, difficult to calculate and it gives rise to very complicated methods. But in particular cases we obtain further information. For instance, for quadratic equations (Example 1) $\alpha(x) = 1$. Then super-Halley method has order four for such equations. Another interesting example is the following one.

Example 3 (Calculus of the pth root, Gutiérrez and Hernández [11]). To calculate the pth root of a positive number a is equivalent to solve the equation $f(t) = 0$, with $f(t) = t^p - a$. In this case,

$L_{f'}(x) = (p-2)/(p-1)$ and then

$$\alpha(x) = \frac{2p-1}{3(p-1)}.$$

The corresponding method can also be used to approximate the pth root of an operator A. This type of problems are related with many systems of partial differential equations. In particular, the solution of the following problem

$$x^{(\text{iv})}(t) + Ax(t) = 0,$$

$$x(0) = x_0$$

is $x(t) = \exp(-A^{1/4}t)x_0$. Here we have to calculate the fourth-root of a matrix A.

Finally, for the C-methods the order of convergence can be increased by taking $C = C(x) = [(1 - L_{f'}(x))/3]/2$. In the above cases C is constant: $C = \frac{1}{2}$ for quadratic equations, $C = \frac{1}{4}$ for the calculus of inverses and $C = (2p-1)/(6(p-1))$ for the calculus of pth roots.

4. A numerical experiment

In this section, we are interested in pointing out a class of equations where the third-order methods studied are a good alternative to the Newton and two-step methods.

We shall consider an important special case of integral equation, the Hammerstein equation (see [15])

$$u(s) = \psi(s) + \int_0^1 H(s,t)f(t,u(t))\,dt. \tag{7}$$

These equations are related with boundary value problems for differential equations. For some of them, third-order methods using second derivatives are useful for their effective (discretized) solution.

The discrete version of (7) is

$$x^i = \psi(t_i) + \sum_{j=0}^m \gamma_j H(t_i,t_j)f(t_j,x^j), \quad i = 0,1,\dots,m, \tag{8}$$

where $0 \leqslant t_0 < t_1 < \cdots < t_m \leqslant 1$ are the grid points of some quadrature formula $\int_0^1 f(t)\,dt \approx \sum_{j=0}^m \gamma_j f(t_j)$, and $x^i = x(t_i)$.

Let us consider the Hammerstein equation studied in [20]:

$$x(s) = 1 - \frac{1}{4}\int_0^1 \frac{s}{t+s}\frac{1}{x(t)}\,dt, \quad s \in [0,1]. \tag{9}$$

Using the trapezoidal rule of integration with step $h = 1/m$, we obtain the following system of nonlinear equations:

$$0 = x^i - 1 + \frac{1}{4m}\left(\frac{1}{2}\frac{t_i}{t_i+t_0}\frac{1}{x^0} + \sum_{k=0}^n \frac{t_i}{t_i+t_k}\frac{1}{x^k} + \frac{1}{2}\frac{t_i}{t_i+t_m}\frac{1}{x^m}\right),$$

$$i = 0,1,\dots,m, \tag{10}$$

where $t_j = j/m$.

Table 1
Exact solution of (10) with $m = 20$

$9.658340375548916e - 001$	$8.383952084700058e - 001$
$9.418615742240362e - 001$	$8.331433414479326e - 001$
$9.231204383172876e - 001$	$8.283148314256238e - 001$
$9.077427356874352e - 001$	$8.238581489055439e - 001$
$8.947533453995619e - 001$	$8.197301553818098e - 001$
$8.835609377538635e - 001$	$8.158943977068417e - 001$
$8.737737155176020e - 001$	$8.123198157064024e - 001$
$8.651162715111133e - 001$	$8.089797476185554e - 001$
$8.573866000336692e - 001$	$8.058511541752204e - 001$
$8.504316843674818e - 001$	$8.029140058401950e - 001$
$8.441326442432634e - 001$	

Table 2
$x_0 = 1.5$, l_∞-error, $m = 20$

Iter.	Newton	Chebyshev	Halley	2-step
1	0.0378	0.0180	0.0175	0.0042
2	$2.33e - 04$	$1.26e - 06$	$1.03e - 06$	$4.05e - 09$
3	$9.98e - 09$	0	0	0
4	0			

Table 3
$x_0 = 1.5$, l_∞-error, $m = 20$

Iter.	$C = \frac{1}{4}$	$C = \frac{1}{2}$	$C = 1$	$C = 2$	$C = 4$
1	0.0175	0.0171	0.0162	0.0144	0.0108
2	$1.03e - 06$	$8.26e - 07$	$4.79e - 07$	$1.43e - 08$	$2.82e - 07$
3	0	0	0	0	0

In this case, the second Fréchet derivative is diagonal by blocks. In particular, the two-step method is more expensive.

We consider $m = 20$ in the quadrature trapezoidal formula. The exact solution is given in Table 1 and it is computed numerically by Newton method.

In Tables 2–4, we compare the obtained results with those of the two-step method proposed in [16].

Summing up, in this paper we have studied a wide class of third-order methods, those evaluating the second Fréchet derivative. We established their geometric interpretation, and we analyzed their behavior in those cases where the evaluation of the second derivative is not very time consuming.

Table 4
$x_0 = 1.5$, l_∞-error, $m = 20$

Iter.	$C = 6$	$C = 8$	$C = 10$	$C = 12$	$C = 14$
1	0.0072	0.0036	$2.98e - 04$	0.0035	0.0071
2	$1.79e - 07$	$3.79e - 08$	$1.77e - 11$	$4.54e - 08$	$4.93e - 07$
3	0	0	0	0	0

References

[1] C. Altman, Iterative methods of higher order, Bull. Acad. Poll. Sci. (Sér. Sci. Math., Astr. Phys.) IX (1961) 62–68.

[2] S. Amat, S. Busquier, V.F. Candela, F.A. Potra, Convergence of third order iterative methods in Banach spaces, Vol. 16, U.P. Cartagena, 2001, preprint.

[3] I.K. Argyros, D. Chen, Results on the Chebyshev method in Banach spaces, Proyecciones 12 (2) (1993) 119–128.

[4] V.F. Candela, A. Marquina, Recurrence relations for rational cubic methods I: the Halley method, Computing 44 (1990) 169–184.

[5] V.F. Candela, A. Marquina, Recurrence relations for rational cubic methods II: the Chebyshev method, Computing 45 (1990) 355–367.

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

[7] J.A. Ezquerro, J.M. Gutiérrez, M.A. Hernández, M.A. Salanova, Resolución de ecuaciones de Riccati algebraicas mediante procesos iterativos de tercer orden, Proceedings of XVI CEDYA—VI CMA, Las Palmas de Gran Canarias, Spain, 1999, pp. 1069–1076 (in Spanish).

[8] W. Gander, On Halley's iteration method, Amer. Math. Monthly 92 (1985) 131–134.

[9] J.M. Gutiérrez, M.A. Hernández, A family of Chebyshev–Halley type methods in Banach spaces, Bull. Austral. Math. Soc. 55 (1997) 113–130.

[10] J.M. Gutiérrez, M.A. Hernández, An acceleration of Newton's method: super-Halley method, Appl. Math. Comput. 117 (2001) 223–239.

[11] J.M. Gutiérrez, M.A. Hernández, Calculus of nth roots and third order methods, Nonlinear Anal. 47 (2001) 2875–2880.

[12] M.A. Hernández, A note on Halley's method, Numer. Math. 59 (3) (1991) 273–279.

[13] J.M. Martínez, Practical quasi-Newton methods for solving nonlinear systems, J. Comput. Appl. Math. 124 (1–2) (2000) 97–121.

[14] A. Melman, Geometry and convergence of Euler's and Halley's methods, SIAM Rev. 39 (4) (1997) 728–735.

[15] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[16] F.A. Potra, V. Pták, Nondiscrete induction and iterative processes, in: Research Notes in Mathematics, Vol. 103, Pitman, Boston, 1984.

[17] G.S. Salehov, On the convergence of the process of tangent hyperbolas, Dokl. Akad. Nauk SSSR 82 (1952) 525–528 (in Russian).

[18] T.R. Scavo, J.B. Thoo, On the geometry of Halley's method, Amer. Math. Monthly 102 (1995) 417–426.

[19] J.F. Traub, Iterative Methods for Solution of Equations, Prentice-Hall, Englewood Cliffs, NJ, 1964.

[20] W. Werner, Newton-like methods for the computation of fixed points, Comput. Math. Appl. 10 (1) (1984) 77–86.

[21] T. Yamamoto, On the method of tangent hyperbolas in Banach spaces, J. Comput. Appl. Math. 21 (1988) 75–86.