Monografías del Semin. Matem. García de Galdeano. **27:** 315–322, (2003).

Adapted construction of iterative processes for calculus the N-th root.

M. A. Hernández and N. Romero

Dpt. Mathematics and Computation. University of La Rioja.

C/ Luis de Ulloa s/n. 26004 Logroño, Spain.

E-mail: mahernan@dmc.unirioja.es, natalia.romero@dmc.unirioja.es

Abstract

In this paper, we construct a modification of Newton's method to accelerate the convergence of this method to the approximation of the positive n-th root of a positive real number. From this modification, we can define a new iterative process with prefixed order $q \in \mathbb{N}$, $q \geq 2$.

keywords: Iterative processes, order of convergence, convergence semilocal, n-th root.

A.M.S Classification: 65H05, 65B99.

Supported in part by a grant from the DGES (ref. PB-98-0198) and a grant from the University of La Rioja (ref. API-00/B16).

1 Introduction

The problem of the approximation of the n-th root of a positive real number R, has been widely studied by different authors (see [1],[2]). This problem is equivalent to consider the function $f(t) = t^n - R$ and solve the equation

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

We denote by s a solution of this non-linear equation in a given interval. There are many ways to solve (1). So, for instance, in the particular case of the calculation of square roots, we have the famous Heron's formula (75 b.C. approx.):

$$t_{k+1} = \frac{1}{2}\left(t_k + \frac{R}{t_k}\right), \qquad k \geq 0.$$

According to some authors (see [3]), this algorithm was known by the Mesopotamian civilizations almost two thousand years before Christ. Heron also obtained formulas to calculate cubic roots and higher ones.

The French algebrist F. Viète (1540-1603), studied the problem of the approximation of roots in non-linear equations. His method was based on the consideration of a polynomical equation, with real coefficients and one real root in the following form:

$$a_1 t + a_2 t^2 + a_3 t^3 + \cdots + a_n t^n = a_0.$$

Then, for the initial approximation t_0, yield $\sum_{i=1}^{n} a_i t_0^i < a_0$, it is considered that $t = t_0 + s_1$ and we obtain the equation $\sum_{i=1}^{n} b_i s_1^i = b_0$, from which we get another approximation in the same conditions of t_0 and we denote it by t_1. So, for a recurrent procedure we can obtain an approach to the solution of the polynomical equation. This approach of approximation plus correction has been essential in the development of the construction of processes to approximate the solutions of an equation. One of the methods more widely used is the Newton's method [4], which is defined by the following iterative process: Given an initial approximation t_0, we consider

$$t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)}, \qquad k \geq 0.$$

This iterative process still follows the approach of approximation plus correction already considered by Viète. This method and its extension to the solution of systems of non-linear equations are the basis of the most frequently used techniques to solve non-linear algebraic equations. This iterative process has the advantage of being easy to apply for any non-linear equation, but it needs the derivability of the function. This construction has been smoothed considering approximations of the derivate in the following way:

$$f'(t_k) \sim \frac{f(t_k) - f(t_{k-1})}{t_k - t_{k-1}}.$$

In this case, we obtain the Secant method [4]. It is an established fact [5] that the convergence of the Newton method is quadratic, at least for a t_0 sufficiently close to the solution. A modern approach is the one proposed by Dubeau [5], who applied the Newton's method to functions

$$p(t) = t^{\beta - n}(t^n - R),$$

where $\beta \in \mathbb{R}$ is looked for in order to get cubic convergence. In fact, the value $\beta = (n+1)/2$ has been obtained as the appropiate for this purpose.

2 Construction of iterative processes with a fixed convergence order

In this section our aim is to construct new iterative processes to obtain an approximation to the solution of (1). The construction of the methos is adapted according to the order of convergence that we want to obtain. This problem has been studied by Neta [6], in the case of square roots. This construction is based on Gander's idea [7], which is applied to equation (1). Gander uses the following result, with f and H having a sufficient number of continuous derivatives in a neighborhood of s.

Let s be a simple zero of f and H any function such that $H(0) = 1$, $H'(0) = \frac{1}{2}$ and $|H''(0)| < \infty$. The iteration $t_{k+1} = F(t_k)$ with $F(t) = t - H(L_f(t))\frac{f(t)}{f'(t)}$, where $L_f(t) = \frac{f(t)f''(t)}{f'(t)^2}$, is of third order.

According to the result, we suggest the family of iterative processes in the form:

$$t_{k+1} = G(t_k) = t_k - \left(1 + \frac{1}{2}L_f(t_k) + \sum_{i\geq 2} a_i L_f(t_k)^i\right)\frac{f(t_k)}{f'(t_k)}, \tag{2}$$

for an initial approximation t_0, is formed by iterative processes with at least cubic convergence. In study [8], is considered an uniparametric family of iterative methods, included in the previous family, which let them to obtain, an iterative method with order four, fixing to the value of the parameter. Now, if we consider algorithm (2), we can truncate the series and obtain then the following algorithm:

$$t_{k+1} = G(t_k) = t_k - \left(1 + \frac{1}{2}L_f(t_k) + \sum_{i=2}^{q-2} a_i L_f(t_k)^i\right)\frac{f(t_k)}{f'(t_k)}, \quad k \geq 0, \tag{3}$$

for an initial approximation t_0, which gives us a $(q-3)$-parametric family of iterative processes. Then, we try to construct a process which allows us to obtain the highest order of convergence to solve (1), and then calculate the n-th root of a positive real R number. For this, we consider the first elemental result of the analysis.

Lemma 1

Let be a function $h : [a, b] \longrightarrow \mathbb{R}$ with $h \in C^{(k+2}(a, b)$. If s exist, with $s \in (a, b)$ and $h(s) = 0$, then the function $F(t) = h(t)^{k+1}$ has at least its first k derivates null in s.

Proof.

From the assumptions, we can consider $h(t) = (t-s)^m g(t)$, where $g : [a, b] \longrightarrow \mathbb{R}$ with $g(s) \neq 0$ and $m \geq 1$. Then, it follows that $F(t) = (t-s)^{m(k+1)} g(t)^{k+1}$ and, from Leibnitz's formula to calculate the successive derivates in the product of funtions, we obtain:

$$F^{(j}(t) = \sum_{i=0}^{j} \binom{j}{i} ((t-s)^{m(k+1)})^{(i}(g(t)^{k+1})^{(j-i} =$$

$$= \sum_{i=0}^{j} \binom{j}{i} ((k+1)m)((k+1)m-1)\cdots((k+1)m-i+1)((t-s)^{m(k+1)-i})(g(t)^{k+1})^{(j-i}.$$

Therefore, for $1 \leq j \leq (k+1)m-1$, it follows that $F^{(j}(s) = 0$.

Besides, for $j = (k+1)m$, we get $F^{((k+1)m)}(s) = ((k+1)m)!$, and since $m \geq 1$, the result is obtained. ■

Summing up these results and the Gander's result, we will consider the iterative processes of (3), which are at least cubically convergent, but now we want to increase the order of convergence, using the parameters a_i, $i = 2, \cdots, q-2$, when they are applied to solve equation (1). In [9], we see that, for a free parameter, we get order of convergence four. Then, in these conditions, we wonder whether it is possible, to obtain the optimum q order of convergence. It is known that, for an iterative process in the form $t_{k+1} = G(t_k)$ to approach the solution of (1), if G is a sufficiently differentiable function, provides a convergence method of q-th order if:

$$G(s) = s, \qquad G'(s) = G''(s) = G^{(3}(s) = \cdots = G^{(q-1}(s) = 0, \qquad G^{(q}(s) \neq 0,$$

where s is the root of (1). Then, applying the Gander's result, we get

$$G(s) = s, \qquad G'(s) = G''(s) = 0.$$

To study the remaining conditions, we consider the following equalities:

$$\frac{d}{dt}(L_f(t))\frac{f(t)}{f'(t)} = [1 + L_f(t)L_{f'}(t) - 2L_f(t)]L_f(t), \tag{4}$$

$$\frac{d}{dt}\left(\frac{f(t)}{f'(t)}\right) = 1 - L_f(t). \tag{5}$$

Then, we obtain an exprexion of $G'(t)$ depending on $L_f(t)$. As $L_f(s) = 0$, applying Lemma 1, we can to obtain conditions upon the parameters a_i, according to the order of convergence that we want to obtain.

Theorem 2

The iterative process of (3) has q order of convergence if

$$a_i = \frac{(in-1)((i-1)n-1)\cdots(2n-1)}{(i+1)!(n-1)^{i-1}}, \qquad 2 \leq i \leq q-2. \tag{6}$$

Proof.

From (3), we get

$$G'(t) = 1 - \left(\frac{1}{2} + \sum_{i=2}^{q-2} a_i i L_f(t)^{i-1}\right) L'_f(t)\frac{f(t)}{f'(t)} - \left(1 + \frac{1}{2}L_f(t) + \sum_{i=2}^{q-2} a_i L_f(t)^i\right)(1 - L_f(t)).$$

Hence, using equalities (4) and (5) we obtain:

$$G'(t) = \left(\frac{3}{2} - \frac{1}{2}L'_f(t) - 3a_2\right) L_f(t)^2 + \left(2q - 3 - (q-2)L'_f(t)\right) a_{q-2} L_f(t)^{q-1} +$$

$$+\sum_{i=3}^{q-2}\left[((2i-1)-(i-1)L'_f(t))a_{i-1}-(1+i)a_i\right]L_f(t)^i.$$

If we use (6) and

$$L_{f'}(t)=\frac{n-2}{n-1},$$

then it follows

$$G'(t)=\frac{((q-1)n-1)((q-2)n-1)\cdots(2n-1)}{(q-1)!(n-1)^{q-2}}L_f(t)^{q-1}. \tag{7}$$

Next, applying Lemma 1, to $h(t)=L_f(t)=(t-s)g(t)$, we have:

$$G''(s)=G^{(3}(s)=\cdots=G^{(q-1}(s)=0, \qquad G^{(q}(s)\neq 0.$$

The proof is then complete. ■

Therefore, we have obtained the optimum situation according to the parameters. Then, from (1), this is equivalent to fix n and define a new iterative process with prefixed q order $q\in\mathbb{N}$ $q\geq 2$, for an initial t_0, in the following way:

$$t_{k+1}=t_k-\left(1+\frac{1}{2}L_f(t_k)+\sum_{i=2}^{q-2}\frac{(in-1)((i-1)n-1)\cdots(2n-1)}{(i+1)!(n-1)^{i-1}}L_f(t_k)^i\right)\frac{f(t_k)}{f'(t_k)},\ k\geq 0. \tag{8}$$

3 Monotonous Convergence

In this section, we will study the semilocal monotonous convergence of the iterative process that we have constructed in the previous section, which is defined in (8).

Note that, for the iterative process $t_{k+1}=G(t_k)$, we have obtained the value of $G'(t)$ in (7). Therefore, from the previous reasoning and considering the sign of $L_f(t)$ it is easy to prove the following result:

Theorem 3

Let t_0 be with $f(t_0)>0$. The iterative process given in (7) defines a $\{t_k\}$ monotonous sequence which converges to the solution s with q order of convergence.

Proof. Note that since $R>0$, we can consider two situations for n: even or odd. When n is even, there are two roots r_1 and r_2 such that $r_2=-r_1$. Therefore, there are two possible cases: $t_0<r_1$ or $t_0>r_2$.

If $t_0>r_2$: First, we observe, for a_i given in (6), that

$$t_1-t_0=-\left(1+\frac{1}{2}L_f(t_0)+\sum_{i=2}^{q-2}a_iL_f(t_0)^i\right)\frac{f(t_0)}{f'(t_0)}\leq 0,$$

and then $t_1\leq t_0$.

Besides, since $G'(t) > 0$ in (s, t_0), it follows immediately that:

$$t_1 - s = G(t_0) - G(s) = G'(\theta_0)(t_0 - s) \geq 0, \quad \theta_0 \in (s, t_0),$$

and then $t_1 > s$.

Supposing by an analogous reasoning, that $t_{k-1} > t_k$ and $t_k > s$ we obtain $t_{k+1} < t_k$ and $t_{k+1} > s$. Then, there exists $l = \lim_{k \to +\infty} t_k$ and, by letting $k \to +\infty$ in (8), it follows $f(l) = 0$ and therefore $l = s$.

If $t_0 < r_1$, then $L_f(t_0) > 0$ and the proof is analogous.

When n is odd, as $f(t_0) > 0$, then $t_0 > s$ and $L_f(t_0) > 0$, and therefore the proof is similar to the one applied when n is even.

Besides, according to Theorem 3, iterative process (8) has q order of convergence. ■

4 Numerical tests

In this section, from a simple problem such as the calculation of 4-th root for the number $R = 5040$, we observe the behavior of the new iterative processes defined in (8). We choose different starting points: $t_0 = 100$, $t_0 = 1000$ and $t_0 = 5040$, and consider several orders of convergence: $q = 25$, $q = 100$, $q = 200$ and $q = 500$. The results obtained are shown in tables 1, 2, 3 and 4.

k	$q = 25$	$q = 25$	$q = 25$
0	100	1000	5040
1	36.74074352765773	367.2594078713632	1850.987341155527
2	13.78793737712009	134.8797661648172	679.7924898159161
3	8.432497797757524	49.54189526835151	249.6602243561512
4	8.425731861221042	18.31600608010673	91.69101404317729
5		8.699152481929406	33.69358878768427
6		8.425731861221042	12.75408517346861
7			8.426787834656201
8			8.425731861221042

Table 1. Order $q = 25$

k	$q = 100$	$q = 100$	$q = 100$
0	100	1000	5040
1	25.88912937297498	258.4619174345599	1302.647847549595
2	8.697071398569527	66.8050413187296	336.6848238695849
3	8.42573186122104	17.40997673500065	87.02131700814318
4	8.425731861221042	8.426343403916963	22.55680503055836
5		8.425731861221042	8.496264562007763
6			8.425731861221042

Table 2. **Order** $q = 100$

k	$q = 200$	$q = 200$	$q = 200$
0	100	1000	5040
1	21.7893786702938	217.1693682186719	1094.533250464947
2	8.428058184376935	47.16960671889299	237.6990704001984
3	8.42573186122104	10.90692375403867	51.62634535346908
4	8.425731861221042	8.425731861221042	11.72388342496529
5			8.425731861221042

Table 3. **Order** $q = 200$

k	$q = 500$	$q = 500$	$q = 500$
0	100	1000	5040
1	17.40667310616728	172.627448849747	870.0416139602313
2	8.425731861221051	29.82832538422122	150.1931580361902
3	8.425731861221042	8.438272160150252	25.97010873665306
4		8.425731861221042	8.42637570583592
5			8.425731861221042

Table 4. **Order** $q = 500$

In view of the results, see Figure 1, we can deduce that, from a fixed order and according to the starting point considered, it is not necessary to increase the order of convergence, since the computational cost increases and the speed of convergence is not improved.

According to the tests carried out, it seems that an order of convergence between 100 and 200 is enough to obtain optimum speed of convergence. The study of the optimum speed of convergence is open to further studies.

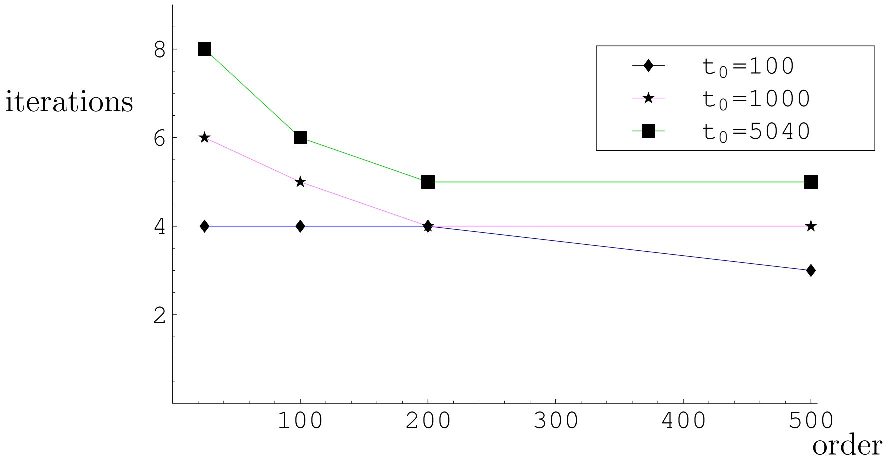

Figure 1. **Orders and iterations.**

References

[1] J. Gerlach, Accelerated Convergence in Newton's Method, SIAM Rev., **36**, (1994), pp. 272–276.

[2] B. Kalantari and I. Kalantari, High Order Iterative Methods for Approximating Square Roots, BIT, **36**, (1996), pp. 395–399.

[3] R. J. Knill, A Modified Babylolonian Algorithm, Amer. Math. Monthly, **99**, (1992), pp. 734–737.

[4] T. J. Ypma, Historical Development of the Newton-Raphson Method, SIAM Rev., **37**, (1995), pp. 531–551.

[5] F. Dubeau, n-th Root Extraction: Double Iteration Process and Newton's Method, J. Comput. Apll. Math., **91**, (1998), pp. 191–198.

[6] B. Neta, Several New Methods for Solving Equations, Inter. J. Computer Math., **23**, (1964), pp. 265–282.

[7] W. Gander, On Halley's Iteration Method, Amer. Math. Monthly, **92**, (1985), pp. 131–134.

[8] J. M. Gutiérrez, M. A. Hernández and M. A. Salanova, A family of Chebyshev-Halley Type Methods in Banach Spaces, Bull. Austr. Math. Soc., **55**, (1997), pp. 113–130.

[9] J. M. Gutiérrez, M. A. Hernández and M. A. Salanova, Calculus of n-th roots and third order Iterative Methods, Nonlinear Analysis, **47**, (2001), pp. 2875–2880.

[10] A. M. Ostrowski, Solution of equations in Euclidean and Banach Spaces, Academic Press., (1973), pp. 32–55.