A technique to composite a modified Newton's method for solving nonlinear equations

Miquel Grau-Sánchez and José Luis Díaz-Barrero

Technical University of Catalonia, Department of Applied Mathematics II and III.
Jordi Girona 1-3, Omega, 08034 Barcelona, Spain.
E-mail address: miquel.grau@upc.edu, jose.luis.diaz@upc.edu

Abstract

A zero-finding technique for solving nonlinear equations more efficiently than they usually are with traditional iterative methods in which the order of convergence is improved is presented. The key idea in deriving this procedure is to compose a given iterative method with a modified Newton's method that introduces just one evaluation of the function. To carry out this procedure some classical methods with different orders of convergence are used to obtain root-finders with higher efficiency index.

Keywords: Newton method; nonlinear equations; iterative methods; order of convergence; efficiency index.

Mathematics Subject Classification: 65H05, 41A25

1   Introduction

No doubt that Newton's method is one of the best root-finding methods for solving nonlinear equations. Recent results improving the classical formula at the expense of an additional evaluation of the function, an additional evaluation of the first derivative or a change in the point of evaluation can be found in the literature on the subject (see [9, 10, 2] and the references therein). In these works the order of convergence and the efficiency index in the neighborhood of a simple root have been improved.

Using the technique that consists in composing a modification of Newton's method with an iterative method we obtain a root-finder for solving nonlinear equations with improved order of convergence and efficiency index. The key idea to improve or even double the order is to use only one additional evaluation of the function instead of the two evaluations needed when applying composition with Newton's method, as it is well-known.

Currently, IEEE 64-bit floating-point arithmetic is sufficient for the most commonly applications in order to obtain the accuracy desired. But, it is increasing the number of applications where it is required to use a higher level of numeric precision [1]. Namely, evaluating orthogonal polynomials, high-precision solution of ODE's, divergent asymptotic series, discrete dynamical systems, experimental mathematics, supernova

1

simulations, climate modeling, and nonlinear oscillator theory among others. So, adaptive multi-precision arithmetics facilities are most appropriate in a modern large-scale scientific computing environment.

2 Main result

Let $f(x) = 0$ be a nonlinear equation where $f : D \subset \mathbf{R} \longrightarrow \mathbf{R}$ is sufficiently smooth in a neighborhood I of a simple root α. Let $\phi(x)$ be an iterative function of order of convergence p in I obtained using $f(x), f'(x), \ldots, f^{(p-1)}(x)$. Hereafter, a technique that consists in an iterative method in two steps, is presented. Namely,

$$
\begin{aligned}
z_n &= \phi(x_n), & (1) \\
x_{n+1} &= z_n - f(z_n)\, g'_q, & (2)
\end{aligned}
$$

where in (2) the factor $1/f'(z_n)$ in the classical Newton's method has been replaced by g'_q, which is an approximation of the derivative of the inverse function of f. That is, if $w_n = f(z_n)$, then $z_n = g(w_n)$. Therefore, $g'(w_n) = 1/f'(z_n)$. As we will see later on, this approximation is given by

$$
g'_q = q\, \frac{g(w_n) - g(y_n)}{w_n - y_n} + \sum_{k=1}^{q-1} \frac{k-q}{k!}\, g^{(k)}(y_n)\, (w_n - z_n)^{k-1}, \tag{3}
$$

where $y_n = f(x_n)$ and p and q are integers such that $p \geq q \geq 2$.

Recalling that (1) is of pth order of convergence and that we have computed the function f and its derivatives up to order equal to $p-1$ at point x_n, to analyze the order of the two-step iterative method given by (1) and (2), we state and prove the following main result:

Theorem 1. *Let e and E be the errors $e_n = x_n - \alpha$ and $E_n = z_n - \alpha = K\,e^p + O(e^{p+1})$ in sequences $\{x_n\}$ and $\{z_n\}$ respectively. Then the order of the iterative method defined by (1)–(2) is equal to $p + q$. More precisely,*

$$
|e_{n+1}| = \begin{cases} |B_{q+1}\, f'(\alpha)^q\, K\, e^{p+q}| + O(e^{p+q+1}), & \text{if } p > q, \quad \text{and} \\[2mm] \left|[(-1)^q\, B_{q+1}\, f'(\alpha)^q + A_2\, K]\, K\, e^{2p}\right| + O(e^{2p+1}), & \text{if } p = q, \end{cases}
$$

where $A_k = \dfrac{f^{(k)}(\alpha)}{k!\, f'(\alpha)}$, and $B_k = \dfrac{g^{(k)}(0)}{k!\, g'(0)}$, $k \geq 2$.

Proof. Putting y instead of y_n and w instead of w_n, and considering Taylor's developments of the functions $g(w)$ and $g'(w)$ in powers of $w - y$, we obtain

$$
\begin{aligned}
g(w) &= \sum_{i=0}^{q} \frac{g^{(i)}(y)}{i!}\, (w - y)^i + \frac{g^{(q+1)}(\xi)}{(q+1)!}\, (w - y)^{q+1}, & (4) \\[2mm]
g'(w) &= \sum_{j=1}^{q} \frac{g^{(j)}(y)}{(j-1)!}\, (w - y)^{j-1} + \frac{g^{(q+1)}(\eta)}{q!}\, (w - y)^q, & (5)
\end{aligned}
$$

2

where ξ and η lie between y and w. From (4) we get $g^{(q)}(y)$ that after putting it into (5) yields

$$
\begin{aligned}
g'(w) &= \sum_{j=1}^{q-1} \frac{g^{(j)}(y)}{(j-1)!} (w-y)^{j-1} + \frac{g^{(q+1)}(\eta)}{q!} (w-y)^q \\
&\quad + \frac{q}{w-y} \left[g(w) - \sum_{i=0}^{q-1} \frac{g^{(i)}(y)}{i!} (w-y)^i - \frac{g^{(q+1)}(\xi)}{(q+1)!} (w-y)^{q+1} \right] \\
&= g'_q + T_q,
\end{aligned}
\tag{6}
$$

where g'_q is given in (3), as claimed before, and

$$
T_q = \left(\frac{g^{(q+1)}(\eta)}{q!} - \frac{q\, g^{(q+1)}(\xi)}{(q+1)!} \right) (w-y)^q.
\tag{7}
$$

From (7) and developing T_q in powers of $w - y$, we have

$$
T_q = \frac{g^{(q+1)}(y)}{(q+1)!} (w-y)^q + O\left((w-y)^{q+1} \right),
\tag{8}
$$

and from

$$
g(y) = \alpha + g'(0) \left(y + \sum_{k=2}^{q+1} B_k\, y^k + O(y^{q+2}) \right),
$$

we have $g^{(q+1)}(y) = g'(0)\, [(q+1)!\, B_{q+1} + O(y)]$.

Similarly, writing $g(w) = g'(0)\, \left[w + B_2\, w^2 + O(w^3) \right]$, we get

$$
g'(w) = g'(0)\, \left[1 + 2\, B_2\, w + O(w^2) \right].
\tag{9}
$$

If we develop (8) in Taylor's series at point 0, then we obtain

$$
\begin{aligned}
T_q &= g'(0)\, B_{q+1}\, (w-y)^q + O(y^{q+1}) \\
&= (-1)^q\, g'(0)\, B_{q+1}\, y^q + O(y^{q+1}).
\end{aligned}
\tag{10}
$$

Substituting (9) and (10) into (6) yields

$$
\begin{aligned}
g'_q &= g'(w) - T_q \\
&= g'(0)\, \left[1 + 2\, B_2\, w + O(w^2) + (-1)^{q+1} B_{q+1}\, y^q + O(y^{q+1}) \right].
\end{aligned}
$$

Now setting

$$
\begin{aligned}
y = f(x) &= f'(\alpha) \left(e + \sum_{k=2}^{q+1} A_k\, e^k + O(e^{q+2}) \right), \\
w = f(z) &= f'(\alpha) \left(E + A_2\, E^2 + O(E^3) \right),
\end{aligned}
$$

we obtain

$$
\begin{aligned}
g'_q &= g'(0)\, \left[1 + (-1)^{q+1} B_{q+1}\, f'(\alpha)^q\, e^q + 2\, f'(\alpha)\, B_2\, E + O(e^{q+1}) \right] \\
&= g'(0)\, \left[1 + (-1)^{q+1} B_{q+1}\, f'(\alpha)^q\, e^q - 2\, A_2\, E + O(e^{q+1}) \right],
\end{aligned}
$$

where in the last expression we have put $f'(\alpha) B_2 = -A_2$.

Subtracting α from both sides of (2) we get e_{n+1}. Assuming that $p > q$, from the previous expression of g_q', we get

$$
\begin{aligned}
e_{n+1} & = E - \left(E + O(E^2)\right)\left(1 + (-1)^{q+1} f'(\alpha)^q B_{q+1} \, e^q + O(e^{q+1})\right) \\
& = (-1)^q f'(\alpha)^q B_{q+1} \, e^q \, E + O(e^{q+1} E).
\end{aligned}
$$

On the other hand, if $p = q$, then

$$
\begin{aligned}
e_{n+1} & = E - \left(E + A_2 E^2 + O(E^3)\right)\left(1 + (-1)^{q+1} f'(\alpha)^q B_{q+1} \, e^q - 2 A_2 E + O(e^{q+1})\right) \\
& = (-1)^q f'(\alpha)^q B_{q+1} \, e^q \, E + A_2 E^2 + O(e^{q+1} E).
\end{aligned}
$$

Replacing E by $E = K \, e^p + O(e^{p+1})$ the statement follows. $\square$

For $q = 3$ equation (3) was used by Kou et al. in [6]. Other contributions related to family (3) can also be found in [7, 5].

Previously, we have set $f'(\alpha) B_2 = -A_2$. This relation can be easily proven. From a theorem of Jabotinsky [4], we have

$$
f'(\alpha)^q B_{q+1} = \frac{1}{(q+1)!} \sum (-1)^r (q+r)! \prod_{\ell=2}^{q+1} \frac{A_\ell^{\beta_\ell}}{\beta_\ell!},
$$

with the sum is taken over all nonnegative integers β_ℓ such that $\sum_{\ell=2}^{q+1} (\ell-1)\beta_\ell = q$, and where $r = \sum_{\ell=2}^{q+1} \beta_\ell$. The proof of this theorem can also be found in [9]. The values of B_{q+1} in terms of A_j's, for $1 \le q \le 4$, are presented in Table 1.

Table 1: Values of $f'(\alpha)^q B_{q+1}$ in terms of A_j

q	$f'(\alpha)^q B_{q+1}$
1	$-A_2$
2	$2 A_2^2 - A_3$
3	$-5 A_2^3 + 5 A_2 A_3 - A_4$
4	$14 A_2^4 - 21 A_2^2 A_3 + 6 A_2 A_4 + 3 A_3^2 - A_5$

We have used the classical definition of efficiency index given in [9]. That is, $EI = m^{1/r}$, where m is the local order of convergence of the method and r is the number of evaluations of the functions per step. Considering the improvement in the order obtained in Theorem 1 the efficiency index is increased considerably. In the case in which the first step in the iterative method is of pth order and there are p evaluations of the functions per iteration, the efficiency index is $EI = p^{1/p}$. By increasing the value of $q \ge 2$ and applying Theorem 1, we obtain $EI = (p+q)^{1/p+1}$. In Table 2 several values for efficiency are given.

Table 2: Efficiencies

	$p = 2$	$p = 3$	$p = 4$
Method $\phi(x)$	$2^{1/2} \approx 1.414$	$3^{1/3} \approx 1.442$	$4^{1/4} \approx 1.414$
$q = 2$	$4^{1/3} \approx 1.587$	$5^{1/4} \approx 1.495$	$6^{1/5} \approx 1.431$
$q = 3$	——	$6^{1/4} \approx 1.565$	$7^{1/5} \approx 1.476$
$q = 4$	——	——	$8^{1/5} \approx 1.516$

3 Some related methods

In this section, some methods that give the best efficiency indexes for $2 \leq p \leq 4$ are constructed. Notice that in Table 2 the best efficiency index correspond to the case when $q = p$. The expression of the asymptotic constant error for known methods is given. Furthermore, we have also computed symbolically and in a different way, the asymptotic constant error for the related methods presented and they agree with the results obtained using Theorem 1.

• For $p = 2$, we choose Newton's method as the $z = \phi(x)$ method. If we write

$$z_n = \psi_2^2(x_n) = x_n - u(x_n) \quad \text{and} \quad g_2' = 2\,[y_n, w_n]_g - g'(x_n),$$

where $u(x_n) = \dfrac{f(x_n)}{f'(x_n)}$ and $[y_n, w_n]_g = \dfrac{g(w_n) - g(y_n)}{w_n - y_n} = \dfrac{z_n - x_n}{f(z_n) - f(x_n)}$, then

$$x_{n+1} = z_n - f(z_n)\left(2\,\frac{z_n - x_n}{f(z_n) - f(x_n)} - \frac{1}{f'(x_n)}\right),$$

or

$$x_{n+1} = \psi_2^4(x_n) = z_n - \frac{f(x_n) + f(z_n)}{f(x_n) - f(z_n)}\,\frac{f(z_n)}{f'(x_n)}.$$

Recall that the expression of error in Newton's method is $E = A_2\,e_n^2 + O_3$. In the method described here the order goes from 2 to 4 and the difference error equation is

$$e_{n+1} = A_2\left(3\,A_2^2 - A_3\right)e_n^4 + O_5,$$

which agrees with the result of Theorem 1 for this particular case.

• For $p = 3$, we use Chebyshev's method [2] as the $z = \phi(x)$ method. If we write

$$z_n = \psi_3^3(x_n) = x_n - \left(1 + \frac{1}{2}\,L(x_n)\right)u(x_n),$$

and

$$\begin{aligned}
g_3' &= 3\,[y_n, w_n]_g - 2\,g'(y_n) - \frac{1}{2}\,g''(y_n)\,(w_n - y_n) \\
&= 3\,\frac{z_n - x_n}{f(z_n) - f(x_n)} - \frac{2}{f'(x_n)} + \frac{f''(x_n)}{2\,f'(x_n)^3}\left(f(z_n) - f(x_n)\right),
\end{aligned}$$

5

where $L(x_n) = \dfrac{f''(x_n)}{f'(x_n)} \, u(x_n)$, then we have $x_{n+1} = \psi_3^6(x_n) = z_n - f(z_n)\, g_3'$. Note that the error in Chebyshev's method is $E = \left(2\, A_2^2 - A_3\right) e_n^3 + O_4$. The error difference equation in this improved method is

$$e_{n+1} = \left(2\, A_2^2 - A_3\right) \left(7\, A_2{}^3 - 6\, A_2 A_3 + A_4\right) e_n^6 + O_7,$$

agreeing again with Theorem 1.

- For $p = 4$, the method $z = \phi(x)$ considered is Schröeder's method [8]. Writing

$$z_n = \psi_4^4(x_n) = x_n - \left(1 + \frac{1}{2}\, L(x_n) \; - \; \frac{1}{6}\, M(x_n)\, u(x_n)^2\right) u(x_n),$$

and

$$
\begin{aligned}
g_4' &= 4\, [y_n, w_n]_g - 3\, g'(y_n) - g''(y_n)\, (w_n - y_n) - \frac{1}{6}\, g'''(y_n)\, (w_n - y_n)^2 \\
&= 4\, \frac{z_n - x_n}{f(z_n) - f(x_n)} - \frac{3}{f'(x_n)} + \frac{f''(x_n)}{f'(x_n)^3}\, (f(z_n) - f(x_n)) \\
&\quad + \frac{1}{6}\left(\frac{f'''(x_n)}{f'(x_n)^4} - \frac{3\, f''(x_n)^2}{f'(x_n)^5}\right) (f(z_n) - f(x_n))^2 \,,
\end{aligned}
$$

where

$$M(x_n) = \frac{f'''(x_n)}{f'(x_n)} - 3\left(\frac{f''(x_n)}{f'(x_n)}\right)^2 ,$$ then we have $x_{n+1} = \psi_4^8(x_n) = z_n - f(z_n)\, g_4'.$

The error in Schröeder's method is $E = \left(5\, A_2^3 - 5\, A_2\, A_3 + A_4\right) e_n^4 + O_5$. The improved method presented here is of 8-th order and the error equation is

$$e_{n+1} = \left(5\, A_2^3 - 5\, A_2\, A_3 + A_4\right) \left(19\, A_2{}^4 - 26\, A_2{}^2 A_3 + 7\, A_2 A_4 + 3\, A_3{}^2 - A_5\right) e_n^8 + O_9,$$

agreeing again with it was obtained in Theorem 1.

Table 3: Test functions, their roots and their initial points

function	α	x_0
$f_1(x) = x^3 - 3x^2 + x - 2$	2.893289	2.5
$f_2(x) = x^3 + \cos x - 2$	1.172578	1.5
$f_3(x) = 2\sin x + 1 - x$	2.380061	2.5
$f_4(x) = (x+1)\, e^{-x} - 1$	0.557146	1.0
$f_5(x) = e^{x^2 + 7x - 30} - 1$	3.0	2.94
$f_6(x) = e^{-x} + \cos(x)$	1.746140	1.5
$f_7(x) = x - 3\ln x$	1.857184	2.0

Table 4: Iteration number and total number of function evaluations (TNFE)

	ψ_2^2	ψ_2^4	ψ_3^3	ψ_3^5	ψ_3^6	ψ_4^4	ψ_4^6	ψ_4^7	ψ_4^8
$f_1(x)$	13	7	9	6	6	7	6	5	5
$f_2(x)$	13	7	8	6	5	7	5	5	5
$f_3(x)$	11	6	8	5	5	6	5	4	4
$f_4(x)$	13	7	8	6	5	7	5	5	5
$f_5(x)$	14	8	9	6	6	7	6	6	5
$f_6(x)$	11	6	8	5	5	6	5	5	4
$f_7(x)$	12	6	8	6	5	6	5	5	4
Iter	87	47	58	40	37	46	37	35	32
TNFE	174	**141**	174	160	**148**	184	185	175	160

4   Numerical experiments and comparison

We have tested the preceding methods with seven functions using the Maple computer algebra system. We have computed the root of each function for initial approximation x_0, and we have defined at each step of the iterative method the length of the floating point arithmetic with multi-precision given by

$$\texttt{Digits} := \rho \times \left[-\log |e_k| + 2 \right],$$

where ρ is the order of the method which extends the length of the mantissa of the arithmetic, and $[x]$ is the largest integer $\leq x$. The iterative method is stopped when $|e_k| = |x_k - \alpha| < 10^{-\eta}$, where $\eta = 3000$ and α is the root. If in the last step of any iterative method it is necessary to increase the number of digits beyond 3000, then it is done. Table 3 shows the expression of the functions tested, the initial approximation x_0 which is the same for all the methods, and the approximation of root α with seven significant digits. The functions tested are the same as those presented in [3]. Table 3 shows the functions; the initial approximation, which is the same for all the methods; and the root with seven significant digits.

In Table 4, for each method and function, the number of iterations needed to compute the root to the level of precision described is shown. The notation works as follows: Newton's iterative method and the modified method are written as ψ_2^2 ($p = 2$) and ψ_2^4 ($p = 2$, $q = 2$). Chebyshev's methods are represented by ψ_3^3 ($p = 3$), ψ_3^5 ($p = 3$, $q = 2$) and ψ_3^6 ($p = 3$, $q = 3$). For Schröeder's method and the modified method we have ψ_4^4 ($p = 4$), ψ_4^6 ($p = 4$, $q = 2$), ψ_4^7 ($p = 4$, $q = 3$) and ψ_4^8 ($p = 4$, $q = 4$). In a compact way the notation used is ψ_p^{p+q}. Notice that the low cost of the iteration functions ψ_2^4 and ψ_3^6, which show higher efficiency index than the other methods considered. In general the results are excellent: the order is maximized and the total number of function evaluations is lowest for the iterative methods ψ_2^4 and ψ_3^6.

Finally, we conclude that the methods ψ_2^4 and ψ_3^6 presented in this paper are competitive

with other efficient equation solvers, such as Newton's, Chebyshev's and Schröeder's methods (ψ_2^2, ψ_3^3 and ψ_4^4 respectively).

5 Concluding remarks

A technique for accelerating the order of convergence of a given iterative process with an additional evaluation of the function is implemented. Furthermore, we have analyzed the new schemes obtained from three particular cases: Newton's, Chebyshev's and Schröeder's methods. Order of convergence and efficiency index have been improved in all these cases. The results have been computationally tested on a set of functions.

Due to the fact that when the order of convergence of any iterative method is high, we need to carry out the computations for testing it with an enlarged mantissa. A multiprecision and adaptive floating-point arithmetics with low computing time must be used in all the calculations, as we have done in this work.

References

[1] Bailey, D.H., Borwein, J.M., High-Precision Computation and Mathematical Physics, XII Advanced Computing and Analysis Techniques in Physics Research, (2008) (to appear). http://crd.lbl.gov/~dhbailey/dhbpapers/dhb-jmb-acat08.pdf

[2] Grau, M., Díaz-Barrero, J.L.: An improvement of the Euler-Chebyshev iterative method, J. Math. Anal. Appl. **315** (2006) 1–7.

[3] Grau-Sánchez, M.: Improvement of the efficiency of some three-step iterative like-Newton methods, Numer. Math. **107**(2007) 131–146.

[4] Jabotinsky, E.: Representation of functions by matrices, Proc. Amer. Math. Soc. **4** (1953) 546–553.

[5] Kou, J., Li, Y., Wang, X.: A family of fifth-order iterations composed of Newton and third-order methods, Appl. Math. Comput. **186** (2007) 1258–1262.

[6] Kou, J., Li, Y.: On Chebyshev-Halley methods with sixth-order convergence for solving non-linear equations, Appl. Math. Comput. **190** (2007) 126–131.

[7] Kung, H.T., Traub, J.F.: Optimal order of one-point and multipoint iteration, J. Assoc. Comput. Mach. **21** (1974) 643–651.

[8] Schröder, E.: Über unendlich viele Algorithmen zur Auflösung der Gleichungen, Math. Ann. **2** (1870) 317–365.

[9] Traub, J.F.: Iterative methods for the solution of equations, 310pp., Prentice-Hall, Englewood Cliffs, New Jersey, (1964).

[10] Weerakoon, S., Fernando, T.G.I.: A variant of Newton's method with accelerated third-order convergence, Appl. Math. Lett. **13** (2000) 87–93.