LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-648574

# An Easy Method To Accelerate An Iterative Algebraic Equation Solver

J. Yao

January 17, 2014

## Disclaimer



This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

# An Easy Method To Accelerate An Iterative Algebraic Equation Solver

**Yao, Jin**

**Lawrence Livermore National Laboratory, California, USA**

**Abstract**

This article proposes to add a simple term to an iterative algebraic equation solver with an order $n$ convergence rate, and to raise the order of convergence to $(2n-1)$. In particular, a simple algebraic equation solver with the $5^{th}$ order convergence but uses only 4 function values in each iteration, is described in details. When this scheme is applied to a Newton-Raphson method of the quadratic convergence for a system of algebraic equations, a cubic convergence can be achieved with an low overhead cost of function evaluation that can be ignored as the size of the system increases.

# THE ORDER OF CONVERGENCE

An iterative equation solver for a set of algebraic equations $\vec{F}(\vec{x}) = 0$ is said to have an order of convergence $n$ when

$$|\vec{x}_{k+1} - \vec{x}_k| = O(|\vec{x}_k - \vec{x}_{k-1}|^n)$$

at the $k^{th}$ iteration. An almost equivalent definition is that

$$|\vec{F}(\vec{x}_{k+1})| = O(|\vec{x}_{k+1} - \vec{x}_k|^n)$$

in the case that the *jacobian* of the system $J \neq 0$ at the solution. The order of convergence of an iterative solver is a measurement of how fast it converges to the true solution.

Conventional methods for solving a regular non-linear equation $f(x) = 0$ of a single variable have the property that the number of function calls required in an iteration cycle equals the order of convergence. For example, a bi-section method $x_{k+1} = 0.5(x_k + x_{k-1})$ which evaluates $f(x_k)$ at each iteration to determine which line-section the solution falls into, is first order; a Newton-Raphson method with

$x_{k+1} = x_k - f(x_k)/f'(x_k)$ which calculates two function values $f(x)$ and $f'(x)$ is second order. The ratio $\mathbf{R}$ between the order of convergence and the number of function values required for an iteration is usually 1.

The following new solution method has $\mathbf{R} = \mathbf{5/4}$, so it is faster than other conventional iterative methods. There are two steps with this method. The first step is the well-known Halley's method[1]. Let $x_k$ be the $n^{th}$ guess for the root. One solves the equation (assuming $f'' \neq 0$)

$$f(x_k) + f'(x_k)\delta + \frac{1}{2}f''(x_k)\delta^2 = 0, \tag{1}$$

and the two roots are explicitly expressed as

$$\delta = -\frac{1}{f''(x_k)}\left(f'(x_k) \pm \sqrt{f'(x_k)^2 - 2f(x_k)f''(x_k)}\right).$$

For consistency with the Newton-Raphson method when the quadratic term vanishes, we pick only one root and it can be written as

$$\delta = \frac{sgn(f'(x_k))}{f''(x_k)}\left(\sqrt{(f'(x_k))^2 - 2f(x_k)f''(x_k)} - |f'(x_k)|\right),$$

in order to avoid the singularity as $f'(x_k) \to 0$. Note that the solution of eq.(1) implies $f(x_k + \delta) = O(\delta^3)$. The above step uses three function calls.

The next step uses one more function call to gain two more orders of convergence. One adds a term $f(x_k + \delta)$ into eq.(1), and solves

$$f(x_k) + f(x_k + \delta) + f'(x_k)\Delta + \frac{1}{2}f''(x_k)\Delta^2 = 0. \tag{2}$$

The solution is similar to what is obtained in the first step

$$\Delta = \frac{sgn(f'(x_k))}{f''(x_k)}\left(\sqrt{(f'(x_k))^2 - 2(f(x_k) + f(x_k + \delta))f''(x_k)} - |f'(x_k)|\right).$$

Finally, let $x_{k+1} = x_k + \Delta$ for completion of the current iteration cycle.

One computes only four function values $f(x_k), f'(x_k), f''(x_k)$, and $f(x_k + \delta)$. However, the above scheme is fifth order convergent as shown below.

From a *Taylor's expansion* one obtains

$$f(x_k + \Delta) = f(x_k) + f'(x_k)\Delta + \frac{1}{2}f''(x_k)\Delta^2 + \frac{1}{6}f^{[3]}(x_k)\Delta^3 + O(\Delta^4).$$

From eq.(2), it is equal to $-f(x_k+\delta)+f^{[3]}(x_k)\Delta^3/6+O(\Delta^4)$. However, from eq.(1) and the Taylor expansion of $f(x_k+\delta)$, the above estimate becomes

$$\frac{1}{6}f^{[3]}(x_k)(\Delta^3-\delta^3)+O(\Delta^4-\delta^4)=(\Delta-\delta)O(\Delta^2,\delta^2).$$

By subtracting eq.(1) from eq.(2) one arrives at

$$(\Delta-\delta)(f'(x_k)+O(\delta))=-f(x_k+\delta)=O(\delta^3). \tag{3}$$

It tells us that $\Delta$ and $\delta$ are of the same order and

$$(\Delta-\delta)=O(\delta^3).$$

One clearly sees that

$$f(x_{k+1})=f(x_k+\Delta)=O(\Delta^5).$$

Therefore the method is $5^{th}$ order convergent, however employs only 4 function values.

The fast convergent scheme described above is more stable than Newton's method in the case of $f'(x)=0$ at the solution. However the order of convergence is reduced from 5 to 4 because when $f'=0$, eq.(3) gives $(\delta-\Delta)=O(\delta^2)$ instead of $O(\delta^3)$. In practice if a $x_k$ is not close to the solution, the term under the squared root $f'(x_k))^2-2f(x_k)f''(x_k)$ may become negative and break the iteration. In this case this term can be set to zero to keep the computation going. When $x_k$ is close to the solution $x^*$, if $f'$ is finite, because $f\to 0$ the term in the square root would be non-negative; if $f'=0$ at the solution, setting this term to zero (if it becomes negative) would give $\delta=-f'(x_k)/f''(x_k)$, and which is similar to a Newton's step by the L'Hospital's rule.

# A NUMERICAL EXAMPLE

We take an simple equation $cos(x)=x$ to demonstrate the efficiency of the proposed scheme. Its numerical solution is $x^*=0.739085133215160641638918505$..... Because a trigonometrical function costs a lot to evaluate, it is certainly desired to use fewer function values for a specified order of accuracy. All numerical evaluations below are done with MATHEMATICA with an accuracy of 100 digits.

Let $f = x - cos(x)$ one has $f' = 1 + sin(x)$ and $f'' = cos(x)$. We take an innocent initial guess that $x_0 = 0$.

Now perform the first iteration, by first solving $f(x_0) + f'(x_0)\delta + \frac{1}{2}f''(x_0)\delta^2 = 0$, or

$$-1 + \delta + \delta^2/2 = 0$$

one finds that (take the first 25 digits)

$$\delta = 0.73205080756887729352744634...$$

Then perform the second step of this iteration by solving

$$f(x_0) + f(x_0 + \delta) + f'(x_0)\Delta + \frac{1}{2}f''(x_0)\Delta^2 = 0$$

to 25 digits to obtain $\Delta = 0.73882397464992265839862270...$ Therefore $x_1 = x_0 + \Delta = \Delta$ here (with $x_0 = 0$). Be aware that this is off the exact solution by about only $2 \times 10^{-4}$.

Now perform the second iteration by first solving $f(x_1) + f'(x_1)\delta + \frac{1}{2}f''(x_1)\delta^2 = 0$ to obtain that $\delta = 0.00026115856404338119688100463655...$

Now practice the second half of this iteration by solving

$$f(x_1) + f(x_1 + \delta) + f'(x_1)\Delta + \frac{1}{2}f''(x_1)\Delta^2 = 0,$$

one obtains $\Delta = 0.00026115856523798323402958...$ and which makes

$$x_2 = x_1 + \Delta = 0.73908513321516064163891850 5,$$

and the value of the equation is now $-2.74365 \times 10^{-20}$. Because the value of $f'$ is order $O(1)$, the accuracy of the root must also be $O(10^{-20})$.

We have observed that, with one iteration, 4 digit accuracy is obtained and with two iterations, one obtains 20 digits of accuracy and the order of convergence is $20/4 = 5$, the same as mathematically proven. However, only $four$ function values are used in each iteration that $f(x), f'(x), f''(x)$, and $f(x + \delta)$.

# GENERALIZATION OF THE SCHEME

In the case that an iterative equation solver of $n^{th}$ order convergence based on a finite *Taylor's expansion* by solving

$$\sum_{i=0}^{n-1} f^{[i]}(x_k)(\delta_k)^i = 0, \tag{4}$$

with $\delta_k = x_{k+1} - x_k$ can be improved with solving again

$$f(x_k + \delta_k) + \sum_{i=0}^{n-1} f^{[i]}(x_k)(\Delta_k)^i = 0, \tag{5}$$

and taking $x_{k+1} = x_k + \Delta_k$.

By taking the difference between eq.(4) and eq.(5) (and obtaining $(\Delta_k - \delta_k) = O(\delta_k^n)$), one arrives at $f(x_{k+1}) = O(\Delta_k^{2n-1})$ following an approach similar to the previous proof for convergence of the $fifth$ order convergent scheme. Thus with $(n+1)$ function values evaluated, a Taylor expansion based equation solver can achieve an order $2n-1$ convergence rate with adding a single term, provided the root of the polynomial expansion in eq.(4, 5) is relatively easy to find. This is true when the evaluation of function values is much more expansive than evaluation of power terms. The $5^{th}$ order convergent scheme eq.(1, 2) obtained in previous discussions can be employed to find the root of a polynomial efficiently.

For a system of equations, the most commonly used root finding method is the Newton's method, i.e. to obtain the solution of $\vec{F}(\vec{x}) = 0$, with $\vec{F} = (F_1, F_2, .....F_M)$ and $\vec{x} = (x_1, x_2, .....x_M)$ ($M$ a positive integer) by solving

$$F(\vec{x}_k) + \vec{\nabla} F(\vec{x}_k) \cdot \vec{\delta}_k = 0 \tag{6}$$

with $\vec{x}_{k+1} = \vec{x}_k + \vec{\delta}_k$, and this scheme is *second* order convergent. By adding the term $\vec{F}(\vec{x}_k + \vec{\delta}_k)$ and solve again that

$$\vec{F}(\vec{x}_k + \vec{\delta}_k) + \vec{F}(x_k) + \vec{\nabla} F(x_k) \cdot \vec{\Delta}_k = 0 \tag{7}$$

with $\vec{x}_{k+1} = \vec{x}_k + \vec{\Delta}_k$. The improved scheme is a *third* order convergent one (proof similar as before).

This convergence rate is obtained with no evaluation of second order derivatives, but $M$ function evaluations in addition to the cost of the Newton's method. Consider the Newton's method requires $M$ function values and $M^2$ derivatives, as $M$ becomes big, the improved scheme is considerably more efficient. The additional cost is relatively insignificant, but provides an extra order of convergence. Not only this, the inverse matrix of the gradient computed in the Newton's step can also be reused.

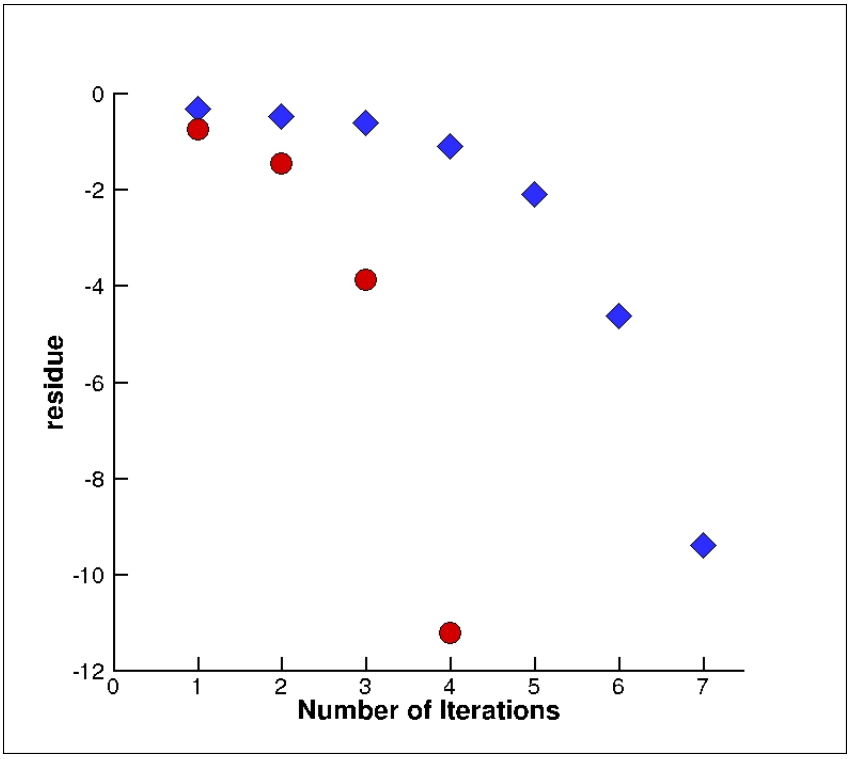


Figure 1: *The convergence map (residue logarithm vs. number of iterations) of the Newton's method (diamonds, represent a parabola) and the proposed method (circles, represent a cubic curve) for the equation system eq.(9).*

# MORE NUMERICAL EXAMPLES

This time we demonstrate the convergence rate of the new scheme for a high order Taylor expansion scheme

$$f(x_k) + f'(x_k)\delta + \frac{f''(x_k)}{2}\delta^2 + \frac{f'''(x_k)}{6}\delta^3 + \frac{f''''(x_k)}{24}\delta^4 = 0 \tag{8}$$

(which is $5^{th}$ order convergent) with the previous example $f(x) = cos(x) - x$. We take a initial guess that has 4 digits of accuracy that $x_0 = 0.7388$, and the above *quartic* polynomial equation has the solution

$$\delta = 0.000285133215160641661631833398416 1...$$

and $x_1 = x_0 + \delta$ gives $f(x_1) = -1.05768 \times 10^{-20}$. This accuracy is expected because the order of convergence here is $5 = 20/4$, the ratio between digits of accuracy, with this $5^{th}$ order Newton's scheme. Now we add the term $f(x_1)$ to eq.(8) and solve

$$f(x_1) + f(x_0) + f'(x_0)\Delta + \frac{f''(x_0)}{2}\Delta^2 + \frac{f'''(x_0)}{6}\Delta^3 + \frac{f''''(x_0)}{24}\Delta^4 = 0$$

and expect the result, with the $(2n-1)^{th}$ order of convergence in theory, to have $9 \times 4 = 36$ digits of accuracy ($n = 5$ is the convergence order of eq.(8), and 4 is the number of effective digits of $x_0$). Indeed, we find

$$\Delta = 0.000285133215160641655312087673873403313044341483 2...$$

and the modified solution $x_1^* = x_1 + \Delta$ gives $f(x_1^*) = 1.17214333 \times 10^{-36}$. Therefore the order of convergence with the proposed scheme is numerically $36/4 = 9$ for the proposed scheme, just as proven in the previous section (taking $n = 5$).

Finally we apply the new scheme to a system of algebraic equations

$$f(x,y) = cos(\pi x)e^y + sin(\frac{x^2+y^2}{2}), \qquad g(x,y) = sin(\pi y)e^x + cos(\frac{x^2+y^2}{2}). \quad (9)$$

This system has a solution at $(x, y) = (1, 0)$. With the initial guess $(x, y) = (0.7, 0.3)$, the solution can be obtained by taking either a Newton iteration scheme

$$f(x_k) + f_x\delta x + f_y\delta y = 0, \qquad g(x_k) + g_x\delta x + g_y\delta y = 0,$$

to solve for $x_{k+1} = x_k + \delta x$, $y_{k+1} = y_k + \delta y$; or with the proposed scheme that

$$f(x_{k+1}) + f(x_k) + f_x\Delta x + f_y\Delta y = 0, \qquad g(x_{k+1}) + g(x_k) + g_x\Delta x + g_y\Delta y = 0,$$

to solve for

$$x_{k+1}^* = x_k + \Delta x, \qquad y_{k+1}^* = y_k + \Delta y$$

as the modified solution at the $(k+1)^{th}$ iteration.

The residues from both the methods are plotted in figure 1. The proposed scheme, being *third* order, clearly converges much faster than the Newton's method of *second* order convergence.

# CONCLUSION

For a *Taylor's expansion* based iterative equation solver that is $n^{th}$ order convergent, the order of convergence can be raised to $2n-1$ by adding a single term to the expansion. When the proposed scheme is applied to the Halley's scheme, only 4 evaluations of function values are required for a $5^{th}$ order convergence. When applied to the Newton's method for a system of $M$ equations, a *third* order of convergence can be obtained with only $M$ evaluations of function values, in addition to the $M + M^2$ evaluations of function values for the *second* order convergence of the Newton's method. This indicates the proposed scheme provides a 50% speed up in average over the Newton's method for $M \gg 1$.

# ACKNOWLEDGMENT

The author would like to express appreciation to Dr. Robert Rieben, Dr. Brian Pudliner, and Dr. Frank Graziani of LLNL for reviewing this paper. The author also would like to thank Dr. David Miller of LLNL for detailed evaluation and insightful comments about this method.

# REFERENCE

[1]. Weisstein, Eric W. "Halley's Irrational Method." From *MathWorld* - A Wolfram Web Resource.

[2]. Press W. H., Teukolsky S. A., Vetterling W. T., and Flannery B. P., *Numerical Recipes in C,* second edition, Cambridge University Press, (1997).

*This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.*