GUARANTEED REAL ROOTS CONVERGENCE OF FITH ORDER POLYNOMIAL AND HIGHER

by
Farid A. Chouery[1], P.E.

Introduction:

In finding the roots of the fifth order polynomial, we find the available iterative algorithm do not give a guarantee of finding the roots. For example, Newton's or Halley's iteration formula runs into difficulties when $f(x)$ and $f'(x)$ or $f'(x)$ and $f''(x)$ are simultaneously near zero. This problem happens not only when the root makes $f(x)$, $f'(x)$ and $f''(x)$ zero but what if during the iteration process one of the values x_n also gives $f'(x_n) = 0$ in Newton's iteration or both $f'(x_n) = f''(x_n) = 0$ in Halley's. This condition throws the improved x_{n+1} into infinity. Thus fishing around x_n for a new x_{n+1} is required with no guaranties, causing additional iterations that still may stumble to a new values x_n that gives $f'(x_n) = 0$ in Newton's iteration or both $f'(x_n) = f''(x_n) = 0$ in Halley's. Similarly, this problem also happens in Lageree's Method. We also note in the convergence algorithm of Lin-Baistow if the coefficient in the denominator is zero it throws the improved x_{n+1} into infinity. Again they all do not give a guaranteed algorithm of convergence. Another problem is the initial value x_0 it is at the programmer own risk it may cause divergence. This value has to be suitable to cause convergence. So how do we pick the initial value?

We seek a guaranteed algorithm for the fifth order polynomial that has a prescribed initial value x_0. This new algorithm is at least of fourth order convergence and most of the time is of fifth order convergence. So it is expected to converge most of the time in five iterations depending on the polynomial coefficients.

Setting up the solution:

We know with higher order Taylor expansion it will represents a better approximation of the function $f(x)$ than Newton or Halley's method. Newton follows the slope of a line to find x_{n+1} value. We will follow a 4th order polynomial curve instead and is obtained from Taylor expansion, thus:

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(x_n)}{2}(x - x_n)^2 + \frac{f'''(x_n)}{6}(x - x_n)^3 + \frac{f^{(4)}(x_n)}{24}(x - x_n)^4 \quad \ldots.(1)$$

Which will give 5th order convergence. We note that $\frac{f^{(5)}(x_n)}{5!} = 1.0$ and $\frac{f^{(n)}(x_n)}{n!} = 0$ for $n \geq 6$. So

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(x_n)}{2}(x - x_n)^2 + \frac{f'''(x_n)}{6}(x - x_n)^3 + \frac{f^{(4)}(x_n)}{24}(x - x_n)^4 + (x - x_n)^5 \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Thus Eq. 1 lacks one more term to bring us back to the original 5th order polynomial of Eq. 2 where:

---

[1] Electrical/Structural/Foundation Professional Engineer, BSEE, MSEE, MSE (Structural) from the University of Washington, WA. President/CEO of FAC SYSTEMS INC. Seattle, WA (www.facsystems.com)

$$f(x) = x^5 + ax^3 + bx^2 + cx + d = 0 \quad \ldots\ldots\ldots\ldots \text{(3)}$$

Assuming we are trying to find the roots of the equation:

$$y^5 + py^4 + qy^3 + ry^2 + sy + t = 0 \quad \ldots\ldots\ldots\ldots \text{(4)}$$

and we have substituted $y = x - p/5$ to give:

$$a = -10\left(\frac{p}{5}\right)^2 + q$$

$$b = 20\left(\frac{p}{5}\right)^3 - 3q\left(\frac{p}{5}\right) + r$$

$$c = -15\left(\frac{p}{5}\right)^4 + 3q\left(\frac{p}{5}\right)^2 - 2r\left(\frac{p}{5}\right) + s$$

$$d = 4\left(\frac{p}{5}\right)^5 - q\left(\frac{p}{5}\right)^3 + r\left(\frac{p}{5}\right)^2 - s\left(\frac{p}{5}\right) + t$$

the coefficients in Eq. 3. We do not need to prove that Eq. 1 is an approximation to Eq. 2 or Eq. 3. It is generally understood that if we follow the polynomial curve of Eq. 1 it would give a closer value to the actual root; more closer than Newton iteration because for any x_n it is a better approximation of the function than Newton's approximation (see Figure 1). These practical assumptions can be fairly accepted by most mathematicians. All would agree that Taylor series with missing term or two gives a closer approximation of any function $f(x)$ provided that Taylor Series Remainder term is small enough and so the root x_{n+1} must be close to the root and no proof is needed (if this is still a question then just compare values with various approximated function). When solving the Quartic equation of Eq. 1, in the iteration we pick the closest root of Eq. 1, $x^* = (x - x_n)$, that is closest to zero so $f(x_n) = 0$ in Eq. 1 and the final x_n becomes a root of $f(x)$.

We start our algorithm by first checking if $f^{(4)}(x) \neq 0$ at any root of $f(x)$ otherwise we will run into problems in finding the roots of Eq.1 when dividing by $f^{(4)}(x_n)$ provided x_n is in the neighborhood of x^*. $f^{(4)}(x) = 5 \cdot 4 \cdot 3 \cdot 2x = 0$ at $x = 0$. This mean we can never have an $x_n = 0$. The only way this can happens is when $d = 0$. Thus, the first check is finding out if $d = 0$, if so the root is $x = 0$ and we find the remainder of the roots by dividing Eq. 3 by x and proceed to find the root of the Quartic equation. If d is very close to zero but not zero and the selected initial value, or some other reason it causes $f^{(4)}(x_i) = 0$ then treat Eq. 1 to a lower order polynomial for only that value to find $x - x_i$ using the equation bellow

$$f(x) = f(x_i) + f'(x_i)(x - x_i) + \frac{f''(x_i)}{2}(x - x_i)^2 + \frac{f'''(x_i)}{6}(x - x_i)^3 \quad \ldots\ldots\ldots\ldots \text{4A}$$

If $f'''(x_i)$ is also zero for the same value then treat Eq. 4A to a lower order polynomial for only that value to find $x - x_i$ using the equation bellow

$$f(x) = f(x_i) + f'(x_i)(x - x_i) + \frac{f''(x_i)}{2}(x - x_i)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 4B$$

If $f''(x_i)$ is also zero for the same value then treat Eq. 4B to a lower order polynomial for only that value to find $x - x_i$ using the equation bellow

$$f(x) = f(x_i) + f'(x_i)(x - x_i) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 4C$$

$f'(x_i)$ is also zero then use Eq. 2 for a better approximation we replace x_n by x_i and solve for x.

$$x^* = x_{i+1} = x_i - [f(x_i)]^{\frac{1}{5}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 4D$$

Obviously if is $f(x_i)$ is also zero then $x = x_i$ is the root. This guarantees[2] conversion for the selected initial value specified in the next section.

Another assumption we have is that the Taylor 4th order polynomial, Eq. 1, has a real root and can never give imaginary roots. This assumption can be contradicted when $f'(x)$ has no roots or all the roots of $f'(x)$ gives $f(x) > 0$. To show the assumption that when $f'(x)$ has no roots the Taylor 4th order polynomial Eq.1 has imaginary roots. Divide Eq. 1 by $(x - x_n)^4$ and let $u = 1/(x - x_n)$ yields:

$$h(u) = f(x_n)u^4 + f'(x_n)u^3 + \frac{f''(x_n)}{2}u^2 + \frac{f'''(x_n)}{6}u + \frac{f^{(4)}(x_n)}{24} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

Similarly divide Eq. 2 $(x - x_n)^5$ yields:

$$g(u) = f(x_n)u^5 + f'(x_n)u^4 + \frac{f''(x_n)}{2}u^3 + \frac{f'''(x_n)}{6}u^2 + \frac{f^{(4)}(x_n)}{24}u + 1 = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

If we multiply Eq. 5 by u we find the Eq. 5 and Eq.6 are identical functions except Eq. 6 is translated in the y-axis by 1. If $f'(x) \neq 0$ or $f'(x)$ has no real root then Eq. 2 has one real root, so is Eq. 6 it would have one real root. Thus, Eq. 5 when multiplied by u to become $uh(u)$ would have one real root. This was done to match Eq. 6 and it would only be translated by 1. Since, $uh(u)$ =0 has one real root namely u=0 then $uh(u)$ would have no other real root. Thus, $h(u)$ has no real roots or Eq.1 has no real roots and they are all imaginary. All this because $f'(x)$ has no real root. This

[2] The problem $f'(x) = 0$ in Newton's iteration or $f'(x) = f''(x) = 0$ Halley's iteration for any polynomial can be avoided by adding a higher term for a better approximation of Eq.1 and find $x_{i+1} = x_i + \left[-\frac{m!\, f(x_i)}{f^{(m)}(x_i)}\right]^{\frac{1}{m}}$ where $f^{(m)}(x_i)$ is selected to be not zero and the term in the bracket is greater then zero if m is even.

problem can also happen if $f(x)$ has only one root. A sufficient test is to determine if $k(x) = f(x) - (x - x_0)^5$ has a real root, where x_0 is an appropriate initial value used in Eq. 1

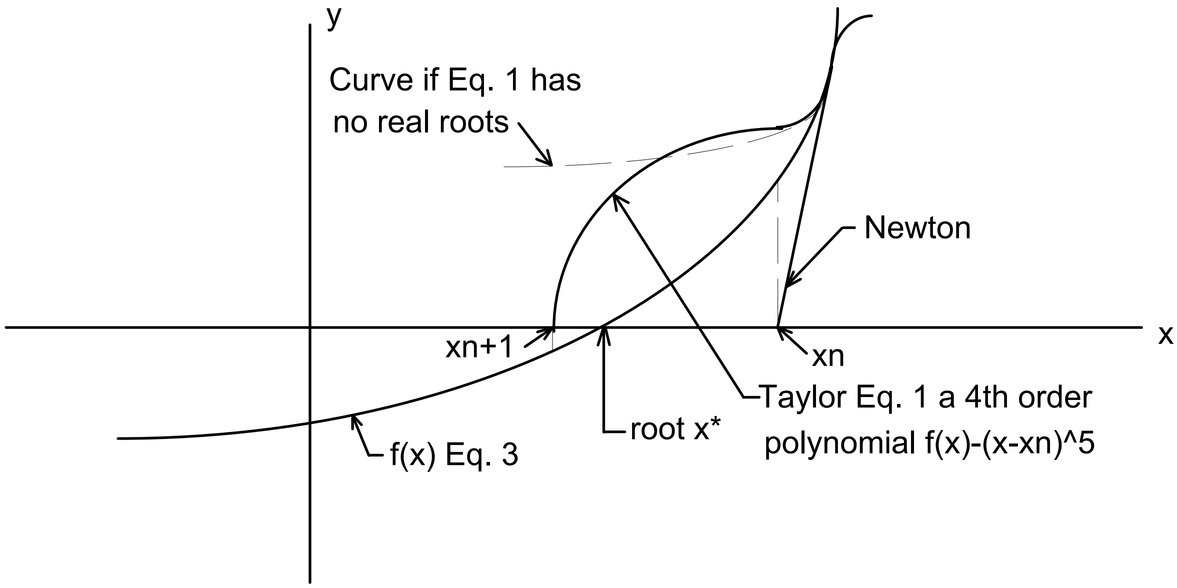


FIGURE 1

and $k(x)$ matches Eq. 1 (this can be derived by using Eq. 2 for $f(x)$ and subtracting $(x - x_n)^5$ to become Eq. 1). If so then all of $f(x) - (x - x_n)^5$ has at least one real roots This can be easily seen graphically as in figure 2 when translating the same function x^5 by x_n then subtracting from $f(x)$. We saying if $(x - x_0)^5$ intersect once then if it is shifted it will intersect at least once more. (Note: $f'(x_n) = k'(x_n)$). Thus we can construct the curve in Figure 1. Another preferred alternative than doing the test for the 5th order polynomials is to go ahead and find the roots of $f'(x)$ and $f''(x)$, (Note: we need to do that anyway in finding the initial condition described in forgoing section) and see if $f'(x)$ roots are imaginary. If the test[3] or the roots $f'(x)$ shows at least

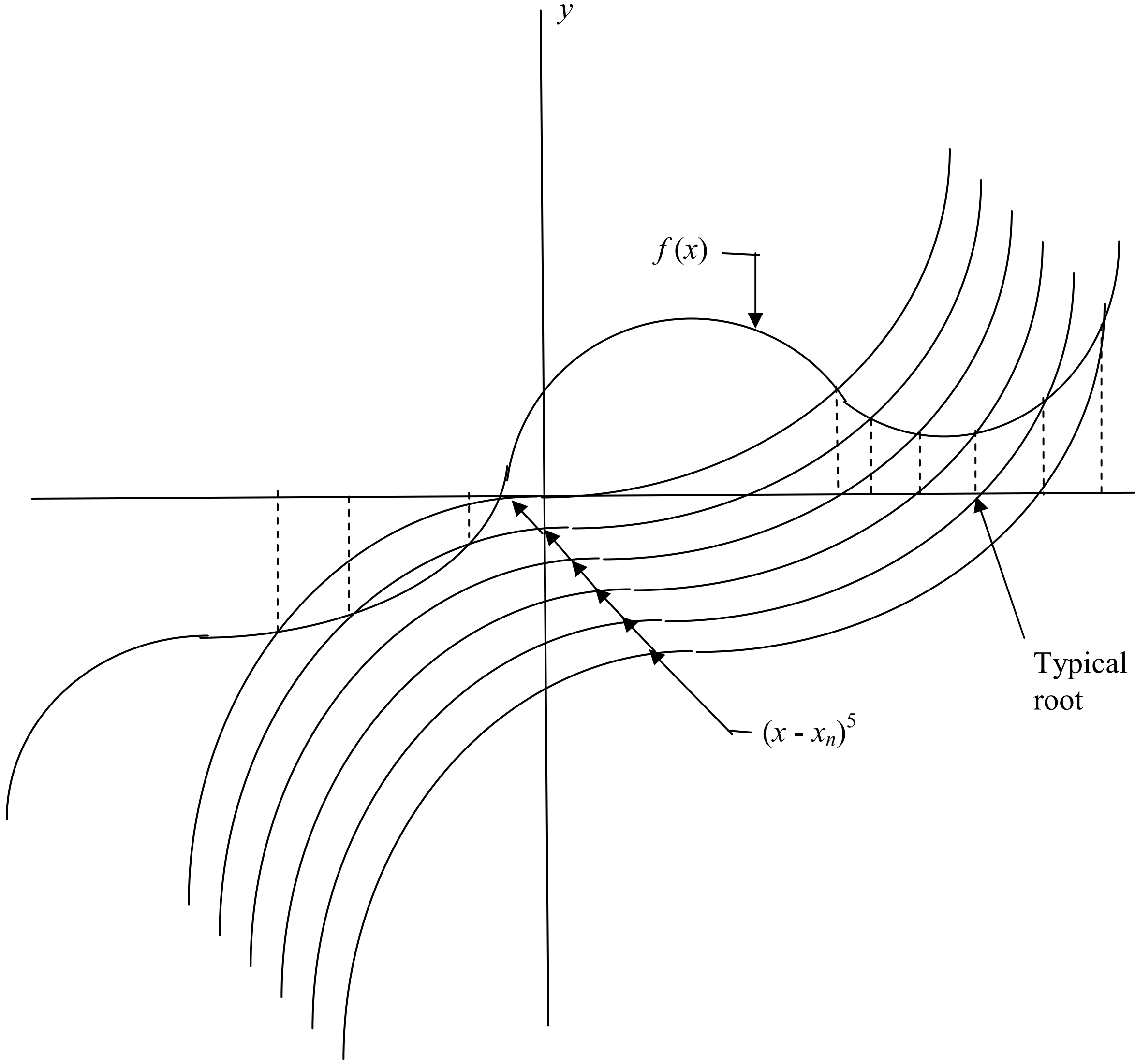
[3] Note: this test is useful because for example if $f(x)$ is an odd function and has a root and is a higher order polynomial or even if it is not a polynomial that has been approximated by a Taylor polynomial (For example if $f(x) = f(0) + f'(0)\, x / 1! + f''(0)\, x^2 / 2! + f'''(0)\, x^3 / 3! + \ldots\ldots + f^{(n)}(0)\, x^n / n!$, where n is odd and $f^{(n)}(0)$ is approximated numerically with high accuracy with an acceptable remainder. In this case $f'(x_n)$, $f''(x_n)$, $f'''(x_n)$, $f^{(4)}(x_n)$, … does not need to be approximated with high accuracy similar to the Secant Method with more initial values). So the test becomes:

$$k(x) = f(x) - \sum_{i=5}^{n} \frac{f^{(i)}(x_0)}{i!}(x - x_0)^i$$

$$= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \frac{f'''(x_0)}{6}(x - x_0)^3 + \frac{f^{(4)}(x_0)}{24}(x - x_0)^4$$

If $k(x)$ has a real root using the appropriate initial value x_0 then all of $k(x)$ has a real root when replacing x_0 by x_n and tells us the Quartic Eq. 1 has no imaginary roots for all x_n This can be concluded since we are subtracting from $f(x)$ a translated function by x_n. In the case of an even function $f(x)$ it is safer to use Eq.7.
If the initial condition in all cases are guest at instead without using the above guaranteed conversions for finding the roots of $f(x)$ then guessing the initial value is permitted and legally and can be defended in a court of law, in the case of an uneducated person in the jury questions the engineer's result for guessing a number. The reason is it can be permitted because it can be verified without guessing with our guaranteed conversions of roots and compared with other roots - please see commentary in the end of the paper.

one real root then Eq.1 will always have at least one real root. If the selected test $f(x)-(x-x_0)^5$ has no real root without checking $f'(x)$ then it is



best to avoid this situation and use a cubic Taylor approximation instead of using Eq.1 or we use the following equation:

$$f(x)=f(x_n)+f'(x_n)(x-x_n)+\frac{f''(x_n)}{2}(x-x_n)^2+\frac{f'''(x_n)}{6}(x-x_n)^3 \quad \text{.......................... (7)}$$

Where, the order of convergence drops to 4th order and in the iteration we pick the closest root

$(x-x_n)$ to zero in the cubic equation.. In this case we start our algorithm by first checking if $f'''(x)\neq 0$ at any root of $f(x)$ where x_n is in the neighborhood of x otherwise we will run into problems in finding the roots of Eq.7 when dividing by $f'''(x_n)$.

$f'''(x)=5\cdot 4\cdot 3\cdot x^2+a=0$ at $x=\pm\sqrt{-a}$. If $a>0$ there is no need to worry $f'''(x_n)$ could never be zero, other wise check and see if $f(\pm\sqrt{-a}) = 0$, if true the root is found and we proceed using synthetic division to find the remainder roots of the Quartic equation. If the root is very close to $\pm\sqrt{-a}$ but not that value and the selected initial value or for some other reason causes $f'''(x_i) = 0$ then treat Eq. 1 to a lower order polynomial for only that value to find $x - x_i$ by using Eq. 4B. If $f''(x_i)$ is also zero use Eq. 4C. If $f'(x_i)$ is also zero then use Eq. 2 for a better approximation without the forth order term by replacing x_n by x_i and solve for x.

$$x = x_{i+1} = x_i - [f(x_i)]^{\frac{1}{5}} \quad \text{.......} \quad (8)$$

Obviously if $f(x_i)$ is also zero then $x = x_i$ is the root. This guarantees conversion for the selected initial value specified in the next section.

Finding the initial value x_0:

After making the above check, we proceed by finding the roots of $f'(x) = 0$ and find the zero slopes. Second we find the roots of $f''(x) = 0$, for the inflection points. These roots can be easily found using the Quartic and Cubic polynomial solutions. Let x'_{min} and x'_{max} be the most negative and the most positive root of $f'(x)$ if any and let x''_{min} and x''_{max} be the most negative and the most positive root of $f''(x)$ if any. By inspection if $f(x'_{min}) > 0$ then our root $x^* < x'_{min}$ and if $f(x'_{max}) < 0$ then our root $x^* > x'_{max}$. Since we know the upper and lower bound ($-M$, M):

$$M = \min[1 + \max(a,b,c,d), \max(1, |a| + |b| + |c| + |d|)] \quad \text{and} \quad -M < x^* < M$$

Then the initial value can be taken as

$$x_0 = \frac{[\min(x'_{min}, x''_{min}) - M]}{2} \quad \text{if} \quad f[\min(x'_{min}, x''_{min})] > 0 \;, -M < x^* < \min(x'_{min}, x''_{min})$$

$$x_0 = \frac{[x''_{min} + x'_{min}]}{2} \quad \text{if} \quad f[\min(x'_{min}, x''_{min})] < 0 \text{ and } f(x'_{min}) > 0 \;, \; x''_{min} < x^* < x'_{min}$$

If x'_{min} do not exist set $x'_{min} = x''_{min}$ in finding x_0 .... (9)

or

$$x_0 = \frac{[\max(x'_{max}, x''_{max}) + M]}{2} \quad \text{if} \quad f[\max(x'_{max}, x''_{max})] < 0 \;, M > x^* > \max(x'_{max}, x''_{max})$$

$$x_0 = \frac{[x''_{max} + x'_{max}]}{2} \quad \text{if} \quad f[\max(x'_{max}, x''_{max}] > 0 \text{ and } f(x'_{max}) < 0 \;, \; x''_{max} > x^* > x'_{max}$$

If x'_{max} do not exist set $x'_{max} = x''_{max}$ in finding x_0 .... (10)

If none of the above is satisfied then use for next consecutive hump at x_i' and x_{i+1}' where $f(x_i') > 0$ and $f(x_{i+1}') < 0$ with $x_i' < x_j'' < x_{i+1}'$

$$x_0 = \frac{x_j'' + x_i'}{2} \quad \text{if } f(x_j'') < 0 \quad \text{and} \quad x_0 = \frac{x_j'' + x_{i+1}'}{2} \quad \text{if } f(x_j'') > 0 \quad \ldots\ldots\ldots\ldots (11)$$

This selected value of x_0 will guarantee convergence. If there are other intermediate x' and x" roots that maybe suitable and do not involve $\pm M$ the algorithm will pick the faster values if M is very large. The process can be easily done by sorting the combine data of x_i' and x_i" and investigate $f(x_i')$ and $f(x_i'')$. The proof of conversions after selected initial conditions can be seen graphically using Newton's iteration slope line for a lower curve approximation of Eq. 1 or Eq. 7.

Real Roots of Higher Order Polynomials:

If we go to the 6th order polynomial, we find using the above 5th order polynomial solution we can find the slopes and the inflection points. If there is one root for $f'(x') = 0$ and $f(x') > 0$ then all the roots of $f(x)$ are all imaginary, other wise we can proceed with Eq.7 to find the real roots. However, before iterating we need to make sure $f'''(x) \neq 0$ when using Eq.7 this procedure for finding the new x_{i+1} is similar to the 5th order procedure. Thus, the real roots of the 6th order polynomial are guaranteed to be extracted. If we go to the 7th order polynomial again using the solution of the 6th order we can extract the slopes and the inflection points. If there is no zero slopes found or all the roots of $f'(x)$ are imaginary, then there is only one root. The solution is similar to the 5th order polynomial and the real roots are guaranteed to be extracted. Finally, if we go to the 8th order we find similarities to the 6th order in knowing if they are all the roots are imaginary. With this we leave it to the reader to construct a guaranteed algorithm for higher order polynomial. It is realized that in order to write the algorithm of polynomial all the previous algorithm of lower order has to be programmed. It is a very nested challenging program for higher order polynomials and probably an expert programmer may be attracted to the problem.

Commentary and Hints to Guarantee Imaginary Roots and multi-dimension functions:

For the 5th order polynomial that has imaginary roots substitute for $x = u + iv$ in Eq. 3 where u and v are real numbers and v cannot be zero. When collecting the real and imaginary in Eq. 3 it forces finding the root (u^*, v^*) of the following two polynomial:

$$\begin{aligned} f(u,v) &= u^5 + au^3 + bu^2 + cu + d - v^2\left(10u^3 + 3au + b\right) + 5v^4 u = 0 \\ g(u,v) &= 5u^4 + 3au^2 + 2bu + c - v^2(u^2 + a) + v^4 = 0 \end{aligned} \quad \ldots\ldots\ldots\ldots (12)$$

Since for any v we can guaranty a root u in both equations and since there is a u that makes both equations of Eq. 12 intersects (this can be seen because one equation is a 5th order in u and the other is a 4th order in u) then there is a solution u that satisfies both equations.

Now the initial value v_0 can be selected by taking $\frac{\partial f}{\partial v}=0, \frac{\partial^2 f}{\partial v^2}=0$ giving four sets of equation in u related to v^2 as follows:

$$\begin{aligned}
v^2 &= \frac{1}{10}\left(10u^3+3au+b\right) && \text{for } \frac{\partial f}{\partial v} \\
v^2 &= \frac{1}{30}\left(10u^3+3au+b\right) && \text{for } \frac{\partial^2 f}{\partial v^2} \\
v^2 &= \frac{1}{2}\left(u^2+a\right) && \text{for } \frac{\partial g}{\partial v} \\
v^2 &= \frac{1}{6}\left(u^2+a\right) && \text{for } \frac{\partial^2 g}{\partial v^2}
\end{aligned} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (13)$$

We see that $\frac{\partial^2 f}{\partial v^2} < \frac{\partial f}{\partial v}$ and $\frac{\partial^2 g}{\partial v^2} < \frac{\partial g}{\partial v}$ for every u. Also $\pm M$ can be calculated from Eq. 12 yields:

$$M = \min\left\{1+\max\left[-\frac{10u^3+3au+b}{5u}, \frac{u^5+au^3+bu^2+cu+d}{5u}\right], \max\left[1, \left|\frac{10u^3+3au+b}{5u}\right|, \left|\frac{u^5+au^3+bu^2+cu+d}{5u}\right|\right]\right\}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (14)$$

If we use Eq. 13 and 14 as we done with Eq. 9 and 10 to select v_0^2 as a function of u and substitute in Eq. 12 and find a u_0 using a guaranteed conversion algorithm for each equation (such as using the Author's or Improved Newton's for $f\,'(x) = 0$ problem or Halley's for $f\,'(x) = f\,''(x) = 0$ problem). Now pick any u_0 then the initial condition (u_0, v_0) for guaranteed root are found. Then we can use a two dimensional Newton's algorithm with the Taylor series expansion to find the roots. We know now how to avoid the pit falls that was in one dimension Newton's of $f\,'(x) = 0$ or Halley's for $f\,'(x) = f\,''(x) = 0$ by using higher order terms in the series at that value. This procedure can be used for a two dimensional Newton's algorithm.

We can conclude from this example that for multi-dimension function we can derive a prescribed initial condition – it will take some work but we have a way to do it - that guarantees conversions using a Taylor series expansion algorithm for multi-dimension that avoids pit falls like as in $f\,'(x) = f\,''(x) = 0$ in one dimension then the solution avoided guessing for the initial condition and becomes a closed solution.

With that said, guessing the initial value should be permitted and legally and can be defended in a court of law, in the event of an uneducated person in the jury questions the engineer's, scientists etc., for his or her result for guessing a number. The reason it can be permitted because it can be verified without guessing with our guaranteed conversions of roots and compared with other roots.

The reason this problem was addressed was the author has solved several fundamental unsolved problems in Structural and Mechanics for Large deflection of Beams, Beam Buckling and Plates that requires finding zeros for a multi-dimension functions. Because this problem is involves in many areas in engineering the guessing question had to be put rest for the benefits of all.