Department of Mechanical Engineering, University of Bath

Mathematics 2 ME10305

Solution Sheet — Root finding and iteration schemes.

Q1. The aim for this question is to repeat some of the techniques used in the lectures to find roots of equations.

(a) Use a suitable sketch to find the number of roots there are likely to be of the equation, $f(x) = x^3 - 2x + 1 = 0$.

(b) Use two *ad hoc* iteration schemes to determine the roots of the equation.

(c) Use the Newton–Raphson scheme with $x_0 = 1.1$ as the initial iterate to find one of those roots.

(d) Taking this root, use the perturbation method to determine how quickly each method used converges to that root.

A1. A sketch of x^3 and $2x - 1$ on the same graph using a few integer values of x is sufficient to show that $x = 1$ is one root. At $x = 0$ we have $x^3 > 2x - 1$ whereas when x is very large but negative we have $x^3 < 2x - 1$. This means that there is a second real root for a negative value of x. A cubic equation must have 3 roots, but since 2 are real, so must the third be.

One ad hoc method is to make the x–term as the subject of the equation, and use that to form an iteration scheme:

$$x = \frac{x^3 + 1}{2} \qquad \text{and hence we set} \qquad x_{n+1} = \frac{x_n^3 + 1}{2}.$$

Choosing 0 as the initial guess yields the sequence, (rounded here to 4 decimal places): 0, 0.5, 0.5625, 0.5889, 0.6022, 0.6092, 0.6130, 0.6152, 0.6164, 0.6171, 0.6175, 0.6177, 0.6179, 0.6179, 0.6180. To 6 decimal places the root is 0.618034, which is $\frac{1}{2}(\sqrt{5} - 1)$. Convergence is very slow for this method. Other initial guesses will be found either to converge to this root, or to diverge.

The second ad hoc method makes the x^3 term the subject of the equation. Hence

$$x = (2x - 1)^{1/3} \qquad \text{and hence we set} \qquad x_{n+1} = (2x_n - 1)^{1/3}.$$

We get the following sequences on choosing 2 and -1 as the initial guesses:

$$2 \rightarrow 1.4422 \rightarrow 1.2352 \rightarrow 1.1371 \rightarrow 1.0841 \cdots \rightarrow 1$$

$$-1 \rightarrow -1.4422 \rightarrow -1.6439 \rightarrow -1.6246 \rightarrow -1.6197 \cdots \rightarrow -1.618034.$$

The cube root guarantees that very large initial guesses are reduced substantially in size, and one of these two roots will always be obtained using this method. The root found using the first method cannot be found using this second method, as illustrated by the following:

$$0.619 \rightarrow 0.6197 \rightarrow 0.6210 \rightarrow 0.6231 \cdots \rightarrow 1.$$

The three roots are 1, 0.618034, and -1.618034.

The Newton-Raphson method is $x_{n+1} = x_n - f(x_n)/f'(x_n)$. In this case we have

$$x_{n+1} = x_n - \frac{x^3 - 2x + 1}{3x^2 - 2} \qquad = \frac{2x^3 - 1}{3x^2 - 2}. \tag{1}$$

Using 1.1 as the initial guess we have

$$1.1 \to 1.0196319 \to 1.0010468 \to 1.0000033 \to 1.0000000$$

which is a much improved rate of convergence. This method also recovers the other two roots:

$$0.5 \to 0.6 \to 0.6173913 \to 0.6180331 \to 0.6180340$$

$$-2 \to -1.7 \to -1.6230885 \to -1.6180551 \to -1.6180340.$$

These different schemes clearly have different properties. Our ad hoc methods seem to converge for every other root, with divergence away from the intervening roots. On the other hand the Newton-Raphson method appears to find all possible roots. But given the form of the Newton-Raphson formula in equation (1), above, where $f'(x_n)$ appears in the denominator, we must avoid choosing an initial guess close to where $f'(x) = 0$.

Convergence

For the first ad hoc method, set $x_n = 1 + \epsilon$ where ϵ is very small, i.e. we are perturbing about $x = 1$. Using the iteration formula we get

$$x_{n+1} = \frac{(1+\epsilon)^3 + 1}{2} \simeq \frac{2 + 3\epsilon}{2} = 1 + \tfrac{3}{2}\epsilon,$$

thereby implying divergence from $x = 1$. For the second method we have

$$x_{n+1} = \Big(2(1+\epsilon) - 1\Big)^{1/3} = (1 + 2\epsilon)^{1/3} \simeq 1 + \tfrac{2}{3}\epsilon,$$

which implies convergence. Finally the Newton-Raphson method yields

$$x_{n+1} = \frac{2(1 + 3\epsilon + 3\epsilon^2 + \epsilon^3) - 1}{3(1 + 2\epsilon + \epsilon^2} \simeq \frac{1 + 6\epsilon + 6\epsilon^2}{1 + 6\epsilon + 3\epsilon^2} = 1 + \frac{3\epsilon^2}{1 + 6\epsilon + 3\epsilon^2} \simeq 1 + 3\epsilon^2.$$

Therefore, if ϵ is very small, then convergence is quadratic.

Q2. Find the only real root of the cubic $x^3 - x^2 - x - 1 = 0$ correct to six significant figures. Use any method you like.

A2. Of the three possible ad hoc schemes I decided to opt for the following,

$$x_{n+1} = \Big(x^2 + x + 1\Big)^{1/3},$$

because the cube root is likely to inhibit any wild excursions of the iterates. Using an initial iterate of 1, I got the answer 1.83929 after about 15 iterations, although I needed a few more just to make absolutely sure. On using a Fortran program written in quadruple precision, the answer which is correct to 12 decimal places is 1.839286755214.

I chose $x_0 = 1$ for the ad hoc scheme on a whim, although it did feel as though that was in the right ball park. Unfortunately, when one uses the Newton-Raphson method,

$$x_{n+1} = x_n - \frac{x_n^3 - x_n^2 - x_n - 1}{3x_n^2 - 2x_n - 1},$$

then there is a trap because the denominator is zero when $x_n = 1$. Therefore I used $x_0 = 3$. This gave the answer correct to over 30 decimal places in 7 iterations. That answer is

$$1.839286755214161132551852564653286 7,$$

in case you were wondering.

Q3. Use a suitable sketch to show that $f(x) = e^{-x} - x = 0$ has only one root. Use both the possible ad hoc schemes and the Newton-Raphson method to find that root. Analyze the approach to the solution for all three methods by setting $x_n = X + \epsilon_n$ where X is the solution of $f(X) = 0$, i.e. it satisfies $e^{-X} = X$.

A3. Sketching x and e^{-x} shows very easily that the curves cross once and once only since x has a positive slope and e^{-x} a negative slope.

The two prospective ad hoc schemes are,

$$x_{n+1} = e^{-x_n} \quad \text{(Scheme 1)}, \qquad x_{n+1} = -\ln x_n \quad \text{(Scheme 2)}.$$

For the first scheme we choose $x_0 = 0.5$ as the first iterate. We get,

$$0.5 \to 0.6065307 \to 0.5452392 \to 0.5797031 \to 0.5600646 \to 0.5711721 \cdots$$

and it takes anther 19 iterations to get the solution, 0.5671433, correct to seven decimal places. It is interesting to note that the successive iterates oscillate in their convergence.

For the second scheme we have,

$$0.5 \to 0.6931472 \to 0.3665129 \to 1.0037215 \to -0.0037146$$

and at this point the iteration scheme fails because then has to find the logarithm of a negative number.

The Newton Raphson method is

$$x_{n+1} = x_n - \frac{e^{-x} - x}{-e^{-x} - 1} = \cdots = \frac{(x+1)e^{-x}}{e^{-x} + 1}.$$

On choosing 1 as the initial guess we obtain,

$$1 \to 0.5378828 \to 0.5669870 \to 0.5671433 \to 0.5671433.$$

Now we analyze the approach to the root by setting $x_n = X + \epsilon$. For ad hoc scheme 1 we have,

$$x_{n+1} = e^{-X+\epsilon} = e^{-X}(1 - \epsilon + \cdots) = X(1 - \epsilon + \cdots) \simeq X - X\epsilon.$$

Here we have used the fact that $e^{-X} = X$ to simplify the appearance of the terms. Given that $|X| < 1$ this means that we have linear convergence. But given that the coefficient of ϵ is negative, it means that the iterates oscillate about the root as they approach it.

For the second ad hoc scheme we have,

$$x_{n+1} = -\ln(X+\epsilon) = -\ln X(1+\epsilon/X) = -\ln X - \ln(1+\epsilon/X) \simeq X - \epsilon/X.$$

Again we have used $-\ln X = X$ to simplify this, and the final line is obtained using a two-term binomial expansion of the logarithm. The coefficient of ϵ is negative but greater than 1 in magnitude, and therefore we have an oscillatory divergence.

Finally, for the Newton-Raphson scheme we have the following, which is a bit lengthy....

$$\begin{aligned}
x_{n+1} &= \frac{x_n+1)e^{-x_n}}{1+e^{-x_n}} \\
&= \frac{(X+1+\epsilon)e^{-X+\epsilon}}{1+e^{-X+\epsilon}} \\
&= X\frac{(X+1+\epsilon)e^{\epsilon}}{1+Xe^{\epsilon}} && \text{using } e^{-X} = X \\
&= X\frac{(X+1+\epsilon)(1-\epsilon+\frac{1}{2}\epsilon^2\cdots)}{(X+1)-\epsilon X+\frac{1}{2}X\epsilon^2+\cdots} \\
&= X\Big[\frac{(X+1)-\epsilon X+\frac{1}{2}(X-1)\epsilon^2+\cdots}{(X+1)-\epsilon X+\frac{1}{2}X\epsilon^2+\cdots}\Big] \\
&= X\Big[1-\frac{\epsilon^2}{2(X+1)}+\cdots\Big] \\
&= X - \frac{X}{2(X+1)}\epsilon^2+\cdots
\end{aligned}$$

This shows that we have quadratic convergence.

Q4. Use the Newton-Raphson method to find the first 4 positive roots of $f(x) = x\sin x - 1 = 0$. Rough locations of the roots may be obtained using a suitable sketch.

A4. A sketch of $\sin x$ and $1/x$ shows that one root is below $\pi/2$, the next between $\pi/2$ and π, the third just above 2π and the fourth just below 3π. The Newton-Raphson scheme is

$$x_{n+1} = x_n - \frac{x_n\sin x_n - 1}{x_n\cos x_n + \sin x_n} = \frac{x_n^2\cos x_n + 1}{x_n\cos x_n + \sin x_n}.$$

We get the following sequences using the initial guesses, $\pi/2$, π, 2π and 3π:

$$\pi/2 \to 1 \to 1.1147287 \to 1.1141571 \to 1.1141571$$

$$\pi \to 2.8232828 \to 2.774129 \to 2.7726062 \to 2.7726047$$

$$2\pi \to 6.4423403 \to 6.439118 \to 6.439117$$

$$3\pi \to 9.3186747 \to 9.3172433 \to 9.3172429$$

— the final numbers above being the fully converged solutions.

Q5. So let us create a general perturbation analysis of the convergence of the Newton-Raphson method towards a double root. We'll fix the roots t o be at $x = 0$ reflects a general situation perfectly, and therefore we will consider $f(x) = x^2 g(x)$ where $g(0) \neq 0$. Write down the Newton-Raphson formula for this $f(x)$, and use a perturbation analysis to determine how quickly the iteration scheme will converge to $x = 0$. What happens when we have $f(x) = x^m g(x)$ where m is a positive integer?

A5. If we set $f(x) = x^2 g(x)$ then $f'(x) = 2xg(x) + x^2 g'(x)$. Therefore the Newton-Raphson scheme is

$$\begin{aligned} x_{n+1} &= x_n - \frac{x_n^2 g(x_n)}{2x_n g(x_n) + x_n^2 g'(x_n)} \\ &= x_n - \frac{x_n g(x_n)}{2g(x_n) + x_n 2g'(x_n)}. \end{aligned}$$

The double root is at $x = 0$, and therefore we will set $x = \epsilon$. If this is substituted into the Newton-Raphson formula, then we obtain,

$$\begin{aligned} x_{n+1} &= \epsilon - \frac{\epsilon g(\epsilon)}{2g(\epsilon) + \epsilon g'(\epsilon)} \\ &= \frac{2\epsilon g(\epsilon) + \epsilon^2 g'(\epsilon) - \epsilon g(\epsilon)}{2g(\epsilon) + \epsilon g'(\epsilon)} \\ &= \frac{\epsilon g(\epsilon) + \epsilon^2 g'(\epsilon)}{2g(\epsilon) + \epsilon g('\epsilon)} \\ &\simeq \tfrac{1}{2}\epsilon. \end{aligned}$$

Therefore errors are halved every iteration for the general case of a double root.

A similar analysis for the general multiple root, $f(x) = x^m g(x)$, gives

$$x_n = \epsilon \qquad \Rightarrow \qquad x_{n+1} = \frac{m-1}{m}\epsilon.$$

Therefore the linear convergence gets worse as the multiplicity of the root increases.

Q6. [This question is best tackled using some suitable software to undertake the computations.]

The objective is to find the zeros of the function, $f(x) = x^{1/3} - \ln x$, where it is no secret that any such zeros must be positive. Use both of the possible *ad hoc* methods and the Newton-Raphson method to find these zeros. I am not sure that it will be useful to sketch this function, but trialling a few tentative values of x is a good start.

A6. It is clear that $f(1) = 1$ and $f(10) = -0.14815$ and so there must be a root between $x = 1$ and $x = 10$.

It is also known that $\ln x$ grows more slowly than any positive power of x, and therefore $f(x)$ should become positive again at some point. If one tries $x = 100$, then $f(100) = 0.03642$ which is positive again. So a second root must lie between $x = 10$ and $x = 100$. So we could use these three values of x as starting iterates for the three schemes.

So we'll define the three schemes:

$$A: \ x_{n+1} = (\ln x_n)^{1/3}, \qquad B: \ x_{n+1} = e^{x_n^{1/3}}, \qquad NR: \ x_{n+1} = x_n - \frac{x_n^{1/3} - \ln x}{\frac{1}{3}x^{-2/3} - x^{-1}}.$$

The following shows the values to which the three schemes converged and the numner of iterations (shown in brackets) taken to achieve 8 significant figures of accuracy.

x_0	Scheme A	Scheme B	N.R.
1	ln(-ve)	6.405672 (42)	6.405672 (7)
10	93.354461 (60)	6.405672 (41)	6.405672 (6)
70	93.354461 (52)	6.405672 (49)	93.354461 (4)
100	93.354461 (48)	$\longrightarrow \infty$	93.354461 (3)

So the Scheme A iterations eventually yield a negative number, and therefore the iterations terminate due to not being able to find its logarithm. It is also clear that Scheme A obtains just one of the roots while Scheme B obtains the other.

The Newton-Raphson scheme gives both roots and quadratic convergence is obtained, as expected. I have included the extra value, $x_0 = 70$ because there is a watershed value of x at which $f'(x) = 0$. For this example, $x = 27$ is the watershed. When $x_0 < 27$ then the lower root is found, while the upper root is found when $x_0 > 27$. When x_0 is very close to 27, then x_1 will be miles away (which is even further than kilometers....). Perhaps megaparsecs..... More precisely, if $x_0 = 27 + \delta$ where $|\delta| \ll 1$, then I challenge you to show that $x_1 \approx 647/\delta$.

The idea for this question came from: `https://www.youtube.com/watch?v=MZEstWFl-Cc`

Q7. [If you have access to a machine/software which can compute with complex numbers, then you may undertake this question, should you wish.]

Write down the Newton-Raphson scheme for $f(x) = x^2 + 1$. Now use $x_0 = 0.5 + 0.5j$ as the initial iterate. To what value does the Newton-Raphson scheme converge?

Use the same method for finding the square root of $2j$. Use $x_0 = 1 + 0j$ in this case.

A7. The Newton-Raphson scheme is

$$x_{n+1} = x_n - \frac{x_n^2 + 1}{2x_n}.$$

The iterations are,

n	x_n
0	$0.500000 + 0.500000j$
1	$-0.250000 + 0.750000j$
2	$0.075000 + 0.975000j$
3	$-0.001716 + 0.997304j$
4	$0.000005 + 1.000002j$
5	$0.000000 + 1.000000j$
6	$0.000000 + 1.000000j$

This converges to $x = j$, perhaps not surprisingly.

For the second problem we need to use $f(x) = x^2 - 2j$, and hence,

$$x_{n+1} = x_n - \frac{x_n^2 - 2j}{2x_n}.$$

The iterations are,

n	x_n
0	$1.000000 + 0.000000j$
1	$0.500000 + 1.000000j$
2	$1.050000 + 0.900000j$
3	$0.995588 + 0.999029j$
4	$1.000007 + 0.999998j$
5	$1.000000 + 1.000000j$
6	$1.000000 + 1.000000j$

This example shows that we are, in effect, solving for two coupled equations in two unknowns, namely the real and imaginary parts of x. Although this is nice and perhaps quite surprising, a different route is required when solving five equations in five unknowns. We will not be covering this aspect, but it is an essential part of solving ODE BVPs.

Q8. [This is a project-style of question. It is lengthy and intricate, but it ends up with an algebraic equation to solve for which the Newton-Raphson method is well-suited. The background application is on the vibrations of a beam.]

First, an introduction to Ordinary Differential Eigenvalue problems. I'll summarise the process first with a 2nd order ODE, and your job will be to apply the same ideas to a 4th order ODE.

The vibrations of a taut string are described by the wave equation, and eventually one obtains the ODE,

$$\frac{d^2y}{dx^2} + \omega^2 y = 0, \qquad \text{subject to } \ y(0) = 0, \ y(1) = 0.$$

The value, y, is a displacement, like that of a violin string, and the boundary conditions represent a zero displacement at both ends, which is what one expects of a violin. Clearly $y = 0$ satisfies the ODE and boundary conditions, but we've heard violins and therefore we need nonzero solutions. The value, ω, is related to the frequency of vibration of the string, and nonzero solutions (eigensolutions!) arise for certain frequencies only, and it is these values which we seek (eigenvalues!). The analysis proceeds as follows.

The general solution is $y = A\cos\omega x + B\sin\omega x$. Given that $y(0) = 0$ we must therefore have $A = 0$, which means that we now have $y = B\sin\omega x$. Application of $y(1) = 0$ yields,

$$B\sin\omega = 0.$$

We can't have $B = 0$ because that means that string has no displacement, and that defeats the purpose of the analysis. So we must have $\omega = n\pi$, where n is a positive integer; these values of ω are called the eigenvalues of the ODE. For a chosen value of n, the associated disturbance shape is $y = B\sin n\pi x$ where B is arbitary; these are the eigensolutions.

Your task, should you wish to take it on, is to use a similar analysis of the corresponding equation for a beam, namely,

$$\frac{d^4y}{dx^4} - \omega^4 y = 0 \qquad \text{subject to } \; y(0) = y'(0) = 0, \; y(1) = y'(1) = 0.$$

The boundary conditions are consistent with those of a cantilever: zero displacement and zero slope.

(a) Use the substitution, $y = e^{\lambda x}$, to write down the general solution in terms of four functions and with four arbitrary constants. Where you have to choose between exponentials and hyperbolic functions, I would advise the hyperbolics on this occasion. Sorry.

b) Now apply the boundary conditions to get four algebraic equations. The following will be a somewhat arduous trek. The aim is to try to eliminate three of the arbitrary constants in order to have an equation involving the last arbitrary constant and an expression involing ω. If this has worked correctly you should get

$$\cos\omega \, \cosh\omega = 1. \tag{1}$$

(c) Sketch both $\cos\omega$ and $1/\cosh\omega$ to estimate where the first root of Eq. (1) might be. (Ignore the obvious one at $x = 0$ which actually yields nothing of any use!)

(d) Apply Newton-Raphson to find this first value of ω. Again, ω is the freqency of vibration of the beam and, given how much more constrained the beam is compared with the string, you should obtain a higher lowest frequency here. In the solutions I will provide the first four values of ω and the corresponding shapes of vibration.

A8. We are solving,

$$\frac{d^4y}{dx^4} - \omega^4 y = 0 \qquad \text{subject to } \; y(0) = y'(0) = 0, \; y(1) = y'(1) = 0.$$

If we substitute $y = e^{\lambda x}$ then the Auxiliary Equation is,

$$\lambda^4 = \omega^4.$$

Therefore, $\lambda = \pm\omega, \pm\omega j$. Given the hint in the question, I will write the general solution in the form,

$$y = A\cos\omega x + B\sin\omega x + C\cosh\omega x + D\sinh\omega x. \tag{1}$$

We'll also need the derivative:

$$y' = \omega\Big(-A\sin\omega x + B\cos\omega x + C\sinh\omega x + D\cosh\omega x\Big). \tag{2}$$

(b) Applying the four boundary conditions:

Eq. (1) at $t = 0$	$0 = A + C$
Eq. (2) at $t = 0$	$0 = B + D$
Eq. (1) at $t = 1$	$0 = A\cos\omega + B\sin\omega + C\cosh\omega + D\sinh\omega$
Eq. (2) at $t = 1$	$0 = -A\sin\omega + B\cos\omega + C\sinh\omega + D\cosh\omega.$

Clearly we have $C = -A$ and $D = -B$ and therefore the two remaining equations become,

$$A(\cos\omega - \cosh\omega) + B(\sin\omega - \sinh\omega) = 0, \tag{3}$$
$$A(-\sin\omega - \sinh\omega) + B(\cos\omega - \cosh\omega) = 0. \tag{4}$$

Now we can obtain A in terms of B in two ways and equate them, or B in terms of A in two ways and equate them. Or, even better, we may write this out in matrix/vector form:

$$\begin{pmatrix} (\cos\omega - \cosh\omega) & (\sin\omega - \sinh\omega) \\ (-\sin\omega - \sinh\omega) & (\cos\omega - \cosh\omega) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Nonzero solutions require the determinant to be zero, and hence

$$(\cos\omega - \cosh\omega)^2 + (\sin\omega - \sinh\omega)(\sin\omega + \sinh\omega) = 0. \tag{5}$$

Using $\cos^2 + \sin^2 = 1$ and $\cosh^2 - \sinh^2 = 1$, this simplifies greatly to yield,

$$\cos\omega\cosh\omega = 1,$$

as desired.

(c) If one sketches both $\cos\omega$ and $1/\cosh\omega$, then we are seeking where these graphs cross in order to obtain the value of ω that we need. Clearly $\omega = 0$ is one, but this one leads nowhere. Both functions descend as ω increases from zero. The next possible occasion when the curves can cross is when $\cos\omega$ is positive again. The value of $\cos\omega$ passes zero when $\omega = \frac{3}{2}\pi$, at which point $1/\cosh\omega$ is very small. So I would say that this would be a good starting iterate for the Newton-Raphson scheme.

(d) The Newton-Raphson scheme is,

$$\omega_{n+1} = \omega_n - \frac{\cos\omega_n\cosh\omega_n - 1}{\cos\omega_n\sinh\omega_n - \sin\omega_n\cosh\omega_n}.$$

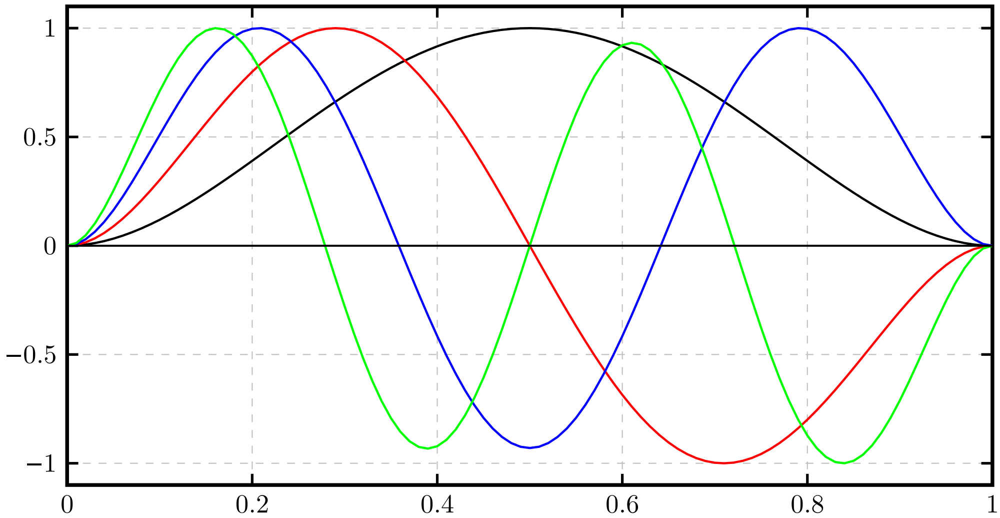
The first four values of ω are,

$$\begin{array}{c} 4.730041 \\ 7.853207 \\ 10.995252 \\ 14.137212 \end{array}$$

which should be compared with,

$$\begin{array}{c} 3/2\pi = 4.712389 \\ 5/2\pi = 7.853982 \\ 7/2\pi = 10.995574 \\ 9/2\pi = 14.137167 \end{array}$$

and which shows how close the values of ω are to odd multiples of $\frac{1}{2}\pi$.



Mode 1: Black
Mode 2: Red (one node)
Mode 3: Blue (two nodes)
Mode 4: Green (three nodes)

Just to say that, if had been considering a pin-jointed beam for which the boundary conditions are that $y = y'' = 0$ at $x = 0$ and $x = 1$, then the natural frequencies would be π, 2π, 3π and so on, which are exactly the same as for the taut string.

D.A.S.R. 15/03/2021