Problem 4: Analyzing $L_6(x)$

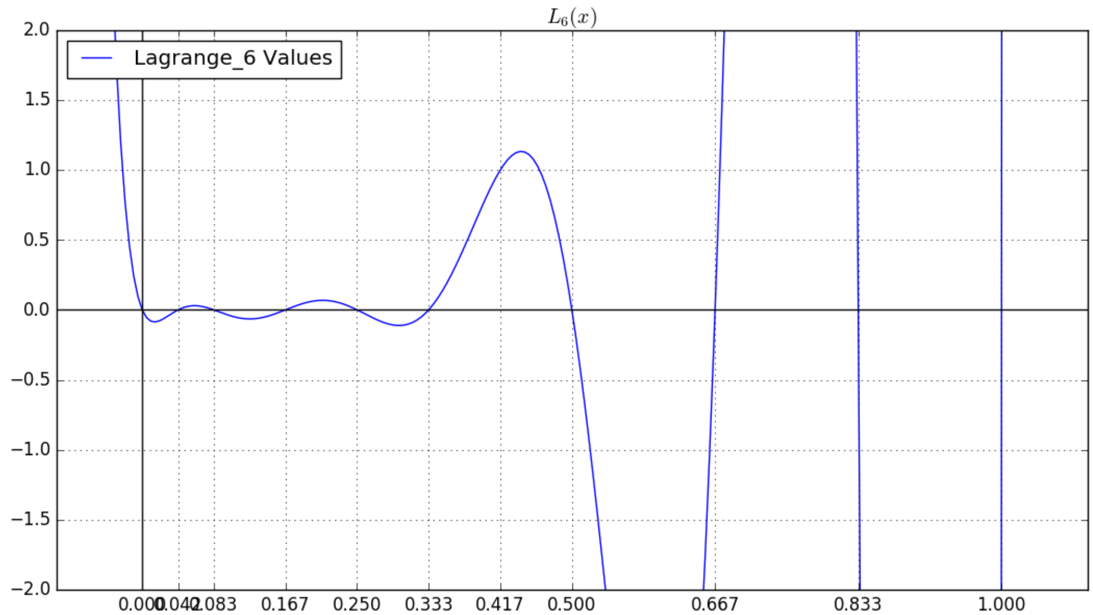

Figure 1. A plot of $L_6(x)$ that uses the nodes from problem 3. The left half is very well behaved and close to what we mighti ntuitively picture the curve as. The right half is extremely erratic. This behavior is attributed to the spacing of the nodes. The nodes on the left half are spaced closer together, while the nodes on the right half are spaced farther apart.

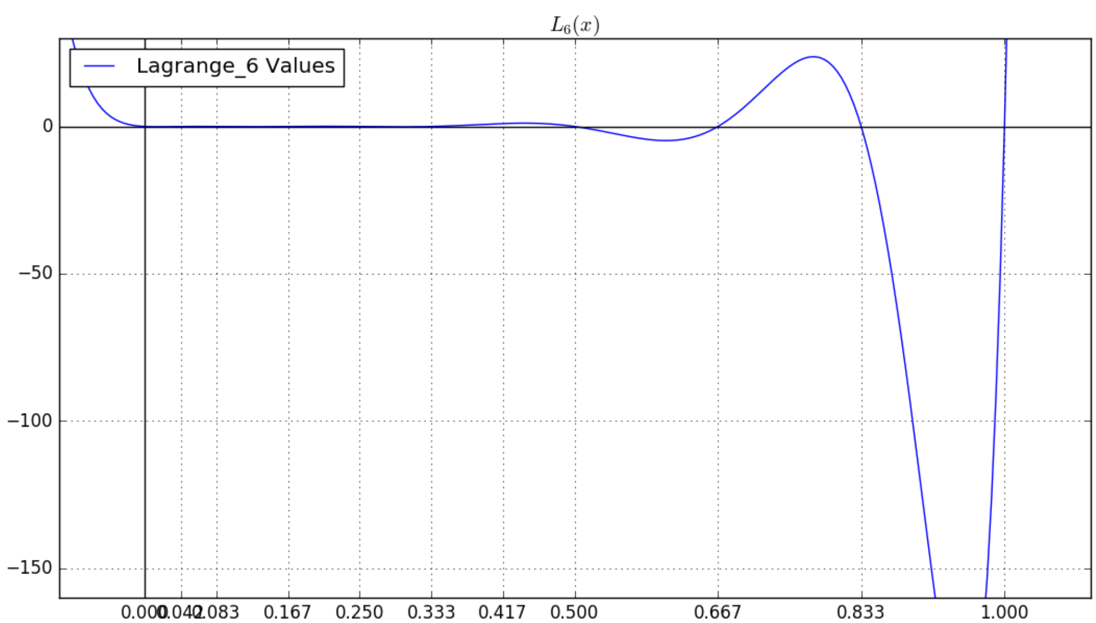

Figure 2. The same curve from figure 1 zoomed out. This shows the magnitudes of the peaks and emphasizes the scale of the deviation from zero of the right half.

Figure 1 shows $L_6(x)$ for the nodes in problem 3. With the knowledge gained from that problem, this graph is not unexpected. The interval that contains closer-spaced nodes is close to *0*. The curve varies wildly for the intervals that contain nodes that are more spread apart. These are the nodes that are greater than *0.5*. We might be able to correlate behavior between the graphs of $L_6(x)$ and the interpolant. We know from problem 3 that successive nodes that are farther apart introduce large errors. The graph of $L_6(x)$ shows explosive behavior for these nodes. However, for the nodes that are close, the curve is like a calm ripple. Larger spaced nodes seem to make the graph look like monsterous waves. Combining these two observations, we may be able to conclude that the more an interval deviates from 0, the greater the error in the approximation for values of x in that interval.

Figure 2 shows a zoomed out version of the graph in figure 1. We can appreciate the magnitudes of the peaks of the curve when $x > 0.5$. Observing that larger local extrema may produce larger approximation errors is clearer from this point of view.

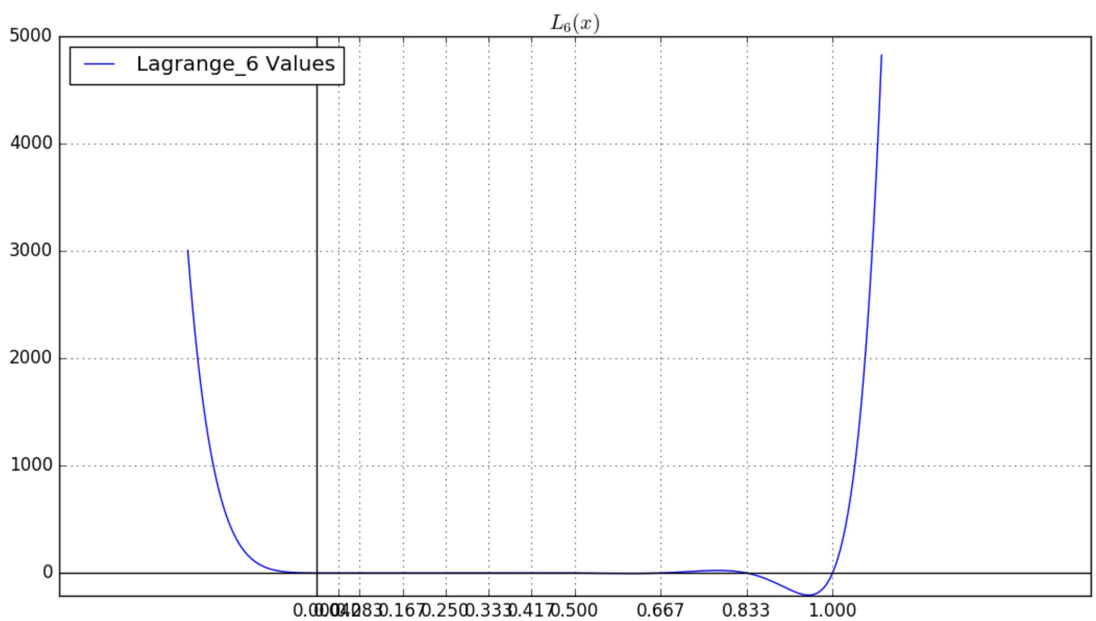

Figure 3. The same curve from figure 1 zoomed out to show the behavior of the curve when x is beyond the interval bounded by the nodes. Both sides approach infinity.

Figure 3 shows what happens when we try to evaluate a value of x that is outside the interval bounded by the first and last nodes. The curve explodes and increases upwards to infinity. Saying that approximations for x in this range would be unreliable is a tremendous understatement.

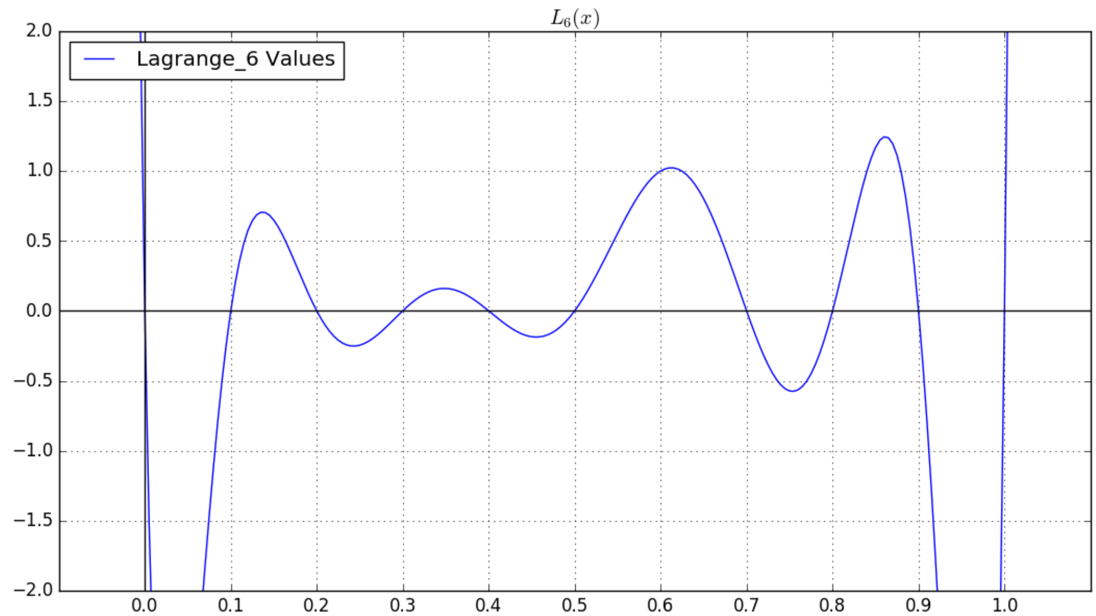

Figure 4. A plot of $L_6(x)$ that uses nodes from *0.0* to *1.0*, at intervals of *0.1*. The center region of the curve is what we expect to see: close to zero except at x_6. The curve noticably deviates from zero near the edges. It also deviates from zero on the left edge, a region where the curve in figure 1 stayed close to zero.

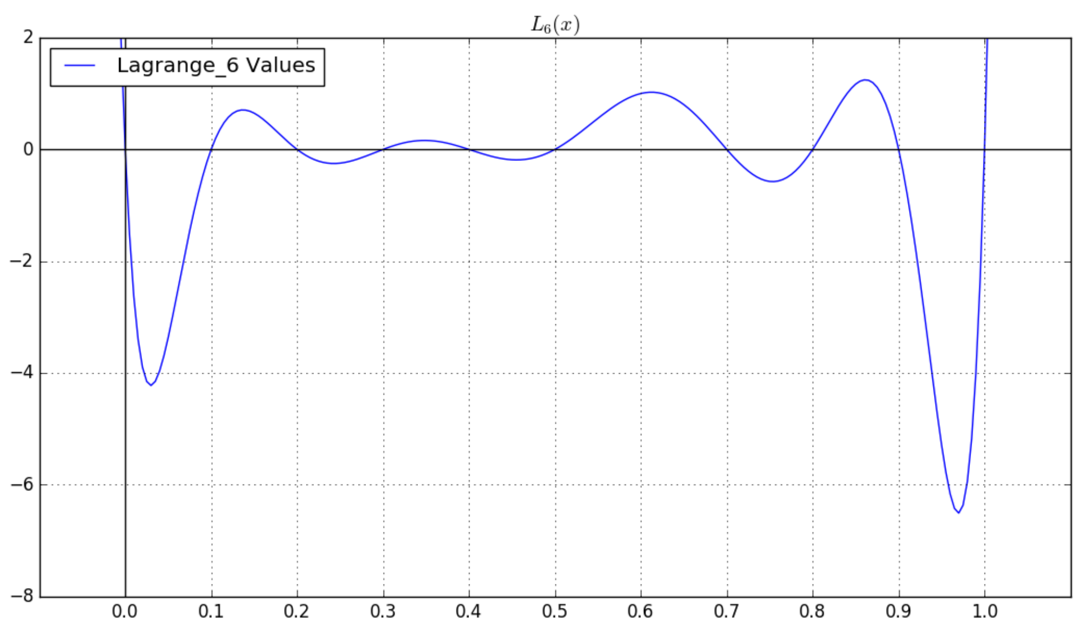

Figure 5. The same curve as in figure 4 zoomed out. The behavior at the edges, although not as bad as in figure 1, is significant.

Figure 4 shows $L_6(x)$ using nodes that are equally spaced by *0.1* units, from *0.0* to *1.0*. There are several noticable differences from the curve in figure 1. First is that the wild variations when $x > 0.5$ are gone. This is presumably due to using nodes that are closer to eachother. However, a surprising difference is the behavior when $x < 0.2$. In this interval, the curve in figure 2 is very close to zero, but in figure 5, the curve deviates considerably from it. This is strange because the nodes are all separated by *0.1*, so we would expect to see uniform behavior throughout the curve. One reason for the difference might be that the nodes used for the curve in figure 1 were separated by an even smaller distance than *0.1*. Moreover, the first and second nodes had the smallest magnitude of separation. These two factors might be what cause figure 1's curve to be so well-bahaved in this interval.

We see this deviation from zero again in figure 4 when $x > 0.8$. Although the deviation is much more tame than in figure 1, it still exists.

These irregularites in the edges highlight the behavior near the center: the interval *[0.2, 0.8]* is close to zero (except at $x_6 = 0.6$ which is where it should increase to *1)*. A hypothesis for this change in behavior again features the spacing of the nodes. Perhaps separating all nodes throughout the entire interval by the same distance is not such a good idea after all. Although this strategy seems to work fine for the nodes in the center, it seems to fail for the nodes on the edge of the interval. More care might be needed in choosing the distance of separation for these nodes. The curve in figure 1 placed its first three nodes *1/24* units apart from eachother, and its next three nodes *1/12* units. If we employ a similar tactic to the nodes for this curve, that is, gradually increasing the spacing of the nodes as they approach the center, we get the curve in figure 6. Notice that the large oscillations at the edges are now gone. This result supports the hypothesis that equally spacing the nodes throughout the entire interval is not a good strategy for getting optimal approximations with Lagrange interpolation.

A great deal of insight to the behavior of Lagrange polynomials has been gained from these results. Mainly, that the placement of the nodes is extremely important. Although we cannot say with certainty yet, a good strategy for choosing nodes seems to be equally spacing them near the center, and gradually decreasing the spacing near the edges.

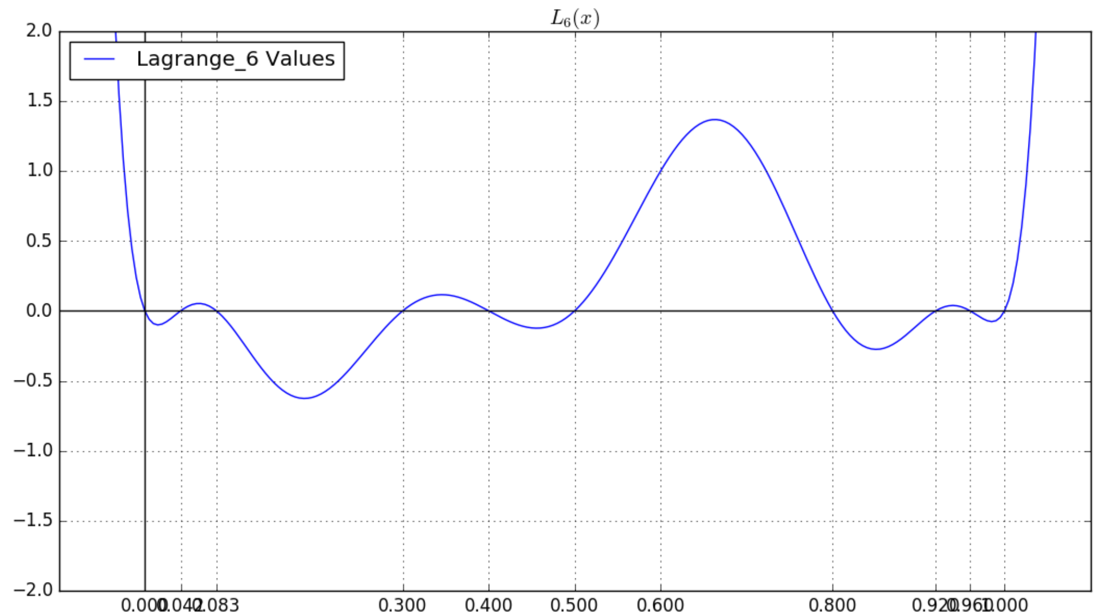

Figure 6. A plot of $L_6(x)$ that uses a combination of the nodes in the curves of figures 1 and 4: *0, 1/24, 2/24, 3/10, 4/10, 5/10, 6/10, 8/10, 9,2/10, 9.6/10, 1*. The edges no longer oscillate at high magnitudes. The curve is close to zero throughout the entire interval, except near x_6 where it should reach *1*.

Problem 4: Analyzing π(x)

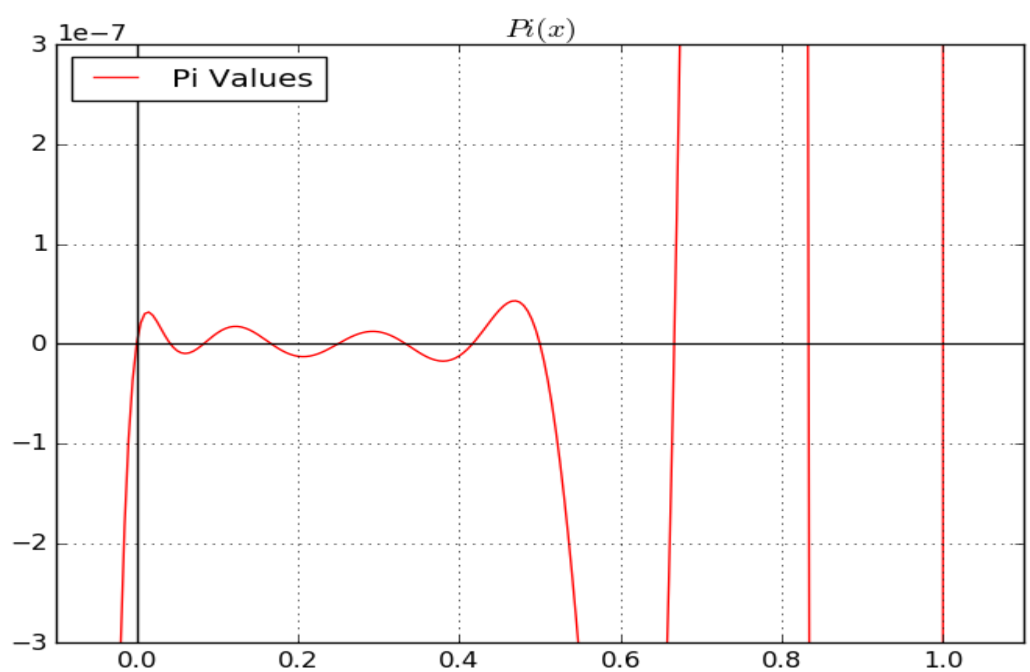

Figure7. A zoomed-in plot of $\pi(x)$ that uses the nodes from problem 3.

Similarly to the $L_6(x)$ plot that uses the nodes from problem 3, the left half (x <= 0.5) behaves very well and the right does not. This is as well due to the spacing of the nodes. The closer the nodes, the better the curve behaves.

One thing noticeable when comparing $L_6(x)$ and π(x) is that in order for us to see the similarity in the curves we have to really zoom-in for π(x). This is a good thing, as the smaller π(x) is, the smaller the error in the approximation is. We can also immediately see why $L_6(x)$ and the Lagrange interpolant from problem 3 that used the same nodes fell apart for $x > 0.5$. In this region, π(x), and thus the approximation error, is relatively large.

π(x) is used to calculate the error for Lagrange interpolation, so we would like it to be as close to zero as it can be throughout the entire interval. Recalling that π(x) = $(x - x_0)(x - x_1)...(x - x_n)$, we can see why the graph in figure 7 behaves the way it does. When $x < 0.5$, there is an x_k for which $(x - x_k)$ is small. This is because in this interval, there are more nodes that are close to eachother. The maximum distance that x can be from the nearest node is small. Thus, the product will be small as well.

On the other hand, when $x > 0.5$, the maximum distance that it can be from the nearest node increases. This is clear because the nodes are more spaced out in this interval, so that the smallest $(x - x_k)$ term will be larger than when $x < 0.5$. As a result, π(x) will be relatively large.

These observations lead us to believe that the smaller the intervals between nodes, the closer π(x) will be to zero throughout the entire interval, thus, improving the accuracy of the approximation.

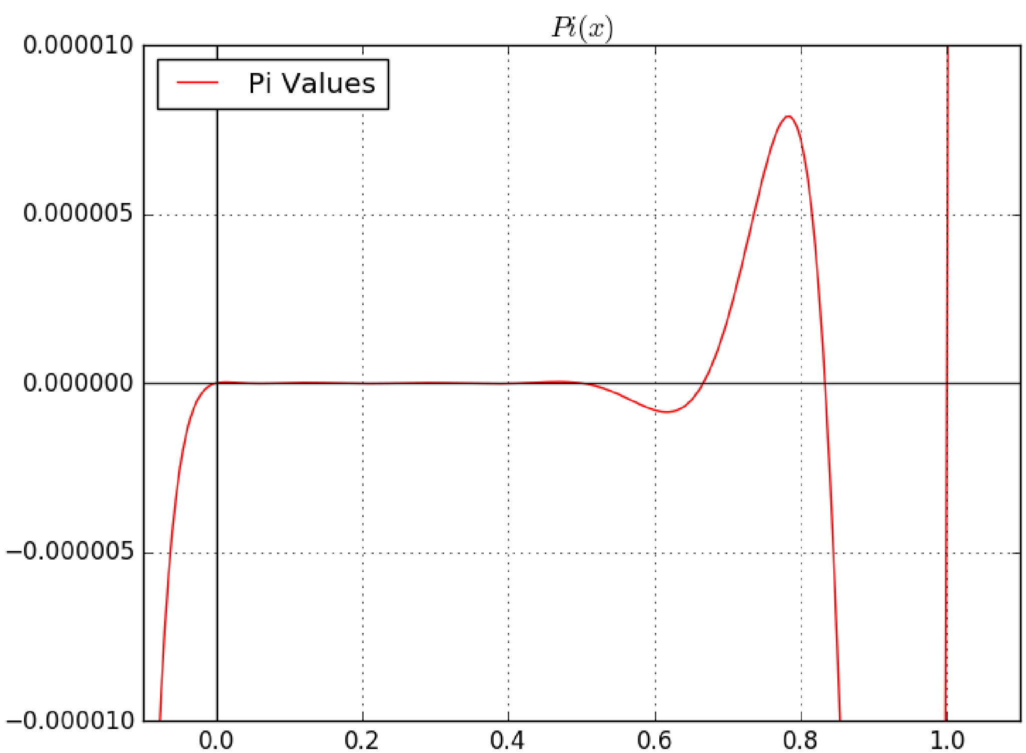

Figure 8. Zoomed-out curve from Figure 7.

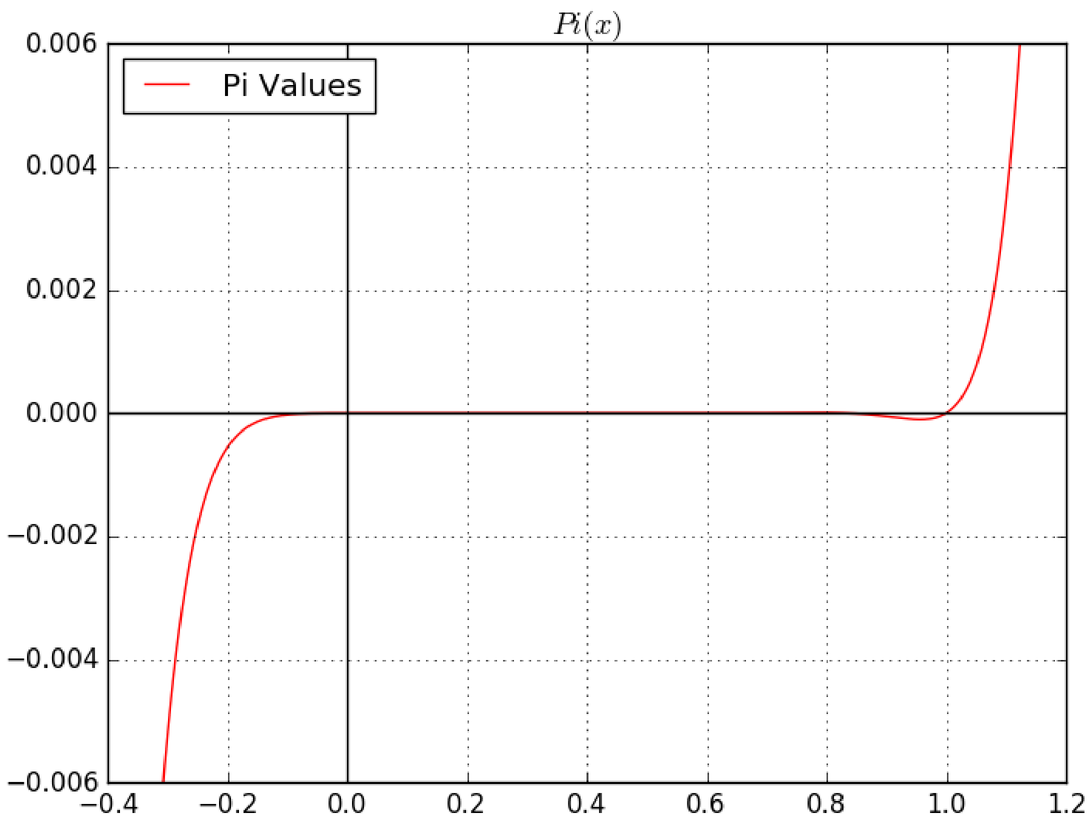

Figure 9. Same curve as in Figure 7 but showing the behavior when x is beyond the interval bounded by the nodes.

When we try to get the values outside of the interval of nodes the curve explodes to positive infinity when greater than 1 as $L_6(x)$ does. But when trying to evaluate with x less than 0 the curve, contrary to $L_6(x)$ is goes to negative infinity. In conclusion, we can say that in order to get a correct approximation we have to keep the nodes closer together and in between 0 and 1. Any other value would produce big errors.

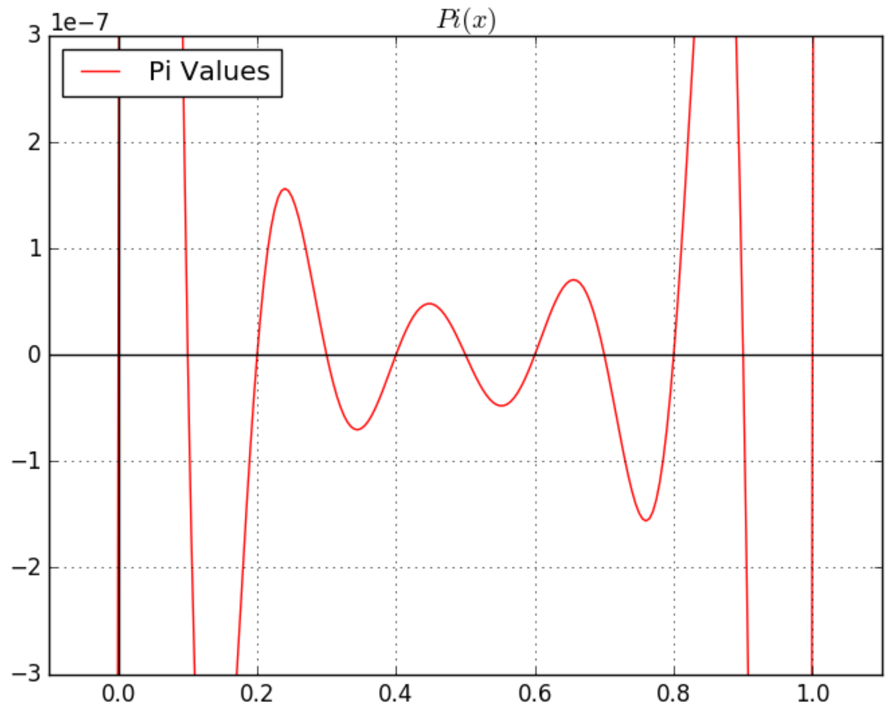

Figure 10. Zoomed-in plot of $\pi(x)$ that uses nodes from 0.0 to 1.0 at intervals of 0.1.

In figure 10 we see a plot of $\pi(x)$ that uses the nodes from 0.0 to 1.0, at intervals of 0.1. The center region of the curve is close to zero and the farther away you go the more it deviates.

The curve is symmetric about 0.5. This is due to the fact that the nodes are equally spaced and that its degree is odd. The value of $\pi(x)$ when $x > 0.5$ is just the negative of $\pi(x)$ when $x < 0.5$. Realizing this, we can expect $\pi(x)$ to be symmetric about the y-axis when its degree is even and the nodes are equally spaced. The degree of $\pi(x)$ is determined by the number of nodes used.

This curve is much like that of $L_6(x)$'s for the same nodes (figure 4). That is, it increasingly deviates from zero as x approaches the edges of the interval. This means that our approximations are good for values of x in the center, but bad for values close to the edges. As we saw for $L_6(x)$, a remedy for this might be to choose nodes that are closer together at the edges, which would hopefully bring down the error in those regions. Support for this hypothesis is given figure 7, in which the left edge of the curve is very close to zero due to the closer spacing of the nodes in that region.

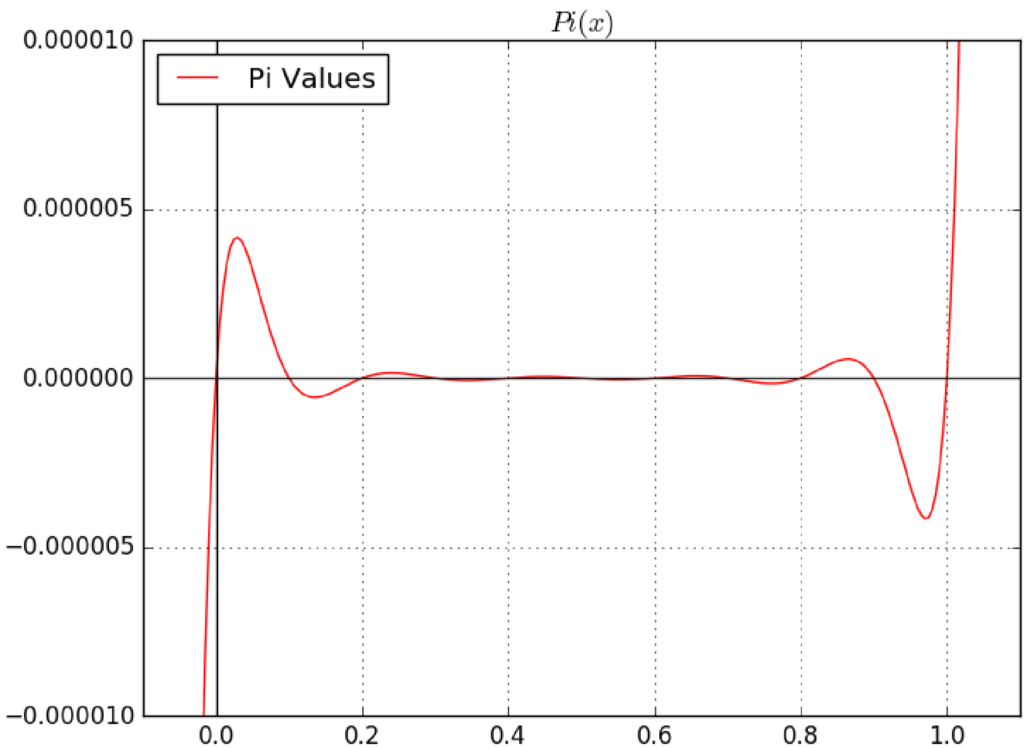

Figure 11. Zoomed-out plot of Figure 10.

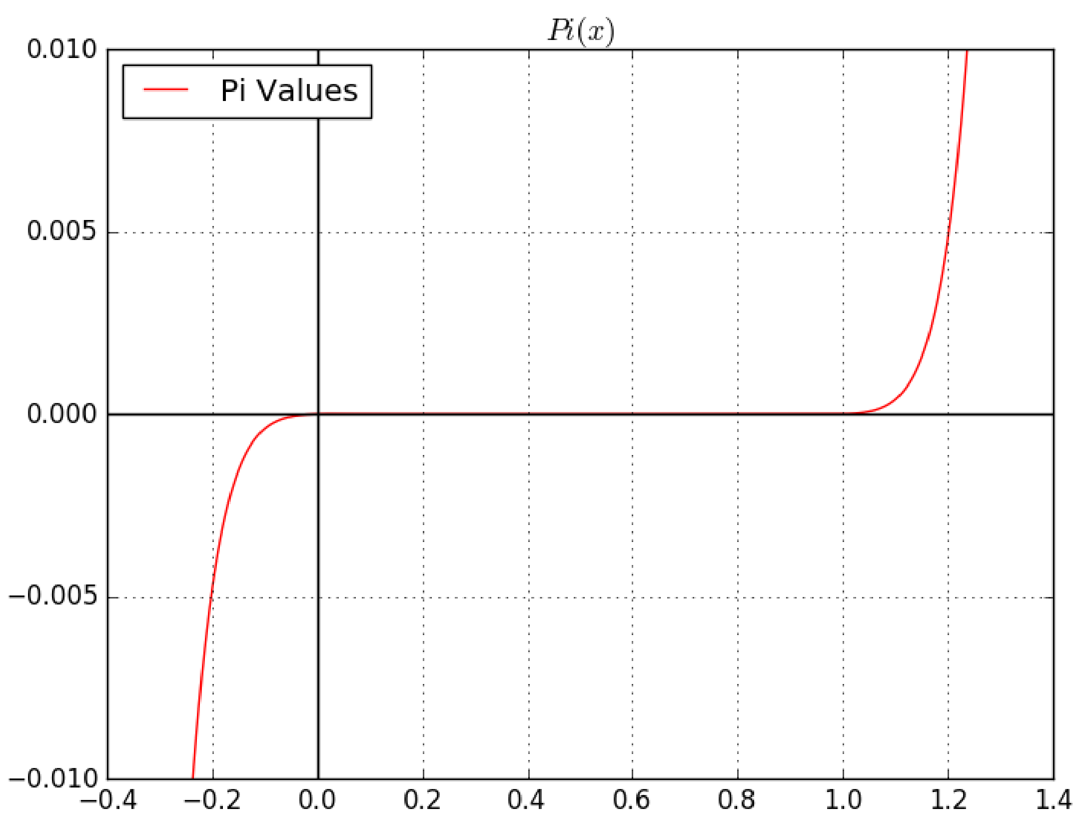

Figure 12. Same curve as in Figure 10 but showing the behavior when x is beyond the interval bounded by the nodes.

This curve in comparison to the $L_6(x)$ that uses the second set of nodes from (0.0 to 1.0, at intervals of 0.1), behaves much better and as expected. Since the nodes are located at the same distance, the farther away you get from the center the more it deviates.

In conclusion, we can see many similarities between the graphs of $\pi(x)$ and $L_6(x)$ for both sets of nodes. They exhibit similar behavior for the same sets of nodes. They both have the same problems, which can be fixed by using the same solution. If we have the liberty to choose the nodes for Lagrange interpolation, perhaps the best way to start is with $\pi(x)$, so to ensure that the set of nodes results in as small an error as possible.