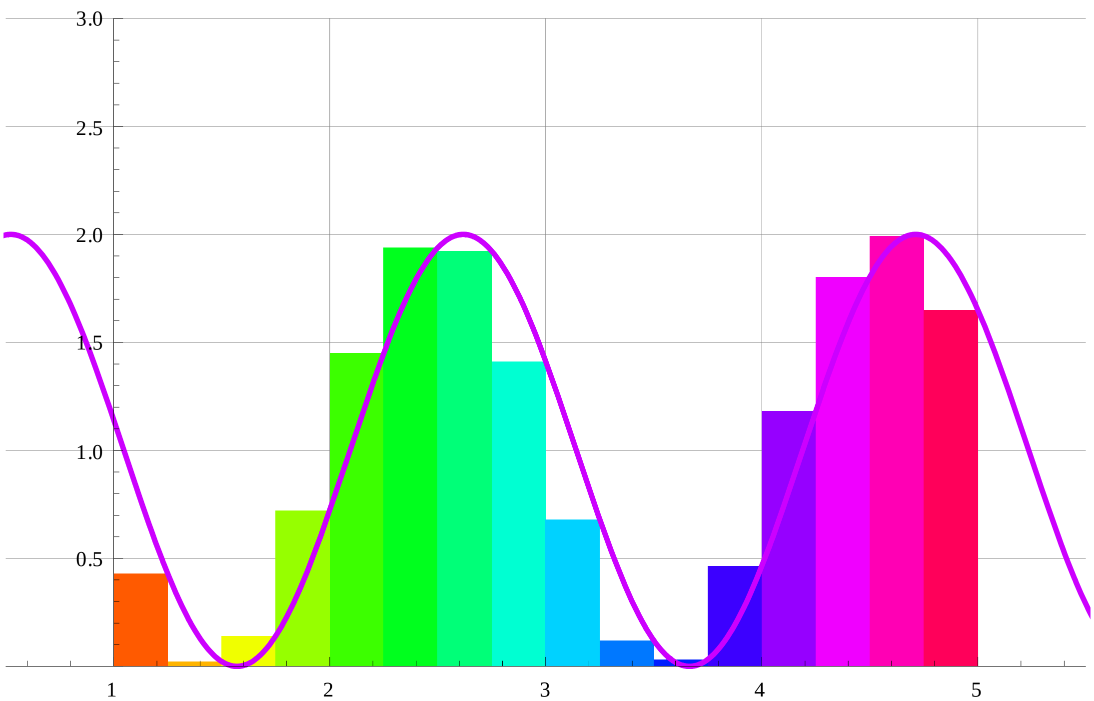
Riemann Sums Methods

While we will be mostly using the right endpoint of each interval to calculate Riemann sums in this course, this is not the only way to do it. Indeed, there are probably more ways than we can count that will satisfy the Riemann hypothesis, so we will only look at five of the standards. To be clear, any partition of the interval [*a,b*] that has each [x_i,x_{i+1}] get smaller as *n* gets larger is sufficient, but irregularly sized intervals, at least for simple functions, tend to make things harder. In general, we are looking for methods that are conceptually straightforward and are not too difficult to calculate. For polynomials, this means regularly spaced intervals.

The five methods we will be considering are: the right endpoint, the left endpoint, maximum, minimum, and the mean (midpoint). This last method is essentially the Trapezoidal rule, which we will develop a general formula for later in the course.

Let's begin with the simplest method for polynomials, **the right endpoint rule.** To refresh our memories, we are looking to fill in the components of the formula $f(x) \approx \sum_{i=1}^{n} f(x_i)\Delta x_i$. (In the limit, of course, we get true equality, but for some of our methods, we won't be developing a general rule and will only we working with finite n, so let's keep the ≈ where it is.) We first need to find Δx_i . In all our methods, we will be using regularly spaced intervals, and so $\Delta x_i = \frac{b-a}{n}$. Consider the graph below of $f(x) = \sin(3x) + 1$ on the interval [1,5]. On the image below, n=16. (We are using 16 rectangles here for the sake of illustration. You will never be asked to calculate that many by hand.) Each $\Delta x_i = \frac{5-1}{16} = \frac{4}{16} = \frac{1}{4}$. This is the width of each rectangle.



The $f(x_i)$ represents the height of the function on the right end of each interval. We need to calculate, therefore, $\sum_{i=1}^{16} f\left(1+\frac{1}{4}i\right)=\left[f\left(1+\frac{1}{4}\right)+f\left(1+\frac{2}{4}\right)+f\left(1+\frac{3}{4}\right)+\ldots+f\left(1+\frac{16}{4}\right)\right]$. Recall that for the graph above, $f(x)=\sin(3x)+1$, thus

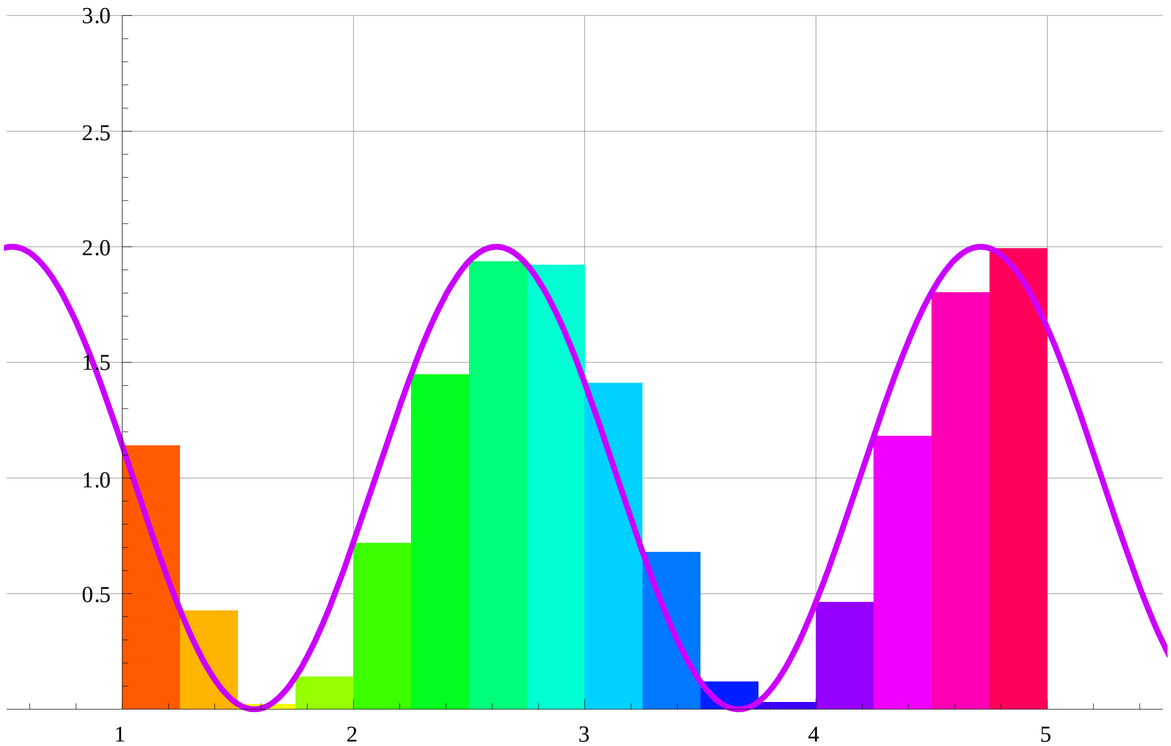
$$\sum_{i=1}^{16}\sin\left(3+\frac{3}{4}i\right)+1=\left[\sin\left(3+\frac{3}{4}\right)+\sin\left(3+\frac{6}{4}\right)+\sin\left(3+\frac{9}{4}\right)+\ldots+\sin\left(3+\frac{48}{4}\right)+16\right]\approx 15.96204195$$

(notice that each value of x_i goes up ¼, and so the 3(1/4) is what we increase by in each sine function; the 16 on the end is the sum of the 16 successive +1s). So then to estimate the region we calculate $f(x)\approx\sum_{i=1}^{n} f(x_i)\Delta x_i$ which is just $\sum_{i=1}^{16}\left[\sin\left(3+\frac{3}{4}i\right)+1\right]\frac{1}{4}\approx 3.99$.

Recall that for general *n*, we do more algebra than brute calculation. We can break the process down into six steps.

1. Find $\Delta x_i=\frac{b-a}{n}$. The values for *a* and *b* must be given, so this reduces to $\frac{k}{n}$, where *k* is some known value.
2. Find $x_i=a+i\Delta x_i$, or in this case $x_i=a+\frac{ki}{n}$.
3. Find an expression for $f(x_i)=f\left(a+\frac{ki}{n}\right)$. This involves replacing the expression for x_i in the formula and simplifying the expression. For instance, if $f(x)=2x-5$ then $f(x_i)=2\left(a+\frac{ki}{n}\right)-5=\frac{2ki}{n}+(2a-5)$. Remember that the *a* value is always given, so this will reduce to a polynomial in *i* alone (we treat n like a constant for now). The more complex the polynomial, the more involved the algebra will be here.
4. Write the expression $\sum_{i=1}^{n} f(x_i)\Delta x_i$ using the given values from #1 and #3.
5. For the general *n* case, and only when we have polynomials, we can proceed to find an exact value by using our summation formulas to replace the *i*'s in our expression. We will then only have *n* left, and we reduce. At this point, we can estimate the function for any value of n.
6. To find the exact value we take the limit as n goes to infinity: $f(x)=\lim_{n\to\infty}\sum_{i=1}^{n} f(x_i)\Delta x_i$.

The left endpoint rule is very similar to the right endpoint rule except that we are working from the left endpoint. This means we start counting at *i*=0 and go to *n-1*, or we start at *i*=1 and replace *i* with *i-1* everywhere in all our expressions. When we are working with general *n*, we just replace *i* with *i-1*. This will make the algebra in Step #3 a bit more challenging, but other aspects remain the same. When doing this for a given n, it may be conceptually easier to think of starting from 0 and leaving out the last endpoint. Consider the graph of our function with the left endpoint rule used to make the rectangles.



Notice now that for each rectangle, it's the height of the function on the left side of the region that is determining the height of the rectangle. Our Δx_i doesn't change, and the values of the function don't change either except that we use i=0 instead of *i=n*.

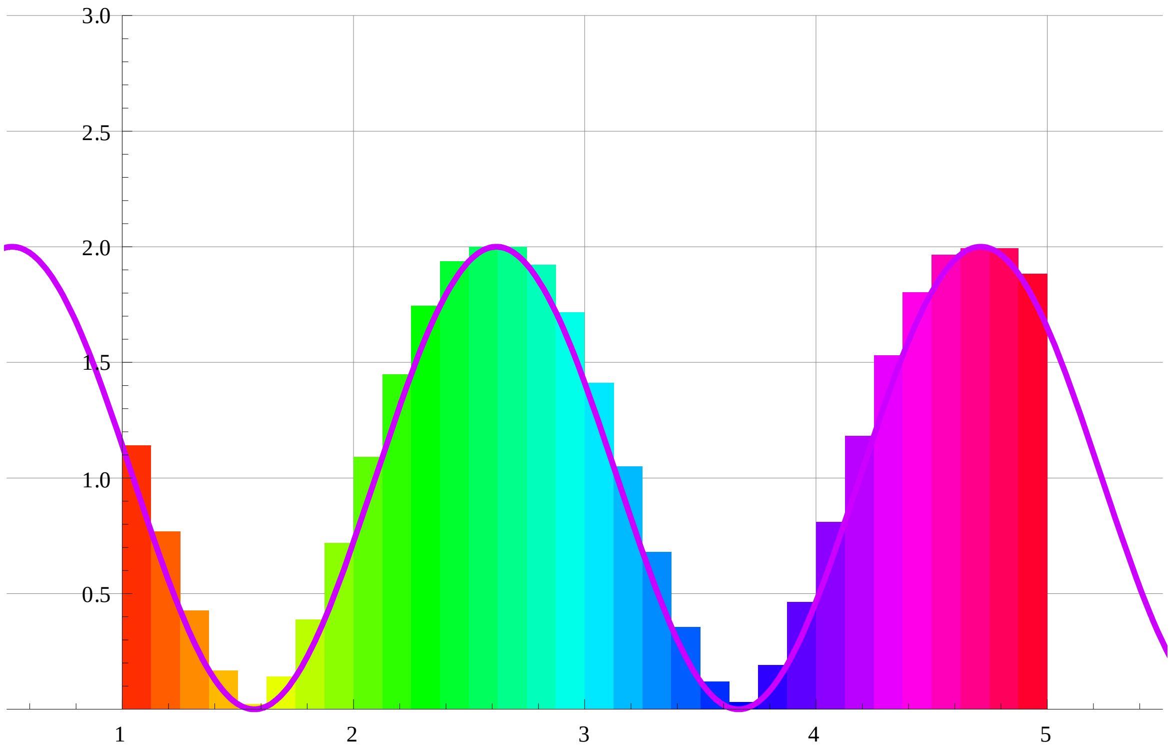
$\sum_{i=0}^{15} f\left(1+\frac{1}{4}i\right) = \left[f(1) + f\left(1+\frac{1}{4}\right) + f\left(1+\frac{2}{4}\right) + \ldots + f\left(1+\frac{15}{4}\right)\right]$ Which becomes

$\sum_{i=0}^{15} \sin\left(3+\frac{3}{4}i\right)+1 = \left[\sin(3) + \sin\left(3+\frac{3}{4}\right) + \sin\left(3+\frac{6}{4}\right) + \ldots + \sin\left(3+\frac{45}{4}\right) + 16\right] \approx 15.45287412$ and

if we multiply by ¼ now, which is the width of our rectangles we get $\sum_{i=0}^{15}\left[\sin\left(3+\frac{3}{4}i\right)+1\right]\frac{1}{4} \approx 3.86$.

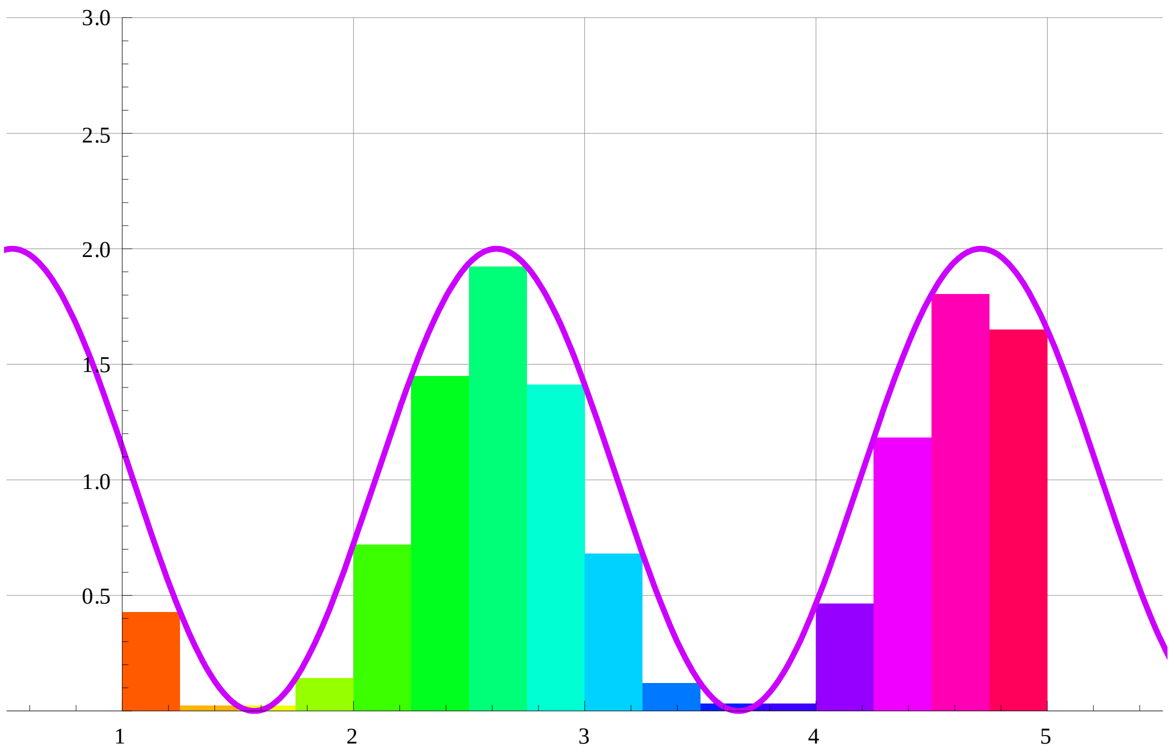
Note: Whether the left endpoint rule or the right endpoint rule will produce a larger value depends on the function. If the function is increasing on the region or decreasing on the region on will consistently come out larger. For a function like this that changes directions several times, which one is larger could depend on the number of n you choose. Notice that some rectangles overestimate the region and some underestimate it from the graphs.

Two methods that are useful theoretically are the maximum method, and the minimum method. These methods are more difficult to execute in practice unless the graph is monotonically increasing or monotonically decreasing. In such cases, they are equivalent to the right endpoint rule or the left endpoint rule (for the case of monotonically increasing graphs, the right endpoint rule is equivalent to the maximum rule, and the left endpoint rule is equivalent to the minimum rule; the reverse is true for monotonically decreasing functions). For a function that changes direction the **maximum rule** states that you choose x_i to be the *x*-value in each interval that gives you the largest value. If there is a critical point in the region, it may be some point interior to the partition and not one of the endpoints. Consider the graph below illustrating the maximum rule.



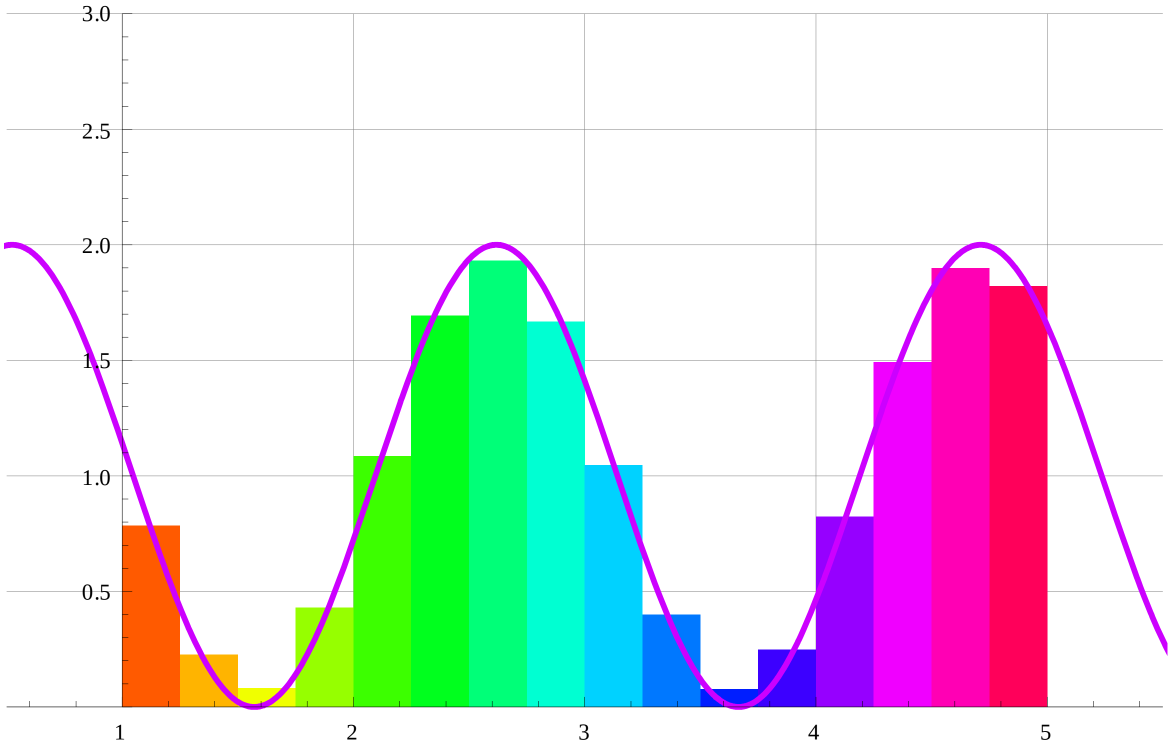
In this example, every rectangle is an overestimate for the region being measured.

In the **minimum rule**, the reverse is true: you want to choose values in the region that make the function value smallest, to generate an underestimate as shown below.



Both of these methods, because they can be difficult to implement are rarely used in numerical calculations, however, they are useful theoretically because the two methods are used in conjunction to prove that Riemann sums converge generally, since both the minimum and the maximums will converge to the same value for very large *n*.

The fifth method is **the midpoint method**. In this case, we are trying to get a better estimate for the region not by using the largest or smallest values in the region (or the leftmost or rightmost endpoints), but instead, we choose the midpoint of the region. As I mentioned earlier, this method has been formalized as the Trapezoidal rule. It can be used for general n with polynomials, but with some difficulty. For evenly spaced intervals, it's a decent compromise between getting better accuracy and using fewer terms.



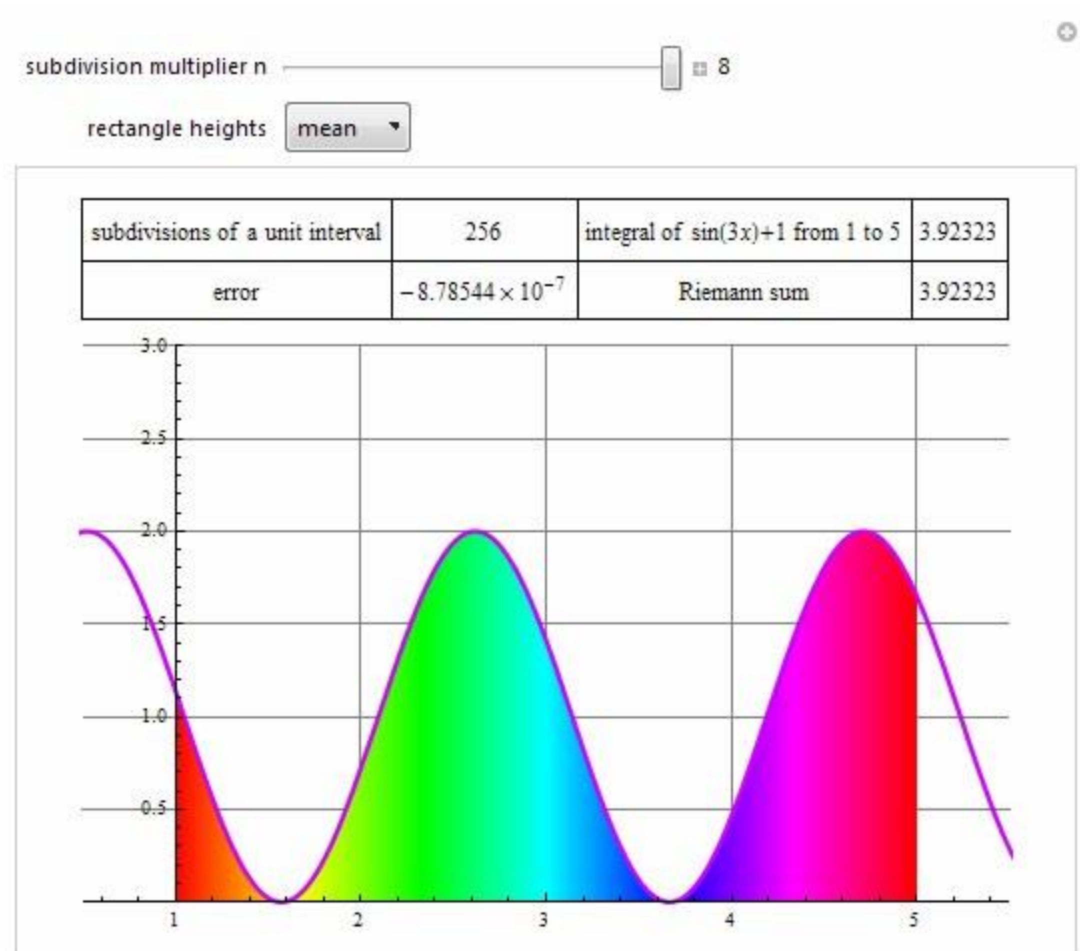
In the graph above, look carefully where the height of the rectangles are measured. They aren't measured at either end, but in the middle (technically, the average of the endpoints). This has the effect of overestimating in part of the region, and underestimating in another part of the same partition region. While certainly not perfect it does prove to be a big improvement.

To calculate this by hand, our Δx_i doesn't change. As before, it's x_i that is changing. The value we want to use in each case is $\frac{x_{i-1}+x_i}{2}$. Another way to thing about this in regularly spaced intervals is $x_i - \frac{\Delta x_i}{2}$. Our calculations then become $\sum_{i=1}^{16} f\left(1+\frac{1}{4}i-\frac{1}{8}\right) = \left[f\left(\frac{7}{8}+\frac{1}{4}\right) + f\left(\frac{7}{8}+\frac{2}{4}\right) + f\left(\frac{7}{8}+\frac{3}{4}\right) + \ldots + f\left(\frac{7}{8}+\frac{16}{4}\right)\right]$. Substituting into the function then

$$\sum_{i=1}^{16}\sin\left(3+\frac{3}{4}i-\frac{3}{8}\right)+1=\left[\sin\left(\frac{21}{8}+\frac{3}{4}\right)+\sin\left(\frac{21}{8}+\frac{6}{4}\right)+\sin\left(\frac{21}{8}+\frac{9}{4}\right)+\ldots+\sin\left(\frac{21}{8}+\frac{48}{4}\right)+16\right]\approx 15.707\ldots$$

And if we multiply by the width of our rectangles we get 3.92686, which isn't a terribly bad estimate. Better than our other ones. The true value of the area (obtained from using even more n or by the Fundamental Theorem of Calculus) is about 3.92323.

The graph below shows the results from Mathematica for n=1024.



Practice: Try calculating the Riemann sums for n=6 and n=5 on the interval [0,2] and [1,4] respectively for the following functions. For each function use the right endpoint rule, the left endpoint rule and the midpoint rule. For each of the polynomials, also find the exact value for the right endpoint rule. Compare your answers with the numerical results you obtain from your calculator. They should be approximately correct.

1. $f(x)=2x-5$
2. $f(x)=x^2+1$
3. $f(x)=x^2-x+6$
4. $f(x)=x^3$
5. $f(x)=\ln(x+1)$