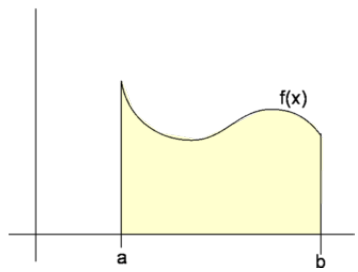
4 Integration

4.1 Areas and Distances

We begin Chapter 4 by trying to solve a simple question: given a function $f(x)$ over an interval (a,b), what is the area of the region under f between a and b?

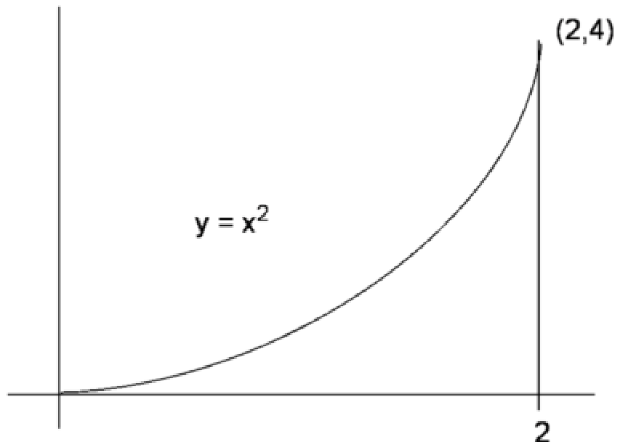


If we had an easy region under a curve, that of a simple polygon, this question is easy to answer – a rectangle has area base time height, a triangle is $\frac{1}{2}bh$, and many polygons can be broken into a sum of triangles. But in our picture above, we have a curved top side, which is not a polygon.

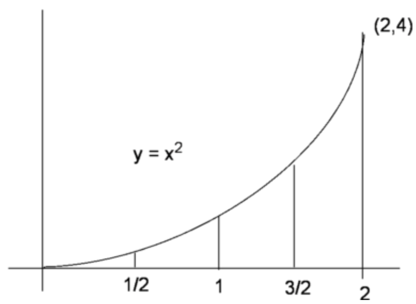
When we were looking for the slope of a tangent line, we took the limit of slopes of secant lines – a limit of an approximation. We do the same here, approximating the area under the curve using the easiest figure we have, the rectangle.

Example 4.1. Estimate the area under the curve $y = x^2$ on $(0,2)$ with 4 rectangles.

Begin by drawing a picture of the function:

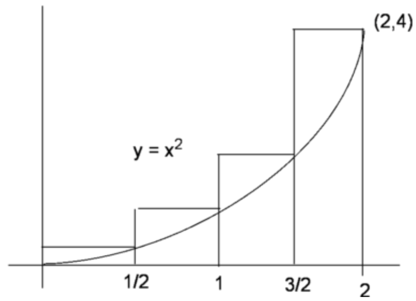


Since we can put the curve entirely inside a rectangle of width 2 and height 4, of course the area under the curve would be smaller than $2 \cdot 4 = 8$, but that's so crap, we can do better. Chop the region under the curve S into four strips, $S_1, ..., S_4$ by drawing three evenly spaced vertical lines $x = \frac{1}{2}$, $x = 1$ and $x = \frac{3}{2}$.



Then, we can make an approximation of this area of these four strips by creating rectangles. One way to do this is to extend the upper-right corner of each strip until we reach the next vertical line, a rectangle whose height is the same as the right edge. Since the function is $f(x) = x^2$, the height of these rectangles will just be the value of $f(x)$ when evaluated at the right endpoint of the intervals $[0, 1/2]$, $[1/2, 1]$, $[1, 3/2]$, $[3/2, 2]$.

Each rectangle has a width of $1/2$, and if we let A_i be the area of the i-th rectangle, we get



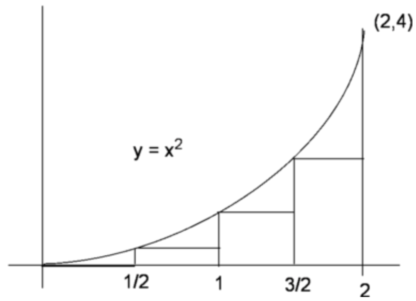
$$\begin{aligned} A_1 &= f\left(\frac{1}{2}\right)\cdot\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right) = \frac{1}{8} \\ A_2 &= f(1)\cdot\left(\frac{1}{2}\right) = (1)^2\cdot\left(\frac{1}{2}\right) = \frac{1}{2} \\ A_3 &= f\left(\frac{3}{2}\right)\cdot\left(\frac{1}{2}\right) = \left(\frac{3}{2}\right)^2\cdot\left(\frac{1}{2}\right) = \frac{9}{8} \\ A_4 &= f(2)\cdot\left(\frac{1}{2}\right) = (2)^2\cdot\left(\frac{1}{2}\right) = 2 \end{aligned}$$

$$R_4 = \frac{1}{8} + \frac{1}{2} + \frac{9}{8} + 2 = \frac{15}{4}.$$

Now because we counted alot of area above the curve in this approximation, we know that

$$A < R_4 = \frac{15}{4}.$$

However, this is not the only approximation we could do. Instead of extending the upper-right corner, we could extend the upper-left corner and create rectangles that way. If the heights of the rectangles are equal to the function evaluated at the left endpoint, we would have the following picture:



These rectangles still have width 1/2, but if we let the area of all four be L_4, then we again add base times height 4 times:

$$L_4 = \left(\frac{1}{2}\right)(0)^2 + \left(\frac{1}{2}\right)\cdot\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)\cdot(1)^2 + \left(\frac{1}{2}\right)\cdot\left(\frac{3}{2}\right)^2 = \frac{7}{4}$$

Just as the area of A was smaller then R_4, we know that the area of A is larger than L_4, so

$$L_4 = \frac{7}{4} < A.$$

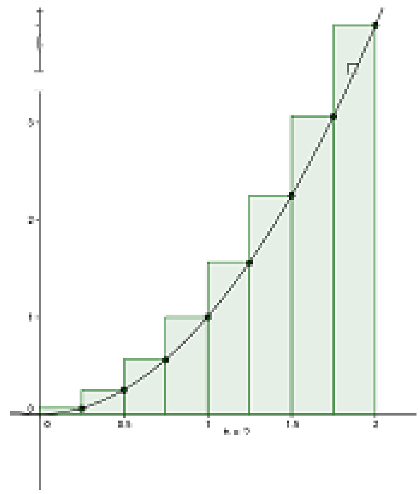
Combining these, we have upper and lower bounds on the area of S:

$$\frac{7}{4} < S < \frac{15}{4}.$$

We could repeat this procedure with more and more rectangles, and the smaller the width of the rectangle it makes sense that the approximation will get better.

Example: Let's go ahead and try to estimate the area under $y = x^2$ on [0,2], but with 8 rectangles. We'll see if we get a better estimate.

1. Using the right hand endpoints



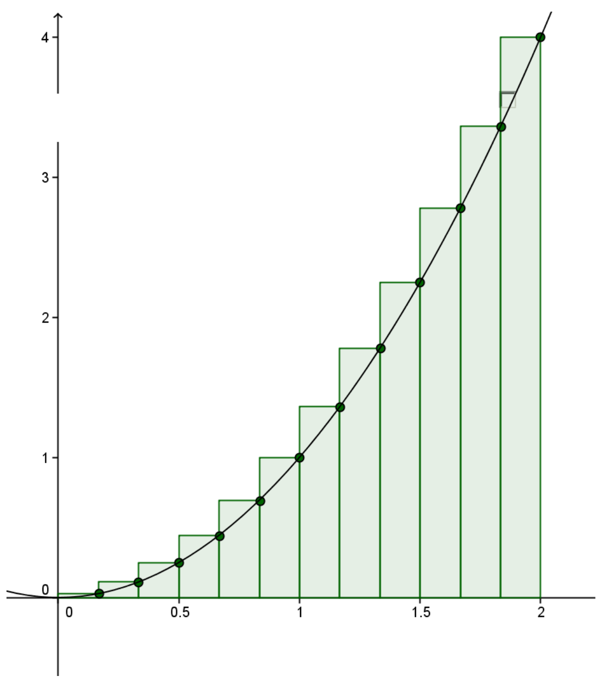
Each rectangle has a width of $1/4$, and if we let A_i be the area of the i-th rectangle, we get

$$
\begin{aligned}
A_1 &= f\left(\frac{1}{4}\right)\cdot\left(\frac{1}{4}\right) = \left(\frac{1}{4}\right)^2\cdot\left(\frac{1}{4}\right) = 0.015625 \\
A_2 &= f\left(\frac{2}{4}\right)\cdot\left(\frac{1}{4}\right) = \left(\frac{2}{4}\right)^2\cdot\left(\frac{1}{4}\right) = 0.0625 \\
A_3 &= f\left(\frac{3}{4}\right)\cdot\left(\frac{1}{4}\right) = \left(\frac{3}{4}\right)^2\cdot\left(\frac{1}{4}\right) = 0.140625 \\
A_4 &= f\left(\frac{4}{4}\right)\cdot\left(\frac{1}{4}\right) = \left(\frac{4}{4}\right)^2\cdot\left(\frac{1}{4}\right) = 0.25 \\
A_5 &= f\left(\frac{5}{4}\right)\cdot\left(\frac{1}{4}\right) = \left(\frac{5}{4}\right)^2\cdot\left(\frac{1}{4}\right) = 0.390625 \\
A_6 &= f\left(\frac{6}{4}\right)\cdot\left(\frac{1}{4}\right) = \left(\frac{6}{4}\right)^2\cdot\left(\frac{1}{4}\right) = 0.5625 \\
A_7 &= f\left(\frac{7}{4}\right)\cdot\left(\frac{1}{4}\right) = \left(\frac{7}{4}\right)^2\cdot\left(\frac{1}{4}\right) = 0.765625 \\
A_1 &= f\left(\frac{8}{4}\right)\cdot\left(\frac{1}{4}\right) = \left(\frac{8}{4}\right)^2\cdot\left(\frac{1}{4}\right) = 1
\end{aligned}
$$

$$R_8 = 0.015625 + 0.0625 + 0.140625 + 0.25 + 0.390625 + 0.5625 + 0.765625 + 1 = 3.1875.$$

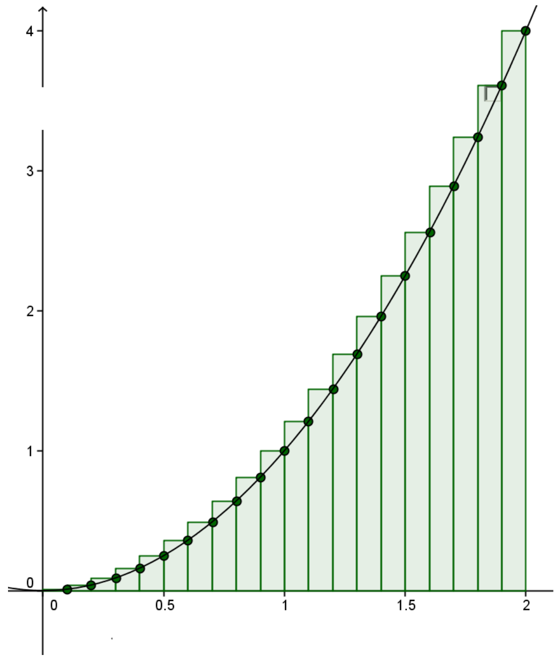
One thing to note is $\frac{7}{4} < 3.1875 < \frac{15}{4}$. This means we do have a better estimate. We can keep doing this to get better estimates. Here's what it would look like by taking more rectangles.

1. 12 rectangles



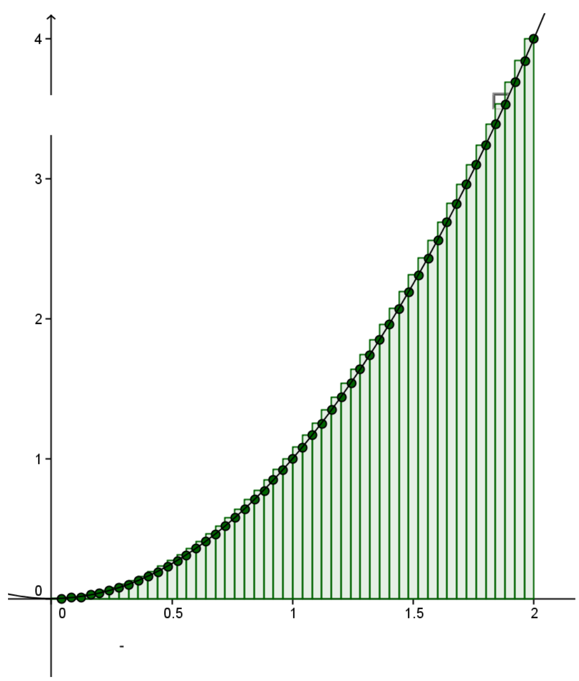
Area ≈ 3.0093

2. 20 rectangles



Area ≈ 2.87

3. 50 rectangles



Area ≈ 2.747

You can see that we are getting better estimates. However, it appears that I'd have to take a lot more rectangles to get really close to 2.666. If I used 100 rectangles, the approximation is 2.7068.

Example 4.2. For the region in the previous problem, take an indefinite number of right-edge rectangles and find

$$\lim_{n\to\infty} R_n.$$

Okay, this is harder to draw. We have the same curve, but we are taking basically infinitely many rectangles to approximate the area. Let's start with just some number n rectangles. Since we are chopping the interval $[0, 2]$ into n even pieces, each rectangle will have the same width. The formula for the width of the rectangles is

$$\Delta x = \frac{b-a}{n}$$

which in our case is,

$$\frac{2-0}{n} = \frac{2}{n}$$

Thus, the right endpoints will be

$$\begin{aligned}
x_1 = a + 1\Delta x &\quad\rightarrow\quad 0 + 1 \cdot \frac{2}{n} = \frac{2}{n} \\
x_2 = a + 2\Delta x &\quad\rightarrow\quad 0 + 2 \cdot \frac{2}{n} = \frac{4}{n} \\
x_3 = a + 3\Delta x &\quad\rightarrow\quad 0 + 3 \cdot \frac{2}{n} = \frac{6}{n} \\
x_i = a + i\Delta x &\quad\rightarrow\quad 0 + i \cdot \frac{2}{n} = \frac{2i}{n} \\
x_n = a + n\Delta x &\quad\rightarrow\quad 0 + n \cdot \frac{2}{n} = \frac{2n}{n} = 2
\end{aligned}$$

The area of the rectangles are

$$\begin{aligned}
\text{1st rectangle:} A_1 &= f(x_1) \cdot \Delta x \\
\text{2st rectangle:} A_2 &= f(x_2) \cdot \Delta x \\
\text{3st rectangle:} A_3 &= f(x_3) \cdot \Delta x \\
\text{and so on...}
\end{aligned}$$

Area of all the rectangles are

$$A \approx f(x_1) \cdot \Delta x + f(x_2) \cdot \Delta x + f(x_3) \cdot \Delta x + ... + f(x_n) \cdot \Delta x$$

So the rectangles are

$$\begin{aligned}
A_1 &= f\left(\frac{2}{n}\right) \cdot \frac{2}{n} = \left(\frac{2}{n}\right)^2 \cdot \frac{2}{n} \\
A_2 &= f\left(\frac{2}{n}\right) \cdot \frac{2}{n} = \left(\frac{4}{n}\right)^2 \cdot \frac{2}{n} \\
A_3 &= f\left(\frac{6}{n}\right) \cdot \frac{2}{n} = \left(\frac{6}{n}\right)^2 \cdot \frac{2}{n} \\
A_i &= f\left(\frac{2i}{n}\right) \cdot \frac{2}{n} = \left(\frac{2i}{n}\right)^2 \cdot \frac{2}{n} \\
... \\
A_n &= f\left(\frac{2n}{n}\right) \cdot \frac{2}{n} = \left(\frac{2n}{n}\right)^2 \cdot \frac{2}{n}
\end{aligned}$$

Let's write it as one long sum and simplify

$$
\begin{aligned}
R_n &= \left(\frac{2}{n}\right)^2 \cdot \frac{2}{n} + \left(\frac{4}{n}\right)^2 \cdot \frac{2}{n} + \left(\frac{6}{n}\right)^2 \cdot \frac{2}{n} + \ldots + \left(\frac{2n}{n}\right)^2 \cdot \frac{2}{n} \\
&= \frac{2}{n} \cdot \frac{1}{n^2} \left(2^2 + 4^2 + 6^2 + \ldots + (2n)^2\right) \\
&\quad \text{factor out } 2^2 = 4 \text{ from every term} \\
&= \frac{8}{n^3} \left(1^2 + 2^2 + 3^2 + \ldots + n^2\right)
\end{aligned}
$$

In order to finish, we need to use a commonly known identity – the sum of the squares of the first n integers:

$$1^2 + 2^2 + 3^2 + 4^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

Thus, we can substitute this into our equation for R_n and we have

$$R_n = \frac{8}{n^3} \cdot \frac{n(n+1)(n+2)}{6} = \frac{4n(n+1)(2n+1)}{3n^3}.$$

Thus, to find the exact value, we need to take a limit as $n \to \infty$:

$$\begin{aligned}
A &= \lim_{n\to\infty} R_n = \lim_{n\to\infty} \frac{4n(n+1)(2n+1)}{3n^3} \\
&= \frac{4}{3}\lim_{n\to\infty} \frac{(n+1)(2n+1)}{n^2} \\
&= \frac{4}{3}\lim_{n\to\infty} \frac{2n^2+3n+1}{n^2} \\
&= \frac{4}{3}\lim_{n\to\infty} \frac{2+3/n+1/n^2}{1} \\
&= \frac{8}{3} \\
&\approx 2.667
\end{aligned}$$

Thus, it seems that the exact area under the curve $f(x) = x^2$ over the interval $(0, 2)$ is $\frac{8}{3}$. This would work for the left endpoint rectangles too, L_n, and this is exactly how we define the area under a curve:

$$A = \lim_{n\to\infty} L_n = \lim_{n\to\infty} R_n.$$

Example 4.3. Estimate the area under the graph of $y = \frac{1}{1+x^2}$ from -2 to 2 using 6 rectangles.

1. Width: $\Delta = \frac{b-a}{n} = \frac{2-(-2)}{6} = \frac{2}{3}$

2. Find $x_1, x_2, x_3, x_4, x_5, x_6$

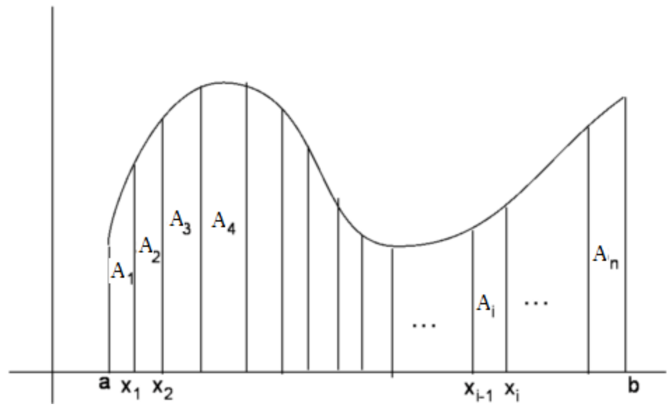
$$\begin{aligned}
x_1 &= a + 1\Delta x = -2 + 1 \cdot \frac{2}{3} = -\frac{4}{3} \\
x_2 &= a + 2\Delta x = -2 + 2 \cdot \frac{2}{3} = -\frac{2}{3} \\
x_3 &= a + 3\Delta x = -2 + 3 \cdot \frac{2}{3} = 0 \\
x_4 &= a + 4\Delta x = -2 + 4 \cdot \frac{2}{3} = \frac{2}{3} \\
x_5 &= a + 5\Delta x = -2 + 5 \cdot \frac{2}{3} = \frac{4}{3} \\
x_6 &= a + 6\Delta x = -2 + 6 \cdot \frac{2}{3} = 2
\end{aligned}$$

3. Now just plug everything in

$$\begin{aligned}
A &= f(x_1) \cdot \Delta x + f(x_2) \cdot \Delta x + f(x_3) \cdot \Delta x + f(x_4) \cdot \Delta x + f(x_5) \cdot \Delta x + f(x_6) \cdot \Delta x \\
&= \frac{1}{1 + (-4/3)^2} \cdot \frac{2}{3} + \frac{1}{1 + (-2/3)^2} \cdot \frac{2}{3} + \frac{1}{1 + (0)^2} \cdot \frac{2}{3} + \frac{1}{1 + (2/3)^2} \cdot \frac{2}{3} + \frac{1}{1 + (4/3)^2} \cdot \frac{2}{3} \\
&+ \frac{1}{1 + (2)^2} \cdot \frac{2}{3} \\
&= \frac{2}{3} [0.36 + 0.6923 + 1 + 0.69203 + 0.36 + 0.2] \\
&= 2.203
\end{aligned}$$

The exact value for the area is approximately 2.2143.

We now consider a more general problem, so suppose we have the following general function $f(x)$ over the interval (a, b), and we divide it into n rectangles:



Again, all of these intervals have equal width, and since the width of the entire interval is $b - a$, the width of each is

$$\Delta x = \frac{b-a}{n}.$$

These n strips divide the interval $[a, b]$ into n subintervals: $[x_0, x_1], [x_1, x_2] \ldots [x_{n-1}, x_n]$. as per our diagram, we have $x_0 = a$ and $x_n = b$. We have the following right endpoints for these intervals:

$$\begin{aligned} x_1 &= a + \Delta x \\ x_2 &= a + 2\Delta x \\ x_3 &= a + 3\Delta x \\ \vdots &= \vdots \end{aligned}$$

If we wish to consider approximating the i^{th} strip of the area, we know it has width Δx and height $f(x_i)$, where x_i if the value of f at the right endpoint. Then, the area of the i^{th} rectangle is

$$A_i = \Delta x f(x_i),$$

and we can get R_n by summing together these rectangles:

$$R_n = f(x_1)\Delta x + f(x_2)\Delta x + \ldots + f(x_n)\Delta x.$$

The more and more rectangles we take, as we let $n \to \infty$, the better the approximation we will get. And this gives us our definition:

Definition 4.1. The *area* of the region A that lies underneath the graph of a positive continuous function f is the limit of the sum of approximating rectangles:

$$A = \lim_{n\to\infty} R_n = \lim_{n\to\infty} \left(f(x_1)\Delta x + f(x_2)\Delta x + \ldots + f(x_n)\Delta x\right).$$

Nice, as long as the function f is continuous, the above limit will always exist. Also, we get out the exact same value if we use the left endpoints instead:

$$A = \lim_{n\to\infty} L_n = \lim_{n\to\infty} \left(xf(x_0)\Delta + f(x_1)\Delta x + \ldots + f(x_{n-1}\Delta x)\right).$$

As a matter of fact, there's nothing special at all about choosing left or right endpoints, all that matters is that we choose some points, as long as it is in the correct interval. So if x_i^* is in $[x_{i-1}, x_i]$, we're good. We call the points $x_1^*, x_2^*, ..., x_n^*$ the sample points, and we get a very general expression for the area under the curve:

$$A = \lim_{n\to\infty} \left(xf(x_1^*)\Delta + f(x_2^*)\Delta x + \ldots + f(x_n^*)\Delta x\right).$$

In order to write this a little cleaner, we use what is known as *sigma* notation for a summation:

$$\sum_{i=1}^{n} f(x_i^*)\Delta x = f(x_1^*)\Delta x + f(x_2^*)\Delta x + \ldots + f(x_n^*)\Delta x.$$

Thus, we can write

$$A = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x$$

$$A = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_{i-1}\Delta x)$$

$$A = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*)\Delta x$$

We will also need these later.

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$

$$\sum_{i=1}^{n} i^3 = \left[\frac{n(n+1)}{2}\right]^2.$$

Example 4.4. Let A be the area under the curve of $f(x) = \sin(x)$ over $[0, \pi]$. Find the expression for A as a limit, without evaluating that limit.

Since $a = 0$ and $b = \pi$, the width of the interval is π. The width of each subinterval is

$$\Delta x = \frac{\pi - 0}{n} = \frac{\pi}{n} \text{ and } x_i = a + i \cdot \Delta x$$

Thus, we have $x_1 = \pi/n$, $x_2 = 2\pi/n$, $x_3 = 3\pi/n$, ..., $x_i = i\pi/n$. The sum of the areas of the approximating rectangles is

$$\begin{aligned} R_n &= \sum_{i=1}^{n} f(x_i)\Delta x \\ &= \Delta x \cdot \sum_{i=1}^{n} \sin(x_i) \\ &= \frac{\pi}{n}\left(\sin\left(\frac{\pi}{n}\right) + \sin\left(\frac{2\pi}{n}\right) + \ldots + \sin\left(\frac{n\pi}{n}\right)\right) \end{aligned}$$

By definition, the area under the curve is

$$A = \lim_{n\to\infty} R_n = \lim_{n\to\infty} \frac{\pi}{n}\left(\sin\left(\frac{\pi}{n}\right) + \sin\left(\frac{2\pi}{n}\right) + \ldots + \sin\left(\frac{n\pi}{n}\right)\right).$$

With sigma notation, we have

$$A = \lim_{n\to\infty} \frac{\pi}{n} \sum_{i=1}^{n} \sin\left(\frac{i\pi}{n}\right).$$

This is NOT an easy sum or limit to try to do by hand, so we won't. Instead, what we can do is approximate this a little - let us do an approximation with $n = 4$ rectangles, each with width $\pi/4$. The four intervals are then

$$\left[0, \frac{\pi}{4}\right], \left[\frac{\pi}{4}, \frac{\pi}{2}\right], \left[\frac{\pi}{2}, \frac{3\pi}{4}\right], \left[\frac{\pi}{4}, \pi\right].$$

We can choose any points we like to try to approximate, but since we only know the exact values of sin with denominators of 2, 3, 4 or 6, we use a right endpoint approximation:

$$\begin{aligned} R_4 &= \sum_{i=1}^{4} f(x_i)\delta x \\ &= \sum_{i=1}^{4} \sin(x_i)\frac{\pi}{4} \\ &= \frac{\pi}{4}\left(\sin\left(\frac{\pi}{4}\right) + \sin\left(\frac{\pi}{2}\right) + \sin\left(\frac{3\pi}{4}\right) + \sin(\pi)\right) \\ &= \frac{\pi}{4}\left(\frac{\sqrt{2}}{2} + 1 + \frac{\sqrt{2}}{2} + 0\right) \\ &= \frac{\pi}{4}\left(\sqrt{2} + 1\right) \end{aligned}$$

But, this may not be so good, so let us consider a left endpoint approximation, and we will average the two:

$$\begin{aligned}
L_4 &= \sum_{i=1}^{4} f(x_{i-1})\delta x \\
&= \sum_{i=0}^{3} \sin(x_i)\frac{\pi}{4} \\
&= \frac{\pi}{4}\left(\sin(0) + \sin\left(\frac{\pi}{4}\right) + \sin\left(\frac{\pi}{2}\right) + \sin\left(\frac{3\pi}{4}\right)\right) \\
&= \frac{\pi}{4}\left(0 + \frac{\sqrt{2}}{2} + 1 + \frac{\sqrt{2}}{2}\right) \\
&= \frac{\pi}{4}\left(\sqrt{2} + 1\right)
\end{aligned}$$

Oh, that's weird! They're the same! This does NOT mean that this is the exact answer. Much more likely is that the two approximations have the same error.

4.1.1 The Distance Problem

Suppose that we know the velocity of an object at all times, meaning that we know the value of $v(t)$ for all possible values of t in the domain of the function. If we have constant velocity, then the distance is very easy to compute, we simply take distance equals velocity times time.

$$d = vt$$

But, besides that, it isn't so simple.

Example 4.5. Suppose the mileage on a car doesn't work and we want to estimate the distance we drive over a minute-long time interval. We get readings off the speedometer every 10 seconds and record them as:

Time (sec)	0	10	20	30	40	50	60
Velocity (ft/sec)	22	34	46	42	38	28	34

(Note that if the units are not consistent, you must change them so that they are). In order to find how far we have traveled, we can assume that the velocity does not change over a 10 second period, and then jumps at the 10 second mark. For example, we can assume that for the first 10 seconds, we travel at 22 ft/sec. Thus, we travel

$$10 \cdot 22 = 220$$

feet in the first 10 seconds. Then, for the next 10 seconds, we assume we travel at a constant speed of 34 ft/sec, so we have another

$$10 \cdot 34 = 340$$

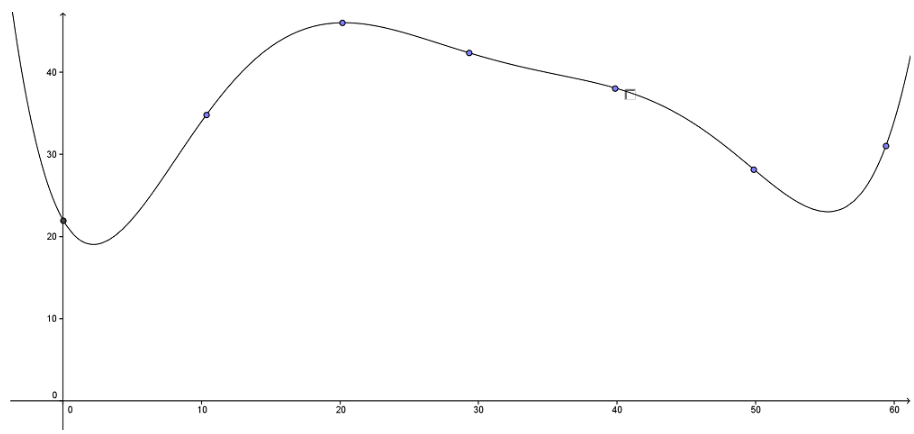
feet in the second 10 seconds. If we add up all the left endpoints ($t = 0, 10, 20, 30, 40, 50$) we get

$$D = 10(22) + 10(34) + 10(46) + 10(42) + 10(38) + 10(28) = 2100.$$

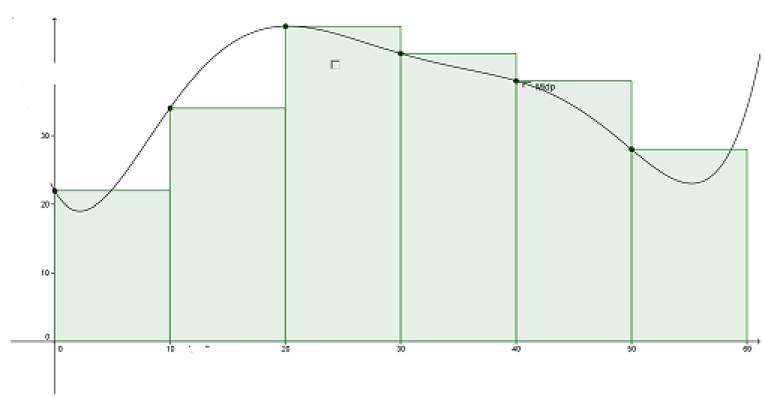
we could just as well add up all the velocities on the right endpoints, and get an approximation that way:

$$D = 10(34) + 10(46) + 10(42) + 10(38) + 10(28) + 10(34) = 2220.$$

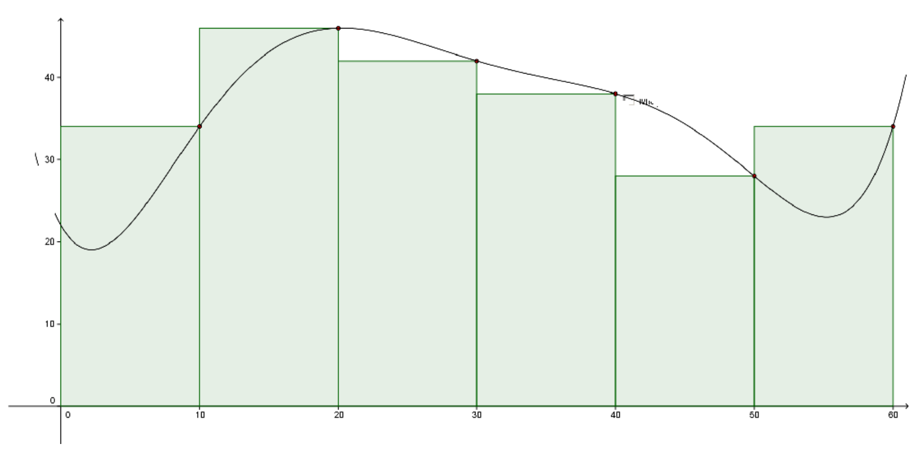
This should remind you of the sums we took in the area under the curve example. If we shorten our intervals, and take readings every 5, 2 or 1 second, we would get a more accurate distance. Simply, what we are doing is taking the area underneath the velocity curve with rectangles equally spaced apart.



1. Using the Left Hand Endpoints



2. Using the Right Hand Endpoints



In general, suppose that we have an object moving with velocity $v = f(t)$ over the interval $[a, b]$ and we need to assume, for now, that $f(t) \geq 0$ for all $t \in [a, b]$. We take readings for velocity at times $a = t_0, t_1, t_2, ..., t_n = b$. If we let all of the times be equally spaced, then the time between any two consecutive readings is

$$\Delta t = \frac{b-a}{n}.$$

Since the velocity during the first time interval is approximately $f(t_0)$, the distance traveled is $f(t_0) \cdot \Delta t$ – this is true for any of the time intervals, and thus, the total distance traveled is approximately

$$f(t_0) \cdot \Delta t + f(t_1) \cdot \Delta t + f(t_2) \cdot \Delta t + \ldots + f(t_{n-1}) \cdot \Delta t.$$

Alternately, if we consider right endpoints, then the total distance traveled is

$$f(t_1) \cdot \Delta t + f(t_2) \cdot \Delta t + f(t_3) \cdot \Delta t + \ldots + f(t_n) \cdot \Delta t.$$

And since the more often we measure velocity, the more accurate our distance becomes, the distance is the limit of the expression above:

$$d = \lim_{n\to\infty} \sum_{i=1}^{n} f(t_{i-1})\Delta t = \lim_{n\to\infty} \sum_{i=1}^{n} f(t_i)\Delta t,$$

the exact same setup as the area under a curve problem.