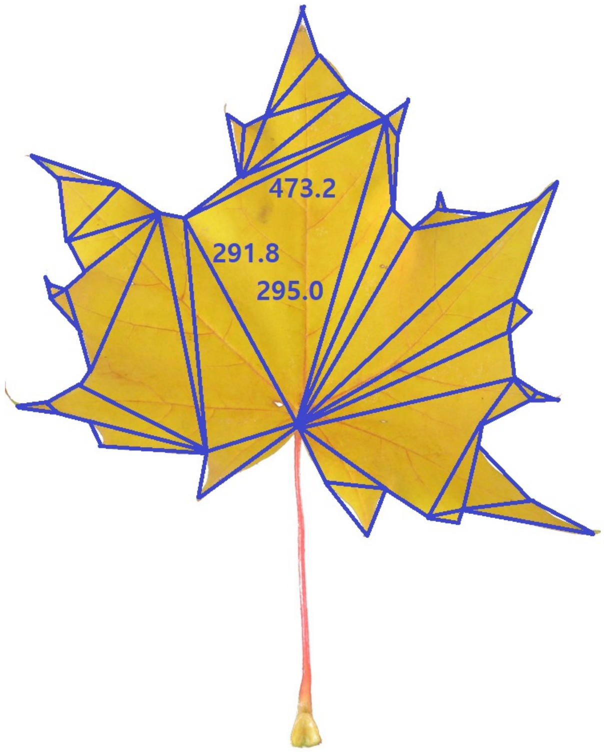
Q: How do we find the area of irregular shape?



A: Triangularization! The area of the leaf can be found by adding the areas of smaller triangles. For a smaller triangle, we can find the area if we can measure three sides.

One of the triangles is measured to have 291.8, 295.0, and 473.2. Using Heron's formula, we can find the area.

Heron's Formula for a triangle $\triangle ABC$ with the sides lengths a, b, and c:

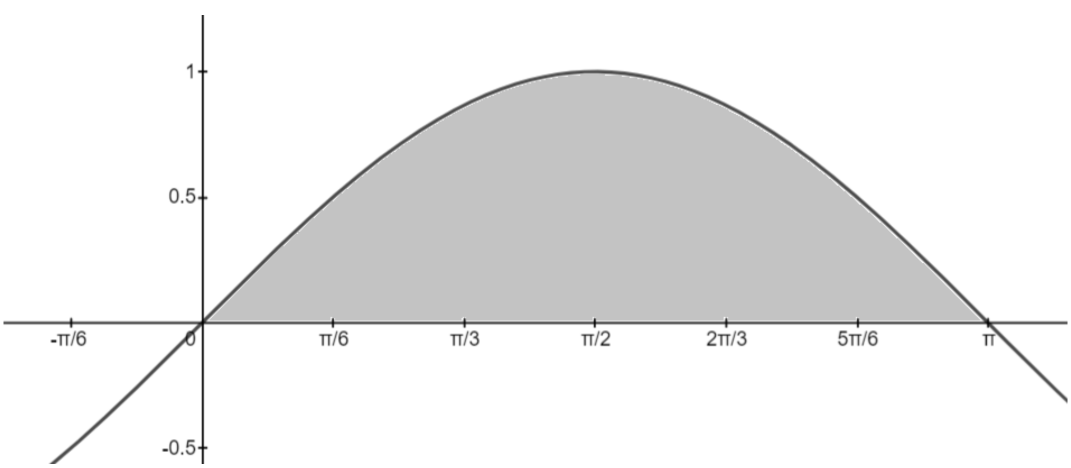
Semi-perimeter: $s = (a + b + c)/2$
Area: $A = \sqrt{s(s-a)(s-b)(s-c)}$

$s = (291.8 + 295.0 + 473.2)/2 = 530$

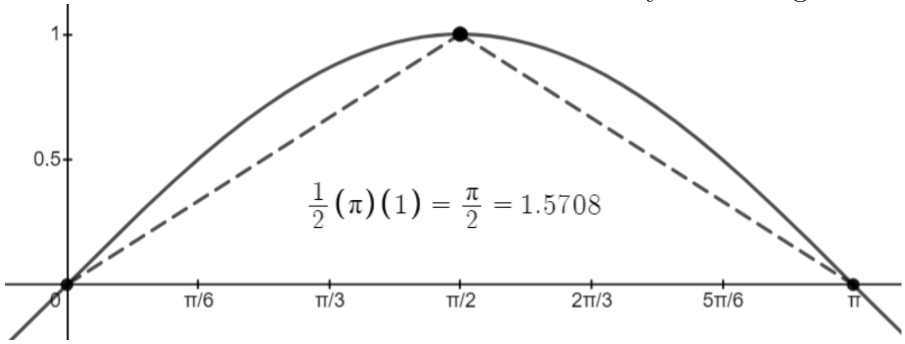
$$A = \sqrt{530(530 - 291.8)(530 - 295.0)(530 - 473.2)} = \sqrt{1685131608} = 41050.4$$

Goal: We want to estimate the area between a curve of a function and x-axis.

For instance, what is the area underneath the portion of the sine graph from $x = 0$ to $x = \pi$.

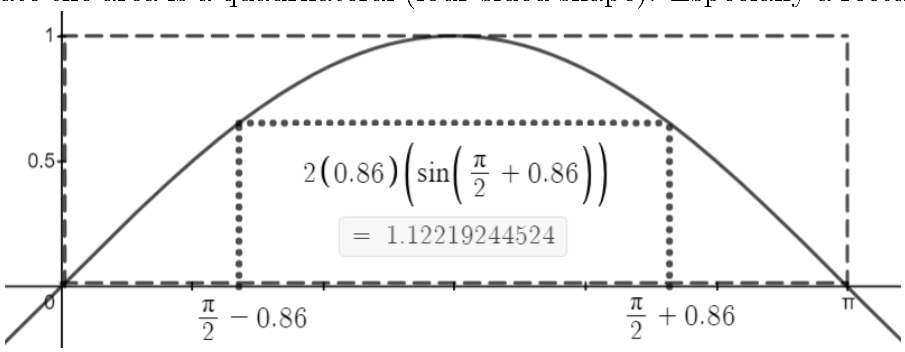


We can get the estimate of the area by using one really large triangle connecting $(0, 0)$, $(\frac{\pi}{2}, 1)$, and $(\pi, 0)$, which yields the area $\approx \frac{1}{2}(\pi)(1) = 1.5708$. This estimate is clearly underestimation as there are some leftover area not covered by the triangle. We can make a better esti-



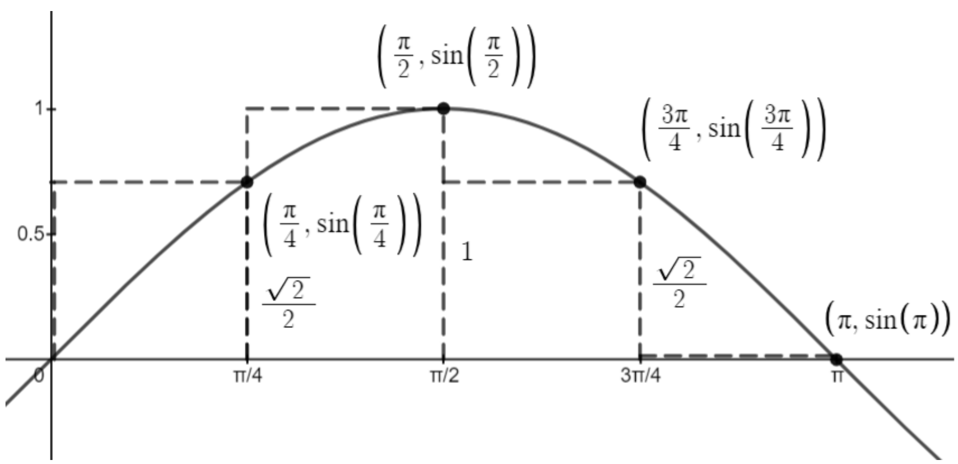
mate if we can further triangularize the leftover areas. More triangles are used, better the estimate will be.

Although there are many algorithms for triangularizing a shape out there, it is still far from being simple enough to be formulated as a formula. The next shape we can consider to estimate the area is a quadrilateral (four-sided shape). Especially a rectangle. Its area is simple.



We can consider two rectangles. The most obvious rectangle is the one that has the region inscribed, and its area is $(\pi)(1) = \pi$. The other one is the rectangle with the maximum area that is inscribed in the region, and its area is $2(0.86)(\sin(\frac{\pi}{2} + 0.86)) = 1.1222$. One is an overestimate, and the other is a underestimate.

Last Suggestion: How about using more than one rectangle? We start with four rectangles.



We first partition the real number line from $x = 0$ to $x = \pi$ into four equal lengths, so the markers of the partition are $x = 0$, $x = \frac{\pi}{4}$, $x = 2 \cdot \frac{\pi}{4} = \frac{\pi}{2}$, $x = 3 \cdot \frac{\pi}{4} = \frac{3\pi}{4}$, and $x = 4 \cdot \frac{\pi}{4} = \pi$. Above the closed intervals $[0, \frac{\pi}{4}]$, $[\frac{\pi}{4}, \frac{\pi}{2}]$, $[\frac{\pi}{2}, \frac{3\pi}{4}]$, and $[\frac{3\pi}{4}, \pi]$, imagine four rectangles drawn so that the portion underneath the curve of $\sin(x)$ over the closed interval $[0, \pi]$ can be estimated.

Q: How do we determine the height of a rectangle?

There could be infinitely many ways to determine the height of each rectangle, but for the sake of "consistency," we use the right-end number of each closed interval to determine the height of the rectangle. For instance, for the closed interval $[0, \frac{\pi}{4}]$, the right-end number is $\frac{\pi}{4}$. Then we draw a rectangle whose top is drawn passing through the point $(\frac{\pi}{4}, \sin(\frac{\pi}{4}))$ over the closed interval $[0, \frac{\pi}{4}]$. The area of such a rectangle is $(\frac{\pi}{4} - 0)\sin(\frac{\pi}{4}) = \frac{\pi}{4} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}\pi}{8}$. For the rest of the intervals, if we use the right-end number to determine the height, then

$$\text{Area} \approx \left(\frac{\pi}{4} - 0\right)\sin\left(\frac{\pi}{4}\right) + \left(\frac{\pi}{2} - \frac{\pi}{4}\right)\sin\left(\frac{\pi}{2}\right) + \left(\frac{3\pi}{4} - \frac{\pi}{2}\right)\sin\left(\frac{3\pi}{4}\right) + \left(\pi - \frac{3\pi}{4}\right)\sin(\pi)$$

Since we divide the closed interval $[0, \pi]$ into four equal lengths, the base of the rectangle should be all $\frac{\pi}{4}$. Then

$$\begin{aligned}
\text{Area} &\approx \frac{\pi}{4} \cdot \sin\left(\frac{\pi}{4}\right) + \frac{\pi}{4} \cdot \sin\left(\frac{\pi}{2}\right) + \frac{\pi}{4} \cdot \sin\left(\frac{3\pi}{4}\right) + \frac{\pi}{4} \cdot \sin(\pi) \\
&= \frac{\pi}{4}\left(\sin\left(\frac{\pi}{4}\right) + \sin\left(\frac{\pi}{2}\right) + \sin\left(\frac{3\pi}{4}\right) + \sin(\pi)\right) \\
&= \frac{\pi}{4}\left(\frac{\sqrt{2}}{2} + 1 + \frac{\sqrt{2}}{2} + 0\right) \\
&= \frac{\pi(1+\sqrt{2})}{4} \quad (\approx 1.8961)
\end{aligned}$$

The first two rectangles overestimated the region, but last two rectangles underestimated. Just from an inspection of the graph, it seems that we underestimated the region overall.

Q: How can we improve the estimate?

One can imagine that the portion being overestimated and underestimated can be reduced if we use more rectangles. So let us use n rectangles.

Markers: First, we need to partition the closed interval $[0, \pi]$ into n equal closed intervals. The first marker will be 0, and let us use the notation $x_0 = 0$ (initial number). Then the second marker will be $x_1 = \frac{\pi}{n}$. The next one is $x_2 = \frac{\pi}{n} + \frac{\pi}{n} = 2 \cdot \frac{\pi}{n}$. The next one is $x_3 = \frac{2\pi}{n} + \frac{\pi}{n} = 3 \cdot \frac{\pi}{n}$. In this way, the very last one would be $x_n = n \cdot \frac{\pi}{n} = \pi$. For instance, if we use $n = 16$ rectangles, the markers are $x_0 = 0$, $x_1 = \frac{\pi}{16}$, $x_2 = \frac{2\pi}{16}$, $x_3 = \frac{3\pi}{16}$, …, and $x_{16} = \frac{16\pi}{16} = \pi$.

Intervals: There are n closed intervals starting from $[x_0, x_1]$, $[x_1, x_2]$, $[x_2, x_3]$, …, $[x_{n-1}, x_n]$ or $[0, \frac{\pi}{n}]$, $[\frac{\pi}{n}, \frac{2\pi}{n}]$, $[\frac{2\pi}{n}, \frac{3\pi}{n}]$, …, $[\frac{(n-1)\pi}{n}, \frac{n\pi}{n}]$.

Base: The base of the rectangle should be all equal to $\frac{\pi}{n}$.

Heights: If we use the right-end number from each closed interval, the heights are given by $\sin(x_1)$, $\sin(x_2)$, $\sin(x_3)$, ..., $\sin(x_n)$ or

$$\sin\left(\frac{\pi}{n}\right), \sin\left(\frac{2\pi}{n}\right), \sin\left(\frac{3\pi}{n}\right), \ldots, \sin\left(\frac{n\pi}{n}\right)$$

Area: Finally the area would be

$$R_n = \frac{\pi}{n}\left(\sin\left(\frac{\pi}{n}\right) + \sin\left(\frac{2\pi}{n}\right) + \sin\left(\frac{3\pi}{n}\right) + \cdots + \sin\left(\frac{n\pi}{n}\right)\right)$$

The letter R stands for the "right-end," and n represents the number of the rectangles. Before, we found out that

$$R_4 = \frac{\pi(1+\sqrt{2})}{4} \approx 1.8961 \quad \text{and} \quad R_6 = \frac{\pi(2+\sqrt{3})}{6} \approx 1.9541$$

Here are some more of estimates for different n values with a help from WolframAlpha:

n	R_n	WolframAlpha command	Symbol from WolframAlpha
10	1.98352	(Pi/10)*Sum[Sin[k*(Pi/10)],k,1,10]	$\frac{\pi}{10}\sum_{k=1}^{10}\sin(k\cdot\frac{\pi}{10})$
100	1.99984	(Pi/100)*Sum[Sin[k*(Pi/100)],k,1,100]	$\frac{\pi}{100}\sum_{k=1}^{100}\sin(k\cdot\frac{\pi}{100})$
500	1.99999	(Pi/500)*Sum[Sin[k*(Pi/500)],k,1,500]	$\frac{\pi}{500}\sum_{k=1}^{500}\sin(k\cdot\frac{\pi}{500})$

Q: Any guess on what this estimate will be close to as $n \to \infty$?

Sigma Notation

Instead of writing

$$R_{10} = \frac{\pi}{10}\left(\sin\left(\frac{\pi}{10}\right) + \sin\left(2\cdot\frac{\pi}{10}\right) + \sin\left(3\cdot\frac{\pi}{10}\right) + \cdots + \sin\left(10\cdot\frac{\pi}{10}\right)\right)$$

we can write using Σ (capital letter 'sigma' in Greek alphabet) notation.

$$R_{10} = \frac{\pi}{10}\boxed{\sum_{k=1}^{10}\sin\left(k\cdot\frac{\pi}{10}\right)}$$

It is a succinct way of writing a sum of many expressions that has some patterns that can be described in a nice expression. The sigma notation is written as

$$\sum_{\text{running index} = \text{beginning index value}}^{\text{ending index value}} \quad \text{expression usually involving the running index}$$

Then the notation represents the sum of the expression by evaluating the expression with the running index starting from the beginning index value to the ending index value "incrementing by 1." For instance,

$$\sum_{k=1}^{10} k = 1+2+3+4+5+6+7+8+9+10$$

$$\sum_{k=1}^{10}(2k+1) = (2\cdot 1+1)+(2\cdot 2+1)+(2\cdot 3+1)+\cdots+(2\cdot 10+1)$$

$$\sum_{k=6}^{10} \pi = \pi+\pi+\pi+\pi+\pi$$

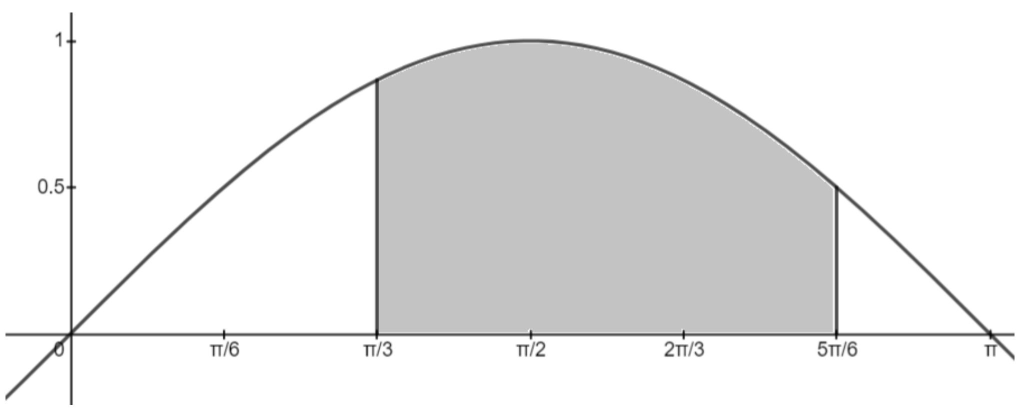
Then

$$R_n = \frac{\pi}{n}\sum_{k=1}^{n} \sin\left(k\cdot\frac{\pi}{n}\right)$$

The area under the curve of $y=\sin(x)$ over the closed interval $[0,\pi]$ is best estimated by

$$\lim_{n\to\infty} R_n = \lim_{n\to\infty}\frac{\pi}{n}\sum_{k=1}^{n} \sin\left(k\cdot\frac{\pi}{n}\right) \text{ which we suspected to be 2}$$

Example 1 Find the area under the curve of $y=\sin(x)$ over the closed interval $[\frac{\pi}{3},\frac{5\pi}{6}]$.



Let us try $n=10$ rectangles. The equal width of the interval is obtained by

$$\Delta x = \frac{\text{length of the closed interval}}{n} = \frac{\frac{5\pi}{6}-\frac{\pi}{3}}{10} = \frac{\pi}{20}$$

Markers: The first marker is $x_0=\frac{\pi}{3}$ (not zero any longer). The second marker is $x_1=\frac{\pi}{3}+\frac{\pi}{20}$. The next marker is $x_2=\frac{\pi}{3}+2\cdot\frac{\pi}{20}$. The rest are

$$x_3=\frac{\pi}{3}+3\cdot\frac{\pi}{20},\quad x_4=\frac{\pi}{3}+4\cdot\frac{\pi}{20},\quad \ldots,\quad x_9=\frac{\pi}{3}+9\cdot\frac{\pi}{20},\quad x_{10}=\frac{\pi}{3}+10\cdot\frac{\pi}{20}$$

Intervals: The closed intervals are $[x_{k-1}, x_k]$ where $k = 1, 2, \ldots, 10$ or
$[\frac{\pi}{3} + (k-1) \cdot \frac{\pi}{20}, \frac{\pi}{3} + k \cdot \frac{\pi}{20}]$ where $k = 1, 2, \ldots, 10$.

Base: $\Delta x = \frac{\pi}{20}$.

Heights: Using the right-end number, the height is $\sin(\frac{\pi}{3} + k \cdot \frac{\pi}{20})$ where $k = 1, 2, \ldots, 10$.

Area: The area is estimated using

$$R_{10} = \frac{\pi}{20} \sum_{k=1}^{10} \sin\left(\frac{\pi}{3} + k \cdot \frac{\pi}{20}\right) \approx 1.33447$$

Right-End Rectangle (RER)

There is nothing special about $\frac{\pi}{3}$, $\frac{5\pi}{6}$, $\sin(x)$, or $n = 10$. Suppose that we want to estimate the area under the curve of $y = f(x)$ over the closed interval $[a, b]$ using n rectangles.

Base: The equal length of the closed intervals is

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

Markers: $x_0 = a$, $x_1 = a + \Delta x$, $x_2 = a + 2 \cdot \Delta x$, …, $x_n = a + n \cdot \Delta x = b$. Or

$$\boxed{x_k = a + k \cdot \Delta x, \text{ where } k = 1, 2, \ldots, n}$$

Heights: $f(x_1)$, $f(x_2)$, $f(x_3)$, …, $f(x_n)$. Or

$$\boxed{f(x_k) = f(a + k \cdot \Delta x)}$$

Area: The area is estimated using

$$\boxed{R_n = \Delta x \sum_{k=1}^{n} f(x_k) = \Delta x \sum_{k=1}^{n} f(a + k \cdot \Delta x)}$$

This is called the **right-end rectangle** estimate R_n.

Example 2 Estimate the area under the curve $f(x) = x^2$ over the closed interval $[2, 5]$. Use $n = 6$ rectangles.

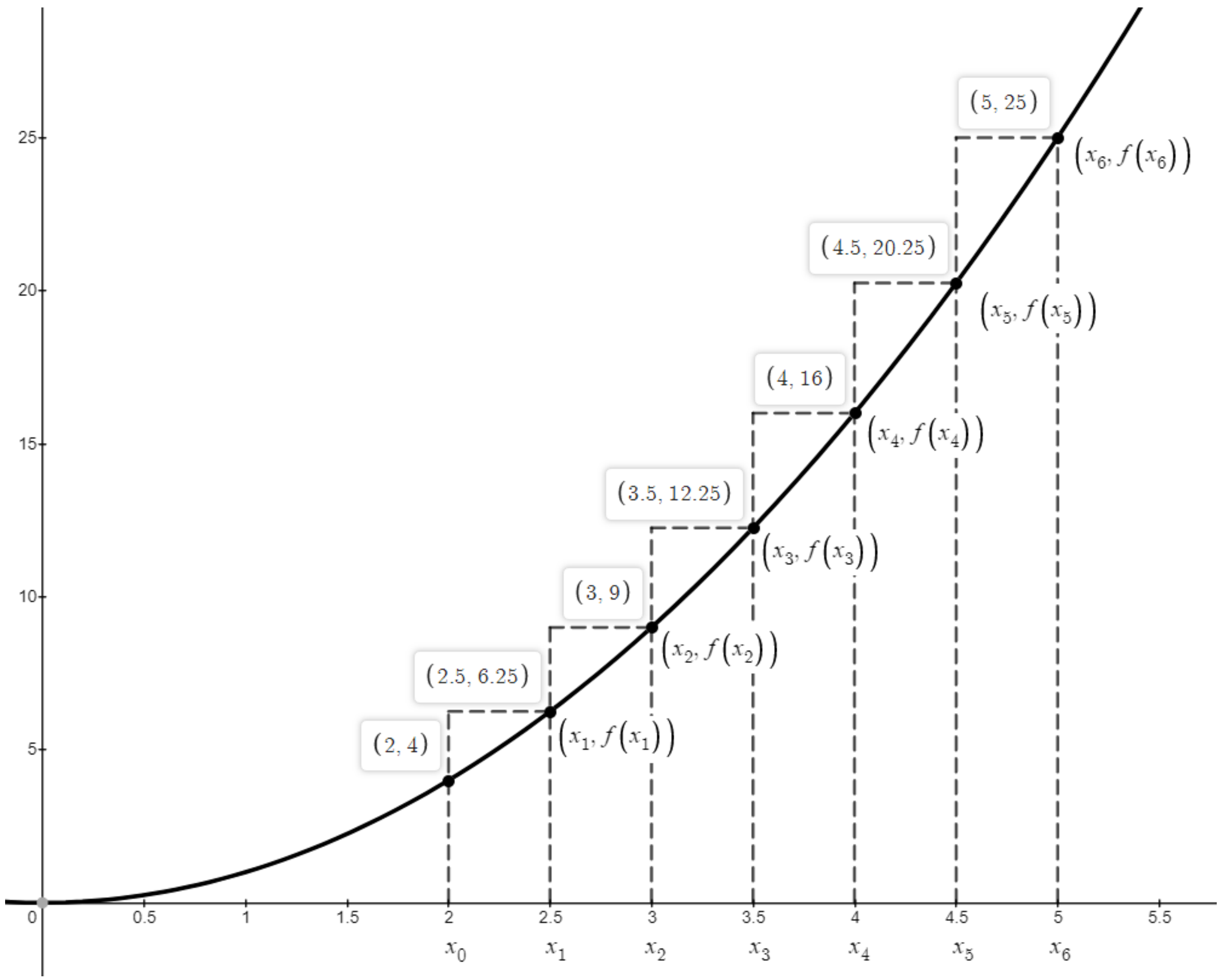
Base: The equal length of the closed intervals is $\Delta x = \frac{5-2}{6} = \frac{1}{2}$

Markers: $x_0 = 2$, $x_1 = 2 + \frac{1}{2}$, …, $x_6 = 2 + 6 \cdot \frac{1}{2} = 5$.
Or $x_k = 2 + k \cdot \frac{1}{2} = 2 + \frac{k}{2}$ where $k = 1, 2, 3, 4, 5, 6$.

Heights: $f(x_k) = (x_k)^2 = (2 + \frac{k}{2})^2$

Area: The area is estimated using the right-end rectangle estimate

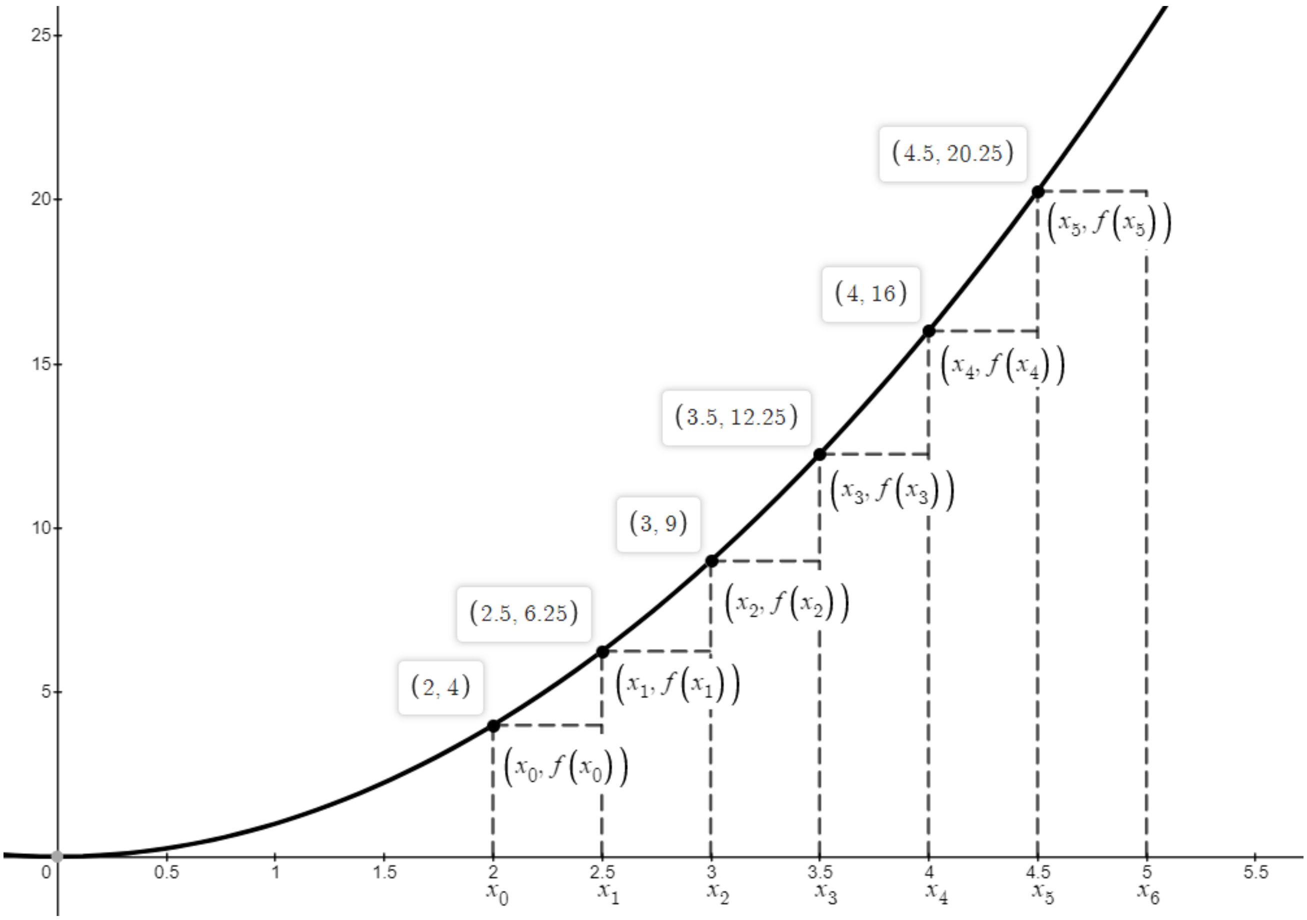
$$\begin{aligned} R_6 &= \frac{1}{2}\sum_{k=1}^{6}\left(2+\frac{k}{2}\right)^2 \\ &= \frac{1}{2}\left(\left(2+\frac{1}{2}\right)^2+\left(2+\frac{2}{2}\right)^2+\left(2+\frac{3}{2}\right)^2+\left(2+\frac{4}{2}\right)^2+\left(2+\frac{5}{2}\right)^2+\left(2+\frac{6}{2}\right)^2\right) \\ &= \frac{355}{8} \quad = \quad 44.375 \end{aligned}$$



Apparently this estimate is overestimation of the actual area.

Left-End Rectangle (LER)

Estimating the area using the left-end numbers would not be too different from using the right-end numbers. Here is the what the picture would look like for the area we estimated using the left-end rectangles.



The only difference is that it is missing the very last rectangle from RER and a new rectangle is introduced at the beginning. In the summation notation, the left-end rectangle method would estimate the area

$$L_6 = \frac{1}{2}\sum_{k=0}^{5}\left(2+\frac{k}{2}\right)^2$$

Look how the expressions are exactly same as the RER method except the beginning index value and the ending index value are changed from $k=1$ through 6 to $k=0$ through 5. From the picture, we can see that this estimation would be underestimation, so the computed value would be smaller than the result of the RER method.

$$\begin{aligned} L_6 &= \frac{1}{2}\sum_{k=0}^{5}\left(2+\frac{k}{2}\right)^2 \\ &= \frac{1}{2}\left(\left(2+\frac{0}{2}\right)^2+\left(2+\frac{1}{2}\right)^2+\left(2+\frac{2}{2}\right)^2+\left(2+\frac{3}{2}\right)^2+\left(2+\frac{4}{2}\right)^2+\left(2+\frac{5}{2}\right)^2\right) \\ &= \frac{271}{8} \quad = \quad 33.875 \end{aligned}$$

Only way to close this gap between the RER and the LER estimates is to consider larger n. Why not, for now, just finding them using n without specifying the number of rectangles.

Example 3 Estimate the area under the curve $f(x) = x^2$ over the closed interval $[2, 5]$. Use n rectangles.

We start with the RER estimate.

Base: The equal length of the closed intervals is $\Delta x = \frac{5-2}{n} = \frac{3}{n}$

Markers: $x_0 = 2$, $x_1 = 2 + \frac{3}{n}$, ..., $x_n = 2 + n \cdot \frac{3}{n} = 5$.
Or $x_k = 2 + k \cdot \frac{3}{n} = 2 + \frac{3k}{n}$ where $k = 1, 2, \ldots, n$.

Heights: $f(x_k) = (x_k)^2 = (2 + \frac{3k}{n})^2$

Area: The area is estimated using the RER and LER methods

$$R_n = \frac{3}{n}\sum_{k=1}^{n}\left(2+\frac{3k}{n}\right)^2 \quad \text{and} \quad L_n = \frac{3}{n}\sum_{k=0}^{n-1}\left(2+\frac{3k}{n}\right)^2$$

We would like to believe that the actual area can be found by

$$\lim_{n\to\infty} R_n = \lim_{n\to\infty}\frac{3}{n}\sum_{k=1}^{n}\left(2+\frac{3k}{n}\right)^2$$

or

$$\lim_{n\to\infty} L_n = \lim_{n\to\infty}\frac{3}{n}\sum_{k=0}^{n-1}\left(2+\frac{3k}{n}\right)^2$$

For these limits to be coinciding as a single value, we need to see that their difference (the gap between the RER and the LER estimates) vanishes to zero.

$$R_n - L_n = \frac{3}{n}\sum_{k=1}^{n}\left(2+\frac{3k}{n}\right)^2 - \frac{3}{n}\sum_{k=0}^{n-1}\left(2+\frac{3k}{n}\right)^2$$

is just nothing but the last rectangle of the RER estimate minus the first rectangle of the LER estimate, i.e.

$$R_n - L_n = \frac{3}{n}(5)^2 - \frac{3}{n}(2)^2 = \frac{3(5^2 - 2^2)}{n}$$

and its limit as $n \to \infty$ is

$$\lim_{n\to\infty}(R_n - L_n) = \lim_{n\to\infty}\frac{63}{n} = 0$$

In general, when estimating the area using the RER and the LER methods, we know that

$$\lim_{n\to\infty}(R_n - L_n) = \lim_{n\to\infty}\frac{b-a}{n}(b^2 - a^2) = \lim_{n\to\infty}\frac{(b-a)(b^2-a^2)}{n} = 0$$

It is far from an actual rigorous proof why the area can be found by taking the limit of these RER or LER estimates, it is okay for now.

In fact, the RER and LER estimates are just examples of more general estimate called the **Riemann sum**.

Riemann Sum

Let $f(x)$ be a continuous function on the closed interval $[a, b]$. We will assume that $f(x) \geq 0$ on $[a, b]$ for now, i.e. the portion of the curve of $f(x)$ on $[a, b]$ is in fact above the x-axis.

Closed intervals: Partition the closed interval $[a, b]$ into n closed intervals: $[x_0, x_1]$, $[x_1, x_2]$, …, $[x_{n-1}, x_n]$. Here, $x_0 = a$ and $x_n = b$. It is not required to partition into intervals with equal length, but for this course, we will assume that. That is, $x_k - x_{k-1} = \frac{b-a}{n} = \Delta x$ is constant where $k = 1, 2, \ldots, n$.

Sample points: On a closed interval $[x_{k-1}, x_k]$, we let x_k^* be any number in the interval. In the case of the RER, we let $x_k^* = x_k$ (right-end number). For the LER, we let $x_k^* = x_{k-1}$ (left-end number). But x_k^* can be any number as long as it is picked from $[x_{k-1}, x_k]$.

Heights: On a closed interval $[x_{k-1}, x_k]$, we use $f(x_k^*)$ as the height of the rectangle whose base is the closed interval. So its area is $\Delta x \cdot f(x_k^*)$.

Area: The area under the curve of $f(x)$ over the closed interval $[a, b]$ is estimated by

$$\Delta x \sum_{k=1}^{n} f(x_k^*)$$

We will find out in later section that this sum called the **Riemann sum** gives the actual area under the curve of $f(x)$ over the closed interval $[a, b]$

$$\text{Area} = \lim_{n\to\infty} \Delta x \sum_{k=1}^{n} f(x_k^*)$$

as $n \to \infty$.

For sample points, if x_k^* is chosen on the closed interval $[x_{k-1}, x_k]$ so that $f(x_k^*)$ is the absolute maximum, then the Riemann sum is called the **upper Riemann sum**. If x_k^* is chosen on the closed interval $[x_{k-1}, x_k]$ so that $f(x_k^*)$ is the absolute minimum, then the Riemann sum is called the **lower Riemann sum**. We know that the absolute extrema exist as f is assumed to be continuous.

In the last example, as f is always increasing, the RER estimate gave us the upper Riemann sum and the LER estimate gave us the lower Riemann sum. However, this may not be true for general class of functions.

The Distance Problem

Q: Why bother with the area under the curve?

Recall a formula from Pre-Algebra class.

$$\text{Distance} = \text{Rate} \times \text{Time}$$

The same formula still applies in Calculus, but we need to make little adjustment. First, the rate will not be just some speed of a car, but it will be a velocity, given usually as a function $v(t)$ as a function of time t. Since it is velocity, its sign matters. Time t is still the elapsed time from $t = 0$.

The distance is now a **displacement**. It is a relative position of a moving body with respect to the initial position of the body. For instance, if a body moved towards the positive direction 3 feet away from its initial position, then its displacement is 3 feet. However, if a body moved towards the positive direction 3 feet ways from its initial position and came back (moving towards the negative direction) 3 feet. Then its displacement is 0 feet. Then the formula becomes

$$\text{Displacement} = \text{Velocity} \times \text{Time}$$

The area under the velocity function $v(t)$ over the closed interval $[a, b]$ is then the displacement of the moving body from $t = a$ to $t = b$.

Example 4 A rocket is launched at a launch station 5 meters above the ground. Its velocity is given by $v(t) = 90t$ (the acceleration boosts up the velocity by 90 m/s every second). Describe what happens in the first 5 seconds. Since the velocity is a nice linear function, we can easily find the area under the curve. No need to estimate. The area under

$v(t)$ over the closed interval $[0, 5]$ is just the area of the triangle with the base $5 - 0 = 5$ and the height $v(5) = 450$. Then the area is $\frac{1}{2}(5)(450) = 1125$ meters. Does this mean the rocket is 1125 meters above the ground? No. The area under the velocity over time interval gives us a "displacement," which is a relative position with respect to the position of the rocket at $t = 0$, which is 5 meters above the ground. Hence, the absolute position of the rocket at $t = 5$ is $5 + 1125 = 1130$ meters above the ground.

If the rocket was launched from a underground silo that is 200 meters below the ground level with the same velocity, then the absolute position of the rocket would be $-200 + 1125 = 925$ meters above the ground at $t = 5$.

The example also illustrates that only knowing velocity does not tell us about the absolute position of the rocket.

Example 5 A rocket is launched at a launch station 5 meters above the ground. Its velocity is given by $v(t) = 90t - 3t^2$ (some sort of force is pushing the rocket from above). Describe what happens in the first 10 seconds, 15 seconds, and 30 seconds.

We need to estimate the area under the curve of $v(t) = 90t - 3t^2$ over the closed interval $[0, 10]$. Let us use the RER with 10 rectangles.

$$R_{10} = \frac{10-0}{10}\sum_{k=1}^{10}(90(0+k) - 3(0+k)^2) = 3795$$

So the rocket's absolute position at $t = 10$ is $5 + 3795 = 3800$ meters above the ground.

On the interval $[0, 15]$,

$$R_{10} = \frac{15-0}{10}\sum_{k=1}^{10}\left(90\left(0+\frac{3k}{2}\right) - 3\left(0+\frac{3k}{2}\right)^2\right) = 7239.375$$

So the rocket's absolute position at $t = 15$ is $5 + 7239.375 = 7244.375$ meters above the ground.

On the interval $[0, 30]$,

$$R_{10} = \frac{30-0}{10}\sum_{k=1}^{10}(90(0+3k) - 3(0+3k)^2) = 13365$$

So the rocket's absolute position at $t = 30$ is $5 + 13365 = 13370$ meters above the ground.

Assigned Exercises: (p 375) 1 - 7 (odds), 13, 15, 17, 21, 23