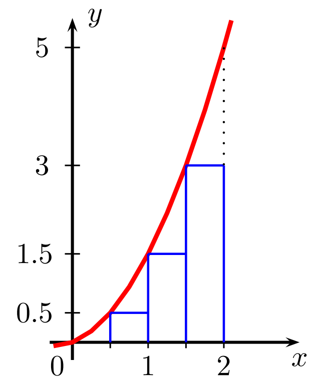
MA123, Chapter 11: Estimating Definite Integrals

Chapter Goals:	• Estimate the value of a definite integral using left endpoints, right endpoints, midpoints, or trapezoids. • Understand the formal definition of the definite integral. (Optional)
Assignments:	Assignment 20 Assignment 21

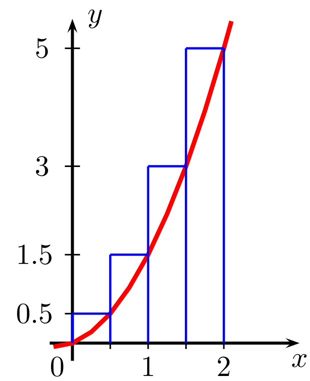
General philosophy: In this chapter, we will learn to estimate the definite integral for functions where we cannot use geometry to compute the areas. *The key idea is to notice that the value of the function does not vary very much over a small interval, and so it is approximately constant over a small interval.* We will use the areas of particular rectangles or trapezoids to estimate the integrals. One particular type of estimation using rectangles is called a Riemann sum.

Example 1: Estimate the area under the graph of $y = x^2 + \frac{1}{2}x$ for x between 0 and 2 in two different ways:

(a) Subdivide the interval $[0, 2]$ into four equal subintervals and use the left endpoint of each subinterval as "sample point."



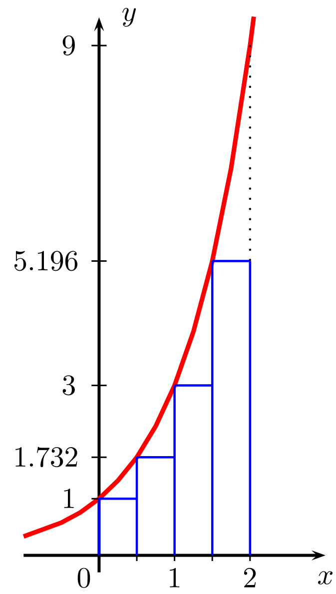
(b) Subdivide the interval $[0, 2]$ into four equal subintervals and use the right endpoint of each subinterval as "sample point."



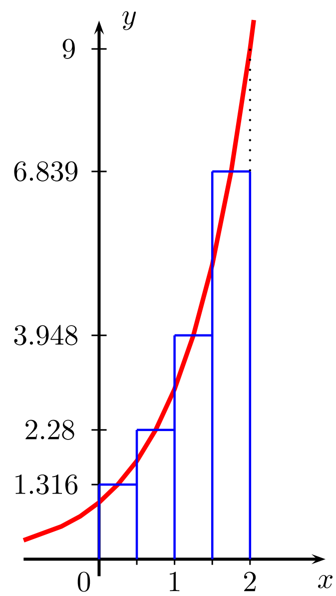
Dividing $[0, 2]$ into four equal pieces is an example of a **partition** of the interval $[0, 2]$. Most of our partitions will use equal-width subdivisions, though that is not required.

(c) Find the difference between the two estimates (right endpoint estimate minus left endpoint estimate).

Example 2: Estimate the area under the graph of $y = 3^x$ for x between 0 and 2.



(a) Use a partition that consists of four equal subintervals of $[0, 2]$, and use the left endpoint of each subinterval as a sample point.



(b) Use a partition that consists of four equal subintervals of $[0, 2]$, and use the midpoint of each subinterval as a sample point.

Note:

In the previous two examples we systematically chose the value of the function at a particular point of each subinterval. However, since the guiding idea is that we are assuming that the values of the function over a small subinterval do not change by very much, then we could take the value of the function at <u>any</u> point of the subinterval as a good sample or representative value of the function. We could also have chosen small subintervals of different lengths. However, we are trying to establish a systematic procedure that works well in general.

Getting better estimates:

We can only expect the previous answers to be approximations of the correct answers. This is because the values of the function do change on each subinterval, even though they do not change by much.

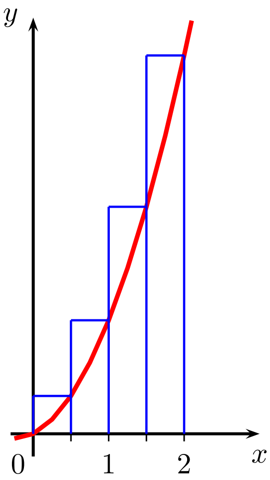
If, however, we replace the subintervals we used by "smaller" subintervals we can reasonably expect the values of the function to vary by much less on each thinner subinterval. Thus, we can reasonably expect that the area of each thinner vertical strip under the graph of the function to be more accurately approximated by the area of these thinner rectangles. Then if we add up the areas of all these thinner rectangles, we should get a much more accurate estimate for the area of the original region.

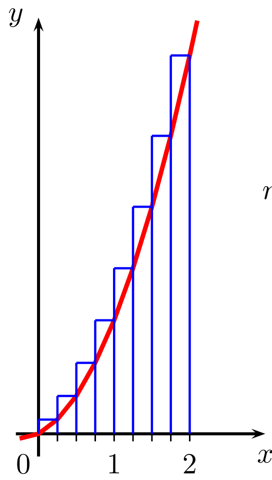
Here is Example 1(b), revisited:

$$y = x^2 + \frac{1}{2}x \quad \text{on} \quad [0, 2]$$

$n = 4$ equal subintervals

Area ≈ 5





$$y = x^2 + \frac{1}{2}x \quad \text{on} \quad [0, 2]$$

$n = 8$ equal subintervals

Area ≈ 4.3125

We will see later that the exact value of the area under consideration in Example 1 is $\dfrac{11}{3} \approx 3.66$.

Example 3: We could estimate the area under the graph of $y = \dfrac{1}{x}$ for x between 1 and 31 by dividing the interval $[1, 31]$ into 30 equal subintervals and using the left endpoint of each subinterval as sample point. Next, we could estimate the area using the right endpoint as sample point.
Find the difference between the two estimates (left endpoint estimate minus right endpoint estimate).

Example 4: Suppose you estimate the area under the graph of $f(x) = x^3$ from $x = 4$ to $x = 24$ by adding the areas of rectangles as follows: partition the interval into 20 equal subintervals and use the right endpoint of each interval to determine the height of the rectangle. What is the area of the 15^{th} rectangle?

Example 5: Suppose you estimate the area under the graph of $f(x) = \frac{1}{x}$ from $x = 12$ to $x = 112$ by adding the areas of rectangles as follows: partition the interval into 50 equal subintervals and use the left endpoint of each interval to determine the height of the rectangle. What is the area of the 24^{th} rectangle?

Example 6: An object travels in a straight line, and we would like to estimate how far the object traveled during the time interval $0 \le t \le 5$, but we only have the following information about the velocity of the object:

time (sec)	0	1	2	3	4	5
velocity (m/sec)	-3	-1	-4	1	3	6

(a) Using left endpoints as sample points, estimate the object's total displacement over the time interval.

(b) Using left endpoints as sample points, estimate the object's total distance traveled over the time interval.

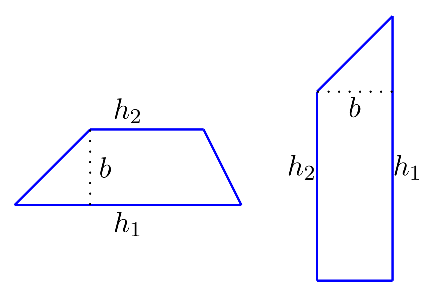
Example 7: The rate (in liters per minute) at which water drains from a tank is recorded at two-minute intervals. Use the average of the left- and right-endpoint approximations to estimate the total amount of water drained during the first 6 min.

t min	0	2	4	6
l/min	48	46	44	42

Trapezoids versus rectangles:

We recognize that the average of the left and right estimates is more accurate than either individually. Instead of computing each separately, we can achieve this in one calculation. We do this by using trapezoids instead of rectangles. Recall that the area of a trapezoid is

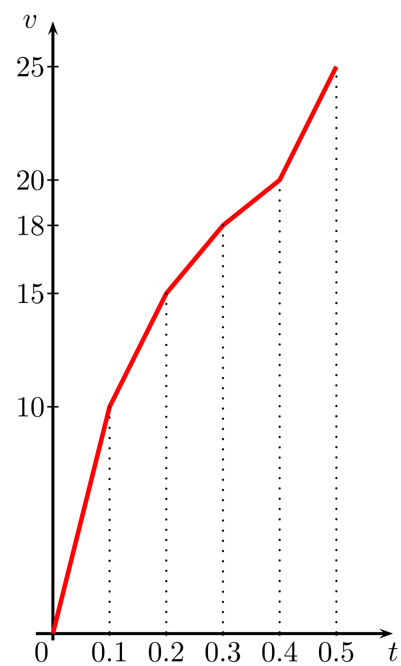
$$\text{Area of a trapezoid} = \frac{(h_1 + h_2) \cdot b}{2}.$$



Example 8: A train travels in a straight westward direction along a track. The velocity of the train varies, but it is measured at regular time intervals of 1/10 hour. The measurements for the first half hour are

time	0	0.1	0.2	0.3	0.4	0.5
velocity	0	10	15	18	20	25

where the velocity in the table is given in miles per hour. Compute the total distance traveled by the train during the first half hour by assuming the velocity is a linear function of t on the subintervals.



Example 9: Suppose you are given the following data points for a function $f(x)$:

x	1	2	3	4
$f(x)$	2	5	8	12

If f is a linear function on each interval between the given points, find $\int_1^4 f(x)\,dx$.

⋆ THE FOLLOWING MATERIAL IS OPTIONAL ⋆

► **Sigma (Σ) notation:** In approximating areas we have encountered sums with many terms. A convenient way of writing such sums uses the Greek letter Σ (which corresponds to our capital S) and is called *sigma notation*. More precisely, if $a_1, a_2, \ldots, a_n$ are real numbers we denote the sum

$$a_1 + a_2 + \cdots + a_n$$

by using the notation

$$\sum_{k=1}^{n} a_k.$$

The integer k is called an *index* or *counter* and takes on (in this case) the values $1, 2, \ldots, n$.
For example,

$$\sum_{k=1}^{6} k^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 = 1 + 4 + 9 + 16 + 25 + 36 = 91$$

whereas

$$\sum_{k=3}^{6} k^2 = 3^2 + 4^2 + 5^2 + 6^2 = 9 + 16 + 25 + 36 = 86.$$

The idea we have used so far is to break up or subdivide the given interval $[a, b]$ into lots of little pieces, or subintervals, on each of which the variable x, and thus the function $f(x)$, does not change much. The technical phrase for doing this is to "partition" $[a, b]$.

Definition of a partition: A *partition* of an interval $[a, b]$ is a collection of points $\{x_0, x_1, x_2, \ldots, x_{n-1}, x_n\}$, listed increasingly, on the x-axis with $a = x_0$ and $x_n = b$. That is: $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$. These points subdivide the interval $[a, b]$ into n *subintervals*: $[a, x_1]$, $[x_1, x_2]$, $[x_2, x_3]$, $\ldots$, $[x_{n-1}, b]$.
The k-th subinterval is thus of the form $[x_{k-1}, x_k]$ and it has *length* $\Delta x_k = x_k - x_{k-1}$.

Assumption: Set $\|P\| = \max_{1\le i\le n} \{\Delta x_i\}$. We will always assume that our partition P is such that $\|P\| \to 0$ as $n \to \infty$. In other words, we always assume that the length of the longest (and as a consequence of all) subinterval(s) tend(s) to zero whenever the number of subintervals in our partition P becomes very large.

► **The definite integral:**
Let $f(x)$ be a function defined on an interval $[a, b]$. Partition the interval $[a, b]$ in n subintervals of lengths $\Delta x_1, \ldots, \Delta x_n$, respectively. For $k = 1, \ldots, n$ pick a representative point p_k in the corresponding k-th subinterval. The definite integral of the function f from a to b is defined as

$$\lim_{n\to\infty} \sum_{k=1}^{n} f(p_k) \cdot \Delta x_k = \lim_{n\to\infty} \Big(f(p_1) \cdot \Delta x_1 + f(p_2) \cdot \Delta x_2 + \cdots + f(p_n) \cdot \Delta x_n \Big) = \lim_{\|P\|\to 0} \sum_{k=1}^{n} f(p_k) \cdot \Delta x_k$$

and it is denoted by

$$\int_a^b f(x)\,dx.$$

The sum $\sum_{k=1}^{n} f(p_k) \cdot \Delta x_k$ is called a *Riemann sum* in honor of the German mathematician Bernhard Riemann (1826-1866), who developed the above ideas in full generality. The symbol $\int$ is called the *integral sign*. It is an elongated capital S, of the kind used in the 1600s and 1700s. The letter S stands for the summation performed in computing a Riemann sum. The numbers a and b are called the *lower and upper limits of integration*, respectively. The function $f(x)$ is called the *integrand* and the symbol dx is called the *differential* of x. You can think of the dx as representing what happens to the term Δx in the limit, as the size Δx of the subintervals gets closer and closer to 0.

Note: The role of x in a definite integral is the one of a *dummy variable*. In fact, $\int_a^b x^2\,dx$ and $\int_a^b t^2\,dt$ have the same meaning. They represent the same number.

Note: We recall from Chapter 3 that a limit does not necessarily exist. However:

Theorem: Let $f(x)$ be a continuous function on the interval $[a, b]$ then $\int_a^b f(x)\,dx$ exists. That is, the limit used in the definition of the definite integral exists.