Week 6: Defining and Estimating Integrals as Areas

Goals:

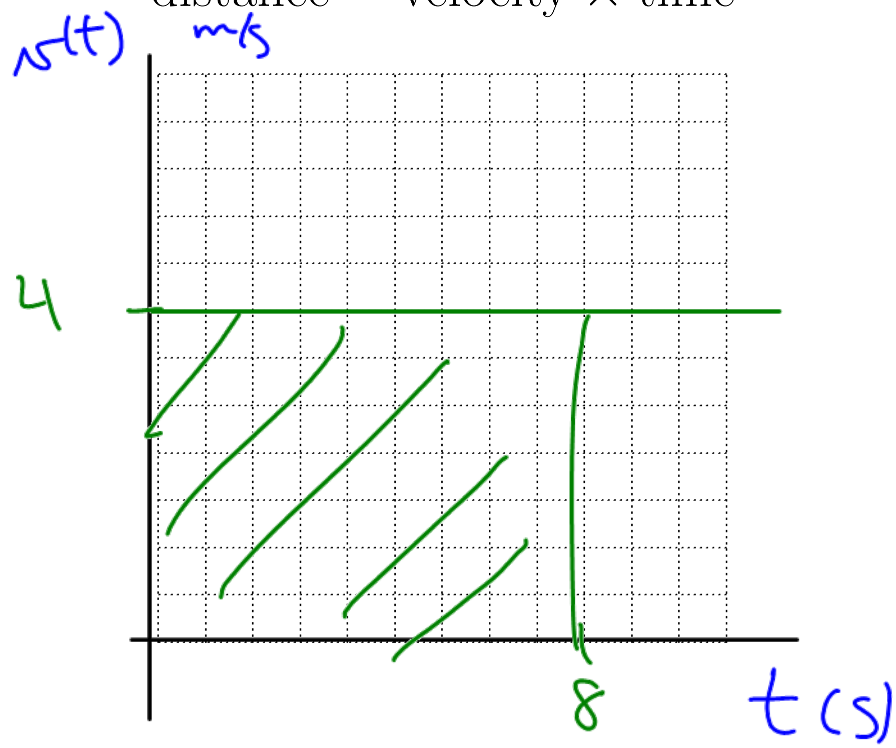
- Determine how to calculate the area described by a function.
- Define the **definite integral**.
- Explore the relationship between the definite integral and area.
- Identify properties of definite integrals
- Introduce the Fundamental Theorem of Calculus
- Compute simple anti-derivatives and definite integrals

Distance and Velocity

Recall that if we measure distance x as a function of time t, the velocity is determined by differentiating $x(t)$, i.e. finding the slope of the graph.
Alternatively, suppose we begin with a graph of the velocity with respect to time. How can we determine what distance will be traveled? Does it "appear" in the graph somehow?
Let's begin with the simple case of constant velocity...

$$v(t) = \frac{d}{dt} x(t)$$

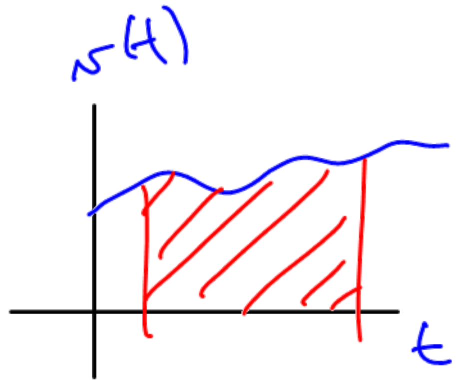
distance = velocity × time



dist $= (4\ m/s)(8\ s)$
$= 32\ m$

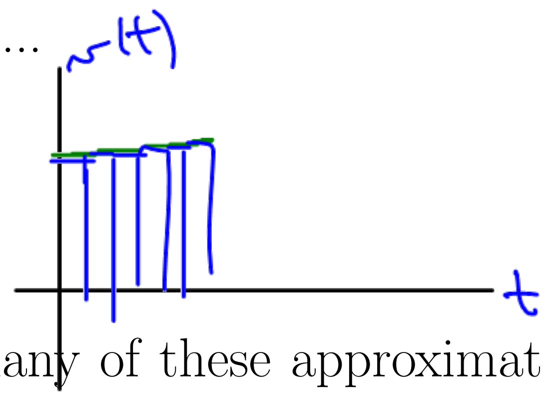
velocity is slope on position graph

For a constant velocity, the distance traveled in a certain length of time was simply the area of the rectangle underneath the velocity vs. time graph.



What if the velocity is changing?

We can't determine the exact distance traveled, but maybe we can *estimate* it. Let's assume that the velocity is not changing too quickly, so over a short amount of time it's roughly constant. We know how to find the distance traveled in that short time...



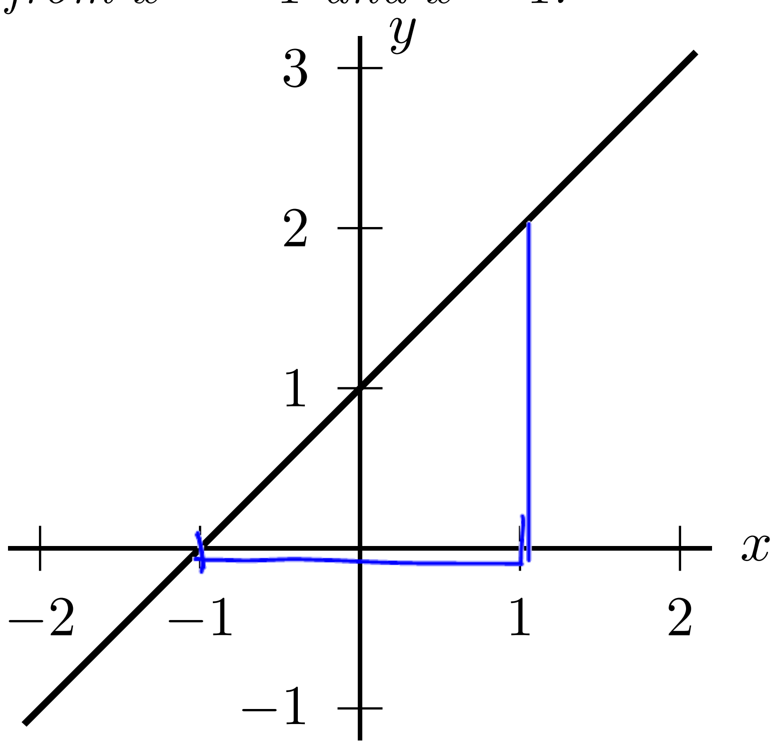
Making many of these approximations, we could come up with a rough estimate of the total distance. How does this estimate relate to the graph?

Calculating Areas

It appears that the **distance traveled** is the **area under the graph of velocity**, even when the velocity is changing. We'll see exactly why this is true very soon.

If we are simply interested in the area under a graph, without any physical interpretation, we can already do so if the graph creates a shape that we recognize.

Example: *Calculate the area between the x-axis and the graph of $y = x + 1$ from $x = -1$ and $x = 1$.*



area = area of triangle

$= \frac{1}{2}(\text{base})(\text{height})$

$= \frac{1}{2}(2)(2)$

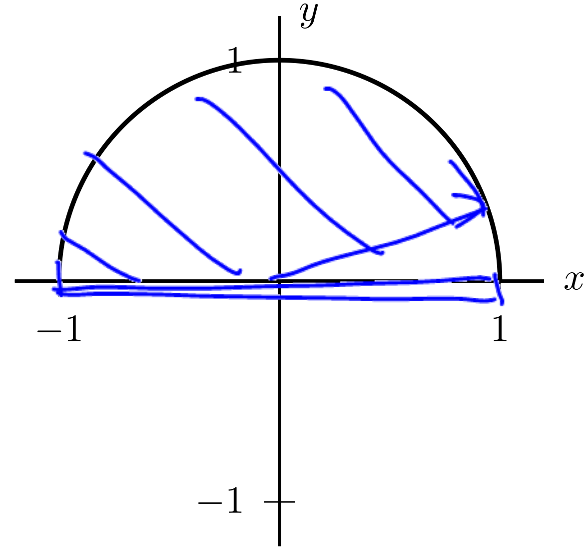
$= 2$ sq units.

Example: *Calculate the area between the x-axis and the graph of* $y = \sqrt{1-x^2}$ *from* $x = -1$ *to* $x = 1$.

$y = \sqrt{1-x^2} \iff y^2 = 1 - x^2$

or $x^2 + y^2 = 1$

circle w/ radius 1



$$\text{Area} = \frac{1}{2}(\text{circle area})$$

$$= \tfrac{1}{2}\left(\pi (1)^2\right)$$

$$= \pi/2 \text{ sq units}$$

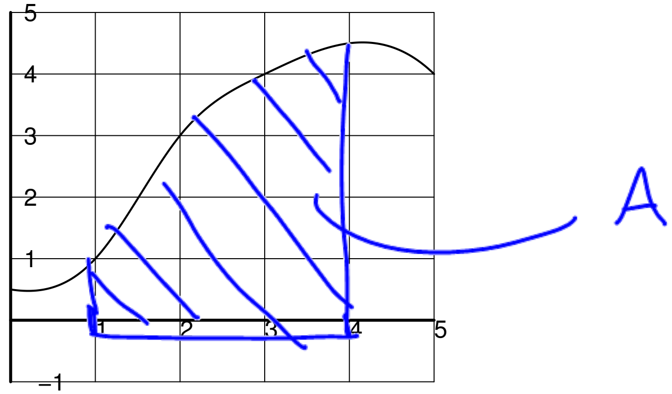
What shapes do you know, right now, for which you can calculate the exact area?

Estimating Areas

Unfortunately, many or most arbitrary areas are essentially impossible to find the area of when the shape isn't a simple composition of triangles, rectangles, or circles. In these cases, we must use less direct methods. We start by making an *estimate* of the area under the graph using shapes whose area is easier to calculate.

Suppose we are trying to find the area underneath the graph of the function $f(x)$ given below between $x = 1$ and $x = 4$.

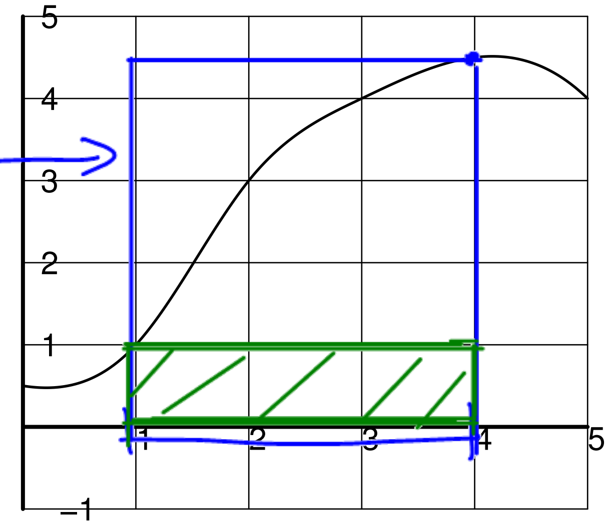
Draw this area on the graph below, and label it area A.



We can make a rough estimation of the area by drawing a rectangle that completely contains the area, or a rectangle that is completely contained by the area. *Calculate this overestimate and underestimate for the area A.*

$(3)(4.5) =$ 13.5 sq units

rectangle area $> A$ area ?



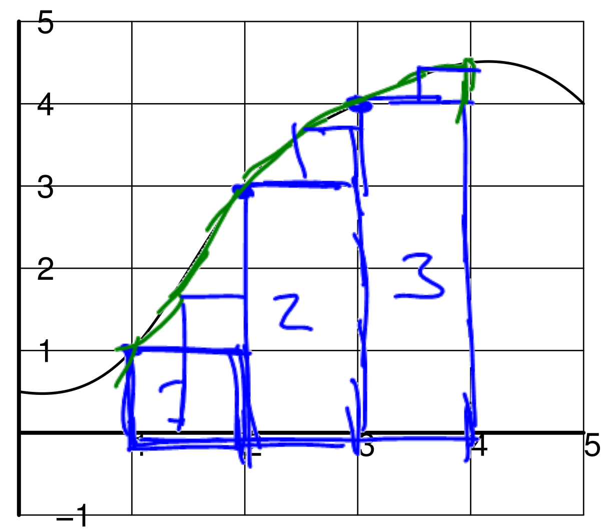
A area $>$ Rectangle area $= (3)(1)$

3 sq units

This one-rectangle estimate is very crude. We can improve our estimates by using smaller rectangles. E.g. We can divide the interval from $x = 1$ to $x = 4$ into **3 intervals**, each of width 1, and use different rectangle heights on each interval.

Estimate the area A by using 3 rectangles of width 1. Use the function value at the left *edge of the interval as the height of each rectangle.*

$$\begin{aligned} A \\ \text{Area} &\approx 3 \text{ rectangle areas} \\ &= (h_1)(w_1) + h_2 \cdot w_2 + h_3 w_3 \\ &= 1(1) + 3(1) + 4(1) \\ &= 1 + 3 + 4 \\ &= 8 \text{ sq units.} \end{aligned}$$

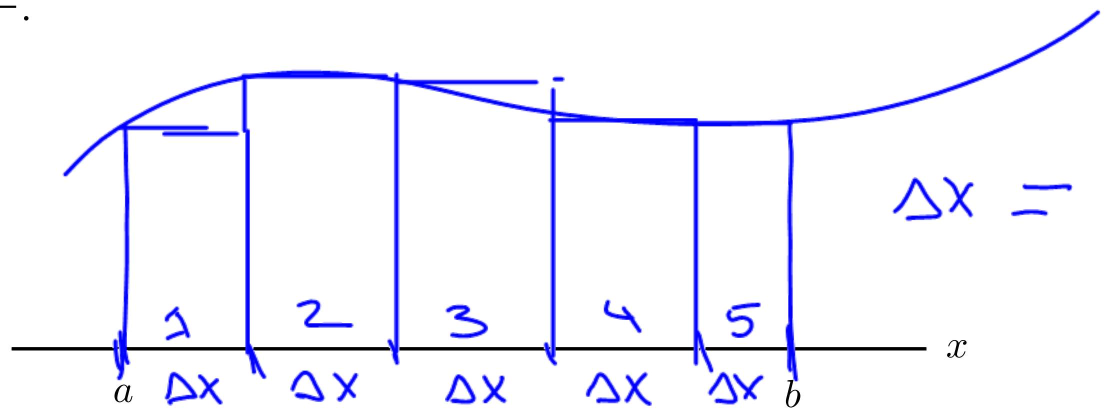


\# rectangles

Generalizing Area Estimates - $LEFT(n)$ and $RIGHT(n)$

We can repeat the rectangle-building process for any number of rectangles, and we expect that our estimation of the exact (curved) area will get better the more rectangles we use. The method we used earlier, choosing the height of the rectangles based on the function value at the left edge, is called the **left hand sum**, and is denoted *LEFT(n)* if we use n rectangles.

Suppose we are trying to estimate the area under the function $f(x)$ from $x = a$ to $x = b$ via the left hand sum with n rectangles. Then the width of each rectangle will be $\Delta x = \dfrac{b-a}{n}$.

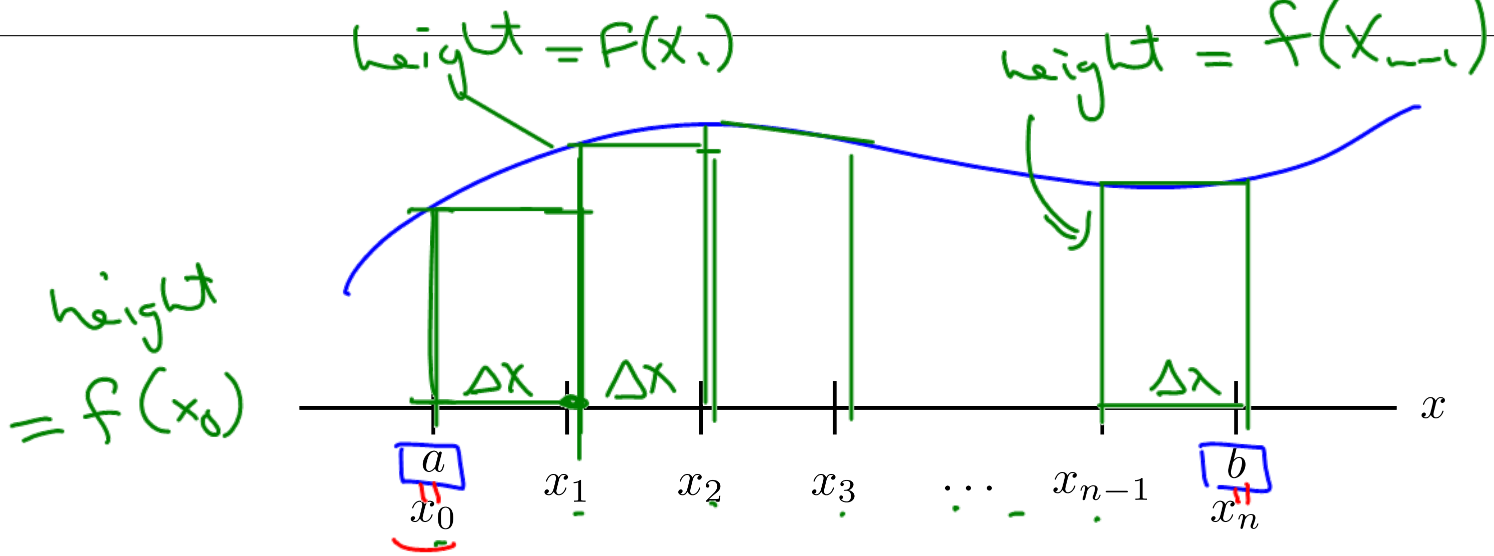


$\Delta x = \dfrac{b-a}{5}$ = width of rectangles

$n = 5$

If we label the endpoints of the intervals to be $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$, then the formula for the left hand sum will be

$$\begin{aligned} LEFT(n) = & f(x_0)\Delta x + \\ & f(x_1)\Delta x + \ldots + \\ & f(x_{n-1})\Delta x \\ & = \sum_{i=1}^{n} f(x_{i-1})\Delta x. \end{aligned}$$

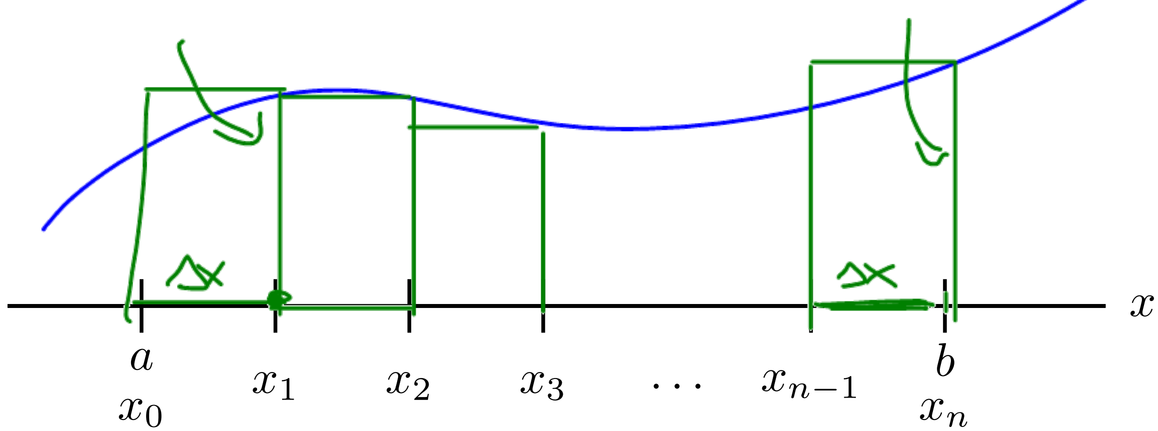


We have a similar definition for the **right hand sum**, or *RIGHT(n)*, calculated by taking the height of each rectangle to be the height of the function at the right hand endpoint of the interval.

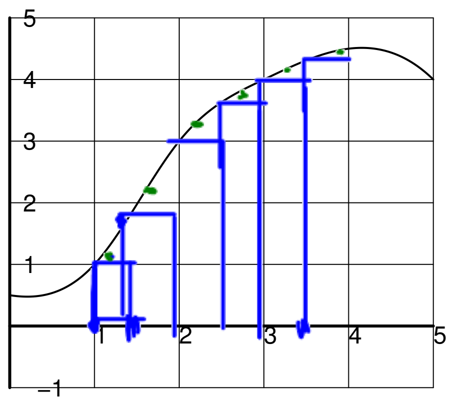
$$\begin{aligned} RIGHT(n) = & f(x_1)\Delta x + \\ & f(x_2)\Delta x + \ldots + \\ & f(x_n)\Delta x = \\ & \sum_{i=1}^{n} f(x_i)\Delta x. \end{aligned}$$

height = $f(x_1)$

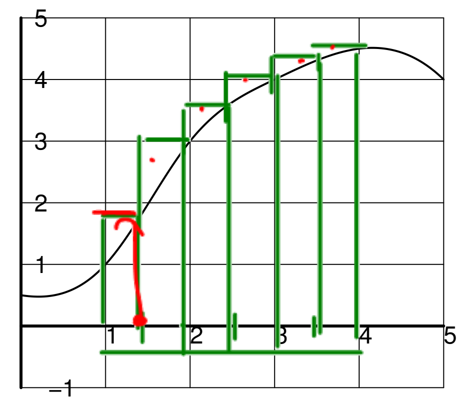
height = $f(x_n)$



Calculate LEFT(6) and RIGHT(6) for the function shown, between $x = 1$ and $x = 4$. You will need to estimate some rectangle heights from the graph.



$$\Delta x = \frac{4-1}{6} = \frac{b-a}{n}$$



LEFT(6)

h · w

$= f(1)(0.5) + f(1.5)(0.5)$
$+ f(2)(0.5) + f(2.5)(0.5)$
$+ f(3)(0.5) + f(3.5)(0.5)$

$= 0.5(\underbrace{1 + 1.6 + 3 + 3.5 + 4 + 4.2}_{\text{6 terms}})$

$= 8.65$ sq units.

Note: errors from video corrected here.

RIGHT(6)

h · w

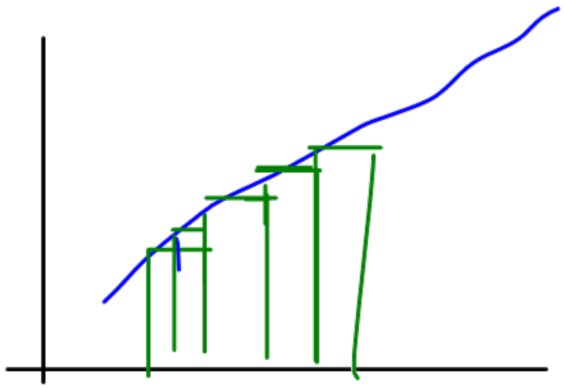
$= f(1.5)(0.5) + f(2)(0.5)$
$+ f(2.5)(0.5) + f(3.0)(0.5)$
$+ f(3.5)(0.5) + f(4)(0.5)$

$= 0.5(1.6 + 3.0 + 3.5 + 4.0 + 4.2 + 4.5)$

$= 10.4$ sq units

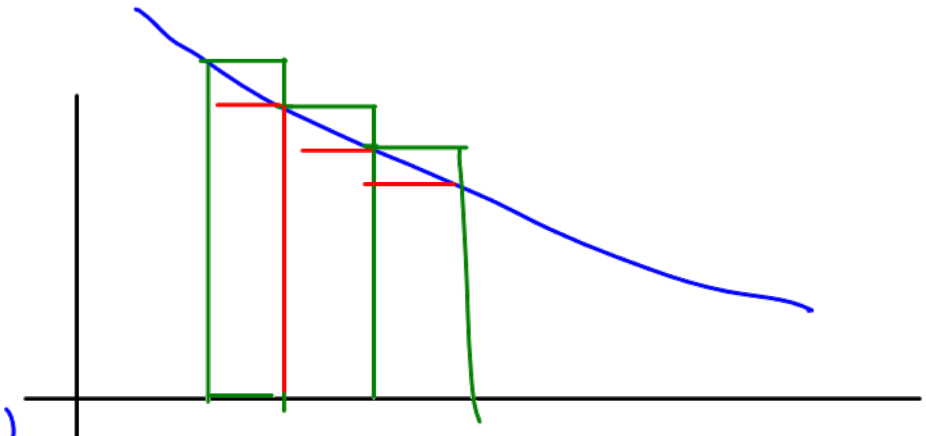
6 terms

In general, when will LEFT(n) be greater than RIGHT(n)?



LEFT(n) < RIGHT(n)

if f(x) is increasing

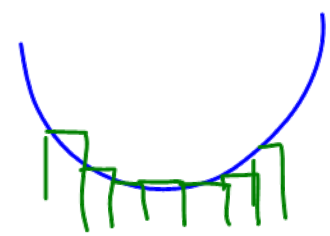


LEFT(n) > RIGHT(n)

if f(x) is decreasing

When will LEFT(n) be an overestimate for the area?

when f(x) is decreasing



When will LEFT(n) be an underestimate?

when f(x) is increasing

Riemann Sums

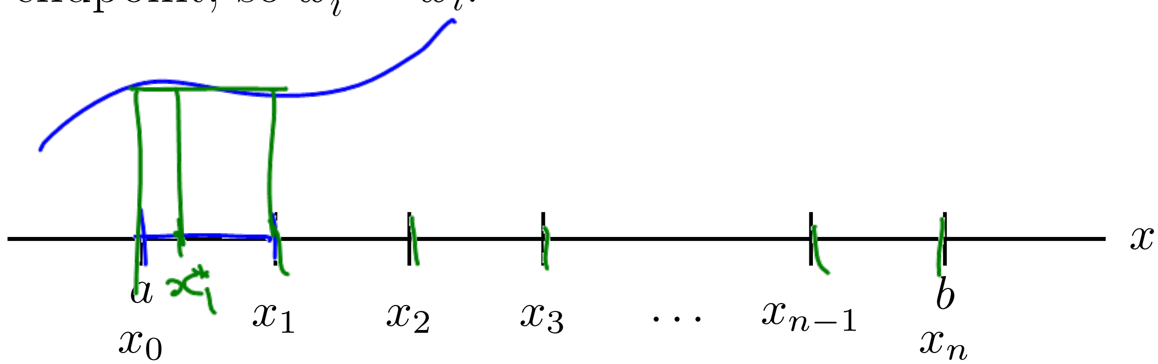
Area estimations like $LEFT(n)$ and $RIGHT(n)$ are often called **Riemann sums**, after the mathematician Bernahrd Riemann (1826-1866) who formalized many of the techniques of calculus. The general form for a Riemann Sum is

h · w

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

h · w

where each x_i^* is some point in the interval $[x_{i-1}, x_i]$. For $LEFT(n)$, we choose the left hand endpoint of the interval, so $x_i^* = x_{i-1}$; for $RIGHT(n)$, we choose the right hand endpoint, so $x_i^* = x_i$.



The common property of all these approximations is that they involve

- a sum of rectangular areas, with
- widths (Δx), and
- heights ($f(x_i^*)$)

There are other Riemann Sums that give slightly better estimates of the area underneath a graph, but they often require extra computation. We will examine some of these other calculations a little later.

The Definite Integral

We observed that as we increase the number of rectangles used to approximate the area under a curve, our estimate of the area under the graph becomes more accurate. This implies that if we want to calculate the **exact area**, we would want to use a **limit**.

exact

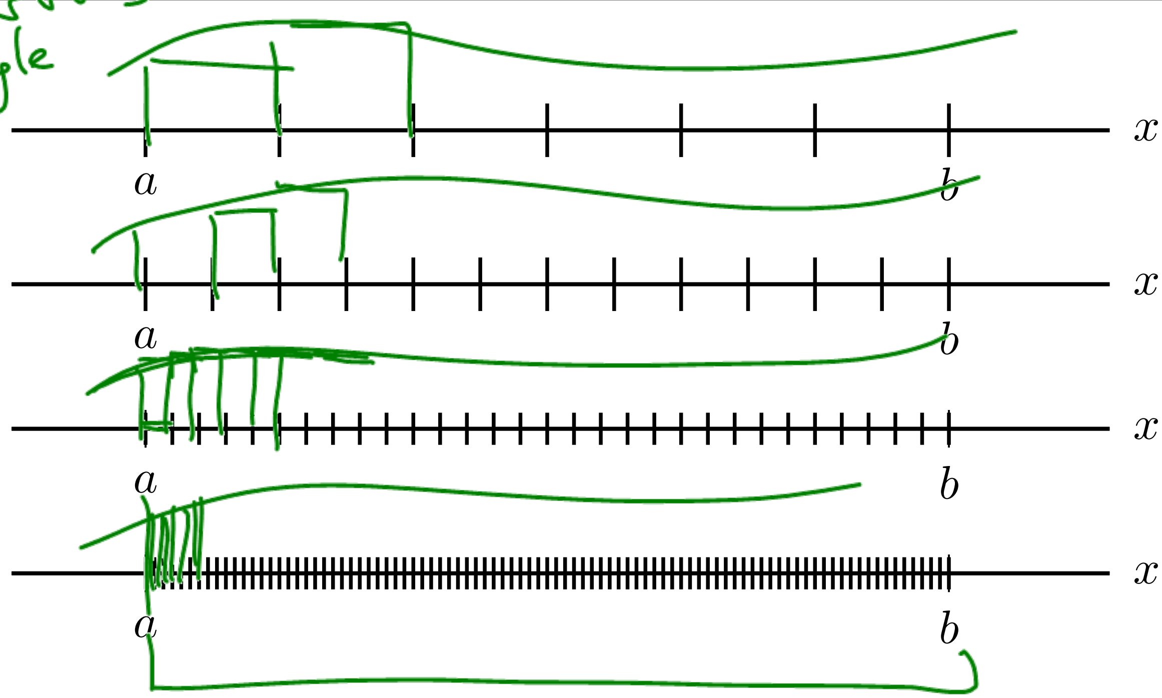
The area underneath the graph of $f(x)$ between $x = a$ and $x = b$ is equal to

$$\lim_{n\to\infty} LEFT(n) = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_{i-1})\Delta x, \text{ where } \Delta x = \frac{b-a}{n}.$$

intervals

rectangle

$\Delta x \to 0$



This limit is called the *definite integral* of $f(x)$ from a to b, and is equal to the area under curve whenever $f(x)$ is a non-negative continuous function. The definite integral is written with some special notation.

Notation for the Definite Integral

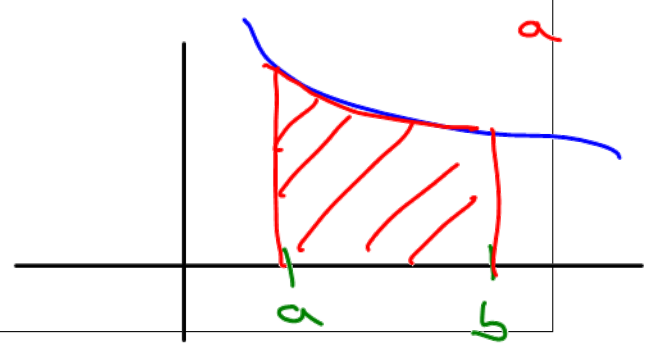
The definite integral of $f(x)$ between $x = a$ and $x = b$ is denoted by the symbol

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

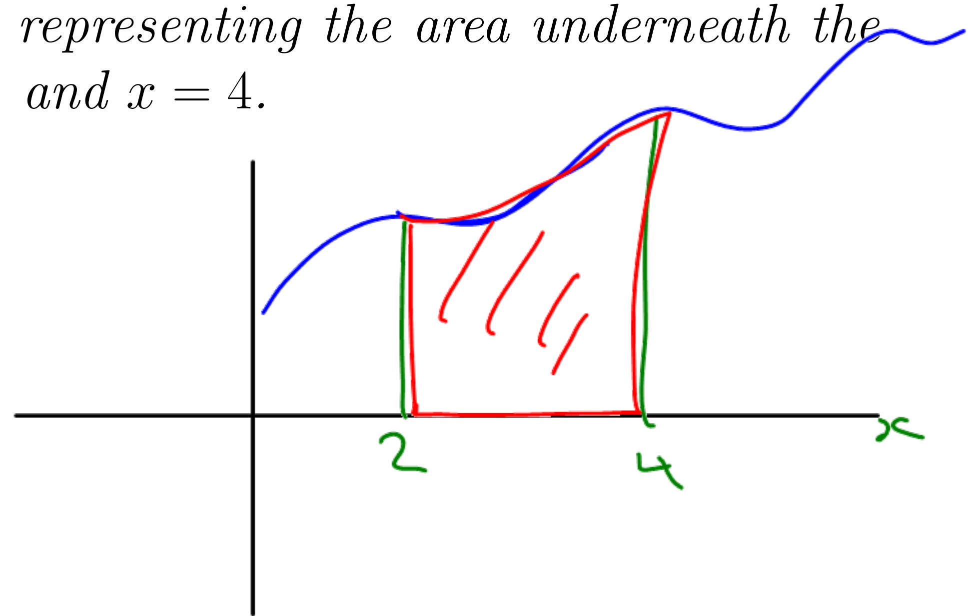
We call a and b the **limits of integration** and $f(x)$ the **integrand**. The dx denotes which variable we are using; this will become important for using some techniques for calculating definite integrals. Note that this notation shares the same common structure with Riemann sums:

area $= \int_a^b f(x)dx$

- a sum ($\int$ sign) $\qquad \sum_{i=1}^{n} f(x_i^*)\,\Delta x$
- widths (dx), and
- heights ($f(x)$)



Example: *Write the definite integral representing the area underneath the graph of* $f(x) = (x + \cos x)$ *between* $x = 2$ *and* $x = 4$.

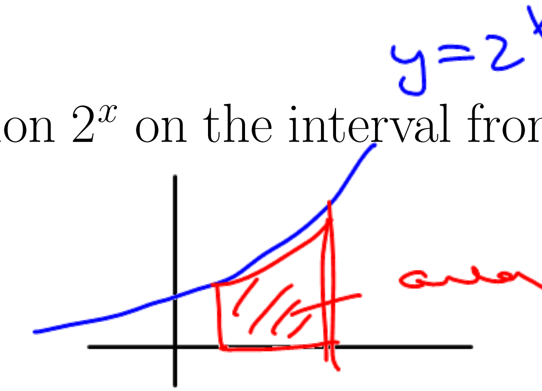


$$\text{area} = \int_2^4 (x + \cos(x))\, dx$$

The Definite Integral - $LEFT(n)$ vs $RIGHT(n)$ as $n \to \infty$

We might be concerned that we defined the area and the definite integral using the left hand sum. Would we get the same answer for the definite integral if we used the right hand sum, or any other Riemann sum? In fact, the limit using *any* Riemann sum would give us the same answer.

Let us look at the left and right hand sums for the function 2^x on the interval from $x = 1$ to $x = 3$.



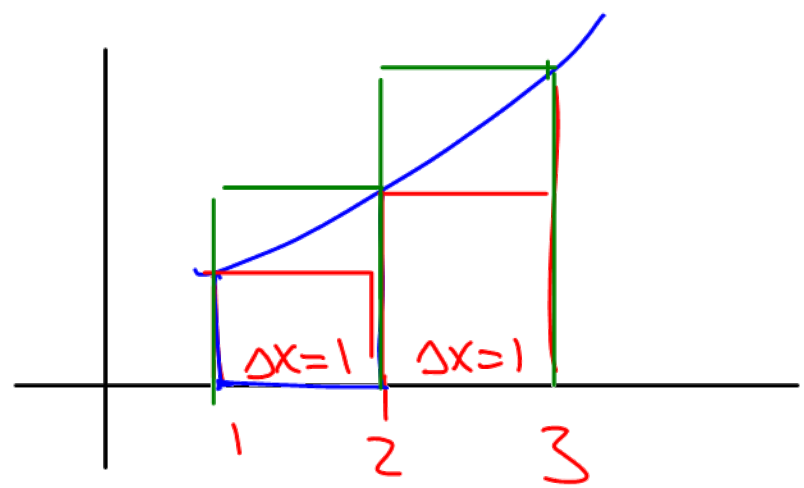
Calculate $LEFT(2) - RIGHT(2)$ *for* $\displaystyle\int_1^3 2^x\,dx$.

That is, how big is the difference between these two estimates of the area under $y = 2^x$ *over* $x = 1 \ldots 3$*?*

$LEFT(2) = 2^1(1) + 2^2(1)$

$h \cdot w + h \cdot w$

$= 6$



$RIGHT(2) = 2^2(1) + 2^3(1)$

$= 12$

$\text{Diff} = 12 - 6 = 6$

Calculate $LEFT(4) - RIGHT(4)$ *for* $\displaystyle\int_1^3 2^x\,dx$.

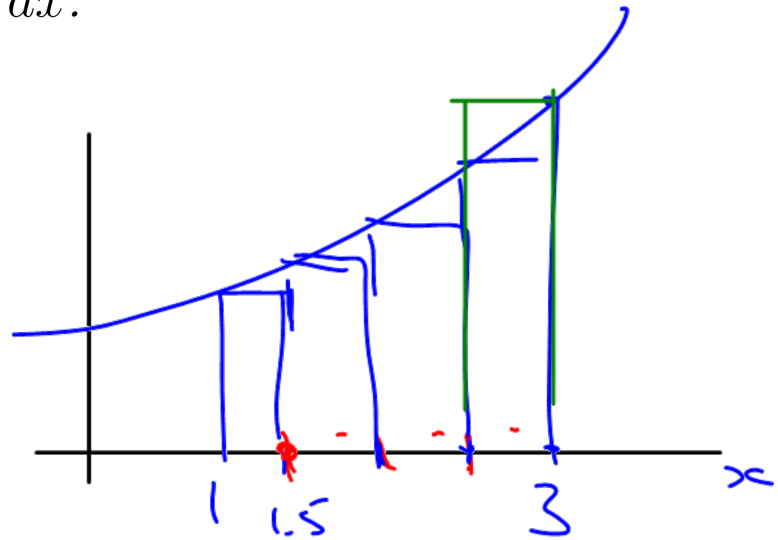
LEFT(4)

$= 2^1 \cdot (0.5) + (2^{1.5})(0.5)$

$2^2 (0.5) + (2^{2.5})(0.5)$

$= 0.5(14.485)$

$= 7.243$



$\Delta x = \dfrac{3-1}{4} = \dfrac{2}{4} = \dfrac{1}{2}$

RIGHT(4)

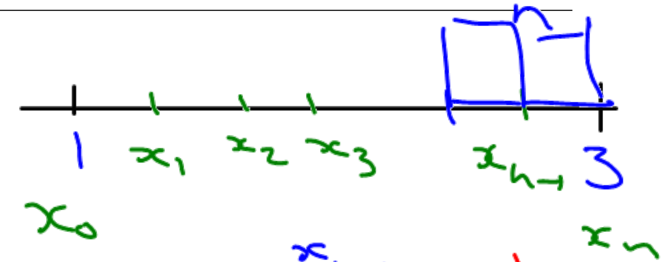
$= 2^{1.5}(0.5) + 2^2(0.5) + 2^{2.5}(0.5) + 2^3(0.5)$

$= 10.243$

Diff $= 3$

$$\Delta x = \frac{3-1}{n}$$

Calculate $LEFT(n) - RIGHT(n)$ *for* $\int_1^3 2^x\,dx.$



$$LEFT(n) = \left(2^1 \cdot \Delta x + 2^{x_1} \cdot \Delta x + 2^{x_2} \cdot \Delta x + \ldots + 2^{x_{n-1}} \cdot \Delta x\right)$$

$$RIGHT(n) = 2^{x_1} \Delta x + 2^{x_2} \cdot \Delta x + \ldots + 2^{x_{n-1}} \cdot \Delta x + 2^3 \cdot \Delta x$$

$$\text{Difference} = RIGHT(n) - LEFT(n)$$

$$= 2^3 \Delta x - 2^1 \Delta x$$

$$= \Delta x\,(8-2)$$

$$= \Delta x\,(6)$$

What will the limit of this $LEFT(n) - RIGHT(n)$ difference be as $n \to \infty$?

$LEFT(n) - RIGHT(n) = -\Delta x\,(6)$

as $n \to \infty$

$\Delta x = \frac{3-1}{n} \to 0$

$\Rightarrow LEFT(n) - RIGHT(n) \to 0$ as well

$\to LEFT(n) = RIGHT(n)$ for $n \to \infty$

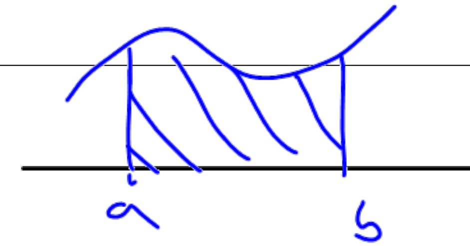
What does this tell us about what would happen if we defined the definite integral in terms of the right hand sum?

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} LEFT(n) \text{ vs. } \lim_{n\to\infty} RIGHT(n)?$$

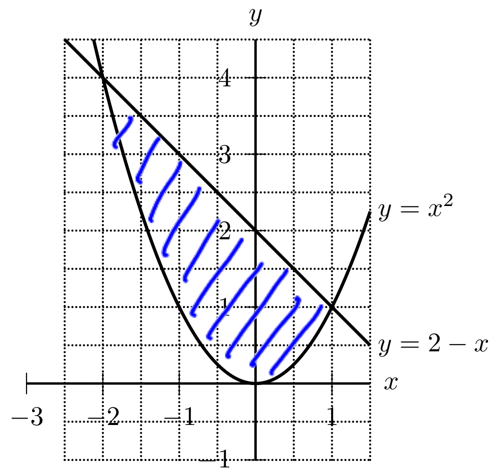
Same answer

Key: # of intervals $\to \infty$

More on the Definite Integral and Area

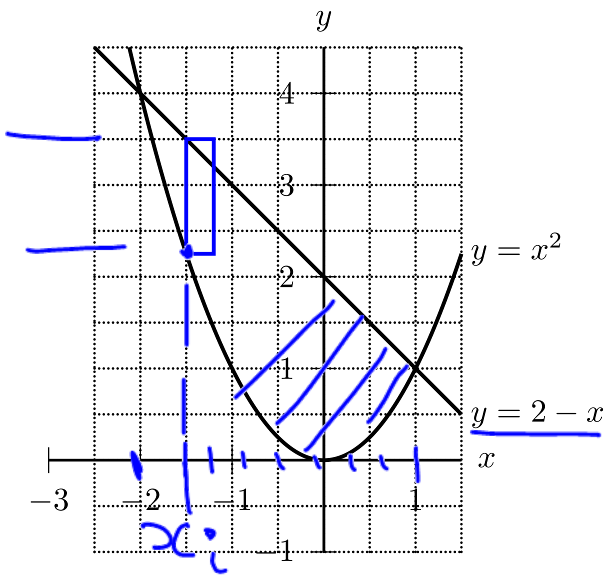


We can use the definite integral to calculate other areas, as well. Suppose we want to find the area between the curves $y = x^2$ and the line $y = 2 - x$. It is easy to see that the two intersect as shown in the following graph.



We can again calculate this area by estimating via rectangles and the taking the limit to get the definite integral.

$y_{top} - y_{lower} =$ height



If we estimate this area using the left hand sum, what will be the height of the rectangle on the interval $[x_i, x_{i+1}]$*?*

1. height $= x_i^2 - (2 - x_i)$

2. height $= (2 - x_i) - x_i^2$

3. height $= (2 - x_i) + x_i^2$

Write the formula for $LEFT(n)$*, using this height as the function value. What will* Δx *be?*

$$\sum(\text{height})\cdot(\text{width})$$

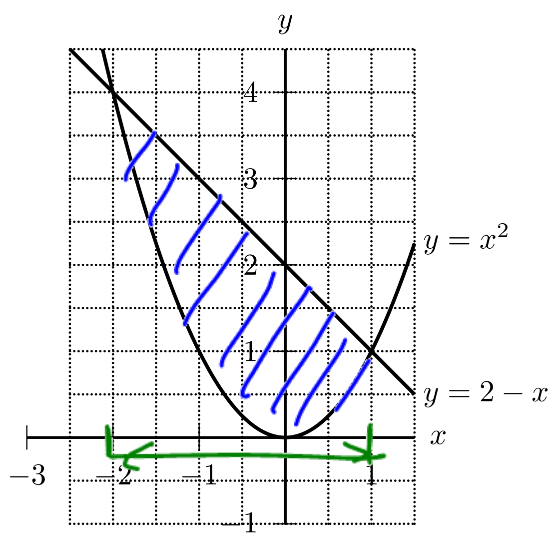
$$= \sum_{i=1}^{n}\Big[\underbrace{(2-x_{i-1})-(x_{i-1})^2}_{\text{height}}\Big]\underbrace{(\Delta x)}_{\text{width}}$$

$$\Delta x = \frac{1-(-2)}{n} = \frac{3}{n}$$

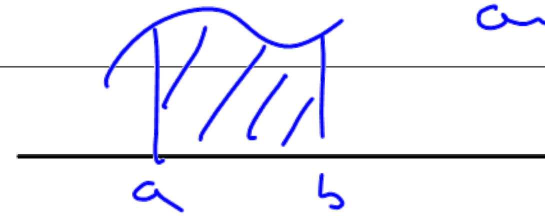
$$\lim_{n\to\infty} \quad \text{or} \quad \lim_{\Delta x\to 0}$$

Write the definite integral representing the area of this region.

$$\int_{x=-2}^{x=1}\Big[\underbrace{(2-x)}_{y_{up}}-\underbrace{x^2}_{y_{lower}}\Big]\,dx$$



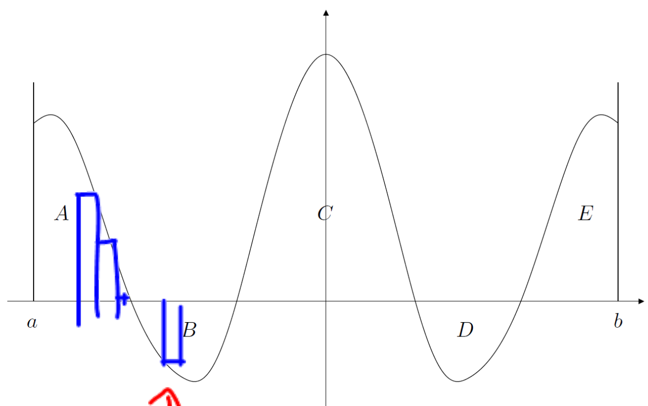
Negative Integral Values



$\text{area} = \int_a^b f(x)\,dx$

So far we have only dealt with positive functions. Will the definite integral still be equal to the area underneath the graph if $f(x)$ is always negative? What happens if $f(x)$ crosses the x-axis several times?

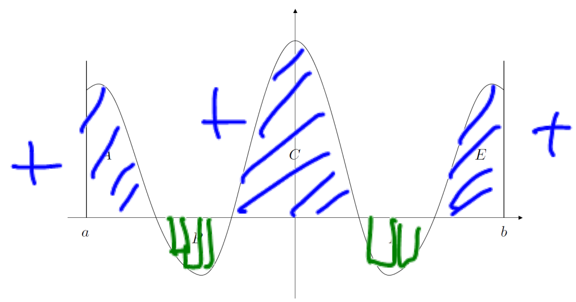
Example: *Suppose that $f(t)$ has the graph shown below, and that A, B, C, D, and E are the areas of the regions shown.*



$f(x_i)$ negative

If we were to partition $[a, b]$ into small subintervals and construct a corresponding Riemann sum, then the first few terms in the Riemann sum would correspond to the region with area A, the next few to B, etc.

$\sum f(x_i) \cdot \Delta x$



$\sum f(x_i) \cdot \Delta x$

+ on A, C, E

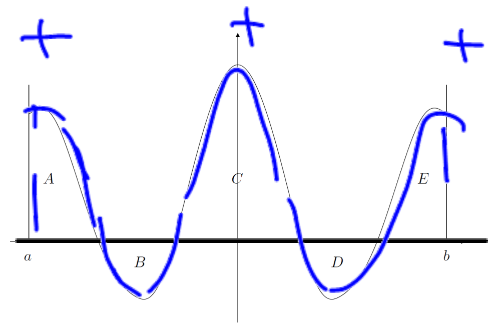
Which of these sets of terms have positive values?

A, C, E

Which of these sets have negative values?

B, D

$\sum f(x_i)\cdot\Delta x$



Express the integral $\int_a^b f(t)\,dt$ *in terms of the (positive) areas* A, B, C, D, *and* E.

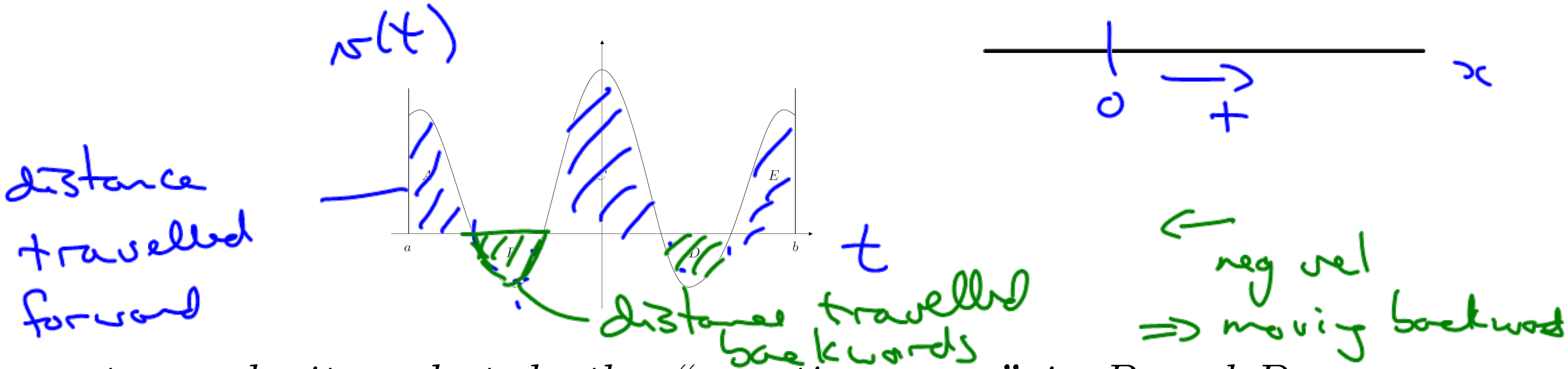
always positive

(a) $\int_a^b f(t)\,dt = \mathrm{A} + \mathrm{B} + \mathrm{C} + \mathrm{D} + \mathrm{E}$

(b) $\int_a^b f(t)\,dt = \mathrm{A} - \mathrm{B} + \mathrm{C} - \mathrm{D} + \mathrm{E}$

(c) $\int_a^b f(t)\,dt = -\mathrm{A} + \mathrm{B} - \mathrm{C} + \mathrm{D} - \mathrm{E}$

(d) $\int_a^b f(t)\,dt = -\mathrm{A} - \mathrm{B} - \mathrm{C} - \mathrm{D} - \mathrm{E}$



If $f(t)$ represents a velocity, what do the "negative areas" in B and D represent?

(a) The areas B and D represent **negative positions**.

(b) The areas B and D represent **backwards motion**.

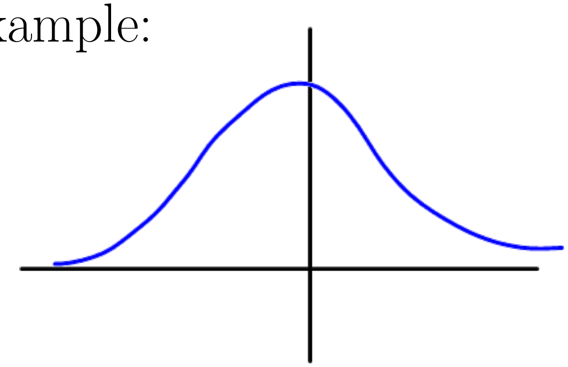
(c) The areas B and D represent **distance travelled backwards**.

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

Better Approximations to Definite Integrals

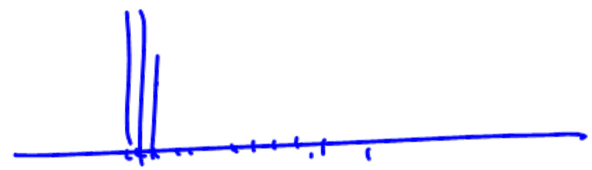
- We saw how to *approximate* definite integrals using $LEFT(n)$ and $RIGHT(n)$ Riemann sums. Unfortunately, these estimates are very crude and inefficient.
- Even when we have sophisticated techniques for evaluating integrals, these methods will not apply to all functions, for example:

  · $\displaystyle\int e^{-x^2}\,dx$ - used in probability

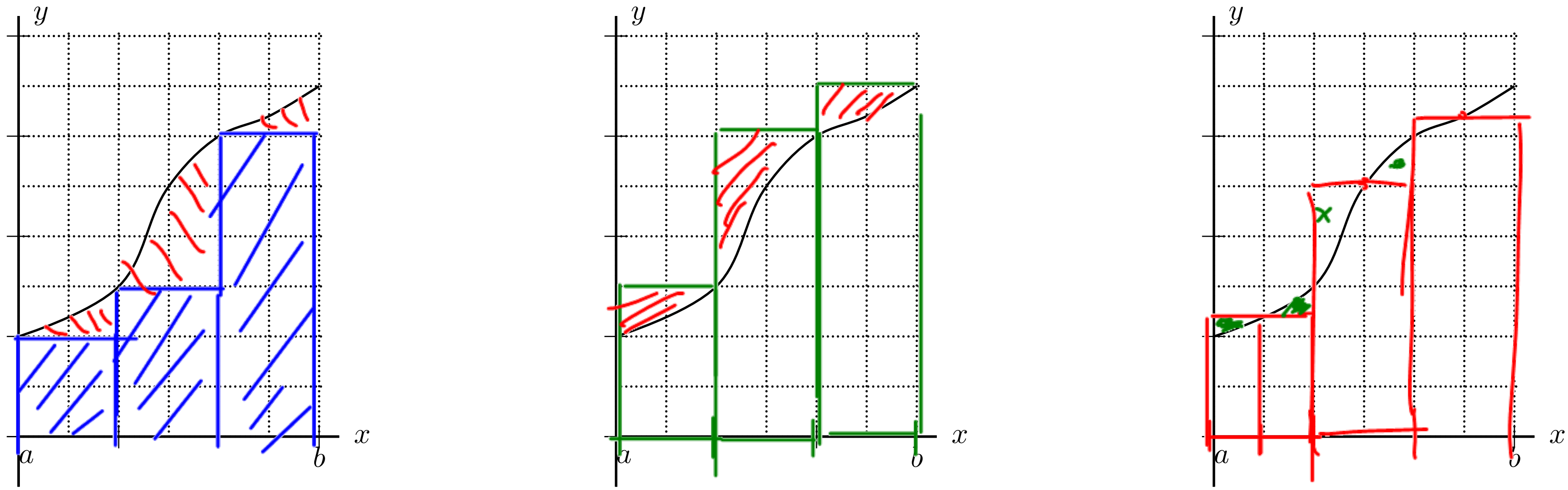
  · $\displaystyle\int \frac{\sin(x)}{x}\,dx$ - used in optics

- To evaluate definite integrals of such functions, we could use left or right hand Riemann sums. However, it would be preferable to develop *more accurate estimates*.
- But more accurate estimates can always be made by using more rectangles.
- More precisely, we want to develop more *efficient* estimates: estimates that are *more* accurate for *similar amounts of work*.



Midpoint Rule

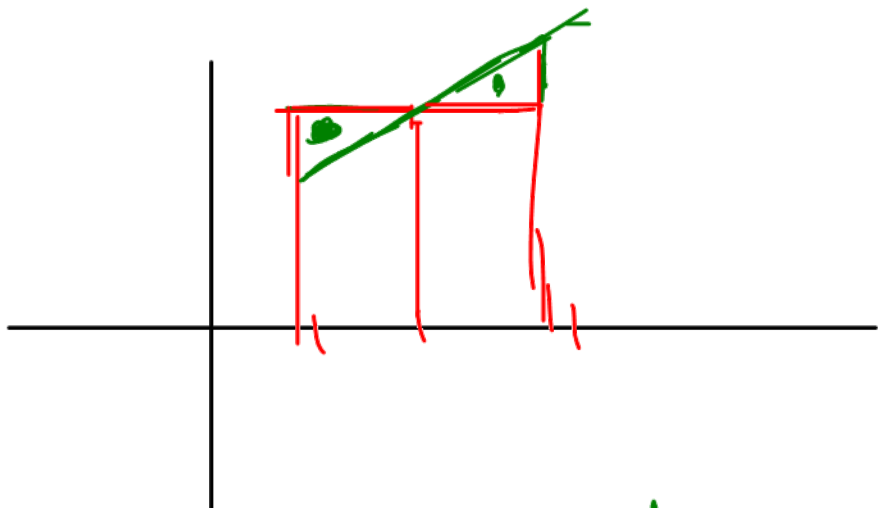
The accuracy of a Riemann sum calculation will usually improve if we choose the midpoint of each subdivision rather than the right or left endpoints.



Compare the accuracy of the left-hand, right-hand and midpoint rules for estimating the area on the interval.

more accurate.

For what kinds of functions f will the midpoint rule always give a value that is exactly equal to the integral?

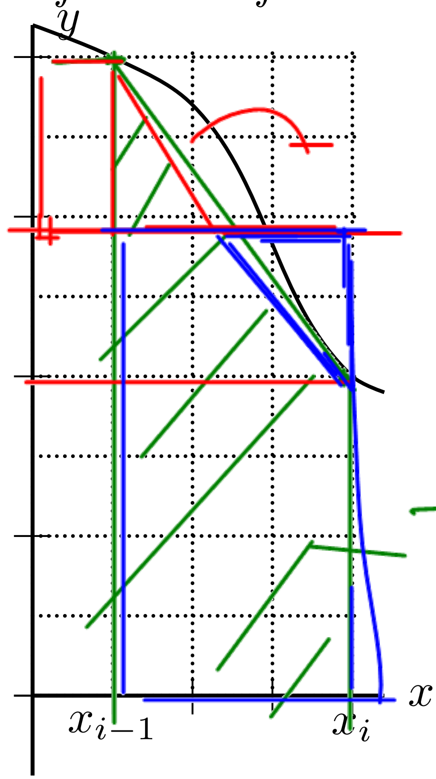


linear functions $\rightarrow$ Midpoint sum $= \int_a^b f(x)\,dx$

The Trapezoidal Rule

The Midpoint Rule is only one possible variation on Riemann sums. Another approach is to use a shape other than a rectangle to estimate the area on an interval. For the appropriately named Trapezoidal Rule, we use a trapezoid on each interval.

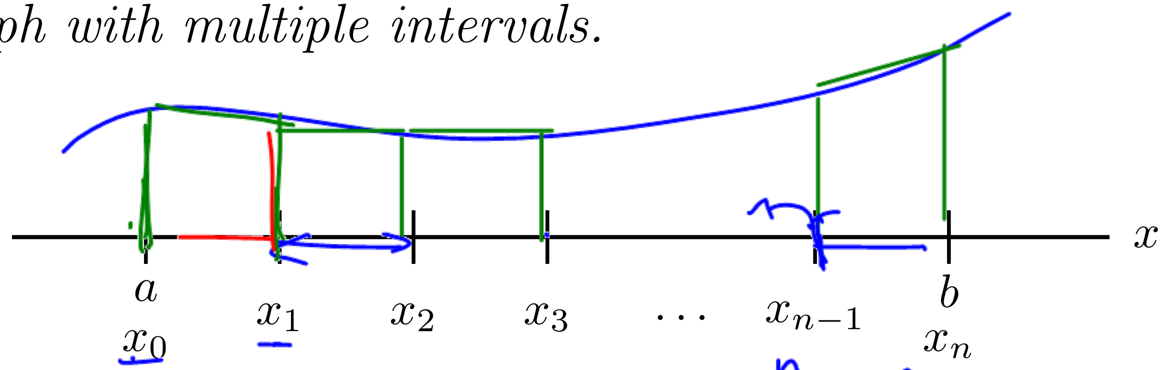
Sketch a trapezoidal approximation to the area under the graph, and then write a formula for the area of the single trapezoid.



$$\text{area} = \left(\begin{array}{c}\text{avg of}\\ \text{two heights}\end{array}\right)(\Delta x)$$

$$= \left(\frac{f(x_{i-1}) + f(x_i)}{2}\right)\Delta x$$

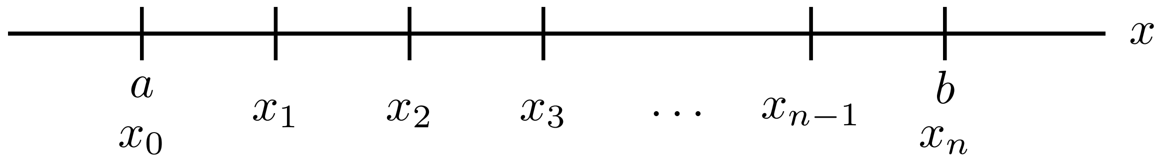
Write a formula for the full Trapezoid Rule (written $TRAP(n)$), estimating the entire area under a graph with multiple intervals.



$$TRAP(n) = \left(\frac{f(x_0)+f(x_1)}{2}\right)\Delta x + \left(\frac{f(x_1)+f(x_2)}{2}\right)\Delta x + \cdots + \left(\frac{f(x_{n-1})+f(x_n)}{2}\right)\Delta x = \sum_{i=1}^{n}\left(\frac{f(x_{i-1})+f(x_i)}{2}\right)\Delta x$$

$$= \Delta x\left[\frac{f(x_0)}{2} + \frac{\cancel{2}f(x_1)}{\cancel{2}} + \frac{\cancel{2}f(x_2)}{\cancel{2}} + \cdots + \frac{\cancel{2}f(x_{n-1})}{\cancel{2}} + \frac{f(x_n)}{2}\right]$$

Alternative: Express the trapezoidal rule $TRAP(n)$ in terms of the left and right hand Riemann sums ($LEFT(n)$ and $RIGHT(n)$).

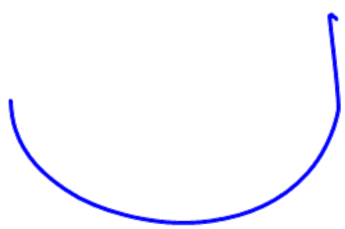


$$TRAP(n) = \frac{LEFT(n) + RIGHT(n)}{2}$$

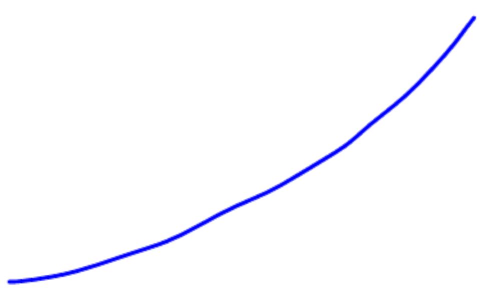
$$= \text{average of } LEFT(n), RIGHT(n)$$

'The trapezoidal rule is especially accurate for functions f for which $|f''(x)|$ is small for all x.' Explain from an intuitive point of view why you would expect this statement to be correct.

$f'' \to$ curvature / concavity



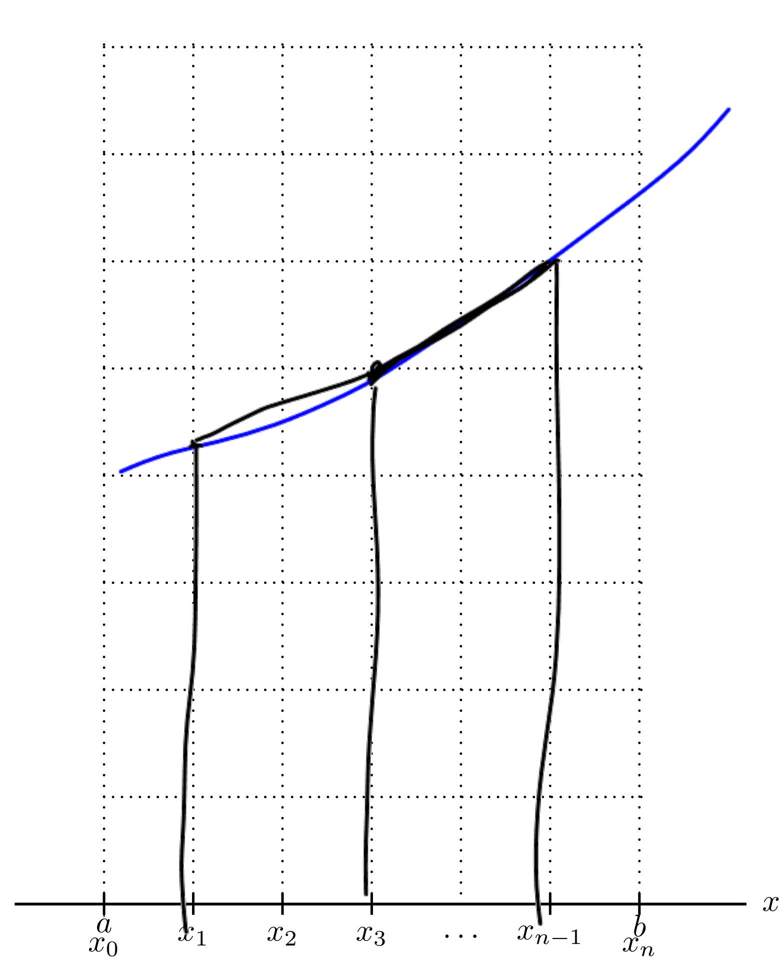
large f''



small f''

$\Downarrow$

closer to linear



Use the trapezoidal rule to estimate $\int_0^{10} f(x)\,dx$*, if we have measured the values in the following table for* $f(x)$.

$\Delta x = 2 - 0$

x	0	2	4	6	8	10
f(x)	1	3	4	5	4	2

TRAP(n)

$$= \underbrace{\Delta x}_{\text{widths}} \left(\underbrace{\frac{f(0)}{2} + f(2) + f(4) + f(6) + f(8) + \frac{f(10)}{2}}_{\text{sum of average heights}} \right)$$

$$= 2\left(\frac{1}{2} + 3 + 4 + 5 + 4 + \frac{2}{2}\right)$$

$$= 2(17.5)$$

$$= 35$$