Bachelor’s degree in Bioinformatics

Integrals

Prof. Renato Bruni

bruni@diag.uniroma1.it

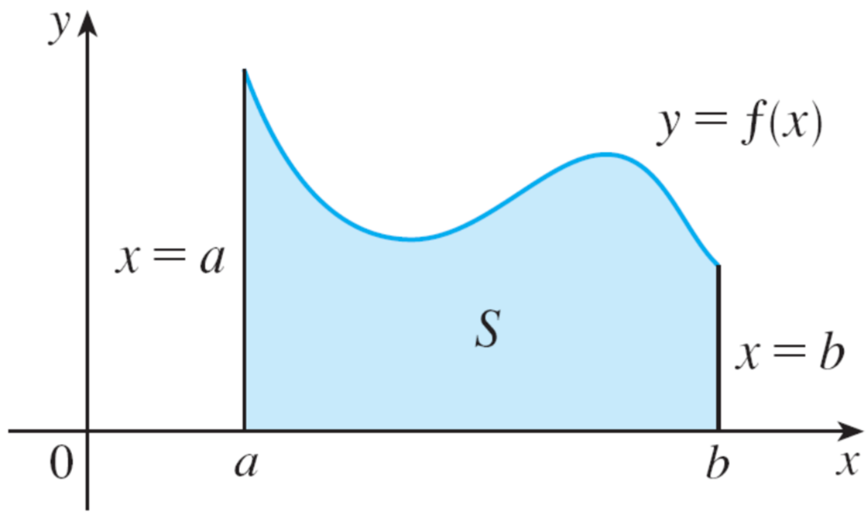
Department of Computer, Control, and Management Engineering (DIAG)

“Sapienza” University of Rome

The Area Problem

We now want to solve the *area problem*: Find the area of the region S that lies under the curve $y = f(x)$ from a to b

S is bounded by the graph of a continuous function f [where $f(x) \geq 0$], the vertical lines $x = a$ and $x = b$, and the x-axis



$$S = \{(x, y) \mid a \leq x \leq b, 0 \leq y \leq f(x)\}$$

The Area Problem

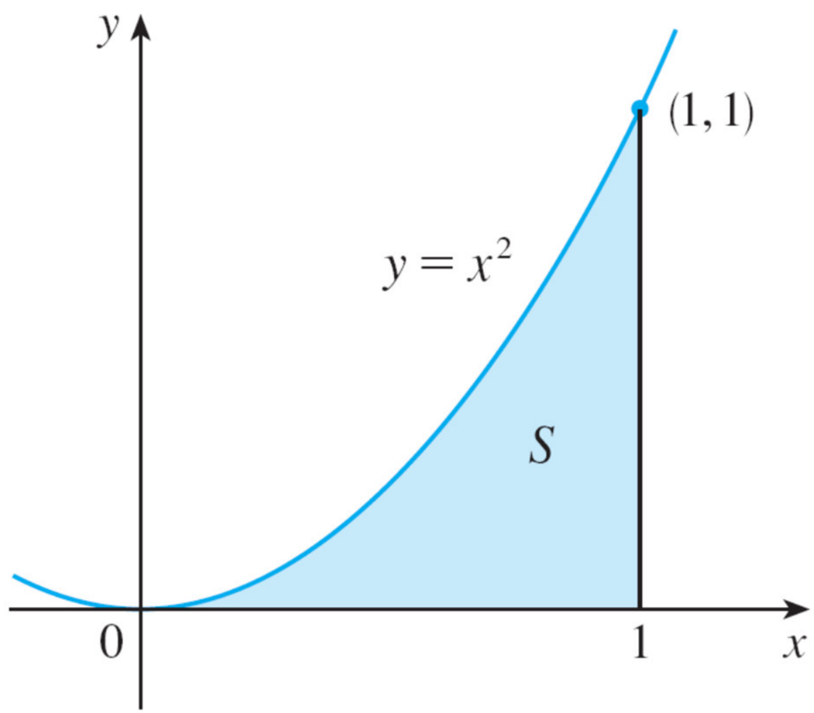
It is not easy to find the area of a region with curved sides. We all have an intuitive idea of what the area of a region is, but now we try to give an exact definition of area

Recall that in defining a tangent we first approximated the slope of the tangent line by slopes of secant lines and then we took the limit of these approximations

We use a similar technique for areas. We first approximate the region S by rectangles and then we take the limit of the areas of these rectangles as we increase the number of rectangles

Example

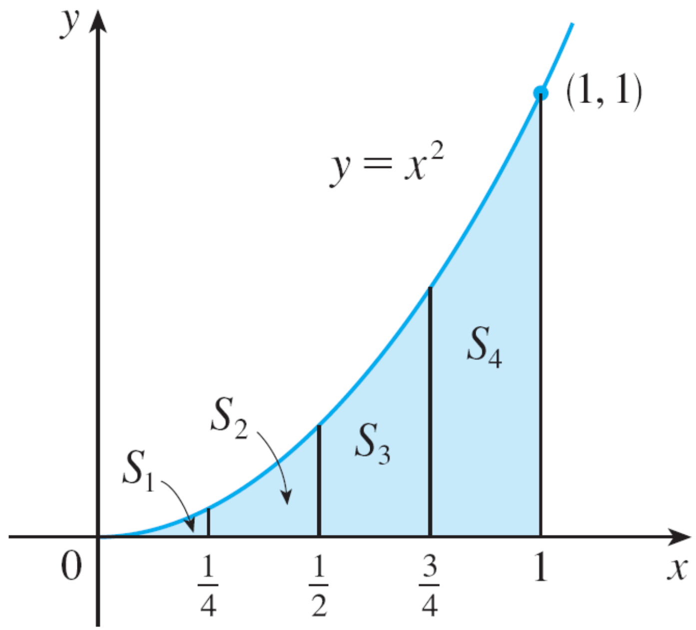
Use rectangles to estimate the area under the parabola $y = x^2$ from 0 to 1 (the parabolic region S)



Example – *Solution*

We first notice that the area of S must be somewhere between 0 and 1 because S is contained in a square with side length 1

Suppose we divide S into four strips S_1, S_2, S_3, S_4 by drawing the vertical lines $x = 1/4$, $x = 1/2$, and $x = 3/4$



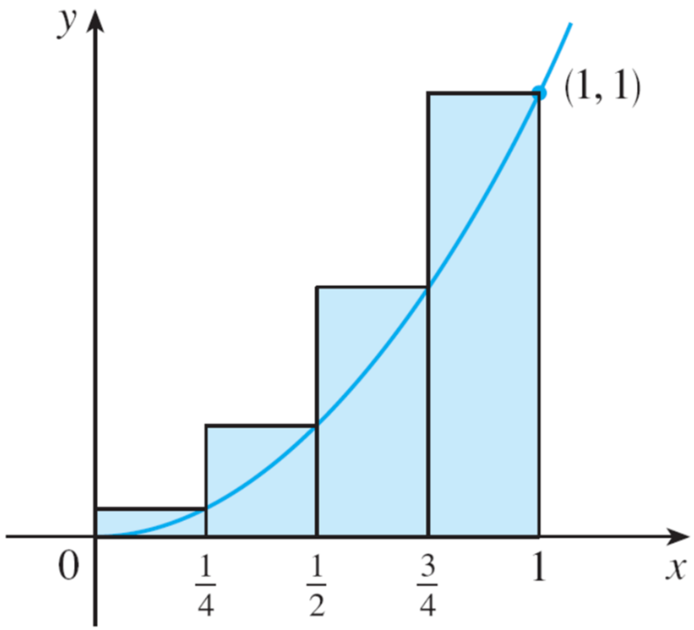
Example – *Solution*

cont'd

We can approximate each strip by a rectangle that has the same base as the strip and whose height is the same as the right edge of the strip

In other words, the heights of these rectangles are the values of the function $f(x) = x^2$ at the *right* endpoints of the subintervals

$$\left[0, \tfrac{1}{4}\right], \left[\tfrac{1}{4}, \tfrac{1}{2}\right], \left[\tfrac{1}{2}, \tfrac{3}{4}\right], \left[\tfrac{3}{4}, 1\right].$$



Each rectangle has width 1/4 and the heights are $\left(\tfrac{1}{4}\right)^2, \left(\tfrac{1}{2}\right)^2, \left(\tfrac{3}{4}\right)^2, 1^2$

Example – *Solution*

cont'd

If we let R_4 be the sum of the areas of these approximating rectangles, we get

$$R_4 = \frac{1}{4} \cdot \left(\frac{1}{4}\right)^2 + \frac{1}{4} \cdot \left(\frac{1}{2}\right)^2 + \frac{1}{4} \cdot \left(\frac{3}{4}\right)^2 + \frac{1}{4} \cdot 1^2$$

$$= \frac{15}{32}$$

$$= 0.46875$$

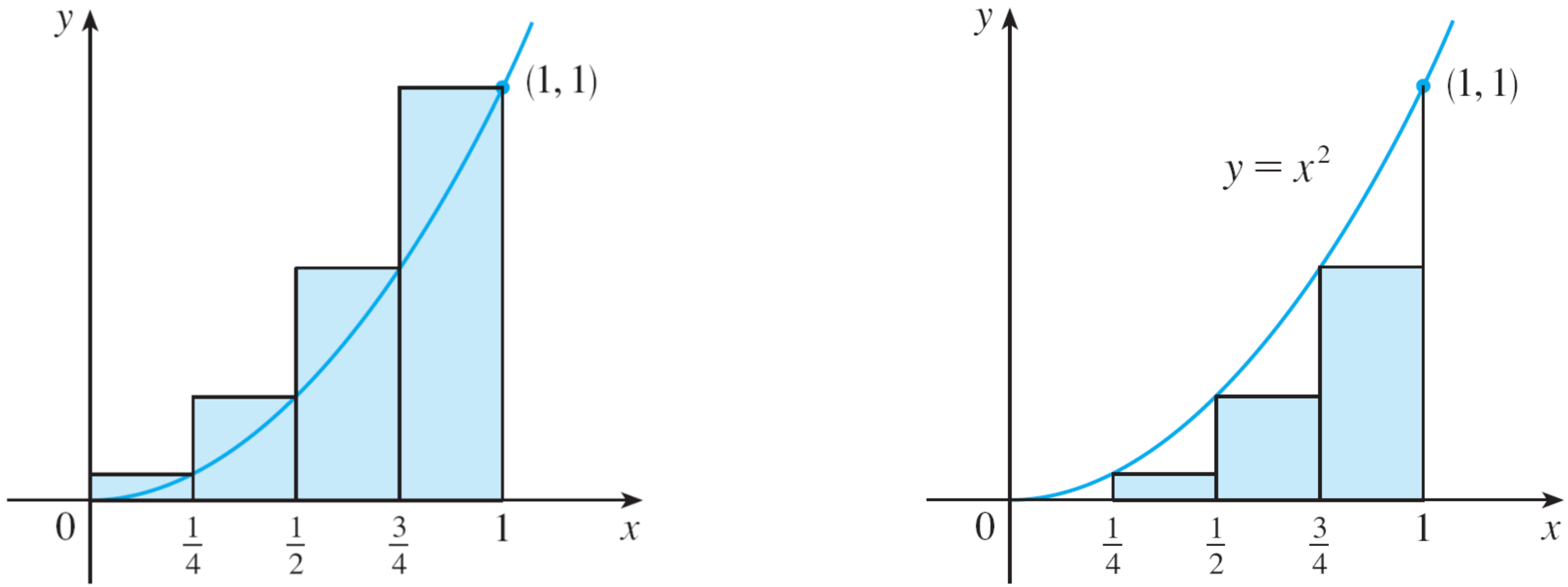
We see that the real area A of S is a bit smaller than R_4, so

$$A < 0.46875$$

Example – *Solution*

cont'd

Instead of using the rectangles on the left we could use the smaller rectangles on the right, whose heights are the values of f at the *left* endpoints of the subintervals (the leftmost rectangle has collapsed because its height is 0)



Example – *Solution*

cont'd

The sum of the areas of these approximating rectangles is

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

$$= \frac{7}{32}$$

$$= 0.21875$$

We see that the area of S is larger than L_4, so we now have lower and upper estimates for A:

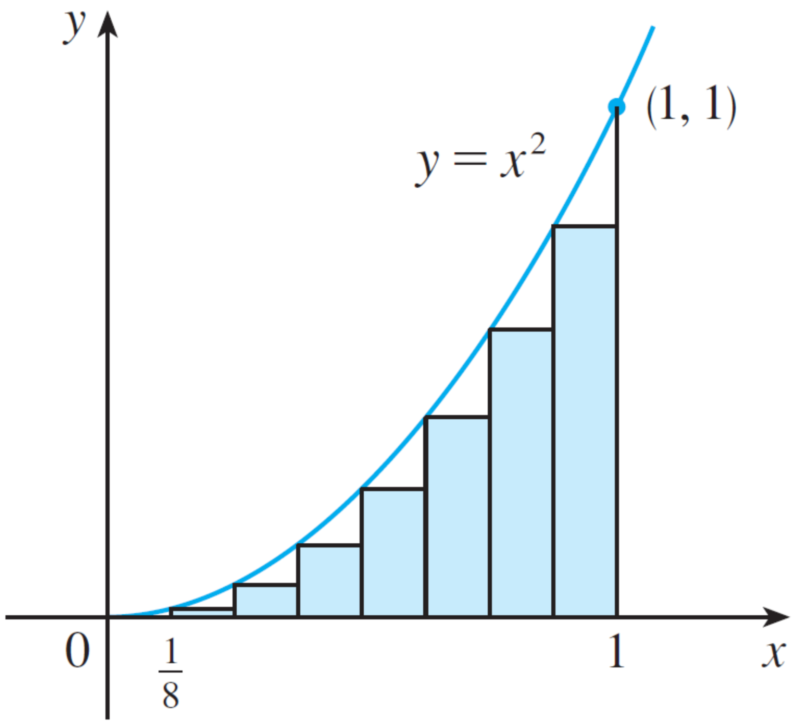
$$0.21875 < A < 0.46875$$

We can repeat this procedure with a larger number of strips

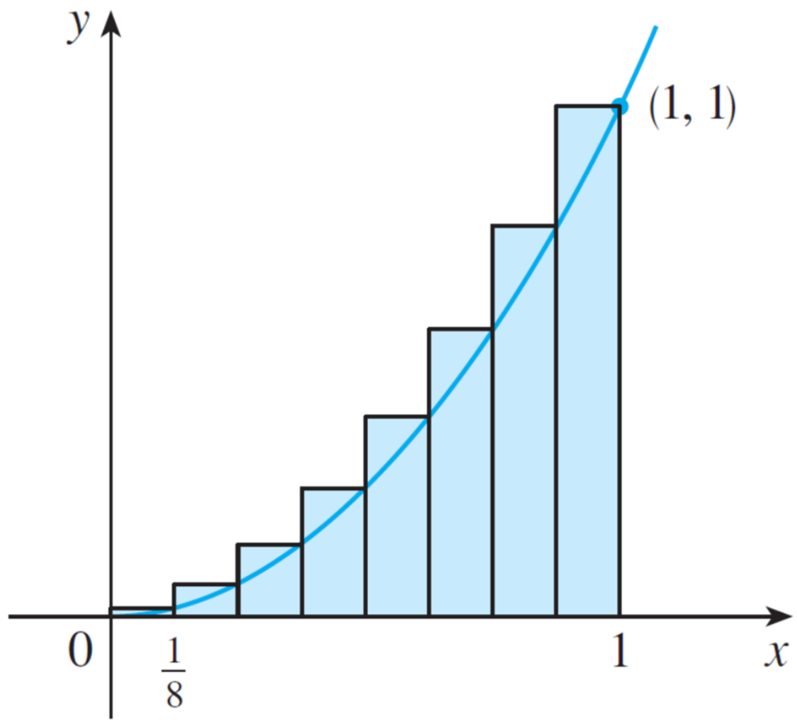
Example – *Solution*

cont'd

Here is what happens when we divide the region *S* into eight strips of equal width



(a) Using left endpoints



(b) Using right endpoints

Approximating *S* with eight rectangles

Example – *Solution*

cont'd

By computing the sum of the areas of the smaller rectangles (L_8) and the sum of the areas of the larger rectangles (R_8), we obtain better lower and upper estimates for A:

$$0.2734375 < A < 0.3984375$$

So one possible answer to the question is to say that the true area of S lies somewhere between 0.2734375 and 0.3984375

We could obtain better estimates by increasing the number of strips

Example – *Solution*

cont'd

The table at the right shows the results of similar calculations using n rectangles whose heights are found with left endpoints (L_n) or right endpoints (R_n)

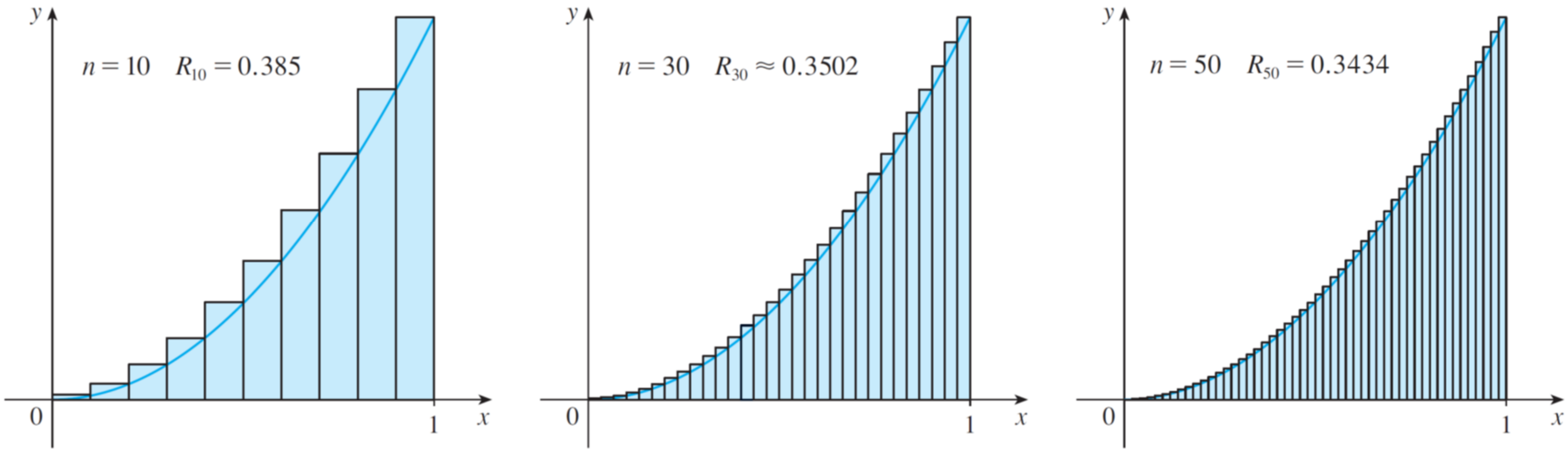
n	L_n	R_n
10	0.2850000	0.3850000
20	0.3087500	0.3587500
30	0.3168519	0.3501852
50	0.3234000	0.3434000
100	0.3283500	0.3383500
1000	0.3328335	0.3338335

In particular, we see by using 50 strips that the area lies between 0.3234 and 0.3434. With 1000 strips we narrow it down even more: A lies between 0.3328335 and 0.3338335

A good estimate is obtained by averaging these numbers:
$A \approx 0.3333335$

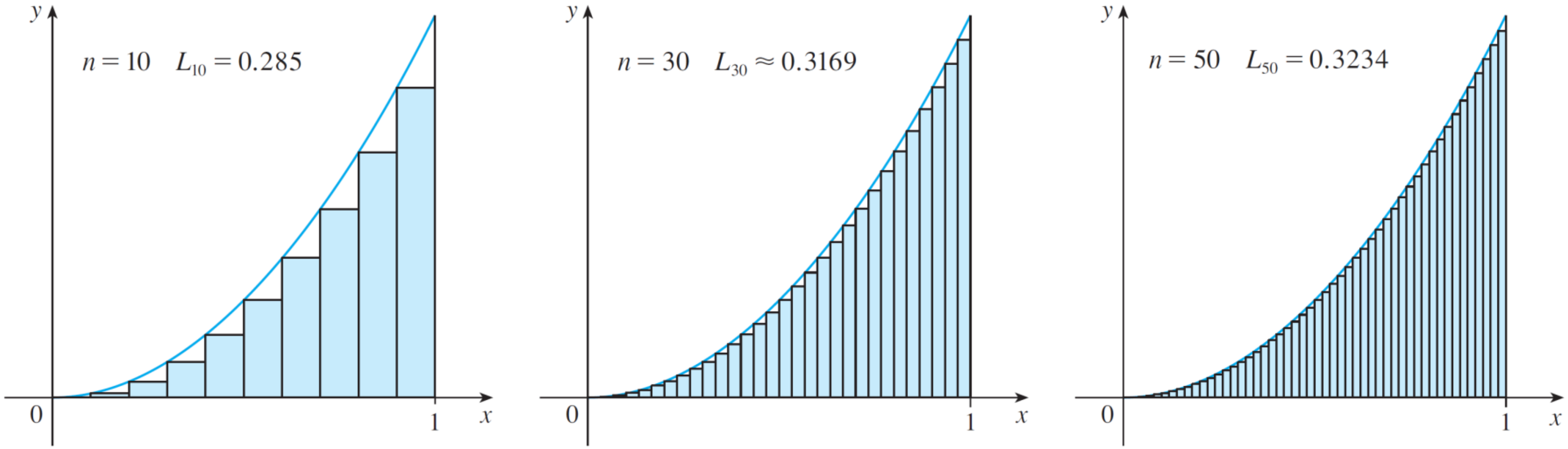
Approximating the Area

Therefore, as n increases, both L_n and R_n become better and better approximations to the area of S



Right endpoints produce upper sums because $f(x) = x^2$ is increasing

Approximating the Area



Left endpoints produce lower sums because $f(x) = x^2$ is increasing

Therefore we *define* the area A to be the **limit** of the sums of the areas of the approximating rectangles, that is,

$$A = \lim_{n \to \infty} R_n = \lim_{n \to \infty} L_n = \frac{1}{3}$$

General case of area

Hence, in order to evaluate the area S under a generic function $f(x) \geq 0$, we divide S into n stripes S_i, each of equal width $\Delta x = (b-a)/n$

So, interval $[a, b]$ is divided into n subintervals, with $x_0 = a$ and $x_n = b$

$$[x_0, x_1], \quad [x_1, x_2], \quad [x_2, x_3], \quad \ldots, \quad [x_{n-1}, x_n]$$

The right endpoints of the subintervals are

$$x_1 = a + \Delta x,$$

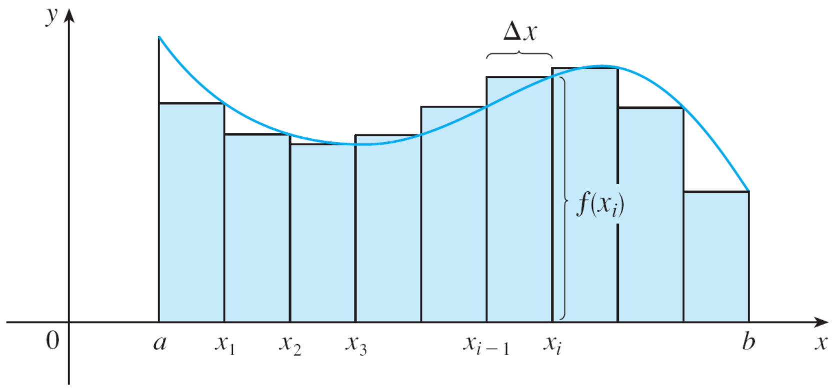
$$x_2 = a + 2\,\Delta x,$$

$$x_3 = a + 3\,\Delta x,$$

$$\vdots$$

General case of area

Now we approximate the i-th strip S_i by a rectangle with width Δx and height $f(x_i)$, which is the value of f at the right endpoint

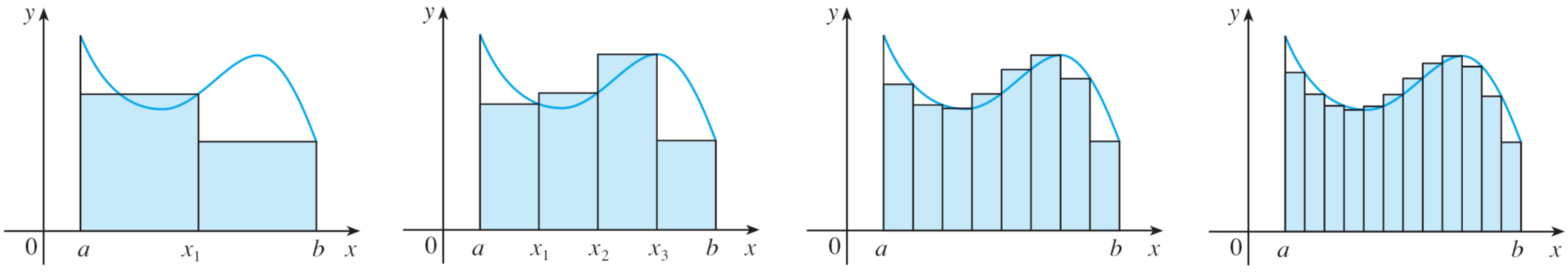


Then the area of the i-th rectangle is $f(x_i)\,\Delta x$. The area of S is approximated by the sum of the areas of these rectangles, which is

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

General case of area

This approximation improves as the number of strips increases, see for example the cases of n = 2, 4, 8, and 12
So, we take the value for $n \to \infty$



(a) $n = 2$ (b) $n = 4$ (c) $n = 8$ (d) $n = 12$

General case of area

Therefore we define the area *A* of the region *S* in the following way.

2 Definition The **area** A of the region S that lies under the graph of the continuous function f is the limit of the sum of the areas of approximating rectangles:

$$A = \lim_{n \to \infty} R_n = \lim_{n \to \infty} [f(x_1)\,\Delta x + f(x_2)\,\Delta x + \cdots + f(x_n)\,\Delta x]$$

It can be proved that this limit always exists, since we are assuming that *f* is continuous.

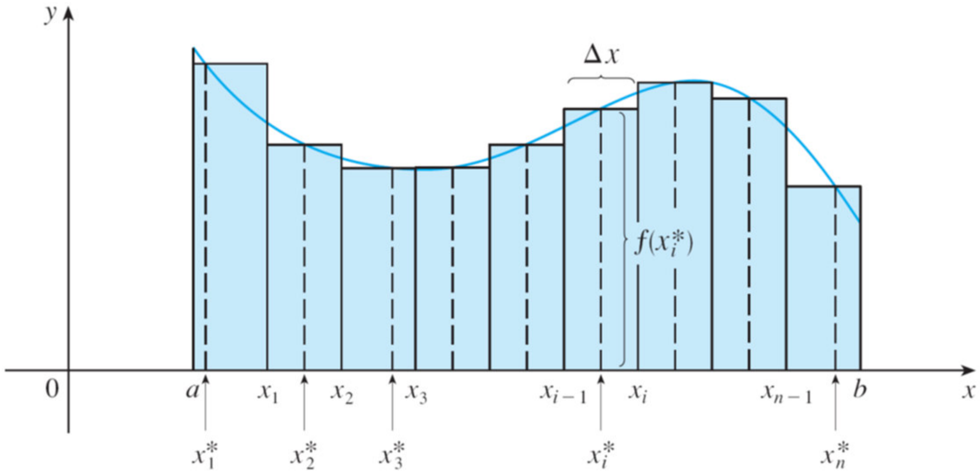
Moreover, it can be shown that we get ***the same value*** if we use left endpoints:

3 $$A = \lim_{n \to \infty} L_n = \lim_{n \to \infty} [f(x_0)\,\Delta x + f(x_1)\,\Delta x + \cdots + f(x_{n-1})\,\Delta x]$$

General case of area

In fact, instead of using left endpoints or right endpoints, we could take the height of the i-th rectangle to be the value of f at ***any*** number x_i* in the i-th subinterval $[x_{i-1}, x_i]$. We call the numbers $x_1*, x_2*, \ldots, x_n*$ the **sample points**

Approximating rectangles when the sample points are not chosen to be endpoints:



So a more general expression for the area of S is

$$A = \lim_{n \to \infty} [f(x_1^*)\,\Delta x + f(x_2^*)\,\Delta x + \cdots + f(x_n^*)\,\Delta x]$$

The Definite Integral

We have seen that a limit of this form is used to solve the area problem

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*)\,\Delta x = \lim_{n\to\infty} [f(x_1^*)\,\Delta x + f(x_2^*)\,\Delta x + \cdots + f(x_n^*)\,\Delta x]$$

It turns out that this same type of limit occurs in a wide variety of situations even when *f* is not necessarily a positive function

So, we give a name to this special type of limit, too

The Definite Integral

2 Definition of a Definite Integral If f is a function defined for $a \leqslant x \leqslant b$, we divide the interval $[a, b]$ into n subintervals of equal width $\Delta x = (b - a)/n$. We let $x_0 \, (= a), x_1, x_2, \ldots, x_n \, (= b)$ be the endpoints of these subintervals and we let $x_1^*, x_2^*, \ldots, x_n^*$ be any **sample points** in these subintervals, so x_i^* lies in the ith subinterval $[x_{i-1}, x_i]$. Then the **definite integral of f from a to b** is

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

provided that this limit exists and gives the same value for all possible choices of sample points. If it does exist, we say that f is **integrable** on $[a, b]$.

Notation

In the notation $\int_a^b f(x)\,dx,$ the function *f*(*x*) is called the **integrand** and *a* and *b* are called the **limits of integration** (*a* is the **lower limit** and *b* is the **upper limit**)

The procedure of calculating an integral is called **integration**

The symbol *dx* has no meaning by itself; it simply indicates that the independent variable is *x*

The symbol $\int$ was introduced by Leibniz and is called an **integral sign**

It is an elongated *S* and was chosen because an integral is a limit of sums

Notation

Note: The definite integral $\int_a^b f(x)\,dx$ is a number; it does not depend on *x.* In fact, we could use any letter in place of *x* without changing the value of the integral:

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \int_a^b f(r)\,dr$$

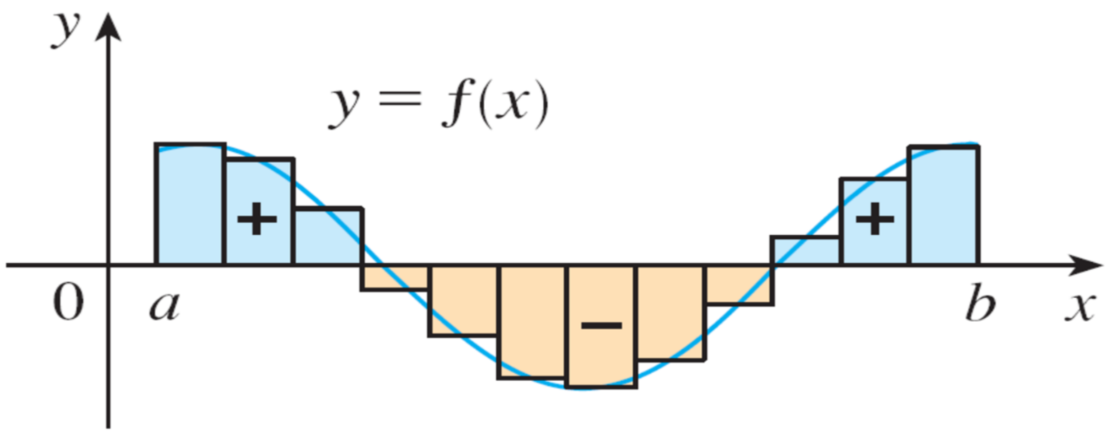
Note: The sum $\sum_{i=1}^{n} f(x_i^*)\,\Delta x$

is called a **Riemann sum** after the German mathematician Bernhard Riemann

The Definite Integral

If $f(x) \geq 0$, we have seen that the definite integral can be interpreted as the area under the curve $y = f(x)$

If $f(x)$ takes both positive and negative values, then the Riemann sum is the sum of the areas of the rectangles that lie above the *x*-axis and the *negatives* of the areas of the rectangles that lie below the *x*-axis



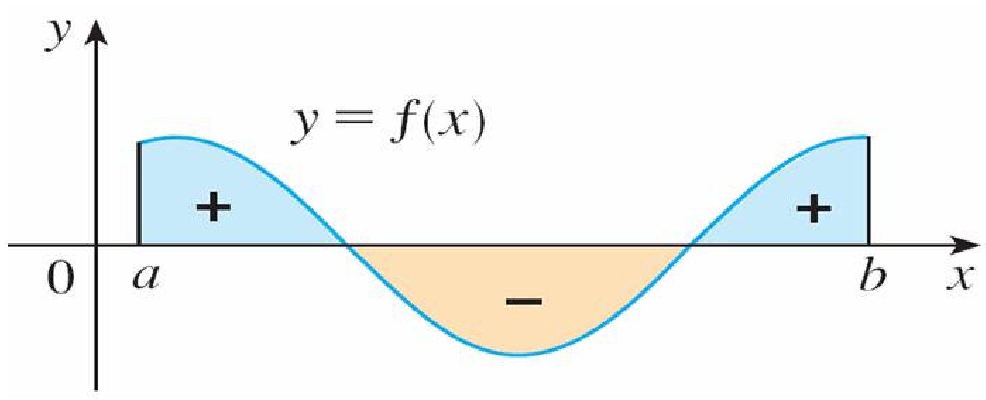
$\Sigma f(x_i^*) \Delta x$ is an approximation to the net area

The Definite Integral

When we take the limit of such Riemann sums, we get the situation illustrated here. A definite integral can be interpreted as a **net area**, that is, a difference of areas:

$$\int_a^b f(x)\,dx = A_1 - A_2$$

where A_1 is the area of the region above the x-axis and below the graph of f, and A_2 is the area of the region below the x-axis and above the graph of f



$\int_a^b f(x)\,dx$ is the net area

Integrability

Note: We have defined the definite integral for an integrable function, but not all functions are integrable. However, the most commonly occurring functions are in fact integrable:

$\boxed{3}$ **Theorem** If f is continuous on $[a, b]$, or if f has only a finite number of jump discontinuities, then f is integrable on $[a, b]$; that is, the definite integral $\int_a^b f(x)\,dx$ exists.

If *f* is integrable on [*a*, *b*], then the limit in Definition 2 exists and gives the same value no matter how we choose the sample points x_i^*

More on Notation

When Leibniz chose the notation for an integral, he chose the ingredients as reminders of the limiting process

In general, when we write

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

we replace $\lim \Sigma$ by $\int$, x_i^* by x, and Δx by dx

Properties of Definite Integrals

When we defined the definite integral $\int_a^b f(x)\,dx$, we implicitly assumed that $a < b$

But the definition as a limit of Riemann sums makes sense even if $a > b$

Notice that if we reverse *a* and *b*, then Δx changes from $(b - a)/n$ to $(a - b)/n$. Therefore

$$\boxed{\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx}$$

Properties of Definite Integrals

If $a = b$, then $\Delta x = 0$ and so

$$\int_a^a f(x)\,dx = 0$$

We now develop some basic properties of integrals that are useful to evaluate integrals in a simple manner. We assume that *f* and *g* are continuous functions

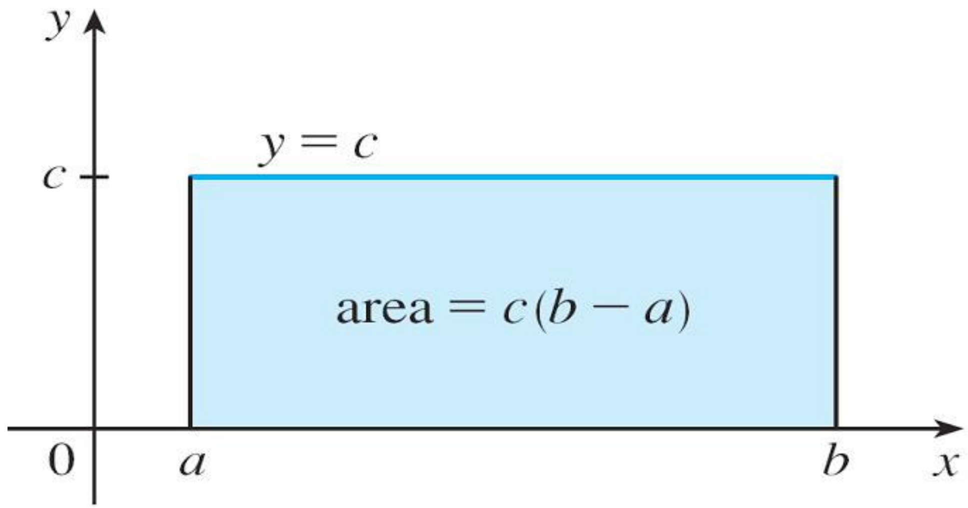
Properties of the Integral

1. $\int_a^b c\,dx = c(b-a)$, where c is any constant
2. $\int_a^b [f(x)+g(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$
3. $\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$, where c is any constant
4. $\int_a^b [f(x)-g(x)]\,dx = \int_a^b f(x)\,dx - \int_a^b g(x)\,dx$

Properties of Definite Integrals

Property 1 says that the integral of a constant function $f(x) = c$ is the constant times the length of the interval

If $c > 0$ and $a < b$, this is to be expected because $c(b - a)$ is the area of the shaded rectangle



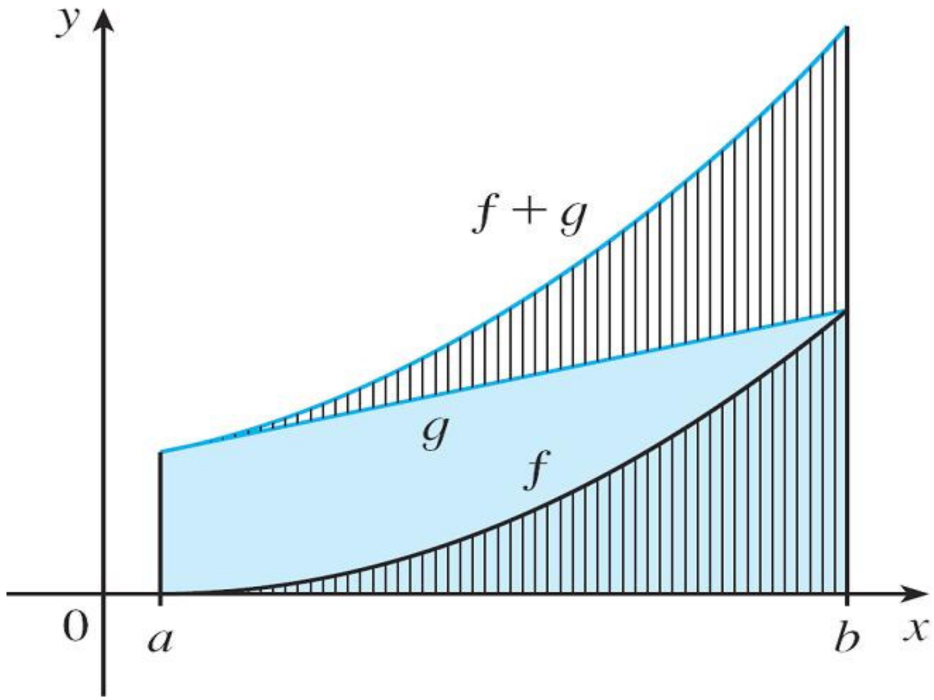
$$\int_a^b c\,dx = c(b - a)$$

Properties of Definite Integrals

Property 2 says that the integral of a sum is the sum of the integrals

For positive functions it says that the area under $f + g$ is the area under f plus the area under g

We can see why this is true from the geometrical view



$$\int_a^b [f(x) + g(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$

Properties of Definite Integrals

In general, Property 2 follows from the fact that the limit of a sum is the sum of the limits:

$$\int_a^b [f(x) + g(x)]\,dx = \lim_{n \to \infty} \sum_{i=1}^{n} [f(x_i) + g(x_i)]\,\Delta x$$

$$= \lim_{n \to \infty} \left[\sum_{i=1}^{n} f(x_i)\,\Delta x + \sum_{i=1}^{n} g(x_i)\,\Delta x \right]$$

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

$$= \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$

Properties of Definite Integrals

Property 3 can be proved in a similar manner and says that the integral of a constant times a function is the constant times the integral of the function

In other words, a constant (but *only* a constant) can be taken in front of an integral sign

Property 4 is proved by writing $f - g = f + (-g)$ and using Properties 2 and 3 with $c = -1$

Example

Use the properties of integrals to evaluate $\int_0^1 (4 + 3x^2)\,dx.$

Solution:

Using Properties 2 and 3 of integrals, we have

$$\int_0^1 (4 + 3x^2)\,dx = \int_0^1 4\,dx + \int_0^1 3x^2\,dx$$

$$= \int_0^1 4\,dx + 3\int_0^1 x^2\,dx$$

Example – *Solution*

cont'd

We know from Property 1 that

$$\int_0^1 4\,dx = 4(1 - 0)$$

and we have seen (when introducing the area problem) that

$$\int_0^1 x^2\,dx = \tfrac{1}{3}.$$

So

$$\int_0^1 (4 + 3x^2)\,dx = \int_0^1 4\,dx + 3\int_0^1 x^2\,dx$$

$$= 4 + 3 \cdot \tfrac{1}{3}$$

$$= 5$$

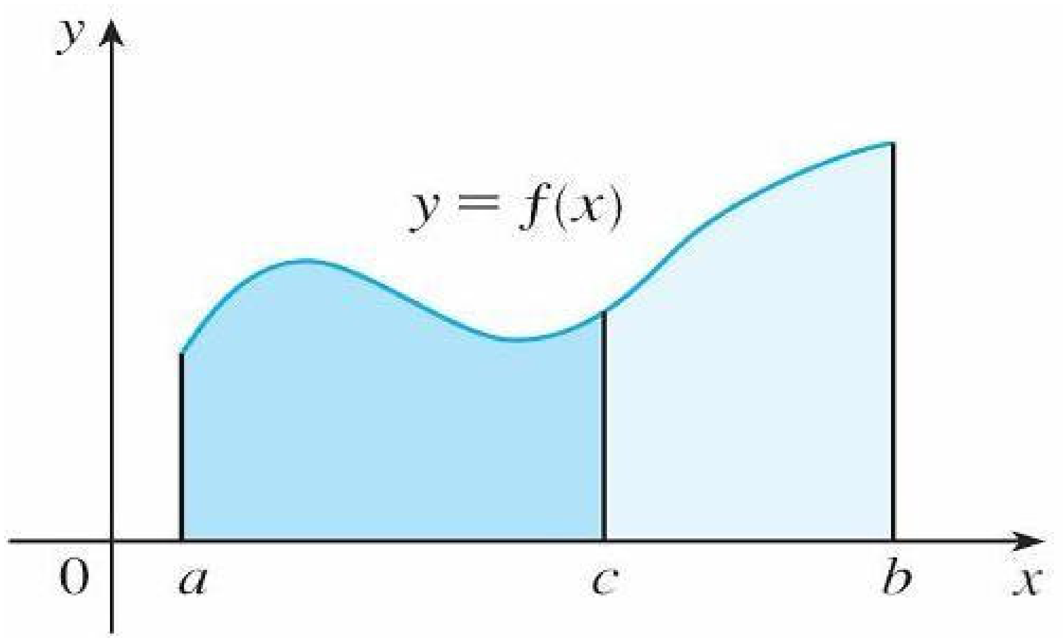
Properties of Definite Integrals

The next property tells us how to combine integrals of the same function over adjacent intervals:

5. $$\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx$$

Properties of Definite Integrals

For the case where $f(x) \geq 0$ and $a < c < b$ Property 5 can be seen from the geometric view: the area under $y = f(x)$ from a to c plus the area from c to b is equal to the total area from a to b



Properties of Definite Integrals

Properties 1–5 are true whether $a < b$, $a = b$, or $a > b$. The following properties, in which we compare sizes of functions and sizes of integrals, are true only if $a \leq b$

Comparison Properties of the Integral

6. If $f(x) \geq 0$ for $a \leq x \leq b$, then $\int_a^b f(x)\,dx \geq 0$.

7. If $f(x) \geq g(x)$ for $a \leq x \leq b$, then $\int_a^b f(x)\,dx \geq \int_a^b g(x)\,dx$.

8. If $m \leq f(x) \leq M$ for $a \leq x \leq b$, then

$$m(b - a) \leq \int_a^b f(x)\,dx \leq M(b - a)$$

Properties of Definite Integrals

If $f(x) \geq 0$, then $\int_a^b f(x)\,dx$ represents the area under the graph of f, so the geometric interpretation of Property 6 is simply that areas are positive (It also follows directly from the definition because all the quantities involved are positive)

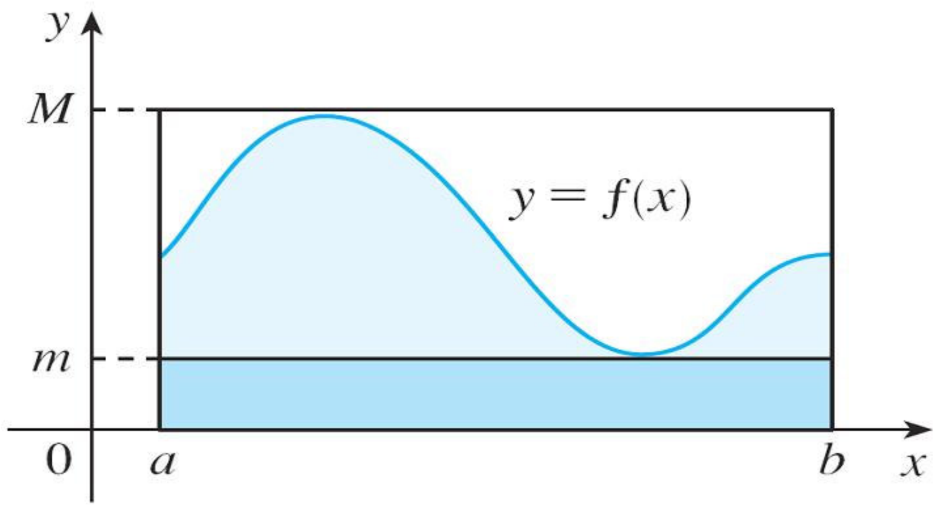
Property 7 says that a bigger function has a bigger integral

It follows from Properties 6 and 4 because $f - g \geq 0$

Properties of Definite Integrals

Property 8 is illustrated for the case where $f(x) \geq 0$

If f is continuous we could take m and M to be for example the absolute minimum and maximum values of f on [a, b]



In this case, it says that the area under the graph of f is greater than the area of the rectangle with height m and smaller than the area of the rectangle with height M

This is useful when we want a rough estimate of an integral

The Fundamental Theorem of Calculus

Now we will introduce the Fundamental Theorem of Calculus. It establishes a connection between the two branches of calculus: differential calculus and integral calculus

It gives the precise inverse relationship between the derivative and the integral

The Integral as a Function

Consider a function *g* defined as follows

$$g(x) = \int_a^x f(t)\,dt$$

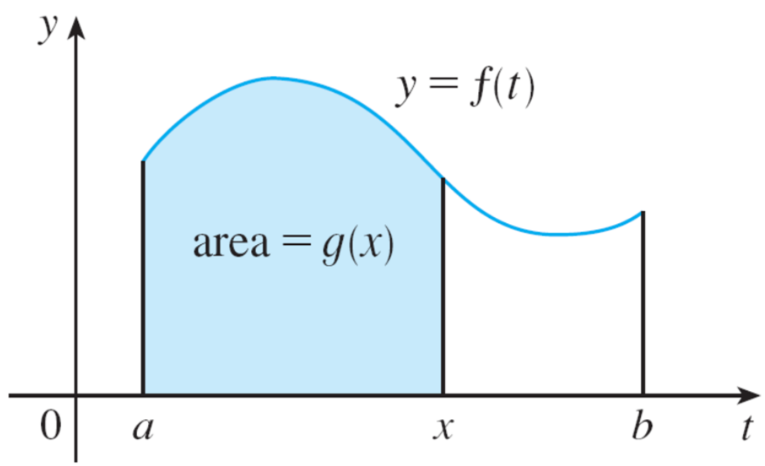
where *f* is a continuous function on [*a*, *b*] and *x* varies between *a* and *b*. Observe that *g* depends only on *x*, which appears as the variable upper limit in the integral

When *x* is a fixed number, then the integral $\int_a^x f(t)\,dt$ is a definite number

But if we then let *x* vary, the number $\int_a^x f(t)\,dt$ also varies and defines a function of *x* denoted by *g*(*x*)

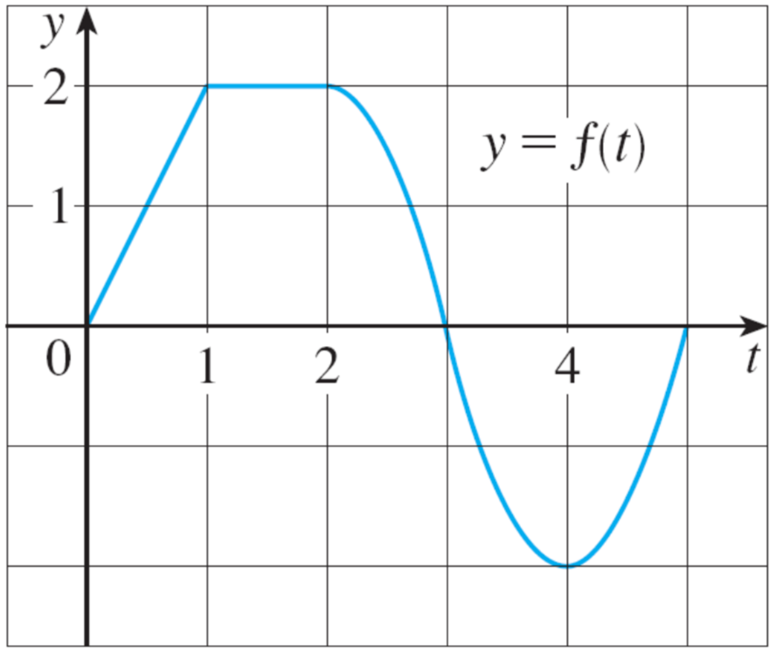
The Integral as a Function

If f happens to be a positive function, then $g(x)$ can be interpreted as the area under the graph of f from a to x, where x can vary from a to b (think of g as the "area so far" function)



Example 1

If f is the function whose graph is this, and $g(x) = \int_0^x f(t)\,dt$, find the values of $g(0)$, $g(1)$, $g(2)$, $g(3)$, $g(4)$, and $g(5)$, then sketch a rough graph of g

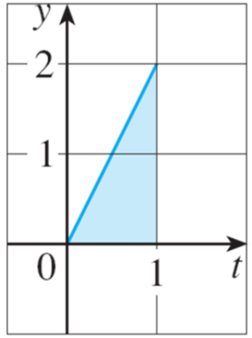


Example 1 – *Solution*

First we notice that $g(0) = \int_0^0 f(t)\,dt = 0$

We see that $g(1)$ is the area of a triangle:

$$g(1) = \int_0^1 f(t)\,dt = \tfrac{1}{2}(1 \cdot 2) = 1$$



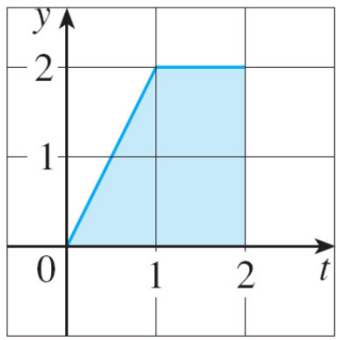
$g(1) = 1$

To find $g(2)$ we add to $g(1)$ the area of a rectangle:

$$g(2) = \int_0^2 f(t)\,dt$$

$$= \int_0^1 f(t)\,dt + \int_1^2 f(t)\,dt$$

$$= 1 + (1 \cdot 2) = 3$$



$g(2) = 3$

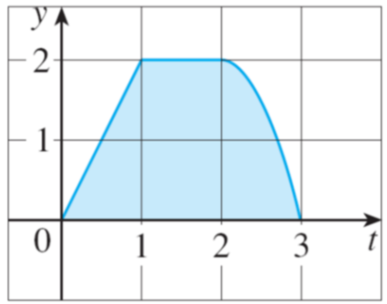
Example 1 – *Solution*

cont'd

If we estimate that the area under f from 2 to 3 is about 1.3, then

$$g(3) = g(2) + \int_2^3 f(t)\,dt$$

$$\approx 3 + 1.3 = 4.3$$

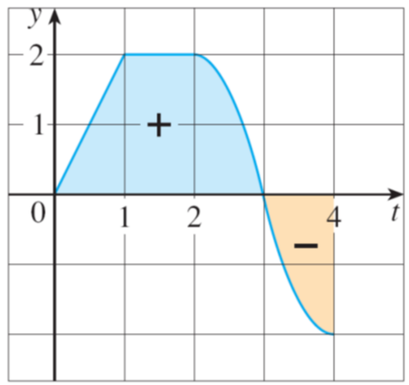


$g(3) \approx 4.3$

For $t > 3$, $f(t)$ is negative and so we start subtracting areas:

$$g(4) = g(3) + \int_3^4 f(t)\,dt$$

$$\approx 4.3 + (-1.3) = 3.0$$



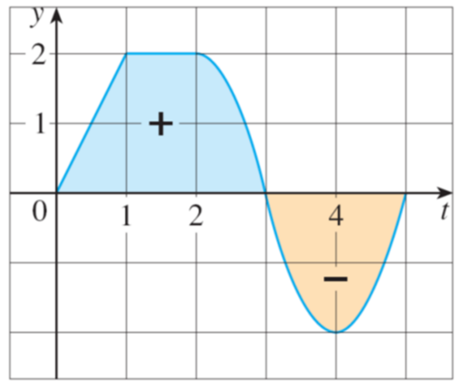
$g(4) \approx 3$

Example 1 – *Solution*

cont'd

$$g(5) = g(4) + \int_4^5 f(t)\,dt$$

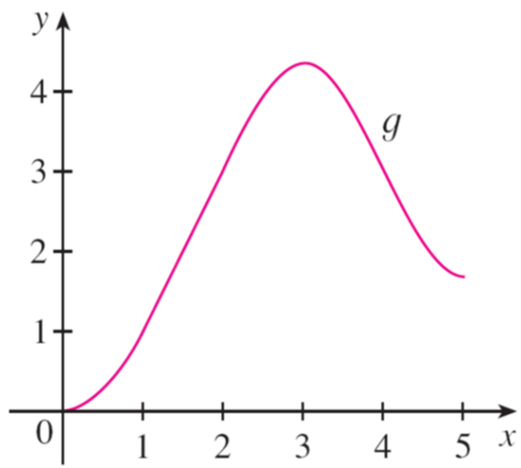
$\approx 3 + (–1.3) = 1.7$



$g(5) \approx 1.7$

We use these values to sketch the graph of *g*

Notice that, because *f*(*t*) is positive for $t < 3$, we keep adding area for $t < 3$ and so *g* is increasing up to $x = 3$, where it attains a maximum value. For $x > 3$, *g* decreases because *f*(*t*) is negative

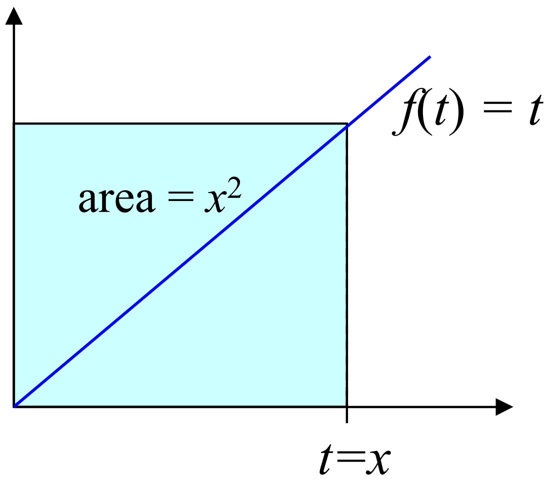


$g(x) = \int_a^x f(t)\,dt$

The Integral as an Antiderivative

Now we want to find what is the relationship between *f*(*t*) and *g*(*x*). If we take for example $f(t) = t$ and $a = 0$, then we have:

$$g(x) = \int_0^x t\,dt = \frac{x^2}{2}$$



Therefore $g'(x) = x$, that is, $g' = f$

In other words, if *g* is defined as the integral of *f* , then *g* turns out to be an antiderivative of *f*, at least in this case

The Fundamental Theorem of Calculus

The fact that this is true, even when *f* is not necessarily positive, is the first part of the Fundamental Theorem of Calculus

The Fundamental Theorem of Calculus, Part 1 If f is continuous on $[a, b]$, then the function g defined by

$$g(x) = \int_a^x f(t)\,dt \qquad a \leqslant x \leqslant b$$

is continuous on $[a, b]$ and differentiable on (a, b), and $g'(x) = f(x)$.

Using Leibniz notation for derivatives, we can write this theorem as

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$$

Roughly speaking, this equation says that if we first integrate *f* and then differentiate the result, we get back to the original function *f*

Example 2

Find the derivative of the function $g(x) = \int_0^x \sqrt{1 + t^2}\, dt.$

Solution:

Since $f(t) = \sqrt{1 + t^2}$ is continuous, Part 1 of the Fundamental Theorem of Calculus gives

$$g'(x) = \sqrt{1 + x^2}$$

The Fundamental Theorem of Calculus

The second part of the Fundamental Theorem of Calculus provides a much simpler method to evaluate integrals

The Fundamental Theorem of Calculus, Part 2 If f is continuous on $[a, b]$, then

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

where F is any antiderivative of f, that is, a function such that $F' = f$.

Example:

$$\int_1^e \frac{1}{x}\,dx = \ln e - \ln 1 = 1 - 0 = 1$$

Integral and Derivative as inverses

Bringing together the two parts of the Fundamental Theorem we have

The Fundamental Theorem of Calculus Suppose f is continuous on $[a, b]$.

1. If $g(x) = \int_a^x f(t)\,dt$, then $g'(x) = f(x)$.
2. $\int_a^b f(x)\,dx = F(b) - F(a)$, where F is any antiderivative of f, that is, $F' = f$.

We noted that Part 1 can be rewritten as

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$$

which says that if f is integrated and then the result is differentiated, we arrive back at the original function f

Integral and Derivative as inverses

Since $F'(x) = f(x)$, Part 2 can be rewritten as

$$\int_a^b F'(x)\,dx = F(b) - F(a)$$

This version says that if we take a function F, first differentiate it, and then integrate the result, we arrive back at the original function F, but in the form $F(b) - F(a)$

Taken together, the two parts of the Fundamental Theorem of Calculus say that differentiation and integration are inverse processes: each undoes what the other does

Why the fundamental theorem holds?

It is easy to see that $g(x) = \int_a^x f(t)dt$ is a function of *x*, but why *g'*(*x*) = *f*(*x*)? Lets compute the derivative of *g*

$$g'(x) = \lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\int_a^{x+\Delta x} f(t)dt - \int_a^x f(t)dt}{\Delta x} = \lim_{\Delta x \to 0} \frac{\int_x^{x+\Delta x} f(t)dt}{\Delta x}$$

Let *m* and *M* be the min and the max of *f* in [*x*, *x*+Δ*x*]

$$m\Delta x \le \int_x^{x+\Delta x} f(t)dt \le M\Delta x \qquad m \le \frac{\int_x^{x+\Delta x} f(t)dt}{\Delta x} \le M$$

So *g'*(*x*) is between *m* and *M*, and for Δ*x* → 0 it is *f*(*x*)

$$g'(x) = \lim_{\Delta x \to 0} \frac{\int_x^{x+\Delta x} f(t)dt}{\Delta x} = f(x)$$

Indefinite Integrals

Both parts of the Fundamental Theorem establish connections between antiderivatives and definite integrals. Part 1 says that if *f* is continuous, then $\int_a^x f(t)\,dt$ is an antiderivative of *f*. Part 2 says that $\int_a^b f(x)\,dx$ can be found by evaluating $F(b) - F(a)$, where *F* is an antiderivative of *f*

Because of the relation given by the Fundamental Theorem between antiderivatives and integrals, the notation $\int f(x)\,dx$ is traditionally used for an antiderivative of *f* and is called an **indefinite integral**

Always remember: ***a definite integral is a value*** (e.g. 3)***, an indefinite integral is a function*** (e.g. x^2)

Indefinite Integrals

The indefinite integral is the family of functions (all with the same behavior but vertically translated) which are anti-derivatives of a given function, or in other words the most general antiderivative of the given function

$$\int f(x)dx = g(x)+C \quad \text{with} \quad g'(x)=f(x)$$

For example, the indefinite integral of x^2 is

$$\int x^2\,dx = \frac{x^3}{3} + C \qquad \text{because} \qquad \frac{d}{dx}\left(\frac{x^3}{3} + C\right) = x^2$$

Table of Indefinite Integrals

1 Table of Indefinite Integrals

$$\int cf(x)\,dx = c\int f(x)\,dx \qquad \int [f(x)+g(x)]\,dx = \int f(x)\,dx + \int g(x)\,dx$$

$$\int k\,dx = kx + C$$

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C \quad (n \neq -1) \qquad \int \frac{1}{x}\,dx = \ln|x| + C$$

$$\int e^x\,dx = e^x + C \qquad \int a^x\,dx = \frac{a^x}{\ln a} + C$$

$$\int \sin x\,dx = -\cos x + C \qquad \int \cos x\,dx = \sin x + C$$

$$\int \sec^2 x\,dx = \tan x + C \qquad \int \csc^2 x\,dx = -\cot x + C$$

$$\int \sec x \tan x\,dx = \sec x + C \qquad \int \csc x \cot x\,dx = -\csc x + C$$

$$\int \frac{1}{x^2+1}\,dx = \tan^{-1} x + C \qquad \int \frac{1}{\sqrt{1-x^2}}\,dx = \sin^{-1} x + C$$

$$\int \sinh x\,dx = \cosh x + C \qquad \int \cosh x\,dx = \sinh x + C$$

Those formulas are valid on intervals where *f* is defined and continuous

Pay attention to the integration interval

It is a mistake to write

$$\int_{-1}^{3} \frac{1}{x^2}\,dx = -\frac{1}{x}\bigg|_{-1}^{3} = -\frac{1}{3} - 1 = -\frac{4}{3}$$

Because $f(x)=1/x^2$ is not continuous on [-1,3]

Thus we write

$$\int \frac{1}{x^2}\,dx = -\frac{1}{x} + C$$

with the understanding that it is valid on the interval $(0,\infty)$ or on the interval $(-\infty, 0)$

Example

Evaluate $\int_1^9 \frac{2t^2 + t^2\sqrt{t} - 1}{t^2}\, dt.$

Solution:

First we need to write the integrand in a simpler form by carrying out the division:

$$\int_1^9 \frac{2t^2 + t^2\sqrt{t} - 1}{t^2}\, dt = \int_1^9 (2 + t^{1/2} - t^{-2})\, dt$$

$$= 2t + \frac{t^{3/2}}{\frac{3}{2}} - \frac{t^{-1}}{-1}\Bigg]_1^9$$

$$= 2t + \tfrac{2}{3}t^{3/2} + \frac{1}{t}\Bigg]_1^9$$

Example – *Solution*

cont'd

$$= \left(2 \cdot 9 + \tfrac{2}{3} \cdot 9^{3/2} + \tfrac{1}{9}\right) - \left(2 \cdot 1 + \tfrac{2}{3} \cdot 1^{3/2} + \tfrac{1}{1}\right)$$

$$= 18 + 18 + \tfrac{1}{9} - 2 - \tfrac{2}{3} - 1$$

$$= 32\tfrac{4}{9}$$

The Substitution Rule

Because of the Fundamental Theorem, to evaluate integrals we need to compute antiderivatives

But our antidifferentiation formulas don't tell us how to evaluate integrals such as

$$\int 2x\sqrt{1 + x^2}\, dx$$

To find this integral we need to work on it: we change from the variable *x* to a new variable *u*

The Substitution Rule

To find $\int 2x\sqrt{1+x^2}\,dx$:

- We note that the function under the square root sign, $\mathrm{g(x)} = 1 + x^2$, has derivative $g'(x) = 2x$
- Both $g(x)$ and $g'(x)$ appear in the integral
- We can try to compute the indefinite integral, if we write it as:

$$\int f\big(g(x)\big)\, g'(x)\ dx$$

and recall the chain rule fomula for derivatives

The Substitution Rule

Observe that if $F' = f$, then

$$\int F'(g(x))\, g'(x)\, dx = F(g(x)) + C$$

because, by the Chain Rule,

$$\frac{d}{dx}\,[F(g(x))] = F'(g(x))g'(x)$$

If we make the "change of variable" or "substitution" $u = g(x)$, then we have

$$\int F'(g(x))g'(x)\, dx = F(g(x)) + C = F(u) + C = \int F'(u)\, du$$

or, writing $F' = f$, we get

$$\int f(g(x))g'(x)\, dx = \int f(u)\, du$$

The Substitution Rule

Thus we have proved the following rule

> $\boxed{4}$ **The Substitution Rule** If $u = g(x)$ is a differentiable function whose range is an interval I and f is continuous on I, then
>
> $$\int f(g(x))g'(x)\,dx = \int f(u)\,du$$

Notice that the Substitution Rule for integration was proved using the Chain Rule for differentiation

Notice also that, if $u = g(x)$, then $du = g'(x)\,dx$, so a way to remember the Substitution Rule is to think of dx and du as differentials

The Substitution Rule

Thus the Substitution Rule says: **It is permissible to operate with *dx* and *du* after integral signs as if they were differentials.**

To find $\int \sqrt{x^2+1}\ 2x\ dx$, we set $u = x^2 + 1$,

$du = g'(x)\,dx = 2x\,dx$, *and*

$$\int \sqrt{x^2+1}\ 2x\ dx = \int \sqrt{u}\,du = \int u^{1/2}\,du =$$

$$= \frac{2}{3}\,u^{3/2} + C = \frac{2}{3}\left(x^2+1\right)^{3/2} + C =$$

Example

Find $\int x^3 \cos(x^4 + 2)\, dx$

Solution:

We make the substitution $u = x^4 + 2$ because its differential is $du = 4x^3\, dx$, which, apart from the constant factor 4, occurs in the integral

Thus, using $x^3\, dx = \frac{1}{4}\, du$ and the Substitution Rule, we have

$$\int x^3 \cos(x^4 + 2)\, dx = \int \cos u \cdot \tfrac{1}{4}\, du$$

$$= \tfrac{1}{4} \int \cos u\, du$$

Example – *Solution*

cont'd

$$= \frac{1}{4} \sin u + C$$

$$= \frac{1}{4} \sin(x^4 + 2) + C$$

Notice that at the final stage we had to return to the original variable x

Definite integrals by substitution

When evaluating a *definite* integral by substitution, two methods are possible. One method is to evaluate the indefinite integral first and then use the Fundamental Theorem. For example

$$\int_0^4 \sqrt{2x+1}\,dx = \int \sqrt{2x+1}\,dx\,\Big]_0^4 =$$

$$½ \int 2\sqrt{2x+1}\,dx\,\Big]_0^4 = ½ \int \sqrt{u}\;2\,dx\,\Big]_0^4 = ½ \int u^{½}\,du\,\Big]_0^4 =$$

$$\frac{1}{2}\,\frac{u^{3/2}}{3/2}\Big]_0^4 = \frac{1}{2}\,\frac{2}{3}\,u^{3/2}\Big]_0^4 = \frac{1}{3}u^{3/2}\Big]_0^4 = \frac{1}{3}(2x+1)^{3/2}\Big]_0^4 =$$

$$= \tfrac{1}{3}(9)^{3/2} - \tfrac{1}{3}(1)^{3/2} = \tfrac{1}{3}(27 - 1) = \tfrac{26}{3}$$

Another method is to change the limits of integration when the variable is changed

Definite integrals by substitution

$\boxed{6}$ **The Substitution Rule for Definite Integrals** If g' is continuous on $[a, b]$ and f is continuous on the range of $u = g(x)$, then

$$\int_a^b f(g(x))\, g'(x)\, dx = \int_{g(a)}^{g(b)} f(u)\, du$$

Example

Evaluate $\int_0^4 \sqrt{2x + 1}\, dx$

Solution:

Let $u = 2x + 1$. Then $du = 2\, dx$, so $dx = \frac{1}{2}\, du$

To find the new limits of integration we note that

when $x = 0$, $u = 2(0) + 1 = 1$

and

when $x = 4$, $u = 2(4) + 1 = 9$

Example – *Solution*

cont'd

Therefore

$$\int_0^4 \sqrt{2x+1}\,dx = \int_1^9 \tfrac{1}{2}\sqrt{u}\;du$$

$$= \tfrac{1}{2}\cdot\tfrac{2}{3}u^{3/2}\Big]_1^9$$

$$= \tfrac{1}{3}(9^{3/2} - 1^{3/2})$$

$$= \tfrac{26}{3}$$

Observe that when using $\boxed{6}$ we do *not* return to the variable *x* after integrating. We simply evaluate the expression in *u* between the appropriate values of *u*

Integration of symmetric funct.

The following theorem uses the Substitution Rule for Definite Integrals $\boxed{6}$ to simplify the calculation of integrals of functions that possess symmetry properties

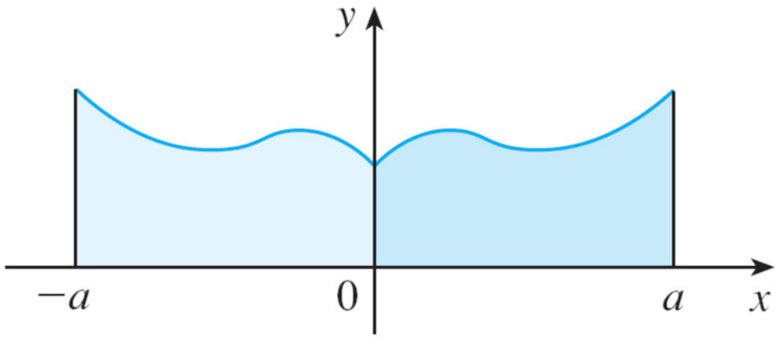
$\boxed{7}$ **Integrals of Symmetric Functions** Suppose f is continuous on $[-a, a]$.

(a) If f is even $[f(-x) = f(x)]$, then $\int_{-a}^{a} f(x)\,dx = 2\int_{0}^{a} f(x)\,dx$.

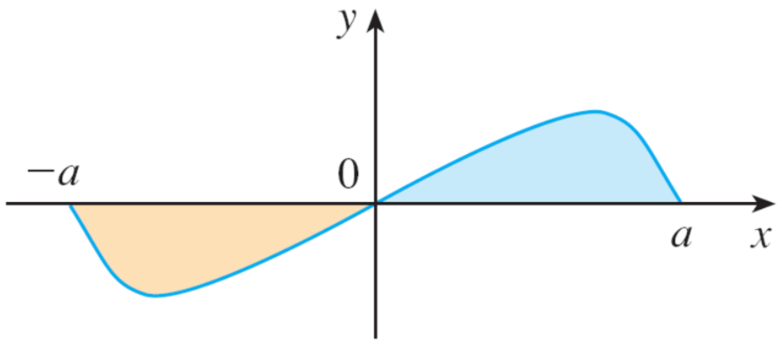
(b) If f is odd $[f(-x) = -f(x)]$, then $\int_{-a}^{a} f(x)\,dx = 0$.

Integration of symmetric funct.

From the graphical point of view:



(a) f even, $\int_{-a}^{a} f(x)\,dx = 2\int_{0}^{a} f(x)\,dx$



(b) f odd, $\int_{-a}^{a} f(x)\,dx = 0$

Part (a) says that the area under $y = f(x)$ from $-a$ to a is twice the area from 0 to a because of symmetry

Part (b) says the integral is 0, because it is the area above x-axis and below the curve minus the area below the axis and above the curve, so the areas cancel

Example 1

Since $f(x) = x^6 + 1$ satisfies $f(-x) = f(x)$, it is even and so

$$\int_{-2}^{2} (x^6 + 1)\,dx = 2\int_{0}^{2} (x^6 + 1)\,dx$$

$$= 2\Big[\tfrac{1}{7}x^7 + x\Big]_0^2$$

$$= 2\left(\tfrac{128}{7} + 2\right)$$

$$= \frac{284}{7}$$

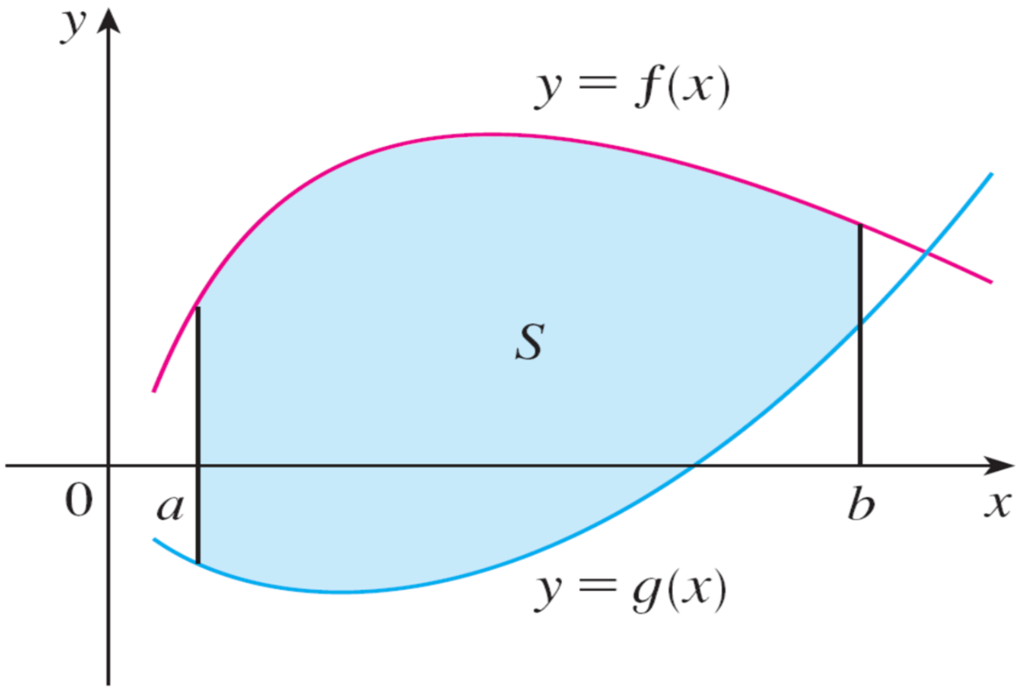
Example 2

Since $f(x) = (\tan x)/(1 + x^2 + x^4)$ satisfies $f(-x) = -f(x)$, it is odd and so

$$\int_{-1}^{1} \frac{\tan x}{1 + x^2 + x^4}\,dx = 0$$

Area between curves

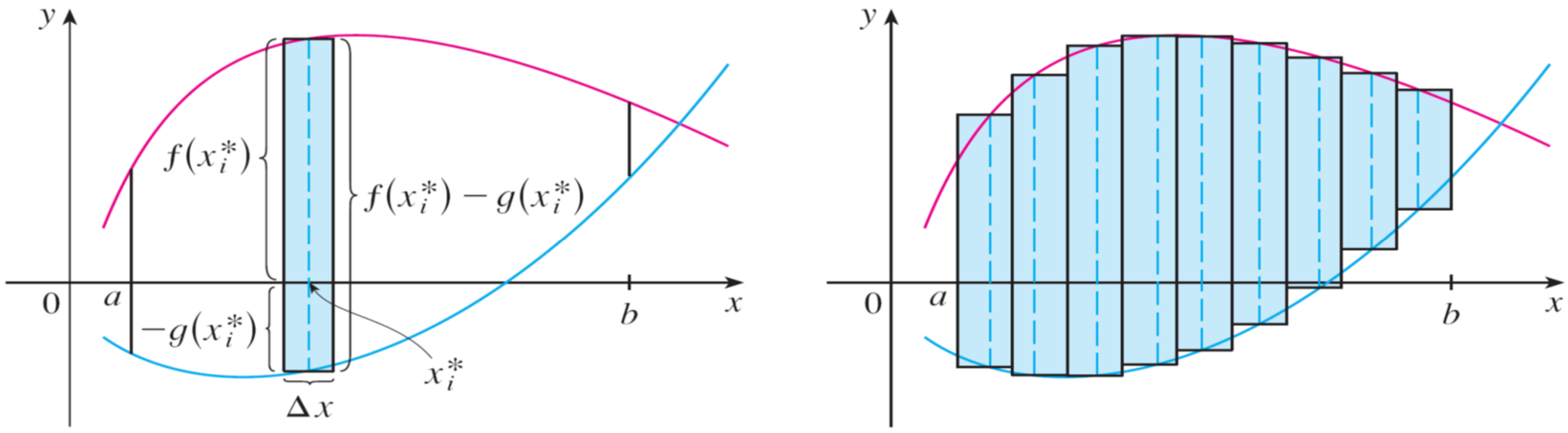
Consider the region S that lies between two curves $y = f(x)$ and $y = g(x)$ and between the vertical lines $x = a$ and $x = b$, where f and g are continuous functions and $f(x) \geq g(x)$ for all x in $[a, b]$



$S = \{(x, y) \mid a \leq x \leq b, g(x) \leq y \leq f(x)\}$

Area between curves

We divide S into n strips of equal width and then we approximate the ith strip by a rectangle with base Δx and height $f(x_i^*) - g(x_i^*)$



(a) Typical rectangle (b) Approximating rectangles

Area between curves

The Riemann sum

$$\sum_{i=1}^{n} [f(x_i^*) - g(x_i^*)]\,\Delta x$$

is therefore an approximation to the area of *S*

This approximation appears to become better and better as $n \to \infty$. Therefore we define the **area** *A* of the region *S* as the limiting value of the sum of the areas of these approximating rectangles

Area between curves

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

We recognize this limit as the definite integral of *f* – *g*

Therefore we have the following formula for area

$\boxed{2}$ The area A of the region bounded by the curves $y = f(x)$, $y = g(x)$, and the lines $x = a$, $x = b$, where f and g are continuous and $f(x) \geqslant g(x)$ for all x in $[a, b]$, is

$$A = \int_a^b [f(x) - g(x)]\, dx$$

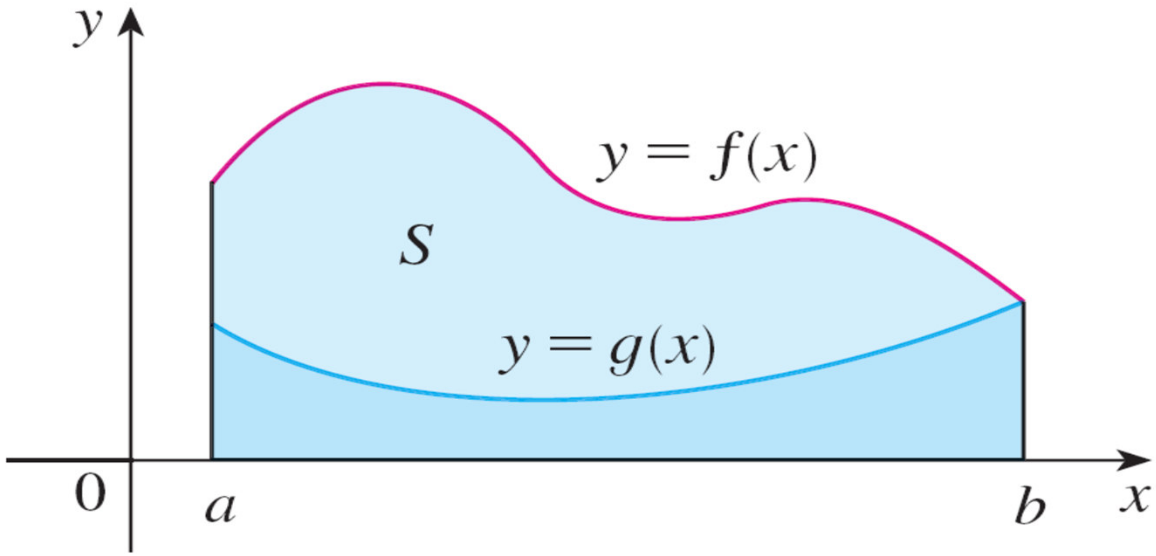
Notice that in the special case where $g(x) = 0$, *S* is just the region under the graph of *f*, as seen

Area between curves

In the case where both *f* and *g* are positive, we can visualize the situation

A = [area under $y = f(x)$] – [area under $y = g(x)$]

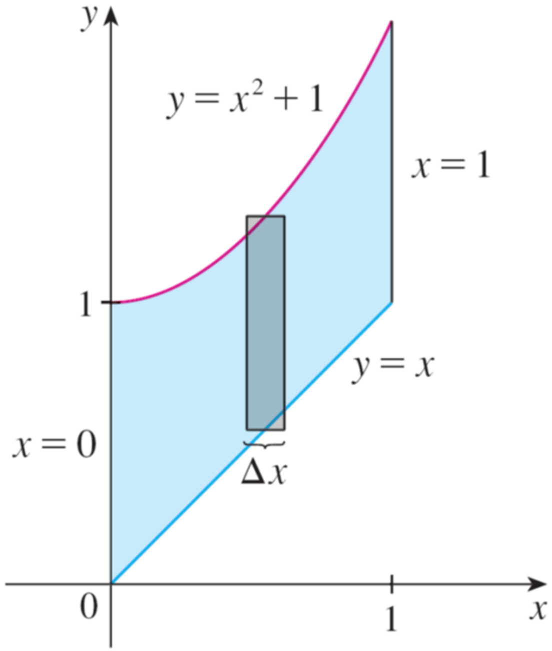
$$= \int_a^b f(x)\,dx - \int_a^b g(x)\,dx = \int_a^b [f(x) - g(x)]\,dx$$



$$A = \int_a^b f(x)\,dx - \int_a^b g(x)\,dx$$

Example

Find the area of the region bounded above by $y = e^x$, bounded below by $y = x$, and bounded on the sides by $x = 0$ and $x = 1$



Solution:

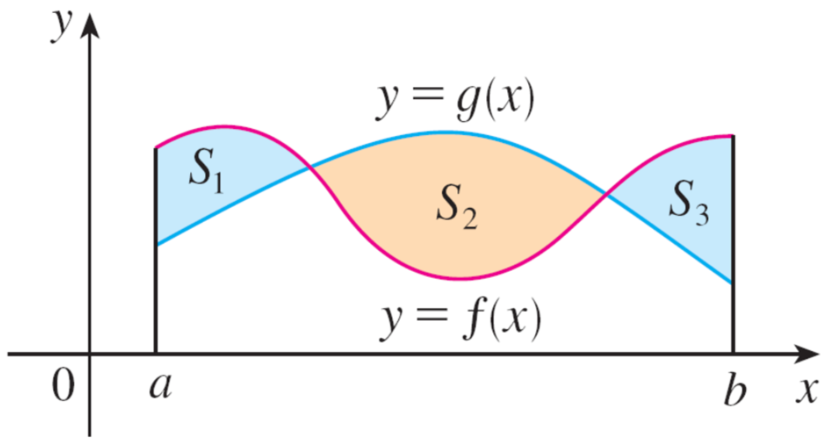
This is the region

we use the formula with $f(x) = e^x$, $g(x) = x$, $a = 0$, and $b = 1$:

$$A = \int_0^1 (e^x - x)\,dx = e^x - \tfrac{1}{2}x^2\Big]_0^1$$

$$= e - \tfrac{1}{2} - 1 = e - 1.5$$

Area between curves



If we want to find the area between the curves $y = f(x)$ and $y = g(x)$ where $f(x) \geq g(x)$ for some values x of but $g(x) \geq f(x)$ for other values of x, then we split the region S into several regions S_1, S_2, . . . with areas A_1, A_2, . . . We then compute the area of the region S as the sum of the areas of S_1, S_2, . . . that is $A = A_1 + A_2 + \ldots$

Since

$$| f(x) - g(x) | = \begin{cases} f(x) - g(x) & \text{when } f(x) \geq g(x) \\ g(x) - f(x) & \text{when } g(x) \geq f(x) \end{cases}$$

Area between curves

Therefore in the general case the area between curves is

$\boxed{3}$ The area between the curves $y = f(x)$ and $y = g(x)$ and between $x = a$ and $x = b$ is

$$A = \int_a^b |f(x) - g(x)|\, dx$$

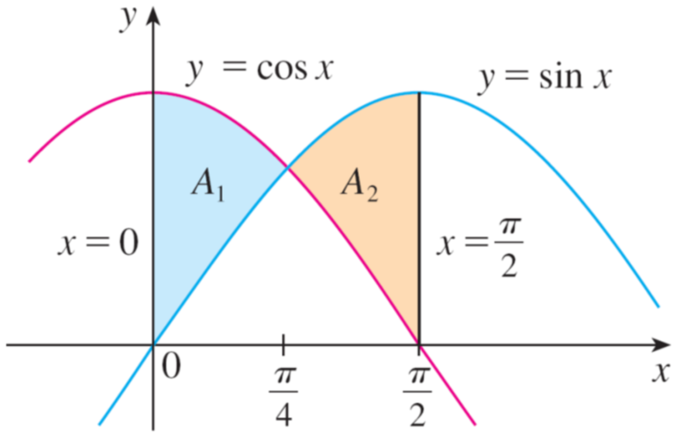
When evaluating this integral, however, we must still split it into integrals corresponding to A_1, A_2,……

Example

Find the area of the region bounded by the curves $y = \sin x$, $y = \cos x$, $x = 0$, and $x = \pi/2$

Solution:

The points of intersection occur when $\sin x = \cos x$, that is, when $x = \pi/4$ (since $0 \leq x \leq \pi/2$). The region is sketched here. Observe that $\cos x \geq \sin x$ when $0 \leq x \leq \pi/4$ but $\sin x \geq \cos x$ when $\pi/4 \leq x \leq \pi/2$.



Example – *Solution*

Therefore the required area is

$$A = \int_0^{\pi/2} |\cos x - \sin x|\, dx = A_1 + A_2$$

$$= \int_0^{\pi/4} (\cos x - \sin x)\, dx + \int_{\pi/4}^{\pi/2} (\sin x - \cos x)\, dx$$

$$= \Big[\sin x + \cos x\Big]_0^{\pi/4} + \Big[-\cos x - \sin x\Big]_{\pi/4}^{\pi/2}$$

$$= \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - 0 - 1\right) + \left(-0 - 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\right)$$

$$= 2\sqrt{2} - 2$$

Example – *Solution*

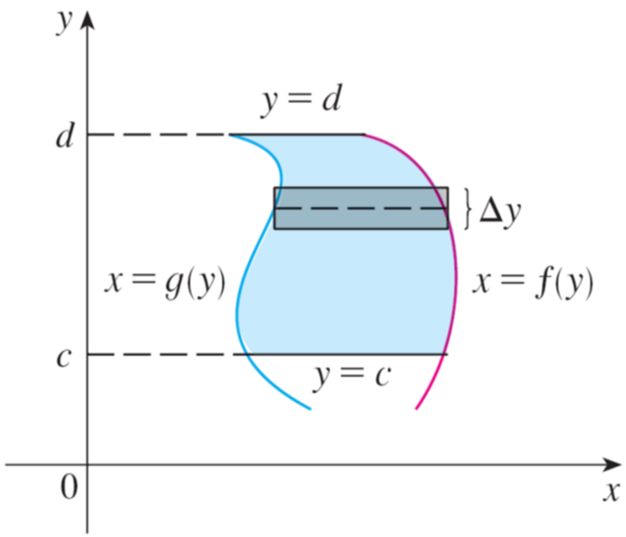
In this particular example we could have saved some work by noticing that the region is symmetric about $x = \pi/4$ and so

$$A = 2A_1 = 2 \int_0^{\pi/4} (\cos x - \sin x)\, dx$$

Area between curves

Some regions are best treated by regarding *x* as a function of *y*. If a region is bounded by curves with equations $x = f(y)$, $x = g(y)$, $y = c$, and $y = d$, where *f* and *g* are continuous and $f(y) \geq g(y)$ for $c \leq y \leq d$ then its area is

$$A = \int_c^d [f(y) - g(y)]\, dy$$



Integration by Parts

Every differentiation rule has a corresponding integration rule. For instance, the Substitution Rule for integration corresponds to the Chain Rule for differentiation. The rule that corresponds to the Product Rule for differentiation is called the rule for *integration by parts*

The Product Rule states that if f and g are differentiable functions, then

$$\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + g(x)f'(x)$$

Integration by Parts

In the notation for indefinite integrals this equation becomes

$$\int [f(x)g'(x) + g(x)f'(x)]\, dx = f(x)g(x)$$

or

$$\int f(x)g'(x)\, dx + \int g(x)f'(x)\, dx = f(x)g(x)$$

We can rearrange this equation as

$$\boxed{1} \qquad \boxed{\int f(x)g'(x)\, dx = f(x)g(x) - \int g(x)f'(x)\, dx}$$

Formula 1 is called the **formula for integration by parts**

Integration by Parts

It is perhaps easier to remember in the following notation

Let $u = f(x)$ and $v = g(x)$. Then the differentials are $du = f'(x)dx$ and $dv = g'(x)dx$, so, by the Substitution Rule, the formula for integration by parts becomes

$$\boxed{2} \qquad \boxed{\int u\, dv = uv - \int v\, du}$$

Example

Find $\int x \sin x \, dx$

Solution Using Formula 1:

Suppose we choose $f(x) = x$ and $g'(x) = \sin x$. Then $f'(x) = 1$ and $g(x) = -\cos x$.(For g we can choose *any* antiderivative of g'.) Thus, using Formula 1, we have

$$\begin{aligned}\int x \sin x \, dx &= f(x)g(x) - \int g(x)f'(x) \, dx \\ &= x(-\cos x) - \int (-\cos x) \, dx \\ &= -x \cos x + \int \cos x \, dx \\ &= -x \cos x + \sin x + C\end{aligned}$$

Example – *Solution*

cont'd

It's wise to check the answer by differentiating it. If we do so, we get $x \sin x$, as expected

Solution Using Formula 2:

Let

$$u = x \qquad dv = \sin x \, dx$$

Then

$$du = dx \qquad v = -\cos x$$

Example – *Solution*

cont'd

and so

$$\int x \sin x \, dx = \int \overbrace{x}^{u} \overbrace{\sin x \, dx}^{dv}$$

$$= \overbrace{x}^{u} \overbrace{(-\cos x)}^{v} - \int \overbrace{(-\cos x)}^{v} \overbrace{dx}^{du}$$

$$= -x \cos x + \int \cos x \, dx$$

$$= -x \cos x + \sin x + C$$

Integration by Parts of definite Integrals

If we combine the formula for integration by parts with Part 2 of Fundamental Theorem of Calculus, we can evaluate definite integrals by parts

Evaluating both sides of Formula 1 between *a* and *b*, assuming *f′* and *g′* are continuous, and using the Fundamental Theorem, we obtain

$$\boxed{6} \qquad \boxed{\int_a^b f(x)g'(x)\,dx = f(x)g(x)\Big]_a^b - \int_a^b g(x)f'(x)\,dx}$$

Can we integrate all elementary functions?

The functions that we have been studied here are called **elementary functions**

These are the polynomials, rational functions, power functions (x^a), exponential functions (a^x), logarithmic functions, trigonometric and inverse trigonometric functions, hyperbolic and inverse hyperbolic functions, and all functions that can be obtained from these by the five operations of addition, subtraction, multiplication, division, and composition

Can we integrate all elementary functions?

For instance, the function

$$f(x) = \sqrt{\frac{x^2 - 1}{x^3 + 2x - 1}} + \ln(\cosh x) - xe^{\sin 2x}$$

is quite complex but still an elementary function

If f is an elementary function, then f' is an elementary function. However, $\int f(x)\,dx$ may not be an elementary function.

Consider for example $f(x) = e^{x^2}$

Can we integrate all elementary functions?

Since f is continuous, its integral exists, and if we define the function F by

$$F(x) = \int_0^x e^{t^2}\, dt$$

then we know from Part 1 of the Fundamental Theorem of Calculus that

$$F'(x) = e^{x^2}$$

Thus $f(x) = e^{x^2}$ has an antiderivative F, but which is? It has been proved that F is **not an elementary function**

This means that no matter how hard we try, we will never succeed in evaluating $\int e^{x^2} dx$ in terms of the functions we know

Can we integrate all elementary functions?

The same can be said of the following integrals:

$$\int \frac{e^x}{x}\,dx \qquad \int \sin(x^2)\,dx \qquad \int \cos(e^x)\,dx$$

$$\int \sqrt{x^3+1}\,dx \qquad \int \frac{1}{\ln x}\,dx \qquad \int \frac{\sin x}{x}\,dx$$

In fact, the majority of elementary functions don't have elementary antiderivatives

You may be assured, though, that the integrals in the exercises are all elementary functions