Contents

CHAPTER 17

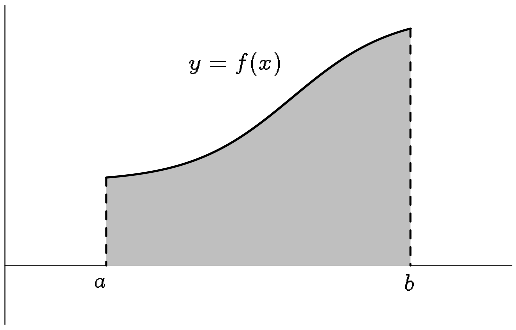
Definite Integrals: Area

A basic type of problem in calculus is to find the area under a curve. More precisely, we want to find the area bounded on the top by the graph of some function $y = f(x)$ with $f(x) \geq 0$, on the bottom by the x-axis, on the left by a vertical line $x = a$, and on the right by another vertical line $x = b$. We call this the "area under $f(x)$ between a and b." We also call it the **definite integral** of $f(x)$ from a to b.

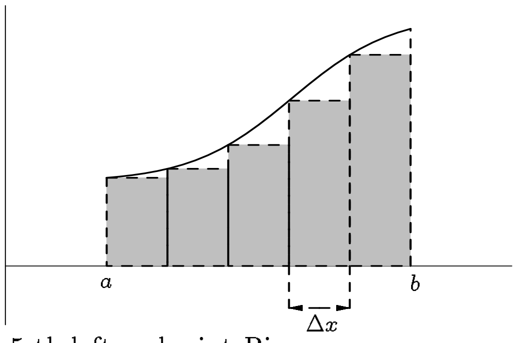
1. Riemann Sums

To understand the meaning of the definite integral it is instructive to discuss a method for obtaining an approximation to the true area area. This method is called the **Riemann sum approximation** (more precisely, the **n-th left endpoint Riemann sum approximation**, where n is some integer). The procedure is:

1. divide the interval from a to b into a certain number n of equal subintervals, each of length $\Delta x = (b - a)/n$;
2. draw a vertical line up to the curve at the endpoints of the subintervals;
3. over each subinterval form a rectangle by going up to the curve at the left endpoint and drawing a horizontal line to the right from that point on the curve;
4. take the sum of the areas of the n thin rectangles.



Area under the curve



5-th left endpoint Riemann sum

For example, if $n = 5$, so that $\Delta x = (b - a)/5$, then the heights of the 5 rectangles are:

$$f(a), \quad f(a + \Delta x), \quad f(a + 2\Delta x), \quad f(a + 3\Delta x), \ \text{and} f(a + 4\Delta x)$$

(see the drawing above). All of the rectangles have the same horizontal dimension Δx. Hence, the total area, which is the 5-th left endpoint Riemann sum, is equal to

$$f(a)\,\Delta x + f(a + \Delta x)\,\Delta x + f(a + 2\Delta x)\,\Delta x + f(a + 3\Delta x)\,\Delta x + f(a + 4\Delta x)\,\Delta x\,.$$

In general, the n-th left endpoint Riemann sum, obtained from the rectangles formed over n small subintervals, is equal to

$$f(a)\,\Delta x + f(a + \Delta x)\,\Delta x + f(a + 2\Delta x)\,\Delta x + \cdots + f(a + (n-1)\Delta x)\,\Delta x\,.$$

Using the summation symbol $\sum$, we can write this sum more compactly as follows:

$$(n\text{-th Riemann Sum}) = \sum_{i=0}^{n-1} f\,(a + i\,\Delta x)\;\Delta x\,,\ \text{where}\Delta x = \frac{(b-a)}{n}\,.$$

Here the summation notation tells us to take all integer values of i starting with $i = 0$ and ending with $i = n - 1$, put each value in place of i in $f(a+i\Delta x)\,\Delta x$, and take the sum.

EXAMPLE 17.1. Find the 5-th left endpoint Riemann sum for the definite integral of x^2 from 0 to 1, i.e., for the area under the parabola x^2 between 0 and 1.

We have $a = 0$, $b = 1$, $\Delta x = 0.2$, and the Riemann sum is

$$\sum_{i=0}^{4}(0.2i)^2 \cdot 0.2 = \left(0^2 + 0.2^2 + 0.4^2 + 0.6^2 + 0.8^2\right) 0.2 = 0.24\,.$$

Later we will see that the actual area is a little larger, namely 1/3. Notice that rather than multiplying each term by 0.2 and then summing, it is faster to factor out the term 0.2, add the remaining terms and then multiply by 0.2.

In general, if $f(x)$ is an increasing function on the interval (a, b), then the left endpoint Riemann sum will be a little less than the true area; but if $f(x)$ is a decreasing function on (a, b), then the left endpoint Riemann sum will be a little greater than the true area.

The larger n is, the closer the Riemann sum will be to the true area. We say that the area is "the limit of the Riemann sums as n approaches ∞" and we write it in symbols this way:

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} f(a + i\Delta x)\,\Delta x$$

In old books, the printed form of the letter "S" is "$\int$" and "infinitely small" values of Δx are written dx. Thus, a good symbol for the limit is

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

This number is called the "**definite integral** of $f(x)$ from a to b." The function $f(x)$ is called the **integrand**, a is called the **lower limit of integration** and b is called the **upper limit of integration**. The letter x is called the **variable of integration**.

In a definite integral it makes no difference what the variable of integration is called. In other words, the area under the function $y = f(z)$ between $z = a$ and $z = b$ means exactly the same thing as the area under the function $y = f(x)$ between $x = a$ and $x = b$. That is, both $\int_a^b f(z)dz$ and $\int_a^b f(x)dx$ are equal to the same thing because in a definite integral the variable of integration only holds a place, and does not affect the answer.

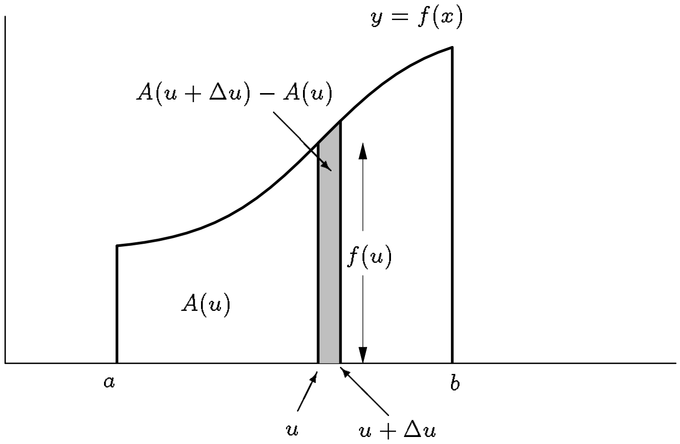
In the next section, we will discover that there is a relation between the <u>definite</u> integral and the **<u>in</u>**definite integral which you studied in the last chapter.

2. Computing the Area Under a Curve

We now describe how to find the exact area under $f(x)$ between a and b. To do this, we introduce an "area function" $A(u)$ as follows: for any u between a and b we set

$$A(u) = \text{area under } f(x) \text{ between } a \text{ and } u.$$

If we could find a formula for the function $A(u)$, then we would just substitute $u = b$



in order to obtain the area between a and b that we want. To find such a formula, we proceed indirectly. Instead of asking for $A(u)$ itself, we ask *how much the area function changes* when we go from u to $u+h$, and then we take the difference quotient $\big(A(u+h) - A(u)\big)/h$. By the definition of the area function,

$$\begin{aligned} A(u+h) - A(u) &= \text{area under } f(x) \text{ between } u \text{ and } u+h \\ &\approx \text{area of the rectangle above the interval } (u, u+h) \text{ of height } f(u) \\ &= h\, f(u). \end{aligned}$$

This approximate equality between the area under the curve and the area of the thin rectangle becomes very close to equality when h is small. Dividing through by h, we find that

$$\frac{A(u+h) - A(u)}{h} \approx f(u),$$

with the approximate equality becoming an equality in the limit as h goes to 0. That is, the limit of the difference quotient — the derivative of $A(u)$ — is equal to $f(u)$:

$$\frac{dA}{du} = f(u).$$

To see what this says, imagine that the area under the curve to the right of u is covered over by a card. The derivative of the area function $A(u)$ means the rate at which area is being uncovered as you move the card to the right. Clearly this depends on how wide the area is at the edge of the card — and this is precisely $f(u)$.

Once we know that dA/du is $f(u)$, this means that $A(u)$ is an antiderivative of $f(u)$. So suppose that we find the antiderivative $\int f(u)du = F(u) + C$, using the methods in the last section. Our function $A(u)$ must correspond to a certain value of C. Just as before, we need some initial information to determine this value. For the area function $A(u)$ the initial information is: when $u = a$ there is no area, i.e., $A(a) = 0$. Thus, $0 = A(a) = F(a) + C$, and hence $C = -F(a)$. We conclude that $A(u) = F(u) - F(a)$. This means that the total area under the curve between a and b is $A(b) = F(b) - F(a)$. In other words, *the area under $f(x)$ between a and b can be computed by finding an antiderivative of $f(x)$, substituting b and a in that antiderivative, and subtracting.* Using the definite integral notation, we conclude that

$$\text{Area under } f(x) \text{ between } a \text{ and } b = \int_a^b f(x)dx = F(x)\Big|_a^b = F(b) - F(a),$$

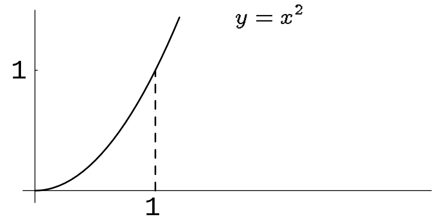
where $F(x)$ is an antiderivative of $f(x)$. Here we used the notation $F(x)\big|_a^b$ to mean "substitute b and a into $F(x)$ and then subtract."

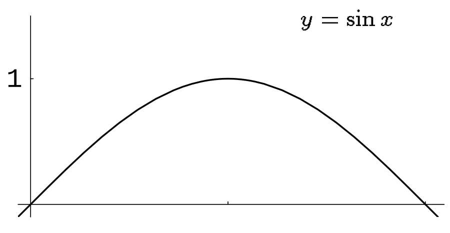
Note: You can tell whether an integral is a definite integral or an indefinite integral by whether or not there are limits of integration on the integral sign. *The answer to a definite integral problem is usually a specific value, not a function with an unknown constant C as in an indefinite integral problem. However, we will see many definite integrals with variable limits and these evaluate to functions of the limits, but without an arbitrary constant stepping in!*

EXAMPLE 17.2. Find the area under (a) $y = x^2$ between 0 and 1, and (b) one arch of the sine curve.

(a) $\int_0^1 x^2 dx = \frac{1}{3}x^3 \big|_0^1 = \frac{1}{3}(1^3 - 0^3) = \frac{1}{3}$.

(b) An arch of the sine curve is the region under $y = \sin x$ between $x = 0$ and $x = \pi$, so its area is $\int_0^\pi \sin x \, dx = -\cos x \big|_0^\pi = -\cos\pi - (-\cos 0) = -(-1) - (-1) = 2$.

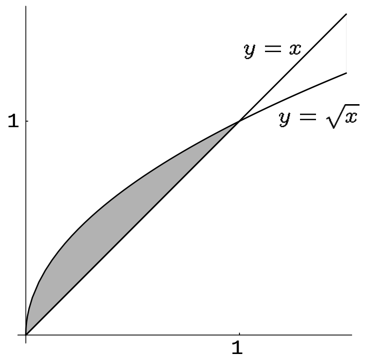




EXAMPLE 17.3. Find the area contained between the curve $y = \sqrt{x}$ and the line $y = x$.

The way to think of the area between two curves is: the area under the top curve between a and b minus the area under the bottom curve between a and b. Here a and b are the x-coordinates of the intersection points. In this case the square-root function and the line $y = x$ intersect at the origin and at the point (1,1). Between 0 and 1 the top curve is $\sqrt{x}$. So the area between them is

$$\int_0^1 \sqrt{x}\, dx - \int_0^1 x\, dx = \frac{2}{3}x^{3/2}\Big|_0^1 - \frac{1}{2}x^2\Big|_0^1 = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.$$



3. Substitution

In integrals evaluated using a u-substitution, as before you must change **EVERYTHING** from the old variable x (or t) to the new variable u. In the case of a definite integral, "everything" includes the limits of integration. You get the new limits of integration (the u-limits of integration) by asking the question: "as x goes from a to b the variable u goes from (what?) to (what?)." To answer this question, substitute a in place of x in the formula that gives u in terms of x; this gives the lower u-limit of integration. Then substitute b in place of x in the same formula to get the upper u-limit of integration. Here are some examples:

EXAMPLE 17.4. Find the area under (a) $y = \sqrt[3]{7x-6}$ between 1 and 2, and (b) $y = \sin\left(\frac{\pi x}{2}\right)\sqrt{1+\cos\left(\frac{\pi x}{2}\right)}$ between 0 and 1.

(a) In $\int_1^2 \sqrt[3]{7x-6}\,dx$ we make the substitution $u = 7x - 6$, so that $du = 7\,dx$ and $dx = \frac{1}{7}du$. This transforms the integral to

$$\int \sqrt[3]{u}\frac{1}{7}\,du = \frac{1}{7}\int u^{1/3}\,du\,.$$

What about the limits of integration? When x is equal to the lower limit of integration 1, we have $u = 7 \cdot 1 - 6 = 1$; whereas when x is equal to the upper limit of integration 2, we have $u = 7 \cdot 2 - 6 = 8$. Thus, our new limits of integration are 1 and 8. So the area is

$$\frac{1}{7}\int_1^8 u^{1/3}du = \frac{1}{7}\cdot\frac{3}{4}u^{4/3}\bigg|_1^8 = \frac{3}{28}\left(8^{4/3} - 1^{4/3}\right) = \frac{3}{28}(16-1) = \frac{45}{28}\,.$$

(b) If we make the substitution $u = 1 + \cos\left(\frac{\pi x}{2}\right)$, so that $du = -\frac{\pi}{2}\sin\left(\frac{\pi x}{2}\right)\,dx$, we see that as x goes from 0 to 1 the variable u goes from 2 to 1. We obtain:

$$\begin{aligned}\int_0^1 \sin\left(\frac{\pi x}{2}\right)\sqrt{1+\cos\left(\frac{\pi x}{2}\right)}dx &= \int_2^1 \sqrt{u}\left(-\frac{2}{\pi}\right)\,du \\ &= -\frac{2}{\pi}\cdot\frac{2}{3}u^{3/2}\bigg|_2^1 \\ &= -\frac{4}{3\pi}\left(1 - 2^{3/2}\right) = \frac{4}{3\pi}(2\sqrt{2}-1) = 0.776\,.\end{aligned}$$

4. Some Properties of Definite Integrals

1. Like indefinite integrals, they are linear. This means, for example, that any constant inside the integral can be pulled outside (as we did in Example 17.4(a) with the constant 1/7). Also, if you have the definite integral from a to b of $f(x) + g(x)$, it makes no difference whether you work with the sum all at once, or find the definite integral of $f(x)$ and $g(x)$ separately and then add the results. (The same observation applies to $f(x) - g(x)$.)
2. If c is a number between a and b, then the area between a and b is equal to the sum of the area between a and c and the area between c and b:
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx\,.$$
Actually, when you use the information from the next item below, we can get that the above formula is valid even when c is not between a, b.
3. If the lower limit of integration is greater than the upper limit, i.e., if b is on the bottom and a is on the top (so you are traversing the area under the curve from right to left rather than from left to right), then the effect is to put in a negative sign. This is because $\int_b^a f(x)dx = F(a) - F(b) = -(F(b) - F(a))$. In Example 17.4(b), after making the u-substitution we obtained an integral of the form $\int_2^1$, which was negative. (We also had a negative constant in front of the integral, so that our final answer was positive.)
4. If the function $f(x)$ is negative, then $\int_a^b f(x)dx$ is negative (assuming that $b > a$), i.e., area below the x-axis is counted as negative area. For example, if we compute $\int_0^{2\pi} \sin x\,dx$ (see Example 17.2(b)), we obtain $-\cos x\ \big|_0^{2\pi} = 0$ as our answer. This makes sense, because $\int_\pi^{2\pi} \sin x\,dx$ is an arch of the sine curve under the x-axis, and so has area -2. By property (2) above, we have $\int_0^{2\pi} \sin x\,dx = \int_0^{\pi} \sin x\,dx + \int_\pi^{2\pi} \sin x\,dx = 2 + (-2) = 0$, i.e., the two equal arches, one above and one below the x-axis, cancel.

5. Application to Degree-Days

The concept of "degree-days" arises when one wants to estimate heating costs. Suppose that in the winter you keep a building heated to a constant temperature of 65°F. Let $f(t)$ be the outside temperature. Then *the cost of heating the building between time $t = a$ and time $t = b$ is roughly proportional to the area between the line $y = 65$ and the temperature curve $y = f(t)$ from $t = a$ to $t = b$.* For example, if the temperature outside has the constant value $f(t) = D$, then the area between the two lines $y = 65$ and $y = D$ is simply $(65 - D)(b - a)$ (the units are units of temperature times time, i.e., degree-days). In that case all we're saying is that the heating cost is proportional to the temperature difference times the time interval. But when $f(t)$ is not constant we need to use definite integration to find this area.

EXAMPLE 17.5. Using the temperature function in Example 8.2 in the section on trig functions, find the number of degree-days in January in an average year in Seattle.

Our temperature function is $14\sin\left(\frac{2\pi}{360}(t - 120)\right) + 51$. Subtracting this from 65 and simplifying, we obtain $65 - (14\sin(\cdots) + 51) = 14\Big(1 - \sin(\cdots)\Big)$. Our time interval is 0 to

30 (recall that we supposed that every month has 30 days). So the number of degree-days is

$$\begin{aligned}
\int_0^{30} 14\left\{1-\sin\left(\frac{2\pi}{360}(t-120)\right)\right\}dt &= 14\left\{t-\left(\frac{360}{2\pi}\right)\left(-\cos\left(\frac{2\pi}{360}(t-120)\right)\right)\right\}\Big|_0^{30}\\
&= 14\left\{30+\frac{180}{\pi}\left(\cos\left(-\frac{1}{2}\pi\right)-\cos\left(-\frac{2}{3}\pi\right)\right)\right\}\\
&= 14\left\{30+\frac{180}{\pi}(0-(-0.5))\right\} = 821.
\end{aligned}$$

Thus, based on this information, every January we can get a rough idea of whether to expect high or low heating bills by comparing the total number of degree-days measured in the particular month with the norm for January, which is 821.

6. Velocity and Distance Traveled As Area

In Chapter 16 we saw that the velocity is the antiderivative of the acceleration. Another way to say this is that *velocity is the area under the acceleration graph.* More precisely, the area under $a(t)$ between $t = t_1$ and $t = t_2$ is equal to the amount the velocity increases during this time interval. In mathematical notation:

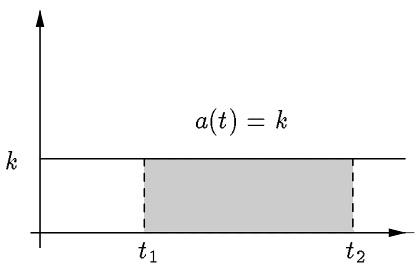
$$v(t_2) - v(t_1) = \int_{t_1}^{t_2} a(t)dt.$$

The graph at the left illustrates this in the special case when $a(t) = k$ is a constant function.

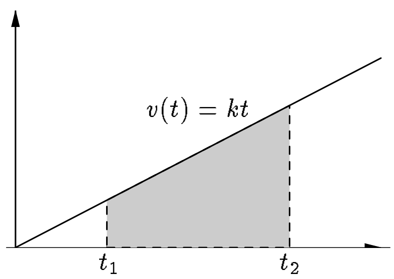
In Chapter 16 we also saw that the distance is the antiderivative of the velocity. In other words, *distance traveled is the area under the velocity graph.* More precisely, the area under $v(t)$ between $t = t_1$ and $t = t_2$ is equal to the distance traveled during this time interval:

$$s(t_2) - s(t_1) = \int_{t_1}^{t_2} v(t)dt.$$

The graph at the right illustrates this in the special case when $v(t) = kt$.



Shaded area is change in velocity during time interval.



Shaded area is distance traveled during time interval.

Note: By "distance traveled" we mean the *net* distance. In other words, the backward distance we travel when the velocity is negative (represented by area that is *beneath* the t-axis) counts with a negative sign, i.e., it cancels some of the positive distance we traveled earlier.

HOMEWORK 17.

1. In each case find the n-th left endpoint Riemann sum (for the indicated value of n), and compare with the true value of the area obtained by the anti-derivative method. In each case draw a picture showing the area in question and the rectangles in the Riemann sum, from which it should be clear why the true value is greater or less than the Riemann sum.

$$\text{(a)} \int_0^2 (4-x^2)dx,\ n=4; \quad \text{(b)} \int_0^2 (4-x^2)dx,\ n=8; \quad \text{(c)} \int_0^4 \sqrt{x}dx,\ n=4;$$
$$\text{(d)} \int_0^{\pi/2} \cos x\, dx,\ n=3; \quad \text{(e)} \int_0^{\pi/2} \cos x\, dx,\ n=6.$$

2. Compute: (a) $\sum_{i=1}^5 i^2$; (b) $\sum_{i=0}^3 \frac{1}{1+i}$; (c) $\sum_{i=0}^3 \cos(i\frac{\pi}{6})$.

3. Find the area under

(a) x^2+x between 1 and 4; (b) $\sqrt{4-3x}$ between 0 and 1;
(c) $1/\sqrt{4-3x}$ between 0 and 1; (d) $1/(2x+1)^3$ between 0 and 2;
(e) $\cos(\pi x)$ between $-1/2$ and $1/2$.

4. Find the area contained

(a) between the curve $y=\sqrt{1-x}$ and the coordinate axes;
(b) under both $\sin x$ and $\cos x$ in the region from $x=0$ to $x=\pi/2$;
(c) between the curve $y=\cos(\frac{\pi}{2}x)$ and the curve $y=x^2-1$;
(d) between the line $y=2$ and the branch of the curve $y=\sec^2 x$ lying between the asymptotes at $\pm\pi/2$.

5. Evaluate the following definite integrals:

$$\text{(a)} \int_1^3 \sqrt{4x-3}dx; \quad \text{(b)} \int_{-1}^1 (x^2+1)^2 dx; \quad \text{(c)} \int_{-1}^1 (x^3+x)^3 dx;$$
$$\text{(d)} \int_0^{0.5} \sin^3(\pi x)\cos(\pi x)dx; \quad \text{(e)} \int_0^{\pi/4} \frac{\sin x\, dx}{\cos^5 x}; \quad \text{(f)} \int_0^3 x\sqrt{9-x^2}dx.$$

6. Find the normal number of degree-days in Fairbanks (a) during an entire year, (b) during the coldest 2-month period (i.e., the time period extending 30 days before and 30 days after the coldest day), and (c) during the warmest 2-month period. Assume that the inside temperature is maintained at 65°F, and use the formula for the normal mean daily temperature that is given in Problem 6 of the homework set on trig functions.

7. In each case, find formulas and draw graphs for $a(t)$ and $v(t)$, and find the net distance traveled during the given time interval. (a) The initial velocity is zero. The acceleration starts at 0, increases linearly to 4 units/sec^2 during the first two seconds, remains constant at 4 units/sec^2 for $2 \le t \le 4$, and drops off linearly to 0 during the last 4-second period $4 \le t \le 8$. (b) The initial velocity is zero. The accleration is the minimum of $\sin t$ and $1-t$ for $0 \le t \le 1$. (Use Newton's method to find the point of intersection of these two functions, maintaining four significant figures of accuracy throughout the problem.)