CHAPTER 1

Numerical Integration: The Trapezoid Rule

Suppose we want to find an integral $\int_a^b f(x)dx$, i.e., the area under the curve $y = f(x)$ between $x = a$ and $x = b$, and suppose that we are unable to evaluate this integral by the antiderivative method. The purpose of this section is to describe an approximation method which, while not giving us the exact value of the integral, can be used to obtain a value which is accurate enough for practical purposes.

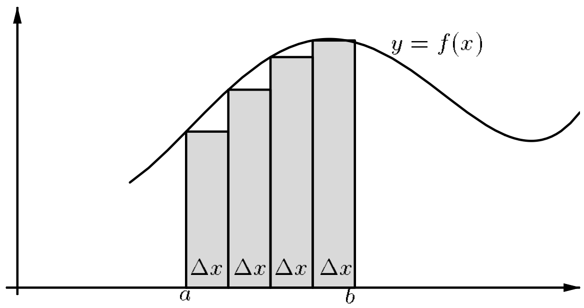
1. Riemann sums

The most basic way to approximate the area under a curve is by **Riemann sums**. A Riemann sum is the total area of a number of thin rectangles that we draw so as to fill out roughly the area under the curve. More precisely, we first describe the **n-th left endpoint Riemann sum**. We choose a number n, and divide the interval $[a, b]$ into n equal subintervals, each of length $(b - a)/n$. We let $\Delta x = (b - a)/n$ denote the length of a subinterval. If n is large, then we have very small subintervals, and in that case it turns out that the Riemann sum we later get will be close to the true value for the area under the curve.

Next, we draw a vertical line up to the curve at each endpoint of the subintervals. The first line has length $f(a)$, the second line has length $f(a + \Delta x)$, the third line has length $f(a + 2\Delta x)$, and so on, until we get to the last line, which has length $f(a + n\Delta x) = f(b)$. Now form a rectangle by drawing a little horizontal line of length Δx from left to right at the top of each of the vertical lines (except for the last one at b).

Each thin rectangle has height equal to the value of $f(x)$ at the *left* endpoint of the little subinterval that forms its base; and each rectangle has width equal to Δx. Thus, the area of the first thin rectangle is $\Delta x \cdot f(a)$, the area of the second rectangle is $\Delta x \cdot f(a + \Delta x)$, the area of the third is $\Delta x \cdot f(a + 2\Delta x)$, and so on; finally, the area of the n-th thin rectangle is

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



Left endpoint Riemann sum

Hence, the sum of the areas of the n rectangles is

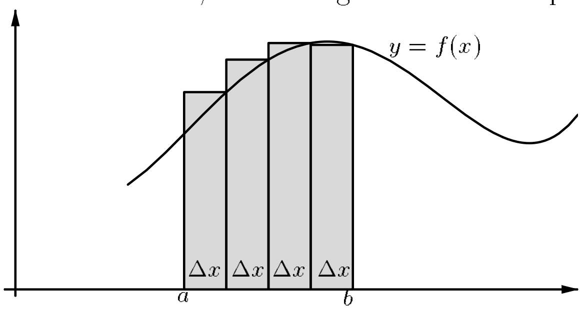
$$\Delta x\Big(f(a) + f(a + \Delta x) + f(a + 2\Delta x) + \cdots + f(a + (n - 1)\Delta x)\Big).$$

This "left endpoint Riemann sum" is approximately equal to the area under the curve. Using the summation notation, we can write:

$$\int_a^b f(x)dx \approx \Delta x \sum_{i=0}^{n-1} f(a + i\Delta x), \text{ where } \Delta x = \frac{b-a}{n}.$$

(To interpret the summation notation, let i start at the value written below the $\sum$, namely 0, and run through the numbers until you get to the top limit of the $\sum$, namely $n-1$; plug each value of i into $f(a + i\Delta x)$ and add the resulting terms.)

Alternately, we could use the *right endpoint Riemann sum*, where we take rectangles that go to the opposite side of the curve from the rectangles in the left endpoint Riemann sum. (In other words, if a rectangle in the left endpoint sum is a little below the curve — which happens when $y = f(x)$ is an increasing function — then the corresponding right endpoint rectangle will be a little above the curve; and if a left endpoint rectangle is a little above the curve — which happens when you have a decreasing function — then the corresponding right endpoint rectangle will be a little below the curve.) In the formula for the right endpoint Riemann sum, the f-values start at $f(a + \Delta x)$ and go through $f(b)$:



Right endpoint Riemann sum

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

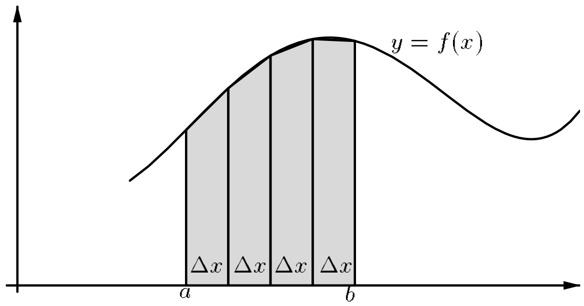
(Compare this with the formula for the left endpoint sum.)

The accuracy of an approximation is greater if we choose n large. Generally, neither the left endpoint nor the right endpoint Riemann sum method is very accurate. That is, to get a value close to the true area under the curve, you need to use very large n.

2. The Trapezoid Rule

A better method can be obtained by *averaging* the left and right endpoint Riemann sums. This is equivalent to drawing a slanted line across the top (see the figure below at the right).

The resulting sum is called the **trapezoid rule**. The trapezoid rule is more accurate than the left or right endpoint Riemann sum. It is easy to find a formula for the trapezoid rule by adding the left endpoint formula and the right endpoint formula and then dividing by 2. That is, we take the sum



The Trapezoidal Rule

$$\Delta x\Big(f(a)+f(a+\Delta x)+f(a+2\Delta x)+\cdots+f(a+(n-1)\Delta x)\Big)+$$
$$+\Delta x\Big(f(a+\Delta x)+f(a+2\Delta x)+\cdots+f(a+(n-1)\Delta x)+f(b)\Big)$$

and divide by 2. The result is the trapezoid rule:

$$\int_a^b f(x)dx \approx \Delta x\Big(\frac{1}{2}f(a)+f(a+\Delta x)+f(a+2\Delta x)+f(a+3\Delta x)+$$
$$+\cdots+f(a+(n-1)\Delta x)+\frac{1}{2}f(b)\Big).$$

Notice the difference with the Riemann sum formula: both $f(a)$ and $f(b)$ appear at the two extremes of the sum, each with coefficient $\frac{1}{2}$.

Example 1.1. Compare the approximations to $\int_1^2 x^{-2}dx$ using the left endpoint Riemann sum, the right endpoint Riemann sum, and the trapezoid rule — all with $n = 4$ subdivisions — with the true value of the integral, which is $-x^{-1}\,\big|_1^2 = 1/2$.

Here $f(x) = 1/x^2$, $\Delta x = 1/4$, and the 4 subintervals have endpoints 1, 5/4, 3/2, 7/4 and 2. We have

$$\text{left Riemann sum} = \frac{1}{4}\left(1+\frac{1}{(5/4)^2}+\frac{1}{(3/2)^2}+\frac{1}{(7/4)^2}\right) = 0.603;$$
$$\text{right Riemann sum} = \frac{1}{4}\left(\frac{1}{(5/4)^2}+\frac{1}{(3/2)^2}+\frac{1}{(7/4)^2}+\frac{1}{2^2}\right) = 0.415;$$
$$\text{trapezoid rule} = \frac{0.603+0.415}{2} = 0.509.$$

Note that the left (or right) Riemann sum is pretty inaccurate, while the trapezoid rule is a lot closer, even though the number $n = 4$ of subdivisions is rather small. We would get an even closer approximation to the true value of the integral if we used a larger value of n.

Example 1.2. Use the trapezoid rule with $n = 6$ subdivisions to find the length of one arch of the sine curve.

In Example 5 of the last section, we saw that an arch of the sine curve $y = f(x) = \sin x$ has length equal to $\int_0^\pi \sqrt{1+(f'(x))^2}dx = \int_0^\pi \sqrt{1+\cos^2 x}\,dx$. We now apply the trapezoid rule to this integral. When using the trapezoid rule, be careful about what function you are substituting for the $f(x)$. In this problem, because we are taking the arclength rather than the area under the sine curve, the function to substitute in the trapezoid rule is <u>not</u> our original function $\sin x$. The $f(x)$ in the trapezoid rule is $\sqrt{1+\cos^2 x}$. That is, it is a *different* $f(x)$ from the $f(x)$ we substituted into the arclength formula $\int_a^b \sqrt{1+(f'(x))^2}dx$.

Be careful not to get confused: the $f(x)$ in the arclength formula is not the $f(x)$ in the trapezoid rule.

Taking $n = 6$, we have $\Delta x = (\pi - 0)/6 = \pi/6$. Thus,

$$\int_0^\pi \sqrt{1+\cos^2 x}\,dx \approx \frac{\pi}{6}\Big(\frac{1}{2}\sqrt{1+\cos^2 0} + \sqrt{1+\cos^2(\pi/6)} + \sqrt{1+\cos^2(\pi/3)}$$
$$+ \sqrt{1+\cos^2(\pi/2)} + \sqrt{1+\cos^2(2\pi/3)} + \sqrt{1+\cos^2(5\pi/6)} + \frac{1}{2}\sqrt{1+\cos^2(\pi)}\Big)$$
$$= \frac{\pi}{6}\Big(\frac{1}{2}\sqrt{2} + \sqrt{7/4} + \sqrt{5/4} + 1 + \sqrt{5/4} + \sqrt{7/4} + \frac{1}{2}\sqrt{2}\Big) = 3.8202.$$

Notice that the straight-line distance from the beginning to the end of the arch is $\pi = 3.1416$; so our answer of 3.8202 for the arclength is reasonable. In doing arclength problems, it is a good idea to compare your answer with a rough estimate for the length in question: if the answer makes no sense, then go over your work again looking for careless errors.

3. Functions Obtained by Interpolating Experimental Data

There are two basic uses of numerical integration: (1) finding the definite integral of a function given by a formula for which you cannot find the antiderivative; and (2) finding the definite integral when the function is not given by a formula at all. The latter situation occurs frequently in practice. Suppose, for example, that you have a function $f(t)$ whose values come from instrument readings at regular time intervals from $t = a$ to $t = b$. You can graph $f(t)$ by tabulating and plotting the points corresponding to the instrument readings, and then drawing a smooth curve connecting these points. In this situation you can use the table of values along with the trapezoid rule to compute $\int_a^b f(t)dt$, or perhaps the integral of some other function made up from $f(t)$. It is generally best to choose the number n of subdivisions to equal the number of time intervals over which you took instrument readings.

Example 1.3. The table below shows the speedometer reading in miles/hour when a car tried to brake before an accident. Use the trapezoid rule to determine how far the car traveled in this time.

t (in seconds)	0	1	2	3	4	5	6	7	8
speed (mph)	70	65	59	52	44	36	27	18	10

First, we note that to convert miles/hour to feet/sec we must multiply by 5280 and divide by 3600. Thus, the distance in feet (using the trapezoid rule with $n = 8$ and $\Delta t = 1$) is

$$\frac{5280}{3600}\cdot\Big(\frac{1}{2}\cdot 70 + 65 + 59 + 52 + 44 + 36 + 27 + 18 + \frac{1}{2}\cdot 10\Big) = 500.13\,\text{ft}.$$

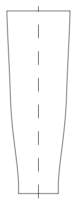
Since your table of speeds is given with only two significant figures of accuracy, in practical terms it is misleading to give the distance traveled to five significant figures. We should simply say that the car traveled 500 ft.

Example 1.4. A wine cork exactly 3 cm high has circular bottom of radius 0.542 cm and circular top of radius 0.859 cm. The following table shows the radius in centimeters of the cork's circular cross-section at intervals of 0.5 cm from bottom to top:

0.542	0.608	0.668	0.723	0.774	0.820	0.859

Find the volume of the cork.

The table shows the distance $g(y)$ from the axis of symmetry (which we take as the y-axis) to the curve that is rotated around the y-axis to get the cork, taken at the endpoints of the 6 intervals of length $\Delta y = 0.5$ from $y = 0$ to $y = 3$. Thus, it makes sense to apply the trapezoid rule with $n = 6$ subdivisions. The function being integrated is $g(y)^2$:

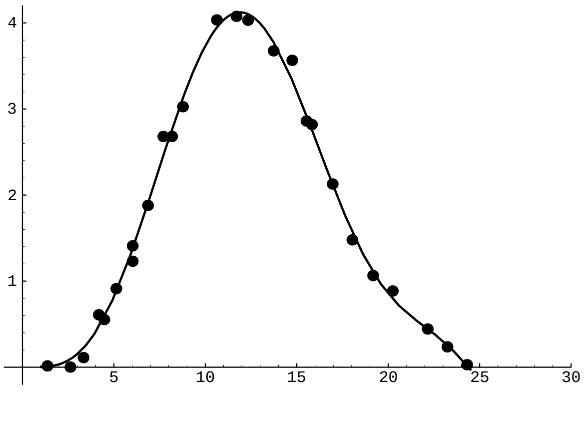


$$\begin{aligned}
\text{Vol} &= \pi \int_0^3 g(y)^2 \, dy \\
&\approx \pi \Delta y \Big(\frac{1}{2} g(0)^2 + g(0.5)^2 + g(1)^2 + g(1.5)^2 + g(2)^2 + g(2.5)^2 + \frac{1}{2} g(3)^2 \Big) \\
&= \frac{\pi}{2} \Big(\frac{1}{2} \cdot 0.542^2 + 0.608^2 + 0.668^2 + 0.723^2 + 0.774^2 + 0.820^2 + \frac{1}{2} \cdot 0.859^2 \Big) \\
&= 4.91 \, \text{cm}^3.
\end{aligned}$$

Example 1.5. (Measuring cardiac output.[1]) The rate R at which a person's heart pumps blood is called the person's **cardiac output**. The normal rate is about 5 liters/min, but the cardiac output can be quite different if the person has certain illnesses or is engaged in strenuous activity.

One way to measure R is to inject a known amount D of dye in a vein near the heart. The dyed blood then circulates once through the system in a little less than half a minute. Meanwhile, the concentration of dye in the blood, measured in milligrams per liter, can be monitored in the aorta, where the first of the dye will appear a few seconds after injection. The table below at the left shows concentration readings in mg/liter taken at 1-second intervals until all of the dye has passed the monitoring instrument. Here $t = 0$ is the instant when $D = 5.0$ mg of dye was injected into the blood. The graph below on the right is obtained by plotting the points in the table and interpolating a smooth curve between those points.

t	$C(t)$	t	$C(t)$
0	0.0	13	4.0
1	0.0	14	3.8
2	0.0	15	3.7
3	0.1	16	2.9
4	0.6	17	2.2
5	0.9	18	1.5
6	1.4	19	1.1
7	1.9	20	0.9
8	2.7	21	0.8
9	3.0	22	0.5
10	3.7	23	0.3
11	4.0	24	0.0
12	4.1		



We now derive a relationship between the cardiac output R, the amount of dye D, and the integral under the $C(t)$ curve. To do this, we define a function $D(t)$ to be the amount

[1]Taken from *Measuring Cardiac Output*, by B. Horelick and S. Koont, Project UMAP, Education Development Center, 1978.

of dye that has passed by the monitoring point within the first t seconds. Then dD/dt is the rate at which dye is passing. We have

$$\frac{dD}{dt} = R \cdot C(t),$$

i.e., *the number of mg per sec of dye is equal to the number of liters per second of blood times the number of mg of dye per liter of blood.* This is an important practical principle, which we shall encounter again when we treat mixing problems: *the rate (in mass per unit time) at which an impurity (such as dye) is flowing past a point is equal to the product of the rate (in liters per unit time) that the liquid is flowing past and the concentration (mass per liter).* For example, if 2 liters per second are going by, and each liter contains 3 mg of impurity, then 6 mg per second of impurity are going by.

Returning to the dye, we now have an expression for the derivative of our function $D(t)$. Then the total dye that passes — which is the amount D that we originally injected — is equal to the definite integral of that expression, i.e.,

$$\frac{dD}{dt} = R\,C(t) \text{ implies that } D = D(24) - D(0) = \int_0^{24} R\,C(t)\,dt = R\int_0^{24} C(t)\,dt,$$

where we took the constant R outside of the integral. Suppose we find the integral of $C(t)$. If we knew R, then this formula would tell us what D is. As it is, we do not know R, but we do know D. So we want to use this formula to solve for R in terms of D:

$$R = \frac{D}{\int_0^{24} C(t)\,dt} = \frac{5.0}{\int_0^{24} C(t)\,dt}.$$

It remains to find the integral of $C(t)$, which we do using the trapezoid rule with $n = 24$ subdivisions (so that $\Delta t = 1$ sec):

$$\int_0^{24} C(t)\,dt \approx \frac{1}{2}f(0) + f(1) + f(2) + f(3) + \cdots + f(23) + \frac{1}{2}f(24) = 44.1\,.$$

Finally, $R = 5.0/44.1 = 0.113$ liters/sec $= 6.8$ liters/min. Notice that if $f(a) = 0$ and $f(b) = 0$, as is the case in this example, then there is no difference between the left (or right) endpoint Riemann sum and the trapezoid rule.

HOMEWORK 1.

1. Suppose that at time $t = 10$ sec an object is traveling at 30.0 m/sec. Its acceleration is measured at two-second intervals until time $t = 20$, with the following results (the units of acceleration are m/sec^2):

t	10	12	14	16	18	20
$a(t)$	2.32	2.43	2.52	2.59	2.65	2.70

Use the trapezoid rule to find the velocity of the object at time $t = 20$.

2. Using $n = 4$ subdivisions, see how close to the true value of $\int_0^{\pi/2} \sin x \, dx$ you get using (a) the left endpoint Riemann sum, (b) the right endpoint Riemann sum, (c) the trapezoid rule, (d) Find how much closer you get to the true value using the trapezoid rule with 10 rather than 4 subdivisions.

3. Find the arclength of the curve $y = x^3$ from $x = 0$ to $x = 2$ using the trapezoid rule with $n = 4$ subdivisions.

4. Find the arclength of the ellipse $4x^2 + y^2 = 4$ between the points $(0, 2)$ and $(0.5, \sqrt{3})$ by (a) solving for y as a function of x and applying the trapezoid rule with $n = 5$ subdivisions to $\int_a^b \sqrt{1 + (f'(x))^2} dx$; and (b) writing the ellipse parametrically, finding the values of θ corresponding to the two points $(0, 2)$ and $(0.5, \sqrt{3})$, and applying the trapezoid rule with $n = 5$ subdivisions to the integral for the arclength of a parametrized curve.

5. In Example 3 of the last section, suppose you wanted to find the distance traveled by the point P during the 5 seconds from when the wheel leaves the roof to when it hits the ground. (a) Explain why, even with the trapezoid rule, you cannot expect to get an accurate answer with a small value of n such as $n = 10$. Now suppose that we are interested only in the first 0.04 sec. Use the trapezoid rule with $n = 4$ subdivisions to find the distance traveled during the first 0.04 sec by (b) the center of the wheel, and (c) the point P.

6. The table below shows the horizontal and vertical velocities of a particle at 1-second intervals for the first 8 seconds. Use the trapezoid rule to find (a) the straight-line distance between the particle's location at time $t = 0$ and at time $t = 8$, and (b) the actual distance traveled along its path between $t = 0$ and $t = 8$.

t	0	1	2	3	4	5	6	7	8
dx/dt	7.02	6.59	5.96	5.23	4.40	3.67	2.74	1.81	1.08
dy/dt	4.12	5.21	5.84	6.15	6.24	6.23	6.09	5.80	5.45