Introduction to the Riemann Integral

Tom Davis
tomrdavis@earthlink.net
http://www.geometer.org/mathcircles
July 17, 2011

Abstract

This article will explain the meaning of the Riemann integral for people who do not know anything about calculus. We will not prove many of the results, but we will try to give an intuitive idea not only of what it means, but why it is important.

1 Introduction

Suppose you drive a car at 60 miles per hour for one hour. How far do you go? To obtain the answer, simply multiply the two values: 60 miles per hour and 1 hour, to obtain 60 miles. But how often have you been on a car trip where the car never changed speed? Never, right? Real vehicles change speed all the time, they don't suddenly change speed, and that makes the simple multiplication of rate times time to obtain distance much less useful.

But to see how we might handle the problem of looking at travel with a variable rate of speed, lets consider how we might approach the following slightly more difficult problem:

Your car travels at 60 mph for one hour, then at 20 mph for another hour, and finally at 40 mph for two more hours. How far do you go? To solve it, you simply look at the time periods individually and add the distance traveled in all three:

$$D = 60 \times 1 + 20 \times 1 + 40 \times 2 = 160 \text{ miles}.$$

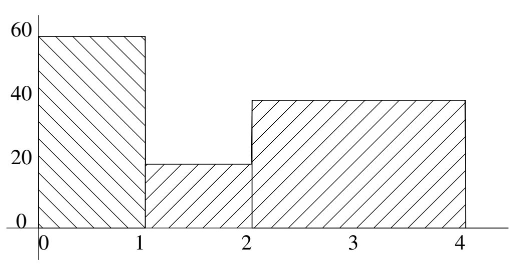

Figure 1: Traveling at a Variable Speed

Figure 1 illustrates the situation graphically: the horizontal axis is time, going from zero to four hours, and the vertical axis is the velocity. Notice that when you break the axis into three chunks the area of the rectangles corresponds to the distance: their height is the velocity and their width is the time. The area is the width times the height.

In a situation where the velocity actually varies continuously, the same principle works. The area under the velocity curve represents the total distance traveled. Thus we have one good example of a situation where it would be very useful to be able to calculate the area under a curve. That is exactly what the Riemann integral allows us to do.

2 Integral as Area

The most general form of the Riemann integral looks something like this:

$$\int_a^b f(x)\mathrm{d}x. \tag{1}$$

Often you will see the general function $f(x)$ or the variables a and b in Equation 2 replaced by some specific function or values, as in the following examples:

$$\int_0^5 x^3\mathrm{d}x \qquad \int_1^t \frac{\log x}{1+x}\mathrm{d}x \qquad \int_{-\infty}^{\infty} e^{-x^2}\mathrm{d}x.$$

In Equation 1 a and b are just numbers, and $f(x)$ stands for any function of a single variable. The funny integral sign ($\int$) and the "$\mathrm{d}x$" at the end basically go together in somewhat the same way that an open parenthesis "(" is usually matched with a ")" in mathematical equations.

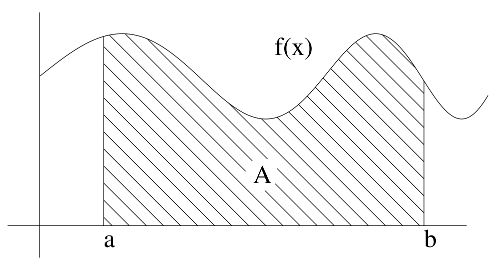

Figure 2: General Riemann Integral

If you draw the graph of the function $f(x)$ as in Figure 2 and a and b indicate points on the $x-axis$ as in the figure, then area of the shaded region indicated by A in the figure *is* the value of the integral in Equation 1. In other words, the Riemann integral represents the area under the curve $f(x)$ between the points a and b. For now, we'll assume that the curve $f(x)$ always lies above the x-axis, so there is no ambiguity about what "under the curve" means.

As a specific example, let's look at Figure 3. The specific function plotted is the parabola $y = x^2/2$. The area indicated in the figure is the area under the parabola from $x = 0$ to $x = 1$. The way to write that area as a Riemann integral is:

$$\int_0^1 \frac{x^2}{2}\mathrm{d}x.$$

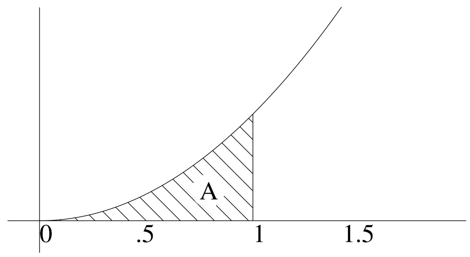

Figure 3: Area Under a Parabola

There are other types of integral besides the Riemann integral[1], but in this article, we will only deal with Riemann integration, so here we will use the terms “Riemann integral” and “integral” interchangeably. Similarly, “integration” and “Riemann integration” will mean the same thing.

3 Simple Integral Evaluations

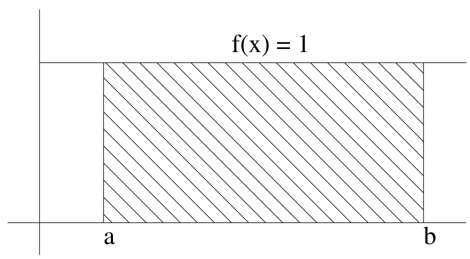

Figure 4: Area Under a Constant Function

There are a few simple functions whose integral we can evaluate using formulas we already know. The easiest is the constant function $f(x) = 1$. See Figure 4.

In this case the “curve” $f(x)$ is just a straight line one unit above the x-axis, so it is obvious that the area from a to b is just the area of the rectangle whose height is 1 and whose width is $b - a$. Thus we know that:

$$\int_a^b 1\mathrm{d}x = b - a.$$

Another simple function whose integral is easy to evaluate are the linear functions $f(x) = kx$, where k is a constant. Figure 5 shows an example where k is approximately equal to $1/2$. Basically, the area we are seeking is the area of the shaded trapezoid and one way to calculate that is as the difference of the areas of two triangles, one with base b and one with base a whose vertices are at the origin. The larger triangle has base b and height kb; the smaller one has base a and height ka. The areas of the two are $kb^2/2$ and

[1] Examples include the Stieltjes integral, the Darboux integral and the Lebesgue integral.

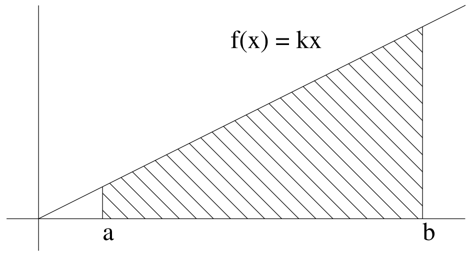

Figure 5: Area Under the Linear Function $f(x) = kx$

$ka^2/2$, respectively. Thus we can conclude that:

$$\int_a^b kx\mathrm{d}x = \frac{kb^2}{2} - \frac{ka^2}{2}. \tag{2}$$

4 Upper and Lower Sums

Without calculus, it is difficult to do many exact evaluations of Riemann integrals, and what we present here is a general method which, by itself, gives a good approximation of an integral. This method also allows us to put error bounds on that estimate, and if we are allowed to use the mathematical concept of a limit, we can, in many cases, provide an exact evaluation of a Riemann Integral.

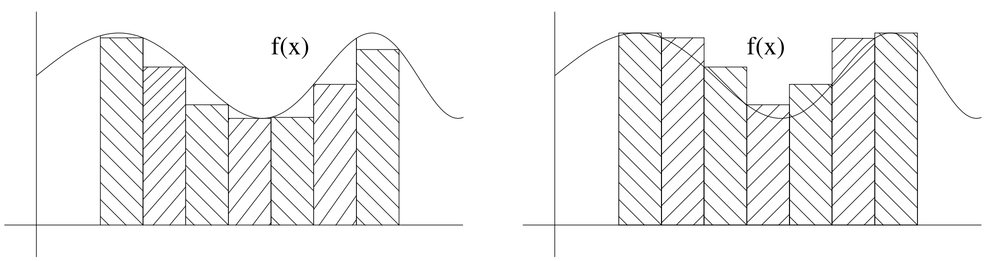

Figure 6: Upper and Lower Sums

Look at the two examples in Figure 6. We would like to determine the area under the curve $y = f(x)$ between two values of x and one method to obtain an estimate is to subdivide the x-axis into a number of equally-spaced intervals[2]. On each of the small intervals, the function $f(x)$ takes on a smallest and a largest value[3].

On the left, we form rectangles whose height is the minimum value of $f(x)$ on each of the small intervals. On the right, the height of the rectangles are the maximum values of $f(x)$ on each interval. If we add up all the areas of the rectangles on the left, since all of them are contained in the area under the curve, that

[2] They do not need to be equally-spaced for the general integral, but for the purposes of an introduction, this is a reasonable way to begin.

[3] Again, this may not be the case for all functions, but for now we will consider only "well-behaved" functions.

sum must be smaller than the area we are seeking. Similarly, on the right, all the rectangle areas include the area we are seeking, so the sum of those rectangle areas is greater than the true area. If we work out both sums, we don't know the true area, but we *do* now have a lower bound and an upper bound for the area.

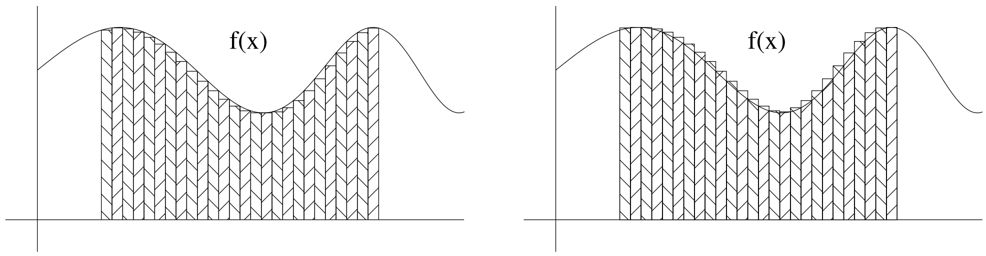

Figure 7: Upper and Lower Sums

In Figure 6 the part of the x-axis of interest was divided only into 7 pieces and as you can see, the area estimate is fairly crude. There is no reason to limit ourselves to 7 subdivisions; in fact, the more subdivisions we make, the better our upper and lower area estimates are likely to be. See Figure 7 as an example.

It would be nice to have a sort of formula to express these upper and lower sums rather than just a picture, so we will make a first pass at this. Imagine that the interval on the x-axis is divided into n equal-sized pieces, each of which has length Δx. For example, if the interval goes from a to b and there are n subdivisions, then $\Delta x = (b - a)/n$.

If we number the intervals from 1 to n, let $\underline{x}_1$ be the value of x where the function $f(x)$ obtains its minimum value in interval 1 (in other words, when the curve is closest to the x-axis), $\underline{x}_2$ where it obtains its minimum value in interval 2, and so on, so that $\underline{x}_i$ is where the function obtains its minimum in interval i. Similarly, let $\overline{x}_i$ be the x-value where $f(x)$ obtains its maximum in interval i.

Then the Riemann lower sum, which is a lower bound for the Riemann integral, is given by:

$$f(\underline{x}_1)\Delta x + f(\underline{x}_2)\Delta x + \cdots + f(\underline{x}_n)\Delta x. \tag{3}$$

If you know the summation notation, we can write Equation 3 as:

$$f(\underline{x}_1)\Delta x + f(\underline{x}_2)\Delta x + \cdots + f(\underline{x}_n)\Delta x = \sum_{i=1}^{n} f(\underline{x}_i)\Delta x.$$

Using exactly the same reasoning, we can conclude that the upper bound to the area, called the Riemann upper sum, is given by:

$$f(\overline{x}_1)\Delta x + f(\overline{x}_2)\Delta x + \cdots + f(\overline{x}_n)\Delta x = \sum_{i=1}^{n} f(\overline{x}_i)\Delta x.$$

So (using the summation notation for compactness), we obtain:

$$\sum_{i=1}^{n} f(\underline{x}_i)\Delta x \leq \int_a^b f(x)\mathrm{d}x \leq \sum_{i=1}^{n} f(\overline{x}_i)\Delta x,$$

and from the previous discussion, it appears that as n gets larger and larger, the lower and upper area approximations get better and better.

At this point, the reason for the integral notation becomes a little clearer. The "Σ" in the sum notation is the Greek letter sigma that corresponds so an English "S". The integral sign ($\int$) looks like an extended "S". Similarly, the Greek letter Delta (Δ) corresponds to the English "D", as in the "$\mathrm{d}x$" in our integrals.

5 Integral Evaluation Using Upper and Lower Sums

Going back to our simple examples, if $f(x) = 1$ is the constant function, as illustrated in Figure 4 we can see that the upper and lower sums will be the same, and will be equal to the exact value of the integral.

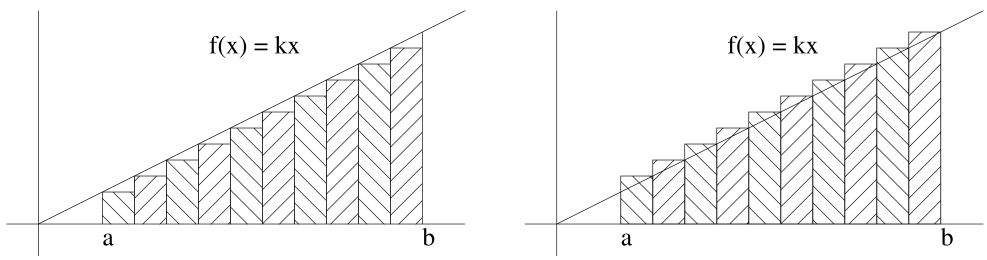

Figure 8: Upper and Lower Sums for $f(x) = kx$

A more interesting example is the linear function $f(x) = kx$ (and we will assume that $k > 0$). See Figure 8. We can see that $f(x) = kx$ obtains its minimum at the left end of each interval and obtains its maximum at the right end of each interval.

If the interval from a to b is divided into n subintervals, then the value at the left end of interval i is

$$f(a + (i-1)(b-a)/n) = k(a + (i-1)(b-a)/n)$$

and the value at the right of interval i is

$$f(a + i(b-a)/n) = k(a + i(b-a)/n)$$

.

The lower sum is thus:

$$ka\Delta x + k\big(a + \frac{b-a}{n}\big)\Delta x + k\big(a + \frac{2(b-a)}{n}\big)\Delta x + \cdots + k\big(a + \frac{(n-1)(b-a)}{n}\big)\Delta x,$$

and the upper sum is:

$$k\big(a + \frac{b-a}{n}\big)\Delta x + k\big(a + \frac{2(b-a)}{n}\big)\Delta x + \cdots + k\big(a + \frac{(n-1)(b-a)}{n}\big)\Delta x + k\big(a + \frac{n(b-a)}{n}\big)\Delta x.$$

A little algebra applied to the upper and lower sums yields a lower sum of:

$$k\big(na + \frac{b-a}{n}(1 + 2 + \cdots + (n-1))\big)\Delta x,$$

and an upper sum of:

$$k\Big(na + \frac{b-a}{n}(1+2+\cdots+n)\Big)\Delta x,$$

where $\Delta x = (b-a)/n$.

We know that $1+2+\cdots+(n-1) = n(n-1)/2$ and that $1+2+\cdots+n = n(n+1)/2$, so combining the equations above, and doing some algebra, we have:

$$\begin{aligned}
k\Big(\frac{(b-a)na}{n} + \frac{(b-a)^2}{n^2}\frac{(n-1)n}{2}\Big) \quad &\leq \quad \int_a^b kx\mathrm{d}x \quad \leq \quad k\Big(\frac{(b-a)na}{n} + \frac{(b-a)^2}{n^2}\frac{n(n+1)}{2}\Big) \\
k\Big((b-a)a + (b-a)^2\frac{(n-1)}{2n}\Big) \quad &\leq \quad \int_a^b kx\mathrm{d}x \quad \leq \quad k\Big((b-a)a + (b-a)^2\frac{(n+1)}{2n}\Big) \\
k\Big((b-a)a + \frac{(b-a)^2}{2}\big(1-\frac{1}{n}\big)\Big) \quad &\leq \quad \int_a^b kx\mathrm{d}x \quad \leq \quad k\Big((b-a)a + \frac{(b-a)^2}{2}\big(1+\frac{1}{n}\big)\Big) \\
k\Big(\big((b-a)a + \frac{(b-a)^2}{2}\big) - \frac{(b-a)^2}{2n}\Big) \quad &\leq \quad \int_a^b kx\mathrm{d}x \quad \leq \quad k\Big(\big((b-a)a + \frac{(b-a)^2}{2}\big) + \frac{(b-a)^2}{2n}\Big) \\
k\Big(\frac{b^2}{2} - \frac{a^2}{2}\Big) - \frac{k(b-a)^2}{2n} \quad &\leq \quad \int_a^b kx\mathrm{d}x \quad \leq \quad k\Big(\frac{b^2}{2} - \frac{a^2}{2}\Big) + \frac{k(b-a)^2}{2n}.
\end{aligned}$$

Notice in the final line above that the integral $\int_a^b kx\mathrm{d}x$ is bounded by the same value: $k(b^2/2 - a^2/2)$ but with a small additional term added to the right and subtracted from the left. Notice that this term has an n in the denominator, so as n gets larger and larger, the error terms get smaller and smaller, and can, in fact, be made to approach zero as closely as desired. We don't exactly have a formal proof here, but the calculation above should make it clear that the upper and lower sums squeeze the Riemann integral closer and closer to $k(b^2/2 - a^2/2)$, which is exactly the value we obtained using the areas of triangles in Section 3, Equation 2.

With even more complicated calculations (but similar to those used above), we can evaluate Riemann integrals for any polynomial function of x. In fact, it turns out that:

$$\int_a^b x^n\mathrm{d}x = \frac{b^{n+1}}{n+1} - \frac{a^{n+1}}{n+1}.$$

If you check back, you will notice that in *every* example we have worked out in detail so far, there corresponds to the function $f(x)$ another function $F(x)$ such that:

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a).$$

There are actually many possible values of F that do this. One way to think of $F(x)$ is as the area under the curve from 0 to x. Then the area under the curve from $x = a$ to $x = b$ is just $F(b) - F(a)$. When you take calculus, you will spend a semester figuring out how to determine $F(x)$ given $f(x)$ — it is not an easy problem.

6 A More Interesting Example

Using the observation in the last paragraph of the previous section, let's work out one more example of an integral, but this time, let's use a function where we don't know the answer. We will find the area under the parabola $f(x) = x^2$ between $x = a$ and $x = b$. We will simplify the problem by finding the area under the parabola from $x = 0$ to $x = b$ which we shall call $F(b)$. Using the observation, the area from $x = a$ to $x = b$ will be $F(b) - F(a)$.

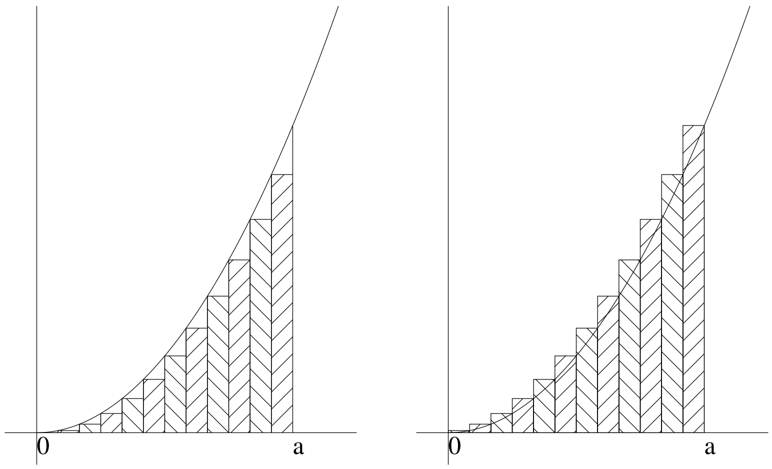

Figure 9: Area Under the Parabola $y = x^2$

The areas represented by the upper and lower Riemann sums are shown in Figure 9. Imagine that there are n subdivisions on the x-axis from 0 to a, each of length $\Delta x = a/n$. As with the linear function, the minimum heights are at the left of the intervals and maximums are at the right. A little work shows that the lower Riemann sum is:

$$(\Delta x)(0^2 + (\Delta x)^2 + (2\Delta x)^2 + \cdots + ((n-1)\Delta x)^2) = (\Delta x)\sum_{k=0}^{n-1}(k\Delta x)^2.$$

Similarly, the upper Riemann sum is:

$$(\Delta x)((\Delta x)^2 + (2\Delta x)^2 + \cdots + (n\Delta x)^2) = (\Delta x)\sum_{k=1}^{n}(k\Delta x)^2.$$

Factoring out all the Δx terms we have the following inequality:

$$(\Delta x)^3\sum_{k=0}^{n-1}k^2 \le \int_0^a x^2 \mathrm{d}x \le (\Delta x)^3\sum_{k=1}^{n}k^2. \qquad (4)$$

We know that:

$$\sum_{k=1}^{n}k^2 = \frac{n(n+1)(2n+1)}{6}$$

so

$$\sum_{k=0}^{n-1} k^2 = \frac{n(n-1)(2n-1)}{6}.$$

Thus, substituting a/n for Δx in the integral inequality Equation 4, we obtain:

$$\frac{a^3 n(n-1)(2n-1)}{6n^3} \leq \int_0^a x^2 \mathrm{d}x \leq \frac{a^3 n(n+1)(2n+1)}{6n^3}.$$

If we expand the fractions above we obtain:

$$\frac{a^3}{3} - \frac{a^3}{2n} + \frac{a^3}{6n^2} \leq \int_0^a x^2 \mathrm{d}x \leq \frac{a^3}{3} + \frac{a^3}{2n} + \frac{a^3}{6n^2}.$$

As in the example with the linear function, the term of interest is $a^3/3$. The other terms have n and n^2 in the denominator, and hence get tiny as n gets large. Thus:

$$\int_0^a x^2 \mathrm{d}x = \frac{a^3}{3},$$

From which we conclude that:

$$\int_a^b x^2 \mathrm{d}x = \frac{b^3}{3} - \frac{a^3}{3}.$$

7 Applications

First we'll take another look at traveling at a variable rate of speed r that we discussed in the introduction.

What really happens is that r changes with time, so instead of being a constant r, r is really a function of t: $r = r(t)$. Now the formula $d = rt$ doesn't work, but the following formula does calculate the distance traveled between times t_1 and t_2:

$$d = \int_{t_1}^{t_2} r(t)\mathrm{d}t. \tag{5}$$

Don't worry about the $\mathrm{d}t$ in place of $\mathrm{d}x$: imagine that you're using a t-axis instead of the usual x-axis.

The reason this works can be made clearer by looking at the Riemann upper and lower sums. On small enough intervals of time, the rate doesn't change much, so the lower bound for the distance traveled in an interval is the slowest rate in that interval multiplied by the time spent in that interval, and an upper bound can be similarly obtained using the the maximum rate in that interval. If we add together all these little chunks of distance which will approximate the distance of the entire trip, we obtain the upper and lower Riemann sums. As the number of intervals is increased, and their size made smaller, these lower and upper limits on the distance will squeeze together to a limiting value which is just the Riemann integral in Equation 5.

As a second example, work = force $\times$ distance. As in the previous example, with a constant force, you just multiply the two. But if force varies with distance, you need to integrate. The gravitational force between two objects with masses m_1 and m_2 is Gm_1m_2/s^2, where s is the distance between them. The number G is just the gravitational constant. To move them apart from distance s_1 to s_2, here is the work W required:

$$W = \int_{s_1}^{s_2} \frac{Gm_1m_2}{s^2}\mathrm{d}s. \tag{6}$$

If you think of the entire work as being broken up into little pieces of work over each little piece of the motion of distance Δs, you will see that the upper and lower Riemann sums provide lower and upper bounds on the work required, and, as in the previous example, the two squeeze together to a value which is indicated by the integral in Equation 6.

Similar integrals can be written down for many, many examples of physical calculations where the physical properties vary with time or distance.

8 General Properties of the Riemann Integral

In this brief introduction, there are many things we have not covered. Here are a few of them:

- What if the functions take on negative values? Both positive and negative values?
- What if $b < a$ in $\int_a^b f(x)\mathrm{d}x$?
- What if the bounds a and b are infinite or the function itself "goes to infinity"?
- What if the function is highly discontinuous?

In a calculus course, a lot of theorems are proved about the Riemann integral, and some are "obvious", like the following:

$$\int_a^b f(x)\mathrm{d}x + \int_b^c f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x.$$

If the above equation is not obvious, it simply states that the total area from a to c is the area from a to b added to the area from b to c, or in terms of Figure 10, that the total area that is indicated by the hashed lines is the sum of areas A and B.

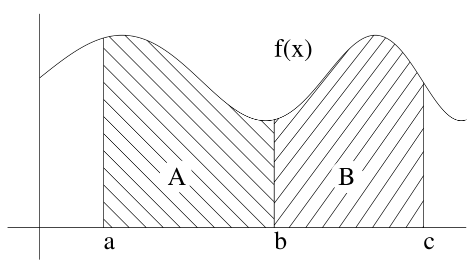

Figure 10: Sum of Areas

Here are some other properties that are "obvious" with the correct figures (which the author has not had time to produce as of this printing).

If k is a constant:

$$\int_a^b kf(x)\mathrm{d}x = k\int_a^b f(x)\mathrm{d}x.$$

For any two functions $f(x)$ and $g(x)$ that are integrable:

$$\int_a^b (f(x) \pm g(x))\mathrm{d}x = \int_a^b f(x)\mathrm{d}x \pm \int_a^b g(x)\mathrm{d}x.$$