The definition of the integral
Math 120 Calculus I
D Joyce, Fall 2013

There are two aspects of the definition that I want to emphasize. First, the degree of formality we use to make the definition. We'll start with an informal definition, then clarify and formalize it.

The second point is more technical. What should be the definition when the function is negative? We'll see, but to begin with, we'll only consider positive functions.

The integral as an area under a curve. Let f be a nonnegative function defined on an interval $[a, b]$, that is, for $x \in [a, b]$, $f(x) \geq 0$. We define the integral of f over the interval $[a, b]$ to be the area of the region above the x-axis, between the lines $x = a$ and $x = b$, and below the curve $y = f(x)$. The notation we use for this integral is due to Leibniz

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

To the degree that we understand what that area is, we understand integrals. This definition was sufficient for a long time as the definition of integrals, but when the concept of function was broadened in the 18th and 19th centuries, it became clear that a more precise definition was needed. There are strange functions (like the characteristic function of the rational numbers) where our preconceived notions of area don't apply. One solution to the problem was to use rectangular approximations.

Rectangular approximations. For a rectangular approximation, partition the interval $[a, b]$ into subintervals. If we partition it into n parts, the first subinterval goes from a to some number x_1, the next from x_1 to some number x_2, and so forth. It makes the notation come out nicer if we let x_0 equal a and x_n equal b.

$$a = x_0 \quad x_1 \quad x_2 \quad x_3 \qquad\qquad\qquad\qquad x_n = b$$
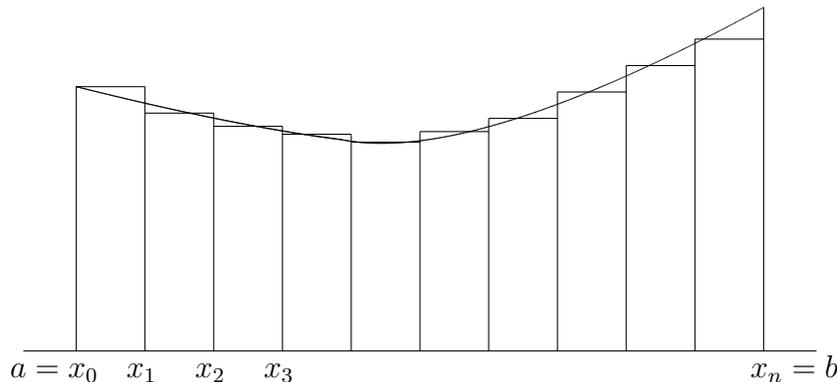
We'll use the symbol P stand for the whole partition, that is,

$$P = \{x_0, x_1, \ldots, x_n\}$$

where $a = x_0 < x_1 < \cdots < x_n = b$. Above each subinterval $[x_{k-1}, x_k]$ draw a rectangle whose height is the value of the function somewhere on that interval, that is, let the height be $y_k = f(c_k)$ where c_k is some number in $[x_{k-1}, x_k]$. That rectangle has area $y_k(x_k - x_{k-1})$. The width of that rectangle can be denoted $\Delta x_k = x_k - x_{k-1}$.

The sum of the areas of these rectangles gives an estimate to the area under the curve. That sum, called a *rectangular estimate* or *Riemann sum*, is

$$\sum_{k=1}^{n} y_k \Delta x_k = \sum_{k=1}^{n} y_k(x_k - x_{k-1}) = y_1(x_1 - x_0) + y_2(x_2 - x_1) + \cdots + y_n(x_n - x_{n-1})$$

There are lots of these rectangular estimates. Some will be close to the area under the curve, some will not. The *norm* of P, denoted $\|P\|$ is one measure of how good a partition P is. It's defined to be the largest width of any of the subintervals in P.

Riemann integrals and integrability. Riemann defined the value of an integral to be the limit of the Riemann sums as the norms of the partition approaches 0. Symbolically,

$$\int_a^b f(x)\, dx = \lim_{\|P\| \to 0} \sum_{k=1}^{n} f(c_k) \Delta x_k.$$

More precisely, $\int_a^b f(x)\, dx = J$ means for each $\epsilon > 0$, there exists a $\delta > 0$ so that for each partition P with $\|P\| < \delta$, and each choice of $c_k \in [x_{k-1}, x_k]$,

$$\left| \sum_{k=1}^{n} f(c_k)\, \Delta x_k - J \right| < \epsilon.$$

Should that limit not exist, then the function is said to be *nonintegrable.*

In all cases where the area under the curve is well understood, Riemann's definition agrees with that area. In some cases, like the characteristic function of the rational numbers, the limit doesn't exist, and those functions are nonintegrable.

Upper and lower rectangular estimates and Darboux's definition. An upper rectangular estimate occurs when the values y_k are chosen so that the they're greater than or equal to all values of $f(x)$ on the corresponding intervals $[x_{k-1}, x_k]$. In that case the union of the rectangles covers all the area under the curve. Therefore, each upper rectangular estimate is greater than or equal to the area of under the curve.

Likewise, a lower rectangular estimate occurs when the values y_k are chosen so that the they're less than or equal to all values of $f(x)$ on the corresponding intervals $[x_{k-1}, x_k]$. And in that case, the union of the rectangles is inside the area under the curve. Therefore, each lower rectangular estimate is less than or equal to the area under the curve.

Darboux defined the integral $\int_a^b f(x)\, dx$ in terms of upper and lower rectangular estimates. If there is only one number which is both less than or equal to all upper rectangular

estimates and greater than or equal to all lower rectangular estimates, then he defined the integral to be that number.

Darboux' and Riemann's definitions are equivalent. That is, a function f has a Riemann integral over an interval $[a, b]$ if and only if it has a Darboux integral over that interval, in which case the two integrals have the same value.

An advantage to Darboux' definition is that it does not explicitly involve limits.

Integrable functions. The functions we're interested in are all integrable. We won't prove the following three theorems, but they show that many functions are integrable.

Theorem. Continuous functions are integrable. In particular, differentiable functions, being continuous, are integrable.

Theorem. Bounded functions which are continuous except at finitely many discontinuites are integrable.

Theorem. Bounded monotone functions (either increasing or decreasing) are integrable.

Negative functions and signed area. Riemann's definition also gives a definition for the case when a function is negative, or when it's sometimes positive and sometimes negative.

If $f(c_k)$ is negative, then $f(c_k)(x_k - x_{k-1})$ is also negative. We can interpret that as a "signed area" of the rectangle which is the negation of the actual area of the rectangle.

When f is a negative function, then all the rectangles in a rectangular approximation are below the x-axis, so they all contribute the negations of their areas to the Riemann sum. Thus, the resulting integral has a negative value, the negation of the area below the x-axis and above the curve $y = f(x)$ and between the vertical lines $x = a$ and $x = b$.

When f is sometimes positive and sometimes negative, then the areas above the x-axis are counted positively, while the areas below the x-axis are counted negatively.

Thus, we can say integrals give *signed areas* in general.

Math 120 Home Page at `http://math.clarku.edu/~djoyce/ma120/`