A Darboux-Style Definition of the Lebesgue Integral

Please read, or at least skim, the complementary handout *A Precise Definition of the Definite Integral* before reading this one. It describes a version of the Riemann integral Jean-Gaston Darboux (1842-1917) that is a little simpler than the original version due to Bernhard Riemann (1826-1866). By way of comparison and contrast, this handout gives a definition of the Lebesgue integral in a form as similar as possible to Darboux's version of the Riemann integral.

Recall that, ntuitively, the definite integral $\int_a^b f(x)\,dx$ represents the area between $y = f(x)$ and $y = 0$ for $a \leq x \leq b$, weighted so that area below $y = 0$ is subtracted and area above $y = 0$ is added. Riemann's basic idea for defining the definite integral was to approximate the area between $y = f(x)$ and $y = 0$ by rectangles; as one makes the rectangles narrower and increases their number, one should get better approximations to the area in question. This approach works pretty well as long as the interval in question is finite, the function in question is bounded on the interval, and has only finitely many discontinuities on the interval. It doesn't work properly in many other cases,. For example the function $f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$ is not Riemann-integrable on any interval $[a, b]$. (Well, unless $a = b]$ ...) The Lebesgue integral handles a wider rangle of functions properly than the Riemann integral, including the function just mentioned.

Preliminaries. We'll need to make a few subsidiary definitions and set up some terminology and notation. The first is actually a basic property of the real numbers.

> **Fact.** If A is a non-empty set of real numbers which has an upper bound, then A has a *least upper bound* or *supremum*, often denoted by $\sup(A)$. Similarly, if a non-empty set of real numbers has a lower bound, then A has a *greatest lower bound* or *infimum*, often denoted by $\inf(A)$.

For example, consider the set

$$A = \left\{ \frac{1}{n+1} \;\middle|\; n \geq 1 \right\} \cup \left\{ \frac{n}{n+1} \;\middle|\; n \geq 1 \right\} = \left\{ \dots, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \dots \right\}.$$

It is not hard to see that A has greatest lower bound $\inf(A) = 0$ and least upper bound $\sup(A) = 1$. In this case, $\inf(A)$ and $\sup(A)$ are not themselves in A. In general, they may or may not be be. For example, the interval $[-2, 3]$ includes both its greatest lower bound -2 and its least upper bound 3, the interval $[-2, 3)$ includes its greatest lower bound but not its least upper bound, the interval $(-2, 3]$ does the reverse of the last, and $(-2, 3)$ includes neither.

> **Definition.** Suppose $a < b$. A *partition* of the interval $[a, b]$ is a set of points $P = \{ t_0, t_1, t_2, \dots, t_n \}$ such that $a = t_0 < t_1 < t_2 < \cdots < t_n = b$.
>
> A function $f(x)$ is *bounded* on $[a, b]$ if it is defined on $[a, b]$ and there are real numbers m and M such that $m \leq f(x) \leq M$ for all $a \leq x \leq b$.

Suppose $f(x)$ is bounded on $[a,b]$ and $P = \{ t_0, t_1, t_2, \dots, t_n \}$ is a partition of $[a,b]$. For each i with $1 \le i \le n$, let $m_i = \inf\{ f(x) \mid t_{i-1} \le x \le t_i \}$ and $M_i = \sup\{ f(x) \mid t_{i-1} \le x \le t_i \}$. Then the *lower sum* of $f(x)$ for P is

$$L(f,P) = \sum_{i=1}^{n} m_i \left(t_i - t_{i-1}\right) = m_1 \left(t_1 - t_0\right) + m_2 \left(t_2 - t_1\right) + \cdots m_n \left(t_n - t_{n-1}\right) ,$$

and the *upper sum* of $f(x)$ for P is

$$U(f,P) = \sum_{i=1}^{n} M_i \left(t_i - t_{i-1}\right) = M_1 \left(t_1 - t_0\right) + M_2 \left(t_2 - t_1\right) + \cdots M_n \left(t_n - t_{n-1}\right) .$$

Some comments are in order here. First, $m_i \left(t_i - t_{i-1}\right)$ and $M_i \left(t_i - t_{i-1}\right)$ are the weighted areas of rectangles such that

$$m_i \left(t_i - t_{i-1}\right) \le \text{weighted area between } y = f(x) \text{ and } y = 0 \text{ on } \left[t_{i-1}, t_{[}\right] \le M_i \left(t_i - t_{i-1}\right) .$$

Second, we need $f(x)$ to be bounded on $[a,b]$ when defining the upper and lower sums in order to ensure that the numbers m_i and M_i are defined. Third, we did not assume $f(x)$ was continuous. Every continuous function on a closed interval $[a,b]$ will, of course, be bounded, but so will any function that has a finite number of removable or jump discontinuities in $[a,b]$. A function with infinitely many discontinuities on $[a,b]$ might not be bounded, and a function with a vertical asymptote at some point in the interval is guaranteed not to be.

A couple of technically useful facts about upper and lower sums of $f(x)$ on a partition P of $[a,b]$ are given in the following results.

Lemma. Suppose $f(x)$ is bounded on $[a,b]$ and P and Q are partitions of $[a,b]$ such that every point of P is also a point of Q. (So the extra points of Q subdivide (at least some of) the pieces that P divides $[a,b]$ into.) Then $L(f,P) \le L(f,Q) \le U(f,Q) \le U(f,P)$.

Proposition. Suppose $f(x)$ is bounded on $[a,b]$ and P and R are any two partitions of $[a,b]$. Then $L(f,P) \le U(f,R)$.

Corollary. If $f(x)$ is bounded on $[a,b]$, then

$$\begin{aligned} &\sup\{ L(f,P) \mid P \text{ a partition of } [a,b] \} \\ &\le \inf\{ U(f,P) \mid P \text{ a partition of } [a,b] \} . \end{aligned}$$

THE DEFINITE INTEGRAL. We can define the fool thing at last:

Definition. A function $f(x)$ bounded on $[a,b]$ is said to be *integrable* on $[a,b]$ if $\sup\{ L(f,P) \mid P \text{ a partition of } [a,b] \} = \inf\{ U(f,P) \mid P \text{ a partition of } [a,b] \}$.

This number is the *definite integral* of $f(x)$ on $[a,b]$, denoted by $\displaystyle\int_a^b f(x)\,dx$.

That is,

$$\int_a^b f(x)\,dx = \sup\{\, L(f,P) \mid P \text{ a partition of } [a,b] \,\} \\ = \inf\{\, U(f,P) \mid P \text{ a partition of } [a,b] \,\}$$

if the sup and inf are equal, and $\int_a^b f(x)\,dx$ is undefined if they are not equal.

One potential problem with this definition is that the least upper and greatest lower bounds involved are not that easy to work with directly. The following result lets us work with something a little more concrete, at the cost of some epsilonics.

Theorem. Suppose $f(x)$ is bounded on $[a, b]$. Then $f(x)$ is integrable on $[a, b]$ if and only if for every $\varepsilon > 0$ there is a partition P of $[a, b]$ such that

$$U(f, P) - L(f, P) < \varepsilon.$$

Armed with the definitions and facts above, one can proceed to prove the usual basic properties of the definite integral relatively easily and move on to the Fundamental Theorem of Calculus, which gives a more practical tool for computing most common definite integrals by exploting the connection with antiderivatives.

For a pretty detailed development of much of this material, please consult any of the four editions of *Calculus* by Michael Spivak, one of the best-written mathematics textbooks anywhere. I ~~stole~~ ~~borrowed~~ took creative inspiration for most of the above from this book.

Reference

1. *Calculus* (Third Edition), by Michael Spivak, Publish or Perish, 1994. ISBN 0-914098-89-6