# Notes on the Lebesgue Integral

by

Francis J. Narcowich

Septemmber, 2014[1]

## 1 Introduction

In the definition of the Riemann integral of a function $f(x)$, the $x$-axis is partitioned and the integral is defined in terms of limits of the Riemann sums $\sum_{j=0}^{n-1} f(x_j^*)\Delta_j$, where $\Delta_j = x_{j+1} - x_j$. The basic idea for the Lebesgue integral is to partition the $y$-axis, which contains the range of $f$, rather than the $x$-axis.

This seems like a "dumb" idea at first. Shouldn't the two ways end up giving the same integral? Most of time this is the case, but Lebesgue was after integrating some functions for which the Riemann integral doesn't exist; for example, the Dirichlet function, which is defined on $[0,1]$:

$$\chi(x) = \begin{cases} 0 & x \text{ rational}, \\ 1 & x \text{ irrational}. \end{cases} \tag{1}$$

Lebesgue's reasoning was that there were uncountably many irrationals versus countably many rationals, so the "area" should be 1. It is is easy to show that the Riemann integral doesn't exist for $\chi$. The integral Lebesgue came up with not only integrates this function but many more. It also has the property that every Riemann integrable function is also Lebesgue integrable.

Many of the common spaces of functions, for example the square integrable functions on an interval, turn out to complete spaces – Hilbert spaces or Banach spaces – if the Riemann integral is replaced by the Lebesgue integral. The idea of splitting the range ($y$-axis) rather than the domain ($x$-axis) turns out to be invaluable in dealing with integrating functions over domains that aren't just real numbers. Such integrals arise in many fields, probability theory for instance. Lebesgue's idea turns out to be a *brilliant* "dumb" idea.

We now turn to the technical details involved in the Lebesgue integral, starting with Lebesgue sums. Choose an increasing sequence of points $\mathcal{P} = \{c \le y_0 < y_1 < y_2 < \cdots < y_n \le d\}$, where the range of $f$ is contained in $[c,d]$. As usual, we set $\|\mathcal{P}\| := \max_{0\le j\le n-1}(y_{j+1} - y_j)$ Let $E_j = \{x \in [a,b] : y_j \le f(x) < y_{j+1}\} = f^{-1}([y_j, y_{j+1}))$ and choose a point $y_j^*$ from each

[1] Revised, 9/5/2021.

interval $[y_j, y_{j+1}]$. (See Figure 1. Note that we have chosen $y_j^* = y_{j+1} = t$.) The corresponding Lebesgue sum is

$$L_{\mathcal{P}, Y^*}(f) := \sum_{j=0}^{n-1} y_j^* \mu(E_j) \tag{2}$$

where $\mu(E_j)$ denotes the "measure" or "length" of the set $E_j$ and $Y^* = \{y_j^*\}_{j=0}^{n-1}$ For this sum to make sense, we need a concept of measure for more sets than just intervals. For example, $\chi^{-1}([1/2, 3/2))$ is the set of all irrational numbers between 0 and 1. It doesn't contain any intervals at all. This leads to the question of how to extend the concept of measure to subsets of the real line that are much more complicated than simple intervals.

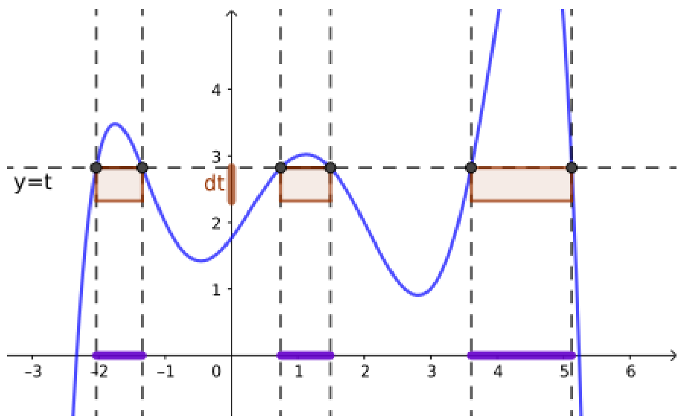


Figure 1: The region where $f$ is defined and a typical horizontal slice. (Downloaded on September 5, 2021: Sławomir Biały at English Wikipedia, CC BY-SA 4.0)

## 2 Measurable Sets

The ordinary idea of the measure/length of an interval can be thought of as a function that assigns to an interval a nonnegative number. In addition, $\mu$ satisfies standard properties; for example, if $I$ and $J$ are disjoint intervals, then $\mu(I \cup J) = \mu(I) + \mu(J)$.

An appropriate generalization of $\mu$ would satisfy the same properties, but for a wider class of subsets of $[a, b]$ than just the collection of intervals. Suppose that $\Sigma$ is a collection of subsets of $[a, b]$ and that we have a function

$\mu : \Sigma \to \mathbb{R}$. This is called a *set function*, because its domain consists of subsets of $[a, b]$. To go further and require that $\mu$ satisfies the same properties as those of length on intervals, we have to put conditions on $\Sigma$. For example, if we want $\mu$ to satisfy $\mu(A \cup B) = \mu(A) + \mu(B)$ for $A \cap B = \emptyset$, then $A \cup B$ must be in $\Sigma$. Here is a list of conditions that we will require $\Sigma$ to satisfy:

1. $\Sigma$ is *non-empty*. It will always contain both $\emptyset$ and $[a, b]$.
2. $\Sigma$ is *closed under complementation*: $A \in \Sigma$ if and only if $A^{\complement} \in \Sigma$.
3. $\Sigma$ is *closed under countable unions*: $\{A_j \in \Sigma\}_{j=1}^{\infty}$ then $\cup_{j=1}^{\infty} A_j \in \Sigma$.

A collection of subsets $\Sigma$ that satisfies that satisfies these conditions is called a *$\sigma$-algebra*.

We can now specify the properties that a set function $\mu : \Sigma \to [0, \infty)$ requires so that it mimics length for intervals. We say that $\mu$ is a ($\sigma$-finite) *measure* if and only if

1. $\mu([a, b]) = b - a$ and $\mu(\emptyset) = 0$.
2. Non-negativity: $\mu(A) \geq 0 \ \forall A \in \Sigma$.
3. Monotonicity: If $A \subseteq B$, then $\mu(A) \leq \mu(B)$.
4. Countable Additivity: If $\{A_j \in \Sigma\}_{j=1}^{\infty}$, with $A_i \cap A_j = \emptyset$, $i \neq j$, then $\mu(\cup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu(A_j)$.

These are fairly general conditions and they form the basis of the field of measure theory. Our aim here is simply to construct the Lebesgue measure. We note that the most natural choice for the class $\Sigma$ would be the one comprising *all* subsets of $[a, b]$. Unfortunately, a theorem whose proof employs the axiom of choice shows that there is no measure for this class. Instead, we will use a procedure that simultaneously constructs $\mu$ and $\Sigma$.

We start with the outer measure of a set, which can be defined for arbitrary subsets of $[a, b]$. Every open set $G$ in $[a, b]$ is the disjoint union of intervals, $G = \cup_i (a_i, b_i)$ and we define the outer measure to be $\mu^*(G) = \sum_i (b_i - a_i)$. This is of course the natural generalization of length to open sets.

We now turn to the general case. Let $A \subseteq [a, b]$ for some finite $a \neq b \in \mathbb{R}$. We define the *outer measure* $\mu^*(A)$ by

$$\mu^*(A) = \inf\{\mu^*(G) : A \subset G \,, G \text{ is open in } [a, b]\}.$$

We can also define the inner measure $\mu_*(A)$:

$$\mu_*(A) = b - a - \mu^*([a,b]\backslash A).$$

It is always true that both the inner and outer measure are nonnegative and that $\mu_*(A) \le \mu^*(A)$.[2] If $\mu_*(A) = \mu^*(A)$, then we say that $A$ is *Lebesgue measurable*[3].

The class $\Sigma$ is then just defined as all the collection of all Lebesgue measurable sets and the Lebesgue measure of $A \in \Sigma$ is defined as $\mu(A) := \mu_*(A) = \mu^*(A)$. It is not hard to show that all open sets and all closed sets are measurable, and that if $A$ is measurable, so is its complement $A^{\complement}$. With more work one can show that the measurable sets form a $\sigma$-algebra and the Lebesgue measure is a non-negative measure defined on $\Sigma$. When we are dealing only with intervals, the Lebesgue measure coincides with the usual concept of length.

The inequality $\mu_*(A) \le \mu^*(A)$ mentioned above is often useful. For instance, a set has measure 0 if for every $\epsilon > 0$ there is an open set $G \supset A$ such that $\mu(G) < \epsilon$. Since $\mu^*(A) = \inf\{\mu^*(G) : A \subset G\} < \varepsilon$ for all $\varepsilon > 0$, we have that $\mu^*(A) = 0$. However, we also have $0 \le \mu_*(A) \le \mu^*(A) = 0$. Therefore, $A$ is measurable and $\mu(A) = \mu_*(A) = \mu^*(A) = 0$. Here is an important example.

**Example 1.** Every countable set[4] has measure 0.

*Proof.* Let $A = \{x_1, x_2, \ldots, x_n, \ldots\}$. Let $I_j = (x_j - \frac{\epsilon}{2^j}, x_j + \frac{\epsilon}{2^j})$. Then, we see that $A \subset \cup I_j$, and hence

$$\begin{aligned}
\mu^*(A) &= \inf\{\mu^*(G) : A \subset G\,, G \text{ is open}\} \\
&\le \mu^*(\cup_j I_j) \le \sum_j \mu^*(I_j) \\
&= \sum_{j=1}^{\infty} (\frac{\epsilon}{2^j} + x_j - x_j + \frac{\epsilon}{2^j}) \\
&= \sum_{j=1}^{\infty} \frac{\epsilon}{2^{j-1}} = \epsilon
\end{aligned}$$

Since this holds for arbitrary $\epsilon > 0$, we see that $A$ is measurable and that $\mu(A) = 0$. □

[2] See Wilcox & Myers, p. 25, Lemma 8.7.

[3] The previous inequality implies that to verify $A$ is measurable only requires showing that $\mu^*(A) \le \mu_*(A)$.

[4] The cantor set is an example of an uncountable set having measure 0.

## 3 Measurable Functions

In order to form the Lebesgue sums in (2), we need to be able to find the measure of sets of the form $f^{-1}([c,d))$. This puts a restriction on $f$.

**Definition 1.** We say the $f : [a,b] \to \mathbb{R}$ is (Lebesgue) *measurable* if and only if for very $c < d$ the set $f^{-1}([c,d))$ is measurable.

There are many conditions equivalent to $f^{-1}([c,d))$ being measurable. In particular if any of these sets is measurable for all $c < d$, then $f$ is a measurable function: $f^{-1}((c,\infty))$, $f^{-1}([c,\infty))$, $f^{-1}((c,d))$, $f^{-1}((c,d])$, $f^{-1}((-\infty,c))$, and other similar sets. It's also useful to know that if, for every open set $G \subseteq [a,b]$, $f^{-1}(G)$ is measurable, then $f$ is measurable. In fact, a definition of measurable functions that applies for more general measure spaces is this: We say that $f : A \to \mathbb{R}$ is *measurable* if for every measurable $E \subset \mathbb{R}$, $f^{-1}(E) \subseteq A$ is a measurable subset of $A$.

Measurable functions may be combined in various ways to obtain other measurable functions. If $f$ and $g$ are measurable, then so are $af + bg$, $f \cdot g$, $f/g$ ($g \neq 0$), $|f|$. Every continuous function is measurable. If $f$ is continuous and the range of $g$ is in the domain of $f$, then $f \circ g$ is measurable. (The reverse composition, $g \circ f$ might not be mensurable. See Wilcox & Myers, p. 56, exercise 20.12.)

In the theory of Lebesgue integration, sets of measure 0 really won't contribute to an integral. Consequently, when we integrate two functions that are different only on a set of measure 0, we will find that their integrals will be the same. This situation occurs frequently enough that we make the definition below.

**Definition 2.** We say that $f$ equals $g$ almost everywhere (a.e.) if and only if the set of $x$ for which $f(x) \neq g(x)$ has measure 0.

A nice example of this is the Dirichlet function $\chi(x)$ defined in (1). Recall that $\chi(x) = 1$, except on $\mathbb{Q}$. Since $\mathbb{Q}$ has measure 0, $\chi = 1$ a.e. Here is one of the more important facts about measurable functions.

**Proposition 1.** *Suppose that $A$ is a measurable set and that $f_n : A \to \mathbb{R}$ is a sequence of measurable functions such that for each $x \in A$ we have $\lim_{n\to\infty} f_n(x) = f(x)$, then $f$ is measurable.*

**Definition 3.** Let $A$ be a measurable set. We say that a measurable function $s : A \to \mathbb{R}$ is *simple* if and only if the range of $s$ has a finite number of values.

The *characteristic* or *indicator* function of a measurable $A$ is defined by

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A. \end{cases}$$

Since the range of this function is $\{0, 1\}$, it is simple. In fact, any simple function can be represented as a finite linear combination of characteristic functions.

**Proposition 2.** *Let $s : A \to \mathbb{R}$ be a simple function with range $\{d_j\}_{j=1}^n$, $n < \infty$, and let $E_j = s^{-1}(\{d_j\})$. Then, $s$ has the form*

$$s = \sum_{j=1}^{n} d_j \chi_{E_j}.$$

*Conversely, if $s$ has the form above, then it is simple.*

The representation for $s$ given above is unique, provided the $E_j = s^{-1}(\{d_j\})$. Otherwise, one may obtain different combinations of characteristic functions that result in the same $s$.

## 4 The Lebesgue Integral

We briefly introduce the concept of the Lebesgue integral in this section, and in addition, discuss some important theorems associated with it. We start by returning to the Lebesgue sum in (2). If $f$ is bounded and measurable on a measurable set $A$, then $\mu(E_j)$ is defined, and so is $L_{\mathcal{P},Y^*}(f) = \sum_{j=0}^{n-1} y_j^* \mu(E_j)$. At this point, we may define upper and lower sums $L_{\mathcal{P}}^{+}(f) = \sum_{j=0}^{n-1} y_{j+1}\mu(E_j)$ and $L_{\mathcal{P}}^{-}(f) = \sum_{j=0}^{n-1} y_j \mu(E_j)$ and then proceed to define the Lebesgue integral $\int_A f(x)dx$ in roughly the same way as one defines the Riemann integral.

Another, somewhat easier, equivalent way is to start out by defining the Lebesgue integral of a simple function $s$ to be

$$\int_A s(x) d\mu = \sum_{j=1}^{n} d_j \mu(E_j), \;\; E_j = s^{-1}(\{d_j\}). \tag{3}$$

Then we may define upper and lower integrals for a bounded, measurable function $f$ via

$$\int_A^+ f(x)d\mu = \inf_s \Big\{ \int_A s(x)dx\colon s(x) \geq f(x),\ s \text{ simple}\Big\}$$
$$\int_A^- f(x)d\mu = \sup_s \Big\{ \int_A s(x)dx\colon s(x) \leq f(x),\ s \text{ simple}\Big\}$$

As usual, if these are equal, then $\int_A f(x)d\mu := \int_A^+ f(x)d\mu = \int_A^- f(x)d\mu$. It turns out that for a bounded measurable function both exist and are equal:

**Proposition 3.** *Let $f$ be a bounded, measurable function on a bounded measurable set $A \subset \mathbb{R}$. Both $\int_A^+ f(x)d\mu$ and $\int_A^- f(x)d\mu$ always exist and are equal. Thus,*

$$\int_A f(x)d\mu = \int_A^+ f(x)d\mu = \int_A^- f(x)d\mu \tag{4}$$

The Lebesgue integral on a measurable set $A$ will be denoted by either $\int_A f(x)\,dx$ or $\int_A f(x)\,d\mu$.

We now turn to unbounded functions – e.g., $f(x) = x^{-1/2}$ on $[0,1]$. For a measurable function $f(x) \geq 0$, whether bounded or not,

$$\sup_s \Big\{ \int_A s(x)dx\colon s(x) \leq f(x),\ s \text{ simple}\Big\},$$

always exists – even if it's infinite. If the sup *is* finite, then we can define the Lebesgue integral of $f$ to be

$$\int_A f(x)d\mu := \sup_s \Big\{ \int_A s(x)dx\colon s(x) \leq f(x),\ s \text{ simple}\Big\}, \tag{5}$$

Remark: One can't use the upper bound formulation because if $f$ is unbounded, one can get *below* the function with a simple function, but *not* above it.

We can write $f$ as the difference of two non-negative functions, $f = f_+ - f_-$, where $f_+ = \frac{1}{2}(f + |f|)$ and $f_- = \frac{1}{2}(|f| - f)$. Also, $|f| = f_+ + f_-$. We say that a measurable function $f$, whether bounded or not, is integrable on $A$ if $\int_A |f|d\mu < \infty$. It's easy to show that if $\int_A |f|d\mu$ exists, then both $\int_A f_+ d\mu$ and $\int_A f_- d\mu$ exist. Consequently, we may define $\int_A f\,d\mu = \int_A f_+\,d\mu - \int_A f_-\,d\mu$. Note that both integrals must be finite for $\int_A f\,d\mu$ to exist.

This definition causes problems for certain Riemann integrable functions. Here's an example for an unbounded interval. The integral $\int_0^\infty \frac{\sin(x)}{x}dx = \pi/2$, in the Riemann sense. However, $\int_0^\infty \frac{|\sin(x)|}{x}dx = \infty$, so the integral does not exist in the Lebesgue sense.

Let $A, B$ denote measurable sets, with $\mu(A), \mu(B) < \infty$. The properties and theorems below are satisfied by the Lebesgue integral.

1. If $f$ is a bounded and measurable function, then the integral exists.
2. Integration is a linear operation. That is, if $f$ and $g$ are both integrable and $a, b \in \mathbb{R}$, then $\int_A (af + bg)\, d\mu = a \int_A f\, d\mu + b \int_A g\, d\mu$.
3. If $\int_A f(x)^2\, dx$ and $\int_A g(x)^2\, dx$ both exist, then $\int_A f(x)g(x)\, dx$ and $\int_A (f(x) + g(x))^2\, dx$ exist.
4. $\int_{A \cup B} f\, d\mu = \int_A f\, d\mu + \int_B f\, d\mu$, where $\mu(A \cap B) = 0$.
5. If $f = g$ almost everywhere, then $\int_A f\, d\mu = \int_A g\, d\mu$.
6. If the Riemann integral of $f$ exists, then the Lebesgue integral exists and the integrals are equal.

**Theorem 1.** *(Monotone Convergence Theorem): Let $\{f_j\}$ be a collection of measurable functions on $A$ that satisfy $0 \leq f_1 \leq f_2 \leq \cdots \leq f_n \leq \ldots$ almost everywhere. Define $f$ as the pointwise limit $f(x) = \lim f_j(x)$. Then, $f$ is measurable and*

$$\lim \int_A f_j\, d\mu = \int_A f\, d\mu.$$

**Theorem 2.** *(Dominated Convergence Theorem): Let $\{f_j\}$ be a set of measurable functions that converges pointwise to a function $f$ and assume there exists an integrable function $g$ such that $|f_j(x)| \leq g(x)$ almost everywhere. Then, $f$ is integrable and*

$$\lim \int_A f_j\, d\mu = \int_A f\, d\mu.$$

**Theorem 3.** *(Fubini's Theorem): Let $f$ be measurable on $A \times B$. If*

$$\int_{A \times B} |f(x, y)|\, d\mu(x, y) < \infty$$

*then $\int_A \int_B f(x, y)\, d\mu(x) d\mu'(y)$ exists and the order of integration may be switched.*

We define the space $L^p([a, b])$ to be the space of $p$ integrable functions for $1 \leq p < \infty$. That is,

$$L^p([a, b]) = \{f : \int_a^b |f(x)|^p < \infty\}.$$

The space $L^p([a,b])$ is a normed space when it is equipped with the norm

$$\|f\|_p = (\int_a^b |f(x)|^p d\mu(x))^{\frac{1}{p}}.$$

For the case $p = \infty$, we define

$$\|f\|_\infty = \text{ ess sup}|f| = \inf\{a \in \mathbb{R} : \mu(\{x : |f|(x) > a\}) = 0\}$$

and we define

$$L^\infty([a,b]) = \{f : \|f\|_\infty < \infty\}.$$

**Example 2.** Just because a function is in $L^\infty$ doesn't necessarily mean that it is bounded in the usual sense. Here is an example of this.

$$f(t) = \begin{cases} 1 & t \in \mathbb{R}\backslash\mathbb{Q} \\ q & t = \frac{p}{q} \end{cases} \tag{6}$$

where $\frac{p}{q}$ is reduced to lowest terms and $q > 0$. This function is unbounded, but $\|f\|_\infty = \text{ ess sup}|f(x)| = 1$, because the set of rational numbers has measure zero. This implies that, although $f$ is unbounded, $f \in L^\infty$.

**Theorem 4.** *The space $L^p([a,b])$ is a complete space for $1 \leq p \leq \infty$.*

This theorem implies that $L^p([a,b])$ is a Banach space. In the special case of $p = 2$, $L^2([a,b])$ is a Hilbert space.

# 5 Examples

**Example 3.** Suppose $f_n(x) = \frac{1}{\sqrt{x+\frac{1}{n}}}$ for $x \in [0,1]$ and $n \geq 1$. Let $f$ denote the pointwise limit of $f_n$. Show that $f$ is integrable.

*Proof.* We see that $f_{n+1}(x) = \frac{1}{\sqrt{x+\frac{1}{n+1}}} \geq \frac{1}{\sqrt{x+\frac{1}{n}}} = f_n(x)$ holds for all $n$ and $\lim_{n\to\infty} f_n(x) = \frac{1}{\sqrt{x}}$ almost everywhere. Then, by the monotone convergence theorem, $\frac{1}{\sqrt{x}}$ is integrable and $\lim_{n\to\infty} \int_0^1 \frac{dx}{\sqrt{x+\frac{1}{n}}} = \int_0^1 \frac{dx}{\sqrt{x}}$. $\square$

**Example 4.** Let $f \in L^1(\mathbb{R})$ and define $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\, dt$. Show that $\hat{f}$ is a continuous function.

*Proof.* By a straightforward computation, we see that

$$\begin{aligned}\hat{f}(\omega+h)-\hat{f}(\omega) &= \int_{-\infty}^{\infty} f(t)(e^{-i(\omega+h)t}-e^{-i\omega t})\,dt \\ &= \int_{-\infty}^{\infty} f(t)e^{-i\omega t}(e^{-iht}-1)dt\end{aligned}$$

The integrand is bounded above by $2|f(t)|$ almost everywhere, and since $f \in L^1$, $2|f(t)|$ is integrable. Applying the dominated convergence theorem results in $\lim_{h\to 0}\hat{f}(\omega+h)=\hat{f}(\omega)$, thus $\hat{f}$ is continuous. $\square$

Previous: Banach spaces and Hilbert spaces
Next: Orthonormal sets & expansions