Integrable functions

Notation. Throughout this section $(X, \mathcal{M}, \mu)$ will denote a $\sigma$-finite measure space.

Definition 1. Let $f : X \to \mathbb{R}$ be a measurable function on $X$. The integral of $f$ with respect to $\mu$ is
\[
\int f \, d\mu := \int f^+ \, d\mu - \int f^- \, d\mu,
\]
provided this makes sense (that is, if the right hand side is not $\infty - \infty$). $f$ is integrable if $\int f \, d\mu \in \mathbb{R}$, and $L^1 = L^1(X, \mathcal{M}, \mu)$ denotes the set of all integrable functions. If $A \in \mathcal{M}$ we define $\int_A f \, d\mu = \int f \chi_A \, d\mu$. When the measure $\mu$ is understood we write $\int f$, and when we want to emphasize the space we write $\int_X f$.

Proposition 2. If $f$ is measurable, then $f$ is integrable if and only if $|f|$ is, in which case $\left| \int f \right| \leq \int |f|$.

Proposition 3. If $f \in L^1$, then $|f| < \infty$ a.e., so $f(x) \in \mathbb{R}$ a.e. $x$.

Proposition 4. If $f$ is measurable and $A \in \mathcal{M}$ then $\int_A f = \int_A f^+ - \int_A f^-$. 

Corollary 5. Let $f$ and $g$ be measurable.

(1) If $\mu(A) = 0$ then $\int_A f = 0$ and $\int_{A^c} f = \int f$.
(2) If $f = 0$ a.e. then $\int f = 0$.
(3) If $f = g$ a.e., then $\int f = \int g$.

Proposition 6. If $f, g \in L^1$ and $c \in \mathbb{R}$ then:

(1) $cf \in L^1$ and $\int cf = c \int f$;
(2) $f + g \in L^1$ and $\int (f + g) = \int f + \int g$.

Proof. (1) Exercise.

(2) There is a subtlety: there might exist $x$ such that $f(x)$ and $g(x)$ are infinite and have opposite signs, so $f(x) + g(x)$ would be undefined. But $f$ and $g$ are real-valued a.e. since they are integrable, so we can change them on a null set (which won’t affect the integrals) to arrange that they are real-valued everywhere. Then we can use a trick: putting $h = f + g$, we have

\[
h^+ - h^- = f^+ - f^- + g^+ - g^-,
\]

so

\[
h^+ + f^- + g^- = h^- + f^+ + g^+,
\]

hence

\[
\int h^+ + \int f^- + \int g^- = \int h^- + \int f^+ + \int g^+,
\]
thus
\[
\int (f + g) = \int (h^+ - h^-) = \int h^+ - \int h^- \\
= \int f^+ - \int f^- + \int g^+ - \int g^- = \int f + \int g.
\]

\[\square\]

**Definition 7.** We call a function *a.e.-defined* if it is defined on the complement of a null set. If \( f \) is a.e.-defined and measurable, then we can extend \( f \) to a measurable function on all of \( X \) in lots of ways, for example by letting it be 0 on the complement of its given domain. Any of these extensions is integrable if and only if all of them are, in which case the value of the integral doesn’t depend upon the choice of extension. We say that an a.e.-defined function is *integrable* if any of its measurable extensions are.

Actually, if \( \mu \) is complete then it doesn’t matter how we extend \( f \).

**Definition 8.** Two a.e.-defined measurable functions \( f \) and \( g \) are *equivalent*, written \( f \sim g \), if \( f = g \) a.e.

It’s not hard to verify that this gives an equivalence relation\(^1\) on the set of a.e.-defined measurable functions.

**Definition 9.** We redefine \( L^1 \) to be the set of *equivalence classes*\(^2\) of a.e.-defined integrable functions under the equivalence relation \( \sim \).

Thus integrable functions which agree a.e. determine the same element of \( L^1 \). However, for convenience we abuse this by continuing to speak of integrable functions as elements of \( L^1 \) — this causes no confusion in practice. Also, from now on we’ll usually drop the cumbersome phrase “a.e.-defined”, so that when we refer to an integrable function on \( X \) we have to keep in mind that it might only be defined almost everywhere. In fact, we’ll do the same for measurable functions in general. We need this flexibility because, for example, we will want to consider an a.e.-limit of a sequence of measurable functions, which might not be everywhere-defined. Also, we want to add measurable functions, and we proceed as we did for integrable ones: assuming \( f \) and \( g \) are integrable and real-valued a.e., we can replace them by equivalent functions that are real-valued everywhere, and the equivalence class of \( f + g \) is well-defined.

**Proposition 10.** \( L^1 \) is a normed space with norm
\[
\|f\|_1 := \int |f|,
\]
and the linear functional \( \int : L^1 \to \mathbb{R} \) is continuous.

**Proof.** First of all, if \( f \in L^1 \) then \( \|f\|_1 = \int |f| \geq 0 \), and
\[
\|f\|_1 = 0 \iff \int |f| = 0 \iff |f| = 0 \text{ a.e.} \\
\iff f = 0 \text{ a.e.} \\
\iff f \text{ is the 0 element of } L^1.
\]

\(^1\)that is, (1) \( f \sim f \), (2) if \( f \sim g \) then \( g \sim f \), and (3) if \( f \sim g \) and \( g \sim h \) then \( f \sim h \)

\(^2\)The *equivalence class* of \( f \) is \( \{g : g \sim f\} \).
Next, in previous computations we’ve verified \( \int |cf| = |c| \int |f| \) and \( \int |f + g| \leq \int |f| + \int |g| \), so \( \| \cdot \|_1 \) is a norm on \( L^1 \). Since the integral is linear on \( L^1 \), continuity follows from
\[
\left| \int f \right| \leq \int |f| = \|f\|_1.
\]
Thus, another reason we need to be using equivalence classes is so that if \( f \in L^1 \) and \( \|f\|_1 = 0 \) then \( f \) is the 0 element of \( L^1 \).

**Corollary 11.** If \( f, g \in L^1 \) and \( \int_A f = \int_A g \) for all \( A \in \mathcal{M} \) then \( f = g \) a.e.

**Proof.** Replacing \( f \) by \( f - g \), without loss of generality we can assume that \( f \in L^1 \) and \( \int_A f = 0 \) for all \( A \in \mathcal{M} \), and we must show that \( f = 0 \) a.e. Put
\[
P = [f^+ \neq 0] \quad \text{and} \quad N = [f^- \neq 0].
\]
Then
\[
\int_P f = \int f^+ = 0,
\]
so \( f^+ = 0 \) a.e. Similarly, \( f^- = 0 \) a.e., so \( f = f^+ - f^- = 0 \) a.e. \( \square \)

**Theorem 12** (Dominated Convergence Theorem). Let \( (f_n) \) be a sequence in \( L^1(X, \mu) \) converging to a measurable function \( f \) a.e. Suppose there exists \( g \in L^1(X, \mu) \) such that for every \( n \in \mathbb{N} \) we have \( |f_n| \leq g \) a.e. Then \( f \in L^1 \) and
\[
\int f_n \to \int f.
\]
Moreover, \( f_n \to f \) in \( L^1 \).

Note that if \( \mu \) is complete or the convergence \( f_n \to f \) is everywhere, then we do not need the hypothesis that \( f \) is measurable.

**Proof.** Since \( |f_n| \leq g \) a.e. for all \( n \), and a countable union of null sets is null, for a.e. \( x \) we have \( |f_n(x)| \leq g(x) \) for all \( n \), hence \( |f(x)| \leq g(x) \). Thus \( |f| \leq g \) a.e., so \( f \in L^1 \).

For the other part, we need a trick: \( g + f_n \geq 0 \) for all \( n \), so by Fatou’s Lemma
\[
\int g + \int f = \int (g + f) = \int \lim (g + f_n)
\]
\[
\leq \lim \inf \int (g + f_n) = \lim \inf \left( \int g + \int f_n \right) = \int g + \lim \inf \int f_n,
\]
hence
\[
\int f \leq \lim \inf \int f_n.
\]
Applying this to \( -f \), we get
\[
-\int f = \int (-f) \leq \lim \inf \int (-f_n) = \lim \inf -\int f_n = -\lim \sup \int f_n,
\]
so
\[
\int f \geq \lim \sup \int f_n.
\]
Therefore we must have
\[ \limsup \int f_n = \liminf \int f_n = \int f. \]
Finally, we have \(|f_n - f| \to 0\) a.e., and \(|f_n - f| \leq 2g\) a.e., so by the above we have
\[ \|f_n - f\|_1 = \int |f_n - f| \to 0. \]

Just as the Monotone Convergence Theorem has a series version, so does the Dominated Convergence Theorem:

**Corollary 13** (Series Version of DCT). If \(\sum f_n\) is a series in \(L^1(X, \mu)\) such that \(\sum \int |f_n| < \infty\), then the series \(\sum f_n\) converges a.e. and in \(L^1\), and
\[ \int \sum f_n = \sum \int f_n. \]

**Proof.** By the Monotone Convergence Theorem
\[ \int \sum |f_n| = \sum \int |f_n| < \infty, \]
so \(\sum |f_n| < \infty\) a.e., hence \(\sum f_n\) converges a.e. For each \(k \in \mathbb{N}\),
\[ \sum_{i=1}^{k} |f_n| \leq \sum_{i=1}^{\infty} |f_n|, \]
and the latter function is integrable, so by the Dominated Convergence Theorem \(\sum_{1}^{\infty} f_n \in L^1\) and
\[ \int \sum_{1}^{\infty} f_n = \lim_k \left( \int \sum_{1}^{k} f_n \right) = \lim_k \sum_{1}^{k} \int f_n = \sum_{1}^{\infty} \int f_n. \]
Finally, the series \(\sum f_n\) actually converges in \(L^1\) by the same reasoning as in the Dominated Convergence Theorem.

When we apply the above result, we’ll sometimes just say “by the Dominated Convergence Theorem”.

**Corollary 14.** \(L^1\) is a Banach space.

**Proof.** It suffices to observe that, by the Dominated Convergence Theorem, every absolutely convergent series \(\sum f_n\) in \(L^1\) converges.

Here’s an unexpected bonus:

**Corollary 15.** If \(f_n \to f\) in \(L^1\) then some subsequence of \((f_n)\) converges to \(f\) a.e.

**Proof.** The proof of the result that a normed space is complete if every absolutely convergent series converges proceeded by replacing a given Cauchy sequence, such as \((f_n)\), by a subsequence \((f_{n_k})\) comprising the partial sums of an absolutely convergent series. By the Dominated Convergence Theorem, every absolutely convergent series in \(L^1\) also converges a.e. Thus \(f_{n_k} \to f\) a.e.
Corollary 16. The integrable simple functions are dense in $L^1(\mathbb{R})$.

Proof. Let $f \in L^1(\mathbb{R})$, and choose a sequence $(\phi_n)$ of simple functions such that $\phi_n \to f$ and $|\phi_n| \leq |f|$ pointwise. Then each $\phi_n$ is integrable, and by the Dominated Convergence Theorem $\phi_n \to f$ in $L^1$. \hfill $\square$

Definition 17. A step function is a linear combination of characteristic functions of bounded intervals in $\mathbb{R}$.

Corollary 18. The step functions are dense in $L^1(\mathbb{R})$.

Proof. By Corollary 16, it suffices to show that for all $A \in \mathcal{L}$ and $\varepsilon > 0$ there exists a finite union $B$ of bounded intervals such that $||\chi_A - \chi_B||_1 < \varepsilon$. But we can find such a $B$ with $m(A \setminus B) < \varepsilon$, and $\int |\chi_A - \chi_B| = m(A \setminus B)$, so we are done. \hfill $\square$

Definition 19. Let $X$ be a metric space. The support of a continuous function $f : X \to \mathbb{R}$ is

$$\text{supp } f := \{f \neq 0\}.$$ We say $f$ has compact support if $\text{supp } f$ is compact (of course), and $C_c(X)$ denotes the set of all such functions.

Corollary 20. $C_c(\mathbb{R})$ is dense in $L^1(\mathbb{R})$.

Proof. By Corollary 18, it suffices to show that if $B$ is a bounded interval then $\chi_B$ can be approximated in the $L^1$ norm by an element of $C_c(\mathbb{R})$. Without loss of generality $B = (a, b)$ with $a < b$. Choose a compact interval $K \subset B$ such that $m(B \setminus K) < \varepsilon$, and choose a continuous function $f : \mathbb{R} \to \mathbb{R}$ such that $\chi_K \leq f \leq \chi_B$ (for example, let $f$ be piecewise linear). Then $f \in C_c(\mathbb{R})$ and

$$\|\chi_B - f\|_1 = \int (\chi_B - f) \leq \int (\chi_B - \chi_K) = m(B \setminus K) < \varepsilon.$$ \hfill $\square$

Corollary 21 (Lebesgue’s Characterization of Riemann Integrability). A bounded real-valued function $f$ on $[a, b]$ is Riemann integrable if and only if it is continuous a.e. Moreover, in this case $f$ is also Lebesgue integrable, and the Riemann and Lebesgue integrals of $f$ coincide.

Proof. For each $n \in \mathbb{N}$ let $\mathcal{P}_n$ be the family of $2^n$ closed intervals of length $2^{-n}(b-a)$ with union $[a, b]$. Define $g_n, h_n : [a, b] \to \mathbb{R}$ as follows: for $x \in [a, b]$, first find $I \in \mathcal{P}_n$ such that $x \in I$. Then define:

- $g_n(x) = h_n(x) = f(x)$ if $x \in \partial I$;
- $g_n(x) = \inf_{y \in I} f(y)$ and $h_n(x) = \sup_{y \in I} f(y)$ if $x \in I^\circ$.
Then 
\[ g_1 \leq g_2 \leq \cdots \leq f \leq \cdots \leq h_2 \leq h_1, \]
and by Darboux’s Theorem \( f \) is Riemann integrable if and only if \( \lim_n \int g_n = \lim_n \int h_n \), in which case the Riemann integral of \( f \) is this common limit.

Observe that \( g_n \) and \( h_n \) are step functions for which \( \int g_n \) equals the “lower sum” of \( f \) associated with the “partition” of \([a,b]\) determined by the endpoints of the intervals in \( \mathcal{P}_n \), and similarly \( \int h_n \) is the “upper sum”, and also that for step functions it is obvious that the Lebesgue and Riemann integrals coincide.

Put \( g = \lim g_n \) and \( h = \lim h_n \). By the Dominated Convergence Theorem, \( g \) and \( h \) are integrable and
\[
\int (h - g) = \lim \int (h_n - g_n) = \lim \int h_n - \lim \int g_n.
\]
Since \( h - g \geq 0 \), we have
\[
\int (h - g) = 0 \iff h - g = 0 \text{ a.e.} \iff g = h \text{ a.e.}
\]

Put \( D = \bigcup_{n=1}^{\infty} \bigcup_{I \in \mathcal{P}_n} \partial I \), which is a countable subset of \([a,b]\). Claim: for \( x \in [a,b] \setminus D \), \( f \) is continuous at \( x \) if and only if \( g(x) = h(x) \). To see this, first assume continuity, and, given \( \varepsilon > 0 \), choose \( \delta > 0 \) such that for all \( y \in [a,b] \), if \( |y - x| < \delta \) then \( |f(y) - f(x)| < \varepsilon \). Choose \( n \) such that \( 2^{-n}(b - a) < \delta \), and find \( I \in \mathcal{P}_n \) such that \( x \in I^c \). Then
\[
|f(y) - f(z)| < 2\varepsilon \quad \text{for all } y, z \in I,
\]
so
\[
0 \leq h(x) - g(x) \leq h_n(x) - g_n(x) \leq 2\varepsilon.
\]
Letting \( \varepsilon \to 0 \), we get \( g(x) = h(x) \). For the converse direction, we essentially reverse the steps: assuming \( g(x) = h(x) \), and given \( \varepsilon > 0 \), there exists \( n \) such that \( h_n(x) - g_n(x) < \varepsilon \), and a subinterval \( I \) associated to \( \mathcal{P}_n \) such that \( x \in I^c \). Let \( \delta \) be the minimum of the distances from \( x \) to the two endpoints of \( I \). Then
\[
|f(x) - f(y)| < \varepsilon \quad \text{if } |x - y| < \delta,
\]
and we’ve shown that \( f \) is continuous at \( x \).

Thus,
\[
f \text{ is Riemann integrable}
\iff g = h \text{ a.e.}
\iff g = h \text{ a.e. on } D^c \text{ (since } m(D) = 0)\]
\iff f \text{ is continuous a.e. on } D^c
\iff f \text{ is continuous a.e.}
\]

For the other part, if these equivalent conditions are satisfied, then \( f = g \text{ a.e.} \) because \( g \leq f \leq h \). Thus \( f \in L^1 \) and
\[
\int f = \int g = \lim \int g_n,
\]
which coincides with the Riemann integral of \( f \). \( \square \)