Let \((X, \mu)\) be a measure space.

**Definition 1.** A function \(f: X \to \mathbb{C}\) is called measurable when \(f^{-1}(U)\) is a measurable set for every Borel set \(U \subset \mathbb{C}\).

**Example 2.** If \(X\) is a topological space, equipped with the \(\sigma\)-algebra of Borel sets, then any continuous function \(X \to \mathbb{C}\) is measurable.

**Theorem 3.** If \(f_n: X \to \mathbb{C}\) is a sequence of measurable functions that converges pointwise to \(f\), then \(f\) is a measurable function.

In particular, since continuity implies measurability (on topological spaces) a pointwise limit of continuous functions is not necessarily continuous but is always measurable.

**Example 4.** Let \(X = [0,1]\) and \(f_n(x) = x^n\). The \(f_n\)'s are continuous and converge pointwise, but the limit function is not continuous: it is 0 on \([0,1)\) and 1 at \(x = 1\). This pointwise limit is measurable.

**Definition 5.** For a subset \(A \subset X\), let \(\xi_A\) be the characteristic function of \(A\):
\[
\xi_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}
\]

**Definition 6.** A simple map \(X \to \mathbb{C}\) is a finite sum \(\sum_{i=1}^{n} c_i \xi_{A_i}\) where \(c_i \in \mathbb{C}\) and each \(A_i\) is a measurable subset of \(X\) that may possibly have infinite measure.

**Definition 7.** A step map \(X \to \mathbb{C}\) is a simple map \(\sum_{i=1}^{n} c_i \xi_{A_i}\) where \(\mu(A_i) < \infty\) for all \(i\).

Simple maps and step maps each take only a finite number of values. The difference between a simple map and a step map is that for a step map the set where it takes each nonzero value has finite measure.

**Example 8.** If \(\mu(A) = \infty\) then \(\xi_A\) is a simple map but not a step map.

For any simple map \(f = \sum_{i=1}^{n} c_i \xi_{A_i}\) we have \(|f| = \sum_{i=1}^{n} |c_i| \xi_{A_i}\), so if \(f\) is a step map then so is \(|f|\). The sets \(A_i\) do not have to be disjoint, but we can always break up the supports of the characteristic functions to assume such a formula has disjoint \(A_i\)'s.

Integration with respect to \(\mu\) is defined on step maps, not simple maps: for a step map \(f = \sum_{i=1}^{n} c_i \xi_{A_i}\), define
\[
\int_X f \, d\mu = \sum_{i=1}^{n} c_i \mu(A_i).
\]

From the properties of measures, this sum is well-defined (independent of the way \(f\) is written as a linear combination of characteristic functions of finite-measure subsets of \(X\)) and integration is linear on step maps.

The following two theorems illustrate the difference between pointwise limits of simple maps and step maps.
Theorem 9. For a function \( f : X \to \mathbb{C} \), the following are equivalent:

1. \( f \) is a pointwise limit of simple maps,
2. \( f \) is a measurable function.

Proof. See property M8 in [1, p. 118].

Notice the conditions in Theorem 9 don’t involve the measure \( \mu \) at all.

Theorem 10. For a function \( f : X \to \mathbb{C} \), the following are equivalent:

1. \( f \) is a pointwise limit of step maps almost everywhere,
2. \( f \) is equal almost everywhere to a measurable function and vanishes outside a \( \sigma \)-finite subset of \( X \).

Proof. See property M11 of [1, pp. 124], taking \( E \) there to be \( \mathbb{C} \).

Notice the three concepts step map, almost everywhere, and \( \sigma \)-finite that appear in Theorem 10 all depend on the measure \( \mu \). We call a function satisfying the conditions of Theorem 10 \( \mu \)-measurable.

The point of Theorem 10 is that the first property is what leads to the possibility \( f \) could be integrable (not all functions are integrable!), while the second property is something you might be able to verify in practice with an individual function. When \( X \) itself is \( \sigma \)-finite (like \( \mathbb{R}^n \) with Lebesgue measure), the second property in Theorem 10 just says \( f \) is equal almost everywhere to a measurable function, so in this case the distinction between pointwise limits of simple maps and pointwise limits of step maps is negligible.

The step maps on \( X \) form a vector space over \( \mathbb{C} \), and for step maps \( f \) and \( g \) we call the number \( \int_X |f - g| \, d\mu \) their \( L^1 \)-distance. It is almost a metric; the only reason it isn’t is that the condition \( \int_X |f - g| \, d\mu = 0 \) is the same as \( f = g \) almost everywhere on \( X \) rather than everywhere on \( X \). Agreling to identify functions equal almost everywhere, we want to complete the space of step maps on \( X \) with respect to this distance.

Here is the basic result that gets integration working, dubbed the “fundamental lemma of integration” in [1, Chap. VI, Section 3].

Theorem 11. If \( \{f_n\} \) is an \( L^1 \)-Cauchy sequence of step maps, then it has a subsequence that converges pointwise almost everywhere and converges uniformly off some set of arbitrarily small measure.

Proof. See [1, Lemma 3.1, p. 129].

Definition 12. Set \( L^1(\mu) = L^1(\mu, \mathbb{C}) \) to be the functions \( X \to \mathbb{C} \) that are a pointwise limit almost everywhere of an \( L^1 \)-Cauchy sequence of step maps on \( X \). Each function in \( L^1(\mu) \) is \( \mu \)-measurable by Theorem 10. We write \( L^1(\mu, \mathbb{R}) \) for the same functions on \( X \) that have real values.

Theorem 13. If \( f \in L^1(\mu) \) is the pointwise limit almost everywhere of the \( L^1 \)-Cauchy sequence of step maps \( \{f_n\} \), then the sequence of integrals

\[ \int_X f_n \, d\mu \]

converges in \( \mathbb{C} \) and the limit is independent of the choice of \( \{f_n\} \).

Proof. See [1, Lemma 3.2, p. 130], whose proof uses Theorem 11.
The limit of the integrals in Theorem 13, depends only on \( f \) and not on the \( f_n \)'s converging pointwise almost everywhere to \( f \), and we write this limit of integrals as \( \int_X f \, d\mu \). Integration with respect to \( \mu \) is a linear map \( L^1(\mu) \to \mathbb{C} \). When \( f \) is a pointwise limit almost everywhere of an \( L^1 \)-Cauchy sequence of step maps, the function \( |f| \) is also such a limit, so \( |f| \in L^1(\mu, \mathbb{R}) \). We also have \( |\int_X f \, d\mu| \leq \int_X |f| \, d\mu \) ([1, p. 132]). Setting

\[
|f|_1 = \int_X |f| \, d\mu
\]

for \( f \in L^1(\mu) \), the function \( d(f, g) = |f - g|_1 \) has the properties of a metric on \( L^1(\mu) \) except it might be zero without \( f = g \) everywhere.

**Example 14.** In Theorem 13, passage to a subsequence is important: an \( L^1 \)-Cauchy sequence of step maps does not have to converge pointwise almost everywhere. To see an example, let \( X = [0, 1] \) and write down a sequence of closed intervals in \([0, 1] \) whose endpoints are fractions with the same denominator and consecutive numerators:

\[0, 1], [0, 1/2], [1/2, 1], [0, 1/3], [1/3, 2/3], [2/3, 1], [0, 1/4], [1/4, 2/4], [2/4, 3/4], [3/4, 1], \ldots\]

Let \( f_n \) be the characteristic function of the \( n \)th such interval. Using Lebesgue measure \( dx \) on \( X \), \( \int_X |f_n(x)| \, dx = \int_X f_n(x) \, dx \to 0 \) as \( n \to \infty \), so \( f_n \) is \( L^1 \)-convergent to 0 (draw a picture so you can see what’s going on!)! Therefore \( \{f_n\} \) is an \( L^1 \)-Cauchy sequence and but for each \( x \in X \) the numbers \( f_n(x) \) oscillate from 0 to 1 and back as \( n \to \infty \) (they are 0 more and more often as \( n \) grows, but not permanently for large \( n \)). Therefore \( L^1 \)-convergence of a sequence of functions does not have to imply pointwise convergence of the sequence of functions anywhere.

**Theorem 15.** If \( f \in L^1(\mu) \) is the \( L^1 \)-limit of a sequence of functions \( f_n \in L^1(\mu) \) then \( f \) is the pointwise limit almost everywhere of some subsequence of the \( f_n \)'s.

**Proof.** Replacing \( \{f_n\} \) by a subsequence, we may assume \( |f_n - f_{n+1}| \leq 1/2^n \) for all \( n \). Then it can be proved such a sequence converges pointwise almost everywhere. See [1, Theorem 5.2, p. 138]. The bound \( 1/2^n \) can be replaced by \( c^n \) for any fixed \( c \in (0, 1) \). \( \square \)

**Example 16.** In Example 14, the subsequence of \( f_n \)'s that are characteristic functions of the intervals \([0, 1/m] \) for \( m \to \infty \) tend pointwise to 0 everywhere in \([0, 1] \) except at \( x = 0 \), which is almost everywhere for Lebesgue measure.

**Corollary 17.** If \( f \) is \( \mu \)-measurable and \( \int_X |f| \, d\mu = 0 \) then \( f = 0 \) almost everywhere.

**Proof.** Let \( f_n = 0 \) for all \( n \), so \( f \) is an \( L^1 \)-limit of the sequence \( \{f_n\} \) for all \( n \). Thus a subsequence of the \( f_n \)'s converges pointwise to \( f \) almost everywhere, so \( f \) is 0 almost everywhere. \( \square \)

Theorem 15 tells us that \( L^1 \)-convergence implies pointwise convergence after passing to a suitable subsequence. The following theorem is something of a converse, giving conditions under which pointwise convergence of a sequence in \( L^1(\mu) \) implies \( L^1 \)-convergence. It is the most important basic result about limits of integrals.

**Theorem 18** (Dominated Convergence Theorem). If \( f_n \in L^1(\mu) \), \( f_n \to f \) pointwise almost everywhere, and there is a \( g \in L^1(\mu, \mathbb{R}) \) such that \( |f_n| \leq g \) almost everywhere on \( X \), then \( f \in L^1(\mu) \) and \( |f_n - f|_1 \to 0 \), so in particular \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \).

**Proof.** See [1, Theorem 5.8, p. 141]. \( \square \)
Any $L^1$-function $g$ in the dominated convergence theorem is said to dominate the $f_n$'s.

**Example 19.** We will use the dominated convergence theorem to prove the intuitively reasonable result that integrating an integrable function over small subsets of $X$ should give small values:

$$f \in L^1(\mu) \implies \lim_{\mu(A) \to 0} \int_A f \, d\mu = 0.$$  

That is, for each $\varepsilon > 0$ we will show there is a $\delta > 0$ such that any measurable subset $A$ of $X$ with $\mu(A) < \delta$ has $|\int_A f \, d\mu| < \varepsilon$. (Here $f$ is fixed.)

The proof is by contradiction. Assume there is no such $\delta$, so we can find a sequence of measurable subsets $Y_n \subseteq X$ with $\mu(Y_n) \to 0$ and $|\int_{Y_n} f \, d\mu| \geq \varepsilon$. For concreteness, we can select the $Y_n$'s so that $\mu(Y_n) \leq 1/2^n$. Let $f_n = f\xi_{Y_n}$, i.e., $f_n$ is $f$ on $Y_n$ and 0 elsewhere. Set

$$Z = \{ x \in X : x \text{ is in } Y_n \text{ for infinitely many } n \}.$$  

This is a measurable subset of $X$ (its complement is the countable union of all intersections of finitely many $Y_n$'s at a time, which is measurable). For any $n$, $Z \subseteq Y_n \cup Y_{n+1} \cup Y_{n+2} \cup \cdots$, so $\mu(Z) \leq \mu(Y_n) + \mu(Y_{n+1}) + \cdots = 1/2^n - 1$. Thus $Z$ has measure 0. If $x \notin Z$ then $x \in Y_n$ for all but finitely many $n$, so $f_n(x) = 0$ for all large $n$ (where “large” depends on $x$). Hence $f_n \to 0$ almost everywhere on $X$ (namely on $X-Z$). Since $|f_n| \leq |f|$ on $X$ and $f \in L^1(\mu)$, by the dominated convergence theorem (in which $g$ there is $|f|$ here and $f$ there is the zero function) we have $\int_X f_n \, d\mu \to \int_X 0 \, d\mu = 0$. But $\int_X f_n \, d\mu = \int_{Y_n} f \, d\mu$ by the definition of $f_n$, so $\int_X f_n \, d\mu| \geq \varepsilon$ for all $n$. Taking $n$ large enough, we get a contradiction.

**Remark 20.** By Example 19, if $f \in L^1(\mu)$ and $A \subseteq X$ satisfies $\mu(A) = 0$ then $\int_A f \, d\mu = 0$. It follows (with some work) that the set-function $A \mapsto \int_A f \, d\mu$ (fixed $f$, varying $A$) is a finite measure on $X$. The converse is also true if $\mu$ is $\sigma$-finite: if $\nu$ is a finite measure on $X$ such that $\nu(A) = 0$ whenever $\mu(A) = 0$ then there is an $f \in L^1(\mu)$ such that $\nu(A) = \int_A f \, d\mu$, and such an $f$ is unique as an $L^1$-function on $X$ (that is, any other function fitting this rule equals $f$ almost everywhere on $X$). This is a special case of the Radon–Nikodym theorem. (Without $\sigma$-finiteness this converse can fail. Try $X = [0,1]$ with $\mu$ being counting measure and $\nu$ being Lebesgue measure.)

**Example 21.** The dominated convergence theorem is only generally valid for sequences of functions, not for nets of functions (i.e., not for functions indexed by directed sets other than the positive integers). For example, partially order the finite subsets $A \subseteq [0,1]$ by inclusion and let $f_A$ be the characteristic function of $A$. Then $0 \leq f_A \leq 1$ and the net $\{f_A\}$ tends to the integrable function 1 pointwise. However, $\int_0^1 f_A(x) \, dx = 0$ for all finite $A$ while $\int_0^1 1 \, dx = 1$. There are situations in real analysis where one wants to deal with a net of integrable functions rather than a sequence of integrable functions (e.g., integrals of the Newton quotients $(f(x+h) - f(x))/h$ as $h \to 0$), and a special argument is needed to justify applying the dominated convergence theorem to such a family of functions.

By definition, every $f : X \to \mathbb{C}$ that is in $L^1(\mu)$ is $\mu$-measurable. Which $\mu$-measurable functions $f : X \to \mathbb{C}$ are in $L^1(\mu)$? The next theorem answers this in terms of $|f|$.

**Corollary 22.** For $\mu$-measurable $f : X \to \mathbb{C}$, $f \in L^1(\mu)$ if and only if $|f| \in L^1(\mu, \mathbb{R})$.

**Proof.** See [1, Corollary 5.9, p. 142].
Note that Corollary 22 is the first time that something special has appeared for positive functions. Most treatments of integration place an unduly heavy emphasis on positive functions right at the start, leading to an unnatural method of defining the integral, by first defining it on nonnegative (measurable) functions, then on real functions using positive and negative parts, and then on complex functions using real and imaginary parts. Here we have described the passage from step maps on $X$ to general integrable functions on $X$ without starting from integration of positive functions.

**Corollary 23.** If $f \in L^1(\mu)$ and $\varphi: X \to \mathbb{C}$ is bounded and measurable then $\varphi f \in L^1(\mu)$.

*Proof.* See [1, Corollary 5.11, p. 143]. \hfill \square

**Example 24.** We will use Corollary 23 and the dominated convergence theorem to prove a “negative” result. Let $C[0,1]$ be the space of continuous functions $[0,1] \to \mathbb{C}$. For $f \in C[0,1]$ and $g \in L^1([0,1])$, the product $fg$ is integrable since $f$ is bounded, so $\int_0^1 fg \, dx$ makes sense. (Note $f$ and $g$ here play the respective roles of $\varphi$ and $f$ from Corollary 23.) In particular, sending $f$ to $\int_0^1 fg \, dx$ is an example of a linear function $C[0,1] \to \mathbb{C}$. So integration against the product of a fixed function in $L^1([0,1])$ provides many examples of continuous linear maps $C[0,1] \to \mathbb{C}$, where $C[0,1]$ is given the sup-norm.

Not all continuous linear maps $C[0,1] \to \mathbb{C}$ have the form $f \mapsto \int_0^1 fg \, dx$. Specifically, $f \mapsto f(0)$ is such a map and we will show there is no $g \in L^1([0,1])$ such that

$$f(0) = \int_0^1 fg \, dx$$

for all $f \in C[0,1]$. That is, evaluation at 0 is not the same as multiplying by a fixed function and then integrating on $[0,1]$.

Assume otherwise: such a $g$ in $L^1([0,1])$ exists. Let $f_n \in C[0,1]$ be the piecewise linear function whose graph is the line segment between the points $(0,1)$ and $(1/n,0)$ and is 0 for $1/n \leq x \leq 1$. Then $|f_n(x)| \leq 1$ for all $x$, so $|f_n(x)g(x)| \leq |g(x)|$. Thus $|f_n g| \leq |g|$ pointwise. For each $x \neq 0$, $f_n(x)g(x) \to 0$ as $n \to \infty$ simply because $g(x)$ is staying put while (for fixed $x$) $f_n(x) = 0$ for $n \gg 0$. Therefore we can apply the dominated convergence theorem (the functions $|f_n g|$ are dominated by $|g|$, which is in $L^1$ by Corollary 22) to say $\int_0^1 f_n g \, dx \to 0$ as $n \to \infty$. But, since each $f_n$ is in $C[0,1]$, by hypothesis $\int_0^1 f_n g \, dx = f_n(0) = 1$ for all $n$.

Measures and their integrals lead to linear maps on spaces of functions. We can also go the other way in important cases, starting with suitable linear maps and recovering a measure, as follows.

**Theorem 25** (Riesz representation theorem). Let $X$ be a locally compact Hausdorff space and let $C_c(X)$ be the continuous functions $X \to \mathbb{C}$ with compact support. Any linear map $L: C_c(X) \to \mathbb{C}$ such that $L(f) \geq 0$ when $f(x) \geq 0$ for all $x$ has the form $L(f) = \int_X f \, d\mu$ for a unique Borel measure $\mu$ on $X$ with the following properties:

1. $\mu(K) < \infty$ for compact $K$ in $X$,
2. for any Borel set $A$ in $X$, $\mu(A) = \inf_{U \supset A} \mu(U)$ over the open $U$ containing $A$,
3. for any open set or $\sigma$-finite Borel set $A$ in $X$, $\mu(A) = \sup_{K \subset A} \mu(K)$ over the compact $K$ in $X$ that are contained in $A$.

**Example 26.** Take $X = [0,1]$, so Theorem 25 says any linear function $C[0,1] \to \mathbb{C}$ that sends nonnegative continuous functions to nonnegative numbers is integration against a
unique Borel measure. Consider the functional \( L(f) = f(0) \) from Example 24. The corresponding measure is the Dirac measure \( \delta_0 \) at 0: \( \delta_0(A) = 1 \) if 0 \( \in \) \( A \) and \( \delta_0(A) = 0 \) if 0 \( \not\in \) \( A \), so \( \int_0^1 f \, d\delta_0 = f(0) \) for continuous \( f \). The Dirac measure at 0 is not of the form \( g \, dx \) for any \( g \in L^1([0,1]) \) since \( \int_A g \, dx = 0 \) when \( A \) has Lebesgue measure 0 (Example 19) but \( \delta_0 \) doesn’t have this property for all \( A \) (take \( A = \{0\} \)). This gives a different solution to Example 24.

If we consider linear maps \( L: C_c(X) \rightarrow \mathbb{C} \) that are continuous (with respect to the sup-norm on \( C_c(X) \)) rather than positive, then there is an analogue of the Riesz representation theorem: such \( L \) have the form \( L(f) = \int_X f \, d\mu \) where \( \mu \) is a complex Borel measure. These two representation theorems for linear maps on \( C_c(X) \) have an overlap, but neither is a special case of the other in general since positive \((i.e.,\ ordinary)\) measures are not always complex measures: the former can assign the value \( \infty \) to subsets while this is ruled out for complex measures.

References