

Measure Theory

1. σ -algebras

A measure is a function which gauges the “size” of subsets of a given set. In general we do not ask that a measure evaluate the size of every subset, since already in the prototypical example of volume in \mathbf{R}^n there exist pathological sets for which the concept of volume is not meaningful. At the same time, the class of subsets of \mathbf{R}^n for which a workable notion of volume can be defined is surprisingly large, due to the fact that it enjoys countable closure properties. That is, it is a σ -algebra in the sense of the following definition.

Definition 1.1. Let X be a set. An *algebra of subsets of X* is a family of subsets of X which includes \emptyset and X and is closed under complements, finite unions, and finite intersections. A *σ -algebra of subsets of X* is a family of subsets of X which includes \emptyset and X and is closed under complements, countable unions, and countable intersections.

A *measurable space* is a set X equipped with a σ -algebra Ω , and a subset $A \subseteq X$ is *measurable* if $A \in \Omega$.

Example 1.2. Let X be any set and let $\Omega = \mathcal{P}(X)$ be the power set of X (the set of all subsets of X). Then Ω is a σ -algebra.

Example 1.3. Let X be an uncountable set and let Ω be the set of all countable and co-countable subsets of X . That is, $A \in \Omega$ if A is countable or $X - A$ is countable. Ω is again a σ -algebra.

Example 1.4. Let X be a metric space and let $\Omega_{Bor}(X)$ be the σ -algebra generated by the open subsets of X . This is the *Borel σ -algebra* on X . Thus, A is Borel measurable if A belongs to every σ -algebra which contains all of the open sets.

Taking complements, we see that the Borel σ -algebra contains every closed set. Going one step further, it also contains all F_σ sets (countable unions of closed sets) and all G_δ sets (countable intersections of open sets).

There is an “internal” description of the σ -algebra generated by a given family of sets, but it is too elaborate to be of much use. Starting with the generating family of sets, one takes all complements, countable unions, and countable intersections to get a larger family; then one takes all complements, countable unions, and countable intersections of that family to get a larger family, and so on. This process can be continued, and in general uncountably many iterations (specifically, \aleph_1 iterations) are required before one obtains a σ -algebra. This is already the case for the Borel σ -algebra on \mathbf{R} , starting with the open sets as a generating family.

Next we consider two basic constructions with σ -algebras, and then a more sophisticated one.

Proposition 1.5. Let X be a set, let Ω be a σ -algebra of subsets of X , and let Y be any measurable subset of X . Then the family of measurable sets contained in Y is a σ -algebra of subsets of Y .

Proposition 1.6. Let (X_n) be a sequence of disjoint sets and for each n let Ω_n be a σ -algebra of subsets of X_n . Then the family

$$\left\{ A \subseteq \bigcup X_n : A \cap X_n \in \Omega_n \text{ for all } n \right\}$$

is a σ -algebra of subsets of $X = \bigcup X_n$.

Definition 1.7. Let (X_n) be a finite or infinite sequence of sets, and for each n let Ω_n be a σ -algebra of subsets of X_n . The *product σ -algebra* is the σ -algebra $\otimes \Omega_n$ of subsets of $\prod X_n$ which is generated by the sets

$$\pi_m^{-1}(A) = \left\{ \vec{x} \in \prod X_n : x_m \in A \right\},$$

for varying values of m and $A \in \Omega_m$. Here π_m is the projection map from the product space onto the factor X_m defined by $\pi_m(\vec{x}) = x_m$.

Proposition 1.8. $\Omega_{Bor}(\mathbf{R}^n) = \otimes_1^n \Omega_{Bor}(\mathbf{R})$ for all n .

Proof. Fix n , let $1 \leq i \leq n$, and consider the family Ω of sets $A \subseteq \mathbf{R}^n$ such that $\pi_i^{-1}(A) \in \Omega_{Bor}(\mathbf{R}^n)$. Since $\Omega_{Bor}(\mathbf{R}^n)$ is a σ -algebra and π^{-1} commutes with unions, intersections, and complements, it follows that Ω is a σ -algebra. Also, if $U \subseteq \mathbf{R}^n$ is open then $\pi_i^{-1}(U)$ is an open subset of \mathbf{R}^n and hence belongs to $\Omega_{Bor}(\mathbf{R}^n)$, so Ω is a σ -algebra that contains every open set, and hence it contains $\Omega_{Bor}(\mathbf{R}^n)$. That is, for every $A \in \Omega_{Bor}(\mathbf{R}^n)$ we have $\pi_i^{-1}(A) \in \Omega_{Bor}(\mathbf{R}^n)$. Since i was arbitrary, this implies that $\otimes_1^n \Omega_{Bor}(\mathbf{R}) \subseteq \Omega_{Bor}(\mathbf{R}^n)$.

Conversely, for each $1 \leq i \leq n$ let (a_i, b_i) be a nonempty, finite open interval in \mathbf{R} . Then

$$\prod_{i=1}^n (a_i, b_i) = \bigcap_{i=1}^n \pi_i^{-1}((a_i, b_i))$$

belongs to $\otimes_1^n \Omega_{Bor}(\mathbf{R})$. But every open subset of \mathbf{R}^n is a countable union of sets of this form, so $\otimes_1^n \Omega_{Bor}(\mathbf{R})$ contains all open subsets of \mathbf{R}^n , and therefore it contains all of $\Omega_{Bor}(\mathbf{R}^n)$. ■

We conclude this section with some variations on the definition of σ -algebras.

Proposition 1.9. Let X be a set and let Ω be a nonempty family of subsets of X .

- (a) If Ω is closed under complements and countable unions then Ω is a σ -algebra.
- (b) If Ω is an algebra and Ω is closed under countable disjoint unions then Ω is a σ -algebra.
- (c) If Ω is an algebra and Ω is closed under countable increasing unions then Ω is a σ -algebra.

Proof. (a) Since Ω is nonempty, it contains some set A , and hence the hypothesis implies that it contains $X = A \cup A^c$ and $\emptyset = X^c$. Furthermore, the identity $\bigcap A_n = (\bigcup A_n^c)^c$ shows that closure under complements and countable unions implies closure under countable intersections.

(b) Suppose Ω is an algebra which is closed under countable disjoint unions and let $A_n \in \Omega$ ($n \in \mathbf{N}$). Let $B_1 = A_1$ and for $n > 1$ let $B_n = A_n - (A_1 \cup \dots \cup A_{n-1})$. Then the B_n are disjoint, and each B_n belongs to Ω since Ω is an algebra. By assumption $\bigcup A_n = \bigcup B_n \in \Omega$, so Ω is closed under arbitrary countable unions, and hence Ω is a σ -algebra by part (a).

(c) Exercise. ■

Exercises.

1.A. Show that $\Omega_{Bor}(\mathbf{R})$ is generated by sets of the form (a, ∞) for $a \in \mathbf{Q}$.

1.B. Let (X_n) be a sequence of sets equipped with σ -algebras Ω_n . Show that $\otimes \Omega_n$ is the σ -algebra generated by the subsets of $\prod X_n$ of the form $\prod A_n$ where each $A_n \in \Omega_n$.

1.C. Prove Proposition 1.9 (c).

2. Measurable functions

In this section we consider the maps which are compatible with σ -algebras.

Definition 2.1. Let X and Y be measurable spaces. A function $f : X \rightarrow Y$ is *measurable* if $f^{-1}(A)$ is measurable in X for every measurable set $A \subseteq Y$.

Proposition 2.2. Let X, Y , and Z be measurable spaces and let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be measurable. Then $g \circ f : X \rightarrow Z$ is measurable.

Proposition 2.3. Let X and Y be metric spaces and let $f : X \rightarrow Y$ be continuous. Then f is measurable with respect to the Borel σ -algebras on X and Y .

Proof. Let Ω be the family of all subsets A of Y such that $f^{-1}(A) \in \Omega_{Bor}(X)$. Then Ω contains all open subsets of Y since f is continuous, and Ω is a σ -algebra since $\Omega_{Bor}(X)$ is a σ -algebra and f^{-1} commutes with unions, intersections, and complements. Thus Ω contains $\Omega_{Bor}(Y)$, so f is measurable. ■

Proposition 2.4. Let X and Y_n be measurable spaces and equip $\prod Y_n$ with the product σ -algebra. Then a function $f : X \rightarrow \prod Y_n$ is measurable $\Leftrightarrow \pi_n \circ f : X \rightarrow Y_n$ is measurable for all n .

Proof. Exercise. ■

When we consider functions into \mathbf{R} , we will always assume that \mathbf{R} is equipped with the Borel σ -algebra and we will take “measurable” to mean “measurable with respect to $\Omega_{Bor}(\mathbf{R})$.”

Theorem 2.5. Let X be a measurable space.

- (a) If $f, g : X \rightarrow \mathbf{R}$ are measurable then so are $f + g$ and fg .
- (b) Let (f_n) be a pointwise bounded sequence of measurable functions from X into \mathbf{R} . Then $\sup f_n$, $\inf f_n$, $\limsup f_n$, and $\liminf f_n$ (all defined pointwise) are measurable.
- (d) Let (f_n) be a sequence of measurable functions from X into \mathbf{R} . If the pointwise limit $\lim f_n$ exists then it is measurable.

Proof. (a) Suppose $f, g : X \rightarrow \mathbf{R}$ are measurable. Then the function $x \mapsto (f(x), g(x))$ is measurable from X into $\mathbf{R} \times \mathbf{R}$ by Proposition 2.4, and the functions $(s, t) \mapsto s + t$ and $(s, t) \mapsto st$ are continuous from $\mathbf{R} \times \mathbf{R}$ to \mathbf{R} , so they are measurable by Proposition 2.3. Therefore the compositions $f + g : x \mapsto f(x) + g(x)$ and $fg : x \mapsto f(x)g(x)$ are measurable.

(b) Let $g = \inf f_n$ be the pointwise infimum of the f_n . For each $a \in \mathbf{R}$ we have

$$g^{-1}((a, \infty)) = \bigcup_{n=1}^{\infty} f_n^{-1}((a, \infty)),$$

so each of these sets is measurable in X . Thus the subsets A of \mathbf{R} for which $g^{-1}(A)$ is measurable in X constitute a σ -algebra Ω of subsets of \mathbf{R} which contains all intervals of the form (a, ∞) . Taking complements, Ω contains all intervals of the form $(-\infty, b]$, taking intersections it contains all intervals of the form $(a, b]$, and taking countable unions it contains all intervals of the form (a, b) . Taking countable unions again, we see that Ω contains all open sets in \mathbf{R} , and therefore Ω contains $\Omega_{Bor}(\mathbf{R})$. We conclude that g is measurable.

Measurability of $\sup f_n$ is proven similarly, and measurability of $\limsup f_n$ and $\liminf f_n$ follow because

$$\limsup f_n = \inf_m \sup_{n \geq m} f_n \quad \text{and} \quad \liminf f_n = \sup_m \inf_{n \geq m} f_n.$$

(c) If $f_n \rightarrow f$ pointwise then $f = \limsup f_n = \liminf f_n$, so f is measurable by part (c). ■

In measure theory it is often useful to consider functions which take values in the extended real line $\overline{\mathbf{R}} = [-\infty, \infty]$. We equip $\overline{\mathbf{R}}$ with the σ -algebra generated by the intervals $[a, b]$ where $a, b \in \overline{\mathbf{R}}$ and $a < b$. We have the following version of Theorem 2.5 for functions into $\overline{\mathbf{R}}$.

Theorem 2.6. *Let X be a measurable space.*

(a) *Let $f, g : X \rightarrow \overline{\mathbf{R}}$ be measurable and define $f + g$ using the conventions $\infty + a = \infty$ and $-\infty + a = -\infty$ for all $a \in \mathbf{R}$, and $\infty + \infty = \infty$, $\infty - \infty = 0$, and $-\infty - \infty = -\infty$. Then $f + g$ is measurable.*

(b) *Let $f, g : X \rightarrow \overline{\mathbf{R}}$ be measurable and define fg using the conventions $\infty \cdot 0 = -\infty \cdot 0 = 0$, $\pm\infty \cdot a = \pm\infty$ for $a \in (0, \infty]$, and $\pm\infty \cdot a = \mp\infty$ for $a \in [-\infty, 0)$. Then fg is measurable.*

(c) *Let (f_n) be a sequence of measurable functions from X into $\overline{\mathbf{R}}$. Then $\sup f_n$, $\inf f_n$, $\limsup f_n$, and $\liminf f_n$ (all defined pointwise) are measurable.*

(d) *Let (f_n) be a sequence of measurable functions from X into $\overline{\mathbf{R}}$. If the pointwise limit $\lim f_n$ exists then it is measurable.*

Proof. The proof is an easy, but slightly tedious, reduction to Theorem 2.5. The key tool is the following fact. Let $f : X \rightarrow \overline{\mathbf{R}}$ be a function and let $Y = f^{-1}((-\infty, \infty))$. Then f is measurable if and only if $f|_Y$ is a measurable function from Y into \mathbf{R} and Y , $f^{-1}(\infty)$, and $f^{-1}(-\infty)$ are measurable sets (exercise). ■

Exercises.

2.A. Prove Proposition 2.4.

2.B. Let X be a measurable space and let (f_n) be a sequence of measurable functions into \mathbf{R} . Show that $\{x \in X : \lim f_n(x) \text{ exists}\}$ is measurable.

2.C. Let $f : \mathbf{R} \rightarrow \mathbf{R}$ be monotone increasing (i.e., $s < t$ implies $f(s) \leq f(t)$). Show that f is measurable, equipping both domain and range with the Borel σ -algebra.

2.D. Let $f : X \rightarrow \overline{\mathbf{R}}$ be a function and let $Y = f^{-1}((-\infty, \infty))$. Show that f is measurable if and only if $f|_Y$ is a measurable function from Y into \mathbf{R} and Y , $f^{-1}(\infty)$, and $f^{-1}(-\infty)$ are measurable sets.

3. Positive measures

Definition 3.1. Let X be a set.

(a) A function μ from an algebra of subsets of X into $[0, \infty]$ is *finitely additive* if

$$\mu(A_1 \cup \cdots \cup A_n) = \mu(A_1) + \cdots + \mu(A_n)$$

whenever $n \in \mathbf{N}$ and A_1, \dots, A_n are disjoint sets in the algebra.

(b) A *measure* on X is a function μ from a σ -algebra of subsets of X into $[0, \infty]$ such that $\mu(\emptyset) = 0$ and

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

whenever (A_n) is a sequence of disjoint sets in the σ -algebra.

A *measure space* (X, μ) is a set X equipped with a measure μ . It is a *probability measure* if $\mu(X) = 1$; *finite* if $\mu(X) < \infty$; and *σ -finite* if there exists a sequence (A_n) of measurable sets such that $X = \bigcup A_n$ and $\mu(A_n) < \infty$ for all n .

Any of the notations X , (X, μ) , or (X, Ω, μ) can be used for a measure space, depending on how much one wants to make explicit. Sometimes we use the term “positive measure” to avoid confusion with signed measures, which we will introduce later.

Example 3.2. Let X be a set, let $\Omega = \mathcal{P}(X)$ as in Example 1.2, and define $\mu(A)$ to be the cardinality of A if A is finite, and ∞ if A is infinite. This is *counting measure*. It is finite if X is finite and σ -finite if X is countable.

Example 3.3. Let X be an uncountable set, let Ω be the σ -algebra of countable and co-countable sets, and define $\mu(A)$ to be 0 if A is countable and 1 if A is co-countable. This is a finite measure.

The basic properties of measures are collected in the following theorem. Property (a) is called “monotonicity” and property (b) is called “subadditivity.”

Theorem 3.4. Let (X, Ω, μ) be a measure space.

(a) If $A, B \in \Omega$ and $A \subseteq B$ then $\mu(A) \leq \mu(B)$.

(b) If $(A_n) \subseteq \Omega$ then $\mu(\bigcup_1^{\infty} A_n) \leq \sum_1^{\infty} \mu(A_n)$.

(c) If $(A_n) \subseteq \Omega$ and $A_1 \subseteq A_2 \subseteq \cdots$ then $\mu(\bigcup_1^{\infty} A_n) = \lim \mu(A_n)$.

(d) If $(A_n) \subseteq \Omega$, $A_1 \supseteq A_2 \supseteq \cdots$, and $\mu(A_N) < \infty$ for some N , then $\mu(\bigcap_1^{\infty} A_n) = \lim \mu(A_n)$.

Proof. (a) If $A \subseteq B$ then $\mu(B) = \mu(B - A) + \mu(A) \geq \mu(A)$.

(b) Let $(A_n) \subseteq \Omega$. Let $B_1 = A_1$ and for $n > 1$ define $B_n = A_n - (A_1 \cup \cdots \cup A_{n-1})$. Then

$$\mu\left(\bigcup A_n\right) = \mu\left(\bigcup B_n\right) = \sum \mu(B_n) \leq \sum \mu(A_n),$$

using part (a) and the fact that $B_n \subseteq A_n$.

(c) Assuming $(A_n) \subseteq \Omega$ is an increasing sequence, let $B_1 = A_1$ and for $n > 1$ define $B_n = A_n - A_{n-1}$. Then

$$\mu\left(\bigcup A_n\right) = \mu\left(\bigcup B_n\right) = \sum \mu(B_n).$$

But $\sum_1^{\infty} \mu(B_n) = \lim_{N \rightarrow \infty} \sum_1^N \mu(B_n) = \lim_{N \rightarrow \infty} \mu(A_N)$, which yields the desired result.

(d) For $n \geq N$ define $B_n = A_N - A_n$. Applying (c) to the sequence (B_n) , we obtain $\mu(\bigcup B_n) = \lim \mu(B_n)$; thus

$$\mu\left(\bigcap A_n\right) = \mu(A_N) - \mu\left(\bigcup B_n\right) = \mu(A_N) - \lim \mu(B_n) = \lim \mu(A_n),$$

as desired. ■

The condition $\mu(A_N) < \infty$ for some N in part (d) is necessary. For example, consider \mathbf{N} with counting measure and let $A_n = \{n, n+1, \dots\}$. Then $\mu(A_n) = \infty$ for all n but $\mu(\bigcap A_n) = \mu(\emptyset) = 0$.

Next, we give two simple ways to construct new measure spaces from old ones.

Proposition 3.5. *Let (X, Ω, μ) be a measure space and let $Y \subseteq X$ be a measurable subset of X . Then the restriction of μ to the family of measurable sets contained in Y is a measure on Y .*

Proposition 3.6. *Let (X_n, Ω_n, μ_n) be a disjoint sequence of measure spaces. Then*

$$\mu(A) = \sum \mu_n(A \cap X_n)$$

defines a measure on the σ -algebra on $X = \bigcup X_n$ introduced in Proposition 1.6.

Proof. Exercise. ■

Definition 3.7. Let (X, Ω, μ) be a measure space. A set $A \in \Omega$ of measure zero (i.e., $\mu(A) = 0$) is called a *null set*. A statement $P(x)$ is true for *almost every* x (abbreviated *a.e.* x) if there is a null set A such that $P(x)$ is true for every $x \in X - A$. The measure μ is *complete* if every subset of every null set is measurable.

By part (a) of Theorem 3.4, any measurable subset of a set of measure zero must also have measure zero. However, it is possible that measure zero sets could have non-measurable subsets. We now show that the domain of any measure can always be enlarged so that all subsets of measure zero sets are measurable.

Proposition 3.8. *Let (X, Ω, μ) be a measure space. Define $\overline{\Omega}$ to be all sets of the form $A \cup B$ such that $A \in \Omega$ and B is contained in a null set, and define $\overline{\mu}(A \cup B) = \mu(A)$ for all such A and B . Then $(X, \overline{\Omega}, \overline{\mu})$ is a complete measure space.*

Proof. To see that $\overline{\Omega}$ is closed under countable unions, let $(A_n \cup B_n) \subseteq \overline{\Omega}$ with each $A_n \in \Omega$ and each $B_n \subseteq N_n \in \Omega$ where $\mu(N_n) = 0$. Then

$$\bigcup (A_n \cup B_n) = \bigcup A_n \cup \bigcup B_n;$$

$\bigcup A_n \in \Omega$ since Ω is a σ -algebra, $\bigcup B_n \subseteq \bigcup N_n \in \Omega$, and $\mu(\bigcup N_n) \leq \sum \mu(N_n) = 0$. So $\bigcup (A_n \cup B_n) \in \overline{\Omega}$.

To see that $\overline{\Omega}$ is closed under complementation, let $A \cup B \in \overline{\Omega}$ with $A \in \Omega$ and $B \subseteq N \in \Omega$, $\mu(N) = 0$. Then

$$X - (A \cup B) = (X - (A \cup N)) \cup (N - (A \cup B)) = A' \cup B'$$

where $A' \in \Omega$ and $B' \subseteq N$, so $X - (A \cup B) \in \overline{\Omega}$. By Proposition 1.9 (a), $\overline{\Omega}$ is a σ -algebra.

Next we check that $\overline{\mu}$ is well-defined. Suppose $A_1 \cup B_1 = A_2 \cup B_2$ where B_1 and B_2 are contained in null sets. Then $A_1 \Delta A_2 = (A_1 - A_2) \cup (A_2 - A_1)$ is a null set, so $\mu(A_1) = \mu(A_2)$. This shows that $\overline{\mu}$ is well-defined.

If $(A_n \cup B_n)$ is a disjoint sequence in $\overline{\Omega}$ then (A_n) is a disjoint sequence in Ω and

$$\overline{\mu}\left(\bigcup (A_n \cup B_n)\right) = \mu\left(\bigcup A_n\right) = \sum \mu(A_n) = \sum \overline{\mu}(A_n \cup B_n).$$

Thus $\overline{\mu}$ is a measure. Finally, if $\overline{\mu}(A \cup B) = 0$ then $\mu(A) = 0$, and so any subset of $A \cup B$ is a subset of a null set in Ω and hence belongs to $\overline{\Omega}$. So $\overline{\mu}$ is complete. ■

Finally, we present a remarkable result about convergence.

Theorem 3.9. (Egoroff) *Let (X, μ) be a finite measure space and let f_n and f be measurable scalar-valued functions on X . Suppose $f_n \rightarrow f$ a.e. Then for every $\epsilon > 0$ there exists $A \subseteq X$ such that $f_n \rightarrow f$ uniformly on $X - A$ and $\mu(A) \leq \epsilon$.*

Proof. Let $B = \{x \in X : f_n(x) \rightarrow f(x)\}$. Then $\mu(X - B) = 0$ by hypothesis. Let $\epsilon > 0$. We claim that for any $\alpha > 0$ there exists a set $A \subseteq X$ such that $\mu(A) \leq \epsilon$ and for sufficiently large n we have $|f_n - f| < \alpha$ on $X - A$. To prove the claim, for each n let

$$B_n = \bigcup_{m=n}^{\infty} \{x \in B : |f_m(x) - f(x)| \geq \alpha\}.$$

Since this sequence decreases to \emptyset , we can find $N \in \mathbf{N}$ such that $\mu(B_N) \leq \epsilon$ by Theorem 3.4 (d). Then $n \geq N$ implies that $|f_n - f| < \alpha$ off of B_N . This establishes the claim.

Now for each $k \in \mathbf{N}$, use the claim to find a set A_k such that $\mu(A_k) \leq \epsilon/2^k$ and for sufficiently large n we have $|f_n - f| < 1/k$ on $X - A_k$. Then let $A = (X - B) \cup \bigcup A_k$. We have $\mu(A) \leq \epsilon$ and $f_n \rightarrow f$ uniformly on $X - A$. ■

For a counterexample with $\mu(X) = \infty$, let X be \mathbf{N} with counting measure and let $f_n(k) = 0$ for $k < n$ and 1 for $k \geq n$. Then $f_n \rightarrow 0$ everywhere, but convergence is uniform only on finite measure (i.e., finite) subsets of \mathbf{N} .

Exercises.

3.A. Let (X, μ) be a measure space, suppose μ is σ -finite, and let $A \subseteq X$ be measurable with $\mu(A) = \infty$. Show that for any $N \in \mathbf{N}$ there exists a measurable set $B \subseteq A$ such that $N \leq \mu(B) < \infty$.

3.B. Prove Proposition 3.6.

3.C. Let μ be a complete measure on a set X and suppose $f_n \rightarrow f$ a.e., with each f_n measurable. Show that f is measurable.

4. Outer measures

The most basic tool for constructing nontrivial measures is the technique of outer measures. In this approach we first define a simpler structure called a premeasure, and then we invoke a general result which automatically enlarges the premeasure to a measure.

Definition 4.1. Let X be a set.

(a) A *ring of subsets of X* is a family of subsets of X which includes \emptyset and is stable under finite unions and differences.

(b) A *premeasure* on X is a function μ from a ring of subsets of X into $[0, \infty)$ such that $\mu(\emptyset) = 0$ and $\mu(\bigcup A_n) = \sum \mu(A_n)$ whenever (A_n) is a disjoint sequence of sets in the algebra whose union also belongs to the algebra.

(c) An *outer measure* on X is a function $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$ such that $\mu^*(\emptyset) = 0$, $A \subseteq B$ implies $\mu^*(A) \leq \mu^*(B)$, and $\mu^*(\bigcup A_n) \leq \sum \mu^*(A_n)$ for any sequence $(A_n) \subseteq \mathcal{P}(X)$.

In particular, any premeasure is finitely additive. Note that premeasures are required to take finite values (which is why they are defined on a ring, not on an algebra).

First we show how to get an outer measure from a premeasure. We adopt the convention that the infimum of the empty set is ∞ .

Theorem 4.2. *Let μ be a premeasure defined on a ring Ω_0 of subsets of a set X . For any $A \subseteq X$ define*

$$\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) : \text{each } A_n \in \Omega_0 \text{ and } A \subseteq \bigcup_{n=1}^{\infty} A_n \right\}.$$

Then μ^* is an outer measure and $\mu^*(A) = \mu(A)$ for all $A \in \Omega_0$.

Proof. We get $\mu^*(\emptyset) = 0$ by taking $A_n = \emptyset$ for all n . The condition $A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$ is trivial, since any cover of B will be a cover of A and hence the infimum that defines $\mu^*(B)$ will be taken over a smaller set than the infimum that defines $\mu^*(A)$. If (A_n) is any sequence of subsets of X and $\epsilon > 0$, for each n we can find a cover (A_n^i) of A_n such that

$$\sum_i \mu(A_n^i) \leq \mu^*(A_n) + \epsilon/2^n,$$

and then the family $\{A_n^i : i, n \in \mathbf{N}\}$ is a cover of $\bigcup A_n$ which verifies

$$\mu^* \left(\bigcup_n A_n \right) \leq \sum_{i,n} \mu(A_n^i) \leq \sum_n (\mu^*(A_n) + \epsilon/2^n) = \epsilon + \sum_n \mu^*(A_n).$$

Taking $\epsilon \rightarrow 0$ completes the proof that μ^* is an outer measure.

Let $A \in \Omega_0$. Then by taking $A_1 = A$ and $A_n = \emptyset$ for $n > 1$, we obtain $\mu^*(A) \leq \mu(A)$. Conversely, let (A_n) be any sequence in Ω_0 whose union contains A . Define $B_1 = A \cap A_1$ and

$$B_n = A \cap (A_n - (A_1 \cup \dots \cup A_{n-1}))$$

for $n > 1$. Then the sets B_n are disjoint and their union equals A , so $\mu(A) = \sum \mu(B_n)$ since μ is a premeasure. Thus

$$\sum \mu(A_n) \geq \sum \mu(B_n) = \mu(A),$$

which shows that $\mu^*(A) \geq \mu(A)$. So $\mu^*(A) = \mu(A)$. ■

Now we show how to get a measure from a premeasure. Let $A \Delta B$ be the symmetric difference of A and B defined by $A \Delta B = (A - B) \cup (B - A)$.

Theorem 4.3. *Let μ be a premeasure on a set X . Then μ extends to a complete measure on X .*

Proof. We will assume that X belongs to the ring on which the premeasure is defined, and thus $\mu(X) < \infty$. This result generalizes to the σ -finite case by decomposing X as a disjoint union of finite measure sets, applying the result to each component separately, and then invoking Proposition 3.6.

Let μ^* be the outer measure arising from μ as in Theorem 4.2. Let Ω_0 be the ring of sets on which μ is defined and let Ω be the family of subsets $A \subseteq X$ such that for every $\epsilon > 0$ there exists $A' \in \Omega_0$ with $\mu^*(A \Delta A') \leq \epsilon$. We claim that Ω is a σ -algebra and the restriction of μ^* to Ω is a measure. It is clear that Ω contains Ω_0 , and we know that μ^* agrees with μ on Ω_0 .

We start by showing that Ω is an algebra and μ^* is additive on Ω . Stability of Ω under complements follows from the fact that $\mu^*(A \Delta A') = \mu^*(A^c \Delta A'^c)$ (since $A \Delta A' = A^c \Delta A'^c$). For stability under finite unions let $A, B \in \Omega$, let $\epsilon > 0$, and find $A', B' \in \Omega_0$ such that $\mu^*(A \Delta A'), \mu^*(B \Delta B') \leq \epsilon$. Then $A' \cup B' \in \Omega_0$ and

$$\mu^*((A \cup B) \Delta (A' \cup B')) \leq \mu^*((A \Delta A') \cup (B \Delta B')) \leq \mu^*(A \Delta A') + \mu^*(B \Delta B') \leq 2\epsilon$$

since μ^* is monotone and subadditive. This is enough. Moreover, if A and B are disjoint then $A' \cap B' \subseteq (A \Delta A') \cup (B \Delta B')$, so

$$\mu(A' \cap B') \leq \mu^*((A \Delta A') \cup (B \Delta B')) \leq 2\epsilon,$$

and

$$\mu(A' \cup B') = \mu(A') + \mu(B') - \mu(A' \cap B')$$

since μ is additive. Since $|\mu^*(A \cup B) - \mu^*(A' \cup B')| \leq \mu^*((A \cup B) \Delta (A' \cup B')) \leq 2\epsilon$, this shows that $|\mu^*(A \cup B) - \mu^*(A) - \mu^*(B)| \leq 6\epsilon$, and we conclude that μ^* is additive on Ω .

We now show that Ω is a σ -algebra using Proposition 1.9 (b). Thus, let (A_n) be a disjoint sequence of sets in Ω and let $A = \bigcup A_n$. Since μ^* is finitely additive on Ω we must have $\sum_1^N \mu^*(A_n) \leq \mu(X)$ for all N , and hence $\sum_1^\infty \mu^*(A_n) \leq \mu(X)$. So for any $\epsilon > 0$ we can find N such that $\sum_{N+1}^\infty \mu^*(A_n) \leq \epsilon$, and by the fact that Ω is an algebra we can find $A' \in \Omega_0$ such that $\mu^*((A_1 \cup \dots \cup A_N) \Delta A') \leq \epsilon$. We then have $\mu^*(A \Delta A') \leq 2\epsilon$, and this shows that $A \in \Omega$. Also

$$\left| \mu^*(A) - \sum_1^N \mu^*(A_n) \right| \leq \epsilon$$

since $\mu^*(\bigcup_1^N A_n) = \sum_1^N \mu^*(A_n)$ and $\mu^*(\bigcup_{N+1}^\infty A_n) \leq \sum_{N+1}^\infty \mu^*(A_n) \leq \epsilon$. So μ^* is countably additive on Ω .

The completeness of μ^* on Ω is left as an exercise. ■

Exercises.

4.A. In Theorem 4.3, prove that the restriction of μ^* to Ω is complete.

4.B. Let μ be a premeasure on a set X such that $\mu(X) < \infty$ and let μ^* be the induced outer measure. Define the *inner measure* of any $A \subseteq X$ by $\mu_*(A) = \mu(X) - \mu^*(A^c)$. Show that A belongs to the σ -algebra Ω used in the proof of Theorem 4.3 if and only if $\mu^*(A) = \mu_*(A)$.

5. Lebesgue measure

Let a *half-open box* in \mathbf{R}^n be a set of the form $(a_1, b_1] \times \cdots \times (a_n, b_n]$ where $a_i < b_i$ for each i . Define the *volume* of this box to be $(b_1 - a_1) \cdot \cdots \cdot (b_n - a_n)$. Open and closed boxes are defined similarly, with the same formula for volume. If A is a finite disjoint union of half-open boxes, define $m(A)$ to be the sum of the volumes of the component boxes.

Proposition 5.1. *The family of finite disjoint unions of half-open boxes in \mathbf{R}^n is a ring of sets, and the function m is a premeasure.*

Proof. That the family of finite disjoint unions of half-open boxes is a ring follows from Exercise 5.A. We must now verify that m is well-defined, since the decomposition of a set into disjoint half-open boxes can generally be accomplished in more than one way. But any two such decompositions have a common refinement (just intersect every box in the first decomposition with every box in the second decomposition; the intersection of any two half-open boxes is a half-open box), so it is enough to show that the total volume of any finite disjoint union of half-open boxes equals the total volume of any refinement of it. This is straightforward though notationally tedious.

We want to show that $m(\bigcup A_k) = \sum m(A_k)$ whenever (A_k) is a disjoint sequence of sets in the ring whose union also belongs to the ring. It is immediate from the definition of m that m is finitely additive, and this implies that m is monotone (see the proof of Theorem 3.4 (a)). So for any $N \in \mathbf{N}$ we have

$$\sum_{k=1}^N m(A_k) = m\left(\bigcup_{k=1}^N A_k\right) \leq m\left(\bigcup_{k=1}^{\infty} A_k\right),$$

and taking $N \rightarrow \infty$ yields $\sum m(A_k) \leq m(\bigcup A_k)$. For the reverse inequality, observe that for any set A in the ring and any $\epsilon > 0$ there is a finite disjoint union of closed boxes which is contained in A and whose volume is $m(A) - \epsilon$, and a finite union of open boxes which contains A and whose total volume is $m(A) + \epsilon$. So letting $A = \bigcup A_k$, find a finite disjoint union B of closed boxes with volume $m(A) - \epsilon$ which is contained in A , and for each k find a finite union B_k of open boxes with total volume $m(A_k) + \epsilon/2^k$ which contains A_k . Then the B_k cover B and by compactness we can find a finite subcover. We conclude that

$$m(A) - \epsilon = m(B) \leq \sum m(B_k) = \sum (m(A_k) + \epsilon/2^k) = \epsilon + \sum m(A_k)$$

(defining $m(B)$ and $m(B_k)$ in the obvious way), and taking $\epsilon \rightarrow 0$ yields the desired reverse inequality. ■

Definition 5.2. The measure m on \mathbf{R}^n deduced via Theorems 4.2 and 4.3 from the premeasure defined above is *Lebesgue measure*, and its domain is the *Lebesgue σ -algebra*.

The Lebesgue σ -algebra on \mathbf{R}^n is easily seen to contain every open box and hence every open set. So it contains the Borel σ -algebra. The two σ -algebras will be related more closely in Theorem 5.5 (b).

When working in \mathbf{R}^n we generally assume “measurable” means “Lebesgue measurable.” Note that this conflicts slightly with the convention that measurability of real-valued functions is judged according to the Borel σ -algebra. Actually, for functions from \mathbf{R} to \mathbf{R} , we are interested both in functions which are Borel-Borel measurable and in functions which are Lebesgue-Borel measurable; these are called, respectively, Borel measurable and Lebesgue measurable functions. The composition of two Borel measurable functions is again Borel measurable, but the composition of two Lebesgue measurable functions need not be Lebesgue measurable.

In contrast, Borel-Lebesgue and Lebesgue-Lebesgue measurability are too restrictive to have much value. The Cantor set has measure zero – see Example 5.4 – so every subset of it is Lebesgue measurable. But there is a continuous, 1-1, increasing function from $[0, 1]$ into the Cantor set, and since there exist subsets of $[0, 1]$ which are not measurable – see Example 5.7 – this means that such a function is not Lebesgue-Lebesgue measurable. Thus, not even all continuous, monotone functions are Lebesgue-Lebesgue measurable.

Example 5.3. Let $\vec{a} = (a_1, \dots, a_n) \in \mathbf{R}^n$. Then $\{\vec{a}\} \subset (a_1 - \epsilon, a_1] \times \dots \times (a_n - \epsilon, a_n]$ for any $\epsilon > 0$. By monotonicity, we have $m(\{a\}) \leq \epsilon^n$ for all ϵ , and hence $m(\{a\}) = 0$. By countable additivity, the Lebesgue measure of any countable subset of \mathbf{R}^n is zero.

(This yields a new proof that $[0, 1]$ is uncountable, since its measure is 1.)

Example 5.4. Let K be the middle-thirds Cantor set,

$$K = [0, 1] - \left(\left(\frac{1}{3}, \frac{2}{3} \right) \cup \left(\frac{1}{9}, \frac{2}{9} \right) \cup \left(\frac{7}{9}, \frac{8}{9} \right) \cup \dots \right).$$

Then K is contained in one interval of length 1, in two intervals each of length $1/3$, in four intervals each of length $1/9$, and in general in 2^k intervals each of length $1/3^k$. Thus $m(K) \leq 2^k/3^k$, which goes to zero as $k \rightarrow \infty$, and we conclude that $m(K) = 0$.

This shows that the Cantor set is an uncountable set of measure zero. Since Lebesgue measure is complete (Theorem 4.3), every subset of the Cantor set is Lebesgue measurable. This leads to a simple proof by cardinality that there are more Lebesgue sets than Borel sets.

An interesting variation of Example 5.4 is obtained by removing smaller intervals. Instead of removing one interval of length $1/3$, two of length $1/3^2$, four of length $1/3^3$, etc., instead remove (say) one interval of length $1/3$, two of length $1/2 \cdot 3^2$, four of length $1/4 \cdot 3^3$, etc. Then the total length removed is $\sum_1^\infty \frac{1}{3^k} = 1/2$, so the set left remaining has measure $1/2$. This is a “fat” Cantor set, a set that is homeomorphic to the usual Cantor set but has strictly positive measure.

In the following result, a G_δ set is a countable intersection of open sets and an F_σ set is a countable union of closed sets.

Theorem 5.5. *Let $A \in \mathbf{R}^n$ be Lebesgue measurable. Then*

- (a) $m(A) = \inf\{m(U) : A \subseteq U \text{ and } U \text{ is open}\} = \sup\{m(K) : K \subseteq A \text{ and } K \text{ is compact}\};$
- (b) *we can write $A = A_1 - N_1 = A_2 \cup N_2$ where A_1 is G_δ , A_2 is F_σ , and N_1 and N_2 are null sets;*
- (c) *if $m(A) < \infty$ then for any $\epsilon > 0$ there is a finite family of open boxes B such that $m(A \Delta B) < \epsilon$.*

Proof. (a) If $A \subseteq U$ then $m(A) \leq m(U)$. Conversely, given $\epsilon > 0$, by the definition of outer measure we can find a sequence of half-open boxes B_k such that $A \subseteq \bigcup B_k$ and $m(A) \geq \sum m(B_k) - \epsilon$, and then for each k we can find an open box U_k which contains B_k and whose measure is $\epsilon/2^k$ greater. Then $A \subseteq \bigcup U_k$ and $m(A) \geq \sum m(U_k) - 2\epsilon$, which is enough to prove the first equality.

For the second equality, observe that $m(A) = \lim m(A \cap \text{ball}_N(0))$, so we can assume A is bounded. Let $\epsilon > 0$, let U be an open set containing $\bar{A} - A$ such that $m(U) \leq m(\bar{A} - A) + \epsilon$, and let $K = \bar{A} - U$. Then K is compact, $K \subseteq A$, and

$$m(K) \geq m(\bar{A}) - m(U) \geq m(\bar{A}) - m(\bar{A} - A) - \epsilon = m(A) - \epsilon.$$

This suffices to prove the second equality.

(b) Suppose $m(A) < \infty$. For each $k \in \mathbf{N}$ find an open set U_k such that $A \subseteq U_k$ and $m(A) \geq m(U_k) - 1/k$. Then $A_1 = \bigcap U_k$ is a G_δ set which contains A , and $m(A_1) \leq m(U_k) \leq m(A) + 1/k$ for all k , so $m(A_1) = m(A)$. Thus $N_1 = A_1 - A$ is a null set, and we have $A = A_1 - N_1$. The other decomposition of A follows in a similar way from internal approximation of A by compact sets.

If $m(A) = \infty$, for each $m, k \in \mathbf{N}$ find an open set U_k^m such that $A \cap (\text{ball}_{m+1}(0) - \text{ball}_m(0)) \subseteq U_k^m$ and

$$m(A \cap (\text{ball}_{m+1}(0) - \text{ball}_m(0))) \geq m(U_k^m) - 1/2^m k.$$

Let $A_1 = \bigcap_k \bigcup_m U_k^m$ and proceed as above.

(c) Assume $m(A) < \infty$ and find a sequence (U_k) as in the proof of part (a). Thus $A \subseteq \bigcup U_k$ and $m(A) \geq \sum m(U_k) - 2\epsilon$. Since $\bigcup U_k$ is a countable union of open boxes, for sufficiently large N the set $B = \bigcup_1^N U_k$ is a finite union of open boxes which satisfies $m(B) \geq m(\bigcup U_k) - \epsilon$. We then have

$$m(A \Delta B) = m(A - B) + m(B - A) \leq \epsilon + 2\epsilon = 3\epsilon,$$

which is enough. ■

The property described in Theorem 5.5 (a) known as “regularity”. Observe that Theorem 5.5 (b) shows that every Lebesgue subset of \mathbf{R}^n differs from a Borel subset of \mathbf{R}^n by a null set.

Theorem 5.6.

(a) *Lebesgue measure is translation invariant.*

(b) *Let $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$ be an invertible linear map and let $A \subseteq \mathbf{R}^n$ be Lebesgue measurable. Then $T(A)$ is Lebesgue measurable and $m(T(A)) = |\det(T)|m(A)$.*

(c) *Lebesgue measure is invariant under the action of the orthogonal group $O(n)$.*

Proof. (a) Let $\vec{x} \in \mathbf{R}^n$ and define $T_{\vec{x}} : \mathbf{R}^n \rightarrow \mathbf{R}^n$ by $T_{\vec{x}}\vec{y} = \vec{y} + \vec{x}$. Then $T_{\vec{x}}$ takes the ring generated by half-open boxes to itself, and $T_{\vec{x}}$ preserves the premeasure on this ring. It follows that $T_{\vec{x}}$ preserves the outer measure constructed from this premeasure via Theorem 4.2 and then that it preserves the measure constructed from this outer measure via Theorem 4.3. Thus Lebesgue measure is invariant under $T_{\vec{x}}$.

(b) Write T in matrix form. It is possible to convert T to the identity matrix by a sequence of operations of the following types: multiply a row by a nonzero scalar; subtract one row from another; interchange two rows. It follows that T can be written as a product of maps of the three types

$$\begin{aligned} (x_1, \dots, x_j, \dots, x_n) &\mapsto (x_1, \dots, ax_j, \dots, x_n) \\ (x_1, \dots, x_j, \dots, x_n) &\mapsto (x_1, \dots, x_j - x_k, \dots, x_n) \\ (x_1, \dots, x_j, \dots, x_k, \dots, x_n) &\mapsto (x_1, \dots, x_k, \dots, x_j, \dots, x_n). \end{aligned}$$

Since the determinant is multiplicative, it suffices to prove that $T(A)$ is measurable and that $m(T(A)) = |\det(T)|m(A)$ only for maps T of the above three types. It is easy to check this for maps of the first and third types. For maps of the second type we use translation invariance to check that a box has the same measure as its image under T (since the image can be decomposed into a finite number of pieces and reassembled to match the original box).

(c) $T \in O(n)$ if $\det(T) = \pm 1$, so this follows immediately from part (b). ■

We can use translation invariance to prove that there exist sets which are not Lebesgue measurable.

Example 5.7. (Vitali) For $s, t \in [0, 1)$ define $s \sim t$ if $t - s \in \mathbf{Q}$, and let A be a subset of $[0, 1)$ which contains exactly one element from each equivalence class. Consider the set

$$B = \bigcup_{r \in \mathbf{Q} \cap [0, 1)} A + r.$$

Observe that the sets $A + r$ are mutually disjoint: if $s \in (A + r_1) \cap (A + r_2)$ then $s - r_1$ and $s - r_2$ both belong to A , but their difference is rational, so we must have $r_1 = r_2$. Also, observe that $[0, 1)$ is the disjoint union of $B_1 = B \cap [0, 1)$ and $B_2 = (B \cap [1, 2)) - 1$: these sets are disjoint by an argument similar to the one which showed the sets $A + r$ are disjoint, and every element s of $[0, 1)$ is equivalent to some $t \in A$, yielding $s \in A + (s - t)$ if $s \geq t$ and $s \in A + (1 + s - t) - 1$ if $s < t$.

Suppose A were Lebesgue measurable. Then by translation invariance,

$$m(B) = \sum_{r \in \mathbf{Q} \cap [0, 1)} m(A + r) = \begin{cases} 0 & \text{if } m(A) = 0 \\ \infty & \text{if } m(A) > 0. \end{cases}$$

But $m(B) = m(B_1) + m(B_2) = m([0, 1]) = 1$, a contradiction. So A cannot be measurable.

An even worse phenomenon occurs in \mathbf{R}^3 :

Example 5.8. (Banach-Tarski) There is a set A in \mathbf{R}^3 such that the open ball $\text{ball}_1(0)$ is a union of two sets, both congruent to A , and it is also a union of three sets, each congruent to A .

Exercises.

5.A. Let X be a set and let Ω_0 be a family of subsets of X . Suppose that the difference of any two sets in Ω_0 is a finite disjoint union of sets in Ω_0 . Prove that the family of finite disjoint unions of sets in Ω_0 is a ring.

5.B. Let $F : \mathbf{R} \rightarrow \mathbf{R}$ satisfy (1) $s < t \Rightarrow F(s) \leq F(t)$ and (2) if s_n decreases to s then $F(s_n) \rightarrow F(s)$. Show that there is a Borel measure (i.e., a measure with domain $\Omega_{\text{Bor}}(\mathbf{R})$) μ_F such that $\mu_F((a, b]) = F(b) - F(a)$ for all $a, b \in \mathbf{R}$, $a < b$.

6. Integration of positive functions

Before defining integrals of real-valued functions, we begin in this section with integrals of positive functions. The positive case is easier because convergence is less of an issue in the same way that this is true of series (any sum of positive terms either converges to a finite value or diverges to infinity, and in either case rearranging the terms has no effect). Once we can integrate positive functions, we will define the integrals of real functions by taking differences of positive functions.

Definition 6.1. Let (X, μ) be a measure space.

(a) A *simple function* on X is a function $\phi : X \rightarrow \mathbf{R}$ with finite range. Every such function has a unique *standard form* $\phi = \sum_1^n a_i \chi_{A_i}$ with distinct coefficients a_i such that the A_i are disjoint and $\bigcup A_i = X$.

(b) Let $f : X \rightarrow [0, \infty]$ be measurable. We define the *integral* of f , denoted $\int f$ or $\int f d\mu$ or $\int f(x) d\mu(x)$, to be the (possibly infinite) value

$$\int f d\mu = \sup \left\{ \sum a_i \mu(A_i) \right\}$$

where the supremum is taken over all measurable simple functions ϕ with standard form $\phi = \sum a_i \chi_{A_i}$, such that $0 \leq \phi \leq f$. (As always, we use the convention $0 \cdot \infty = \infty \cdot 0 = 0$.)

(c) For $A \subseteq X$ and $f : X \rightarrow [0, \infty]$ measurable, we define $\int_A f d\mu$ to be the integral of $f|_A$ over the measure space $(A, \mu|_A)$.

We start by observing that the Lebesgue theory of integration subsumes the ordinary theory of summation.

Example 6.2. Let μ be counting measure on \mathbf{N} (Example 3.2). Then every function $f : \mathbf{N} \rightarrow [0, \infty]$ is measurable, and $\int f d\mu = \sum_{n \in \mathbf{N}} f(n)$.

Next we verify some basic properties of integrals of simple functions.

Lemma 6.3. Let (X, μ) be a measure space and let $\phi, \psi \geq 0$ be measurable simple functions with standard forms $\phi = \sum a_i \chi_{A_i}$ and $\psi = \sum b_j \chi_{B_j}$.

(a) If $\phi \leq \psi$ then $\sum a_i \mu(A_i) \leq \sum b_j \mu(B_j)$.

(b) $\int \phi = \sum a_i \mu(A_i)$.

(c) $\int (a\phi + b\psi) = a \int \phi + b \int \psi$ for all $a, b \geq 0$.

(d) The map $A \mapsto \int_A \phi$ is a measure on X .

Proof. (a) Suppose $\phi \leq \psi$. Then by finite additivity of μ we have

$$\sum_i a_i \mu(A_i) = \sum_{i,j} a_i \mu(A_i \cap B_j).$$

On the set $A_i \cap B_j$ the functions ϕ and ψ are constantly equal to a_i and b_j respectively, so $\phi \leq \psi$ implies that if $A_i \cap B_j \neq \emptyset$ then $a_i \leq b_j$. Thus

$$\sum_{i,j} a_i \mu(A_i \cap B_j) \leq \sum_{i,j} b_j \mu(A_i \cap B_j) = \sum_j b_j \mu(B_j),$$

as desired.

(b) This follows directly from part (a) and the definition of the integral.

(c) We have $a\phi + b\psi = \sum_{i,j} (a \cdot a_i + b \cdot b_j) \chi_{A_i \cap B_j}$. Using (b) and finite additivity of μ it follows that

$$\int (a\phi + b\psi) = \sum_{i,j} (a \cdot a_i + b \cdot b_j) \mu(A_i \cap B_j) = a \int \phi + b \int \psi.$$

(d) It is clear that $\int_{\emptyset} \phi = 0$. If (C_n) is a disjoint sequence of measurable sets, let $C = \bigcup C_n$; then

$$\int_C \phi = \sum_i a_i \mu(C \cap A_i) = \sum_{i,n} a_i \mu(C_n \cap A_i) = \sum_n \int_{C_n} \phi,$$

which verifies countable additivity. ■

Note that it follows from part (c) of Lemma 6.3 that the integral of a simple function $\phi = \sum a_i \chi_{A_i}$ equals $\sum a_i \mu(A_i)$ regardless of whether this is the standard form for ϕ .

The following are two fundamental results about convergence. (The third fundamental result will appear in the next section.) Observe first that $f \leq g$ implies $\int f \leq \int g$; this is obvious from the definition of the integral.

Theorem 6.4. (Monotone convergence theorem) *Let (f_n) be a pointwise increasing sequence of measurable functions from a measure space X into $[0, \infty]$. Then*

$$\int \lim f_n = \lim \int f_n.$$

Proof. Let $f = \lim f_n$, taking the limit pointwise, and note that f is measurable by Theorem 2.6 (d). Then $f_n \leq f$ for all n , so $\int f_n \leq \int f$ for all n and hence

$$\lim \int f_n = \sup \int f_n \leq \int f.$$

Conversely, assuming $\int f < \infty$ fix $\epsilon > 0$ and let ϕ be a simple function such that $0 \leq \phi \leq f$ and $\int \phi \geq (1 - \epsilon) \int f$. For each n define

$$A_n = \{x : f_n(x) \geq (1 - \epsilon)\phi(x)\};$$

Then (A_n) is an increasing sequence of measurable sets and $\bigcup A_n = X$. Also

$$\int f_n \geq \int_{A_n} f_n \geq (1 - \epsilon) \int_{A_n} \phi,$$

and taking the limit as $n \rightarrow \infty$ and using Lemma 6.3 (d) we obtain

$$\lim \int f_n \geq (1 - \epsilon) \int \phi \geq (1 - \epsilon)^2 \int f.$$

Taking $\epsilon \rightarrow 0$ yields $\lim \int f_n \geq \int f$. If $\int f = \infty$ essentially the same proof works, but now we choose ϕ such that $\int \phi \geq N$ and let $N \rightarrow \infty$. ■

Corollary 6.5. (Fatou's lemma) *Let (f_n) be any sequence of measurable functions from a measure space X into $[0, \infty]$. Then*

$$\int (\liminf f_n) \leq \liminf \int f_n.$$

Proof. For each n let $g_n = \inf_{k \geq n} f_k$, so that the sequence (g_n) is increasing and $\liminf f_n = \lim g_n$. Thus, the monotone convergence theorem yields

$$\int (\liminf f_n) = \int \lim g_n = \lim \int g_n.$$

For each n we have $g_n \leq f_n$; this implies

$$\lim \int g_n = \liminf \int g_n \leq \liminf \int f_n,$$

which together with the above displayed equation yields the desired inequality. ■

The inequality in Fatou's lemma can be strict. For example, if $f_n = \chi_{[n, n+1]}$ on \mathbf{R} then $f_n \rightarrow 0$ pointwise but $\liminf \int f_n = 1$.

We can now state the basic properties of integrals of positive functions.

Theorem 6.6. *Let X be a measure space and let $f, g, f_n : X \rightarrow [0, \infty]$ be measurable functions.*

- (a) *If $f \leq g$ then $\int f \leq \int g$.*
- (b) *$\int (af + bg) = a \int f + b \int g$ for all $a, b \geq 0$.*
- (c) *$\int \sum f_n = \sum \int f_n$.*
- (d) *The map $A \mapsto \int_A f$ is a measure on X .*

Proof. (a) Trivial. (We already used this fact.)

(b) Let ϕ_n and ψ_n be increasing sequences of simple functions which converge pointwise to f and g , respectively. (For example, we could define

$$\phi_n(x) = \begin{cases} k/2^n & \text{if } k/2^n \leq f(x) < (k+1)/2^n \text{ for some } k < 4^n \\ 2^n & \text{if } f(x) \geq 2^n \end{cases}$$

and ψ_n similarly.) Then $a\phi_n + b\psi_n$ increases pointwise to $af + bg$, so MCT plus Lemma 6.3 (c) imply

$$\int (af + bg) = \lim \int (a\phi_n + b\psi_n) = \lim \left(a \int \phi_n + b \int \psi_n \right) = a \int f + b \int g.$$

(c) Using part (b) and MCT, we have

$$\sum \int f_n = \lim_n \sum_{k=1}^n \int f_k = \lim_n \int \sum_{k=1}^n f_k = \int \sum f_n.$$

(d) $\int_\emptyset f = 0$ is clear. For countable additivity let (A_n) be a disjoint sequence of measurable sets and for each n let $f_n = f \cdot \chi_{A_n}$. Then by part (c)

$$\sum \int_{A_n} f = \sum \int f_n = \int \sum f_n = \int_A f$$

where $A = \bigcup A_n$. ■

Exercises.

6.A. Let X be a measure space and let $f, f_n : X \rightarrow [0, \infty]$ be measurable functions. Suppose $f_n \rightarrow f$ pointwise and $\lim \int f_n = \int f < \infty$. Show that $\int_A f_n \rightarrow \int_A f$ for all measurable $A \subseteq X$. Give an example where this is not true if $\lim \int f_n = \int f = \infty$.

6.B. Let X be a measure space and let $f : X \rightarrow [0, \infty]$ be measurable. Suppose $\int f < \infty$. Given $\epsilon > 0$, show that there is a measurable set $A \subseteq X$ such that $\mu(A) < \infty$ and $\int_A f \geq \int f - \epsilon$.

7. Integration of general functions

Write f^+ and f^- for the positive and negative parts of a function $f : X \rightarrow \overline{\mathbf{R}}$. Thus $f^+ = f \vee 0$ and $f^- = -(f \wedge 0)$ (where \vee denotes supremum and \wedge denotes infimum).

Definition 7.1. Let X be a measure space. A measurable function $f : X \rightarrow \overline{\mathbf{R}}$ is *integrable* if both $\int f^+$ and $\int f^-$ are finite. In this case we define $\int f = \int f^+ - \int f^-$.

Example 7.2. If μ is counting measure on \mathbf{N} then $f : X \rightarrow \mathbf{R}$ is integrable in the above sense if and only if it is absolutely summable, and we have $\int f d\mu = \sum_{n \in \mathbf{N}} f(n)$.

Theorem 7.3. Let X be a measure space and let $f, g : X \rightarrow \mathbf{R}$ be integrable.

(a) $\int (af + bg) = a \int f + b \int g$ for all $a, b \in \mathbf{R}$.

(b) $|\int f| \leq \int |f|$.

Proof. (a) Since f and g are integrable, so is $h = f + g$. Then

$$h^+ - h^- = h = f + g = f^+ - f^- + g^+ - g^-,$$

so $h^+ + f^- + g^- = h^- + f^+ + g^+$. Additivity of integrals of positive functions then implies

$$\int h^+ + \int f^- + \int g^- = \int h^- + \int f^+ + \int g^+,$$

and rearranging yields $\int h = \int f + \int g$.

It is easy to verify $\int af = a \int f$ in the two cases $a \geq 0$, $a \leq 0$ using Theorem 6.6 (b).

(b) This follows from the inequality $|a - b| \leq a + b$ with $a = \int f^+$ and $b = \int f^-$. ■

Next we prove the fundamental theorem on convergence for real-valued functions.

Theorem 7.4. (Dominated Convergence Theorem) Let (f_n) be a sequence of integrable functions which converges pointwise to a function f . Suppose there is an integrable function $g \geq 0$ such that $|f_n| \leq g$ for all n . Then f is integrable and $\int f_n \rightarrow \int f$.

Proof. Since $|f(x)| = \lim |f_n(x)| \leq g(x)$ for all x , f is integrable. Now we verify $\int f_n \rightarrow \int f$. We have $g \pm f_n \geq 0$, so by Fatou's lemma

$$\begin{aligned} \int g + \int f &\leq \liminf \int (g + f_n) = \int g + \liminf \int f_n \\ \int g - \int f &\leq \liminf \int (g - f_n) = \int g - \limsup \int f_n \end{aligned}$$

and hence $\liminf \int f_n \geq \int f \geq \limsup \int f_n$. Thus $\int f_n \rightarrow \int f$. ■

Finally, we record the connection between Lebesgue and Riemann integration. We omit the proof.

Theorem 7.5. Let $a, b \in \mathbf{R}$, $a < b$, and let $f : [a, b] \rightarrow \mathbf{R}$ be bounded. Then f is Riemann integrable \Leftrightarrow the set of discontinuities of f has Lebesgue measure zero, and in this case it is also Lebesgue integrable and the two integrals agree.

Exercises

7.A. Let f and g be integrable. Show that $\int_A f = \int_A g$ for all measurable $A \Leftrightarrow \int |f - g| = 0 \Leftrightarrow f = g$ almost everywhere.

7.B. Let $f : \mathbf{R} \rightarrow \mathbf{R}$ be a Lebesgue to Borel measurable function which is integrable. Show that there is a Borel to Borel measurable function g which equals f almost everywhere. More generally, if (X, μ) is any measure space and $\overline{\mu}$ is the completion of μ , show that every $\overline{\mu}$ -measurable function $f : X \rightarrow \mathbf{R}$ almost everywhere equals a μ -measurable function $g : X \rightarrow \mathbf{R}$. (Hint: approximate by simple functions.)

8. Product measures

Recall the product σ -algebra from Definition 1.7. We call subsets of $X \times Y$ of the form $A \times B$ (with $A \subseteq X$ and $B \subseteq Y$) *rectangles*.

Lemma 8.1. *Let (X, μ) and (Y, ν) be measure spaces and let Ω_0 be the ring of subsets of $X \times Y$ consisting of finite disjoint unions of rectangles $A \times B$ with $A \subseteq X$ and $B \subseteq Y$ measurable and $\mu(A), \nu(B) < \infty$. Then*

$$(\mu \times \nu) \left(\bigcup_1^n (A_i \times B_i) \right) = \sum_1^n \mu(A_i) \nu(B_i)$$

(with the sets $A_i \times B_i$ disjoint) defines a premeasure on $X \times Y$ with domain Ω_0 .

Proof. It follows easily from Exercise 5.A that Ω_0 is an algebra of sets, and well-definedness of $\mu \times \nu$ is proven in the usual way. To verify that $\mu \times \nu$ is a premeasure, suppose that

$$A \times B = \bigcup_{n=1}^{\infty} (A_n \times B_n)$$

(disjoint union); we must show that $\mu(A)\nu(B) = \sum_1^{\infty} \mu(A_n)\nu(B_n)$.

For each n let $f_n : X \rightarrow [0, \infty)$ be the function which takes the value $\nu(B_n)$ on the set A_n and the value 0 off the set A_n . That is, $f_n = \nu(B_n) \cdot \chi_{A_n}$. Then $\int f_n = \mu(A_n)\nu(B_n)$. For any point $x \in A$ we have

$$\sum_{n=1}^{\infty} f_n(x) = \sum \{\nu(B_n) : x \in A_n\} = \nu(B)$$

since $\{B_n : x \in A_n\}$ is a disjoint sequence whose union equals B . Also, for $x \notin A$ we have $f_n(x) = 0$ for all n . Thus $\sum f_n = \nu(B) \cdot \chi_A$ (pointwise), so by Theorem 6.6 (c)

$$\sum_{n=1}^{\infty} \mu(A_n)\nu(B_n) = \sum_{n=1}^{\infty} \int f_n = \int \nu(B) \cdot \chi_A = \mu(A)\nu(B),$$

as desired. ■

The next result now follows immediately from Theorem 4.3.

Theorem 8.2. *Let (X, μ) and (Y, ν) be σ -finite measure spaces. Then there is a complete measure $\mu \times \nu$ on $X \times Y$ whose domain contains the product σ -algebra and which satisfies $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$ for all measurable sets $A \subseteq X$ and $B \subseteq Y$.*

(We only use σ -finiteness to ensure that the product σ -algebra is generated by the ring Ω_0 .)

The measure constructed in Theorem 8.2 is called the *product measure*. It is routine to extend this construction to any finite number of factors. Moreover, it is easy to see that the order in which the factors are taken has no effect. If the number of factors is infinite, however, there is a potentially serious problem. Say that (X_n, μ_n) is a sequence of measure spaces. For each n let A_n be a measurable subset of X_n . We want a product measure to satisfy $\mu(\prod A_n) = \prod \mu_n(A_n)$. But in order for this product to be finite but nonzero, we must have $\mu_n(A_n) = 1$ for all but finitely many n — or at least, their measures must converge to 1 rapidly enough for the product to make sense. For this reason, infinite product measures are typically only defined when each μ_n is a probability measure (Definition 3.1).

We now prove a series of results relating product measures to iterated integrals. We need the following notation. Let X and Y be sets, let $C \subseteq X \times Y$, and let $f : X \times Y \rightarrow \mathbf{R}$ be a function. Then for $x \in X$ and $y \in Y$ we define sets $C_x \subseteq Y$ and $C^y \subseteq X$ by

$$C_x = \{y' \in Y : (x, y') \in C\} \quad \text{and} \quad C^y = \{x' \in X : (x', y) \in C\}$$

and functions $f_x : Y \rightarrow \mathbf{R}$ and $f^y : X \rightarrow \mathbf{R}$ by

$$f_x(y) = f^y(x) = f(x, y).$$

These are called *sections* of C and f .

Theorem 8.3. (Cavalieri's Principle) *Let (X, μ) and (Y, ν) be complete σ -finite measure spaces and let C be a measurable subset of the product measure space $(X \times Y, \mu \times \nu)$. Then the functions $x \mapsto \nu(C_x)$ and $y \mapsto \mu(C^y)$ are measurable and*

$$(\mu \times \nu)(C) = \int \nu(C_x) d\mu(x) = \int \mu(C^y) d\nu(y).$$

Proof. First suppose (X, μ) and (Y, ν) are finite measure spaces. Let $c = (\mu \times \nu)(C)$ and let h be the function $h(x) = \nu(C_x)$. This is well-defined for almost every x by Exercise 8.A. For each $n \in \mathbf{N}$ we can find a disjoint sequence of rectangles $(A_k \times B_k)$ which contains C and whose total measure is at most $c + 1/n$. Letting $f_n = \sum_k \nu(B_k) \cdot \chi_{A_k}$, we then have $h \leq f_n$ and $\int f_n \leq c + 1/n$. So $h \leq f$ and $\int f \leq c$, where $f = \inf f_n$ is measurable.

Applying the same construction to the complement of C , we get a measurable function g satisfying $h \geq g$ and $\int g \geq c$. Thus $c \leq \int g \leq \int f \leq c$, so we must have equality. Then h is sandwiched between two measurable functions that have the same integral, so it follows that h is measurable and $\int h = c$. The analogous statements hold in the other variable by symmetry.

The result extends to the σ -finite case by partitioning into finite measure subsets. ■

Theorem 8.4. (Tonelli) *Let (X, μ) and (Y, ν) be complete σ -finite measure spaces and let $f : X \times Y \rightarrow [0, \infty]$ be measurable. Then the functions $x \mapsto \int f_x d\nu$ and $y \mapsto \int f^y d\mu$ are measurable and*

$$\int f d(\mu \times \nu) = \int \left[\int f(x, y) d\nu(y) \right] d\mu(x) = \int \left[\int f(x, y) d\mu(x) \right] d\nu(y).$$

Proof. The result holds when f is a characteristic function by Theorem 8.3, and it follows for simple functions by linearity. For a general measurable function f , find an increasing sequence of simple functions ϕ_n which converge pointwise to f as in the proof of Theorem 6.6 (b). Then $\int (\phi_n)_x d\nu \rightarrow \int f_x d\nu$ and $\int (\phi_n)^y d\mu \rightarrow \int f^y d\mu$ for all x and y by MCT, so measurability of the functions $x \mapsto \int (\phi_n)_x d\nu$ and $y \mapsto \int (\phi_n)^y d\mu$ implies measurability of the functions $x \mapsto \int f_x d\nu$ and $y \mapsto \int f^y d\mu$. Applying MCT throughout the equations

$$\int \phi_n d(\mu \times \nu) = \int \left[\int \phi_n(x, y) d\nu(y) \right] d\mu(x) = \int \left[\int \phi_n(x, y) d\mu(x) \right] d\nu(y)$$

(using the fact that the inner integrals increase to $\int f_x d\nu$ and $\int f^y d\mu$, as shown above) yields the desired equality of integrals. ■

Theorem 8.5. (Fubini) *Let (X, μ) and (Y, ν) be complete σ -finite measure spaces and let $f : X \times Y \rightarrow \mathbf{R}$ be integrable. Then f_x and f^y are integrable for almost every x and y , respectively, the functions $x \mapsto \int f_x d\nu$ and $y \mapsto \int f^y d\mu$ are integrable, and*

$$\int f d(\mu \times \nu) = \int \left[\int f(x, y) d\nu(y) \right] d\mu(x) = \int \left[\int f(x, y) d\mu(x) \right] d\nu(y).$$

Proof. Since f is integrable we have $\int |f| < \infty$, and Tonelli's theorem then implies that $\int |f|_x d\nu$ and $\int |f|^y d\mu$ are finite almost everywhere. (Otherwise the iterated integral in Tonelli's theorem would be infinite, since a function which is infinite on a set of positive measure must have an infinite integral.) This shows that f_x

and f^y are integrable for almost every x and y , respectively. The remainder of the theorem follows from Tonelli's theorem by decomposing f into its positive and negative parts. ■

Exercises.

8.A. Let (X, μ) and (Y, ν) be complete finite measure spaces. Then the product measure $\mu \times \nu$ is the completion of its restriction to the product σ -algebra from Definition 1.7. (Hint: first show that every subset of $X \times Y$ is contained in a set in the product σ -algebra that has the same outer measure.)

8.B. Let (f_n) be a sequence of measurable functions on a measure space X . Suppose that $\sum \int |f_n|$ is finite. Use product measures to prove that $\sum \int f_n = \int \sum f_n$.

9. Signed measures

We now generalize the notion of a measure to include real-valued measures.

Definition 9.1. Let (X, Ω) be a measurable space. A *signed measure* on (X, Ω) is a function $\mu : \Omega \rightarrow \mathbf{R}$ such that $\mu(\emptyset) = 0$ and if (A_n) is a sequence of disjoint sets in Ω then $\sum \mu(A_n)$ converges absolutely and $\mu(\bigcup A_n) = \sum \mu(A_n)$.

Signed measures are sometimes allowed to take the value ∞ or $-\infty$ (but not both), but we will not need to do this.

In a sense, signed measures reduce to positive measures: every signed measure is the difference of two positive measures with disjoint supports. The precise statement of this result involves the following notion of singularity.

Definition 9.2. Let μ and ν be two signed measures defined on the same measurable space. Then μ and ν are *mutually singular*, written $\mu \perp \nu$, if there exist complementary measurable sets A and B such that $\mu|_B = 0$ and $\nu|_A = 0$. (That is, $\mu(C) = 0$ for all measurable $C \subseteq B$ and $\nu(C) = 0$ for all measurable $C \subseteq A$.) If μ is a positive measure then ν is *absolutely continuous with respect to μ* , written $\nu \ll \mu$, if $\mu(A) = 0$ implies $\nu(A) = 0$ for all measurable A .

Lemma 9.3. *Every signed measure is bounded. If μ is a signed measure and C is a measurable set such that $\mu(C) > 0$ then there exists a measurable subset $B \subseteq C$ such that $\mu(B) \geq \mu(C)$ and $\mu|_B$ is positive.*

Proof. Let μ be a signed measure on a set X and suppose X has subsets of arbitrarily large (positive) measure. We construct a sequence of disjoint subsets $A_n \subseteq X$ as follows. If there is a subset of X whose measure is at least 1 and whose complement contains subsets of arbitrarily large measure, then let A_1 be such a set. Having constructed A_1, \dots, A_n , if there is a subset of $X - (A_1 \cup \dots \cup A_n)$ whose measure is at least 1 and whose complement in $X - (A_1 \cup \dots \cup A_n)$ contains subsets of arbitrarily large measure, then let A_{n+1} be such a set.

This process must terminate at some finite stage; otherwise we would have a disjoint sequence of sets (A_n) , each of which has measure at least 1, and then we would have to have $\mu(\bigcup A_n) = \infty$, contradicting the definition of a signed measure. So there is some stage where $A = X - (A_1 \cup \dots \cup A_n)$ contains subsets of arbitrarily large measure, but for any $B \subseteq A$ with measure at least 1, $A - B$ does not contain subsets of arbitrarily large measure. It follows that any subset of A of measure at least 1 contains subsets of arbitrarily large measure. Now construct an infinite sequence (B_n) as follows. Let B_1 be a subset of A with $\mu(B_1) \geq 1$, and having constructed B_n , let B_{n+1} be a subset of B_n with $\mu(B_{n+1}) \geq \mu(B_n) + 1$. We can do this because each B_n has measure at least 1, and hence contains subsets of arbitrarily large measure. Then the sequence of differences $(B_n - B_{n+1})$ is disjoint and satisfies $\mu(B_n - B_{n+1}) \leq -1$ for all n , which again contradicts countable additivity and the fact that signed measures are assumed not to take the value $-\infty$. We conclude

that X cannot contain subsets of arbitrarily large positive measure, and by symmetry it cannot contain subsets of arbitrarily large negative measure either. Thus μ is bounded.

Now suppose $\mu(C) > 0$ and define a sequence (C_n) of disjoint subsets of C as follows. Let $C_1 = \emptyset$. Inductively, let

$$s_n = \inf\{\mu(B) : B \subseteq C - (C_1 \cup \dots \cup C_n)\}$$

and choose $C_{n+1} \subseteq C - (C_1 \cup \dots \cup C_n)$ such that $\mu(C_{n+1}) \leq \min(0, s_n + 1/n)$.

Let $B = C - \bigcup C_n$. Then

$$\mu(B) = \mu(C) - \sum \mu(C_n) \geq \mu(C)$$

since $\mu(C_n) \leq 0$ for all n . For any $B' \subseteq B$, if $\mu(B') < 0$ then $\mu(B') < -1/n$ for some n , and then $\mu(C_{n+1} \cup B') < \mu(C_{n+1}) - 1/n \leq s_n$, contradicting the definition of s_n . Thus $\mu|_B$ is positive. \blacksquare

Theorem 9.4. (Hahn-Jordan decomposition) *Let μ be a signed measure on a set X .*

(a) *There exist disjoint measurable sets $P, N \subseteq X$ such that $\mu|_P$ and $-\mu|_N$ are both finite positive measures.*

(b) *There is a unique decomposition $\mu = \mu^+ - \mu^-$ such that μ^+ and μ^- are mutually singular finite positive measures.*

Proof. (a) Let $a = \sup\{\mu(A) : A \text{ is measurable}\} < \infty$. We claim that there is a measurable set P such that $\mu|_P$ is positive and $\mu(P) = a$. Given this, let $N = X - P$; then $-\mu|_N$ must be positive, since if $\mu(B) > 0$ for some $B \subseteq N$ then $\mu(P \cup B) = a + \mu(B) > a$, a contradiction. Finiteness of $\mu|_P$ and $\mu|_N$ follows from boundedness of μ .

To prove the claim, let (A_n) be a sequence of measurable sets such that $\mu(A_n) \rightarrow a$. For each n use the lemma to find a subset B_n of A_n such that $\mu(B_n) \geq \mu(A_n)$ and $\mu|_{B_n}$ is positive. (If $\mu(A_n) \leq 0$ then we can take $B_n = \emptyset$.) Then $P = \bigcup B_n$ has the desired properties.

(b) Let P and N be as in (a) and define $\mu^+(A) = \mu(A \cap P)$ and $\mu^-(A) = -\mu(A \cap N)$ for all measurable A , i.e., $\mu^+ = \mu|_P$ and $\mu^- = -\mu|_N$. It is clear that μ^+ and μ^- are mutually singular positive measures such that $\mu = \mu^+ - \mu^-$. If ν^+, ν^- are any other such pair, let P' and N' be complementary measurable sets witnessing the mutual singularity of ν^+ and ν^- . Then $P - P' = P \cap N'$ implies that $\mu(B) = 0$ for all $B \subseteq P - P'$, and similarly for $P' - P$. Thus

$$\mu^+(A) = \mu(A \cap P) = \mu(A \cap P \cap P') = \mu(A \cap P') = \nu^+(A)$$

for all $A \subseteq X$, and similarly $\mu^- = \nu^-$. \blacksquare

For any signed measure μ we define its *total variation* to be the positive measure $|\mu| = \mu^+ + \mu^-$.

Now we turn to the comparison of two measures on a set. The significance of absolute continuity is shown in the following result.

Theorem 9.5. (Lebesgue decomposition) *Let μ be a positive measure and ν a signed measure defined on the same σ -algebra. Then there are signed measures ν_1 and ν_2 such that $\nu_1 \ll \mu$, $\nu_2 \perp \mu$, and $\nu = \nu_1 + \nu_2$. This decomposition of ν is unique.*

Proof. We have $a = \sup\{|\nu|(A) : \mu(A) = 0\} \leq |\nu|(X) < \infty$. Find a sequence of measurable sets (A_n) such that $\mu(A_n) = 0$ for all n and $|\nu|(A_n) \rightarrow a$ and let $A = \bigcup A_n$. Then $\mu(A) = 0$ and $|\nu|(A) = a$.

Let ν_1 be the restriction of ν to the complement of A and let ν_2 be the restriction of ν to A . It is clear that $\nu = \nu_1 + \nu_2$ and that $\nu_2 \perp \mu$. If ν_1 failed to be absolutely continuous with respect to μ then there would be a set B such that $\mu(B) = 0$ but $\nu_1(B) \neq 0$. Without loss of generality $B \cap A = \emptyset$, and $|\nu|(A \cup B) = |\nu|(A) + |\nu|(B) > a$ then contradicts the definition of a . Thus $\nu_1 \ll \mu$.

Let $\nu = \nu'_1 + \nu'_2$ be any other decomposition of ν satisfying $\nu'_1 \ll \mu$ and $\nu'_2 \perp \mu$. Let B be a measurable set such that $\mu(B) = 0$ and ν'_2 is supported on B . Then ν_1 and ν'_1 are both supported on the complement of

$A \cup B$ since μ is zero on $A \cup B$, which implies that $\nu_2 =$ the restriction of ν to $A \cup B = \nu'_2$. Thus $\nu_2 = \nu'_2$ and therefore also $\nu_1 = \nu'_1$, proving that the decomposition is unique. \blacksquare

In light of the preceding result, it is worthwhile to study absolute continuity further. We give two characterizations; the first is completely general, while the second requires that μ be σ -finite.

Proposition 9.6. *Let μ be a positive measure and ν a signed measure defined on the same σ -algebra. Then $\nu \ll \mu \Leftrightarrow$ for every $\epsilon > 0$ there exists $\delta > 0$ such that $\mu(A) < \delta$ implies $|\nu(A)| < \epsilon$.*

Proof. The reverse direction is easy. (If $\mu(A) < \delta$ for all δ then $|\nu(A)| < \epsilon$ for all ϵ .) For the forward direction, suppose the ϵ - δ condition fails. Then there exists a value of ϵ such that the condition fails for all δ . Thus, for each n there is a measurable set A_n such that $\mu(A_n) < 2^{-n}$ but $|\nu(A_n)| \geq \epsilon$. Let

$$A = \limsup A_n = \bigcap_{k=1}^{\infty} \bigcup_{i=k}^{\infty} A_i;$$

then $\mu(A) \leq \mu(\bigcup_k A_k) < 2^{-(k-1)}$ for all k , and hence $\mu(A) = 0$. However, $|\nu|(\bigcup_k A_k) \geq \epsilon$ for all k , which implies $|\nu|(A) \geq \epsilon$ (using Theorem 3.4 (d) and the fact that $|\nu|$ is finite). From the definition of $|\nu|$ it follows that there is a measurable set $B \subseteq A$ such that $\nu(B) \neq 0$, and we have contradicted the assumption of absolute continuity. This completes the proof. \blacksquare

It follows easily from Theorem 6.6 (d) that if μ is a positive measure then for any μ -integrable real-valued function f the map $A \mapsto \int_A f d\mu$ defines a signed measure. (Write f as a difference of two positive integrable functions.) Any measure of this form is clearly absolutely continuous with respect to μ . The following theorem shows that if μ is σ -finite then the converse is true, i.e., every $\nu \ll \mu$ arises in this way.

Theorem 9.7. (Radon-Nikodym) *Let (X, μ) be a σ -finite measure space and let ν be a signed measure which is absolutely continuous with respect to μ . Then there exists a μ -integrable real-valued function f such that*

$$\nu(A) = \int_A f d\mu$$

for all measurable sets A . Moreover, $|\nu|(A) = \int_A |f| d\mu$ for all A .

Proof. By treating the positive and negative parts of ν separately we can reduce the first part of the theorem to the case where ν is a finite positive measure.

Suppose first that μ is also finite. Let S be the set of measurable functions $f : X \rightarrow [0, \infty]$ with the property that $\int_A f d\mu \leq \nu(A)$ for all measurable A , and let $a = \sup\{\int f d\mu : f \in S\}$. Since $\int f d\mu \leq \nu(X)$ for all $f \in S$, we have $a \leq \nu(X) < \infty$. Choose a sequence (f_n) in S such that $\int f_n d\mu \rightarrow a$.

Observe that if f and g belong to S then so does $h = \max(f, g)$, because if $B = \{x : f(x) > g(x)\}$ then for any measurable set A we have

$$\int_A h d\mu = \int_{A \cap B} f d\mu + \int_{A - B} g d\mu \leq \nu(A \cap B) + \nu(A - B) = \nu(A).$$

Thus, replacing f_n with $\max(f_1, \dots, f_n)$ we can assume that the sequence (f_n) is increasing pointwise. Letting $f = \lim f_n$, MCT implies that $\int_A f_n \rightarrow \int_A f$ for all measurable A , so $f \in S$. Also $\int f = \lim \int f_n = a$.

Since $f \in S$, the measure ν' defined by $\nu'(A) = \nu(A) - \int_A f d\mu$ is positive. We must show that $\nu' = 0$. Suppose not; then for some $\epsilon > 0$ the signed measure $\nu' - \epsilon \cdot \mu$ satisfies $(\nu' - \epsilon \cdot \mu)(X) > 0$. By Lemma 9.3 there exists a positive measure set B such that $(\nu' - \epsilon \cdot \mu)(B) > 0$ and the restriction of $\nu' - \epsilon \cdot \mu$ to B is a positive measure. Letting $f' = f + \epsilon \chi_B$ then contradicts maximality of f , since

$$\int_A f' d\mu = \int_A f d\mu + \epsilon \mu(A \cap B) \leq \int_A f d\mu + \nu'(A) = \nu(A)$$

for all A (so that $f' \in S$) but

$$\int f' d\mu = \int f d\mu + \epsilon\mu(B) > a.$$

(Note that $\mu(B) > 0$ since $\nu'(B) > 0$ and $\nu' \ll \mu$.) This contradiction shows that $\nu' = 0$, and we conclude that $\nu(A) = \int_A f d\mu$ for all A . Also f is integrable since $\int f d\mu = \nu(X) < \infty$.

We now have the first part of the theorem, assuming μ is finite. For the second part let ν be a signed measure and suppose $\nu(A) = \int_A f d\mu$ for some μ -integrable function f . Define $\nu_1(A) = \int_A f^+ d\mu$ and $\nu_2(A) = \int_A f^- d\mu$; these are positive measures by Theorem 6.6 (d). The uniqueness statement of Theorem 9.5 now implies that $\nu_1 = \nu^+$ and $\nu_2 = \nu^-$, and the characterization of $|\nu|$ then follows from the fact that $|f| = f^+ + f^-$.

Now suppose μ is only σ -finite and let X be the disjoint union of a sequence of measurable sets A_n such that $\mu(A_n) < \infty$ for all n . Restricting to each A_n , we can find a sequence of functions $f_n : A_n \rightarrow \mathbf{R}$ such that $\nu(A) = \int_A f_n d\mu$ for all n and all measurable $A \subseteq A_n$. We also know that $|\nu|(A_n) = \int_{A_n} |f_n| d\mu$, and summing over n yields $\int |f| d\mu < \infty$ where f is the function on X satisfying $f|_{A_n} = f_n$. Thus f is μ -integrable, and for any measurable set A we have

$$\nu(A) = \sum_n \nu(A \cap A_n) = \sum_n \int_{A \cap A_n} f d\mu = \int_A f d\mu.$$

We have $|\nu|(A) = \int_A |f| d\mu$ similarly. ■

One sometimes writes $d\nu = f d\mu$ or even $f = d\nu/d\mu$ to indicate the relationship $\nu(A) = \int_A f d\mu$ in the Radon-Nikodym theorem. The function f is called the *Radon-Nikodym derivative* of ν with respect to μ . It really behaves like a derivative in the following respect. If f is $|\nu|$ -integrable, define $\int f d\nu = \int f d\nu^+ - \int f d\nu^-$.

Proposition 9.8. *Let μ be a σ -finite measure, let ν be a signed measure, and suppose $\nu \ll \mu$. For any ν -integrable function g the function $g \cdot (d\nu/d\mu)$ is μ -integrable and satisfies*

$$\int g d\nu = \int g \cdot (d\nu/d\mu) d\mu.$$

Proof. We can reduce to the case that ν is positive in the usual way. The conclusion is true when $g = \chi_A$ is a characteristic function by the definition of $d\nu/d\mu$. It is then true for simple functions by linearity, for positive integrable functions by MCT, and for all integrable functions by linearity again. ■

Exercises.

9.A. Find a positive Borel measure ν on \mathbf{R} such that $\nu \neq 0$, $\nu \perp m$ where m denotes Lebesgue measure, and $\nu(\{x\}) = 0$ for all $x \in \mathbf{R}$. (Hint: use Exercise 5.B.)

9.B. Let μ be a positive measure and let f be an integrable real-valued function. Prove that for every $\epsilon > 0$ there exists $\delta > 0$ such that $\mu(A) < \delta \Rightarrow |\int_A f d\mu| < \epsilon$.