REVIEW SHEET: BASICS OF MEASURES AND INTEGRATION

VAUGHN CLIMENHAGA – MATH 6320, FALL 2017

This document contains the essential definitions, examples, and properties that you need to know. It is written in a somewhat informal style with the goal of giving an overview that can be read easily; you should refer to the textbook/notes for careful statements and full details.

*Most of the exercises included here are (in my opinion) straightforward applications of the relevant definitions and have solutions requiring only a few lines. In cases where the solution requires a little more thought and a more creative step is needed, I have marked the exercise with a +. In cases where the solution is relatively straightforward but would be tedious to write down in full detail, I have marked the exercise with a *.*

1. Measures

1.1. Algebras and σ-algebras.

Definition 1.1. Let X be a set and $2^X := \{A \subset X\}$ its power set. An *algebra* on X is a collection $\mathcal{A} \subset 2^X$ that contains X and is closed under complements and finite unions. A σ*-algebra* is an algebra that is also closed under countable unions.

Exercise 1. If the above is taken as the definition of an algebra, then every algebra is also closed under finite intersections, so this definition agrees with the one in Cohn's book. Similarly for σ-algebras.

Exercise 2. (Folland 1.4): Show that an algebra $\mathcal{A}$ is a σ-algebra iff $\mathcal{A}$ is closed under countable increasing unions (i.e., if $\{E_j\}_1^\infty \subset \mathcal{A}$ and $E_1 \subset E_2 \subset \cdots$, then $\bigcup_1^\infty E_j \in \mathcal{A}$).

Proposition 1.2. The intersection of σ-algebras is a σ-algebra.

Definition 1.3. The σ*-algebra generated by* $\mathcal{F} \subset 2^X$ is denoted by $\sigma(\mathcal{F})$ and is the intersection of all σ-algebras containing $\mathcal{F}$. The most important example comes when $X = \mathbb{R}^d$ and $\mathcal{F}$ is the collection of open subsets; then $\sigma(\mathcal{F})$ is the *Borel* σ*-algebra* on $\mathbb{R}^d$ and is denoted $\mathcal{B}(\mathbb{R}^d)$.

Exercise 3. Describe 4 other examples of σ-algebras on $\mathbb{R}$.

Example 1.4. Let $X = \{0,1\}^{\mathbb{N}}$ be the set of all infinite binary sequences. Given a finite binary word $w = w_1 \cdots w_n \in \{0,1\}^n$, let $C(w) = \{x \in X : x_i = w_i \ \forall 1 \leq i \leq n\}$ be the corresponding *cylinder set*. Let $\mathcal{F}_n := \{C(w) : w \in \{0,1\}^k, 0 \leq k \leq n\}$ be the collection of cylinder sets of order $\leq n$, and let $\mathcal{A}_n := \sigma(\mathcal{F}_n)$. If we view $x \in X$ as one specific outcome of an of an infinite sequence of experiments, so that x_n records the outcome of the nth experiment, then $\mathcal{A}_n$ is the σ-algebra consisting of all events which are determined by the first n experiments.

Exercise 4. In the above example, let $\mathcal{F} = \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ be the collection of all cylinder sets, and let $\mathcal{A} = \sigma(\mathcal{F})$. Is it true that $\mathcal{A} = \bigcup_{n\in\mathbb{N}} \mathcal{A}_n$?

1.2. **Measures.**

Definition 1.5. Let X be a set and $\mathcal{A}$ a σ-algebra on X. Then $(X, \mathcal{A})$ is called a *measurable space*. A *measure* on $(X, \mathcal{A})$ is a countably additive function $\mu\colon \mathcal{A} \to [0, +\infty]$ such that $\mu(\emptyset) = 0$. The triple $(X, \mathcal{A}, \mu)$ is a *measure space*.

Example 1.6. Counting measure. The point mass at x for any $x \in X$.

Exercise 5. (Bass 3.4): Prove that $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$.

Definition 1.7. A measure μ is said to be *finite* if $\mu(X) < \infty$ and σ*-finite* if there is a sequence $A_n \in \mathcal{A}$ such that $X = \bigcup_{n\in\mathbb{N}} A_n$ and $\mu(A_n) < \infty$ for all n.

All the measures we consider in this course will be σ-finite.

Proposition 1.8. Measures are monotonic: $A \subset B$ implies $\mu(A) \le \mu(B)$. If $A_1 \subset A_2 \subset \cdots$ then $\mu(\bigcup_n A_n) = \lim_n \mu(A_n)$. If $A_1 \supset A_2 \supset \cdots$ **and** $\mu(A_1) < \infty$ then $\mu(\bigcap_n A_n) = \lim_n \mu(A_n)$.

Exercise 6. (Bass 3.1): Let $(X, \mathcal{A})$ be a measurable space and $\mu\colon \mathcal{A} \to [0, \infty]$ a *finitely* additive function with $\mu(\emptyset) = 0$. Suppose that whenever $A_n \in \mathcal{A}$ is a sequence with $A_n \subset A_{n+1}$ for all n, we have $\mu(\bigcup_n A_n) = \lim_n \mu(A_n)$. Prove that μ is a measure.

1.3. **Outer measures.** Our most important tool for defining measures relies on the following notions.

Definition 1.9. An *outer measure* on X is a function $\mu^*\colon 2^X \to [0, \infty]$ that is monotonic, countably subadditive, and has $\mu^*(\emptyset) = 0$. A set $E \subset X$ is μ^**-measurable* if

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c) \tag{1}$$

for every $A \subset X$.

To remember the definition of μ^*-measurability, it may be helpful to think of $x \in X$ as representing a random point, with $x \in E$ representing the statement that some event occurs, and $x \in E^c$ representing the statement that the event does not occur. Then the terms on the right-hand side of (1) represent the conditional probabilities that an event A occurs, conditioned on whether or not E occurred. The μ^*-measurable events are those that behave well when we use them as conditions.

Theorem 1.10. If μ^* is an outer measure on X, then the collection $\mathcal{A} \subset 2^X$ of μ^*-measurable sets is a σ-algebra, and $\mu = \mu^*|_{\mathcal{A}}$ is a measure.

Example 1.11. One can check that the following formula defines an outer measure on $\mathbb{R}$, called *Lebesgue outer measure*:

$$\lambda^*(A) = \inf\left\{ \sum_i (b_i - a_i) : \{a_i, b_i\}_{i=1}^\infty \subset \mathbb{R}, A \subset \bigcup_{i=1}^\infty (a_i, b_i) \right\}. \tag{2}$$

A λ^*-measurable set is called *Lebesgue measurable*; the collection of such sets is the *Lebesgue σ-algebra* on $\mathbb{R}$. Every open interval is Lebesgue measurable and has $\lambda^*((a, b)) = b - a$. Thus the Lebesgue σ-algebra contains the Borel sets, and the Lebesgue measure λ generalizes the notion of length.

This example is a specific case of the following general framework for constructing outer measures, and hence measures, represented by the schematic here:

$$\boxed{\text{premeasure } \ell \text{ on an algebra } \mathcal{A}_0} \longrightarrow \boxed{\text{outer measure } \mu^* \text{ on } 2^X} \longrightarrow \boxed{\text{measure } \mu \text{ on } \sigma\text{-algebra } \mathcal{A} \text{ of } \mu^*\text{-measurable sets}}$$

Definition 1.12. A *premeasure* on an algebra $\mathcal{A}_0$ is a function $\ell \colon \mathcal{A}_0 \to [0, \infty]$ that has $\mu(\emptyset) = 0$ and for which $\ell(\bigsqcup_{i=1}^\infty A_i) = \sum_i \ell(A_i)$ whenever $A_1, A_2, \dots \in \mathcal{A}_0$ are disjoint sets for which $\bigsqcup_i A_i$ is also in $\mathcal{A}_0$.

Note that premeasures are finitely additive since given disjoint $A_1, \dots, A_n \in \mathcal{A}_0$, we can put $A_i = \emptyset$ for all $i > n$ and then apply the definition. The following result is Theorem 4.16 in Bass's book, and formalizes the procedure outlined in the schematic.

Theorem 1.13 (Carathéodory extension theorem)**.** If $\mathcal{A}_0$ is an algebra on X and ℓ is a premeasure on $\mathcal{A}_0$, then

$$\mu^*(A) = \inf\left\{ \sum_{i=1}^\infty \mu(A_i) : A_1, A_2, \dots \in \mathcal{A}_0, A \subset \bigcup_{i=1}^\infty A_i \right\}$$

defines an outer measure on X with the following properties:

(1) $\mu^*(A) = \ell(A)$ for all $A \in \mathcal{A}_0$;
(2) every $A \in \mathcal{A}_0$ is μ^*-measurable.

The first of these properties shows that μ^* extends ℓ, rather than redefining its values; the second property guarantees that when we obtain a measure μ from μ^* by restricting to the σ-algebra $\mathcal{A}$ of all μ^*-measurable sets, we have $\mathcal{A} \supset \sigma(\mathcal{A}_0)$.

Exercise 7. * Describe the algebra on $\mathbb{R}$ generated by the collection of all open intervals (including intervals of the form (a, ∞) and $(-\infty, b)$), and describe the premeasure on this algebra that leads to Lebesgue outer measure.

Exercise 8. + (Folland 1.30): Given a Lebesgue measurable set $E \subset \mathbb{R}$ with $\lambda(E) > 0$, show that for every $\alpha < 1$ there is an open interval $I \subset \mathbb{R}$ such that $\lambda(E \cap I) > \alpha\lambda(I)$. *Hint: argue by contradiction, cover E by an arbitrary collection of open intervals, and use the definition of Lebesgue measure in* (2).

Exercise 9. ** Describe the algebra on $\mathbb{R}^d$ generated by the collection of all sets of the form $(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_d, b_d)$, where we allow $a_i = -\infty$ and/or $b_i = \infty$. Then describe the premeasure on this algebra that gives each such set the mass $\prod_{i=1}^d (b_i - a_i)$.

The premeasure in the last exercise leads to Lebesgue measure on $\mathbb{R}^d$.

Example 1.14. Let $X = \{0,1\}^{\mathbb{N}}$ be the set of infinite binary sequences, and let $\mathcal{A}_0 = \{\bigsqcup_{i=1}^k C_i : C_1, \dots, C_k \subset X \text{ are cylinder sets}, k \in \mathbb{N}\}$. Then $\mathcal{A}_0$ is an algebra (prove it!) and given $p \in (0,1)$ we can define a premeasure μ_p as follows: if C is the cylinder set associated to a word $w_1 \cdots w_n$, then let

$$\mu_p(C) = p^{\#\{1 \le i \le n : w_i = 0\}} (1-p)^{\#\{1 \le i \le n : w_i = 1\}}.$$

Extending μ_p to $\mathcal{A}_0$ by finite additivity gives a premeasure, which gives first an outer measure and then a measure on X that we call the $(p, 1-p)$-*Bernoulli measure*.

Finally, one can give a complete description of all finite measures on $\mathbb{R}$: the statement of the following proposition does not use outer measures, but the proof does (see Proposition 1.3.8 in Cohn).

Proposition 1.15. If μ is a finite Borel measure on $\mathbb{R}$, then the function $F \colon \mathbb{R} \to [0, \infty)$ defined by $F(x) = \mu((-\infty, x])$ is bounded, non-decreasing, right-continuous, and has $\lim_{x \to -\infty} F(x) = 0$. Conversely, given any function with these properties, there is a unique finite Borel measure μ on $\mathbb{R}$ such that $\mu((-\infty, x]) = F(x)$.

Exercise 10. Given μ and F as in the proposition and $x \in \mathbb{R}$, show that the function F is continuous at x if and only if $\mu(\{x\}) = 0$.

Exercise 11. (Bass 4.6): Let $(X, \mathcal{A}, \mu)$ be a probability space (a measure space for which $\mu(X) = 1$) and let $A_n \in \mathcal{A}$ for every n. Let B consist of those points x that are in infinitely many of the A_n.

(a) Show that $B \in \mathcal{A}$.
(b) If $\mu(A_n) > \delta > 0$ for all n, show that $\mu(B) \ge \delta$.
(c) Prove the *Borel-Cantelli lemma*:[1] If $\sum_n \mu(A_n) < \infty$, then $\mu(B) = 0$.
(d) With $X = [0,1]$ and $\mu = \lambda$, find A_n with $\sum_n \mu(A_n) = \infty$ but $\mu(B) = 0$.

Hint: it will help to consider the sets $B_N := \bigcup_{n \ge N} A_n$.

1.4. Properties of Lebesgue measure on $\mathbb{R}^d$.

Definition 1.16. Let $\mathcal{A}$ be a σ-algebra on $\mathbb{R}^d$ that contains $\mathcal{B}(\mathbb{R}^d)$, and let μ be a measure on $\mathcal{A}$. We say that μ is *regular* if for every $A \in \mathcal{A}$ we have[2]

$$\begin{aligned} \mu(A) &= \inf\{\mu(U) : A \subset U \text{ and } U \text{ is open}\} \\ &= \sup\{\mu(K) : K \subset A \text{ and } K \text{ is compact}\}. \end{aligned}$$

[1]This is a result from probability theory saying that if a sequence of events, such as our A_n's, have summable probabilities, then with probability 1, only finitely many of them occur.

[2]Different authors may use slightly different definitions of 'regular'; indeed, Cohn's definition is not quite the one given here. In all the spaces we will encounter, the definitions are equivalent.

Proposition 1.17. Lebesgue measure on $\mathbb{R}^d$ is regular; in fact, *every* finite Borel measure on $\mathbb{R}^d$ is regular.

Definition 1.18. A set $A \subset \mathbb{R}^d$ is called a G_δ *set* if $A = \bigcap_{n=1}^\infty U_n$ for some sequence of open sets $U_n \subset \mathbb{R}^d$. It is called a F_σ *set* if $A = \bigcup_{n=1}^\infty E_n$ for some sequence of closed sets $E_n \subset \mathbb{R}^d$.

Exercise 12. Prove that if μ is a regular measure on $(\mathbb{R}^d, \mathcal{A})$, then for every $A \in \mathcal{A}$ there is a G_δ set $Y \subset \mathbb{R}^d$ and a F_σ set $Z \subset \mathbb{R}^d$ such that $Z \subset A \subset Y$ and $\mu(Z) = \mu(A) = \mu(Y)$.

Regularity of λ is used in the proof of the following result.

Proposition 1.19. Lebesgue measure is the only measure on $\mathbb{R}^d$ with the property that $\mu\big(\prod_{i=1}^d [a_i, b_i)\big) = \prod_{i=1}^d (b_i - a_i)$ for all a_i, b_i.

Proposition 1.20. Lebesgue measure is translation-invariant, in the sense that if $A \subset \mathbb{R}^d$ is Lebesgue measurable, then so is $A + x$ for every $x \in \mathbb{R}^d$, and $\lambda(A) = \lambda(A + x)$. Moreover, if μ is any translation-invariant Borel measure on $\mathbb{R}^d$, then there is $c > 0$ such that $\mu(A) = c\lambda(A)$ for every Borel A.

Proposition 1.21. The middle-third Cantor set is a compact uncountable set with Lebesgue measure 0.

Exercise 13. * Let $X = \{0,1\}^{\mathbb{N}}$ as in Examples 1.4 and 1.14. Define a map $T\colon X \to \mathbb{R}$ by $T(x_1 x_2 \cdots) = \sum_{n=1}^\infty 2x_n 3^{-n}$.

(a) Prove that $T(X)$ is the middle-third Cantor set and that T is injective.
(b) Prove that if $A \subset \mathbb{R}$ is Borel, then $T^{-1}(A) \subset X$ lies in the σ-algebra generated by all cylinder sets. *Hint: first prove it when A is an interval.*

Exercise 14. Let X be as in the previous exercise and let μ be the $(\frac{1}{2}, \frac{1}{2})$-Bernoulli measure on X from Example 1.14. Define a Borel measure ν on $\mathbb{R}$ by $\nu(A) = \mu(T^{-1}A)$, so that ν gives weight 2^{-n} to each of the 2^n closed intervals that appear in the nth step of the construction of the Cantor set. Let $F\colon \mathbb{R} \to [0,1]$ be the cumulative distribution function of ν given by $F(x) = \nu((-\infty, x])$, as in Proposition 1.15. This function is sometimes called the *Cantor function* or the *devil's staircase*. Prove that:

(a) F is continuous and $F(C) = [0,1]$, where C is the middle-third Cantor set;
(b) $F'(x)$ exists if and only if $x \notin C$;
(c) F' vanishes everywhere that it exists.

Exercise 15. * (Folland 1.32). Given a sequence $a_j \in (0,1)$, show that $\prod_{j=1}^\infty (1 - a_j) > 0$ iff $\sum_{j=1}^\infty a_j < \infty$ by comparing the sum to $\sum_{j=1}^\infty \log(1 - a_j)$. Use this to prove that for every $\beta \in (0,1)$, there is a sequence $a_j \in (0,1)$ such that $\prod_{j=1}^\infty (1 - a_j) = \beta$.

Exercise 16. Construct a set $C \subset [0,1]$ as $C = \bigcap_{n=1}^\infty C_n$, where $C_0 = [0,1]$ and $C_n \subset C_{n-1}$ is a finite union of closed intervals obtained by removing some open interval from the middle of each component interval of C_{n-1}. If we remove the middle-third of each interval at each step, then we get the usual Cantor set. Show that by adjusting the proportion that we remove and using the result of the previous exercise, it is possible to make $\lambda(C)$ take any value in $(0,1)$; we call this a *fat Cantor set*.

Proposition 1.22. There is a subset of $\mathbb{R}$ that is not Lebesgue measurable.

1.5. Completeness.

Definition 1.23. A measure space $(X, \mathcal{A}, \mu)$ is *complete* if for every $A \in \mathcal{A}$ with $\mu(A) = 0$, we have $B \in \mathcal{A}$ for every $B \subset A$. Such a set B is called *μ-null*.

Proposition 1.24. If μ^* is an outer measure on X and $\mathcal{A}$ the σ-algebra of μ^*-measurable sets, then $(X, \mathcal{A}, \mu = \mu^*|_{\mathcal{A}})$ is complete. In particular, Lebesgue measure on the Lebesgue σ-algebra is complete.

Definition 1.25. Given a measure space $(X, \mathcal{A}, \mu)$, the *completion* of $\mathcal{A}$ under μ is the collection $\mathcal{A}_\mu$ of all $A \subset X$ such that $E \subset A \subset F$ for some $E, F \in \mathcal{A}$ with $\mu(F \setminus E) = 0$. Define a measure $\bar{\mu} \colon \mathcal{A}_\mu \to [0, \infty]$ by letting $\bar{\mu}(A) = \mu(F) = \mu(E)$ when A, E, F are as above. Then $\bar{\mu}$ is the *completion* of μ.

Proposition 1.26. For every measure space $(X, \mathcal{A}, \mu)$, the completion $(X, \mathcal{A}_\mu, \bar{\mu})$ is complete.

Proposition 1.27. The Lebesgue σ-algebra is the completion of the Borel σ-algebra under Lebesgue measure.

Exercise 17. Given a measure space $(X, \mathcal{A}, \mu)$, let $\mathcal{N}$ be the collection of μ-null sets. Show that $\mathcal{A}_\mu = \sigma(\mathcal{A} \cup \mathcal{N})$.

2. Integration

2.1. Measurable functions.

Definition 2.1. Given a measurable space $(X, \mathcal{A})$, a function $f \colon X \to [-\infty, \infty]$ is *measurable* if $f^{-1}([-\infty, t]) \in \mathcal{A}$ for all $t \in \mathbb{R}$. When $X = \mathbb{R}^d$ and $\mathcal{A}$ is the Borel (or Lebesgue) σ-algebra, we say that f is *Borel (or Lebesgue) measurable*.

Proposition 2.2. We get the same definition if we replace $[-\infty, t]$ with $[-\infty, t)$, $[t, \infty]$, or $(t, \infty]$. Indeed, we get the same definition if we replace the collection $\{[-\infty, t] : t \in \mathbb{R}\}$ with the collection of all open subsets of $\mathbb{R}$, or the collection of all closed subsets of $\mathbb{R}$, or the collection of all Borel subsets of $\mathbb{R}$.

Exercise 18. (Folland 2.4): Show that if $f \colon X \to \mathbb{R}$ is such that $f^{-1}((r, \infty))$ is measurable for every $r \in \mathbb{Q}$, then f is measurable.

Example 2.3. Continuous functions on $\mathbb{R}^d$ are Borel measurable, as are monotonic functions on $\mathbb{R}$.

Example 2.4. For every $A \in \mathcal{A}$, the characteristic function $\mathbf{1}_A$ is measurable.

Definition 2.5. A *simple function* is a finite linear combination of characteristic functions. Equivalently, it is a function that only takes finitely many values, and it is measurable if $f^{-1}(\alpha) \in \mathcal{A}$ for each of these finitely many values α. We denote collection of all measurable simple functions by $\mathcal{S}$, and non-negative simple functions by $\mathcal{S}_+$.

Proposition 2.6. Measurable functions are well-behaved under standard operations: If $f_1, f_2, \dots$ are measurable then so are $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$, and $\liminf_n f_n$. If f, g are measurable then so are $f+g$, $f-g$, fg, f/g, and αf for every $\alpha \in \mathbb{R}$, provided we restrict each of these to the domain where it is defined. (For example, $f - g$ is undefined on the set of points where $f(x) = g(x) = \infty$, and f/g is undefined on the set of points where $g(x) = 0$.)

Exercise 19. Given a sequence of measurable functions f_n, show that $\{x : \lim_n f_n(x)$ exists$\}$ is measurable.

Exercise 20. * (Folland 2.7): Let $(X, \mathcal{A}, \mu)$ be a measure space. Suppose that we are given a one-parameter family of measurable sets $\{E_\alpha : \alpha \in \mathbb{R}\} \subset \mathcal{A}$ that are nested in the sense that $E_\alpha \subset E_\beta$ whenever $\alpha < \beta$, and that have $\bigcup_\alpha E_\alpha = X$ and $\bigcap_\alpha E_\alpha = \emptyset$. Define $f\colon X \to \mathbb{R}$ by $f(x) = \inf\{\alpha \in \mathbb{R} : x \in E_\alpha\}$. Prove that f is a measurable real-valued function such that $f(x) \le \alpha$ on each E_α and $f(x) \ge \alpha$ on each E_α^c.

Definition 2.7. Given a $[-\infty, \infty]$-valued function f, write $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$ for the positive and negative parts, so $f^\pm \ge 0$ and $f = f^+ - f^-$.

2.2. Integrating nonnegative functions.

Definition 2.8. The integral of a non-negative simple function $f = \sum_{i=1}^m c_i \mathbf{1}_{A_i}$ with respect to a measure μ is defined to be $\int f\,d\mu = \sum_{i=1}^m c_i \mu(A_i)$.

Exercise 21. (Bass 6.1): Suppose we write a non-negative simple function f in two different ways as $f = \sum_{j=1}^m a_j \mathbf{1}_{A_j} = \sum_{i=1}^n b_i \mathbf{1}_{B_i}$. Prove that $\int f$ is well-defined by showing that $\sum_{j=1}^m a_j \mu(A_j) = \sum_{i=1}^n b_i \mu(B_i)$.

Definition 2.9. The integral of a measurable function $f\colon X \to [0, \infty]$ is defined as

$$\int f\,d\mu = \sup\left\{ \int g\,d\mu : g \in \mathcal{S}_+ \text{ and } g \le f \right\}. \tag{3}$$

Proposition 2.10. For every measurable function $f\colon X \to [0, \infty]$, there is a sequence $f_n \in \mathcal{S}_+$ such that $f_1(x) \le f_2(x) \le \cdots$ and $f(x) = \lim_n f_n(x)$ for all $x \in X$. Given any such sequence, we have $\int f_n\,d\mu \to \int f\,d\mu$, where $\int f_n\,d\mu$ is interpreted as the sum from Definition 2.8, and $\int f\,d\mu$ is given by (3).

Proposition 2.11. The integral is linear and monotonic.

Definition 2.12. Two functions f and g are *equal μ-a.e.* if $\{x : f(x) \ne g(x)\}$ is a μ-null set. In this case we write $f = g$ a.e., or $f = g$ μ-a.e. if we need to highlight the measure.

Proposition 2.13. If $f, g\colon X \to [0,\infty]$ are measurable and $f = g$ a.e., then $\int f\,d\mu = \int g\,d\mu$.

Definition 2.14. Given a measure space $(X, \mathcal{A}, \mu)$, a measurable function $f\colon X \to [0,\infty]$, and a set $A \in \mathcal{A}$, we write $\int_A f\,d\mu = \int f \mathbf{1}_A\,d\mu$.

Exercise 22. $+$ With $X, \mathcal{A}, \mu, f$ as above, try to prove that $\nu(A) := \int_A f\,d\mu$ defines a measure on $(X, \mathcal{A})$. Explain why we do not yet have all the tools to prove this: where do you get stuck?

The previous exercise should lead you to the following fundamental question: **under what conditions does convergence of functions imply convergence of their integrals?**

Example 2.15. Pointwise convergence of functions **does not** imply convergence of integrals: $f_n = n\mathbf{1}_{(0,\frac{1}{n}]}$ has $f_n \to 0$ pointwise, but $\int f_n\,d\lambda = 1 \not\to 0 = \int 0\,d\mu$.

Theorem 2.16 (Monotone convergence theorem)**.** If $f_1, f_2, \dots$ are measurable $[0,\infty]$-valued functions such that $f_n \le f_{n+1}$ a.e., then writing $f(x) = \lim_n f_n(x)$ we have $\int f\,d\mu = \lim_n \int f_n\,d\mu$.

The next two results are easy consequences of the MCT (apply the MCT to $g_n = \sum_{k=1}^n f_k$ for the first, and $g_n = \inf_{k\ge n} f_k$ for the second).

Theorem 2.17 (Beppo Levi's theorem)**.** If $f_1, f_2, \dots$ are measurable $[0,\infty]$-valued functions, then $\int \sum_{k=1}^\infty f_k\,d\mu = \sum_{k=1}^\infty \int f_k\,d\mu$.

Theorem 2.18 (Fatou's lemma)**.** If $f_1, f_2, \dots$ are measurable $[0,\infty]$-valued functions, then $\int \varliminf_n f_n\,d\mu \le \varliminf_n \int f_n\,d\mu$.

Note that you can remember the direction of the inequality in Fatou's lemma by remembering the sequence of functions in Example 2.15.

Exercise 23. $+$ Let $\underline{f_n}$ be a sequence of non-negative measurable functions. Is it necessarily true that $\int \varlimsup_n f_n\,d\mu \le \varlimsup_n \int f_n\,d\mu$? If not, give a counterexample.

Exercise 24. Complete the proof you began in Exercise 22.

2.3. The space of integrable functions.

Definition 2.19. A $[-\infty,\infty]$-valued measurable function f is *integrable* if $\int f^+\,d\mu$ and $\int f^-\,d\mu$ are both finite. In this case we define $\int f\,d\mu = \int f^+\,d\mu - \int f^-\,d\mu$. When $\mu = \lambda$ is Lebesgue measure we may write either $\int f\,d\lambda$ or $\int f(x)\,dx$.

Proposition 2.20. A measurable function f is integrable if and only if $|f|$ is integrable, in which case $|\int f\,d\mu| \le \int |f|\,d\mu$.

Definition 2.21. The set of $\mathbb{R}$-valued integrable functions on a measure space $(X, \mathcal{A}, \mu)$ is denoted by $L^1(X, \mathcal{A}, \mu, \mathbb{R})$. We often omit one or more of the parameters in brackets in order to simplify notation.

Exercise 25. + (Bass 8.6) Let μ be a finite measure and $f\colon X \to \mathbb{R}$ measurable. Prove that $f \in L^1(\mu)$ iff $\sum_{n=1}^{\infty} \mu(\{x : |f(x)| \geq n\}) < \infty$.

To review: The definition of integration follows a schematic reminiscent of the one for premeasures, outer measures, and measures.

$\boxed{\begin{array}{l}\text{simple non-negative } f\text{:}\\ \int f = \sum c_i \mu(A_i)\end{array}} \longrightarrow \boxed{\begin{array}{l}\text{measurable non-negative } f\text{:}\\ \int f = \sup\{\int g : 0 \leq g \leq f \\ \qquad\qquad \text{and } g \text{ simple}\}\end{array}} \longrightarrow \boxed{\begin{array}{l}\text{integrable } f\text{:}\\ \int f = \int f^+ - \int f^-\end{array}}$

Proposition 2.22. L^1 is a vector space, and $\int\colon L^1 \to \mathbb{R}$ is a linear map with the property that $\int f\,d\mu \leq \int g\,d\mu$ whenever $f \leq g$ a.e.

Proposition 2.23. If $f\colon \mathbb{R} \to \mathbb{R}$ is Riemann integrable in the sense of undergraduate calculus, then it is integrable in the sense of Definition 2.19, and the two integrals agree.

If we interpret the integral as the area under the graph, then we can think of Riemann integration as slicing this region vertically, and Lebesgue integration (what we have defined here) as slicing it horizontally.

Exercise 26. Given $f \in L^1(\mathbb{R}, \lambda)$ and $a \in \mathbb{R}$, let $f_a(x) = f(x-a)$. Using translation-invariance of Lebesgue measure, prove that $f_a \in L^1(\mathbb{R}, \lambda)$ for all $a \in \mathbb{R}$ and that $\int f_a\,d\lambda = \int f\,d\lambda$.

Theorem 2.24 (Dominated convergence theorem)**.** Suppose $g\colon X \to [0,\infty]$ is integrable and $f, f_n\colon X \to [-\infty,\infty]$ are measurable, with the property that $|f_n| \leq |g|$ a.e. for all n and $f_n \to f$ a.e. Then $\int f_n\,d\mu \to \int f\,d\mu$.

Exercise 27. Let $g \in L^1(\mathbb{R}, \lambda)$ and let $f\colon \mathbb{R} \to \mathbb{R}$ be bounded, measurable, and continuous at 1. Prove that $\lim_n \int_{-n}^{n} f(1 + \frac{x}{n^2}) g(x)\,dx$ exists, and determine its value.

Exercise 28. * (Folland 2.22): Let μ be counting measure on $\mathbb{N}$. Interpret Fatou's lemma, the MCT, and the DCT as statements about infinite series.

Theorem 2.25 (Chebyshev's inequality)**.** Given $t > 0$ and $f \in L^1$, we have

$$\mu(\{x : |f(x)| \geq t\}) \leq \frac{1}{t} \int |f|\,d\mu.$$

This has the following consequences: (1) $\int |f|\,d\mu = 0$ implies $f = 0$ a.e., and (2) if f is $[-\infty,\infty]$-valued and integrable, then $|f| < \infty$ a.e.

3. Convergence

3.1. **Modes of convergence.** Given a set X and a sequence of functions $f_n\colon X \to \mathbb{R}$, there are two obvious senses in which the sequence f_n may be said to converge to a function $f\colon X \to \mathbb{R}$:

(1) *pointwise convergence*, in which $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in X$;
(2) *uniform convergence*, in which $\lim_{n\to\infty} \sup_{x\in X} |f_n(x) - f(x)| = 0$.

The relationship between these is most clearly seen by writing them as statements in first-order logic.

(1) Pointwise: $\forall x \in X\ \forall \varepsilon > 0\ \exists N \in \mathbb{N}$ s.t. $\forall n \geq N$ we have $|f_n(x) - f(x)| < \varepsilon$.
(2) Uniform: $\forall \varepsilon > 0\ \exists N \in \mathbb{N}$ s.t. $\forall n \geq N\ \forall x \in X$ we have $|f_n(x) - f(x)| < \varepsilon$.

Observe that for pointwise convergence, N is allowed to depend on both x and ε, while for uniform convergence it can only depend on ε.

Now consider a measure space $(X, \mathcal{A}, \mu)$, and let $\mathcal{M}$ denote the set of all measurable functions $f\colon X \to \mathbb{R}$. We are interested in studying different ways in which a sequence $f_n \in \mathcal{M}$ can converge to a limit $f \in \mathcal{M}$. In this setting, the notions above become more useful if we modify them slightly to account for the fact that we ignore behavior that only occurs on a null set.

Definition 3.1. f_n *converges to* f *pointwise almost everywhere* if there is a set $A \in \mathcal{A}$ such that $\mu(A) = 0$ and $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in A^c$. We usually write $f_n \to f$ μ-a.e., or $f_n \to f$ a.e., or $f_n \xrightarrow{\text{a.e.}} f$. This is also referred to as almost everywhere convergence or (in probability theory) as almost sure convergence.

Definition 3.2. f_n *converges to* f *uniformly almost everywhere* if there is a set $A \in \mathcal{A}$ such that $\mu(A) = 0$ and $\lim_{n\to\infty} \sup_{x\in A^c} |f_n(x) - f(x)| = 0$. This is also referred to as essentially uniform convergence.

The following definition and exercise justify the fact that essentially uniform convergence is also called L^∞-*convergence* and written $f_n \xrightarrow{L^\infty} f$.

Definition 3.3. The μ-*essential supremum* of a measurable function $f\colon X \to \mathbb{R}$ is

$$\begin{aligned}\mu\text{-}\operatorname{ess\,sup}(f) &:= \sup\{R \in \mathbb{R} : \mu\{x \in X : f(x) > R\} > 0\} \\ &= \inf\{R \in \mathbb{R} : \mu\{x \in X : f(x) > R\} = 0\}.\end{aligned}$$

(The essential infimum is defined similarly.) The L^∞-*norm* of f is

$$\|f\|_\infty = \mu\text{-}\operatorname{ess\,sup}(|f|).$$

Exercise 29. Show that $f_n \to f$ uniformly a.e. if and only if $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$.

Exercise 30. Show that if $f\colon \mathbb{R} \to \mathbb{R}$ is continuous, then λ-$\operatorname{ess\,sup}(f) = \sup(f)$.

It is clear that L^∞-convergence implies a.e.-convergence. The following two examples show that the converse fails; in both cases X is the real line (or a subset of it) with Lebesgue measure λ.

Example 3.4. (Horizontal escape to infinity) Define $f_n\colon \mathbb{R} \to \mathbb{R}$ by $f_n = \mathbf{1}_{[n,n+1]}$. Then $f_n \xrightarrow{\lambda\text{-a.e.}} 0$ but $\|f_n - 0\|_\infty = 1$ for all n.

Example 3.5. (Vertical escape to infinity) Define $f_n\colon [0,1] \to \mathbb{R}$ by $f_n = n\mathbf{1}_{[0,\frac{1}{n}]}$. Then $f_n \xrightarrow{\lambda\text{-a.e.}} 0$ but $\|f_n - 0\|_\infty = n \not\to 0$.

The second of these examples has the property that the convergence becomes uniform if we ignore a small set around 0. This is made precise by the following definition, which lies in between L^∞- and a.e.-convergence.

Definition 3.6. f_n *converges to* f *almost uniformly* if for every $\varepsilon > 0$ there is a set $A \in \mathcal{A}$ such that $\mu(A) < \varepsilon$ and $\lim_{n\to\infty} \sup_{x\in A^c} |f_n(x) - f(x)| = 0$. We may abbreviate this as $f_n \xrightarrow{\text{a.u.}} f$.

Clearly, L^∞-convergence implies almost uniform convergence, but the converse fails, as Example 3.5 shows.

Exercise 31. Show that if $f_n \xrightarrow{\text{a.u.}} f$ then $f_n \xrightarrow{\text{a.e.}} f$.

It follows from Example 3.4 that a.e.-convergence does not in general imply almost uniform convergence. However, this relies on the fact that $\lambda(\mathbb{R}) = \infty$, and for finite measures we have the following important result.

Theorem 3.7 (Egorov's theorem)**.** If $\mu(X) < \infty$ and $f_n \xrightarrow{\text{a.e.}} f$, then $f_n \xrightarrow{\text{a.u.}} f$.

There are two more notions of convergence that we will use.

Definition 3.8. f_n *converges to* f *in* L^1 if $\lim_{n\to\infty} \int |f_n - f| \, d\mu = 0$. We usually write $f_n \xrightarrow{L^1} f$. This is also referred to (in probability theory) as convergence in mean.

Definition 3.9. f_n *converges to* f *in measure* if for every $\varepsilon > 0$ the sequence of sets $A_n^\varepsilon := \{x \in X : |f_n(x) - f(x)| > \varepsilon\}$ satisfy $\lim_{n\to\infty} \mu(A_n^\varepsilon) = 0$. We may write $f_n \xrightarrow{\text{meas}} f$ or $f_n \xrightarrow{\mu} f$. This is also referred to (in probability theory) as convergence in probability.

Exercise 32. Show that if $f_n \xrightarrow{L^1} f$ then $f_n \xrightarrow{\text{meas}} f$.

Exercise 33. Show that if $f_n \xrightarrow{\text{a.u.}} f$ then $f_n \xrightarrow{\text{meas}} f$.

The various exercises above show that the five notions of convergence we have introduced satisfy the following logical relationships.

$$\begin{tikzcd} \boxed{f_n \xrightarrow{L^\infty} f} \arrow[r, Rightarrow] & \boxed{f_n \xrightarrow{\text{a.u.}} f} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \boxed{f_n \xrightarrow{\text{a.e.}} f} \\ \boxed{f_n \xrightarrow{L^1} f} \arrow[r, Rightarrow] & \boxed{f_n \xrightarrow{\text{meas}} f} & \end{tikzcd} \tag{$\star$}$$

In general, there are no implications between these modes of convergence apart from the ones in the diagram. For the top row we already saw this via Examples 3.4 and 3.5. It only remains to show that

(1) a.e.-convergence does not imply convergence in measure;
(2) L^∞-convergence does not imply L^1-convergence;

(3) L^1-convergence does not imply a.e.-convergence.

The first of these is accomplished by Example 3.4, where the horizontal escape to infinity gives $f_n \xrightarrow{\text{a.e.}} 0$ but $\int |f_n|\,d\lambda = 1$ for all n. The second and third claims are justified by the following two examples.

Example 3.10 (Widening)**.** Define $f_n : \mathbb{R} \to \mathbb{R}$ by $f_n = \frac{1}{n}\mathbf{1}_{[0,n]}$. Then $\|f_n\|_\infty \to 0$ but $\int |f_n|\,d\lambda = 1 \not\to 0$.

Example 3.11 (Circling)**.** Define $f_n : [0,1] \to \mathbb{R}$ by $f_n = \mathbf{1}_{I_n}$, where the intervals I_n are chosen as follows: first $I_1 = [0, \frac{1}{2}]$ and $I_2 = [\frac{1}{2}, 1]$, then I_3, I_4, I_5, I_6 are the four intervals $[\frac{k-1}{4}, \frac{k}{4}]$ for $k = 1, 2, 3, 4$, and so on, so the sequence I_n contains every interval of the form $[\frac{k-1}{2^n}, \frac{k}{2^n}]$ for $n \in \mathbb{N}$ and $1 \le k \le 2^n$, ordered so that $\lambda(I_n) \to 0$. Then $\int |f_n|\,d\lambda = \lambda(I_n) \to 0$, but f_n does not converge to 0 λ-a.e., since for every $x \in [0,1]$ there are infinitely many $n \in \mathbb{N}$ such that $x \in I_n$ and hence $f_n(x) = 1$.

In the last example we can think of the intervals I_n as "circling around the interval $[0,1]$ and shrinking". A more refined version of this example is given below as Exercise 36.

The four examples in this section – (H)orizontal, (V)ertical, (W)idening, (C)ircling – are sufficient to remember which implications between the five modes of convergence hold. Indeed, if A and B are two of the five modes of convergence, then either A implies B, or one of the four examples (H), (V), (W), (C) satisfies A but not B.

As Egorov's theorem suggests, there are still some other useful relationships between the five modes of convergence. We saw there that in for a *finite* measure, a.e.-convergence implies almost uniform convergence, so that in fact the two notions coincide in this setting.

Exercise 34. Show that if μ is finite and $f_n \xrightarrow{L^\infty} f$, then $f_n \xrightarrow{L^1} f$.

Of the four fundamental examples, (H) and (W) require the entire positive real line – an infinite measure space – while (V) and (C) require only the unit interval, which has finite measure. The two extra arrows that we can draw in $(\star)$ for finite measures (a.e. implies a.u., and L^∞ implies L^1) are precisely those for which (H) and/or (W) provided a counterexample but (V) and (C) did not. So when $\mu(X) < \infty$, we have the following modification of $(\star)$.

Finite measures:

$$\begin{tikzcd}
\boxed{f_n \xrightarrow{L^\infty} f} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \boxed{\begin{array}{c} f_n \xrightarrow{\text{a.u.}} f \\ f_n \xrightarrow{\text{a.e.}} f \end{array}} \arrow[d, Rightarrow] \\
\boxed{f_n \xrightarrow{L^1} f} \arrow[r, Rightarrow] & \boxed{f_n \xrightarrow{\text{meas}} f}
\end{tikzcd}$$

There are two more results connected to a.e.-convergence that hold for all measure spaces, which are worth mentioning.

Proposition 3.12. If $f_n \xrightarrow{\text{meas}} f$, then there is a subsequence with $f_{n_k} \xrightarrow{\text{a.e.}} f$.

Exercise 35. Show that if μ is finite, then $f_n \xrightarrow{\text{meas}} f$ if and only if every subsequence f_{n_k} has a further subsequence $f_{n_{k_j}}$ such that $f_{n_{k_j}} \xrightarrow{\text{a.e.}} f$.

Finally, although (C) shows that L^1-convergence does not imply a.e.-convergence, it turns out that the implication does hold if the L^1-convergence is fast enough.

Proposition 3.13. If $\sum_{n=1}^{\infty} \int |f_n - f| \, d\mu < \infty$, then $f_n \xrightarrow{\mu\text{-a.e.}} f$.

In fact the summability condition is optimal, in a sense made precise by the following exercise, which refines example (C).

Exercise 36. Given a sequence of real numbers $a_n \geq 0$, let $S_n = \sum_{k=1}^{n} a_k$, and let $J_n = [S_{n-1}, S_n] \subset [0,\infty)$. Let $\pi\colon [0,\infty) \to [0,1)$ take x to its fractional part, and let $I_n = \pi(J_n)$; then let $f_n = \mathbf{1}_{I_n}$ as before. Show that $\int f_n \, d\lambda = a_n$; then show that $f_n \xrightarrow{\lambda\text{-a.e.}} 0$ if and only if $\sum_n a_n < \infty$.

Note that this reduces to example (C) when $a_n = \frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \dots$.

3.2. **Metrics, norms, and inner products.** To put the notions of convergence from the previous section into a more general framework, we start by recalling the notion of convergence in metric spaces from undergraduate real analysis.

Definition 3.14. A *metric space* is a set Ω together with a distance function $d\colon \Omega \times \Omega \to \mathbb{R}$ satisfying the following axioms.

1. *Positivity:* $d(x,y) \geq 0$ for all $x,y \in \Omega$, with equality if and only if $x = y$.
2. *Reflexivity:* $d(x,y) = d(y,x)$ for all $x,y \in \Omega$.
3. *Triangle inequality:* $d(x,y) \leq d(x,z) + d(z,y)$ for all $x,y,z \in \Omega$.

Definition 3.15. Given a sequence x_n in a metric space (Ω, d), we say that x_n converges to $x \in \Omega$ if the following is true:

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \text{ s.t. } \forall n \geq N \text{ we have } d(x_n, x) < \varepsilon.$$

In this case we say that $x_n \to x$ with respect to d, and sometimes write $x_n \xrightarrow{d} x$.

Note that $x_n \xrightarrow{d} x$ if and only if $d(x_n, x) \to 0$ in $\mathbb{R}$. Taking $\Omega = \mathbb{R}$ and $d(x,y) = |x-y|$ recovers the usual definition of convergence of real numbers. More generally one can consider $\mathbb{R}^k$ with the Euclidean distance

$$d(\mathbf{x},\mathbf{y}) = \Big(\sum_{i=1}^{k} |x_i - y_i|^2\Big)^{1/2}, \qquad \mathbf{x} = (x_1,\dots,x_k), \quad \mathbf{y} = (y_1,\dots,y_k). \tag{4}$$

Now consider a sequence $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots \in \mathbb{R}^d$; note that we write the index of the sequence as a bracketed superscript, reserving subscripts for the the indices of the k-tuple, so we write $\mathbf{x}^{(n)} = (x_1^{(n)}, \dots, x_d^{(n)}) \in \mathbb{R}^k$.

Exercise 37. Given a sequence $(\mathbf{x}^{(n)})_n$ in $\mathbb{R}^k$ and a point $\mathbf{x} \in \mathbb{R}^k$, prove that the following are equivalent.

(1) $d(\mathbf{x}^{(n)}, \mathbf{x}) \to 0$.
(2) $\lim_{n\to\infty} \max_{1\le i\le k} |x_i^{(n)} - x_i| = 0$.
(3) $\lim_{n\to\infty} \sum_{i=1}^k |x_i^{(n)} - x_i|$.
(4) $\lim_{n\to\infty} x_i^{(n)} = x_i$ for all $1 \le i \le k$.

Given $\mathbf{v} \in \mathbb{R}^k$, writing $\|\mathbf{v}\|_\infty := \max_{1\le i\le k} |v_i|$ and $\|\mathbf{v}\|_1 := \sum_{i=1}^k |v_i|$. Then the first three conditions in Exercise 37 can be written as $\|\mathbf{x}^{(n)} - \mathbf{x}\|_p \to 0$ for $p = 2, 1, \infty$, respectively. These all fit into the following framework.

Definition 3.16. A *norm* on a real vector space V is a function $\|\cdot\|\colon V \to \mathbb{R}$ satisfying the following axioms.

(1) *Positivity:* $\|v\| \ge 0$ for all $v \in V$, with equality if and only if $v = 0$.
(2) *Homogeneity:* $\|cv\| = |c| \cdot \|v\|$ for all $v \in V$ and $c \in \mathbb{R}$.
(3) *Triangle inequality:* $\|v + w\| \le \|v\| + \|w\|$ for all $v, w \in V$.

Exercise 38. Prove that $\|\cdot\|_\infty$ and $\|\cdot\|_1$ as defined above give norms on $\mathbb{R}^k$.

Proposition 3.17. If $\|\cdot\|$ is a norm on V, then $d(v, w) = \|v - w\|$ defines a metric on V.

Playing around with different norms in finite dimensions does not change the notion of convergence.

Exercise 39. Prove that if $\|\cdot\|$ and $\|\cdot\|'$ are two norms on a finite-dimensional vector space V, then there is $C > 0$ such that $C^{-1}\|v\|' \le \|v\| \le C\|v\|'$ for all $v \in V$.

To prove that $\|\mathbf{v}\|_2 := \big(\sum_{i=1}^k |v_i|^2\big)^{1/2}$ is a norm on $\mathbb{R}^k$, and hence that d from (4) is a metric, the following definition is useful.

Definition 3.18. An *inner product* on a real vector space V is a function $\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{R}$ satisfying the following axioms.[3]

(1) *Positivity:* $\langle v, v\rangle \ge 0$ for all $v \in V$, with equality if and only if $v = 0$.
(2) *Symmetry:* $\langle v, w\rangle = \langle w, v\rangle$ for all $v, w \in V$.
(3) *Linearity:* $\langle cu + v, w\rangle = c\langle u, w\rangle + \langle v, w\rangle$ for all $c \in \mathbb{R}$ and $u, v, w \in V$.

Example 3.19. $\langle \mathbf{v}, \mathbf{w}\rangle = \sum_{i=1}^k v_i w_i$ is an inner product on $\mathbb{R}^k$.

Inner products induce norms by defining $\|v\| = (\langle v, v\rangle)^{1/2}$. This can be proved by the following steps: first expand $\langle v - w, v - w\rangle$ to deduce the following, which is related to Young's inequality.

Lemma 3.20. $\langle v, w\rangle \le \frac{\|v\|^2}{2} + \frac{\|w\|^2}{2}$ for all $v, w \in V$.

Applying this to $v/\|v\|$ and $w/\|w\|$ gives

Proposition 3.21 (Cauchy–Schwarz inequality)**.** $\langle v, w\rangle \le \|v\| \cdot \|w\|$ for all $v, w \in V$.

[3] If V is a complex vector space then $\langle v, w\rangle \in \mathbb{C}$ and instead of symmetry one should assume conjugate symmetry $\langle v, w\rangle = \overline{\langle w, v\rangle}$.

Exercise 40. Prove that if $\langle \cdot, \cdot \rangle$ is an inner product on V, then $\|v\| = (\langle v, v\rangle)^{1/2}$ defines a norm on V. *Hint: For the triangle inequality, apply Cauchy–Schwarz to the right-hand side of* $\langle v + w, v + w\rangle = \langle v, v + w\rangle + \langle w, v + w\rangle$.

Exercise 41. Show that if $\|v\|$ is a norm induced by an inner product, then it satisfies the *parallelogram law*: $\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2$ for all $v, w \in V$.

Exercise 42. Show that for $k \geq 2$, there is no inner product on $\mathbb{R}^k$ that induces the norm $\| \cdot \|_p$ for $p = 1$ or $p = \infty$.

3.3. **More about modes of convergence.** Return to the vector space $\mathcal{M}$ of measurable functions on a measure space $(X, \mathcal{A}, \mu)$. Are there metrics, norms, or inner products on $\mathcal{M}$ that correspond to the notions of convergence described earlier?

First note that since we identify functions that agree a.e., we should actually work not with $\mathcal{M}$ itself, but with the quotient space

$$\mathcal{M}' := \mathcal{M}/\sim, \text{ where } f \sim g \text{ if } f = g \ \mu\text{-a.e.}$$

Each element of $\mathcal{M}'$ is an equivalence class $[f] := \{g \in \mathcal{M} : g = f \ \mu\text{-a.e.}\}$. Equivalently, let $\mathcal{N} = \{f \in \mathcal{M} : f = 0 \ \mu\text{-a.e.}\}$, then $\mathcal{N}$ is a subspace of $\mathcal{M}$, and $\mathcal{M}'$ is the quotient space $\mathcal{M}/\mathcal{N}$. All of our definitions so far are well-behaved if we work with $\mathcal{M}'$ instead of $\mathcal{M}$; for example, if $f, g \in \mathcal{M}$ have $f = g$ μ-a.e., then $\int f\, d\mu = \int g\, d\mu$, so integration is well-defined on elements of $\mathcal{M}'$.

Exercise 43. * Prove that if a sequence $f_n \in \mathcal{M}$ converges to $f \in \mathcal{M}$ in any of the five senses from the previous section, then given any sequence $g_n \in [f_n]$ and any $g \in \mathcal{M}[f]$, we have $g_n \to g$ (in the same sense) if and only if $g \in [f]$.

This exercise shows that we can (and will) work with equivalence classes of functions as the elements of the space on which we construct a metric, norm, etc.

Definition 3.22. Given a σ-finite measure space $(X, \mathcal{A}, \mu)$, let $L^\infty(X, \mathcal{A}, \mu) = \{[f] \in \mathcal{M}' : \|f\|_\infty < \infty\}$, where $\|f\|_\infty = \mu\text{-ess}\sup(|f|)$ as in Definition 3.3. We will often write L^∞ or $L^\infty(X)$ or $L^\infty(\mu)$ when there is no risk of confusion.

Definition 3.23. Given a measure space $(X, \mathcal{A}, \mu)$, let $L^1(X, \mathcal{A}, \mu) = \{[f] \in \mathcal{M}' : f$ is integrable$\}$, and let $\|f\|_1 = \int |f|, d\mu$.

Although elements of L^∞ and L^1 are equivalence classes of functions, we will often abuse notation and simply write $f \in L^1$ (and so on), conflating f and $[f]$. It is easy to show that L^∞ and L^1 are both normed vector spaces, and that

$$f_n \xrightarrow{L^\infty} f \text{ if and only if } \|f_n - f\|_\infty \to 0,$$

$$f_n \xrightarrow{L^1} f \text{ if and only if } \|f_n - f\|_1 \to 0.$$

Thus L^∞ and L^1-convergence both come from norms. What about the other three notions of convergence?

Exercise 44. Suppose (Ω, d) is a metric space, $x \in \Omega$, and $x_n \in \Omega$ is a sequence with the property that every subsequence x_{n_k} has a further subsequence $x_{n_{k_j}}$ that converges to x in the metric d. Prove that $d(x_n, x) \to 0$.

Exercise 45. Use the previous exercise together with Exercise 35 to prove that when $X = [0,1]$ and λ is Lebesgue measure, there is no metric on $\mathcal{M}'$ that induces the notion of a.e.-convergence.

Because a.e.-convergence and almost uniform convergence are equivalent for finite measures, we conclude that neither of these comes from a metric in general. It only remains to consider convergence in measure.

Exercise 46. Let μ be a finite measure on X and define $\phi\colon [0,\infty) \to [0,1)$ by $\phi(r) = \frac{r}{1+r}$. Prove that $d([f],[g]) = \int \phi(|f-g|)\,d\mu$ defines a metric on $\mathcal{M}'$, and that $d([f_n],[f]) \to 0$ if and only if $f_n \xrightarrow{\text{meas}} f$.

Thus convergence in measure is induced by a metric. However, it is not induced by any norm, as the next exercise shows.

Exercise 47. + Consider the unit interval with Lebesgue measure. Let f_n be the sequence of functions from example (C), and prove that for any sequence $c_n \in \mathbb{R}$ we have $c_n f_n \xrightarrow{\text{meas}} 0$. Show that for any norm $\|\cdot\|$ on $\mathcal{M}'$, there is a sequence c_n such that $\|c_n f_n\| \not\to 0$. *Hint: for the second part, use the fact that the constant function 1 can be written as $f_1 + f_2$, and as $f_3 + f_4 + f_5 + f_6$, and so on, in order to get a lower bound on $\|f_n\|$ for some n.*

3.4. L^p spaces.

Definition 3.24. A sequence x_n in a metric space X is a *Cauchy sequence* if

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N} \text{ s.t. } \forall m,n \geq N \text{ we have } d(x_m, x_n) < \varepsilon.$$

X is *complete* if every Cauchy sequence converges.

Definition 3.25. A normed vector space V is a *Banach space* if it is complete in the metric induced by the norm. A Banach space whose norm is induced by an inner product is called a *Hilbert space*.

Exercise 48. Prove that $\mathbb{R}^k$ is complete in the metric induced by the Euclidean norm.

Since all norms on a finite-dimensional space are equivalent, this shows that every finite-dimensional normed vector space is a Banach space. The infinite-dimensional situation is rather more subtle.

Exercise 49. Let ℓ^∞ be the set of all bounded sequences of real numbers, with the norm $\|x\|_\infty := \sup_n |x_n|$. Prove that ℓ^∞ is a Banach space.

Exercise 50. More generally, let μ be σ-finite and prove that $L^\infty(\mu)$ is a Banach space.

Recall the following definitions from topology.

Definition 3.26. Given a metric space Ω, a set $U \subset \Omega$ is *open* if

$$\forall x \in U \ \exists \varepsilon > 0 \text{ s.t. } B(x,\varepsilon) \subset U,$$

where $B(x,\varepsilon) = \{y \in \Omega : d(x,y) < \varepsilon\}$. A set V is *closed* if V^c is open. The *closure* of a set A is the closed set

$$\overline{A} = \{x \in \Omega : B(x,\varepsilon) \cap A \neq \emptyset \ \forall \varepsilon > 0\}.$$

Exercise 51. Given a complete metric space (Ω, d) and a subset $A \subset \Omega$, prove that A is closed if and only if (A, d) is a complete metric space.

Exercise 52. Consider the following sets of sequences:

$$\begin{aligned} c_0 &:= \{x \in \ell^\infty : x_n \to 0\}, \\ c_{00} &:= \{x \in \ell^\infty : \exists N \in \mathbb{N} \text{ s.t. } x_n = 0 \ \forall n \geq N\}. \end{aligned}$$

Show that c_0 and c_{00} are linear subspaces of ℓ^∞, and that c_0 is the closure of c_{00} in the ℓ^∞-norm. Use this to deduce that c_0 is a Banach space and c_{00} is not.

To determine whether a normed vector space is complete, it turns out to be helpful to have a criterion in terms of series convergence.

Definition 3.27. Let V be a normed vector space and $v_1, v_2, \dots \in V$. The series is *absolutely convergent* if $\sum_{n=1}^\infty \|v_n\| < \infty$, and *convergent* if there if $v \in V$ such that the partial sums $w_N = \sum_{n=1}^N v_n$ converge to v in norm: $\|v - w_N\| \to 0$.

Proposition 3.28. Let V be a normed vector space. Then V is complete if and only if every absolutely convergent series in V is convergent.

Example 3.29. Let $(X, \mathcal{A}, \mu)$ be a measure space and fix $1 \leq p < \infty$. The set

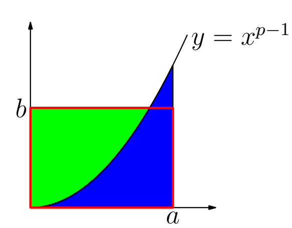
$$L^p(X, \mathcal{A}, \mu) := \left\{ [f] \in \mathcal{M}' : \int |f|^p \, d\mu < \infty \right\}$$

is a vector space. We often write $L^p(X)$ or $L^p(\mu)$ or L^p when there is no risk of confusion.

Definition 3.30. Given $f \in L^p$, we write $\|f\|_p := (\int |f|^p \, d\mu)^{1/p}$ for the *L^p-norm of f*. Of course, to justify this terminology one must prove that this is in fact a norm.

Most properties of a norm can be easily verified for $\|\cdot\|_p$. The only one that is not immediate is the triangle inequality. This takes a little work to prove, but can be done by mimicking the argument for inner product spaces. The following analogue of Lemma 3.20 is crucial.

Lemma 3.31 (Young's inequality)**.** Given $1 \leq p < \infty$, let q be the *conjugate exponent* defined by the condition $\frac{1}{p} + \frac{1}{q} = 1$. Then for every $a, b \geq 0$ we have $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.



Young's inequality can be proved either using concavity of logarithm and rewriting the left-hand side as $e^{\frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q}$, or by comparing areas in the figure shown.

Given $f \in L^p$ and $g \in L^q$, one can write $\hat{f} = f/\|f\|_p$ and $\hat{g} = g/\|g\|_q$, then apply Young's inequality to $|\hat{f}(x)\hat{g}(x)|$ at each $x \in X$ to obtain the following generalization of Cauchy–Schwarz.

Proposition 3.32 (Hölder's inequality)**.** If p, q are conjugate exponents, $1 \leq p, q \leq \infty$, and $f \in L^p$, $g \in L^q$, then $fg \in L^1$ and we have $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

Finally, with a little bit of computation one can mimic the proof of the triangle inequality from Cauchy–Schwarz to prove the triangle inequality for L^p-norms. (The $p = \infty$ case is an easy exercise.)

Proposition 3.33 (Minkowski's inequality)**.** Given $1 \leq p \leq \infty$ and $f, g \in L^p$, we have $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

All of this shows that $L^p(\mu)$ is a normed vector space for every $1 \leq p < \infty$. In fact it is complete as well, by Proposition 3.28 and the following.

Proposition 3.34. If $f_n \in L^p$ are such that $\sum_n \|f_n\|_p < \infty$, then there is $f \in L^p$ such that $\lim_{N\to\infty} \|f - \sum_{n=1}^N f_n\|_p = 0$.

The L^p spaces are one of the fundamental examples of Banach spaces. Some justification for the notation L^∞ is provided by the following.

Exercise 53. Let μ be a finite measure and f a measurable function. Prove that $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$.

It is useful to keep in mind examples of functions which are in some L^p spaces but not others.

Example 3.35. Given $\alpha > 0$, let $f_\alpha(x) = x^{-\alpha}$ for $x \in (0, \infty)$. Then

$$f_\alpha \in L^p([0,1]) \Leftrightarrow \alpha p < 1 \Leftrightarrow \alpha < 1/p,$$
$$f_\alpha \in L^p([1,\infty)) \Leftrightarrow \alpha p > 1 \Leftrightarrow \alpha > 1/p.$$

A special case occurs when $X = \mathbb{N}$ and μ is counting measure; in this case we write ℓ^p for the set of all infinite sequences with $\sum_n |x_n|^p < \infty$. Recall that the case $p = \infty$ appeared already in an exercise.

Exercise 54. Show that $c_{00} \subset \ell^p \subset c_0$, and that both inclusions are strict.

Exercise 55. Show that the closure of c_{00} in the ℓ^p-norm is ℓ^p.

The fact that an arbitrary element of a Banach space can often be approximated with elements taken from a more restrictive subspace was, in some sense, at the heart of many of our earlier results on integration, where we approximated arbitrary measurable functions with simple functions.

Recall that $\mathcal{S}$ denotes the collection of all measurable simple functions on X. This can also be described as the vector space spanned by the set of functions of the form $\mathbf{1}_A$, where $A \subset X$ is measurable.

Proposition 3.36. Given a measure space $(X, \mathcal{A}, \mu)$ and $1 \le p \le \infty$, the subspace $\mathcal{S}$ is dense in $L^p(\mu)$.

Now we consider the specific case $X = \mathbb{R}$ and $\mu = \lambda$.

Definition 3.37. A function $f \colon \mathbb{R} \to \mathbb{R}$ is a *step function with bounded support* if there are $a_0 < a_1 < \cdots < a_n$ such that $f|_{(a_{i-1}, a_i)}$ is constant for each $1 \le i \le n$, and $f(x) = 0$ for all $x < a_0$ and all $x > a_n$.

The set $\mathcal{S}_0$ of all step functions with bounded support is the vector space spanned by functions of the form $\mathbf{1}_{(a,b)}$, where $a < b \in \mathbb{R}$.

Proposition 3.38. $\mathcal{S}_0$ is a dense subspace of $L^p(\mathbb{R}, \lambda)$ for every $1 \le p < \infty$.

This has a couple useful consequences.

Proposition 3.39. The space $C_c(\mathbb{R}) = \{f \colon \mathbb{R} \to \mathbb{R} : f \text{ is continuous and } \exists R > 0 \text{ s.t. } f(x) = 0\ \forall |x| > R\}$ is dense in $L^p(\mathbb{R}, \lambda)$ for every $1 \le p < \infty$.

Exercise 56. Given $f \colon \mathbb{R} \to \mathbb{R}$ and $a \in \mathbb{R}$, define a function $f_a \colon \mathbb{R} \to \mathbb{R}$ by $f_a(x) = f(x - a)$; note that this is the function whose graph is equal to the graph of f, shifted horizontally by a units. Show that if $1 \le p < \infty$, then for any $f \in L^p(\mathbb{R}, \lambda)$ and $b \in \mathbb{R}$ we have $\lim_{a \to b} \|f_a - f_b\|_p = 0$.

Note that these last three results all fail for $p = \infty$.