Analysis for people who don’t like skipping details

Emma Yoneda

2026-01-22

Contents

Part I

Set Theory

Chapter 1

Axioms of Set Theory

1.1 Set Theoretic Axioms

1.2 Operations on Sets

1.3 Relations and Maps

Theorem 1.3.1. (Some identities for preimages): *Let $f : X \to Y$ be a bijective map. Then for all $Y_i \subset Y$, the following identities hold:*

1. *The preimage of a union of sets is the union of the preimages:*

$$f^{-1}\left(\bigcup_{i\in I} Y_i\right) = \bigcup_{i\in I} f^{-1}(Y_i)$$

2. *The preimage of an intersection of sets is the intersection of the preimages:*

$$f^{-1}\left(\bigcap_{i\in I} Y_i\right) = \bigcap_{i\in I} f^{-1}(Y_i)$$

3. *The preimage of the complement of a set is the complement of its preimage:*

$$f^{-1}(Y \setminus Y_i) = X \setminus f^{-1}(Y_i)$$

1.4 Orders

Definition 1.4.1. Let S be a subset of a partially ordered set $(P, \leq)$. Then:

1. A **lower bound** of S is an element $y \in P$ such that, for all $x \in S$, $y \leq x$.
2. An **upper bound** of S is an element $y \in P$ such that, for all $x \in S$, $x \leq y$.

Definition 1.4.1.1: Infimum and Supremum

Let S be a subset of a partially ordered set $(P, \leq)$. Then:

1. A lower bound b of S is called an **infimum**, or **meet** of S if it is the **greatest lower bound** of S, meaning that for all lower bounds y of S, we have

$$y \leq b$$

If b is an infimum of S, we write:

$$\inf S = \inf_{s \in S} s := b$$

2. An upper bound B of S is called a **supremum**, or **join** of S if it is the **least upper bound** of S, meaning that for all upper bounds y of S, we have

$$B \leq y$$

If B is a supremum of S, we write:

$$\sup S = \sup_{s \in S} s := B$$

It is often practical to use a slightly expanded notation that lets us implicitly specify subsets of P fulfilling a property φ:

$$\inf \{p \in P \mid \varphi(p)\} = \inf_{\substack{p \in P \\ \varphi(p)}} p$$

Part II

Algebraic Structures

Chapter 2

Ordered Fields

Definition 2.0.0.1: Ordered Commutative Ring

An **ordered commutative Ring** $(R, \leq)$ is a commutative Ring R equipped with an **ordering relation** $\leq$, such that for all $a, b, c \in R$, we have:

1. $\leq$ defines a *total order* on F. i.e:
   (a) $a \leq a$ (the order is *reflexive*),
   (b) $a \leq b \wedge b \leq c \implies a \leq c$ (the order is *transitive*),
   (c) $a \leq b \wedge b \leq a \implies a = b$ (the order is *antisymmetric*),
   (d) $a \leq b \vee b \leq a$ (the order is *strongly connected*)
2. $a \leq b \implies a + c \leq b + c$
3. $0 \leq a \wedge 0 \leq b \implies 0 \leq ab$

Lemma 2.0.1. *For every ordered commutative Ring $(R, \leq)$ and $a \in R$, we have $-a \leq 0 \leq a$ or $a \leq 0 \leq -a$.*

Proof. Since the order $\leq$ is strongly connected, we have $a \leq 0$ or $0 \leq a$.

1. If $a \leq 0$, then we have $-a + a \leq 0 + -a$, i.e. $a \leq 0 \leq -a$,
2. if $0 \leq a$, then we have $-a + 0 \leq -a + a$, i.e. $-a \leq 0 \leq a$

□

Lemma 2.0.2. *Let $a \in (R, \leq)$. Then $0 \leq a^2$.*

Proof. Since the order $\leq$ is strongly connected, we have $a \leq 0$ or $0 \leq a$.

1. If $0 \leq a$, then we have $0 \leq a \cdot a = a^2$,
2. if $a \leq 0$, then $0 \leq -a$ and we have $0 \leq -a \cdot (-a) = a^2$.

□

Lemma 2.0.3. *Every ordered field has characteristic* 0.

Proof. Assume that F is a field of characteristic p. Then an ordering relation would need to fulfill:

$$1 \leq 1 + 1 \leq \sum_{i=1}^{p} 1 = 0 \leq 1$$

Which implies $0 = 1$. However, by the definition of a field, we have $0 \neq 1$. □

Lemma 2.0.4. *Let* $a \leq b$ *and* $c \geq 0$. *Then* $ac \leq bc$.

Proof. Since $a \leq b$, we have $0 = a - a \leq b - a$. Therefore, we also have $0 \leq (b - a)c = bc - ac$. Adding ac to both sides, we get $ac \leq bc$. □

Corollary 2.0.5. *Let* $a \leq b$. *Then* $a^{-1} \geq b^{-1}$.

Proof.

$$\begin{aligned} & a \leq b \\ \implies & 1 = aa^{-1} \leq ba^{-1} \\ \implies & b^{-1} \leq b^{-1}ba^{-1} = a^{-1} \end{aligned}$$

□

Summary 2.0.6. Let F be an ordered Field and let $a, b, c \in F$ Then all of the following hold:

1. $a \leq a$
2. If $a \leq b$ and $b \leq c$, then $a \leq c$ (transitivity)
3. If $a \leq b$ and $b \leq a$, then $a = b$ (antisymmetry)
4. We always have at least one of $a \leq b$ and $b \leq a$
5. If $a \leq b$, then $a + c \leq b + c$
6. If $0 \leq a$ and $0 \leq b$, then $0 \leq ab$.
7. If $0 \leq a$, then $-a \leq 0$.
8. $0 \leq a^2$
9. If $a \leq b$ and $c \geq 0$, then $ac \leq bc$.
10. If $a \leq b$ and $c \leq 0$, then $ac \geq bc$.

2.1 The Archimedean Property

Definition 2.1.1. Let F be an archimedean ordered field. Then we say that F is **archimedean** if for every $x, y \in F_{>0}$, there exists a natural number n such that

$$nx > y$$

Comment 2.1.2. It follows immediately that if F is non-archimedean, there exists $x, y \in F$ such that for all natural numbers n, we have

$$nx < y$$

which immediately implies

$$n_F = \sum_{k=1}^{n} 1_F = \sum_{k=1}^{n} xx^{-1} = x^{-1}nx < x^{-1}y := y'$$

Therefore, there exists an element y' that is "infinitely large", i.e. it is greater than the image of the embedding of any natural number into the field. It immediately follows that $\frac{1}{y'} < \frac{1}{n_F}$ for all $n \in \mathbb{N}$, meaning that F also contains "infinitely small" elements.

2.2 Why always $\mathbb{R}$?

If you're the kind of person who generally prefers algebra to analysis, you might have always felt unsatisfied by a seeming lack of generality to analysis - why does everyone only ever seem to care about $\mathbb{R}$? The goal of this chapter is to make you feel like you finally have a satisfying answer - we will prove that $\mathbb{R}$ is *the only* ordered field, up to isomorphism, that has the key property that every bounded set has a least upper bound.

Whenever someone gives a definition explicitly concerning $\mathbb{R}$, they are giving a definition concerning ordered fields with the least upper bound property - it just so happens that $\mathbb{R}$ is the only such field!

2.2.1 Subfields of ordered fields

Theorem 2.2.1. *Let F be an archimedean ordered field. Then F is isomorphic to a subfield of the real numbers $\mathbb{R}$.*

This means that $\mathbb{R}$ can be viewed as a "maximal archimedean ordered field". Later we will prove that $\mathbb{R}$ is also unique up to isomorphism, meaning that it is *the* maximal archimedean ordered field. This realization is a key step on our journey of justifying the ubiquity of the real numbers.

2.2.2 The least upper bound property

Definition 2.2.2. Let F be an ordered field. We say that F has the **least upper bound property**, or alternatively that F is **Dedekind complete**, if every subset of F that has an upper bound in F has a least upper bound in F.

Theorem 2.2.3. *F has the least upper bound property if and only if it has the equivalent "greatest lower bound property", i.e. every subset of F that has a lower bound in F has a greatest lower bound in F.*

Theorem 2.2.4. *Let F be a non-archimedean ordered field. Then F does not have the least-upper-bound property.*

Proof. Since F is an ordered field, it must have characteristic 0. Let N_F be the infinite set

$$N_F : \left\{ \sum_{k=0}^{n} 1_F : n \in \mathbb{N} \right\}$$

Since F is non-archimedean, there exists an element x such that for all $n \in N_F$, we have $n < x$. However, for any upper bound b of N_F, we have that for all $n \in N_F$, $b > n + 1 \in N_F$. Therefore, $b - 1$ is also an upper bound, meaning that no least upper bound exists. □

Importantly, this immediately implies that Cauchy-completeness of an ordered field is *not* equivalent to Dedekind-completeness!

Corollary 2.2.5. *Let F be an ordered field. Then F has the least-upper-bound property if and only if it is archimedean and Cauchy complete.*

Theorem 2.2.6. *Every ordered field with the least upper bound property is isomorphic. Therefore the real numbers $\mathbb{R}$ are, up to isomorphism, the only ordered field with the least upper bound property.*

2.2.3 Alternative completeness properties

Part III

Topology

Chapter 3

Metric Spaces and Topological Spaces

3.1 Vocabulary

Definition 3.1.1. Let X be a topological space, $x \in X$, and $V \subset X$. We call V a *neighborhood of* x if there exists an open set $U \subset V$ such that $x \in U$.

Theorem 3.1.2. *Let X be a topological space and let $V \subset X$. Then V is open if and only if for every $x \in V$, V is a neighborhood of x.*

> *Proof.* If V is open then it is trivially a neighborhood of all of its points.
>
> Assume that V is a neighborhood of all its points. Let $U_x \subset V$ be the necessary open set containing $x \in V$ that makes V a neighborhood of x. Then since every U_x is a subset of V we have
>
> $$\bigcup_{x \in V} U_x \subset V$$
>
> and since every $x \in V$ is contained in some U_x we also have
>
> $$V \subset \bigcup_{x \in V} U_x$$
>
> Therefore V is a union of open sets, making it open. □

Definition 3.1.3. Let X be a topological space. We say that a subset of X is F_σ (from French "*fermé*", "closed", and "*somme*", "sum, union") if it is a countable union of closed sets. Dually, we say it is G_δ (from German "*Gebiet*", an old term for "open set", and "*Durschschnitt*", "average, intersection") if it is a countable intersection of open sets.

Theorem 3.1.4. *The complement of a G_δ set is F_σ and vice versa.*

3.2 Sequences and Limits

Definition 3.2.0.1: Liminf and Limsup

Let X be a topological space that is linearly orderable by an order $\leq$. Let x_n be a sequence in X. Then we define:

$$\liminf_{n\to\infty} x_n := \lim_{n\to\infty}\left(\inf_{m\geq n} x_m\right)$$

and

$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty}\left(\sup_{m\geq n} x_m\right)$$

Corollary 3.2.1. *We have:*

$$\liminf_{n\to\infty} x_n = \sup_{n\geq 0}\left(\inf_{m\geq n} x_m\right)$$

and

$$\limsup_{n\to\infty} x_n = \inf_{n\geq 0}\left(\sup_{m\geq n} x_m\right)$$

3.3 Continuity

The notion of continuity is central to analysis (and of key importance to mathematics and general), and one could argue the most important reason why the field of topology is of interest in the first place is because it gives us the most general setting in which we can define a notion of a continuous function. There are many different definitions of continuity in various levels of generality.

Definition 3.3.0.1: Continuous function

Let $f : X \to Y$ be a function between topological spaces. We call f **continuous** if the preimage $f^{-1}(U)$ of any open set U is again an open set.

If the two extremal topologies are involved, continuity of a function is often trivial to verify:

Theorem 3.3.1. *Let $f : X \to Y$ be a function between topological spaces. Assume Y has the trivial topology. Then f is continuous.*

> *Proof.* By definition of the trivial topology, the only open sets in Y are Y itself and the empty set. We have $f^{-1}(Y) = X$, which is open, and $f^{-1}(\emptyset) = \emptyset$, which is also open. □

Theorem 3.3.2. *Let $f : X \to Y$ be a function between topological spaces. Assume X has the discrete topology. Then f is continuous.*

Proof. Every subset of X is open, therefore the preimage f^{-1} of any set must be open. □

Definition 3.3.3. Let $f : X \to Y$ be a function between topological spaces. Let $x \in X$. We call f **continuous at** x if, for any neighborhood $V \subset Y$ of $f(x)$, there exists a neighborhood $U \subset X$ of x such that $f(U) \subset V$.

Lemma 3.3.4. *$f : X \to Y$ is continuous at $x \in X$ if and only if, for every neighborhood $V \subset Y$ of $f(x)$, we have that $f^{-1}(V)$ is a neighborhood of x.*

Proof.

$\Longrightarrow$: If $f(U) \subset V$, then by the definition of preimages we have $U \subset f^{-1}(V)$. Therefore, since U is a neighborhood of x, the superset $f^{-1}(V)$ must be a neighborhood of x as well.

$\Longleftarrow$: If $f^{-1}(V)$ is a neighborhood of x, then $U = f^{-1}(V)$ already fulfills our definition.

□

Theorem 3.3.5. *$f : X \to Y$ is continuous if and only if it is continuous at every point $x \in X$.*

Proof.

$\Longrightarrow$: Let f be continuous and let $x \in X$. Then if V is a neighborhood of $f(x)$, there must exist an open set U such that contains $f(x) \in U \subset V$. Then $f^{-1}(U) \subset f^{-1}(V)$ is an open set containing x, meaning that $f^{-1}(V)$ is a neigborhood of x. Therefore f is continuous at every x

$\Longleftarrow$: Let $V \subset X$ be open. Then $f^{-1}(V)$ is a neighborhood every $x \in f^{-1}(V)$. Therefore, $f^{-1}(V)$ is open.

□

Definition 3.3.5.1: Sequentially continuous functions

Let $f : X \to Y$ be a function between topological spaces. We call f **sequentially continuous at a point** x if, for every sequence x_n such that $\lim_{n\to\infty} x_n = x$, we have $\lim_{n\to\infty} f(x_n) = f(x)$. We say the function is **sequentially continuous** if this condition holds for every point $x \in X$.

This definition most directly captures the intuitive idea that a function is continuous if $f(x)$ gets arbitrarily close to $f(y)$ whenever x gets arbitrarily close to y.

Theorem 3.3.6. *Every continuous function is sequentially continuous.*

Theorem 3.3.7. *If X is first-countable (and we assume the axiom of choice), then any sequentially continuous function is continuous.*

Corollary 3.3.7.1

A function $f : X \to Y$ from a first-countable space X into any topological space Y is continuous if and only if it is sequentially continuous.

In particular, continuity and sequential continuity are equivalent for functions between metric spaces.

Theorem 3.3.7.1: Epsilon-Delta-Criterion

Let $f : M \to N$ be a function between metric spaces. Then f is continuous at a point $x \in M$ if and only if for every $\varepsilon \in \mathbb{R}_{>0}$, there exists a $\delta \in \mathbb{R}_{>0}$ such that for all $y \in M$, we have that

$$d_M(x, y) < \delta \implies d_N(f(x), f(y)) < \varepsilon$$

This is the standard definition of continuity used in most introductory courses in real analysis, since it can be easily defined for $f : \mathbb{R} \to \mathbb{R}$ even if topological spaces and metric spaces haven't been introduced yet. Since it is only defined for functions between metric spaces, it is less general than most of our other definitions, but it has the advantage of often leading to simpler proofs.

Proof. $\Rightarrow$: Assume that f is sequentially continuous at a point x, but that the given condition doesn't hold. Then there exists an $\varepsilon \in \mathbb{R}_{>0}$ such that for every $\delta \in \mathbb{R}_{>0}$, there exists an $x_\delta \in M$ such that

$$d_M(x, x_\delta) \leq \delta, \text{ but } d_N(f(x), f(x_\delta)) \geq \varepsilon$$

Therefore, if we define $\delta_n := \frac{1}{n}$, then the sequence x_{δ_n} converges to x, but the sequence $f(x_{\delta_n})$ doesn't converge to $f(x)$, since $d_N(f(x), f(x_\delta)) \geq \varepsilon > 0$.

$\Leftarrow$: Let x_n be a sequence with $\lim_{n\to\infty} = x$ which fulfills our condition. We need to show $\lim_{n\to\infty} f(x_n) = f(x)$, meaning that for every $\varepsilon \in \mathbb{R}_{>0}$, we need to find an $N \in \mathbb{N}$, such that for all $n \geq N$, we have

$$d_N(f(x_n) - f(x)) < \varepsilon$$

by our epsilon-delta condition, this holds for every x_n such that $d(x_n, x) < \delta$. Since $\lim_{n\to\infty} x_n = x$, we can find an N such that this condition is fulfilled for all $n > N$. Therefore does $f(x_n)$ indeed converge to $f(x)$.

□

Theorem 3.3.8. *X be a topological space and let $A \subset X$. Then, assuming the discrete topology on $\{0, 1\}$, the indicator function $\mathbb{1}_A : X \to \{0, 1\}$ is continuous at a point $x \in X$ if and only if $x \in \operatorname{int}(A)$ or $x \in \operatorname{int}(X \setminus A)$.*

Proof. 1. Let $x \in \operatorname{int}(A)$. Then by definition of the interior of a set there exists an open set $U \subset A$ that contains x. Since $U \subset A$, we have $f(U) = \{1\}$. Therefore, if V is a neighborhood around $f(x) = 1$, then $f^{-1}(V)$ must contain U, making it a neighborhood of x.

2. Let $x \in \operatorname{int}(X \setminus A)$. Then the same argument as before applies, except we have a $U \subset X \setminus A$ with $f(U) = \{0\}$.

3. Let $x \in \partial A$ with $x \in A$. Then $V = \{1\}$ is an open neighborhood of $f(x)$, but $f^{-1}(V) \subset A$. However, since x is on the boundary of A, every open set containing x must contain points in $X \setminus A$. Therefore $f^{-1}(V)$ cannot be a neighborhood of x.

4. Let $x \in \partial A$ with $x \in X \setminus A$. Then the same argument applies to $V = \{0\}$, since $f^{-1}(V)$ cannot contain points in A.

□

Comment 3.3.9. We have to assume the discrete topology on $\{0, 1\}$, since if $\{0\}$ is not open, then the function ends up continuous at points $x \in \partial A \setminus A$, and if $\{1\}$ is not open, then the function ends up continuous at points $x \in \partial A \cap A$.

Corollary 3.3.10. *The characteristic function of the rational numbers (also known as the **Dirichlet function**) is nowhere continuous.*

Proof. Assuming the standard topology on $\mathbb{R}$, the interiors of both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are empty. □

Theorem 3.3.10.1: A function continuous at exactly one point

The function $f : \mathbb{R} \to \mathbb{R}$ with

$$f(x) = x \cdot \mathbb{1}_{\mathbb{Q}}(x) = \begin{cases} x & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

is continuous at 0 and discontinuous at every other point.

Proof. 1. Let V be a neighborhood of $f(0) = 0$. Then by definition, there must be an $\varepsilon > 0$ such that $(-\varepsilon, \varepsilon) \in V$. Then since $f(x) \leq x$, we have $f^{-1}(y) \geq y$, implying that

$$(-\varepsilon, \varepsilon) \subset f^{-1}((-\varepsilon, \varepsilon)) \subset f^{-1}(V)$$

and therefore $f^{-1}(V)$ is a neighborhood of 0.

2. Let $x \in \mathbb{Q} \setminus \{0\}$. Then, since all irrationals get mapped to zero, the preimage of $(\frac{1}{2}x, \frac{3}{2}x)$ only contains rational numbers and therefore cannot be a neighborhood of x.

3. Let $x \notin \mathbb{Q}$. Then the preimage of $(-\frac{1}{2}x, \frac{1}{2}x)$ contains x, but not any rationals between x and $\frac{1}{2}x$, and therefore cannot be a neighborhood of x.

□

Theorem 3.3.10.2: A function only continuous at the irrationals

$T : \mathbb{R} \to \mathbb{R}$, defined as

$$T(x) = \begin{cases} \frac{1}{q} & x \in \mathbb{Q}, x = \frac{p}{q}, p, q \text{ have no common divisors} \\ 0 & x \notin \mathbb{Q} \end{cases},$$

is discontinuous at every rational number and continuous at every irrational number.

Thomae's function has many other names - it is also known the *modified Dirichlet function*, the *Riemann function*, or under more whimsical names such as the *popcorn function*, *raindrop function*, *countable cloud function*, or the *Stars over Babylon* (due to John Horton Conway, one of the coolest mathematicians of all time).

Recall that we call a set F_σ if it is a countable union of closed sets, and that we call a set G_δ if it is a countable intersetion of open sets.

Theorem 3.3.10.3: A function discontinuous at an arbitrary F_σ-set

Let $F = \bigcup_{n\in\mathbb{N}} F_n$ be a countable union of closed sets F_n. For any point $x \in F$, let $n(x)$ be the smallest natural number such that $x \in F_{n(x)}$. Then the function $f_F : \mathbb{R} \to \mathbb{R}$ defined by

$$f_F(x) = \begin{cases} \frac{1}{n(x)} & x \in F, x \in \mathbb{Q} \\ -\frac{1}{n(x)} & x \in F, x \notin \mathbb{Q} \\ 0 & x \notin F \end{cases}$$

is continuous at every $x \in X \notin F$ and discontinuous at every $x \in F$

Corollary 3.3.11. (Functions continuous at an arbitrary G_δ-set): *Since the complement of a G_δ-set is F_σ, we can use the same construction to construct a function that is continuous at an arbitrary G_δ-subset of $\mathbb{R}$.*

Proposition 3.3.12. *Let f be a function between complete metric spaces. Then the set of continuities of f is G_δ and the set of discontinuities of f is F_σ.*

Corollary 3.3.13. *There is no function $f : \mathbb{R} \to \mathbb{R}$ that is only continuous at the rationals.*

Proof. The irrationals are uncountable and the rationals are dense in the reals. Any countable union of closed sets either only contains singleton sets, in which case it is countable, or contains at least one non-singleton interval, in which case it contains rational numbers. Therefore the irrationals are not F_σ and the rationals are not G_δ. □

Chapter 4

Topological Fields

Theorem 4.0.1. *Let F be an ordered field. Then F becomes a topological field if we give it the order topology.*

Theorem 4.0.2. *Limits and field operations*

Corollary 4.0.3. *Let a_n be a sequence in an ordered field F. Let z_n be a zero sequence in F, and let $a \in F$. Then if we have*

$$a_n \geq a - z_n$$

for infinitely many n, it follows that

$$\liminf_{n \to \infty} a_n \geq a$$

Chapter 5

Topological Vector Spaces

Chapter 6

Topological Manifolds

Lemma 6.0.1. *Let M be a topological space. Then the following are equivalent:*

1. *Every point in M has a neighborhood that is homeomorphic to an open subset of $\mathbb{R}^n$.*
2. *Every point in M has a neighborhood that is homeomorphic to an open ball in $\mathbb{R}^n$.*
3. *Every point in M has a neighborhood that is homeomorphic to $\mathbb{R}^n$.*

If M has this property, we call it ***locally Euclidean of dimension*** n*..*

Notably, a topological space M is locally euclidean of dimension 0 if and only if every open subset is homeomorphic to $\mathbb{R}^0 = \{0\}$, i.e. every open set contains a single point, i.e. M is discrete.

Definition 6.0.2. Let M be locally euclidean of dimension n. Let $U \subset M$ be open. Then:

1. We call U a **coordinate domain**,
2. We call any homeomorphism $\varphi : U \to V$ to an open subset $V \subseteq \mathbb{R}^n$ a **coordinate map**,
3. We call the pair (U, φ) is a **coordinate chart**, or just **chart**.

Definition 6.0.3. Let M be a topological space. Then we call M an n**-dimensional topological manifold** if it is:

1. Hausdorff, and
2. second-countable, and
3. locally Euclidean of dimension n.

Some authors omit the latter two conditions, but virtually all important examples of locally euclidean topological spaces do fulfill these properties, and most interesting theorems about topological manifolds require them, so not much is gained by working with a more general definition.

Theorem 6.0.4. *Every open subset of an n-dimensional topological manifold is itself an n-dimensional topological manifold.*

Theorem 6.0.5. *A topological space is a 0-manifold if and only if it is a countable discrete space.*

The two following theorems are of fundamental importance, but the proofs sadly require additional machinery that we will not establish here:

Proposition 6.0.6. *If $m \neq n$, then a nonempty topological space cannot be both an m-manifold and an n-manifold.*

Note that the empty set is explicitly excluded, since it does in fact qualify as a manifold of any arbitrary dimension.

Proposition 6.0.7. *Every topological n-manifold is homeomorphic to a subset of a Euclidean space $\mathbb{R}^k$, where $k \geq n$.*

Corollary 6.0.8. *Every topological manifold is separable and metrizable.*

Chapter 7

Uniform Spaces

Many theorems in analysis require a notion of *uniform convergence*, *uniform continuity*, and so on. These ideas can be easily expressed in a metric space - recall that, for example, a function $f : M \to N$ between metric spaces is uniformly continuous if there exists a $\delta > 0$ such that for every $\varepsilon > 0$, we have that if $d_M(x, y) < \delta$, then $d_N(f(x), f(y)) < \varepsilon$.

Meanwhile, we wouldn't be able to refine the definition of continuity like this in a topological space, since the general structure of the neighborhoods of a topological space might vary wildly at different locations in the space - the important quality of a metric space here is that the notion of distance in a metric space can be applied "uniformly" to pairs of points, no matter where they are located. In this section, we want to define a set of spaces more general than metric spaces, but less general than topological spaces, which shares this important property of "uniformity", which will allow us to generalize many useful properties of metric spaces.

7.1 Diagonal Uniformity

Definition 7.1.1. For any set X, we denote by $\Delta(X)$ the diagonal $\{(x, x) \mid x \in X\}$ in $X \times X$.

Our first definition of a *uniform structure* on a set X is based on the observation that in a metric space, x and y are close together if and only if (x, y) is close to $\Delta(X)$.

Definition 7.1.2. For any pair of subsets U, V of $X \times X$ (which by definition can be viewed as relations on X), we can extend the notion of function composition to these arbitrary relations by defining $U \circ V$ to be the set

$$\{(x, y) \in X \times X \mid \exists z \in X : ((x, z) \in V, (z, y) \in U\}$$

Definition 7.1.3. A **diagonal uniformity** on a set X is a collection $\mathcal{D}(X)$ of subsets of $X \times X$, called **surroundings**, such that:

1. If $D \in \mathcal{D}$, then $\Delta(X) \subset D$,

2. If $D_1, D_2 \in \mathcal{D}$, then $D_1 \cap D_2 \in \mathcal{D}$,

3. If $D \in \mathcal{D}$, then there exists an $E \in \mathcal{D}$ such that $E \circ E \subset D$,

4. If $D \in \mathcal{D}$, then there exists an $E \in \mathcal{D}$ such that $E^{-1} \subset D$

5. If $D \in \mathcal{D}$ and $D \subset E$, then $E \in \mathcal{D}$.

We call a set X equipped with such a structure a **uniform space**.

Example 7.1.4. For any metric space (M, d), the metric d generates a *metric uniformity* by having a surrounding

$$D_\varepsilon^d = \{(x, y) \in M \times M \mid d(x, y) < \varepsilon\}$$

for every $\varepsilon > 0$. Uniformities that can be generated in this way from metrics are called **metrizable**.

Comment 7.1.5. For an arbitrary metric d, the uniformity generated by d is identical to the one generated by a scaled version λd (with $\lambda \in \mathbb{R}^\times$). Therefore different metrics may generate the same uniformity.

Part IV

Differentiation

Chapter 8

Differentiation in Normed Vector Spaces

8.1 The Fréchet Derivative

Definition 8.1.1. (Fréchet Derivative): Let $(V, \|-\|_V, (W, \|-\|_W)$ be normed vector spaces. Let $x \in U \subset V$. Then a map $f : U \to W$ is called **Fréchet differentiable at** x_0, **totally differentiable at** x_0, or just **differentiable at** x_0, if there exists a bounded linear map $A : V \to W$ such that

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - A(h)\|_W}{\|h\|_V} = 0$$

f is called (Fréchet / totally) differentiable if it is differentiable at every point.

Theorem 8.1.2. *If such an A exists, it is unique. We call it the **Fréchet derivative**, **differential**, or **derivative**, of f at x, and denote it as:*

$$Df(x) := A$$

Some comments:

1. In the case $f : \mathbb{R} \to \mathbb{R}$, the linear maps $\mathbb{R} \to \mathbb{R}$ are exactly the maps $x \mapsto cx$, with c constant. Therefore if f is a function $\mathbb{R} \to \mathbb{R}$, then assuming the standard absolute value norm on $\mathbb{R}$, this expression can be rearranged to give us the classic definition of a derivative:

$$\begin{aligned}
&\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - A(h)\|_{\mathbb{R}}}{\|h\|_{\mathbb{R}}} = 0 \\
\iff &\lim_{h \to 0} \frac{|f(x_0 + h) - f(x_0) - c \cdot h|}{|h|} = 0 \\
\iff &\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - c \cdot h}{h} = 0 \\
\iff &\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} - c = 0 \\
\iff &\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = c
\end{aligned}$$

   In this case, we generally write $c := f'(x_0)$.

However, note that under our general definition of a derivative, the derivative is a *map*, meaning that the derivative $Df(x)$ of $f : \mathbb{R} \to \mathbb{R}$ at x" is *technically* not the scalar $f'(x) \in \mathbb{R}$, but instead the linear map

$$\begin{aligned} Df(x) : \mathbb{R} &\to \mathbb{R} \\ t &\mapsto f'(x) \cdot t \end{aligned}$$

2. The definition demands that A be a *bounded* linear map. However, recall that if V is finite-dimensional, every linear map from V is inherently bounded, so this additional constraint is only relevant if V is infinite-dimensional.

3. If V and W are finite-dimensional vector spaces over the same field $\mathbb{F}$, A is a matrix, and we can write $A \cdot h$ instead of $A(h)$.

Definition 8.1.3. We call $f : U \to W$ **continuously differentiable** if the function

$$\begin{aligned} Df : U &\to \hom(V, W) \\ x &\mapsto Df(x) \end{aligned}$$

is continuous. We denote the set of continuously differentiable functions $U \to W$ as $C^1(U, W)$.

Proposition 8.1.4.

1. *Every constant map is totally differentiable with total derivative* 0.

2. *Every bounded linear map F is totally differentiable with total derivative* F.

Theorem 8.1.5. (Differential of Multiplication): *The multiplication operator*

$$\begin{aligned} M : \mathbb{F}^2 &\to \mathbb{F} \\ \vec{x} &\mapsto x_1 \cdot_{\mathbb{F}} x_2 \end{aligned}$$

is differentiable, with derivative:

$$DM(\vec{x}) = (x_2, x_1)$$

Proof. We have:

$$\begin{aligned} &M(\vec{x} + \vec{h}) - M(\vec{x}) - DM(\vec{h}) \\ =&(x_1 + h_1)(x_2 + h_2) - x_1 x_2 - h_1 x_2 - h_2 x_1 \\ =&x_1 x_2 + x_1 h_2 + h_1 x_2 + h_1 h_2 - x_1 x_2 - h_1 x_2 - h_2 x_1 \\ =&h_1 \cdot h_2 \end{aligned}$$

Since norms on finite dimensional vector spaces are equivalent, we can assume the Maximum norm on $\mathbb{F}^2$, and get:

$$\begin{aligned}
&\lim_{\vec{h}\to\vec{0}} \frac{\|h_1 \cdot h_2\|_{\mathbb{F}}}{\left\|\vec{h}\right\|_{\max}} \\
&= \lim_{\vec{h}\to\vec{0}} \frac{\|h_1 \cdot h_2\|_{\mathbb{F}}}{\max\{|h_1|, |h_2|\}} \\
&= \lim_{\vec{h}\to\vec{0}} \frac{\|h_1 \cdot h_2\|_{\mathbb{F}}}{\max\{|h_1|, |h_2|\}} \\
&= \lim_{\vec{h}\to\vec{0}} \frac{|h_1|_{\mathbb{F}}|h_2|_{\mathbb{F}}}{\max\{|h_1|, |h_2|\}} \\
&= \lim_{\vec{h}\to\vec{0}} |\min\{h_1, h_2\}|_{\mathbb{F}} \\
&= 0
\end{aligned}$$

□

Proposition 8.1.6. (Linearity of the differential operator): *Let V, W be normed vector spaces over a field $\mathbb{F}$, and let $F, G : V \supset U \to W$ be totally differentiable at $\vec{x} \in U$. Let $c \in \mathbb{F}$. Then:*

1. $D(cF)(\vec{x}) = c \cdot (DF(\vec{x}))$
2. $D(F + G)(\vec{x}) = DF(\vec{x}) + DG(\vec{x})$

8.2 Divergence and Curl

Definition 8.2.1. (Divergence of a Vector Field): Let

$$\begin{aligned}
f : \mathbb{R}^m \supset S &\to \mathbb{R}^n \\
x &\mapsto (f_1(x), \ldots, f_n(x))
\end{aligned}$$

be continuously differentiable. Then the **divergence** of f is defined to be

$$\operatorname{div} f = \operatorname{tr} Df$$

Comment 8.2.2. If we assume the standard base on $\mathbb{R}^n$, we have:

$$\operatorname{div} f = \operatorname{tr}(Df) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f_i(x)$$

$\operatorname{div} f$ is sometimes written as $\nabla \cdot f$, which is horrendous notation that I will stay far away from.

$\operatorname{div} f$ has a nice physical interpretation: Imagine f as a physical vector field describing the flow of a fluid. Take a neighborhood around a point x, measure the amount of fluid flowing out of that neighborhood, and subtract the amount of fluid flowing into that neighborhood. Then $(\operatorname{div} f)(x)$ is the limiting value of this operation as we let our neighborhoods converge to the point x itself.

This means that x is a *source* iff $(\operatorname{div} f)(x) > 0$, and a *sink* iff $(\operatorname{div} f)(x) < 0$.

Chapter 9

Differentiable Manifolds

9.1 Differentiable Manifolds

Definition 9.1.1. We say a map is "C^n" if its first n derivatives exist and are continuous. If $f \in C^n$ is bijective such that $f^{-1} \in C^n$, then we call f a C^n-**diffeomorphism**.

Under the convention that the 0th derivative of f is f itself, a C^0-diffeomorphism is the same thing as a homeomorphism.

9.2 Inverse Function Theorem

Theorem 9.2.0.1: Inverse Function Theorem

Let X, Y be finite-dimensional real affine spaces, let $U \subset X$ be open and let $f : U \to Y$ be C^n. Then if the differential $Df(p)$ at a point $p \in U$ is invertible, There exists an open set V with $p \in V \subset U$ such that $f|_V$ is a C^n-diffeomorphism.

9.3 Implicit Function Theorem

Part V

Measure and Integration

Chapter 10

The Riemann Integral

Definition 10.0.0.1: Riemann integrability

Let $f : [a, b] \to \mathbb{R}$. Then we call f **Riemann integrable** if it is bounded and its upper and lower Darboux integrals are equal.

Theorem 10.0.1. *Let $f : [a, b] \to \mathbb{R}$. Then f is Riemann integrable if and only if, for every $x \in [a, b]$, the upper and lower limits*

$$\lim_{x \nearrow c} f(x), \qquad \lim_{x \searrow c} f(x)$$

exist.

Note that we do not require the limits to coincide at any given point.

Theorem 10.0.2. *Let $f : [a, b] \to \mathbb{R}$. Then f is Riemann integrable if and only if the set of discontinuities of f can be covered by a countable union of intervals of arbitrarily small length.*

10.1 The Gauge Integral

Chapter 11

Measure Theory

11.1 The Measure Problem

The most basic goal of measure theory is to establish a generalized notion of a "measure function", which assigns a "volume" to a given set. In particular, we would like to establish a function that assigns volume functions to subsets of $\mathbb{R}^n$ and has the following three properties:

1. When given a subset with an easily intuitively definable volume, the volume function should agree with that volume. In particular, the volume of a cuboid should be the product of the lengths of its sides, and the length of a real interval (a, b) should be $b - a$.

2. The volume of a countable disjoint union of sets should be the sum of the individual volumes. This property is generally referred to as *σ-additivity*.

3. The volume should be invariant under isometries, i.e. functions like rotations, translations, and reflections should not change the volume of a set.

We call a σ-additive function a *measure*. It will turn out that there exists exactly one function on $\mathbb{R}^n$, called the *Lebesque measure λ^n*, that fulfills these conditions for a very large family of sets (the so-called "Borel σ-Algebra") - enough to include every "somewhat reasonable" subset of $\mathbb{R}^n$. However, there are still counterexamples.

Proposition 11.1.1. *Every subset of $\mathbb{R}^n$ being Lebesque-measurable is consistent with ZF (without the axiom of choice).*

Theorem 11.1.2. *Assuming the axiom of choice, there exist subsets of $\mathbb{R}^n$ that cannot be assigned a volume without arriving at a contradiction.*

It turns out that this result crucially relies on the full axiom of choice, and in particular is not implied by commonly used weaker forms of the axiom of choice such as the axiom of dependent choice.

The following two subsections deal with two different ways of proving this theorem by construction *non-measurable sets*: The *Vitali Sets*, and the decomposition of a sphere given in the *Banach-Tarski-Paradox*.

11.1.1 Vitali Sets

Proposition 11.1.3. *The relation $x \sim y \Leftrightarrow x - y \in \mathbb{Q}$ is an equivalence relation on the real numbers.*

Theorem 11.1.4. *There exist sets $V \subset [0,1]$ such that for each $r \in \mathbb{R}$, there exists exactly one number $v \in V$ such that $v - r$ is rational. We call such a set a **Vitali Set**.*

> *Proof.* Consider the aforementioned equivalence relation $x \sim y$ on $\mathbb{R}$. Each equivalence class must contain at least one representative also contained in $[0,1]$, since if $x - y = q \in \mathbb{Q}$, we have $x - (y + q) \in \mathbb{Q}$ for all $q \in \mathbb{Q}$, letting us pick a q such that $x - (y + q) \in [0,1]$. Being equivalence classes, they must also be disjoint.
>
> This means we can use the axiom of choice to pick exactly one element of each equivalence class of $\sim$, giving us our set V. □

Lemma 11.1.5. *Let $q_1, q_2, \dots$ be an enumeration of $\mathbb{Q} \cap [-1,1]$. Then for $j \neq k$, we have*

$$(q_j + V) \cap (q_k + V) = \emptyset$$

> *Proof.* Assume the intersection is non-empty. Then there exist $v_1, v_2 \in V$ such that $q_j + v_1 = q_k + v_2$, meaning $v_1 - v_2 \in \mathbb{Q}$. Therefore, v_1 must be in the same equivalence class as v_2. Since V contains exactly one element of each equivalence class, we have $v_1 = v_2$, and therefore we also have $q_j = q_k$, i.e. $j = k$. □

Lemma 11.1.6. *We have*

$$[0,1] \subset \bigcup_{k \in \mathbb{N}} (q_k + V) \subset [-1,2]$$

> *Proof.*
> 1. $\bigcup_{k \in \mathbb{N}} (q_k + V) \subset [-1,2]$ follows trivially from $q_k \in [-1,1]$ and $V \subset [0,1]$.
> 2. $[0,1] \subset \bigcup_{k \in \mathbb{N}} (q_k + V)$ follows from the definition of V, since for every $y \in [0,1]$ we have a unique $v \in V$ such that $y - v := q \in \mathbb{Q}$, and since $y \in [0,1]$ and $v \in [0,1]$, we have $q \in [-1,1]$, i.e. q is contained in our enumeration.
>
> □

Corollary 11.1.7. *Vitali sets are not measurable by any translation-invariant measure.*

Proof. Assume that λ^1 is our desired Lebesque measure on $\mathbb{R}$, which is invariant under isometries and countably additive. Then we have

$$1 = \lambda^1([0,1]) \leq \lambda^1\left(\bigcup_{k\in\mathbb{N}}(q_k + V)\right) \leq \lambda^1([-1,2]) = 3$$

Since our measure is countably addiive and invariant under isometries, we can translate each individual set and preserve the measure:

$$\begin{aligned}\lambda^1\left(\bigcup_{k\in\mathbb{N}}(q_k + V)\right) &= \sum_{k\in\mathbb{N}}\lambda^1(q_k + V)\\ &= \sum_{k\in\mathbb{N}}\lambda^1(V)\end{aligned}$$

Now, if $\lambda^1(V) \leq 0$, we wouldn't have $1 \leq \sum_{k\in\mathbb{N}}\lambda^1(V)$, and if $\lambda^1(V) > 0$, we wouldn't have $\sum_{k\in\mathbb{N}}\lambda^1(V) \leq 3$. Therefore every possible measure we could assign to V leads to a contradiction. □

11.1.2 The Banach-Tarski Paradox

Theorem 11.1.8. *Given any two sets $A, B \subset \mathbb{R}^n$, with $n \geq 3$, such that both A and B have a nonempty interior, there exist disjoint decompositions $A_1 \sqcup \ldots \sqcup A_k = A$ and $B_1 \sqcup \ldots \sqcup B_k = B$ such that for each i, A_i and B_i can be transformed into each other by an isometry.*

Corollary 11.1.9. *The unit sphere can be transformed into two copies of the unit sphere by a finite disjoint decomposition followed by an isometry. The subsets of the decomposition therefore violate the countable union condition we expect from measure functions, and can therefore not be a assigned a meaningful volume.*

11.2 σ-Additivity

Theorem 11.2.1. *Let X be a set and $\mathcal{A} \subseteq \mathcal{P}(X)$. Then the σ-additivity of a function $\mu : \mathcal{A} \to [0,\infty]$ is only well defined if $\mathcal{A}$ contains the empty set and is closed under countable union.*

Corollary 11.2.2. ($\emptyset$ is always a zero set): *Let μ be a σ-additive function $\mathcal{A} \to [0,\infty]$. Then $\mu(\emptyset) = 0$.*

Corollary 11.2.3. (σ-Additivity implies monotonicity): *Let μ be a σ-additive function $\mathcal{A} \to [0,\infty]$. Then $A \subset B$ implies $\mu(A) \leq \mu(B)$.*

11.3 Lattices and Boolean Algebras

We have just seen that we cannot define our desired measure function on the full power set of $\mathbb{R}^n$. This means we will have to work on smaller systems of subsets of a given set. Naturally, we want to find the largest such systems that are still well-behaved enough to allow us to define a sensible notion of measure.

We will arrive at different algebraic structures on subsets of power sets, which will serve as the domains of our measure functions.

In order to gain a full birds-eye view of these definitions, we will first introduce the more general notion of *boolean algebras*:

11.3.1 Boolean Algebras

Definition 11.3.1. A **boolean algebra** is a set X, equipped with two binary operations $\wedge$ and $\vee$, a unary operation $\neg$, and two elements 0 and 1, such that:

1. $\wedge$ and $\vee$ are commutative,
2. 1 is a neutral element of $\wedge$, and 0 is a neutral element of $\vee$,
3. $\wedge$ distributes over $\vee$ and $\vee$ distributes over $\wedge$,
4. $x \wedge \neg x = 0$, and $x \vee \neg x = 1$.

Corollary 11.3.2. *Any boolean algebra also has the following properties:*

1. $\wedge$ *and* $\vee$ *are associative,*
2. $\wedge$ *and* $\vee$ *have the following* ***absorption property****:*

$$a \wedge (a \vee b) = a$$
$$a \vee (a \wedge b) = a,$$

3. $a = b \wedge a$ *if and only if* $a \vee b = b$.

There are three "central" boolean algebras, from which most of the terminology describing them is descended:

Theorem 11.3.3. *The set of* ***propositional formulas*** *forms a boolean algebra, where* 0 *is the logical falsum (*$\perp$*, an unfulfillable formula),* 1 *is the logical verum (*$\top$*, a tautological formula), and* $\wedge$*,* $\vee$ *and* $\neg$ *are logical "and", "or" and "not".*

In computer science and circuit engineering, one often considers the subalgebra of this boolean algebra where every formula is directly evaluated to "0" or "1".

Theorem 11.3.4. (Power set algebra): *The* ***power set*** $\mathcal{P}(X)$ *of any set* X *forms a boolean algebra, where* $0 = \emptyset$*,* $1 = X$*,* $\wedge$ *is the set intersection operation* $\cap$*,* $\vee$ *is the set union operation* $\cup$*, and* $\neg$ *is the set complement operation* $M \to X \setminus M$*.*

Theorem 11.3.5. (Restrictions of Boolean algebras): *Let* $\mathcal{B}$ *be a Boolean algebra on a set* X*, and let* Y *be a subset of* X*. Then the* ***restriction of*** $\mathcal{B}$ ***to*** Y*, defined as*

$$\mathcal{B}|_Y := \{E \cap Y \mid E \in \mathcal{B}\},$$

is a boolean Algebra on Y*.*

Theorem 11.3.6. *If* $Y \in \mathcal{B}$*, then*

$$\mathcal{B}|_Y = \mathcal{B} \cap \mathcal{P}(Y) = \{E \subset Y \mid E \in \mathcal{B}\}$$

Theorem 11.3.7. (Atomic algebra): *Let X be partioned into a union*

$$X = \bigcup_{\alpha \in I} A_\alpha$$

of disjoint sets A_α, which we refer to as atoms. Then this partition forms a Boolean algebra

$$\mathcal{A}((A_\alpha)_{\alpha\in I}) := \left\{ E \;\middle|\; E = \bigcup_{\alpha \in J} A_\alpha, J \subset I \right\}$$

of all the sets that can be represented as a union of atoms.

The power set Algebra on X is exactly the atomic algebra where X is partitioned into singleton atoms.

Theorem 11.3.8. *Atomic algebras are uniquely determined by their atoms, up to relabeling. More precisely: Let $(A_\alpha)_{\alpha\in I}$ and $(B_\beta)_{\beta\in J}$ be two partitions of a set X. Then*

$$\mathcal{A}((A_\alpha)_{\alpha\in I}) = \mathcal{A}((B_\beta)_{\beta\in J})$$

if and only if there exists a bijection $\varphi : I \longrightarrow J$ such that $B_{\varphi(\alpha)} = A_\alpha$ for all $\alpha \in I$.

Theorem 11.3.9. *Every finite Boolean algebra is an atomic algebra.*

Corollary 11.3.10. *Every finite Boolean algebra has cardinality 2^n, where $n \in \mathbb{N}$.*

Corollary 11.3.11. *There is a one-to-one correspondence, up to relabeling, between finite Boolean algebras on a set X and finite partitions of X into non-empty sets.*

Theorem 11.3.12. (Dyadic algebras): *Let $n, i_1, \ldots, i_d \in \mathbb{Z}$. The **dyadic algeba** $\mathcal{D}_n(\mathbb{R}^d)$ at scale 2^{-n} in $\mathbb{R}^d$ is the atomic algebra generated by the products of the half-open dyadic intervals*

$$I_j := \left[\frac{i_j}{2^n}, \frac{i_j + 1}{2^n} \right)$$

of length 2^{-n}.

This algebra consists exactly of the "grid figures" made up of a finite number of "pixels" of length 2^{-n}.

Theorem 11.3.13. (Intersection of Boolean Algebras): *The intersection of a family $(\mathcal{B}_\alpha)_{\alpha\in I}$ of Boolean algebras on a set X is again a Boolean algebra, assuming the convention that, if I is empty, the intersection is the full power set. Furthermore, this Intersection is the finest Boolean algebra that is coarser than every $\mathcal{B}_\alpha$.*

Definition 11.3.14. Let $\mathcal{F}$ be any family of subsets of a set X. Then we define $\langle \mathcal{F} \rangle_{\mathcal{B}}$ to be the intersection of all Boolean algebras that contain $\mathcal{F}$. We call this the **Boolean algebra generated by** $\mathcal{F}$.

Equivalently, $\langle \mathcal{F} \rangle_{\mathcal{B}}$ is the smallest Boolean algebra containing $\mathcal{F}$.

Theorem 11.3.15. *$\mathcal{F}$ is a Boolean algebra if and only if $\langle \mathcal{F} \rangle_{\mathcal{B}} = \mathcal{F}$.*

11.3.2 Lattices

Boolean algebras themselves turn out to be specific instances of *lattices*, which play an important role in order theory and universal algebra.

Definition 11.3.16. A **lattice** is an algebraic structure $(L, \vee, \wedge)$, consisting of a set L, an operation $\vee$, called **join**, and an operation $\wedge$, called **meet**, such that the absorbtion laws $a \vee (a \wedge b) = a$ and $a \wedge (a \vee b) = a$.

Proposition 11.3.17. *Equivalently, a partially ordered set* $(L, \leq)$ *is a lattice if every pair of elements has a least upper bound* $\sup(a, b) := a \vee b \in L$ *and a greatest lower bound* $\inf(a, b) := a \wedge b \in L$.

Definition 11.3.18. We call a lattice **bounded** if there exists a **least element** 0, i.e. 0 fulfills $a \vee 0 = a$, and a **greatest element** 1, which fulfills $a \wedge 1 = a$.

Corollary 11.3.19. *A boolean algebra is a bounded lattice such that meet and join are distributive over each other and such that complements exist.*

11.4 Set Algebras

Definition 11.4.1. Let X be a set and $\mathcal{A} \subset \mathcal{P}(X)$. Then we call $\mathcal{A}$ a **set algebra** if it has the following properties:

1. $\emptyset \in \mathcal{A}$
2. For any $A \in \mathcal{A}$, we have $X \setminus A \in \mathcal{A}$ ($\mathcal{A}$ is closed under the operation of taking complements).
3. For any $F, G \in \mathcal{A}$, we have $F \cup G$ in $\mathcal{A}$ ($\mathcal{A}$ is closed under binary unions).

Corollary 11.4.2. *If* $\mathcal{A}$ *is a set algebra on* X*, it also fulfills the following:*

1. $X \in \mathcal{A}$,
2. *For any* $F, G \in \mathcal{A}$*, we have* $F \cap G$ *in* $\mathcal{A}$,
3. *For any* $A_1, \ldots, A_n \in \mathcal{A}$*, we have* $\bigcup_{i=1}^n A_i \in \mathcal{A}$,
4. *For any* $A_1, \ldots, A_n \in \mathcal{A}$*, we have* $\bigcap_{i=1}^n A_i \in \mathcal{A}$.

Thus, we obtain the following more concise (but less readable) definition of set algebras:

Corollary 11.4.3. *A set algebra is a subalgebra of the power set boolean algebra on* X.

Corollary 11.4.4. *A topology on* X *is "simply" a set algebra on* X *that is closed under arbitrary unions.*

Theorem 11.4.5. ((Stone's Representation Theorem for Boolean Algebras:)): *Every boolean algebra is isomorphic to a set algebra.*

11.4.1 Set Rings

A very important weakening of the concept of set algebras is given by *set rings*, which contain the empty set and are closed under intersection and union, but don't have to contain the full set X or be closed under complements.

Theorem 11.4.6. *Let $\mathcal{A}$ be a set ring. Then it is closed under finite symmetric difference, and forms a ring in the algebraic sense, with symmetric difference as addition and intersection as multiplication. If it contains the full set X, it forms a ring with identity.*

Definition 11.4.7. Let $I \subset \mathbb{R}$. We call I an **interval** if there exist $a, b \in \mathbb{R}$ such that $(a, b) \subset I \subset [a, b]$.

Theorem 11.4.8. *The set of subsets of the real numbers which can be written as a finite union of intervals forms a set ring.*

Theorem 11.4.9. *Let $\mathcal{R}$ and $\mathcal{S}$ be set rings. Then the set of finite unions of cartesian products of elements of R_i and S_i, i.e. of elements of the form,*

$$\bigcup_{i \in \mathbb{N}} R_i \times S_i,$$

is also a set ring, which we will denote by $\mathcal{R} \boxtimes \mathcal{S}$.

Corollary 11.4.10. *The set of finite unions of cuboids in $\mathbb{R}^n$, where we define a cuboid to be a product of arbitrary intervals, i.e. we don't care if the boundary on any particular side is open or closed, forms a set ring. We call sets of this form **Elementary Sets**.*

11.4.2 Set Semirings

Theorem 11.4.11. *Let $\mathcal{S} \subset \mathcal{P}(X)$. We call $\mathcal{S}$ a **set semiring**, or **semiring of sets**, if:*

1. *$\emptyset \in \mathcal{S}$,*
2. *$\mathcal{S}$ is closed under finite intersections,*
3. *For $A, B \in \mathcal{S}$, there exist disjoint sets $S_1, \ldots, S_n \in \mathcal{S}$ such that $A \setminus B = \bigcup_{i=i}^{n} S_i$.*

This means that a set semiring is a weakened form of a set ring where complements are not necessarily contained in the semiring, but can still be "constructed" from elements of the ring. Any set ring is therefore immediately also a set semiring.

Sadly, unlike with rings of sets, there is absolutely no connection between set semirings and the algebraic notion of a semiring - a semiring of sets is exclusively a (semi)(ring of sets), and *not* a (semiring)(of sets). This makes it tempting for me to use an alternative name which makes this distinction more clear, but since I haven't encountered any good alternative names anywhere else (and because I already know I will forget to stick with this convention moving forwards) I will stick with the less than perfect established name.

Set semirings are of fundamental importance to measure theory because the set of cuboids in $\mathbb{R}^n$ forms a set semiring, and we will end up defining our lebesque measure by approximating sets through coverings of the set with cuboids. Of course, we first have to establish a basic theory of set semirings and prove this claim.

Theorem 11.4.12. *The set $\mathcal{I}$ of real intervals forms a set semiring.*

Theorem 11.4.13. *The product of two set semirings is again a set semiring.*

Corollary 11.4.14. *The set $\mathcal{Q}$ of cuboids in $\mathbb{R}^n$ (once again with both open and closed sides allowed) forms a set semiring.*

11.5 Measures on $\mathbb{R}^n$

(TODO: Merge into abstract measure theory chapter)

11.5.1 Jordan Measure

Definition 11.5.1. Let $E \subset \mathbb{R}^n$ be an elementary set. Then we can assign to it the **elementary volume** $\operatorname{vol}(E)$, where the volume of a cuboid is the product of its side lengths and the volume of a finite union of cuboids is the sum of the volumes of the individual cuboids making up E.

Definition 11.5.2. (Inner and outer Jordan Measures): Let $E \subset \mathbb{R}^n$.

1. The **inner Jordan measure** $J_*(E)$ is

$$J_*(E) := \sup_{\substack{Q_i \in \mathcal{Q}, \\ \bigcup_{i=1}^n Q_i \subset E}} \operatorname{vol}(Q)$$

2. The **outer Jordan measure** $J_*(E)$ is

$$J^*(E) := \inf_{\substack{Q_i \in \mathcal{Q}, \\ \bigcup_{i=1}^n Q_i \supset E}} \operatorname{vol}(Q)$$

Definition 11.5.3. (Jordan measurable set, Jordan measure): We call E **Jordan-measurable** if $J_*(E) = J^*(E)$. Then we call $J(E) = J_*(E) = J^*(E)$ the **Jordan measure**.

Theorem 11.5.4. *The Jordan measure is σ-additive.*

Theorem 11.5.5. *The following are equivalent:*

1. *E is Jordan measurable,*
2. *For every $\varepsilon > 0$, there exist elementary sets $A \subset E \subset B$ such that $\operatorname{vol}(B \setminus A) \leq \varepsilon$,*
3. *For every $\varepsilon > 0$, there exists an elementary set A such that $J^*(A \triangle E) \leq \varepsilon$.*

Theorem 11.5.6. *The collection of subsets of $\mathbb{R}^n$ that are either Jordan measurable or have a Jordan-measurable complement form a Boolean algebra, known as the Jordan algebra.*

Theorem 11.5.7. *The Jordan algebra is non-atomic.*

Theorem 11.5.8. ((Regions under continuous Graphs are Jordan measurable)): *Let B be a closed box in $\mathbb{R}^n$, and let $f : B \to \mathbb{R}$ be a continuous function. Then the set*

$$\{(x, t) \mid x \in B, 0 \leq t \leq f(x)\} \subset \mathbb{R}^{n+1}$$

is Jordan measurable.

Theorem 11.5.9. *Triangles are Jordan measurable.*

Theorem 11.5.10. *Convex polytopes in $\mathbb{R}^n$ are Jordan measurable.*

Theorem 11.5.11. *Open and closed Euclidean balls are Jordan measurable.*

Theorem 11.5.12. *Every subset of a Jordan null set is Jordan measurable and also a Jordan null set.*

Theorem 11.5.13. *The sets $[0,1]^2 \setminus \mathbb{Q}^2$ and $[0,1]^2 \cap \mathbb{Q}^2$ are both not Jordan measurable.*

Informally, sets with a lot of "holes" or with very messy, fractal-like boundaries are generally not Jordan-measurable.

Theorem 11.5.14. *There exist countable unions, and countable intersections, of Jordan measurable sets which are not Jordan measurable.*

11.5.2 Lebesque Measure

One can extend the Jordan measure to a significantly larger number of subsets of $\mathbb{R}^n$ by simply allowing countable unions of cuboids (instead of finite unions of cuboids). The Lebesque measure is simply this generalization of the Jordan measure. Formally:

Definition 11.5.15. Let $E \subset \mathbb{R}^n$. The **Lebesque outer measure** of E is given by

$$\lambda^*(E) := \inf_{\substack{E \subset \bigcup_{i=1}^{\infty} Q_n, \\ Q_n \in \mathcal{Q}}} \left\{ \sum_{n=1}^{\infty} |Q_n| \right\},$$

i.e. the Lebesque outer measure of E is the greatest lower bound of the measures of all coverings of E by cuboids.

Theorem 11.5.16. *Let $U \subset \mathbb{R}^n$ be open. Then U is a countable union of cuboids.*

Therefore, we can define Lebesque measurability similarly to Jordan measurability - a set is Lebesque measurable if it is "almost" an open set.

Definition 11.5.17. A set $E \subset \mathbb{R}^n$ is **Lebesque measurable** if for every $\varepsilon > 0$, there exists an open set $U \subset \mathbb{R}^n$ such that $U \subset E$ and $\lambda^*(U \setminus E) \leq \varepsilon$. If E is Lebesque measurable, we refer to $\lambda(E) := \lambda^*(E)$ as the **Lebesque measure** of E. If the dimension n should be emphasized, we sometimes write $\lambda(E)$ as $\lambda^n(E)$.

Theorem 11.5.18. *The Lebesque measure fulfills:*

$$E \subset \bigcup_{i \in N} E_i \implies \lambda(E) \leq \sum_{i \in \mathbb{N}} \lambda(E_i)$$

This is a weakened form of σ-additivity, which is known as σ-subadditivity.

Theorem 11.5.19. *The Lebesque measure defines a σ-additive function on the Lebesque measurable sets.*

Theorem 11.5.20. *The collection of subsets of $\mathbb{R}^n$ that are either Lebesque measurable or have a Lebesque-measurable complement form a Boolean algebra.*

You may have heard in linear algebra that the determinant of a matrix S tells us how the matrix scales the volume of the unit cube. Since the Lebesque measure is defined using volumes of unit cubes, this intuition also holds for the image of any set E under S:

Theorem 11.5.21. (Linear Transformation Equation): *Let $S \in \mathbb{R}^{n \times n}$. Then for all $E \subset \mathbb{R}^n$, we have:*

$$\lambda^n(S(E)) = |\det(S)|\lambda^n(E)$$

11.6 σ-Algebras

The most important type of set algebra for the purposes of measure theory is the *σ-Algebra*, on which we will define the notion of a "measure" in our desired final form.

Definition 11.6.1. Let X be an arbitrary set and $\mathcal{A} \subset \mathcal{P}(X)$. We call $\mathcal{A}$ a **σ-Algebra on** X if:

1. $X \in \mathcal{A}$
2. For all $A \in \mathcal{A}$, we have $X \setminus A \in \mathcal{A}$ ($\mathcal{A}$ is closed under the operation of taking a complement).
3. For all $A_i \in \mathcal{A}$, we have $\bigcup_{i \in \mathbb{N}} \in \mathcal{A}$ ($\mathcal{A}$ is closed under the operationg of taking countable unions).

If $\mathcal{A}$ is a σ-algebra on X, we call $(X, \mathcal{A})$ a **measure space**, and any set $A \in \mathcal{A}$ **$\mathcal{A}$-measurable**.

The "σ" here once again stands for "countable sum", as it also did for σ-additivity and F_σ sets.

Corollary 11.6.2. *Any σ-algebra contains the empty set and is closed under countable intersection.*

Corollary 11.6.3. *A σ-algebra can be more concisely defined as a set algebra that is closed under countable union and intersection, not just finite ones.*

This also trivially makes every σ-algebra a Boolean algebra.

Theorem 11.6.4. *Every atomic algebra is a σ-algebra.*

Corollary 11.6.5. *The set of Lebesque measurable sets contains the Borel σ-slgebra.*

Theorem 11.6.6. *The Lebesque algebra is a σ-algebra.*

Theorem 11.6.7. *The Jordan algebra is not a σ-algebra.*

Theorem 11.6.8. *Just like for Boolean algebras, the restriction $\mathcal{A}|_Y$ of a σ-algebra $\mathcal{A}$ on X to a subset $Y \subset X$ is again a σ-algebra on Y.*

Definition 11.6.9. Let $\mathcal{A} \subset \mathcal{P}(X)$. Then $\langle \mathcal{A} \rangle_\sigma$ denotes the smallest σ-algebra containing $\mathcal{A}$.

Theorem 11.6.10. *We have $\langle \mathcal{F} \rangle_{\mathcal{B}} = \langle \mathcal{F} \rangle_\sigma$ if and only if $\langle \mathcal{F} \rangle_{\mathcal{B}}$ is a σ-algebra.*

11.6.1 The Borel σ-Algebra

Definition 11.6.11. Let X be a topological space. The **Borel σ-algebra $\mathcal{B}[X]$ of** X is the σ-Algebra generated by the open subsets of X.

Theorem 11.6.12. *The Borel σ-Algebra $\mathcal{B}[\mathbb{R}^d]$ is equivalently generated by any of the following:*

1. *The closed subsets of* $\mathbb{R}^d$,
2. *The compact subsets of* $\mathbb{R}^d$,
3. *The open balls of* $\mathbb{R}^d$,
4. *The boxes in* $\mathbb{R}^d$,
5. *The elementary sets in* $\mathbb{R}^d$.

11.6.2 A Lebesque measurable set which is not Borel

Theorem 11.6.13. *There exist Lebesque measurable sets which are not Borel.*

11.7 Measures

With our different subset systems in place, we can finally give a general formal definition of a measure. Along the way, we will encounter outer measures, contents, and premeasures, which are weakened measures that we can use to generate proper ones.

Definition 11.7.1. Let X be a set and μ be a function $\mathcal{P}(X) \to [0, \infty]$. We call μ an **outer measure on** X if it is σ-subadditive.

Frustratingly, the Jordan outer measure J^* is not an outer measure in this sense - it only fulfills finite subadditivity, not the full σ-Subadditivity. However, the Lebesque outer measure λ^* thankfully is an outer measure in this sense (and historically, I assume this definition arose from a desire to generalize the Lebesque outer measure).

Definition 11.7.2. Let μ be an outer measure on a set X. Then we call a subset $A \subset X$ μ**-measurable**, or just **measurable**, if for all $S \subset X$ we have

$$\mu(S) = \mu(S \cap A) + \mu(S \setminus A)$$

The system of all μ-measurable sets is sometimes denoted $\mathcal{M}(\mu)$.

Note that by the subadditivity of outer measures, we already get that the left side is at most as large as the right side, meaning that this condition can equivalently be weakened to

$$\mu(S) \geq \mu(S \cap A) + \mu(S \setminus A).$$

Theorem 11.7.3. *Let* μ *be an outer measure. Then the set of* μ*-measurable sets forms a* σ*-Algebra.*

In order for our disjoint sum condition for "measures" to be well-defined, it already followed that the domain of a measure needs to be a system of sets that contains the empty set and is closed under countable union. This theorem suggests that, at minimum, it should also be a σ-Algebra. This finally leads us to a concrete definition of what a measure should be in general:

Definition 11.7.4. Let $\mathcal{A}$ be a <u>σ-Algebra</u>. Then we call a σ-additive function $\mu : \mathcal{A} \to [0, \infty]$ a **measure**.

Definition 11.7.5. Let $\mathcal{S}$ be a <u>set semiring.</u> Then we call a <u>finitely additive</u> function $\mathcal{S} \to [0, \infty]$ a **content**, and a <u>σ-additive</u> function $\mathcal{S} \to [0, \infty]$ a **premeasure**.

In effect, a premeasure is a measure whose domain might not be as big as it could be. Every measure trivially also defines a content and a premeasure, and every measure defined on $\mathcal{P}(X)$ defines an outer measure.

We can construct outer measures from a very large class of functions by mimicking the construction of the Lebesque outer measure from the elementary Volume function:

Theorem 11.7.6. (Carathéodory Extension): *Let $\mathcal{S}$ be a system of subsets of a set X containing the empty set. Let $\lambda : \mathcal{S} \to [0, \infty]$ be a function such that $\lambda(\emptyset) = 0$. Then the function*

$$\mu(E) := \inf \left\{ \sum_{i=1}^{\infty} \lambda(P_i) \,\middle|\, P_i \in \mathcal{S}, E \subset \bigcup_{i=1}^{\infty} P_i \right\}$$

is an outer measure on X.

Chapter 12

Measurable Functions

12.1 Measurable Functions

Definition 12.1.0.1: Measurable Functions

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. Then we call a map $f : X \to Y$ **$\mathcal{A}$-$\mathcal{B}$-measurable**, or just **measurable**, if the preimage of every measurable set is again measurable, i.e.

$$B \in \mathcal{B} \implies f^{-1}(B) \in \mathcal{A}$$

Once again, note the similarities between this definition and the topological definition of a continuous function. If the sigma-algebra on one of the sets is supposed to be clear from context, many authors only specify one of the the two sigma algebras. For example, for a function $f : X \to \overline{\mathbb{R}}$, many authors talk about $\mathcal{A}$-measurability when they implicitly mean $\mathcal{A} - \mathcal{B}(\overline{\mathbb{R}})$-measurability

Proposition 12.1.1. *The composition of two measurable functions is again measurable.*

Lemma 12.1.2. *Let $(X, \mathcal{A})$, $(Y, \mathcal{B})$ be measurable spaces and $f : X \to Y$. Let $\mathcal{E} \subset \mathcal{B}$. Then*

$$f^{-1}(\langle \mathcal{E} \rangle_\sigma) = \left\langle f^{-1}(\mathcal{E}) \right\rangle_\sigma$$

Corollary 12.1.3. *Let $\mathcal{E}$ be a base of $\mathcal{B}$, i.e. $\langle \mathcal{E} \rangle_\sigma = \mathcal{B}$. Then $f : X \to Y$ is $\mathcal{A}$-$\mathcal{B}$ measurable if and only if*

$$E \in \mathcal{E} \implies f^{-1}(E) \in \mathcal{A},$$

This means we don't need to check the preimage of every single set in $\mathcal{B}$ to show that f is measurable - it suffices to check a base.

Corollary 12.1.4. *Every continuous function between topological spaces is measurable in the corresponding Borel-σ-algebras.*

> *Proof.* Let $f : X \to Y$ be continuous. Then the preimage of every open set of Y is open in X, i.e. contained in the Borel σ-algebra on X, and the open sets of Y form a base of the Borel σ-algebra on Y. □

Theorem 12.1.4.1: Simple criteria for measurability)

Let $(X, \mathcal{A})$ be a measureable space. Let $f : X \to \overline{\mathbb{R}}$. Then the following are equivalent:

1. f is $\mathcal{A}$-$\mathcal{B}(\overline{\mathbb{R}})$ measurable,
2. $\forall c \in \mathbb{R} : \{f > c\} \in \mathcal{A}$,
3. $\forall c \in \mathbb{R} : \{f \geq c\} \in \mathcal{A}$,
4. $\forall c \in \mathbb{R} : \{f < c\} \in \mathcal{A}$,
5. $\forall c \in \mathbb{R} : \{f \leq c\} \in \mathcal{A}$,

Proof.

1. The latter four conditions are equivalent to each other, since:
   (a) $\{f \geq c\} = \bigcap_{k \in \mathbb{N}} \left\{f > c - \frac{1}{k}\right\}$,
   (b) $\{f > c\} = \bigcup_{k \in \mathbb{N}} \left\{f \geq c + \frac{1}{k}\right\}$,
   (c) $\{f < c\} = X \setminus \{f \geq c\}$,
   (d) $\{f \leq c\} = X \setminus \{f > c\}$.
2. The intervals $[c, \infty]$ form a base of $\mathcal{B}(\overline{\mathbb{R}})$, since:
   (a) $\{\infty\} = \bigcap_{k \in \mathbb{N}} [k, \infty]$
   (b) $\{-\infty\} = \bigcap_{k \in \mathbb{N}} [-\infty, -k]$
   (c) $(a, b) = [a, \infty] \setminus ([b, \infty] \cap [-\infty, a])$

□

Theorem 12.1.5. *Let $f, g : X \to \overline{\mathbb{R}}$ be $\mathcal{A}$-$\mathcal{B}(\overline{\mathbb{R}})$ measurable. Then the following sets are contained in $\mathcal{A}$:*

1. $\{f > g\}$,
2. $\{f \geq g\}$,
3. $\{f = g\}$,
4. $\{f \neq g\}$.

Theorem 12.1.6. *Let $f, g : X \to \overline{\mathbb{R}}$ be $\mathcal{A}$-$\mathcal{B}[\overline{\mathbb{R}}]$ measurable. Then the following functions are also $\mathcal{A}$-measurable:*

1. cf, *for all* $c \in \mathbb{R}$,
2. $|f|^p$, *for all* $p \in \mathbb{R}_{>0}$,
3. $f + g$, *assuming the sum is defined everywhere on* X, *i.e. there exists no* $x \in X$ *such that* $f(x) = \infty$ *and* $g(x) = -\infty$ *or vice versa,*
4. $f \cdot g$.

Theorem 12.1.7. *Let $(X, \mathcal{A})$ be a measurable space and $\mathbb{1}_E : X \to \mathbb{R}$ be the indicator function of a set $E \subset X$. Then $\mathbb{1}_E$ is $\mathcal{A}$-$\mathcal{B}[\mathbb{R}]$ measurable if and only if $E \in \mathcal{A}$.*

Proof. We have $\{1\} = (-\infty, 1) \cup (1, \infty) \in \mathcal{B}[\mathbb{R}]$ and $\mathbb{1}_E^{-1}(\{1\}) = E$, so E has to be in $\mathcal{A}$ for $\mathbb{1}_E$ to be measurable.

$E \in \mathcal{A}$ is also a sufficient condition for $\mathbb{1}_E$ to be measurable, since the only other possible preimages are $\emptyset \in \mathcal{A}$, $X \in \mathcal{A}$, $X \setminus E \in \mathcal{A}$ ◻

Theorem 12.1.8. *Let $(X, \mathcal{A})$ be a measurable space, let $D \in \mathcal{A}$, and let $f_k : D \to \mathbb{R}$ be $\mathcal{A}$ measurable. Then the following functions are $\mathcal{A}$-measurable:*

1. $\inf_{n\in\mathbb{N}} f_n$
2. $\sup_{n\in\mathbb{N}} f_n$
3. $\liminf_{n\to\infty} f_n$
4. $\limsup_{n\to\infty} f_n$

Proof. For $s \in \mathbb{R}$, we have:

1. $\left\{(\inf_{n\in\mathbb{N}} f_n) \geq s\right\} = \bigcap_{k=1}^{\infty} \{f_k \geq s\} \in \mathcal{A}$
2. $\left\{(\sup_{n\in\mathbb{N}} f_n) \leq s\right\} = \bigcap_{k=1}^{\infty} \{f_k \leq s\} \in \mathcal{A}$

Therefore $\inf_{n\in\mathbb{N}} f_n$ and $\sup_{n\in\mathbb{N}} f_n$ are measurable. The same argument holds for sup and inf over subsets of $\mathbb{N}$. Therefore, the compositions

$$\liminf_{n\to\infty} f_n = \sup_{n\in\mathbb{N}} \left(\inf_{k\geq n} f_k\right)$$

and

$$\limsup_{n\to\infty} f_n = \inf_{n\in\mathbb{N}} \left(\sup_{k\geq n} f_k\right)$$

are also measurable. ◻

Corollary 12.1.8.1

Let f_n be a sequence of $\mathcal{A}$-measurable functions with a pointwise limit f. Then f is $\mathcal{A}$-measurable.

Chapter 13

Lebesque Integration

13.1 Step Functions

Definition 13.1.1. Let $(X, \mathcal{A})$ be a measurable space. Then we call a function $f : Y \to \mathbb{R}$ a **step function** if it can be represented as a finite linear combination of indicator functions of sets in $\mathcal{A}$, i.e there exist $\alpha_i \in \mathbb{R}$ and $A_i \in \mathcal{A}$ such that:

$$f = \sum_{i \leq k} \alpha_i \cdot \mathbb{1}_{A_i}$$

Proposition 13.1.2. *Step functions form a vector space over $\mathbb{R}$.*

Proposition 13.1.3. *Step functions can only take finitely different values.*

Proposition 13.1.4. *Every step function is (extensionally) equal to a step function over pairwise disjoint sets.*

> *Proof.* Let
>
> $$s = \sum_{i \leq k} \alpha_i \cdot \mathbb{1}_{A_i}$$
>
> Then if $\beta_1, \ldots, \beta_m$ are all of the finitely many different values that s can take, s is by definition equal to
>
> $$t := \sum_{i \leq m} \beta_i \cdot \mathbb{1}_{\{s = \beta_i\}}$$
>
> And since a function cannot take two values at once, the sets $\{s = \beta_i\}$ are disjoint. □

Lemma 13.1.5. (Sum of Step Functions): *Let s_1 and s_2 be step functions*

$$s_1 = \sum_{i=0}^{m} \alpha_i \mathbb{1}_{A_i}, \qquad s_2 = \sum_{j=0}^{n} \beta_j \mathbb{1}_{B_j}$$

defined on pairwise disjoint sets. Then we have:

$$s_1 + s_2 = \sum_{i=0}^{m} \sum_{j=0}^{n} (\alpha_i + \beta_j) \mathbb{1}_{A_i \cap B_j}$$

Proof. 1. Assume s_1 and s_2 are nonnegative step functions

$$s_1 = \sum_{i=0}^{m} \alpha_i \cdot \mathbb{1}_{A_i}, \qquad s_2 = \sum_{j=0}^{n} \beta_j \cdot \mathbb{1}_{B_j}$$

defined on pairwise disjoint sets.

Then:

$$s_1(x) + s_2(x) = \sum_{i=0}^{m} \alpha_i \cdot \mathbb{1}_{A_i}(x) + \sum_{j=0}^{n} \beta_j \cdot \mathbb{1}_{B_j}(x)$$

Since the A_i and B_j are disjoint coverings of X, for every $x \in X$, there exists exactly one i such that $x \in A_i$ and exactly one j such that $x \in B_j$. Therefore, every x contributes exactly one α_i for $x \in A_i$ and one β_j for $x \in B_j$, i.e. $x \in A_i \cap B_j$ and $s_1(x) + s_2(x) = \alpha_i + \beta_j$. Since this is the only intersection which contains x, we can put everything together to get our desired formula:

$$s_1(x) + s_2(x) = \sum_{i=0}^{m} \sum_{j=0}^{n} (\alpha_i + \beta_j) \cdot \mathbb{1}_{A_i \cap B_j}(x)$$

□

The integral of an indicator function is already clear from our intuition: The points contained in the area under an indicator function $\mathbb{1}_A$ are exactly the cartesian product of A with the interval $[0, 1]$, forming a "rectangle with gaps" whose side lengths are 1 and $\mu(A)$. Therefore the integral should be:

Definition 13.1.6. (Lebesque integral of an indicator function):

$$\int_X \mathbb{1}_A \, d\mu = \mu(A)$$

The definition for the integral of a step function should follow naturally - a step function is just a finite sum of scaled characteristic functions, therefore the integral of a step function is the sum of the scaled integrals of the step functions.

Definition 13.1.7. (Lebesque integral of a step function): Let $D \subset X$, $A_i \subset X$ pairwise disjoint, and $\alpha_i \geq 0$. Then:

$$\begin{aligned}
\int_D s \, d\mu &= \int_D \left(\sum_{i \leq k} \alpha_i \cdot \mathbb{1}_{A_i} \right) d\mu \\
&:= \sum_{i \leq k} \left(\alpha_i \cdot \int_D \mathbb{1}_{A_i} \, d\mu \right) \\
&= \sum_{i \leq k} (\alpha_i \cdot \mu(D \cap A_i))
\end{aligned}$$

Theorem 13.1.8. (Linearity of the step function integral):

Theorem 13.1.9. (Monotonicity of the step function integral):

13.2 Defining the Lebesque Integral

We can now use the step function integral to define the integral of more general functions:

Theorem 13.2.0.1: Approximation via step functions

Let $(X, \mathcal{A})$ be a measurable space. Let $f : X \to \overline{\mathbb{R}}$ be a nonnegative $\mathcal{A}$-measurable function. Then there exists a monotonically increasing sequence s_n of nonnegative step functions whose pointwise limit is f.

Construction 1. For $n \in \mathbb{N}$ and $k \in \{0, \dots, n \cdot 2^n\}$, we set:

$$F_{n,k} := \left\{ x \in X \,\middle|\, \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \right\}.$$

Then

$$s_n(x) = \begin{cases} \frac{k}{2^n} & x \in F_{n,k} \\ n & \text{otherwise} \end{cases}$$

are step functions converging to f. □

Construction 2. Define $s_0 = 0$. Then we can inductively define

$$E_n := \left\{ s_{n-1} + \frac{1}{n} \leq f \right\}$$

and

$$s_n := s_{n-1} + \frac{1}{n} \cdot \mathbb{1}_{E_n},$$

meaning we have

$$s_n = \sum_{k=1}^{n} \frac{1}{k} \cdot \mathbb{1}_{E_i}.$$

We now need to show that this series converges pointwise to f. Let $n \in \mathbb{N}$.

1. Assume $x \in E_n$. Then, by definition,

$$f_n(x) = f_{n-1} + \frac{1}{n} \leq f(x)$$

2. Assume $x \notin E_n$. Then, by induction, we have

$$f_n(x) = f_0(x) = 0 \leq f(x).$$

Therefore, we have $f_0 \leq f_1 \leq \dots$ and $f_n \leq f$ for all n, meaning that

$$\lim_{n \to \infty} f_n(x) \leq f(x).$$

If $f(x) = \infty$, then $x \in E_n$ for every n, and we have

$$\lim_{n\to\infty} f_n(x) = \sum_{k=1}^{\infty} \frac{1}{k} = \infty$$

If $f(x) < \infty$, then there must be an infinite number of $n \in \mathbb{N}$ such that $f_{n-1}(x) > f_n(x) - \frac{1}{n}$, implying

$$\lim_{n\to\infty} f_n(x) \geq f(x)$$

□

Definition 13.2.0.1: Lebesque integral of a positive function

Let $f : D \to [0, \infty]$ be $\mathcal{A}$-measurable. Then:

$$\int_D f \, d\mu = \sup_{\substack{s \text{ is a step function,} \\ 0 \leq s \leq f}} \left\{ \int_D s \, d\mu \right\}$$

Definition 13.2.1. (Lebesque integral of an arbitrary measurable function): Let f be $\mathcal{A}$-mesurable and

$$f^+ = f \cdot \mathbb{1}_{f \geq 0}, \qquad f^- = -f \cdot \mathbb{1}_{f<0}$$

Note that $f^+ \geq 0$ and $f^- > 0$. Then, as long as f^+ or f^- have a finite integral, we can define:

$$\int_D f \, d\mu \int_D f^+ \, d\mu - \int_D f^- \, d\mu$$

Corollary 13.2.2. (Integrating over subsets):

$$\int_M f \, d\mu = \int_X f \cdot \mathbb{1}_M \, d\mu = \int_M f \, d\mu|_M$$

It is common to derive from this corollary a slight abuse of notation: Assume that f is not $\mathcal{A}$-measurable, but it is μ-measurable, i.e. $f : D \to \overline{\mathbb{R}}$ such that $\mu(X \setminus D) = 0$ and f is $\mathcal{A}|_D$-measurable. Then it is common to implicitly expand the domain of D to X by setting $f(x) = 0$ on $X \setminus D$, and to therefore write:

$$\int_X f \, d\mu := \int_D f \, d\mu$$

Corollary 13.2.3. (Integrating over zero sets): *Let N be a set such that $\mu(N) = 0$. Then*

$$\int_N f \, d\mu = 0.$$

Definition 13.2.3.1: Integrable Function

We call a function $f : X \to \overline{\mathbb{R}}$ **integrable** with regards to a measure μ if it is μ-measurable and

$$\int_X f \, d\mu \in \mathbb{R}$$

Proposition 13.2.4. (Integrating with the counting measure): *Let X be an arbitrary set. Let* card *be the counting measure on $\mathcal{P}(X)$. Let $f : X \to \overline{\mathbb{R}}$. Then f is integrable with respect to* card *if and only if $\sum_{x \in X} f(x)$ is absolutely convergent, and we have*

$$\int_X f \, d\text{card} = \sum_{x \in X} f(x)$$

Theorem 13.2.4.1: Monotonicity of the Lebesque Integral

Let μ be an outer measure on X such that $f, g : X \to \overline{\mathbb{R}}$ are μ-measurable. Then if $f \leq g$ μ-almost everywhere and $\int_X f \, d\mu > -\infty$, then the integral of g exists and we have

$$\int_X f \, d\mu \leq \int_X g \, d\mu$$

Proof.

1. Assume f, g are nonnegative. Then if s is a step function such that $s \leq f$. Then if we define $t := \mathbb{1}_{f \leq g} \cdot s$, we have $t \leq g$. Furthermore, for all $c \geq 0$, we have:

$$\mu(\{s = c\}) = \mu(\{s = c\} \cap)$$

2.

□

Theorem 13.2.4.2

Let $f, g : X \to \overline{\mathbb{R}}$ and let f be μ-measurable. Then if $g = f$ μ-almost everywhere, then g is μ-measurable, and

$$\int_X g \, d\mu = \int_X f \, d\mu$$

as long as the right integral exists.

Lemma 13.2.5. (Chebyshev Inequality): *Let $f : X \to [0, \infty]$ be μ-measurable with $\int_X f \, d\mu < \infty$. Let $s \in (0, \infty]$. Then:*

$$\mu(\{x : f(x) \geq s\}) \leq \begin{cases} \frac{1}{s} \int_X f \, d\mu & s \in (0, \infty) \\ 0 & s = \infty \end{cases}$$

Corollary 13.2.6. *Let $f : X \to \overline{\mathbb{R}}$ be μ-measurable. Then:*

1. *If $\int_X f \, d\mu < \infty$, $\{x : f(x) = \infty\}$ is a μ-zero set.*
2. *If $f \geq 0$ and $\int_X f \, d\mu = 0$, $\{x : f(x) > 0\}$ is a μ-zero set.*

Theorem 13.2.6.1: Monotone Convergence Theorem

Let $f_n : X \to [0, \infty]$ be μ-measurable such that $f_i \le f_{i+1}$. Let $\lim_{n\to\infty} f_n = f$. Then

$$\int_X \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_X f_n \, d\mu = \sup_{n\in\mathbb{N}} \int_X f_n \, d\mu$$

Theorem 13.2.6.2: Linearity of the Lebesque Integral

Let μ be a measure on X. Let $f, g : X \to \overline{\mathbb{R}}$ and $\alpha, \beta \in \mathbb{R}$. Then we have:

$$\int_X (\alpha f + \beta g) \, d\mu = \alpha \int_X f \, d\mu + \beta \int_X g \, d\mu$$

Corollary 13.2.7. *Let $f : X \to \overline{\mathbb{R}}$ be μ-measurable. Then f is integrable if and only if $|f|$ is integrable.*

Corollary 13.2.8. *Let $f : X \to \overline{\mathbb{R}}$ be μ-measurable. Then, assuming the integral of f exists, we have*

$$\left| \int_X f \, d\mu \right| \le \int_X |f| \, d\mu$$

Corollary 13.2.8.1: Comparison Test for the Lebesque Integral

Let $g : X \to [0, \infty]$ be μ-measurable such that $|f| \le g$ μ-almost everywhere. Then if

$$\int_X g \, d\mu < \infty,$$

f is integrable.

13.3 Comparing Lebesque Integration with Riemann Integration

Theorem 13.3.1. *The Dirichlet function $\mathbb{1}_{\mathbb{Q}}$ is Lebesque integrable, but not Riemann integrable.*

Theorem 13.3.1.1

A Lebesque-integrable function is not necessarily "almost Riemann-integrable". More formally, there exists a function $f : \mathbb{R} \to \mathbb{R}$ with the following properties:

1. f is Lebesque-integrable.
2. There exists no Riemann-integrable function g such that $f = g$ almost everywhere.

Theorem 13.3.1.2

Let $f : [a, b] \to \mathbb{R}$. Then f is Riemann integrable if and only if it is the limit of a uniformly convergent sequence of step functions over Jordan measurable sets.

13.4 Convergence Theorems

13.5 L^p-Spaces

13.6 Density Functions

Theorem 13.6.0.1

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\theta : X \to \overline{\mathbb{R}}$ be nonnegative and μ-measurable. Then the map

$$\begin{aligned} \nu : \mathcal{A} &\to \overline{\mathbb{R}} \\ A &\mapsto \int_A \theta \, d\mu \end{aligned}$$

is a measure, which we denote μ_θ. We call θ the **density of** ν **with respect to** μ.

Corollary 13.6.1. *The following hold for* μ_θ*:*

1. $\mu(A) = 0$ *implies* $\mu_\theta(A) = 0$.
2. *For every nonnegative* μ*-measurable function* f*, we have*
$$\int_X f \, d\mu_\theta = \int_X f \cdot \theta \, d\mu.$$
3. θ *is unique up to equality* μ*-almost everywhere.*

Definition 13.6.1.1

Let μ and ν be measures on $(X, \mathcal{A})$. Then we call ν **absolutely continuous with respect to** μ, which we denote $\nu \ll \mu$, if $\mu(A) = 0$ implies $\nu(A) = 0$.

Lemma 13.6.2. *Let* σ *and* ν *be finite measures on* $(X, \mathcal{A})$ *such that* $\nu(A) \leq \sigma(A)$ *for every* $A \in \mathcal{A}$*. Then there exists a density function* θ *such that* $\nu = \sigma_\theta$*.*

This lemma is actually a special case of the significantly more general Riesz representation theorem for Hilbert spaces, which I sadly do not have the time to go into at this point (but look it up, it's really neat!).

Lemma 13.6.3. *Let* μ, ν *be measures on* $(X, \mathcal{A})$*. Let* $\sigma := \mu + \nu$*. Then if* $f \in \mathcal{L}^*(\sigma)$*, then* $f \in \mathcal{L}^*(\mu)$ *and* $f \in \mathcal{L}^*(\nu)$ *and we have*

$$\int_X f \, d\sigma = \int_X f \, d\mu + \int_X f \, d\nu$$

Theorem 13.6.4. (Mini-Radon-Nikodym): *Let* μ*,* ν *be finite measures on* $(X, \mathcal{A})$ *such that* $\nu \ll \mu$*. Then there exists a density function* $\theta \in L^1(\mu)$ *such that* $\nu = \mu_\theta$*.*

In this case, θ is sometimes also called the **Radon-Nikodym derivative of** ν **with respect to** μ and denoted $\frac{d\nu}{d\mu}$.

Theorem 13.6.5. *Let* μ *and* ν *be finite measures on* $(X, \mathcal{A})$*. Then the following are equivalent:*

1. $\nu \ll \mu$*,*
2. *There exists a density function* $\theta \in L^1(\mu)$ *such that* $\nu = \mu_\theta$*,*
3. *For all* $\varepsilon > 0$*, there exists a* $\delta > 0$ *such that* $\mu(A) < \delta$ *implies* $\nu(A) < \varepsilon$*.*

Proof. (i) $\implies$ (ii) is our mini-Radon-Nikodym theorem. (ii) $\implies$ (iii) follows from absolute continuity of the Lebesque integral. (iii) $\implies$ (i) follows immediately from the definition of $\ll$. $\square$

Theorem 13.6.5.1: Radon-Nikodym

Let μ, ν be σ-finite measures on $(X, \mathcal{A})$ such that $\nu \ll \mu$. Then there exists a density function $\theta \in L^1(\mu)$ such that $\nu = \mu_\theta$.

Definition 13.6.5.1

Let μ and ν be measures on $(X, \mathcal{A})$. Then we call μ and ν **singular with respect to each other**, which we denote $\mu \perp \nu$, if there exists a set $M \in \mathcal{A}$ such that

$$\mu(M) = \nu(X \setminus M) = 0.$$

Theorem 13.6.5.2: Lebesque's decomposition theorem

Let μ and ν be measures on $(X, \mathcal{A})$, and let ν be σ-finite. Then there exists a unique decomposition $\nu = \nu_a + \nu_s$ such that $\nu_a \ll \mu$ and $\nu_a \perp \mu$.

Chapter 14

Integration over Immersed Manifolds

14.1 Product Measures and Fubini's Theorem

14.2 Change of Variables

Lemma 14.2.1. *Let $U \subset \mathbb{R}^n$ and $x_0 \in U$. Let $\varphi : U \to \mathbb{R}^n$ be a function such that $D\varphi(x_0)$ is invertible.*

Then for a sequence $Q_j = Q(x_j, r_j) \subset U$ of cuboids of sidelength r_j with center x_j such that $r_j \to 0$ and $x_0 \in Q_j$, we have:

$$\limsup_{j\to\infty} \frac{\lambda^n(\varphi(Q_j))}{\lambda^n(Q_j)} \leq |\det D\varphi(x_0)|$$

Proof. We can assume $x_0 = 0$ and $\varphi(0) = 0$, since otherwise we can translate space as needed before doing any calculations without breaking any of our assumptions.

1. Assume $D(\varphi(0)) = E_n$. Then, by definition of differentiablity and equivalence of norms on finite-dimensional vector spaces, we have

$$\begin{aligned} 0 &= \lim_{x\to 0} \frac{\|\varphi(x) - \varphi(0) - D\varphi(0)x\|\infty}{\|x\|_\infty} \\ &= \lim_{x\to 0} \frac{\|\varphi(x) - x\|_\infty}{\|x\|_\infty}, \end{aligned}$$

   Let $\varepsilon > 0$. Then, by the definition of convergence, for every x with a sufficiently small norm, we have:

$$\frac{\|\varphi(x) - x\|_\infty}{\|x\|_\infty} \leq \varepsilon,$$

   which means

$$\|\varphi(x) - x\|_\infty \leq \varepsilon\|x\|_\infty.$$

Furthermore, for $x \in Q_j$, we have:

$$\begin{aligned}\|x\|_\infty &= \left\|x - \vec{0}\right\|_\infty \\ &= \|x - x_0\|_\infty \\ &\le \left\|x - x_j\right\| + \left\|x_j - x_0\right\| \\ &\le 2r_j\end{aligned}$$

For sufficiently large j, these imply:

$$\begin{aligned}\|\varphi(x) - x\|_\infty &\le \varepsilon\|x\|_\infty \\ &\le 2\varepsilon r_j.\end{aligned}$$

Further applying the triangle inequality, we get:

$$\begin{aligned}\|\varphi(x) - \varphi(x)\| &\le \|\varphi(x) - x\|_\infty \\ &\quad + \left\|x - x_j\right\|_\infty \\ &\quad + \left\|x_j - \varphi(x_j)\right\|_\infty \\ &\le 2\varepsilon r_j + r_j + 2\varepsilon r_j \\ &\le (1+4\varepsilon)r_j,\end{aligned}$$

which means that φ increases the side length of our cube by a factor of at most $(1+4\varepsilon)$. Therefore, it increases the volume by a factor at most $(1+4\varepsilon)^n$, i.e:

$$\frac{\lambda^n(\varphi(Q_j))}{\lambda^n(Q_j)} \le (1+4\varepsilon)^n$$

Letting $j \to \infty$ and $\varepsilon \searrow 0$, we have

$$\begin{aligned}\limsup_{j\to\infty} \frac{\lambda^n(\varphi(Q_j))}{\lambda^n(Q_j)} &\le \lim_{\varepsilon\to 0}(1+4|\varepsilon|)^n \\ &= 1 \\ &= |\det E_n|\end{aligned}$$

2. Now, let $S := D\varphi(0)$ and $\varphi_0 := S^{-1} \circ \varphi$, i.e. $\varphi = S \circ \varphi_0$. Then $D\varphi_0(0) = E_n$. By the linear transformation equation $\lambda^n(S(E)) = |\det(S)|\lambda^n(E)$ (11.5.21), we have:

$$\begin{aligned}\limsup_{j\to\infty} \frac{\lambda^n(\varphi(Q_j))}{\lambda^n(Q_j)} &= \limsup_{j\to\infty} \frac{\lambda^n(S(\varphi_0(Q_j)))}{\lambda^n(Q_j)} \\ &= |\det S| \limsup_{j\to\infty} \frac{\lambda^n(\varphi_0(Q_j))}{\lambda^n(Q_j)} \\ &\le |\det S| \\ &= |\det D\varphi(0)|\end{aligned}$$

□

Theorem 14.2.1.1: Multivariable Substitution Formula

Let $U \subset \mathbb{R}^n$ be open. Let $\varphi : U \to \mathbb{R}^n$ be C^1. Then if $f : V \to \overline{\mathbb{R}}$ is λ^n-measurable, we have:

$$\int_V f(y)\,dy = \int_{\varphi^{-1}(V)} f(\varphi(x)) \cdot |\det D\varphi(x)|\,dx.$$

Corollary 14.2.2. *Let $U, V \subset \mathbb{R}^n$ be open. Let $\varphi : U \to \mathbb{R}^n$ be C^1. Then if $A \subset U$ is λ^n-measurable, so is $\varphi(A)$, and we have*

$$\lambda^n(\varphi(A)) = \int_A |\det D\varphi(x)|\,dx.$$

Proof. Apply the previous equation to $f = \mathbb{1}_{\varphi(A)}$. □

Example 14.2.3. (The Gaussian Integral): We want to find the area under the Gaussian bell curve pre-normalization, i.e.

$$\int_{\mathbb{R}} e^{-x^2}\,dx$$

To do this, we add an additional dimension and exploit the resulting rotational symmetry. By Fubini's Theorem, we have:

$$\begin{aligned}
\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\,d\lambda^2(x,y) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-(x^2+y^2)}\,d\lambda(y)\right) d\lambda(x) \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-x^2} \cdot e^{-y^2}\,d\lambda(y)\right) d\lambda(x) \\
&= \int_{\mathbb{R}} e^{-x^2} \cdot \left(\int_{\mathbb{R}} e^{-y^2}\,d\lambda(y)\right) d\lambda(x) \\
&= \int_{\mathbb{R}} e^{-x^2}\,d\lambda(x) \cdot \int_{\mathbb{R}} e^{-y^2}\,d\lambda(y) \\
&= \left(\int_{\mathbb{R}} e^{-x^2}\,d\lambda(x)\right)^2
\end{aligned}$$

Therefore, we have

$$\int_{\mathbb{R}} e^{-x^2}\,d\lambda(x) = \sqrt{\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\,d\lambda(x,y)^2}$$

We can now calculate the two-dimensional integral using multivariable substitution to transform to polar coordinates:

$$\begin{aligned}
\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\, d\lambda^2(x,y) &= \int_{(0,\infty)\times(0,2\pi)} \dots \\
&= \int_{(0,\infty)\times(0,2\pi)} re^{-r^2}\, d\lambda^2(r,\omega) \\
&= \int_0^\infty \left(\int_0^{2\pi} re^{-r^2}\, d\omega \right) dr \\
&= \int_0^\infty 2\pi re^{-r^2}\, dr \\
&= 2\pi \int_0^\infty re^{-r^2}\, dr \\
&= 2\pi \int_0^\infty \frac{1}{2} e^{-r^2}\, 2rdr \\
&= 2\pi \int_{-\infty}^0 \frac{1}{2} e^s\, ds \\
&= \pi \int_{-\infty}^0 e^s\, ds \\
&= \pi
\end{aligned}$$

Which means that the area under the bell curve is $\sqrt{\pi}$.

Chapter 15

Generalizing even further

15.1 The Bochner Integral

15.2 The Pettis Integral

Part VI

Summaries

Appendix A

Littlewood's three Principles of Real Analysis

Appendix B

Modes of Convergence

There are many different inequivalent ways in which a series of function $(f_i)_{i\in\mathbb{N}}$ on a common domain X could "converge" to a function f:

1. Assume $f_n : X \to T$, where (T, τ) is a topological space. Then we say f_n converges **pointwise** to f if, for every $x \in X$, $f_n(x)$ converges to $f(x)$, i.e. for every neighborhood U around $f(x)$, all points $f_n(x)$ eventually lie in U for large enough n:

$$\forall x \in X : \forall U \in \mathcal{N}(f(x)) : \exists N \in \mathbb{N} : \\ n \geq N \implies f_n(x) \in U$$

   If we assume functions $f_n : X \to M$, where (M, d_M) is a metric space equipped with the metric topology, then this is equivalent to the statement that the distance $d_M(f(x), f_n(x))$ gets arbitrarily small:

$$\forall x \in X : \forall \varepsilon \in \mathbb{R} : \exists N \in \mathbb{N} : \\ n \geq N \implies d_M(f(x), f_n(x)) \leq \varepsilon$$

2. If we have a metric on the set of functions themselves, we can just have the f_n converge to f directly like any set of points would. Therefore we say that f_n converges to f **in Norm**, if:

$$\forall \varepsilon \in \mathbb{R} : \exists N \in \mathbb{N} : \\ n \geq N \implies d_M(f, f_n) \leq \varepsilon$$

   A special case that is particularly important for real (and functional) analysis is **convergence in L^p-Norm**, i.e.:

$$\forall \varepsilon \in \mathbb{R} : \exists N \in \mathbb{N} : \\ n \geq N \implies \|f - f_n\|_{L^p} \leq \varepsilon,$$

   where:

$$\|f - f_n\|_{L^p} = \left(\int_X |f - f_n|^p \, d\mu \right)^{\frac{1}{p}}$$

3. (TODO: Uniform Spaces)

Let $f_n : X \to M$, where $(M, \leq)$ is a metric space. Then we say f_n converges **uniformly** to if the same condition still holds when we have to choose our N independently of (i.e. before) x:

$$\forall \varepsilon \in \mathbb{R} : \exists N \in \mathbb{N} : \forall x \in X : \\ n \geq N \implies d_M(f(x), f_n(x)) \leq \varepsilon$$

4. Let $f_n : X \to M$ and let μ be a measure on X. Then f_n converges to f **almost uniformly** if there exists a set A_ε such that $\mu(A_\varepsilon) < \varepsilon$ and such that f_n converges to f uniformly on $X \setminus A_\varepsilon$. Note that this does **not** imply that f_n converges uniformly to f *almost everywhere*, since all our A_ε still have positive measure, and so uniform convergence might not hold in the "limit case" where A_ε has to be zero.

Example B.0.1. The sequence $f_n : [0, 1] \to \mathbb{R}$, $x \mapsto x^n$ converges to the zero function almost uniformly, but not uniformly almost everywhere:

1.

2. Let E have zero measure. Then E cannot contain any closed interval as a subset. Therefore, for any m, there must be a point $x_m \in \left[1 - \frac{1}{m}, 1 - \frac{1}{m+1}\right]$ such that $x_m \notin E$. We therefore have:

$$\begin{aligned} \sup_{x \in [0,1] \setminus E} |f_n - 0| &= \sup_{x \in [0,1] \setminus E} |x^n| \\ &\geq f_n(x_m) \\ &\geq f_n\left(1 - \frac{1}{m}\right) \\ &= \left(1 - \frac{1}{m}\right)^n \end{aligned}$$

Therefore, f_n cannot converge uniformly to 0, since for every choice of E and any arbitrarily large $n \geq N$, we can still always find a point $x_m \in [0, 1] \setminus E$ such that $f_n(x_m)$ is arbitrarily close to 1.