Analysis for people who don’t like skipping details

Emma Yoneda

2026-02-15

Contents

Part I

Set Theory

Chapter 1

Relations and Maps

Theorem 1.0.1. (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:*

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

Chapter 2

Orders

Definition 2.0.1. 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 2.0.1.1: 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 3

Ordered Fields

3.1 Basic Properties

Definition 3.1.0.1: 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 3.1.1. *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 3.1.2. *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 3.1.3. *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 3.1.4. *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 3.1.5. *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 3.1.6. 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 $ac \geq bc$.

3.2 The Archimedean Property

Definition 3.2.1. 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 3.2.2. 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.

3.3 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!

3.3.1 Subfields of ordered fields

Theorem 3.3.1. *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.

3.3.2 The least upper bound property

Definition 3.3.2. 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 3.3.3. *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 3.3.4. *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 3.3.5. *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 3.3.6. *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.*

3.3.3 Alternative completeness properties

Part III

Topology

Chapter 4

Metric Spaces and Topological Spaces

4.1 Vocabulary

Definition 4.1.1. 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 4.1.2. *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 4.1.3. 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 4.1.4. *The complement of a G_δ set is F_σ and vice versa.*

4.2 Sequences and Limits

Definition 4.2.0.1: 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 4.2.1. *We have:*

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

and

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

4.3 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 4.3.0.1: 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 4.3.1. *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 4.3.2. *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 4.3.3. 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 4.3.4. *$f : X \to Y$ 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 4.3.5. *$f : X \to Y$ 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 4.3.5.1: 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 4.3.6. *Every continuous function is sequentially continuous.*

Theorem 4.3.7. *If X is first-countable (and we assume the axiom of choice), then any sequentially continuous function is continuous.*

Corollary 4.3.7.1

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 4.3.7.1: 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 4.3.8. *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 $x \in \text{int}(A)$. 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 $x \in \text{int}(X \setminus A)$. 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 4.3.9. 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 4.3.10. *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 4.3.10.1: 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.

$\square$

Theorem 4.3.10.2: A function only continuous at the irrationals

Thomae's function $T : \mathbb{R} \to \mathbb{R}$, 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 4.3.10.3: 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 4.3.11. (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 4.3.12. *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 4.3.13. *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 5

Topological Fields

Theorem 5.0.1. *Let F be an ordered field. Then F becomes a topological field if we give it the order topology.*

Theorem 5.0.2. *Limits and field operations*

Corollary 5.0.3. *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 6

Topological Vector Spaces

Chapter 7

Topological Manifolds

Lemma 7.0.1. *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 7.0.2. 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 7.0.3. 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 7.0.4. *Every open subset of an n-dimensional topological manifold is itself an n-dimensional topological manifold.*

Theorem 7.0.5. *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 7.0.6. *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 7.0.7. *Every topological n-manifold is homeomorphic to a subset of a Euclidean space $\mathbb{R}^k$, where $k \geq n$.*

Corollary 7.0.8. *Every topological manifold is separable and metrizable.*

Chapter 8

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.

8.1 Diagonal Uniformity

Definition 8.1.1. 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 8.1.2. 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 8.1.3. 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 8.1.4. 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 8.1.5. 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 9

Differentiation in Normed Vector Spaces

9.1 The Fréchet Derivative

Definition 9.1.1. (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 9.1.2. *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 9.1.3. 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 9.1.4.

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 9.1.5. (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:

$$\begin{aligned} &M(\vec{x} + \vec{h}) - M(\vec{x}) - DM(\vec{h}) \\ =&(x_1 + h_1)(x_2 + h_2) - x_1x_2 - h_1x_2 - h_2x_1 \\ =&x_1x_2 + x_1h_2 + h_1x_2 + h_1h_2 - x_1x_2 - h_1x_2 - h_2x_1 \\ =&h_1 \cdot h_2 \end{aligned}$$

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 9.1.6. (Linearity of the differentiation operator): *Let V, W be normed vector spaces over a field $\mathbb{F}$, and let $F, G : V \supset U \longrightarrow W$ 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. $D(F + G)(\vec{x}) = DF(\vec{x}) + DG(\vec{x})$

9.2 Differential Operators

9.2.1 Divergence

Definition 9.2.0.1: Divergence of a Vector Field

Let

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

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

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

Comment 9.2.1. If we assume the standard basis on $\mathbb{R}^n$, we have:

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

For the rest of the text, we will always assume the standard basis on $\mathbb{R}^n$, and thus use this expression liberally. $\operatorname{div} f$ 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 $(\operatorname{div} f)(x) < 0$.

Corollary 9.2.2. *Let* $\operatorname{id} : \mathbb{R}^d \to \mathbb{R}^d$ *be the identity function. We have*

$$\operatorname{div}(\operatorname{id}) = d,$$

or, if the argument of the function is included:

$$\operatorname{div}(x) = \operatorname{div}(\operatorname{id}(x)) = d.$$

Theorem 9.2.2.1: Linearity of Divergence

Let $F, G : \mathbb{R}^n \to \mathbb{R}^n$. Let $a, b \in \mathbb{R}$. Then we have:

$$\operatorname{div}(a \cdot F(x) + b \cdot G(x)) = a \cdot \operatorname{div}(F(x)) + b \cdot \operatorname{div}(G(x))$$

Proof.

$$\begin{aligned}
\operatorname{div}(a \cdot F(x) + b \cdot G(x)) &= \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(a \cdot F(x) + b \cdot G(x)) \\
&= \sum_{i=1}^{n} a \cdot \frac{\partial}{\partial x_i} F(x) + b \cdot \frac{\partial}{\partial x_i} G(x) \\
&= \left(a \cdot \sum_{i=1}^{n} \frac{\partial}{\partial x_i} F(x)\right) + \left(b \cdot \sum_{i=1}^{n} \frac{\partial}{\partial x_i} G(x)\right) \\
&= a \cdot \operatorname{div}(F(x)) + b \cdot \operatorname{div}(G(x))
\end{aligned}$$

□

Theorem 9.2.2.2: Product Rule for Divergence

Let $f : \mathbb{R}^n \to \mathbb{R}$. Let $G : \mathbb{R}^n \to \mathbb{R}^n$. Then we have:

$$\operatorname{div}(f \cdot G) = f \cdot \operatorname{div}(G) + \left\langle \vec{\nabla} f, G \right\rangle$$

Proof. This follows from the standard product rule of differentiation:

$$\begin{aligned}
\operatorname{div}(f(x) \cdot G(x)) &= \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x) G_i(x) \\
&= \sum_{i=1}^{n} f(x) \cdot \frac{\partial}{\partial x_i} G_i(x) + G_i(x) \cdot \frac{\partial}{\partial x_i} f(x) \\
&= \left(\sum_{i=1}^{n} f(x) \cdot \frac{\partial}{\partial x_i} G_i(x)\right) + \left(\sum_{i=1}^{n} G_i(x) \cdot \frac{\partial}{\partial x_i} f(x)\right) \\
&= f(x) \cdot \left(\sum_{i=1}^{n} \frac{\partial}{\partial x_i} G_i(x)\right) + \left(\sum_{i=1}^{n} G_i(x) \cdot \frac{\partial}{\partial x_i} f(x)\right) \\
&= f(x) \cdot \operatorname{div}(G(x)) + \left\langle \vec{\nabla} f(x), G(x) \right\rangle
\end{aligned}$$

□

9.2.2 Laplacian

Definition 9.2.2.1: Laplacian

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 **laplacian** of f is defined to be

$$\Delta f = \operatorname{div}(\nabla f)$$

Comment 9.2.3. Once again, assuming the standard base on $\mathbb{R}^n$, we simply have:

$$\Delta f = \operatorname{div}(\nabla f) = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} f_i(x)$$

Chapter 10

First Steps in Differential Geometry

10.1 C^n-Submanifolds of $\mathbb{R}^n$

Definition 10.1.1. 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.

Definition 10.1.1.1: C^n-Manifold

Hi

Theorem 10.1.1.1

Let $M \subset \mathbb{R}^{n+k}$. Then the following are equivalent:

1. M is an n-dimensional C^r-submanifold of $\mathbb{R}^{n+k}$
2. M is locally given as a C^r-level set, i.e. For every $p \in M$ there exists an open set $\Omega \subset \mathbb{R}^{n+k}$ and an r-times continuously differentiable function $f : \Omega \to \mathbb{R}^k$ such that:
   (a) $M \cap \Omega = \{x \in \Omega : f(x) = 0\}$
   (b) For all $x \in \Omega$, we have rank $Df(x) = k$.
3. M

Definition 10.1.1.2: Tangent Vector

Let $M \subset \mathbb{R}^{n+k}$. Let $p \in M$. Then we call a vector $v \in \mathbb{R}^{n+k}$ a *tangent vector of M at p* If there exists a map $\gamma : (-\delta, \delta) \to M$ such that $\gamma(0) = p$ and $D\gamma(0) = v$.

Theorem 10.1.1.2: Tangent Space

Now, let M be a manifold. Then the set of all tangent vectors of M at p forms an n-dimensional vector space, which we call the *tangent space of M at p* and which we denote T_pM.

10.2 Inverse Function Theorem

Theorem 10.2.0.1: 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.

10.3 Implicit Function Theorem

Part V

Measure and Integration

Chapter 11

The Riemann Integral

Definition 11.0.0.1: 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 11.0.1. *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 11.0.2. *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.*

Chapter 12

Some Special Functions

12.1 The Gamma Function

In this section, we discover some basic properties of the so-called gamma function. We will find that it is the canonical extension of the factorial function to positive real numbers. With some knowledge of complex analyis, it can easily be extended to arbitrary complex numbers, but that is beyond the scope of these notes.

Definition 12.1.0.1: Gamma function

For any $x \in \mathbb{R}_{>0}$, the gamma function is defined as the improper integral

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt$$

Theorem 12.1.1. *The gamma function obeys the recurrence relation*

$$\Gamma(x+1) = x \cdot \Gamma(x)$$

Proof. Integration by parts gives:

$$\begin{aligned}
\Gamma(x+1) &= \int_0^\infty t^x e^{-t}\,dt \\
&= \left[-t^x e^{-t}\right]_0^\infty + \int_0^\infty x t^{x-1}e^{-t}\,dt \\
&= 0 - \lim_{n\to\infty}(-t^x e^{-t}) + x\int_0^\infty t^{x-1}e^{-t}\,dt \\
&= x\int_0^\infty t^{x-1}e^{-t}\,dt \\
&= x \cdot \Gamma(x)
\end{aligned}$$

□

Corollary 12.1.2. *For any $n \in \mathbb{N}$, we have*

$$\Gamma(n+1) = n!$$

Proof. Our recurrence relation gives

$$\Gamma(n+1) = n! \cdot \Gamma(1)$$

and we have

$$\begin{aligned}
\Gamma(1) &= \int_0^\infty t^{1-1} e^{-t}\, dt \\
&= \int_0^\infty e^{-t}\, dt \\
&= \left[-e^{-t}\right]_0^\infty \\
&= \left(\lim_{n\to\infty} -e^{-n}\right) - (-e^0) \\
&= 0 - (-1) \\
&= 1
\end{aligned}$$

□

Lemma 12.1.3. *We have:*

$$\Gamma(x) = 2c^x \int_0^\infty t^{2x-1} e^{-ct^2}\, dt$$

Proof. Substituting $t := cu^2$, we get:

$$\begin{aligned}
\Gamma(x) &= \int_0^\infty t^{x-1} e^{-t}\, dt \\
&= \int_0^\infty (cu^2)^{x-1} e^{-cu^2} \cdot 2cu\, du \\
&= \int_0^\infty c^{x-1} u^{2(x-1)} e^{-cu^2} \cdot 2cu\, du \\
&= \int_0^\infty c^x u^{2x-1} e^{-cu^2}\, du \\
&= c^x \cdot \int_0^\infty u^{2x-1} e^{-cu^2}\, du \\
&= c^x \cdot \int_0^\infty t^{2x-1} e^{-ct^2}\, dt
\end{aligned}$$

□

Proposition 12.1.4. *We have* $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Proof Sketch. By the last lemma, we have:

$$\begin{aligned}\Gamma\left(\frac{1}{2}\right) &= 2\int_0^\infty t^{2\cdot\frac{1}{2}-1}e^{-t^2}\ dt\\ &= 2\int_0^\infty e^{-t^2}\ dt\\ &= \int_{-\infty}^\infty e^{-t^2}\ dt\end{aligned}$$

The fact that $\int_{-\infty}^\infty e^{-t^2}\ dt = \sqrt{\pi}$ might be familiar if you have a basic knowledge of probability theory, since this is exactly the reason why π appears in the formula for a normal distribution:

$$\mathcal{N}(\mu,\sigma^2)(t) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)}$$

Setting $\mu = 0$ and $\sigma^2 = \frac{1}{2}$, which was Gauss' original definition for the standard normal distribution, we get:

$$\mathcal{N}(0,1)(t) = \frac{1}{\sqrt{\pi}}e^{-t^2},$$

where the factor $\frac{1}{\sqrt{\pi}}$ appears to normalize the integral of the distribution to 1, since it wouldn't be a probability distribution otherwise.

We will formally prove this fact much later, when we introduce the substitution formula for multidimensional Lebesque integration. "$\square$"

Chapter 13

Measure Theory

13.1 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 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.

4. We want our volume to be a positive real number.

We call a σ-additive function a *measure* - we want to eventually find a meausure on $\mathbb{R}^n$ that also fulfills the other properties, but we want measures to be far more generally applicable, and as such we want to be able to define measures much more general spaces that may not have a defined notion of an interval or an isometry.

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 - enough to include every "somewhat reasonable" subset of $\mathbb{R}^n$. However, there are still counterexamples.

Proposition 13.1.1. *Every subset of $\mathbb{R}^n$ being Lebesque-measurable is consistent with ZF (without the axiom of choice).*

Theorem 13.1.2. *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*.

13.1.1 Vitali Sets

Proposition 13.1.3. *The relation $x \sim y \Leftrightarrow x - y \in \mathbb{Q}$ is an equivalence relation on the real numbers.*

Theorem 13.1.4. *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 13.1.5. *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 13.1.6. *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 13.1.7. *Vitali sets are not measurable by a translation-invariant normed measure.*

Proof. Assume that λ is translation-invariant, countably additive and that $\lambda([a,b]) = b - a$. Then we have:

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

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

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

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

13.1.2 The Banach-Tarski Paradox

Theorem 13.1.8. *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 13.1.9. *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.*

13.2 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*:

13.2.1 Boolean Algebras

Definition 13.2.1. 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 13.2.2. *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 13.2.3. *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 13.2.4. (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 13.2.5. (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 13.2.6. *If $Y \in \mathcal{B}$, then*

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

Theorem 13.2.7. (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 13.2.8. *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 13.2.9. *Every finite Boolean algebra is an atomic algebra.*

Corollary 13.2.10. *Every finite Boolean algebra has cardinality 2^n, where $n \in \mathbb{N}$.*

Corollary 13.2.11. *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 13.2.12. (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 13.2.13. (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 13.2.14. 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 13.2.15. *$\mathcal{F}$ is a Boolean algebra if and only if $\langle \mathcal{F} \rangle_{\mathcal{B}} = \mathcal{F}$.*

13.2.2 Lattices

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

Definition 13.2.16. 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 13.2.17. *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 13.2.18. 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 13.2.19. *A boolean algebra is a bounded lattice such that meet and join are distributive over each other and such that complements exist.*

13.3 Set Algebras

Definition 13.3.1. 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 13.3.2. *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 13.3.3. *A set algebra is a subalgebra of the power set boolean algebra on X.*

Corollary 13.3.4. *A topology on X is "simply" a set algebra on X that is closed under arbitrary unions.*

Theorem 13.3.5. (Stone's Representation Theorem for Boolean Algebras): *Every boolean algebra is isomorphic to a set algebra.*

13.3.1 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 13.3.6. *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 13.3.7. 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 13.3.8. *The set of subsets of the real numbers which can be written as a finite union of intervals forms a set ring.*

Theorem 13.3.9. *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 13.3.10. *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**.*

13.3.2 Set Semirings

Theorem 13.3.11. *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 13.3.12. *The set $\mathcal{I}$ of real intervals forms a set semiring.*

Theorem 13.3.13. *The product of two set semirings is again a set semiring.*

Corollary 13.3.14. *The set Q of cuboids in $\mathbb{R}^n$ (once again with both open and closed sides allowed) forms a set semiring.*

13.4 σ-Algebras

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

Definition 13.4.0.1: σ-Algebra

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

Corollary 13.4.2. *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 13.4.3. *Every atomic algebra is a σ-algebra.*

Theorem 13.4.4. *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.*

Theorem 13.4.5. *Let $X \neq \emptyset$, let $(Y, \mathcal{B})$ be a measurable space and let $f : X \to Y$ be an arbitrary map. Then*

$$f^{-1}(\mathcal{B}) = \left\{ f^{-1}(B) \,\middle|\, B \in \mathcal{B} \right\}$$

is a σ-algebra.

Note however that we don't necessarily have $f^{-1}(\mathcal{B}) = \mathcal{A}$, or even $f^{-1}(\mathcal{B}) \subset \mathcal{A}$! We call maps that fulfill the latter property *measurable*. Measurable maps play a key role in measure theory and especially in integration theory, entirely analogous to the role continuous functions play in topology. We will study them in more detail in a later chapter.

Theorem 13.4.6. *The intersection of arbitrarily many σ-algebras is on the same set X is once again a σ-algebra.*

Corollary 13.4.6.1: Induced σ-Algebra

Let $\mathcal{E} \subset \mathcal{P}(X)$. Then $\langle \mathcal{E} \rangle_\sigma$ denotes the intersection of all σ-algebras containing $\mathcal{E}$.

In particular, note that if $\mathcal{A}$ is a σ-Algebra containing $\mathcal{E}$, we have

$$\langle \mathcal{E} \rangle_\sigma \subseteq \mathcal{A}.$$

Theorem 13.4.7. *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.*

Theorem 13.4.8. *Let X be any set. Let $\mathcal{E}_i \subset \mathcal{P}(X)$ be a family of sets of subsets of X indexed by $i \in I$. Then we have:*

$$\left\langle \bigcup_{i \in I} \langle \mathcal{E}_i \rangle_\sigma \right\rangle_\sigma = \left\langle \bigcup_{i \in I} \mathcal{E}_i \right\rangle_\sigma$$

13.4.1 The Borel σ-Algebra

Definition 13.4.9. 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 13.4.10. *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 $\mathbb{R}^d$.*

13.5 Dynkin Systems

Definition 13.5.0.1: Dynkin System

Let $\mathcal{D} \subset \mathcal{P}(X)$. Then we call $\mathcal{D}$ a **Dynkin system** iff:

1. $X \in \mathcal{D}$
2. $\mathcal{D}$ is closed under complements $B \setminus A$ as long as $A \subset B$
3. $\mathcal{D}$ is closed under coutable pairwise disjoint union

Theorem 13.5.0.1

A Dynkin system is a σ-algebra iff. it is closed under intersection.

Theorem 13.5.1. *Let $\mathcal{E} \subset \mathcal{P}(X)$ be closed under intersection. Then $\langle \mathcal{E} \rangle_{\mathcal{D}} = \langle \mathcal{E} \rangle_{\sigma}$.*

13.6 Monotone Classes

Definition 13.6.0.1: Monotone classes

Let $\mathcal{M} \subset \mathcal{P}(X)$. Then we call $\mathcal{M}$ a **monotone class** if:

1. For any family $(A_i)_{i\in\mathbb{N}} \subset \mathcal{M}$ such that $A_i \subset A_{i+1}$ holds for all $i \in \mathbb{N}$, we have
$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$$
2. For any family $(A_i)_{i\in\mathbb{N}} \subset \mathcal{M}$ such that $A_i \supset A_{i+1}$ holds for all $i \in \mathbb{N}$, we have
$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{M}$$

We denote the smallest monotone class containing a set $\mathcal{S} \subset \mathcal{P}(X)$ as $\langle \mathcal{S} \rangle_{\mathcal{M}}$.

Corollary 13.6.1. *Every σ-algebra is a monotone class.*

Corollary 13.6.2. *Given $\mathcal{F} \subset \mathcal{P}(X)$, we have $\langle \mathcal{F} \rangle_{\mathcal{M}} \subseteq \langle \mathcal{F} \rangle_{\sigma}$*

Recall that a *set algebra* is a set $\mathcal{A} \subset \mathcal{P}(X)$ that contains the empty set and is closed under complements and binary unions.

Theorem 13.6.3. *Let $\mathcal{F} \subset \mathcal{P}(X)$ be a set algebra and a monotone class. Then $\mathcal{F}$ is a σ-algebra.*

Proof. We need to show that $\mathcal{F}$ is closed under countable unions. Let $(F_j)_{j\in\mathbb{N}} \subset \mathcal{F}$ be a family of elements of $\mathcal{F}$. Since $\mathcal{F}$ is closed under finite union, the increasing sequence of sets

$$A_i := \bigcup_{j=1}^{i} F_j$$

must be contained in $\mathcal{F}$. Therefore, since $\mathcal{F}$ is a monotone class, the union

$$\bigcup_{i\in\mathbb{N}} A_i = \bigcup_{i\in\mathbb{N}} \left(\bigcup_{j=1}^{i} F_j \right) = \bigcup_{j\in\mathbb{N}} F_j$$

must also be contained in $\mathcal{F}$. Therefore, $\mathcal{F}$ is closed under countable union, making it a σ-algebra. □

The relevance of monotone classes is given by the following theorem, which gives us an additional way of verifying whether a given subset of a power set forms a σ-algebra:

Theorem 13.6.3.1: Monotone class theorem

Let $\mathcal{A} \subset \mathcal{P}(X)$ be a set algebra. Then

$$\langle\mathcal{A}\rangle_{\mathcal{M}} = \langle\mathcal{A}\rangle_{\sigma}$$

Proof. We have already observed that, since any σ-algebra is a monotone class, we have $\langle\mathcal{A}\rangle_{\mathcal{M}} \subset \langle\mathcal{A}\rangle_{\sigma}$. Therefore, it suffices to show that we also have $\langle\mathcal{A}\rangle_{\mathcal{M}} \supset \langle\mathcal{A}\rangle_{\sigma}$. This holds automatically if $\langle\mathcal{A}\rangle_{\mathcal{M}}$ is a σ-algebra, and since we already know that any monotone class is a σ-algebra if it is a set algebra, it suffices to show that $\langle\mathcal{A}\rangle_{\mathcal{M}}$ is a set algebra.

For any $E \in \langle\mathcal{A}\rangle_{\mathcal{M}}$, let $\mathcal{M}_E$ be the system of "good subsets" of X which can be combined with E as desired, i.e.

$$\mathcal{M}_E := \left\{ F \subset X \;\middle|\; \begin{array}{c} F \setminus E \in \langle\mathcal{A}\rangle_{\mathcal{M}} \\ E \setminus F \in \langle\mathcal{A}\rangle_{\mathcal{M}} \\ E \cup F \in \langle\mathcal{A}\rangle_{\mathcal{M}} \end{array} \right\}$$

1. First, we want to show that $\mathcal{M}_E$ is a monotone class.
2. Next, we want to show that for every $E \in \mathcal{A}$, we have $\langle\mathcal{A}\rangle_{\mathcal{M}} \subset \mathcal{M}_E$. Since $\mathcal{A}$ is a set algebra, every element of $\mathcal{A}$ combines with E as desired, giving us $\mathcal{A} \subset \mathcal{M}_E$. Therefore, since $\mathcal{M}_E$ is a monotone class, we have

$$\langle\mathcal{A}\rangle_{\mathcal{M}} \subset \langle\mathcal{M}_E\rangle_{\mathcal{M}} = \mathcal{M}_E.$$

3. Next, we want to show that for every $F \in \langle\mathcal{A}\rangle_{\mathcal{M}}$, we have $\langle\mathcal{A}\rangle_{\mathcal{M}} \subset \mathcal{M}_F$. We just showed that $\langle\mathcal{A}\rangle_{\mathcal{M}} \subset \mathcal{M}_F$, which means we have $F \in \mathcal{M}_E$ for any $E \in \mathcal{A}$. Since the definition of $\mathcal{M}_E$ is entirely symmetric in E and F, we also have $E \in \mathcal{M}_F$. Therefore, since all $E \in \mathcal{A}$ are contained in $\mathcal{M}_F$, we have $\mathcal{A} \subset \mathcal{M}_F$, implying

$$\langle\mathcal{A}\rangle_{\mathcal{M}} \subset \langle\mathcal{M}_F\rangle_{\mathcal{M}} = \mathcal{M}_F.$$

4. We have just shown that, for any $E, F \in \langle\mathcal{A}\rangle_{\mathcal{M}}$, we have $E \in \mathcal{M}_F$, implying that $\langle\mathcal{A}\rangle_{\mathcal{M}}$ is closed under complementation and binary union. Since $X \in \mathcal{A}$, we also have $X \in \langle\mathcal{A}\rangle_{\mathcal{M}}$. Therefore, $\langle\mathcal{A}\rangle_{\mathcal{M}}$ is a set algebra.

Therefore, $\langle\mathcal{A}\rangle_{\mathcal{M}}$ is indeed both a set algebra and a monotone class, making it a σ algebra. $\square$

13.7 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 13.7.0.1: Outer Measure

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.

Theorem 13.7.1. *Let X be any set. Then the constant function $\mu = 0$ is an outer measure on X, which is known as the* ***trivial measure****.*

Theorem 13.7.2. *Let X be any set. Let $A \subset X$ and $x \in X$. Then the function*

$$\delta_x(A) := \begin{cases} 1 & x \in A \\ 0 & x \in X \setminus A \end{cases}$$

is an outer measure on X, which is known as the ***Dirac measure****.*

Theorem 13.7.2.1: Counting Measure

Let X be any set. Let $A \subset X$. Then the function

$$\text{card}(A) := \begin{cases} |A| & |A| \text{ is finite} \\ \infty & |A| \text{ is infinite} \end{cases}$$

is an outer measure on X, which is known as the **counting measure**.

Definition 13.7.2.1: Measurable Set

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 $S \subseteq X$ we have

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

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

By this definition, μ-measurable sets are exactly the sets for which μ is σ-additive. This means that if we restrict μ to its measurable sets, we get a proper measure.

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 13.7.2.2: Importance of σ-algebras

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

Since the restriction of an outer measure to its measurable sets forms a proper measure, this theorem tells us that σ-algebras give us the most sensible definition for which properties the domain of a measure should fulfill. Therefore, we finally get a proper definition of a measure:

Definition 13.7.2.2: Measure

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

Definition 13.7.3. If $\mathcal{A}$ is a σ-algebra on X, and μ is a measure on $\mathcal{A}$, we call the triple $(X, \mathcal{A}, \mu)$ a **measure space**.

Definition 13.7.4. Let $(X, \mathcal{A}, \mu)$ be a meausure space. Then we call this space a **probability space**, and μ a **probability measure**, iff $\mu(X) = 1$.

Definition 13.7.4.1: Content, Premeasure

Let $\mathcal{S}$ be a set semiring. Then we call a finitely additive function $\mathcal{S} \to [0, \infty]$ a **content**, and a σ-additive 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 is trivially also a content and a premeasure, and every measure defined on $\mathcal{P}(X)$ is an outer measure.

Definition 13.7.4.2: (σ)-finite measure

Let $(X, \mathcal{A}, \mu)$ be a measure space. Then:

1. We call μ **finite** if $\mu(X) < \infty$.
2. We call μ **σ-finite** if X can be written as a countable union of sets of finite measure.

Theorem 13.7.4.1: Measures are σ-subadditive

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $(A_i)_{i \in \mathbb{N}} \subset \mathcal{A}$ be a family of (not necessarily disjoint!) measurable subsets of X. Then we have:

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leq \sum_{i \in \mathbb{N}} \mu(A_i)$$

Theorem 13.7.4.2: Continuity from above and below

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $(A_i)_{i\in\mathbb{N}} \subset \mathcal{A}$. Then we have:

1. **Continuity from below:** If $A_i \subset A_{i+1}$ for every $i \in \mathbb{N}$, we have
$$\lim_{i\to\infty} \mu(A_i) = \mu\left(\bigcup_{i\in\mathbb{N}} A_i\right)$$
2. **Continuity from above:** If $\mu(A_k) < \infty$ for any $k \in \mathbb{N}$ and $A_i \supset A_{i+1}$ for every $i \in \mathbb{N}$, we have
$$\lim_{i\to\infty} \mu(A_i) = \mu\left(\bigcap_{i\in\mathbb{N}} A_i\right)$$

Definition 13.7.5. Let $(X, \mathcal{A}, \mu)$. We call a set $A \in \mathcal{A}$ a ***μ*-zero set**, or simply zero set, iff $\mu(A) = 0$. We denote the set of all μ-zero sets by $\mathcal{N}(\mu)$.

Definition 13.7.5.1: μ-almost everywhere

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let P be some property. We say that P holds ***μ*-almost everywhere on *M***, or that P holds for ***μ*-almost all *x* ∈ *M***, iff the set of $x \in M$ such that $P(x)$ is false is a zero set.

Definition 13.7.5.2: Complete measure

We call a measure μ **complete** if any subset of a μ-zero set is also a μ-zero set.

Note that, in particular, any subset of a μ-zero set needs to be μ-measurable in the first place.

Lemma 13.7.6. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\mathcal{Z}_\mu$ be the system of all subsets of X that are also subsets of a μ-zero set. Then μ is complete if and only if $\mathcal{Z}_\mu \subset \mathcal{A}$.*

Corollary 13.7.7. *Let μ be an outer measure. Then $\mu|_{\mathcal{M}(\mu)}$ is complete.*

Theorem 13.7.7.1: Completion of a Measure

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let
$$\overline{\mathcal{A}} = \left\{A \cup N \;\middle|\; A \in \mathcal{A}, N \in \mathcal{Z}_\mu\right\}.$$
and
$$\overline{\mu}(A \cup N) := \mu(A).$$
Then $\overline{\mathcal{A}}_\mu$ is a σ-algebra and $\overline{\mu}$ is a complete measure on $\overline{\mathcal{A}}_\mu$.

Theorem 13.7.8. (Uniqueness of Completion): *Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $(X, \mathcal{B}, \nu)$ be a complete measure space such that $\mathcal{A} \subset \mathcal{B}$ and $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$. Then we have $\overline{\mathcal{A}}_\mu \subset \mathcal{B}$ and $\overline{\mu} = \nu$ on $\overline{\mathcal{A}}_\mu$.*

13.8 Carathéodory Extension

We can construct outer measures from a very large class of functions by the following construction:

Theorem 13.8.0.1: 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.

Definition 13.8.0.1: Regular Measure

Let μ be an outer measure on X. Then we call μ **regular** iff. for every $M \subset X$, we can find a μ-measurable set $D \supset M$ such that $\mu(M) = \mu(D)$.

Lemma 13.8.1. *The Carathéodory extension of a premeasure is regular.*

Theorem 13.8.1.1: Characterization of measurable sets

Let $\lambda : \mathbb{R} \to [0, \infty]$ be a <u>σ-finite</u> premeasure on a ring $\mathcal{R} \subset \mathcal{P}(X)$. Let $\overline{\lambda}$ be its Carathéodory extension. Then a set $D \subset X$ is μ measurable if and only if either of the following hold:

1. There exists an $E \supset D$ such that $E \in \langle \mathcal{R} \rangle_\sigma$ and $\overline{\lambda}(E \setminus D) = 0$.
2. There exists a $C \subset D$ such that $C \in \langle \mathcal{R} \rangle_\sigma$ and $\overline{\lambda}(D \setminus C) = 0$.

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

13.9.1 Jordan Content

Definition 13.9.1. Let $E \subset \mathbb{R}^n$ be an elementary set. Then we can assign to it the **elementary volume** $\mathrm{vol}(E)$, 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 13.9.2. (Inner and outer Jordan content): Let $E \subset \mathbb{R}^n$.

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

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

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

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

Definition 13.9.3. (Jordan measurable set, Jordan content): We call E **Jordan-measurable** if $J_*(E) = J^*(E)$. Then we call $J(E) = J_*(E) = J^*(E)$ the **Jordan content**.

Theorem 13.9.4. *The Jordan content is σ-additive.*

Theorem 13.9.5. *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 $\mathrm{vol}(B \setminus A) \leq \varepsilon$,*
3. *For every $\varepsilon > 0$, there exists an elementary set A such that $J^*(A \triangle E) \leq \varepsilon$.*

Theorem 13.9.6. *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 13.9.7. *The Jordan algebra is non-atomic.*

Theorem 13.9.8. (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 13.9.9. *Triangles are Jordan measurable.*

Theorem 13.9.10. *Convex polytopes in $\mathbb{R}^n$ are Jordan measurable.*

Theorem 13.9.11. *Open and closed Euclidean balls are Jordan measurable.*

Theorem 13.9.12. *Every subset of a Jordan null set is Jordan measurable and also a Jordan null set.*

Theorem 13.9.13. *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 13.9.14. *There exist countable unions, and countable intersections, of Jordan measurable sets which are not Jordan measurable.*

13.9.2 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 13.9.15. Let $E \subset \mathbb{R}^n$. The **Lebesque outer measure** of E is given by

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

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

Theorem 13.9.16. *We have* $\lambda^0 = \mathrm{card}$ - *the zero-dimensional lebesque measure is the counting measure.*

Proof. $\mathbb{R}^0$ is by definition cartesian product of $\mathbb{R}$ with a set of cardinality 0 i.e. the empty set, as its index set:

$$\mathbb{R}^0 = \prod_{\emptyset} \mathbb{R}$$

By definition of cartesian products, its members must be functions $\emptyset \to \mathbb{R}$. There is exactly one such "empty function", which we will denote $f_\emptyset$. Thus, the only subsets of $\mathbb{R}^0$ are the empty set and $\{f_\emptyset\}$.

1. Since λ^0 and card are measures, we have

$$\lambda^0(\emptyset) = 0 = \mathrm{card}(\emptyset)$$

2. By definition, a 0-dimensional cuboid must be a function $\emptyset \to I$ to an interval $I \subset \mathbb{R}$. Since all functions from the empty set to any set are identical, $f_\emptyset$ qualifies as a cuboid. We thus have:

$$\begin{aligned} \lambda^0(\{f_\emptyset\}) &= \mathrm{vol}^0(\{f_\emptyset\}) \\ &= \prod_{\emptyset} \mathrm{vol}(I) \\ &= 1 \\ &= \mathrm{card}(\{f_\emptyset\}) \end{aligned}$$

Where the last step follows since empty products evaluate to the multiplicative identity, in this case the number 1.

□

Theorem 13.9.17. *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 13.9.18. 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 13.9.19. *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 13.9.20. *The Lebesque measure defines a σ-additive function on the Lebesque measurable sets.*

Theorem 13.9.21. *The collection of subsets of $\mathbb{R}^n$ that are either Lebesque measurable or have a Lebesque-measurable complement form a Boolean algebra.*

Theorem 13.9.22. *There exist Lebesque measurable sets which are not Borel.*

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 13.9.22.1: Linear Transformation Equation

Let $S \in \mathbb{R}^{n \times n}$. Then for all $E \subset \mathbb{R}^n$, we have:

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

Definition 13.9.22.1: Smith-Volterra-Cantor Set

The Smith-Volterra-Cantor set is the set obtained by removing the middle 1/8 of the interval $[0, 1]$, and then iteratively removing the middle 1/8 of any remaining intervals.

Theorem 13.9.22.2: Compact does not imply Jordan measurable

The Smith-Volterra-Cantor set is compact, but not Jordan measurable.

Chapter 14

Measurable Functions

Definition 14.0.0.1: 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 σ-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.

Corollary 14.0.0.1

The composition of two measurable functions is again measurable.

Lemma 14.0.1. *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 14.0.1.1

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 14.0.2. *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 14.0.2.1: 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 14.0.2.2

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 14.0.3. *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 14.0.4. *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 $\emptyset \in \mathcal{A}$, $X \in \mathcal{A}$, $X \setminus E \in \mathcal{A}$ $\square$

Theorem 14.0.5. *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. $\square$

Corollary 14.0.5.1

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

Definition 14.0.5.1: μ-measurable

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $D \in \mathcal{A}$. Then we call a function $f : D \to \overline{\mathbb{R}}$ μ-measurable iff:

1. $X \setminus D$ is a μ-zero set, and
2. f is $\mathcal{A}|_D$-measurable.

Lemma 14.0.6. *Let $f : D \to \overline{\mathbb{R}}$ be μ-measurable. Then the function*

$$g(x) = \begin{cases} f(x) & x \in D \\ 0 & x \notin D \end{cases}$$

is $\mathcal{A}$-measurable.

Lemma 14.0.6.1

Let $(X, \mathcal{A}, \mu)$ be a **complete** measure space. Let f be μ-measurable. Then every function g such that $g = f$ μ-almost everywhere is μ-measurable.

Theorem 14.0.6.1: μ-measurable pointwise limits

Let $(X, \mathcal{A}, \mu)$ be a **complete** measure space. Let $(f_k)_{k \in \mathbb{N}}$ be a family of μ-measurable functions. Then if f_k converges pointwise to a function f μ-almost everywhere, f is μ-measurable.

Chapter 15

Lebesque Integration

15.1 Step Functions

Definition 15.1.1. 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 15.1.2. *Step functions form a vector space over $\mathbb{R}$.*

Proposition 15.1.3. *Step functions can only take finitely different values.*

Proposition 15.1.4. *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. $\square$

Lemma 15.1.5. (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 15.1.6. (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 15.1.7. (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 15.1.8. (Linearity of the step function integral):

Theorem 15.1.9. (Monotonicity of the step function integral):

15.2 Defining the Lebesque Integral

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

Theorem 15.2.0.1: 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 $k \in \{0, \ldots, n \cdot 2^n\}$, we set:

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

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

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

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,

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

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

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

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

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

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 15.2.0.1: 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 15.2.1. (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 15.2.2. (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 15.2.3. (Integrating over zero sets): *Let N be a set such that $\mu(N) = 0$. Then*

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

Definition 15.2.3.1: 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 15.2.4. (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*

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

Theorem 15.2.4.1: 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.

$\square$

Theorem 15.2.4.2

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 15.2.5. (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 15.2.6. *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 15.2.6.1: Monotone Convergence Theorem

Let $f_n : X \to [0, \infty]$ be μ-measurable such that $f_i \leq f_{i+1}$. 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 15.2.6.2: 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 15.2.7. *Let $f : X \to \overline{\mathbb{R}}$ be μ-measurable. Then f is integrable if and only if $|f|$ is integrable.*

Corollary 15.2.8. *Let $f : X \to \overline{\mathbb{R}}$ be μ-measurable. Then, assuming the integral of f exists, we have*

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

Corollary 15.2.8.1: Dominated Convergence Theorem

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

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

f is integrable.

15.3 Comparing Riemann Integration and Lebesque Integration

Theorem 15.3.1. *The Dirichlet function $\mathbb{1}_{\mathbb{Q}}$ is Lebesque integrable, but not Riemann integrable.*

Theorem 15.3.1.1

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.

Proof. Consider the indicator function $\mathbb{1}_{SVC}$ of the Smith-Volterra-Cantor set. □

Theorem 15.3.1.2

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.

15.4 Convergence Theorems

Theorem 15.4.0.1: Levi's Stronger Monotone Convergence Theorem

The monotone convergence theorem continues to hold if $f_n \leq f_{n+1}$ and $f_n \nearrow f$ only hold μ-almost everywhere.

Corollary 15.4.0.1: Exchange of Integrals and Sums

Let f_n be nonnegative and μ-measurable. Then

$$\int_X \left(\sum_{n=1}^{\infty} f_n\right) d\mu = \sum_{n=1}^{\infty}\left(\int_x f_n \, d\mu\right)$$

Theorem 15.4.0.2: Fatou's Lemma

Let (f_n) be a sequence of μ-measurable functions $X \to \overline{\mathbb{R}}$. Let $g \in \mathcal{L}^1$. Then if we have $f_n \geq g$ μ-almost everywhere for all $n \in \mathbb{N}$, we have

$$\int_X \left(\liminf_{n\to\infty} f_n\right) d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu$$

Theorem 15.4.0.3: Lebesque's Dominated Convergence Theorem

Let f, f_k be μ-measurable functions such that $f_k \to f$ μ-almost everywhere. Assume there exists a nonnegative Function $g \in \mathcal{L}^1$ such that $\sup_k |f_k| \leq g$ μ-almost everywhere. Then we have $f \in \mathcal{L}^1$ and

$$\int_X f \, d\mu = \lim_{k\to\infty} \int_X f_k \, d\mu$$

Lemma 15.4.1. (Continuity of Parametrized Integrals): *Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $U \subset \mathbb{R}^n$ be open. Let $f : U \times X \to \mathbb{R}$. Then*

$$\varphi : U \to \mathbb{R}$$
$$\varphi(y) = \int_X f(y, x) \, d\mu(x)$$

is continuous if the following hold:

1. *$f(\cdot, x)$ is continuous in U for μ-almost all $x \in X$,*
2. *$f(y, \cdot)$ is μ-measurable for all $y \in U$,*
3. *there exists a function $g \in \mathcal{L}^1(\mu)$ such that for all $y \in U$ and μ-almost all $x \in X$, we have*

$$\|f(y, x)\| \leq g(x)$$

Theorem 15.4.1.1: Differentiation under the Integral Sign

Let $(X, \mathcal{A}, \mu)$ be a measure space, $U \subset \mathbb{R}^n$ open, and $f : U \times X \to \mathbb{R}$. Then

$$\varphi(y) := \int_X f(y, x)\, d\mu(x)$$

is continuously differentiable with

$$\frac{\partial}{\partial y_i} \int_X f(y, x)\, d\mu(x) = \int_x \frac{\partial}{\partial y_i} f(y, x)\, d\mu(x)$$

if the following hold:

1. $f(\cdot, x)$ is continuously differentiable in U for almost all x,
2. $f(y, \cdot)$ is μ-measurable for almost all $y \in U$,
3. There exists a function $g \in \mathcal{L}^1(\mu)$ such that

   $$\left| \frac{\partial}{\partial y_i} f(y, x) \right| \leq g(x)$$

   for all $y \in U$, μ-almost all $x \in X$ and $i \in \{1, \ldots, n\}$.

Theorem 15.4.2. (Absolute continuity of the Lebesque integral): *Let $f \in \mathcal{L}^1(\mu)$. Let $\varepsilon > 0$. Then there exists a $\delta > 0$ such that, for all $A \in \mathcal{A}$:*

$$\mu(A) < \delta \implies \int_A |f|\, d\mu < \varepsilon$$

15.5 L^p-Spaces

Definition 15.5.0.1: $\mathcal{L}^p$

Let $1 \leq p < \infty$. We denote by $\mathcal{L}^p = \mathcal{L}^p(X, \mathcal{A}, \mu)$ the set of all μ-measurable functions f such that

$$\int_X |f|^p \, d\mu < \infty$$

Definition 15.5.0.2: L^p **Norm**

Let $1 \leq p < \infty$ and $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$. We call

$$\|f\|_p = \sqrt[p]{\int_X |f|^p \, d\mu}$$

The L^p-Norm, or just p-Norm of f.

Note that, if we view integration as the natural "continuous analogue of summation"[1] and f as a vector with uncountably many components, this is a direct generalization of the familiar p-norm of vectors, such as the euclidean norm

$$\|\vec{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Theorem 15.5.0.1: L^∞**-Norm**

We have

$$\|f\|_\infty := \lim_{p \to \infty} \|f\|_p = \inf_{s>0} \mu\{|f| > s\}$$

Lemma 15.5.1. (Young Inequality): *Let $p, q \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then, for nonnegative $a, b \geq 0$, we have:*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Proof. From the concavity of the logarithm, we get:

$$\ln\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \geq \frac{1}{p} \ln(a^p) + \frac{1}{q} \ln(b^q)$$

Which we can immediately simplify to get:

$$\frac{1}{p} \ln(a^p) + \frac{1}{q} \ln(b^q) = \ln(a) + \ln(b) = \ln(ab)$$

Exponentiating both sides gives us our desired result. □

[1]We can actually formalize this by going the other way - summation is just integration over a discrete space, using the counting measure as the canonical measure.

Comment 15.5.2. Keep in mind that:

$$\begin{aligned}\frac{1}{p} + \frac{1}{q} &= 1\\ \iff \frac{1}{q} &= 1 - \frac{1}{p}\\ &= \frac{p-1}{p}\\ \iff q &= \frac{p}{p-1}\end{aligned}$$

Now, letting $\frac{1}{\infty} = 0$, we get:

Corollary 15.5.2.1: Hölder Inequality

Let $p, q \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Let $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^q$. Then we have $fg \in \mathcal{L}^1$ and, in particular,

$$\left|\int_X f(x) \cdot g(x)\, d\mu(x)\right|$$
$$\leq \|fg\|_1 \leq \|f\|_p \cdot \|g\|_q$$

Proof.

$p = 1$: In this case, we have $q = \infty$. We have

$$g(x) < \inf_{s>0} \{\mu(|g| > s) = 0\} = \|g\|_\infty$$

μ-almost everywhere, and thus

$$\begin{aligned}\int_X |f(x) \cdot g(x)|\, d\mu(x) &\leq \int_X |f(x)| \cdot \|g\|_\infty\, d\mu(x)\\ &= \int_X |f(x)|\, d\mu(x) \cdot \|g\|_\infty\\ &= \|f\|_1 \cdot \|g\|_\infty\end{aligned}$$

$p > 1$: Assume that $\|f\|_p \cdot \|g\|_q > 0$. Then, applying the Young inequality with $a = \dfrac{|f(x)|}{\|f\|_p}$ and $b = \dfrac{|g(x)|}{\|g\|_q}$, we get

$$\begin{aligned}\frac{f(x)g(x)}{\|f\|_p \|g\|_q} &\leq \frac{|f(x)||g(x)|}{\|f\|_p \|g\|_q}\\ &\leq \frac{|f(x)|^p}{p\|f\|_p^p} + \frac{|g(x)|^q}{q\|g\|_q^q}\end{aligned}$$

Integrating both sides gives:

$$\begin{aligned}
\frac{\|fg\|_1}{\|f\|_p\|g\|_q} &= \frac{1}{\|f\|_p\|g\|_q} \cdot \int_X f(x)g(x)\ d\mu(x) \\
&\leq \frac{1}{\|f\|_p\|g\|_q} \cdot \int_X |f(x)||g(x)|\ d\mu(x) \\
&\leq \frac{1}{p\|f\|_p^p} \cdot \int_X |f(x)|^p\ d\mu(x) + \frac{1}{q\|g\|_q^q} \cdot \int_X |g(x)|^q\ d\mu(x) \\
&= \frac{1}{p\|f\|_p^p} \cdot \|f\|_p^p + \frac{1}{q\|g\|_q^q} \cdot \|g\|_q^q \\
&= \frac{1}{p} + \frac{1}{q} \\
&= 1
\end{aligned}$$

Which gives us our desired inequality.

□

Lemma 15.5.3. *$\mathcal{L}^p$ is closed under addition.*

Proof.

1. $p = 1$: Follows from the triangle inequality on $\mathbb{R}$:

$$\begin{aligned}
\|f+g\|_1 &= \int_X |f(x) + g(x)|\ d\mu(x) \\
&\leq \int_X |f(x)| + |g(x)|\ d\mu(x) \\
&= \int_X |f(x)|\ d\mu(x) + \int_X |g(x)|\ d\mu(x) \\
&= \|f\|_1 + \|g\|_1
\end{aligned}$$

2. $p \in (1, \infty)$: Since the function $t \mapsto t^p$ is convex on $[0, \infty)$, we have:

$$\begin{aligned}
|f+g|^p &= 2^p \left|\frac{f+g}{2}\right|^p \\
&\leq 2^{p-1}(|f|^p + |g|^p)
\end{aligned}$$

3. $p = \infty$: We have $|f(x)| \leq \|f\|_\infty$ and $|g(x)| \leq \|g\|_\infty$ for μ-almost all x, which immediately gives us

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty$$

μ-almost everywhere and thus

$$\|f+g\|_\infty = \inf_{s>0} \left\{|f+g| \leq s\ \mu\text{-almost everywhere}\right\} \leq \|f\|_\infty + \|g\|_\infty$$

□

We now want to show that the L^p norm defines a seminorm, which we will then easily be able to turn into a proper norm.

Theorem 15.5.3.1: Minkowski Inequality

Let $f, g \in \mathcal{L}^p$. Then we have

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Proof. We already showed the cases $p = 1$ and $p = \infty$ while proving $\mathcal{L}^p$ is closed under addition. Thus, let $p \in (1, \infty)$. Then the triangle inequality on $\mathbb{R}$ gives:

$$\begin{aligned} |f(x) + g(x)|^p &= |f(x) + g(x)| \cdot |f(x) + g(x)|^{p-1} \\ &\leq (|f(x)| + |g(x)|) \cdot |f(x) + g(x)|^{p-1} \\ &= |f(x)||f(x) + g(x)|^{p-1} + |g(x)||f(x) + g(x)|^{p-1} \end{aligned}$$

Applying the Hölder inequality with $q = \frac{p}{p-1}$ gives:

$$\begin{aligned} &\int_X |f(x)||f(x) + g(x)|^{p-1} \, d\mu(x) \\ &\leq \|f\|_p \cdot \left\| |f + g|^{p-1} \right\|_q \\ &= \|f\|_p \cdot \left(\int_X |f(x) + g(x)|^{(p-1)q} \, d\mu(x) \right)^{\frac{1}{q}} \\ &= \|f\|_p \cdot \left(\int_X |f(x) + g(x)|^p \, d\mu(x) \right)^{\frac{p-1}{p}} \\ &= \|f\|_p \cdot \|f + g\|_p^{p-1} \end{aligned}$$

Switching the roles of f and g similarly gives:

$$\int_X |g(x)||f(x) + g(x)|^{p-1} \, d\mu(x) \leq \|g\|_p \cdot \|f + g\|_p^{p-1}$$

And putting everything together, we get:

$$\begin{aligned} \|f + g\|_p^p &= \int_X |f(x) + g(x)|^p \, d\mu(x) \\ &\leq \int_X |f(x)||f(x) + g(x)|^{p-1} \, d\mu(x) + \int_X |g(x)||f(x) + g(x)|^{p-1} \, d\mu(x) \\ &\leq \|f\|_p \cdot \|f + g\|_p^{p-1} + \|g\|_p \cdot \|f + g\|_p^{p-1} \\ &= \|f + g\|_p \cdot \|f + g\|_p^{p-1} \end{aligned}$$

From which we get our desired result by dividing out the factor of $\|f + g\|_p^{p-1}$. □

Corollary 15.5.3.1

The $\mathcal{L}^p$-norm is a seminorm on $\mathcal{L}^p$. In particular, we have:

1. $\|f\|_p \geq 0$, with $\|f\|_p = 0$ iff. $f = 0$ μ-almost everywhere
2. $\|\lambda f\|_p = |\lambda| \cdot \|f\|_p$
3. $\|f + g\|_p \leq \|f\|_p + \|g\|_p$

Corollary 15.5.3.2: L^p-space

Define L^p-space to be the quotient space of $\mathcal{L}^p$ under μ-almost-everywhere equality. Then $\|\cdot\|_p$ forms a proper norm on L^p.

Lemma 15.5.4. *Let $1 \leq p < \infty$. Let $u_j \in L^p$ and*

$$f_k = \sum_{j=1}^{k} u_j$$

Then, if $\sum_{j=1}^{\infty} \|u_j\|p < \infty$, we have:

1. *The pointwise limit $f(x) = \lim_{k\to\infty} f_k(x)$ exists for μ-almost all $x \in X$,*
2. *$f \in L^p$,*
3. $\|f - f_k\|_p \to 0$

Lemma 15.5.4.1: Convergence in L^p

If we have $f_k \to f$ in L^p Norm for $1 \leq p \leq \infty$, then there exists a subsequence f_{k_i} such that $f_{k_i} \to f$ pointwise μ-almost everywhere.

Theorem 15.5.4.1: Riesz-Fischer

For $1 \leq p \leq \infty$, $(L^p, \|\cdot\|_p)$ is a Banach space.

Definition 15.5.5. Let $1 \leq p < \infty$. Let the sequence $(f_n) \subset L^p$ converge pointwise μ-almost everywhere against the μ-measurable function f. Let

$$\nu(A) := \limsup_{n\to\infty} \int_A |f_n|^p \, d\mu$$

Then we call f_k **uniformly integrable** iff, for any $A \in \mathcal{A}$, we have:

1. $\forall \varepsilon > 0 : \exists \delta > 0 : \forall A \in \mathcal{A} : \mu(A) < \delta \implies \nu(A) < \varepsilon$
2. $\forall \varepsilon > 0 : \exists E \in \mathcal{A} : \mu(E) < \infty, \nu(X \setminus E) < \varepsilon$

Theorem 15.5.5.1: Vitali Convergence Theorem

Let $1 \leq p < \infty$. Let the sequence $(f_n) \subset L^p$ converge pointwise μ-almost everywhere against the μ-measurable function f. Then the following are equivalent:

1. $f \in L^p$ and $\|f_n - f\|_p \to 0$,
2. f_k is uniformly integrable.

Definition 15.5.6. Let $\Omega \subset \mathbb{R}^n$ be open. Then the **support of a function** $f : \Omega \to \mathbb{R}$ is the set

$$\operatorname{spt} f := \overline{\{x \in \Omega : f(x) \neq 0\}}$$

The space of k times continuously differentiable functions with compact support in Ω is denoted $C_c^k(\Omega)$.

Theorem 15.5.6.1: $C_c^0(\Omega)$ is dense in $L^p(\Omega)$

Let $\Omega \subset \mathbb{R}^n$ be open. Let $1 \leq p < \infty$. Then, for every $f \in L^p(\Omega)$, there exists a sequence f_k in $C_c^0(\Omega)$ such that $\|f - f_k\|_p \to 0$.

15.6 Density Functions

Theorem 15.6.0.1

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 15.6.1. *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 15.6.1.1

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 15.6.2. *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 15.6.3. *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 15.6.4. (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 15.6.5. *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 $\ll$. □

Theorem 15.6.5.1: 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 15.6.5.1

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 15.6.5.2: 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 16

Integration in $\mathbb{R}^n$

16.1 Products of Measure Spaces

In this section, we will be interested in finding a canonical σ-Algebra on cartesian products of measurable spaces.

Definition 16.1.0.1: General Products in Category Theory

Let C be a category. Let X_1 and X_2 be objects of C. Then the **product of X_1 and X_2** is an object $X_1 \times X_2$ equipped with two morphisms $\pi_1 : X \to X_1$, $\pi_2 : X \to X_2$ with the property that for every object Y and every pair of morphisms $f_1 : Y \to X_1$, $f_2 : Y \to X_2$, there exists a unique morphism $f : Y \to X_1 \times X_2$ such that:

1. $f_1 = \pi_1 \circ f$
2. $f_2 = \pi_2 \circ f$

We call this property the **universal property of product objects.**

In the case of sets, this definition leads us to the familiar cartesian products, where π_1 and π_2 are the projections to the first and second component of the product respectively. We can generalize the cartesian product to uncountably infinite products as follows:

Definition 16.1.0.2: Cartesian Product of sets

Let I be any set. Let $(X_i)_{i\in I}$ be a Family of nonempty sets indexed by I, i.e. there exists a function assigning to each element $i \in I$ an element of $(X_i)_{i\in I}$. Then the cartesian product of $(X_i)_{i\in I}$ is the unique set:

$$\prod_{i\in I} X_i := \left\{ x : I \to \bigcup_{i\in I} X_i \;\middle|\; x_i := x(i) \in X_i \text{ for all } i \in I \right\}$$

Written more simply: Each element of $\prod_{i\in I} X_i$ is a function assigning to each $i \in I$ an element $x(i) \in X_i$, where we generally write the argument in the index ($x_i := x(i)$), to stay consistent with the established notation in the countable case.

This definition is simply a reinterpretation of the definition of finite cartesian products: It amounts to viewing, for example, the tuple $T := (x, y, z)$ as a map:

$$T : \{1, 2, 3\} \to \{x, y, z\}$$

such that $T(1) = x$, $T(2) = y$, and $T(3) = z$.

Definition 16.1.0.3: Projections

Let $J \subset I$. Then we define the projection of I to J as the map

$$\pi_J := \pi_J^I : \prod_{i \in I} X_i \to \prod_{j \in J} X_j$$
$$x \mapsto \left(x|_J : J \to \bigcup_{j \in J} X_j \right)$$

which restricts the domain of a given element x of the product to the subset J.

In particular, if J is a set consisting of a single element j, we have:

$$\pi_j := \pi_{\{j\}}^I : \prod_{i \in I} X_i \to X_j$$
$$x \mapsto x_j$$

We can now define a product σ-algebra of an indexed family of σ-algebras as the sigma algebra fulfilling the universal property of product objects. We are working in the category of measurable spaces, where morphisms are measurable functions.

Therefore, our definition simply amounts to defining product sigma algebras as the σ-algebras induced by the projection functions, which in turn is the smallest σ-algebra such that all projections are measurable.

Definition 16.1.0.4: Product σ-Algebra

Let $(X_i, \mathcal{A}_i)_{i \in I}$ be a family of measurable spaces. Then the product σ-algebra on the cartestian product $\prod_{i \in I} X_i$ is the smallest σ-Algebra $\bigotimes_{i \in I} \mathcal{A}_i$ such that every projection π_j is $\bigotimes_{i \in I} \mathcal{A}_i - \mathcal{A}_j$-measurable.

Explicitly, this means we have:

$$\bigotimes_{i \in I} \mathcal{A}_i = \left\langle \bigcup_{i \in I} \pi_i^{-1}(\mathcal{A}_i) \right\rangle_\sigma$$

Note that this is also entirely analogous to the definition of the product topology, which is the coarsest (i.e. "smallest") topology such that projections are continuous.

If the cardinality of I is a small finite number, we may sometimes write out the product σ-algebra as:

$$\bigotimes_{i \in I} \mathcal{A}_i = \mathcal{A}_1 \otimes \mathcal{A}_2 \otimes \dots$$

Theorem 16.1.0.1

Let $(X, \mathcal{A})$ and $(Y_i, \mathcal{B}_i)$ be measurable spaces. Let $g : X \to \prod_{i\in I} Y_i$. Then g is $\mathcal{A} - \bigotimes_{i\in I} \mathcal{B}_i$-measurable if and only if all maps $\pi_i \circ g$ are $\mathcal{A} - \mathcal{B}_i$-measurable.

Proof.

1. Assume g is measurable. We know that π_i are measurable by definition, and we know that compositions of measurable functions are measurable. Therefore, $\pi_i \circ g$ are measurable.

2. Assume $\pi_i \circ g$ are measurable. Then we have:

$$\begin{aligned} g^{-1}\left(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{A}_i)\right) &= \bigcup_{i\in I} g^{-1}(\pi_i^{-1}(\mathcal{A}_i)) \\ &= \bigcup_{i\in I} (\pi_i \circ g)^{-1}(\mathcal{A}_i)) \subset \mathcal{A} \end{aligned}$$

Where the last step follows from the measurability of $\pi_i \circ g$. Therefore, since $\bigcup_{i\in I} \pi_i^{-1}(\mathcal{A}_i)$ forms a basis of $\bigotimes_{i\in I} \mathcal{A}_i$, we have that g is measurable.

□

As a corollary, we get that our product space actually fulfills the universal property of product objects:

Corollary 16.1.1. *Let $(X, \mathcal{A})$ and $(Y_i, \mathcal{B}_i)_{i\in I}$ be measurable spaces. Let $g_i : X \to Y_i$ be a family of $\mathcal{A} - \mathcal{B}_i$-measurable functions. Then there exists a $\mathcal{A} - \mathcal{B}_i$-measurable function $g : X \to \prod_{i\in I} Y_i$ such that, for each $i \in I$, $g_i = \pi_i \circ g$.*

Proof. Define g as the cartesian product of all g_i:

$$\begin{aligned} g : I &\to \left(X \to \prod_{i\in I} Y_i\right) \\ i &\mapsto g_i \end{aligned}$$

Then our previous lemma tells us that, since all $g_i = \pi_i \circ g$ are measurable, g is measurable as well. □

Corollary 16.1.2. *Let $(X_i, \mathcal{A}_i)_{i\in I}$ be measurable spaces, and let $J \subset I$ be nonempty. Then the projection π_J is $\bigotimes_{i\in I} \mathcal{A}_i - \bigotimes_{j\in J} \mathcal{A}_j$-measurable.*

Definition 16.1.3. (Embedding into a larger product space): Let $p_j \in X_j$. Then we define the **embedding of $\prod_{i\in I\setminus j} X_i$ into $\prod_{i\in I} X_i$ by x_j** to be the map e_{x_j} which takes a product $(x_i)_{i\in I\setminus j}$ and "adds in x_j as a new component", i.e. it maps $(x_i)_{i\in I\setminus j}$ to the product $(x'_i)_{i\in I}$ with $x'_j = x_j$ and $(x'_i)_{i\in I\setminus j} = (x_i)_{i\in I\setminus j}$.

Definition 16.1.3.1: Cuts of subsets

Let $M \subseteq \prod_{i\in I} X_i$, and let $x_j \in X_j$. Then we define the **cut of M through x_j** to be the set $M^{x_j} = e_{x_j}^{-1}(M)$.

More explicitly, M^{x_j} is the set such that for for every $y \in M$, we have $y|_{I\setminus\{j\}} \in M^{x_j}$ if and only if $y_j = x_j$.

Corollary 16.1.4. *Let $(X_i, \mathcal{A}_i)_{i\in I}$ be measurable spaces, let $M \in \bigotimes_{i\in I} \mathcal{A}_i$, and let $x_j \in X_j$ for $j \in I$. Then we have*

$$M^{x_j} \in \bigotimes_{i\in I\setminus j} \mathcal{A}_i$$

Proof. Let $x_j \in X_j$, $j \in I$. We want to show that e_{x_j} is measurable, which would follow immediately if we could show that all projections $\pi_i^I \circ e_{x_j}$ are measurable.

1. Let $i \in I\setminus \neq j$. Then $\pi_i^I \circ e_{x_j} = \pi_i^{I\setminus j}$, which is measurable.
2. Let $i = j$. Then $\pi_i^I \circ e_{x_j} = \pi_j^I \circ e_{x_j}$, which is just the constant map onto x_j, which is measurable.

Since all projections are measurable, e_{x_j} itself must also be measurable. Therefore, since M is measurable, the preimage $M^{x_j} = e_{x_j}^{-1}(M)$ must also be measurable. □

Note that, in particular, this tells us that if $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$ are measurable spaces, then for any $M \in \mathcal{A}_1 \otimes \mathcal{A}_2$, $x_1 \in X_1$, and $x_2 \in X_2$, we have

$$M^{x_2} = \{x \in X_1 \mid (x, x_2) \in M\} \in \mathcal{A}_1$$

and

$$M^{x_1} = \{x \in X_2 \mid (x_1, x) \in M\} \in \mathcal{A}_2$$

Theorem 16.1.4.1: A basis of the product σ-Algebra

Let $(X_i, \mathcal{A}_i)_{i\in I}$ be a family of measurable spaces. For each $i \in I$, let $\mathcal{E}_i$ be a basis of $\mathcal{A}_i$. Then we have

$$\bigotimes_{i\in I} \mathcal{A}_i = \left\langle \bigcup_{i\in I} \pi_i^{-1}(\mathcal{E}_i) \right\rangle_\sigma .$$

Proof. Recall that, for any map f and any system of subsets $\mathcal{S} \subset \mathcal{A}$ of a σ-algebra $\mathcal{A}$, we have

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

and

$$\left\langle \bigcup_{i\in I} \mathcal{S}_i \right\rangle_\sigma = \left\langle \bigcup_{i\in I} \langle \mathcal{S}_i \rangle_\sigma \right\rangle_\sigma ,$$

which gives us:

$$\begin{aligned}
\bigotimes_{i\in I} \mathcal{A}_i &= \left\langle \bigcup_{i\in I} \pi_i^{-1}(\mathcal{A}_i) \right\rangle_\sigma \\
&= \left\langle \bigcup_{i\in I} \pi_i^{-1}(\langle \mathcal{E}_i \rangle_\sigma) \right\rangle_\sigma \\
&= \left\langle \bigcup_{i\in I} \left\langle \pi_i^{-1}(\mathcal{E}_i) \right\rangle_\sigma \right\rangle_\sigma \\
&= \left\langle \bigcup_{i\in I} \pi_i^{-1}(\mathcal{E}_i) \right\rangle_\sigma
\end{aligned}$$

□

If I is finite, we get the following representation of the product σ-algebra:

Corollary 16.1.5. *For each $i \in \{1, \dots, n\}$, let $(X_i, \mathcal{A}_i)$ be a measurable space such that $\mathcal{E}_i$ is a basis of $\mathcal{A}_i$ and such that X_i can be represented as a countable union of basis elements $E_i^k \in \mathcal{E}_i$. Then*

$$Q_0 := \left\{ \prod_{i=1}^n E_i \,\middle|\, E_i \in \mathcal{E}_i \right\}$$

is a basis of $\bigotimes_{i=1}^n \mathcal{A}_i$.

Proof. We need to show $\langle Q_0 \rangle_\sigma = \bigotimes_{i=1}^n \mathcal{A}_i$.

1. Let $E_i \in \mathcal{E}_i$. We know that

$$\pi_i^{-1}(E_i) = E_i \times \prod_{j\neq i} X_j ,$$

which gives us the following representation for the basis Q we just established in the last theory:

$$Q = \bigcup_{i=1}^n \pi_i^{-1}(\mathcal{E}_i) = \bigcup_{i=1}^n \left\{ E_i \times \prod_{j\neq i} X_j \,\middle|\, E_i \in \mathcal{E}_i \right\}$$

Furthermore, $\cap$-stability of σ-algebras gives us that any element

$$\prod_{i=1}^n E_i = \bigcap_{i=1}^n \pi_i^{-1}(E_i)$$

of Q_0 must be contained in $\langle Q \rangle_\sigma$.

Therefore, we have

$$\langle \mathcal{Q}_0 \rangle_\sigma \subset \langle \mathcal{Q} \rangle_\sigma = \bigotimes_{i=1}^{n} \mathcal{A}_i$$

2. Now, let $E_i \in \mathcal{E}_i$ and $k \in \mathbb{N}$. By definition of $\mathcal{Q}_0$, we have

$$E_i \times \prod_{j \neq i} E_j^k \in \mathcal{Q}_0,$$

which gives us:

$$\pi_i^{-1}(E_i) = \bigcup_{k=1}^{\infty} \left(E_i \times \prod_{j \neq i} E_j^k \right) \in \langle \mathcal{Q}_0 \rangle_\sigma ,$$

and by extension:

$$\mathcal{Q} = \bigcup_{i=1}^{n} \pi_i^{-1}(E_i) \subset \langle \mathcal{Q}_0 \rangle_\sigma$$

which means we also have

$$\bigotimes_{i=1}^{n} \mathcal{A}_i = \langle \mathcal{Q} \rangle_\sigma \subset \sigma(\mathcal{Q}_0).$$

□

Corollary 16.1.6. *The Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ is the product of the Borel σ-algebra $\mathcal{B}(\mathbb{R})$, i.e. we have:*

$$\mathcal{B}(\mathbb{R}^n) = \bigotimes_{i=1}^{n} \mathcal{B}(\mathbb{R})$$

Proof. We know that the set $\mathcal{I}$ of real intervals forms a basis of $\mathcal{B}(\mathbb{R})$. Given this basis, the set $\mathcal{Q}_0$ is exactly the set of n-dimensional real cuboids, which forms a basis of $\mathcal{B}(\mathbb{R}^n)$. Therefore, we have:

$$\mathcal{B}(\mathbb{R}^n) = \langle \mathcal{Q}_0 \rangle_\sigma = \bigotimes_{i=1}^{n} \mathcal{B}(\mathbb{R})$$

□

For the final part of this chapter, we establish some notation for commonly encountered subsets of product σ-algebras.

Definition 16.1.7. Let I be a set. Then we denote by $\mathcal{P}_0(I)$ the finite, nonenmpty subsets of I.

Definition 16.1.7.1: Measurable Cuboids

Let $(X_i, \mathcal{A}_i)_{i\in I}$ be a family of measurable spaces. Then we define the set of **measurable cuboids in $\prod_{i\in I} X_i$** as:

$$Q := \bigcup_{J\in\mathcal{P}_0(I)} \left\{ \prod_{j\in J} A_j \times \prod_{I\setminus J} X_i \,\middle|\, A_j \in \mathcal{A}_j \right\},$$

i.e. a finite number of the sides of our cuboid can be arbitrary measurable sets, and the remaining sides must correspond to the remaining domain sets $(X_i)_{i\in I\setminus J}$.

Definition 16.1.7.2: Cylindrical Sets

Let $(X_i, \mathcal{A}_i)_{i\in I}$ be a family of measurable spaces. Then we define the set of **cylinder sets in $\prod_{i\in I} X_i$** as:

$$\begin{aligned}\mathcal{Z} &:= \bigcup_{J\in\mathcal{P}_0(I)} \left\{ A_J \times \prod_{i\in I\setminus J} X_i \,\middle|\, A_J \in \bigotimes_{j\in J} \mathcal{A}_j \right\} \\ &= \bigcup_{J\in\mathcal{P}_0(I)} \pi_J^{-1}\left(\bigotimes_{j\in J} \mathcal{A}_j\right),\end{aligned}$$

i.e. the base of the cylinder is a measurable set $A_J \subset \prod_{i\in I} X_i$ and all other sides are once again the remaining domain sets $(X_i)_{i\in I\setminus J}$.

Note that if we view an everyday geometric cylinder $Z \subset \mathbb{R}^3$ as the product of of the circle at its base with a finite real interval, it doesn't actually qualify as a cylindrical set in $\mathbb{R}^3$, since it would have to be a product of the base circle with all of $\mathbb{R}$, i.e. only "infinite" cylinders immediately qualify.

By technicality, normal cylinders are still cylindrical sets though, since we can set $J = I$ and have view the entire cylinder as the base, without adding any additional "sides".

Therefore, this definition is not very useful if I is finite, since every measurable set is a cylindrical with itself as its base and every cylindrical set is a product of measurable sets, i.e. measurable.

Theorem 16.1.7.1

Let $(X_i, \mathcal{A}_i)_{i\in I}$ be a family of measurable spaces. Then the measurable cuboids in $\prod_{i\in I} X_i$ and the cylindrical sets in $\prod_{i\in I} X_i$ form bases of $\bigotimes_{i\in I} \mathcal{A}_i$.

Proof. We will show $Q \subset \mathcal{Z} \subset \bigotimes_{i\in I} \mathcal{A}_i \subset \langle Q\rangle_\sigma$, which implies $\langle Q\rangle_\sigma \subset \langle \mathcal{Z}\rangle_\sigma \subset \bigotimes_{i\in I} \mathcal{A}_i \subset \langle Q\rangle_\sigma$, proving the theorem.

1. $\mathcal{Q} \subset \mathcal{Z}$: Let $A \in \mathcal{Q}$. Then A is given as

$$A = \prod_{j\in J}^{n} A_j \times \prod_{i\notin J} X_i$$
$$= \pi_J^{-1}\left(\prod_{j\in J} A_j\right).$$

   Therefore, since $\prod_{j\in J} A_j \in \bigotimes_{j\in J} \mathcal{A}_j$, A is a cylinder set.

2. $\mathcal{Z} \subset \bigotimes_{i\in I} \mathcal{A}_i$: Recall that π_J is $\bigotimes_{i\in I} \mathcal{A}_i - \mathcal{A}_J$-measurable. Therefore, for any $J \in \mathcal{P}_0(I)$, we have that $\pi_J^{-1}(\mathcal{A}_j) \subset \bigotimes_{i\in I} \mathcal{A}_i$, which means any cylinder set is contained in $\bigotimes_{i\in I} \mathcal{A}_i$.

3. $\bigotimes_{i\in I} \mathcal{A}_i \subset \langle \mathcal{Q} \rangle_\sigma$: For any $i \in I$ and $A_i \in \mathcal{A}_i$, we have:

$$\pi_i^{-1}(A_i) = A_i \times \prod_{j\neq i} X_j \in \mathcal{Q},$$

   i.e. $\pi_i^{-1}(\mathcal{A}_i) \subset \mathcal{Q}$ for all $i \in I$, which means we also have:

$$\bigcup_{i\in I} \pi_i^{-1}(\mathcal{A}_i) \subset \mathcal{Q}.$$

   By definition of $\bigotimes_{i\in I} \mathcal{A}_i$, we get:

$$\bigotimes_{i\in I} \mathcal{A}_i = \left\langle \bigcup_{i\in I} \pi_i^{-1}(\mathcal{A}_i) \right\rangle_\sigma \subset \langle \mathcal{Q} \rangle_\sigma .$$

$\square$

16.2 Product Measures and Fubini's Theorem

In this section, given two measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$, we want to derive a canonical measure $\mu \otimes \nu$ on $\mathcal{A} \otimes \mathcal{B}$. Our intuition already tells us what the measure of a cuboid should be: Given a cuboid

$$Q = A \times B \subset X \times Y,$$

where $A \in \mathcal{A}$ and $B \in \mathcal{B}$, we want the measure of the cuboid to be the product of the measures of its sides:

$$(\mu \otimes \nu)(M) = \mu(A) \cdot \nu(B)$$

Recall that, for any $M \in X \times Y$, the cuts through M by $x \in X$ and $y \in Y$ are simply defined as

$$M^x = \{y \in Y \mid (x, y) \in M\},$$
$$M^y = \{x \in X \mid (x, y) \in M\}.$$

Theorem 16.2.0.1

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. Let $M \in \mathcal{A} \otimes \mathcal{B}$. Then:

1. The function $x \mapsto \nu(M^x)$ is $\mathcal{A}$-measurable,
2. The function $y \mapsto \mu(M^y)$ is $\mathcal{B}$-measurable,
3. We have

$$\int_X \nu(M^x)\, d\mu(x) = \int_Y \mu(M^y)\, d\nu(y)$$

Proof. Let $\mathcal{S}$ be the system of all sets $M \in \mathcal{A} \otimes \mathcal{B}$ for which this theorem holds.

Let $\mathcal{Q}$ be the system of all measurable cuboids in $X \times Y$, which we have just established are simply given by $Q = A \times B$ such that $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

Let $\mathcal{F}$ be the system of finite unions of cuboids (which by our definition of cuboids inherently includes complements of cuboids). Since we count $X \times Y$ as a cuboid, $\mathcal{F}$ forms a set algebra.

We will show that $\mathcal{S}$ is a monotone class containing $\mathcal{F}$, which will give us

$$\mathcal{A} \otimes \mathcal{B} = \langle \mathcal{Q} \rangle_\sigma = \langle \mathcal{F} \rangle_\sigma = \langle \mathcal{F} \rangle_M \subseteq \mathcal{S}$$

since $\mathcal{S} \subseteq \mathcal{A} \otimes \mathcal{B}$ by definition, this implies $\mathcal{A} \otimes \mathcal{B} = \mathcal{S}$.

The proof the $\mathcal{S}$ is a monotone class is however left as an exercise to the reader :) $\square$

Theorem 16.2.0.2: Product measure

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu$ be σ-finite measure spaces. Then there exists exactly one measure $\mu \otimes \nu$ on $\mathcal{A} \otimes \mathcal{B}$ such that for any $A \in \mathcal{A}$ and $B \in \mathcal{B}$, we have

$$(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$$

This measure is given by

$$(\mu \otimes \nu)(M) = \int_X \nu(M^x)\, d\mu(x) = \int_Y \mu(M^y)\, d\mu(y)$$

Corollary 16.2.0.1: Cavalieri's Principle

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. Then, for all $M \in \mathcal{A} \otimes \mathcal{B}$, we have

$$\begin{aligned}(\mu \otimes \nu)(M) &= \int_X \nu(M^x)\, d\mu(x) \\ &= \int_X \int_Y \chi_M(x, y)\, d\nu(y)\, d\mu(x) \\ &= \int_Y \mu(M^Y)\, d\nu(x) \\ &= \int_Y \int_X \chi_M(x, y)\, d\mu(x)\, d\nu(y)\end{aligned}$$

Corollary 16.2.1. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. Then, for any $M \in \mathcal{A} \otimes \mathcal{B}$, the following are equivalent:*

1. $(\mu \otimes \nu)(M) = 0$,
2. $\mu(M^y) = 0$ *ν-almost everywhere,*
3. $\nu(M^x) = 0$ *μ-almost everywhere.*

Lemma 16.2.2. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. Let $f : X \times Y \to \mathbb{R}$ be $\mathcal{A} \otimes \mathcal{B}$ measurable. Then the function $f(x, \cdot) : Y \to \mathbb{R}$ is $\mathcal{B}$-measurable for all $x \in X$.*

Theorem 16.2.2.1: Fubini's Theorem

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. Let $f \in \mathcal{L}^*(\mu \otimes \nu)$. Then:

$$\begin{aligned}\int_{X \times Y} f\, d(\mu \otimes \nu) &= \int_X \int_Y f(x, y)\, d\nu(y)\, d\mu(x) \\ &= \int_Y \int_X f(x, y)\, d\mu(x)\, d\nu(y)\end{aligned}$$

Corollary 16.2.3. *The volume of the open ball*

$$B_r^n(0) = \left\{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\|_2 < r\right\}.$$

is given by:

$$\lambda^n(B_r^n(0)) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} \cdot r^n = \begin{cases} \frac{\pi^k}{k!} \cdot r^n & n = 2k \\ \frac{\pi^k}{\prod_{i=1}^k \frac{1}{2}+k-i} \cdot r^n & n = 2k+1 \end{cases}$$

In particular:

1. The 1-dimensional unit ball is simply the interval $(-1, 1)$ and thus has length 2,
2. The 2-dimensional unit ball is a circle of radius 1 and has area π
3. The 3-dimensional unit ball has volume $\frac{4\pi}{3}$
4. The 4-dimensional unit ball has volume $\frac{\pi^2}{2}$
5. The 5-dimensional unit ball has volume $\frac{8\pi^2}{15}$
6. ...

Proof. By Cavalieri's principle and the linear transformation formula for the Lebesque measure, we get:

$$\begin{aligned}
\lambda^n(B_r^n(0)) &= \int_{\mathbb{R}} \lambda^{n-1}(B_r^n(0)^y)\, dy \\
&= \int_{(-r,r)} \lambda^{n-1}(B_r^n(0)^y)\, dy \\
&= \int_{-1}^{1} = \det\left(\sqrt{r^2 - y^2} \cdot I_n\right) \cdot \lambda^{n-1}\left\{x \in \mathbb{R}^{n-1} \mid \|x\|_2 < 1\right\} dy \\
&= r^n \cdot \int_{-1}^{1} \left(\sqrt{r^2 - y^2}\right)^{n-1} \cdot \lambda^{n-1}\left\{x \in \mathbb{R}^{n-1} \mid \|x\|_2 < 1\right\} dy \\
&= r^n \cdot \lambda^{n-1}(B_1^{n-1}(0)) \cdot \int_{-1}^{1} \left(\sqrt{1 - y^2}\right)^{n-1} dy
\end{aligned}$$

We can thus get our result by solving the integral

$$I_n := \int_{-1}^{1} \left(\sqrt{1 - y^2}\right)^{n-1} dy$$

Substituting $\cos(\theta) = y$ gives:

$$\begin{aligned}
I_n &= \int_{-1}^{1} \left(\sqrt{1 - y^2}\right)^{n-1} dy \\
&= \int_{-1}^{1} \left(\sqrt{1 - \cos(\theta)^2}\right)^{n-1} \cdot \sin(\theta)\, d\theta \\
&= \int_{0}^{\pi} \sin^n(\theta)\, d\theta
\end{aligned}$$

Now, partial integration gives:

$$\begin{aligned}
I_n &= \int_0^\pi \sin^n(\theta)\, d\theta \\
&= \int_0^\pi \sin(\theta) \cdot \sin^{n-1}(\theta)\, d\theta \\
&= \left[-\cos(\theta) \cdot \sin^{n-1}(\theta)\right]_0^\pi + (n-1) \cdot \int_0^\pi \cos^2(\theta) \cdot \sin^{n-2}(\theta)\, d\theta \\
&= (n-1) \cdot \int_0^\pi \cos^2(\theta) \cdot \sin^{n-2}(\theta)\, d\theta \\
&= (n-1) \cdot \int_0^\pi (1-\sin^2) \cdot \sin^{n-2}(\theta)\, d\theta \\
&= (n-1) \cdot (I_{n-2} - I_n),
\end{aligned}$$

which means:

$$\begin{aligned}
I_n &= (n-1) \cdot I_{n-2} - (n-1) \cdot I_n \\
\implies nI_n &= (n-1) \cdot I_{n-2} \\
\implies I_n &= \frac{n-1}{n} \cdot I_{n-2}
\end{aligned}$$

Unrolling this recursive formula, we get:

$$\begin{aligned}
I_{2k} &= I_0 \cdot \prod_{i=1}^{n} \frac{2i-1}{2i} \\
I_{2k+1} &= I_1 \cdot \prod_{i=1}^{n} \frac{2i}{2i+1}
\end{aligned}$$

We can easily calculate I_0 and I_1 explicitly:

$$\begin{aligned}
I_0 &= \int_0^\pi \sin^0(\theta)\, d\theta \\
&= \int_0^\pi 1\, d\theta \\
&= \pi
\end{aligned}$$

$$\begin{aligned}
I_1 &= \int_0^\pi \sin(\theta)\, d\theta \\
&= (-\cos(\pi)) - (-\cos(0)) \\
&= 2
\end{aligned}$$

Thus, we have:

$$\begin{aligned}
I_{2k} &= \pi \cdot \prod_{i=1}^{k} \frac{2i-1}{2i} \\
I_{2k+1} &= 2 \cdot \prod_{i=1}^{k} \frac{2i}{2i+1}
\end{aligned}$$

This gives us:

$$\begin{aligned}
I_{2k+1}I_{2k} &= 2\pi \cdot \prod_{i=1}^{k} \frac{2i-1}{2i} \cdot \prod_{i=1}^{k} \frac{2i}{2i+1} \\
&= 2\pi \cdot \prod_{i=1}^{k} \frac{2i-1}{2i+1} \\
&= \frac{2\pi}{2k+1} \\
&= \frac{\pi}{k+\frac{1}{2}}
\end{aligned}$$

$$\begin{aligned}
I_{2k}I_{2k-1} &= 2\pi \cdot \prod_{i=1}^{k} \frac{2i-1}{2i} \cdot \prod_{i=1}^{k-1} \frac{2i}{2i+1} \\
&= 2\pi \cdot \frac{2k-1}{2k} \cdot \prod_{i=1}^{k-1} \frac{2i-1}{2i} \cdot \prod_{i=1}^{k-1} \frac{2i}{2i+1} \\
&= 2\pi \cdot \frac{2k-1}{2k} \cdot \frac{1}{2k-1} \\
&= \frac{\pi}{k}
\end{aligned}$$

We know that

$$\lambda(B_1^1(0)) = \lambda((-1,1)) = 2$$

And, recalling the fact that λ^0 is the counting measure, we also know that

$$\lambda^0(B_1^0(0)) = \lambda^0(\{()\}) = 1$$

Thus, at long last, we can calculate the volumes of our unit balls:

$$\begin{aligned}
\lambda^{2k}(B_1^{2k}(0)) &= (I_{2k}I_{2k-1}) \cdot (I_{2k-2}I_{2k-3}) \cdot \ldots \cdot (I_2 I_1) \cdot \lambda^0(B^0(0)) \\
&= \left(\prod_{i=1}^{k} \frac{\pi}{i}\right) \cdot 1 \\
&= \frac{\pi^k}{k!} = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}
\end{aligned}$$

$$
\begin{aligned}
\lambda^{2k+1}(B_1^{2k+1}(0)) &= (I_{2k+1}I_{2k}) \cdot (I_{2k-1}I_{2k-2}) \cdot \ldots \cdot (I_3I_2) \cdot \lambda^1(B^1(0)) \\
&= \left(\prod_{i=1}^{k} \frac{\pi}{i + \frac{1}{2}} \right) \cdot 2 \\
&= \frac{\pi^k}{\prod_{i=1}^{k} i + \frac{1}{2}} \cdot 2 \\
&= \frac{\pi^k}{\prod_{i=0}^{k} i + \frac{1}{2}} \\
&= \frac{\pi^k \cdot \sqrt{\pi}}{\left(\prod_{i=0}^{k} i + \frac{1}{2}\right) \cdot \sqrt{\pi}} \\
&= \frac{\pi^{\frac{n}{2}}}{\left(\prod_{i=0}^{k} i + \frac{1}{2}\right) \cdot \Gamma\left(\frac{1}{2}\right)} \\
&= \frac{\pi^{\frac{n}{2}}}{\left(\prod_{i=1}^{k+1} i - \frac{1}{2}\right) \cdot \Gamma\left(\frac{1}{2}\right)} \\
&= \frac{\pi^{\frac{n}{2}}}{\Gamma\left(k + \frac{1}{2} + 1\right)} \\
&= \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)}
\end{aligned}
$$

Which gives us

$$\lambda^n(B_r^n(0)) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} \cdot r^n$$

as desired. □

Proposition 16.2.4. *The* 5*-dimensional unit ball has the highest volume out of any unit ball. As the dimension increases, the volume of the unit ball converges to zero.*

Proof Sketch. Plugging small natural numbers into our formula, we see that the formula is monotonically increasing for $n \leq 5$ and monotonically decreasing for $n \geq 5$. Furthermore, we have $\Gamma(\frac{n}{2} + 1) \geq \Gamma(k + 1) = k!$, which grows asymptotically faster than $\pi^{n/2} \leq \sqrt{\pi} \cdot \pi^k$. "□"

Lemma 16.2.4.1

Let $j \in \{1, \ldots, n\}$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be continously differentiable with $f, \partial_j f \in L^1(\mathbb{R}^n)$. Then

$$\int_{\mathbb{R}^n} \partial_j f(x)\, d\lambda^n(x) = 0$$

Proof. 1. First, assume there is an $R > 0$ such that $f(x) = 0$ for all $\|x\|_2 \geq R$. Let $x = (y, x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$. Then by Fubini we get:

$$\begin{aligned}\int_{\mathbb{R}^n} \partial_j f(x)\, d\lambda^n(x) &= \int_{\mathbb{R}^n} \partial_j f(y, x_n)\, d\lambda^n(y, x_n) \\ &= \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \partial_j f(y, x_n)\, d\lambda^1(x_n)\, d\lambda^{n-1}(y).\end{aligned}$$

For $j = n$, we apply the fundamental theorem of calculus and our assumption that $f(x) = 0$ for all $\|x\|_2 \geq R$ and get:

$$\begin{aligned}\int_{\mathbb{R}} \partial_j f(y, x_n)\, d\lambda^1(x_n) &= \lim_{x_n \to \infty} f(y, x_n) - f(y, -x_n) \\ &= 0\end{aligned}$$

By induction, the same holds for all j.

2. Now, let $\varphi : \mathbb{R} \to \mathbb{R}$ be a smooth function fulfilling:

$$\varphi(t) = \begin{cases} 1 & |t| \leq 1, \\ 0 & |t| \geq 2. \end{cases}$$

And let

$$\begin{aligned}\varphi_R &: \mathbb{R}^n \to \mathbb{R} \\ \varphi_R(x) &= \varphi\left(\frac{\|x\|_2}{R}\right)\end{aligned}$$

Then we have

$$(\varphi_R f)(x) = 0$$

for all $\|x\|_2 \geq 2R$, meaning that by the product rule of differentiation and step 1 we get:

$$\begin{aligned}\int_{\mathbb{R}^n} \varphi_R(x) \cdot \partial_j f(x)\, d\lambda^n(x) + \int_{\mathbb{R}^n} \partial_j \varphi_R(x) \cdot f(x)\, d\lambda^n(x) \\ = \int_{\mathbb{R}^n} \partial_j(\varphi_R(x) \cdot f(x))\, d\lambda^n(x) \\ = 0,\end{aligned}$$

i.e.

$$\int_{\mathbb{R}^n} \varphi_R(x) \cdot \partial_j f(x)\, d\lambda^n(x) = -\int_{\mathbb{R}^n} \partial_j \varphi_R(x) \cdot f(x)\, d\lambda^n(x)$$

Now, notice we also have $\varphi_R(x) = 1$ for $\|x\| \leq R$ - in particular, we have $\varphi_R \to 1$ and $\partial_j \varphi_R \to 0$ pointwise on $\mathbb{R}^n$ for $R \to \infty$, which means we have

$$\int_{\mathbb{R}^n} \partial_j f(x)\, d\lambda^n(x) = \int_{\mathbb{R}^n} \lim_{R \to \infty} \varphi_R(x) \cdot \partial_j f(x)\, d\lambda^n(x)$$

and

$$\int_{\mathbb{R}^n} \lim_{R \to \infty} \partial_j \varphi_R(x) \cdot f(x)\, d\lambda^n(x) = 0$$

Finally, we also have:

$$\left|\varphi_R \cdot (\partial_j f)\right| \leq \left|\partial_j f\right| \in L^1(\mathbb{R}^n)$$

and

$$\begin{aligned}\left|(\partial_j \varphi_R) \cdot f\right| &\leq \frac{\max|\varphi'|}{R} \cdot |f| \\ &\leq (\max|\varphi'|) \cdot |f| \in L^1(\mathbb{R}^n)\end{aligned}$$

Thus, we can apply Lebesque's dominated convergence theorem to pull out the limits and connect the two equalities, giving us:

$$\begin{aligned}\int_{\mathbb{R}^n} \partial_j f(x)\, d\lambda^n(x) &= \int_{\mathbb{R}^n} \lim_{R\to\infty} \varphi_R(x) \cdot \partial_j f(x)\, d\lambda^n(x) \\ &= \lim_{R\to\infty} \int_{\mathbb{R}^n} \varphi_R(x) \cdot \partial_j f(x)\, d\lambda^n(x) \\ &= -\lim_{R\to\infty} \int_{\mathbb{R}^n} \partial_j \varphi_R(x) \cdot f(x)\, d\lambda^n(x) \\ &= -\int_{\mathbb{R}^n} \lim_{R\to\infty} \partial_j \varphi_R(x) \cdot f(x)\, d\lambda^n(x) \\ &= 0\end{aligned}$$

□

Corollary 16.2.4.1: Partial Integration-ish Formula

Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable such that $f, \partial_j f \in L^p(\mathbb{R}^n)$ and $g, \partial_j g \in L^q(\mathbb{R}^n)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then we have

$$\int_{\mathbb{R}^n} (\partial_j f(x)) \cdot g(x)\, d\lambda^n(x) = 0 - \int_{\mathbb{R}^n} f(x) \cdot (\partial_j g(x))\, d\lambda^n(x)$$

Proof. Applying the product rule of differentiation, and then the last lemma, we get:

$$\begin{aligned}\int_{\mathbb{R}^n} (\partial_j f(x)) \cdot g(x)\, d\lambda^n(x) + \int_{\mathbb{R}^n} f(x) \cdot (\partial_j g(x))\, d\lambda^n(x) \\ = \int_{\mathbb{R}^n} \partial_j (f(x) \cdot g(x))\, d\lambda^n(x) \\ = 0\end{aligned}$$

□

16.3 Change of Variables

Lemma 16.3.1. *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:

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

For sufficiently large j, these imply:

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

Further applying the triangle inequality, we get:

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

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:

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

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

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

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 $\lambda^n(S(E)) = |\det(S)|\lambda^n(E)$ (13.9.2), we have:

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

□

Theorem 16.3.1.1: 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 16.3.1.1

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 $f = \mathbb{1}_{\varphi(A)}$. □

Definition 16.3.1.1: Gram Matrix of a Diffeomorphism

Let $\varphi : U \to V$ be a C^k-diffeomorphism between open subsets of $\mathbb{R}^n$. Then the **Gram matrix** $g(\varphi) \in C^{k-1}(U, \mathbb{R}^{n\times n})$ is given by:

$$g(\varphi)(x) = D\varphi(x)^\top D\varphi(x),$$

or equivalently:

$$g(\varphi)_{ij}(x) = \left\langle \frac{\partial}{\partial x_i}\varphi(x), \frac{\partial}{\partial x_j}\varphi(x) \right\rangle$$

Since $g(\varphi)(x)$ is the product of a matrix with its own transpose, it is symmetric and strictly positive definite. This lets us alternatively write the multivariable substitution formula as:

$$\int_V f(y)dy = \int_{\varphi^{-1}(V)} f(\varphi(x))\sqrt{\det g(\varphi)(x)}\, dx$$

16.4 Alternative coordinate systems on $\mathbb{R}^2$ and $\mathbb{R}^3$

16.4.1 Polar Coordinates

A commonly used alternative coordinate system on $\mathbb{R}^2$ is given by **polar coordinates**, which describes a point $x \in \mathbb{R}^2$ based on its distance r from the origin and the *polar angle* ω between x and $(0, r)^\top$. These might be familiar from dealing with complex numbers, where any $z \in \mathbb{C}$ can be represented as

$$z = r \cdot e^{\omega i}$$

Definition 16.4.0.1: Polar Coordinates on $\mathbb{R}^2$

Let $U = (0, \infty) \times (0, 2\pi)$ and $V = \mathbb{R}^2 \setminus \{(x, 0)^\top : x \geq 0\}$. The **polar coordinate transformation** is the map:

$$\begin{aligned} \varphi : U &\to V \\ \begin{pmatrix} r \\ \omega \end{pmatrix} &\mapsto \begin{pmatrix} r \cos \omega \\ r \sin \omega \end{pmatrix} \end{aligned}$$

which maps a point given in polar coordinates to its corresponding point in the cartesian coordinate system.

The Jacobian of the polar coordinate transform is:

$$\begin{aligned} D\varphi(r, \omega) &= \begin{pmatrix} \frac{\partial}{\partial r} r \cos \omega & \frac{\partial}{\partial \omega} r \cos \omega \\ \frac{\partial}{\partial r} r \sin \omega & \frac{\partial}{\partial \omega} r \sin \omega \end{pmatrix} \\ &= \begin{pmatrix} \cos \omega & -r \sin \omega \\ \sin \omega & r \cos \omega \end{pmatrix} \end{aligned}$$

Which means the determinant is:

$$\begin{aligned} \det D\varphi(r, \omega) &= \det \begin{pmatrix} \cos \omega & -r \sin \omega \\ \sin \omega & r \cos \omega \end{pmatrix} \\ &= r \cos(\omega)^2 + r \sin(\omega)^2 \\ &= r \end{aligned}$$

Corollary 16.4.0.1: Polar Coordinate Integral Transform

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be integrable. Then we have:

$$\int_{\mathbb{R}^2} f(x, y)\, d\lambda(x, y) = \int_0^{2\pi} \int_0^{\infty} r \cdot f(r \cos(\omega), r \sin(\omega))\, dr\, d\omega$$

Proposition 16.4.1. *The inverse of the polar coordinate transform φ is:*

$$\varphi^{-1}(x, y) = \begin{cases} \left(r,\ \arccos\left(\frac{x}{r}\right)\right)^\top & y \geq 0 \\ \left(r,\ 2\pi - \arccos\left(\frac{x}{r}\right)\right)^\top & y < 0 \end{cases}$$

where $r = \sqrt{x^2 + y^2}$.

Example 16.4.2. (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:

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

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

Example 16.4.3. (Beta function): For $p, q \in \mathbb{R}_{>0}$, the beta function B is defined as:

$$B(p,q) = \int_0^1 t^{p-1} \cdot (1-t)^{q-1}\, dt$$

The beta function is of interest primarily because of its close relationship to the gamma function:

Theorem 16.4.4. *We have:*

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

Proof. Substituting $t = g(\omega) := \sin(\omega)^2$ into the definition of the beta function gives us:

$$\begin{aligned} g'(\omega) &= \frac{\mathrm{d}}{\mathrm{d}\omega} \sin(\omega)^2 \\ &= \sin(\omega)\cos(\omega) + \cos(\omega) + \sin(\omega) \\ &= 2\sin(\omega)\cos(\omega) \end{aligned}$$

and:

$$1 = \sin(\omega)^2 \Leftrightarrow \omega = \frac{\pi}{2},$$

which means we have:

$$\begin{aligned} B(p,q) &= \int_0^1 t^{p-1} \cdot (1-t)^{q-1}\, dt \\ &= \int_0^{\pi/2} \sin(\omega)^{2p-2} \cdot (1-\sin(\omega)^2)^{q-1} \cdot 2\sin(\omega)\cos(\omega)\, d\omega \\ &= \int_0^{\pi/2} \sin(\omega)^{2p-2} \cdot \cos(\omega)^{2q-2} \cdot 2\sin(\omega)\cos(\omega)\, d\omega \\ &= 2 \cdot \int_0^{\pi/2} \sin(\omega)^{2p-1} \cdot \cos(\omega)^{2q-1}\, d\omega \end{aligned}$$

Meanwhile, performing a reverse substitution with $t := g(r) = r^2$ in the definition of the gamma function gives:

$$\begin{aligned} \Gamma(m) &= \int_0^\infty t^{m-1} e^{-t}\, dt \\ &= \int_0^\infty (r^2)^{m-1} e^{-r^2} \cdot 2r\, dr \\ &= 2\int_0^\infty r^{2m-1} e^{-r^2}\, dr \end{aligned}$$

Now, we get:

$$\begin{aligned} B(p,q)\Gamma(p+q) &= 2\int_0^\infty B(p,q) \cdot r^{2p+2q-1} \cdot e^{-r^2}\, dr \\ &= 4\int_0^\infty \int_0^{\pi/2} (\sin\omega)^{2p-1} \cdot r^{2p-1} \cdot (\cos\omega)^{2q-1} \cdot r^{2q-1} \cdot r \cdot e^{-r^2}\, d\omega\, dr \\ &= 4\int_0^\infty \int_0^{\pi/2} (r\sin\omega)^{2p-1} \cdot (r\cos\omega)^{2q-1} \cdot r \cdot e^{-r^2}\, d\omega\, dr \end{aligned}$$

By applying the polar coordinate integral transform in reverse, we get:

$$\begin{aligned}
B(p,q)\Gamma(p+q) &= 4\int_0^{\infty}\int_0^{\pi/2}(r\sin\omega)^{2p-1}\cdot(r\cos\omega)^{2q-1}\cdot r\cdot e^{-r^2}\,d\omega\,dr \\
&= 4\int_0^{\infty}\int_0^{\infty} x^{2p-1}\cdot y^{2q-1}\cdot e^{-\sqrt{x^2+y^2}^2}\,dy\,dx \\
&= 4\int_0^{\infty}\int_0^{\infty} x^{2p-1}\cdot y^{2q-1}\cdot e^{-x^2-y^2}\,dy\,dx \\
&= \left(2\int_0^{\infty} x^{2p-1}e^{-x^2}\,dx\right)\cdot\left(2\int_0^{\infty} y^{2q-1}e^{-y^2}\,dy\right) \\
&= \Gamma(p)\Gamma(q)
\end{aligned}$$

□

16.4.2 Spherical Coordinates

The most natural generalization of polar coordinates to $\mathbb{R}^3$ are *spherical coordinates*, in which $\theta \in (0, \pi)$ describes the *polar angle* between a point P and the z-Axis and $\omega \in (0, 2\pi)$ describes the *azimuthal angle* between the x-axis and the projection P' of P onto the xy-plane. We can derive spherical coordinates from two dimensional coordinates as follows:

1. Take a point $P = (x, y, z) \in \mathbb{R}^3$. Notice that $z = r\cos(\theta)$.
2. Project P to a point $P' = (x, y)$. The distance of P' from $(0, 0)$ is exactly $\rho := r\sin\theta$. Therefore, P' has polar coordinates (ρ, ω), which gives us:

$$P = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \rho\cos\omega \\ \rho\sin\omega \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} r\sin\theta\cos\omega \\ r\sin\theta\sin\omega \\ r\cos\theta \end{pmatrix}$$

Definition 16.4.4.1: Spherical Coordinates on $\mathbb{R}^3$

Let $U = (0, \infty) \times (0, \pi) \times (0, 2\pi)$ and $V = \mathbb{R}^3 \setminus \{(x, 0, z)^\top : x \geq 0\}$. The **spherical coordinate transformation** is the map:

$$\begin{aligned} \varphi : U &\to V \\ \begin{pmatrix} r \\ \theta \\ \omega \end{pmatrix} &\mapsto \begin{pmatrix} r\sin\theta\cos\omega \\ r\sin\theta\sin\omega \\ r\cos\theta \end{pmatrix} \end{aligned}$$

which maps a point given in spherical coordinates to its corresponding point in the cartesian coordinate system.

The Jacobian of the spherical coordinate transform is:

$$\begin{aligned} D\varphi(r, \theta, \omega) &= \begin{pmatrix} \frac{\partial}{\partial r} r\sin\theta\cos\omega & \frac{\partial}{\partial \theta} r\sin\theta\cos\omega & \frac{\partial}{\partial \omega} r\sin\theta\cos\omega \\ \frac{\partial}{\partial r} r\sin\theta\sin\omega & \frac{\partial}{\partial \theta} r\sin\theta\sin\omega & \frac{\partial}{\partial \omega} r\sin\theta\sin\omega \\ \frac{\partial}{\partial r} r\cos\theta & \frac{\partial}{\partial \theta} r\cos\theta & \frac{\partial}{\partial \omega} r\cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \sin\theta\cos\omega & r\cos\theta\cos\omega & -r\sin\theta\sin\omega \\ \sin\theta\sin\omega & r\cos\theta\sin\omega & r\sin\theta\cos\omega \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} \end{aligned}$$

And a long, tedious and unenlightening calculation eventually gives the determinant as:

$$\det D\varphi(r, \theta, \omega) = r^2\sin(\theta)$$

Which gives us:

Corollary 16.4.4.1: Spherical Coordinate Integral Transform

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be integrable. Then we have:

$$\begin{aligned} &\int_{\mathbb{R}^3} f(x, y, z)\, d\lambda^3(x, y, z) \\ &= \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} r^2\sin\theta \cdot f(r\sin\theta\sin\omega, r\sin\theta\cos\omega, r\cos\theta)\, dr\, d\theta\, d\omega \end{aligned}$$

Example 16.4.5. (Another derivation of the volume of B_R^3):

$$\begin{aligned}
\lambda^3(B_R^3) &= \int_{\mathbb{R}^3} \mathbb{1}_{B_R^3}(x,y,z)\, d\lambda^3(x,y,z) \\
&= \int_{\mathbb{R}^3} \mathbb{1}_{\left\{\sqrt{x^2+y^2+z^2} \,<\, R\right\}}(x,y,z)\, d\lambda^3(x,y,z) \\
&= \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \mathbb{1}_{\{r\,<\,R\}}(r)\cdot r^2\sin\theta\, dr\, d\theta\, d\omega \\
&= \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \mathbb{1}_{(0,R)}(r)\cdot r^2\sin\theta\, dr\, d\theta\, d\omega \\
&= \left(\int_0^{\infty} \mathbb{1}_{(0,R)}(r)\cdot r^2\, dr\right)\cdot\left(\int_0^{\pi}\sin\theta\, d\theta\right)\cdot\left(\int_0^{2\pi} 1\, d\omega\right) \\
&= \left(\int_0^{R} r^2\, dr\right)\cdot\left(\int_0^{\pi}\sin\theta\, d\theta\right)\cdot\left(\int_0^{2\pi} 1\, d\omega\right) \\
&= \frac{1}{3}R^3\cdot(-\cos(\pi)+\cos(0))\cdot 2\pi \\
&= \frac{4}{3}\pi R^3
\end{aligned}$$

Example 16.4.6. (Gravitational Potential of a Sphere:): The gravitational potential that a body $B \subset \mathbb{R}^3$ with mass distribution $m : B \to \mathbb{R}$ exerts on a point $x \in \mathbb{R}^3 \setminus B$ is given by:

$$V(x) = -G\cdot\int_B \frac{m(r)}{\|x-y\|_2}\, d\lambda^3(y)$$

We will derive a closed form expression for the case $B = B_R^3$. If we assume that the mass distribution is rotationally symmetric, we can replace $m : \mathbb{R}^3 \to \mathbb{R}$ with a function $\widetilde{m} = m(\|y\|_2) : \mathbb{R} \to \mathbb{R}$. We will commit a slight abuse of notation here and let m be this new function $\widetilde{m}$ from now on. Because of this rotational symmetry, we also have $V(x) = V(0,0,\|x\|)$.

Now, if y has spherical coordinates (r,θ,ω), then assuming y lies outside of our ball we can use the Pythagorean theorem to get:

$$\begin{aligned}
\|x-y\|_2^2 &= (\|x\|_2 - r\cos\theta)^2 + r^2\sin^2\theta \\
&= \|x\|_2^2 + r^2 - 2r\|x\|_2\cos\theta
\end{aligned}$$

Thus, we have:

$$\begin{aligned}
V(x) &= -G\cdot\int_{B_R^3} \frac{m(r)}{\|x-y\|_2}\, d\lambda^3(y) \\
&= -G\int_0^{2\pi}\int_0^{\pi}\int_0^{R}\left(\frac{m(r)}{\sqrt{\|x\|_2^2+r^2-2r\|x\|_2\cos\theta}}\cdot r^2\sin\theta\right) dr\, d\theta\, d\omega \\
&= -2\pi G\int_0^{R} r^2\cdot m(r)\cdot\left(\int_0^{\pi}\frac{\sin\theta}{\sqrt{\|x\|_2^2+r^2-2r\|x\|_2\cos\theta}}\, d\theta\right) dr
\end{aligned}$$

Substituting $t = g(\theta) = -\cos\theta$ into the inner integral, we get:

$$\begin{aligned}
&\int_0^\pi \frac{\sin\theta}{\sqrt{\|x\|_2^2 + r^2 - 2r\|x\|_2\cos\theta}}\, d\theta \\
=&\int_{-1}^1 \frac{1}{\sqrt{\|x\|_2^2 + r^2 + 2r\|x\|_2 t}}\, dt \\
=&\left[\frac{1}{r\cdot\|x\|_2}\cdot\sqrt{\|x\|_2^2 + r^2 + 2r\|x\|_2 t}\right]_{t=-1}^{t=1} \\
=&\frac{1}{r\cdot\|x\|_2}\cdot\left(\sqrt{\|x\|_2^2 + r^2 + 2r\|x\|_2} - \sqrt{\|x\|_2^2 + r^2 - 2r\|x\|_2}\right) \\
=&\frac{1}{r\cdot\|x\|_2}\cdot((\|x\|_2 + r) - \|x\|_2 - r) \\
=&\frac{2}{\|x\|_2}
\end{aligned}$$

Therefore, we have:

$$V(x) = \frac{4\pi}{\|x\|_2}\cdot\int_0^R m(r)r^2\, dr$$

Now, we can once again use the rotational symmetry of m to describe the total mass M of our body in terms of spherical coordinates:

$$\begin{aligned}
M &= \int_{B_R^3(0)} m(y)\, d\lambda^3(y) \\
&= \int_0^{2\pi}\int_0^\pi\int_0^R m(r)r^2\sin\theta\, dr\, d\theta\, d\omega \\
&= 4\pi\int_0^R m(r)r^2\, dr
\end{aligned}$$

Thus, the total mass distribution of our sphere is simply:

$$V(x) = -\frac{GM}{\|x\|}$$

which is exactly the same mass distribution we would get if we concentrated the entire mass of our sphere at its center. This relation is part of a theorem in classical mechanics known as the **shell theorem**.

16.4.3 Cylindrical Coordinates

Alternatively, one can trivially extend polar coordinates into three dimensions by simply leaving the z coordinate unchanged, leading to *cylindrical coordinates*:

Definition 16.4.6.1: Cylindrical Coordinates on $\mathbb{R}^3$

Let $U = (0, \infty) \times (0, 2\pi) \times \mathbb{R}$ and $V = \mathbb{R}^3 \setminus \{(x, 0, z)^\top : x \geq 0\}$. The **spherical coordinate transformation** is the map:

$$\varphi : U \to V$$
$$\begin{pmatrix} r \\ \omega \\ z \end{pmatrix} \mapsto \begin{pmatrix} r\cos\omega \\ r\sin\omega \\ z \end{pmatrix}$$

which maps a point given in cylindrical coordinates to its corresponding point in the cartesian coordinate system.

Since the cylindrical coordinate transform leaves the z coordinate unchanged, its Jacobian and its determinant end up being almost identical:

$$\begin{aligned} D\varphi(r, \omega, z) &= \begin{pmatrix} \frac{\partial}{\partial r} r\cos\omega & \frac{\partial}{\partial \omega} r\cos\omega & \frac{\partial}{\partial z} r\cos\omega \\ \frac{\partial}{\partial r} r\sin\omega & \frac{\partial}{\partial \omega} r\sin\omega & \frac{\partial}{\partial z} r\sin\omega \\ \frac{\partial}{\partial r} z & \frac{\partial}{\partial \omega} z & \frac{\partial}{\partial z} z \end{pmatrix} \\ &= \begin{pmatrix} \cos\omega & -r\sin\omega & 0 \\ \sin\omega & r\cos\omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \det D\varphi(r, \omega, z) &= \det \begin{pmatrix} \cos\omega & -r\sin\omega & 0 \\ \sin\omega & r\cos\omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \det \begin{pmatrix} \cos\omega & -r\sin\omega \\ \sin\omega & r\cos\omega \end{pmatrix} \\ &= r\cos(\omega)^2 + r\sin(\omega)^2 \\ &= r \end{aligned}$$

Corollary 16.4.6.1: Cylindrical Coordinate Integral Transform

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be integrable. Then we have:

$$\int_{\mathbb{R}^3} f(x, y, z)\, d\lambda(x, y, z) = \int_{\mathbb{R}} \int_0^{2\pi} \int_0^\infty r \cdot f(r\cos(\omega), r\sin(\omega), z)\, dr\, d\omega\, dz$$

Proposition 16.4.7. *The inverse of the cylindrical coordinate transform φ is:*

$$\varphi^{-1}(x, y, z) = \begin{cases} \left(r,\ \arccos\left(\frac{x}{r}\right),\ z\right)^\top & y \geq 0 \\ \left(r,\ 2\pi - \arccos\left(\frac{x}{r}\right),\ z\right)^\top & y < 0 \end{cases}$$

where $r = \sqrt{x^2 + y^2}$.

Theorem 16.4.7.1: Volume of a Solid of Revolution

Let S be a solid constructed by rotating a function $\rho(z)$ around the z axis, i.e.

$$S = \left\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 < \rho(z)^2\right\}$$

Then we have:

$$\lambda^3(S) = \pi \int_{\mathbb{R}} \rho(z)^2 \, d\lambda^1(z)$$

Proof.

$$\begin{aligned}
\lambda^3(S) &= \int_{\mathbb{R}^3} \mathbb{1}_{\left\{x^2+y^2 \,<\, \rho(z)^2\right\}}(x, y, z) \, d\lambda^3(x, y, z) \\
&= \int_{\mathbb{R}} \int_0^{2\pi} \int_0^{\infty} r \cdot \mathbb{1}_{\left\{r^2 \,<\, \rho(z)^2\right\}}(r, \omega, z) \, dr \, d\omega \, dz \\
&= \int_{\mathbb{R}} \int_0^{2\pi} \int_0^{\infty} r \cdot \mathbb{1}_{\{r \,<\, \rho(z)\}}(r, \omega, z) \, dr \, d\omega \, dz \\
&= \int_{\mathbb{R}} \int_0^{2\pi} \int_0^{\rho(z)} r \, dr \, d\omega \, dz \\
&= \int_{\mathbb{R}} \int_0^{2\pi} \frac{1}{2}\rho(z)^2 \, d\omega \, dz \\
&= \int_{\mathbb{R}} \pi\rho(z)^2 \, dz \\
&= \pi \int_{\mathbb{R}} \rho(z)^2 \, dz
\end{aligned}$$

□

16.5 The Surface Area Measure

Definition 16.5.1. Let $U \supset \mathbb{R}^n$ be open. We call a map $f : U \to \mathbb{R}^{n+k}$ an *immersion* iff f is continuously differentiable and $Df(x)$ is invertible.

Definition 16.5.2. Let $M \subset \mathbb{R}^{n+k}$ be an n-dimensional C^1 submanifold. Then a *local parametrization of* M is an injective immersion $f : \mathbb{R}^n \supset U \to M \subset \mathbb{R}^{n+k}$

Definition 16.5.2.1: Gram Matrix

Let $U \subset \mathbb{R}^n$ and $f : U \to \mathbb{R}^{n+k}$. Then the *Gram matrix* $g_f(x)$ is the matrix

$$g_f(x) = Df(x)^\top Df(x)$$

Our goal now is to find a sensible definition of *surface area*, i.e. of an n-dimensional measure associated with an object embedded into $n+k$-dimensional space.

Corollary 16.5.2.1: Area of an Embedded Diffeomorphism of a Set

Let $f = S \circ \varphi$, where $\varphi : U \to V$ is a diffeomorphism with $U, V \subset \mathbb{R}^n$ and S is a linear isometry. Then we have

$$\lambda^n(\varphi(E)) = \int_E \sqrt{\det g_f}$$

This theorem suggests that it is sensible to use the same definition for arbitrary immersions:

Definition 16.5.2.2: Area of an Immersion of a Set

Let $f \in C^1(U, \mathbb{R}^{n+k})$ be an n-dimensional immersion. Let $E \subset U$ be λ^n-measurable. Then the n-dimensional area of $f(E)$ is given by:

$$A(f, E) = \int_E \sqrt{\det g_f}$$

$\sqrt{\det g_f}$ is known as the *Jacobian of* f, and often denoted Jf, but I dislike notation that hides the actual calculation being done behind multiple layers of nested notation and will thus be sticking to $\sqrt{\det g_f}$.

Corollary 16.5.2.2: Length of a Regular Curve

An Immersion $f : I = (a, b) \to \mathbb{R}^n$ is known as a *regular curve*. The one-dimensional area of the curve, i.e. its length, is given by

$$L(f) := A(f, I) = \int_a^b \|Df(t)\|_2 \, dt$$

Proof. We have:

$$Df(t) = \begin{pmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{pmatrix}$$

and thus

$$g_f(t) = Df(t)^\top Df(t) = \|Df(t)\|_2^2$$

□

Note that if some areas of the curve are covered multiple times, then this length formula counts them multiple times as well.

Corollary 16.5.2.3: Area of a Regular Surface

An immersion $f : U \supset \mathbb{R}^2 \to \mathbb{R}^3$, with U open, is known as a *regular surface*. Its area is given by

$$\begin{aligned} A(f) := A(f, U) &= \int_U \sqrt{\left\|\frac{\partial}{\partial x} f\right\|_2^2 \cdot \left\|\frac{\partial}{\partial y} f\right\|_2^2 - \left\langle \frac{\partial}{\partial x} f, \frac{\partial}{\partial y} f \right\rangle^2} \, d\lambda^2(x, y) \\ &= \int_U \left\| \frac{\partial}{\partial x} f \times \frac{\partial}{\partial y} f \right\| d\lambda^2(x, y) \end{aligned}$$

Proof. We have

$$Df = \begin{pmatrix} \frac{\partial}{\partial x} f(x, y) & \frac{\partial}{\partial y} f \end{pmatrix},$$

therefore our Gram matrix is

$$g_f = \begin{pmatrix} \|\frac{\partial}{\partial x} f\|_2^2 & \left\langle \frac{\partial}{\partial x} f, \frac{\partial}{\partial y} f \right\rangle_2 \\ \left\langle \frac{\partial}{\partial x} f, \frac{\partial}{\partial y} f \right\rangle & \left\|\frac{\partial}{\partial y} f\right\|_2^2 \end{pmatrix},$$

giving our desired result for $\sqrt{\det g_f}$. □

Theorem 16.5.2.1: Area of a Graph

Let $U \subset \mathbb{R}^n$ be open, $u : U \to \mathbb{R}^k$ and

$$\begin{aligned} f : U &\to \mathbb{R}^{n+k} \\ x &\mapsto (x, u(x)) \end{aligned}$$

Then we have:

$$A(f, U) = \int_U \sqrt{\det(E_n + Du(x)^\top Du(x))}\, d\lambda^n(x)$$

If $k = 1$, i.e. $U \subset \mathbb{R}^n$, $u : U \to \mathbb{R}$, then we can simplify significantly and get:

$$A(f, U) = \int_U \sqrt{1 + \|Du(x)\|^2}\, d\lambda^n(x)$$

Theorem 16.5.2.2: Surface Measure

Let M be an n-dimension C^1-submanifold of $\mathbb{R}^{n+k}$ locally parametrized by $f : U \to M$. Let $E \subset M$. Then there exists a measure ω_M such that

$$\omega_M(E) = A(f, E)$$

Theorem 16.5.2.3: Surface of Revolution

Let S be a solid of revolution generated by $\rho(z)$. Then the surface area of S is:

$$\omega_S(S) = 2\pi \int_{\mathbb{R}} \rho(z) \sqrt{1 + \left(\frac{\mathrm{d}}{\mathrm{d}z}\rho(z)\right)^2}\, dz$$

Proof. This follows from our formula for the surface area of a regular surface. However, as additional practice, let us derive the formula from scratch. We can parametrize S up to a zero set via:

$$f : \begin{pmatrix} \theta \\ z \end{pmatrix} \mapsto \begin{pmatrix} \rho(z)\cos(\theta) \\ \rho(z)\sin(\theta) \\ z \end{pmatrix}$$

Which gives us:

$$\begin{aligned} Df(\theta, z) &= \begin{pmatrix} \frac{\partial}{\partial\theta}\rho(z)\cos(\theta) & \frac{\partial}{\partial z}\rho(z)\cos(\theta) \\ \frac{\partial}{\partial\theta}\rho(z)\sin(\theta) & \frac{\partial}{\partial z}\rho(z)\sin(\theta) \\ \frac{\partial}{\partial\theta} z & \frac{\partial}{\partial z} z \end{pmatrix} \\ &= \begin{pmatrix} -\rho(z)\sin(\theta) & \cos(\theta)\cdot\frac{\partial}{\partial z}\rho(z) \\ \rho(z)\cos(\theta) & \sin(\theta)\cdot\frac{\partial}{\partial z}\rho(z) \\ 0 & 1 \end{pmatrix} \end{aligned}$$

This gives us the Gram matrix:

$$\begin{aligned}
g_f(\theta, z) &= \begin{pmatrix} -\rho(z)\sin(\theta) & \rho(z)\cos(\theta) & 0 \\ \cos(\theta)\cdot\frac{\partial}{\partial z}\rho(z) & \sin(\theta)\cdot\frac{\partial}{\partial z}\rho(z) & 1 \end{pmatrix} \cdot \begin{pmatrix} -\rho(z)\sin(\theta) & \cos(\theta)\cdot\frac{\partial}{\partial z}\rho(z) \\ \rho(z)\cos(\theta) & \sin(\theta)\cdot\frac{\partial}{\partial z}\rho(z) \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \rho(z)^2\cdot(\sin^2(\theta)+\cos^2(\theta)) & \begin{matrix} -\rho(z)\sin(\theta)\cos(\theta)\frac{\mathrm{d}}{\mathrm{d}z}\rho(z) \\ +\rho(z)\sin(\theta)\cos(\theta)\frac{\mathrm{d}}{\mathrm{d}z}\rho(z) \end{matrix} \\ \begin{matrix} -\rho(z)\sin(\theta)\cos(\theta)\frac{\mathrm{d}}{\mathrm{d}z}\rho(z) \\ +\rho(z)\sin(\theta)\cos(\theta)\frac{\mathrm{d}}{\mathrm{d}z}\rho(z) \end{matrix} & (\cos^2(\theta)+\sin^2(\theta)\Big(\frac{\partial}{\partial z}\rho(z)\Big)^2+1 \end{pmatrix} \\
&= \begin{pmatrix} \rho(z)^2 & 0 \\ 0 & \Big(\frac{\partial}{\partial z}\rho(z)\Big)^2+1 \end{pmatrix}
\end{aligned}$$

and thus:

$$\begin{aligned}
\sqrt{\left|\det g_f(\theta, z)\right|} &= \sqrt{\rho(z)^2\left(\frac{\partial}{\partial z}\rho(z)\right)^2+\rho(z)^2} \\
&= \rho(z)\sqrt{1+\frac{\partial}{\partial z}\rho(z))^2}
\end{aligned}$$

Which leaves us with:

$$\begin{aligned}
\omega_S(S) &= \int_0^{2\pi}\int_{\mathbb{R}}\rho(z)\sqrt{1+\frac{\partial}{\partial z}\rho(z))^2}\,dz\,d\theta \\
&= 2\pi\int_{\mathbb{R}}\rho(z)\sqrt{1+\frac{\partial}{\partial z}\rho(z))^2}\,dz
\end{aligned}$$

□

Example 16.5.3. (Gabriel's Horn): We want to show that the solid of rotation G generated by $z \in [1, \infty)$ and $\rho(z) = \frac{1}{z}$ has finite volume, but infinite surface area.

1. We already showed in the section on cylindrical coordinates that the volume is simply:

$$\begin{aligned}
\lambda^3(G) &= \pi\cdot\int_1^\infty \rho(z)^2\,dz \\
&= \pi\cdot\int_1^\infty \frac{1}{z^2}\,dz \\
&= -\frac{\pi}{2}\cdot\left[\frac{1}{z}\right]_{z=1}^{z=\infty} \\
&= -\frac{\pi}{2}\cdot(0-1) \\
&= \frac{\pi}{2}
\end{aligned}$$

2. We have just shown that the surface area is given by:

$$
\begin{aligned}
\omega_S(S) &= 2\pi \int_{\mathbb{R}} \rho(z)\sqrt{1 + \frac{\partial}{\partial z}\rho(z))^2}\, dz \\
&= 2\pi \int_1^\infty \frac{1}{z}\sqrt{1 - \frac{1}{z^2}}\, dz \\
&> 2\pi \int_1^\infty \frac{1}{z}\, dz \\
&= 2\pi(\ln(\infty) - \ln(1)) \\
&= \infty
\end{aligned}
$$

Theorem 16.5.3.1: Surface Integral

Let $M \subset \mathbb{R}^{n+k}$ be an n-dimensional C^1 submanifold of $\mathbb{R}^{n+k}$, and let $M = \bigsqcup_{i\in I} M_i$ be a decomposition of M into countably many pairwise disjoint measure sets such that $M_i \subset V_i$ for a local parametrization $f_i : U_i \to V_i$. Then for any measurable function $u : M \to \overline{\mathbb{R}}$, we have:

$$\int_M u(x)\, d\omega_M(x) = \sum_{i\in I} \int_{f_i^{-1}(M_i)} u(f_i(x)) \cdot \left|\sqrt{\det g_{f_i}(x)}\right| d\lambda^n(x)$$

In particular, if we can find a local parametrization $f : U \to V$ such that $M \setminus V$ is a zero set, we have:

$$\int_M u(x)\, d\omega_M(x) = \int_{f^{-1}(M)} u(f(x)) \cdot |\det Df(x)|\, d\lambda^n(x)$$

Theorem 16.5.3.2: Onion Formula

Integrating a function f over $\mathbb{R}^n$ is the same thing as integrating over all surface integrals of the function over all balls B_r^n, i.e. we have:

$$\int_{\mathbb{R}^n} f(x)\, d\lambda^n(x) = \int_0^\infty \left(\int_{\partial B_r^n} f(y)\, d\omega_{\partial B_r^n}(y)\right) dr$$

Note that via multivariable substitution we also get:

$$\int_0^\infty \left(\int_{\partial B_r^n} f(y)\, d\omega_{\partial B_r^n}(y)\right) dr = \int_0^\infty \left(\int_{\partial B_1^n} r^n f(ry)\, d\omega_{\partial B_1^n}(y)\right) dr$$

Corollary 16.5.4. *We have:*

$$\omega_{\partial B_1^n}(\partial B_1^n) = (n+1) \cdot \lambda^{n+1}(B_1^{n+1}),$$

i.e. the surface area of the n-dimensional unit ball is $(n+1)$ *times the volume of the* $n+1$*-dimensional unit ball.*

Proof. Applying the onion formula and then Fubini twice, we get:

$$
\begin{aligned}
\lambda^{n+1}(B_1^{n+1}) &= \int_{\mathbb{R}^n} \mathbb{1}_{B_1^n}(x)\, d\lambda^n(x) \\
&= \int_0^\infty \int_{\partial B_r^n} \mathbb{1}_{B_1^n}(y)\, d\omega_{\partial B_r^n}(y)\, dr \\
&= \int_{\partial B_r^n} \int_0^\infty \mathbb{1}_{B_1^n}(y)\, dr\, d\omega_{\partial B_r^n}(y) \\
&= \int_{\partial B_r^n} \int_0^1 1\, dr\, d\omega_{\partial B_r^n}(y) \\
&= \int_0^1 \int_{\partial B_r^n} 1\, d\omega_{\partial B_r^n}(y)\, dr \\
&= \int_0^1 \omega_{\partial B_r^n}(\partial B_r^n)\, dr \\
&= \int_0^1 r^n \cdot \omega_{\partial B_1^n}(\partial B_1^n)\, dr \\
&= \frac{\omega_{\partial B_1^n}(\partial B_1^n)}{n+1}
\end{aligned}
$$

□

16.6 The Divergence Theorem

Definition 16.6.1. Let $\Omega \in \mathbb{R}^n$. We say Ω has a C^1 boundary if $\partial\Omega$ is an $n-1$ dimensional C^1-manifold such that Ω locally lies on only one side of $\partial\Omega$.

More precisely, Ω has a C^1 boundary iff for every point $p \in \partial\Omega$, there exists an open set $U \subset \mathbb{R}^n$, an open interval $I \subset \mathbb{R}$, an a C^1-function $u : U \to I$ such $p \in U \times I$ and

$$\Omega \cap (U \times I) = \{(x, y) \in U \times I : y < u(x)\}$$

Lemma 16.6.2. *Let $\Omega \in \mathbb{R}^n$ be open with C^1 boundary. Then we have:*

1. $\partial\Omega \cap (U \times I) = \{(x, y) \in U \times I : y = u(x)\}$
2. $(\mathbb{R}^n \setminus \overline{\Omega}) \cap (U \times I) = \{(x, y) \in U \times I : y \geq u(x)\}$

In particular, $\partial\Omega$ is an $n-1$ dimensional manifold given via the graph criterion as $(x, u(x))$.

Theorem 16.6.2.1: Outer Normal Field

Let $\Omega \subset \mathbb{R}^n$ be open with C^1 boundary. Then there exists exactly one map $\nu_\Omega : \partial\Omega \to \mathbb{R}^n$ such that:

1. $\nu_\Omega(p) \perp T_p(\partial\Omega)$,
2. $\|\nu_\Omega(p)\|_2 = 1$,
3. $\exists \varepsilon > 0 : t \in (0, \varepsilon) \implies p + t\nu_\Omega(p) \notin \Omega$.

This map is continuous and unique and is known as the *outer normal field of* Ω. If u is given as in the definition of C^1 boundary, and $p = (x, u(x))$, then we have:

$$\nu_\Omega(p) = \frac{(-Du(x), 1)^\top}{\sqrt{1 + \|Du(x)\|_2^2}}$$

Lemma 16.6.2.1: Partition of Unity

Let W_i, $i \in I$ be an open cover of a compact set $K \subset \mathbb{R}^n$. Then there exists a finite family of smooth functions $\chi_j : \mathbb{R}^n \to \mathbb{R}^n$ with compact support such that

1. For all $x \in K$, we have $\sum_{j \in J} \chi_j(x) = 1$
2. For all $j \in J$, there exists an i such that $\operatorname{spt}\chi_j \subset W_i$.

Theorem 16.6.2.2: Divergence Theorem / Gauss's Theorem

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with C^1 boundary and outer normal $\nu_\Omega : \partial\Omega \to \mathbb{R}^n$. Let $f : \overline{\Omega}\mathbb{R}$ be continuously differentiable. Then we have:

$$\int_\Omega \operatorname{div} f(x)\, d\lambda^n(x) = \int_{\partial\Omega} \langle f(x), \nu(x) \rangle \; d\omega(x).$$

Theorem 16.6.2.3: Multidimensional Partial Integration

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with C^1 boundary and outer normal $\nu_\Omega : \partial\Omega \to \mathbb{R}^n$. Let $f : \mathbb{R}^n \to \mathbb{R}$ and $G : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable. Then we have:

$$\begin{aligned} &\int_\Omega f \cdot \operatorname{div}(G)\, d\lambda^n(x) \\ =& \int_{\partial\Omega} f \cdot \langle G, \nu_\Omega\rangle \; d\omega(x) - \int_\Omega \left\langle \vec{\nabla} f, G \right\rangle \; d\lambda^n(x) \end{aligned}$$

Proof. By the product rule for divergence, we have:

$$\begin{aligned} & f(x) \cdot \operatorname{div}(G(x)) + \left\langle \vec{\nabla} f(x), G(x) \right\rangle \\ =& \operatorname{div}(f(x) \cdot G(x)) \end{aligned}$$

Now, simply apply the divergence theorem, together with basic properties of scalar products, to get:

$$\begin{aligned} &\int_\Omega f(x) \cdot \operatorname{div}(G(x)) + \left\langle \vec{\nabla} f(x), G(x) \right\rangle \; d\lambda^n(x) \\ =& \int_\Omega \operatorname{div}(f(x) \cdot G(x))\, d\lambda^n(x) \\ =& \int_{\partial\Omega} \langle f(x) \cdot G(x), \nu_\Omega(x)\rangle \; d\omega(x) \\ =& \int_{\partial\Omega} f(x) \cdot \langle G(x), \nu_\Omega(x)\rangle \; d\omega(x), \end{aligned}$$

which we can rearrange to get our desired identity. □

If we plug $G(x) := \vec{\nabla} g(x)$ into this formula, we get:

Corollary 16.6.2.1: Green's First Identity

Let $\Omega \subset \mathbb{R}^n$ be open and bounded with C^1 boundary. Let $f : \overline{\Omega} \to \mathbb{R}$ be continuously differentiable and let $g : \overline{\Omega} \to \mathbb{R}$ be twice continuously differentiable. Then we have:

$$\begin{aligned} &\int_\Omega f \cdot \Delta g + \left\langle \vec{\nabla} f, \vec{\nabla} g \right\rangle \; d\lambda^n \\ =& \int_{\partial\Omega} f \cdot \left\langle \nu_\Omega, \vec{\nabla} g \right\rangle \; d\omega \\ &\left(= \int_{\partial\Omega} f \cdot \frac{\partial}{\partial \nu_\Omega} g \; d\omega(x) \right) \end{aligned}$$

Theorem 16.6.2.4: Green's Second Identity

Let $\Omega \subset \mathbb{R}^n$ be open and bounded with C^1 boundary. Let $f : \overline{\Omega} \to \mathbb{R}$ and $g : \overline{\Omega} \to \mathbb{R}$ be twice continuously differentiable. Then we have:

$$\begin{aligned}&\int_{\Omega} f \cdot \Delta g - g \cdot \Delta f \; d\lambda^n \\ =&\int_{\partial\Omega} f \cdot \left\langle \nu_\Omega, \vec{\nabla} g \right\rangle - g(x) \cdot \left\langle \nu_\Omega(x), \vec{\nabla} f(x) \right\rangle \; d\omega(x)\end{aligned}$$

Proof. This is just two instances of Green's first identity subtracted from each other:

$$\begin{aligned}&\int_{\Omega} f(x) \cdot \Delta g(x) - g(x) \cdot \Delta f(x) \; d\lambda^n \\ =&\int_{\Omega} f(x) \cdot \Delta g(x) + \left\langle \vec{\nabla} f(x), \vec{\nabla} g(x) \right\rangle \; d\lambda^n \\ -&\int_{\Omega} g(x) \cdot \Delta f(x) + \left\langle \vec{\nabla} f(x), \vec{\nabla} g(x) \right\rangle \; d\lambda^n \\ =&\int_{\partial\Omega} f(x) \cdot \left\langle \nu_\Omega(x), \vec{\nabla} g(x) \right\rangle \; d\omega(x) \\ -&\int_{\partial\Omega} g(x) \cdot \left\langle \nu_\Omega(x), \vec{\nabla} f(x) \right\rangle \; d\omega(x) \\ =&\int_{\partial\Omega} f(x) \cdot \left\langle \nu_\Omega(x), \vec{\nabla} g(x) \right\rangle - g(x) \cdot \left\langle \nu_\Omega(x), \vec{\nabla} f(x) \right\rangle \; d\omega(x)\end{aligned}$$

□

Chapter 17

Convolutions

Convolution is a way of creating a function by taking a weighted average of two functions. Among other things, it can be used to smooth a given function.

Lemma 17.0.1. *$\tau_h : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be the function that translates a vector by the vector h, i.e. $\tau_h(x) = x + h$. Let $f \in L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$. Then the following hold:*

1. $f \circ \tau_h \in L^p(\mathbb{R}^n)$
2. $\|f \circ \tau_h\|_p = \|f\|_p$
3. $\lim\limits_{h \to 0} \|f \circ \tau_h - f\|_p = 0$

Definition 17.0.1.1: Convolution

Let $1 \leq p < \infty$, $f \in L^p(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^n$. The *convolution of f and g* is defined to be:

$$
\begin{aligned}
&f * g : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}} \\
&(f * g)(x) = \int_{\mathbb{R}^n} f(x - y) g(y) \, d\lambda^n(y)
\end{aligned}
$$

Theorem 17.0.2. *Let $1 \leq p < \infty$, $f \in L^p(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^n$. Then we have:*

1. $f * g \in L^p(\mathbb{R}^n)$
2. $\|f * g\|_p \leq \|f\|_p \|g\|_1$
3. $f * g = g * f$

Theorem 17.0.3. (Approximation through convolution): *Let $\eta \in L^1(\mathbb{R}^n)$ such that*

$$\int_{\mathbb{R}^n} \eta(x)\, d\lambda^n(x).$$

Let $\rho > 0$ and

$$\eta_\rho(x) := \rho^{-n} \cdot \eta\left(\frac{x}{\rho}\right).$$

Let $1 \leq p < \infty$ and $f \in L^p(\mathbb{R}^n)$. Then we have:

1. $f * \eta_\rho \in L^p(\mathbb{R}^n)$
2. $\left\|f * \eta_\rho\right\|_p \leq \|f\|_p \cdot \|\eta\|_1$
3. $\left\|(f * \eta_\rho) - f\right\|_p \to 0$

Definition 17.0.3.1: Multi-index Notation

Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$. Then we define:

1. $|\alpha| = \sum_{i=1}^{n} \alpha_i$
2. $\partial^\alpha f(x) = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f(x)$

Theorem 17.0.3.1: Smoothing

Let $\eta \in C^k(\mathbb{R}^n)$ such that for $|\alpha| \leq k$, we have

$$\|\partial^\alpha \eta\|_{C^0(\mathbb{R}^n)} \leq K.$$

Let $f \in L^1(\mathbb{R}^n)$. Then we have:

1. $f * \eta \in C^k(\mathbb{R}^n)$,
2. $\partial^\alpha(f * \eta) = f * (\partial^\alpha \eta)$,
3. $\|\partial^\alpha(f * \eta)\|_{C^0(\mathbb{R}^n)} \leq K\|f\|_1$

Smoothing lets us strengthen our result on the density of continous functions in $L^p(\Omega)$:

Corollary 17.0.3.1: Density of C_0^∞ in $L^p(\Omega)$

Let $\Omega \subset \mathbb{R}^n$ be open. Let $1 \leq p < \infty$. Then, for every $f \in L^p(\Omega)$, there exists a sequence $f_k \in C_0^\infty(\Omega)$ such that $\|f - f_k\|_{L^p(\Omega)} \to 0$.

Definition 17.0.3.2: Locally integrable Function

Let $\Omega \subset \mathbb{R}^n$ open and $1 \leq p \leq \infty$. Then we call a function $f : \Omega \to \mathbb{R}$ *locally integrable* iff for any compact set $K \subset \Omega$, we have $\mathbb{1}_K \cdot f \in L^p(\Omega)$. We denote the set of all locally integrable functions $f : \Omega \to \mathbb{R}$ by

$$L^p_{\text{loc}}(\Omega)$$

Lemma 17.0.3.1: Fundamental Lemma of the Calculus of Variations

Let $\Omega \subset \mathbb{R}^n$ be open. Let $f \in L^1_{\text{loc}}(\Omega)$ such that

$$\forall \varphi \in C_0^\infty(\Omega) :$$
$$\varphi \geq 0 \implies \int_\Omega f\varphi \, d\lambda^n(x) \geq 0$$

Then we have $f(x) \geq 0$ for λ^n-almost all $x \in \Omega$.

Corollary 17.0.3.2: There is no identity of convolution.

There exists no function $\eta \in L^1(\mathbb{R}^n)$ such that for all $f \in C_0^\infty(\mathbb{R}^n)$, we have $f * n = f$.

Chapter 18

The Bochner 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.