

Analysis for Algebraists

Emma Yoneda

2025-12-25

Contents

I Set Theory	3
II Algebraic Structures	4
1 Ordered Fields	5
1.1 The Archimedean Property.....	7
1.2 Why always \mathbb{R} ?	7
1.2.1 Subfields of ordered fields.....	7
1.2.2 The least upper bound property.....	8
1.2.3 Alternative completeness properties.....	8
III Topology	9
2 Metric Spaces and Topological Spaces	10
2.1 Vocabulary	10
2.2 Continuity	11
3 Topological Vector Spaces	16
3.1 Normed Vector Spaces	16
3.2 Banach Spaces.....	16
3.3 Hilbert Spaces.....	16
3.4 Topological Vector Spaces.....	16
4 Uniform Spaces	17
4.1 Diagonal Uniformity	17
IV Differentiation and Integration	19
5 Differentiation	20
5.1 The Frechét Derivative.....	20
5.2 Frechét Spaces.....	20
5.3 The Gateaux Derivative.....	20
6 Manifolds and Diffeomorphisms	21

6.1	Inverse Function Theorem.....	21
7	Measure Theory	22
7.1	The Measure Problem.....	22
7.1.1	Vitali Sets.....	23
7.1.2	The Banach-Tarski Paradox	24
7.2	Lattices and Boolean Algebras.....	24
7.2.1	Boolean Algebras.....	25
7.2.2	Lattices	25
7.3	Set Algebras.....	26
7.3.1	σ -Algebras	26
7.3.2	Set Rings	27
7.3.3	Semirings of Sets	27
7.4	Measure functions	27
7.5	The Lebesgue Measure.....	28
8	Integration	29
8.1	The Lebesgue Integral.....	29
8.2	The Bochner Integral	29

Part I

Set Theory

Part II

Algebraic Structures

Chapter 1

Ordered Fields

Definition 1.0.1. 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 1.0.2. 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 1.0.3. 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 1.0.4. *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 1.0.5. *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 1.0.6. *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 1.0.7. 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 $ac \geq bc$.

1.1 The Archimedean Property

Definition 1.1.1. Let F be an archimedean ordered field. Then we say that F is **archimedean** if for every $x, y \in F_{>0}$, there exists a natural number n such that

$$nx > y$$

Comment 1.1.2. It follows immediately that if F is non-archimedean, there exists $x, y \in F$ such that for all natural numbers n , we have

$$nx < y$$

which immediately implies

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

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

1.2 Why always \mathbb{R} ?

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

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

1.2.1 Subfields of ordered fields

Theorem 1.2.1. Let F be an archimedean ordered field. Then F is isomorphic to a subfield of the real numbers \mathbb{R} .

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

1.2.2 The least upper bound property

Definition 1.2.2. Let F be an ordered field. We say that F has the **least upper bound property**, or alternatively that F is **Dedekind complete**, if every subset of F that has an upper bound in F has a least upper bound in F .

Theorem 1.2.3. F has the least upper bound property if and only if it has the equivalent "greatest lower bound property", i.e. every subset of F that has a lower bound in F has a greatest lower bound in F .

Theorem 1.2.4. Let F be a non-archimedean ordered field. Then F does not have the least-upper-bound property.

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

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

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

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

Corollary 1.2.5. Let F be an ordered field. Then F has the least-upper-bound property if and only if it is archimedean and Cauchy complete.

Theorem 1.2.6. Every ordered field with the least upper bound property is isomorphic. Therefore the real numbers \mathbb{R} are, up to isomorphism, the only ordered field with the least upper bound property.

1.2.3 Alternative completeness properties

Part III

Topology

Chapter 2

Metric Spaces and Topological Spaces

2.1 Vocabulary

Definition 2.1.1. 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 2.1.2. 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. \square

Definition 2.1.3. 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 "Durschchnitt", "average, intersection") if it is a countable intersection of open sets.

Theorem 2.1.4. *The complement of a G_δ set is F_σ and vice versa.*

2.2 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 2.2.1. Let $f : X \rightarrow Y$ 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 extreme topologies are involved, continuity of a function is often trivial to verify:

Theorem 2.2.2. *Let $f : X \rightarrow Y$ 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. \square

Theorem 2.2.3. *Let $f : X \rightarrow Y$ be a function between topological spaces. Assume X has the discrete topology. Then f is continuous.*

Proof. Every subset of X is open, therefore every preimage of f must be open. \square

Definition 2.2.4. Let $f : X \rightarrow Y$ 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 2.2.5. *$f : X \rightarrow Y$ 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.

\implies : 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.

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

□

Theorem 2.2.6. *$f : X \rightarrow Y$ is continuous if and only if it is continuous at every point $x \in X$.*

Proof.

\implies : 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 neighborhood of x . Therefore f is continuous at every x

\impliedby : 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 2.2.7. Let $f : X \rightarrow Y$ be a function between topological spaces. We call f **sequentially continuous at a point x** if, for every sequence x_n such that $\lim_{n \rightarrow \infty} x_n = x$, we have $\lim_{n \rightarrow \infty} f(x_n) = f(x)$. 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 2.2.8. *Every continuous function is sequentially continuous.*

Theorem 2.2.9. *If X is first-countable (and we assume the axiom of choice), then any sequentially continuous function is continuous.*

Corollary 2.2.10. *A function $f : X \rightarrow Y$ 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 2.2.11 (Epsilon-Delta-Criterion). *Let $f : M \rightarrow N$ 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 $f : \mathbb{R} \rightarrow \mathbb{R}$ 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 $\lim_{n \rightarrow \infty} x_n = x$ which fulfills our condition. We need to show $\lim_{n \rightarrow \infty} f(x_n) = f(x)$, 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 $\lim_{n \rightarrow \infty} x_n = x$, we can find an N such that this condition is fulfilled for all $n > N$. Therefore does $f(x_n)$ indeed converge to $f(x)$. \square

Theorem 2.2.12. *X be a topological space and let $A \subset X$. Then, assuming the discrete topology on $\{0, 1\}$, the characteristic function $\chi_A : X \rightarrow \{0, 1\}$ is continuous at a point $x \in X$ if and only if $x \in \text{int}(A)$ or $x \in \text{int}(X \setminus A)$.*

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 . \square

Comment 2.2.13. 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 2.2.14. *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. \square

Theorem 2.2.15. (A function continuous at exactly one point): *The function $f : \mathbb{R} \rightarrow \mathbb{R}$ with*

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

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 $(-\varepsilon, \varepsilon) \subset V$. 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 2.2.16. (A function only continuous at the irrationals) *Thomae's function* $T : \mathbb{R} \rightarrow \mathbb{R}$, 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 intersection of open sets.

Theorem 2.2.17. (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 $f_F : \mathbb{R} \rightarrow \mathbb{R}$ 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 2.2.18. (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 2.2.19. 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 2.2.20. There is no function $f : \mathbb{R} \rightarrow \mathbb{R}$ 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 . \square

Chapter 3

Topological Vector Spaces

3.1 Normed Vector Spaces

3.2 Banach Spaces

3.3 Hilbert Spaces

3.4 Topological Vector Spaces

Chapter 4

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 $f : M \rightarrow N$ 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.

4.1 Diagonal Uniformity

Definition 4.1.1. 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 4.1.2. 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 4.1.3. 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 4.1.4. 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 4.1.5. 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 and Integration

Chapter 5

Differentiation

5.1 The Frechét Derivative

5.2 Frechét Spaces

5.3 The Gateaux Derivative

Chapter 6

Manifolds and Diffeomorphisms

6.1 Inverse Function Theorem

Definition 6.1.1. We say a map is " \mathcal{C}^n " if its first n derivatives exist and are continuous. If $f \in \mathcal{C}^n$ is bijective such that $f^{-1} \in \mathcal{C}^n$, then we call f a \mathcal{C}^n -**diffeomorphism**.

Theorem 6.1.2. (Inverse Function Theorem) *Let X, Y be finite-dimensional real affine spaces, let $U \subset X$ be open and let $f : U \rightarrow Y$ be \mathcal{C}^n . Then if the differential d_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 \mathcal{C}^n -diffeomorphism.*

Chapter 7

Measure Theory

7.1 The Measure Problem

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

1. When given a subset with an easily intuitively definable volume, the volume function should agree with that volume. In particular, the volume of a cuboid should be the product of the lengths of its sides, and the length of a real interval (a, b) should be $b - a$.
2. The volume of a countable disjoint union of sets should be the sum of the individual volumes.
3. The volume should be invariant under isometries, i.e. functions like rotations, translations, and reflections should not change the volume of a set.

It will turn out that there exists exactly one function, called the *Lebesgue 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 7.1.1. *Every subset of \mathbb{R}^n being Lebesgue-measurable is consistent with ZF (without the axiom of choice).*

Theorem 7.1.2. *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*.

7.1.1 Vitali Sets

Proposition 7.1.3. *The relation $x \sim y \Leftrightarrow x - y \in \mathbb{Q}$ is an equivalence relation on the real numbers.*

Theorem 7.1.4. *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 . \square

Lemma 7.1.5. *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. $j = k$. \square

Lemma 7.1.6. *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. \square

Corollary 7.1.7. *Vitali sets are non-measurable.*

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

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

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

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

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

7.1.2 The Banach-Tarski Paradox

Theorem 7.1.8. *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 $A_1 \sqcup \dots \sqcup A_k = A$ and $B_1 \sqcup \dots \sqcup B_k = B$ such that for each i , A_i and B_i can be transformed into each other by an isometry.*

Corollary 7.1.9. *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 assigned a meaningful volume.*

7.2 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 only contain measurable sets.

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*:

7.2.1 Boolean Algebras

Definition 7.2.1. 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 7.2.2. Any boolean algebra also has the following properties:

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

$$\begin{aligned} a \wedge (a \vee b) &= a \\ a \vee (a \wedge b) &= a, \end{aligned}$$

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 7.2.3. 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 7.2.4. 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 $M \rightarrow X \setminus M$.

7.2.2 Lattices

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

Definition 7.2.5. 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 7.2.6. 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 7.2.7. 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 7.2.8. A boolean algebra is a bounded lattice such that meet and join are distributive over each other and such that complements exist.

7.3 Set Algebras

Definition 7.3.1. 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 7.3.2. 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 $A_1, \dots, A_n \in \mathcal{A}$, we have $\bigcup_{i=1}^n A_i \in \mathcal{A}$,
4. For any $A_1, \dots, A_n \in \mathcal{A}$, 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 7.3.3. A set algebra is a subalgebra of the power set boolean algebra on X .

Corollary 7.3.4. A topology on X is "simply" a set algebra on X that is closed under arbitrary unions.

Theorem 7.3.5. (Stone's Representation Theorem for Boolean Algebras:) Every boolean algebra is isomorphic to a set algebra.

7.3.1 σ -Algebras

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

Definition 7.3.6. 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 $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$ (\mathcal{A} is closed under the operation of taking countable unions).

If \mathcal{A} is a σ -Algebra on X , we call (X, \mathcal{A}) a **measure space**.

Corollary 7.3.7. Any σ -algebra contains the empty set and is closed under countable intersection.

Corollary 7.3.8. A σ -algebra can be more concisely defined as a set algebra that is closed under countable union and intersection, not just finite ones.

7.3.2 Set Rings

Theorem 7.3.9. Let $\mathcal{A} \subset \mathcal{P}(X)$ such that $\emptyset \in \mathcal{A}$ and such that \mathcal{A} is closed under symmetric difference and intersection. Then \mathcal{A} forms a ring (in the general algebraic sense), where addition is symmetric difference, multiplication, intersection, additive inverses are set complements and the neutral element of multiplication is X . We call such an \mathcal{A} a **ring of sets** or **set ring**.

Corollary 7.3.10. Set rings are closed under finite union and intersection.

7.3.3 Semirings of Sets

Theorem 7.3.11. Let $\mathcal{S} \subset \mathcal{P}(X)$. We call \mathcal{S} a **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 $S_1, \dots, S_n \in \mathcal{S}$ such that $A \setminus B = \bigcup_{i=1}^n S_i$.

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.

Sadly, unlike with rings of sets, there is absolutely no connection between semirings of sets and the algebraic notion of a semiring - a semiring of sets is exclusively a semi(ring of sets), and *not* a (semiring) of sets.

Semirings of sets are of fundamental importance to measure theory primarily because the *lebesgue measure* we will end up defining on \mathbb{R}^n is constructed through the approximation of sets by cuboids, which form a semiring - a complement of two cuboids is not necessarily a cuboid, but it can always be constructed from a finite union of cuboids.

7.4 Measure functions

Definition 7.4.1. Let X be an uncountable set and $f : X \rightarrow [0, \infty]$. Then we define

$$\sum_{x \in X} f(x) = \sup_{|F| < \infty} \sum_{x \in F} f(x)$$

It may seem odd that we can sum over an uncountable set by simply summing over the finite subsets - this is partly justified by the following lemma:

Lemma 7.4.2. If all terms of a sum are positive, and uncountably many of these terms are non-zero, then the sum diverges.

Proof. Assume that a sum over a set converges, i.e.

$$\sum_{x \in X} f(x) = L \in \mathbb{R}$$

Let $S_n = \{x \in X : f(x) > \frac{1}{n}\}$ for $n \in \mathbb{N}$. Then we have:

$$\begin{aligned} L &= \sum_{x \in X} f(x) \\ &\geq \sum_{x \in S_n} f(x) \\ &> \sum_{x \in S_n} \frac{1}{n} \\ &= \frac{|S_n|}{n} \end{aligned}$$

So we have $|S_n| < nL$ for all $n \in \mathbb{N}$, meaning that all S_n are finite. The set of all non-zero terms is given by:

$$S = \{x \in X \mid f(x) > 0\} = \bigcup_{n \in \mathbb{N}} \left\{ x \in X \mid f(x) > \frac{1}{n} \right\}$$

And since S is a countable union of finite sets, it must be countable. Therefore a finite limit L can't exist if S is uncountable. \square

Definition 7.4.3. If our sum is absolutely convergent, we can also define

$$\sum_{x \in X} f(x) = \sum_{x \in X, f(x) \geq 0} f(x) - \sum_{x \in X, f(x) \leq 0} |f(x)|$$

7.5 The Lebesgue Measure

Chapter 8

Integration

8.1 The Lebesque Integral

Theorem 8.1.1. *Let X be an arbitrary set. Let c be the counting measure on $\mathcal{P}(X)$. Let $f : X \rightarrow \overline{\mathbb{R}}$ be a c -measurable function. Then*

$$\int_X f \, dc = \sum_{x \in X} f(x)$$

8.2 The Bochner Integral