Notes on (pretty much) all of mathematics

Krzysztof Szymański
338068@uwr.edu.pl

July 5, 2025

Contents

Chapter 1

Naive set theory

Lemma 1.1. *Let*

$$f : X \to X$$

be a function from a finite *nonempty set into itself. Then, you can uniquely decompose the set X into disjoint parts*

$$X = X_1 \cup X_2$$

such that f restricted to X_1 is a permutation and every element from X_2 eventually gets to X_1, that is

$$\forall x_2 \in X_2 \exists k\, f^k \in X_1.$$

Proof. Define a decreasing (TODO: why) sequence of sets by

$$X^0 = 0, X^{k+1} = f[X^k].$$

Every set in the sequence is finite and nonempty. Since the sequence is decreasing, it must stabilize eventually. Let k_0 be the first index at which it stabilizes, i.e.

$$X^k = X^{k+1} = f[X^k].$$

Then

$$f : X^k \to X^k$$

is surjective, so by a previous lemma it is bijective. We take

$$X_1 = X^k,\, X_2 = X \setminus X^k$$

to form the desired decomposition.

$\square$

Visualization If we want to visualize how such a function looks, let us first recall a visualization of permutations – they form cycles! The elemnts outside of the *permutation set* X_1 form trees, whose roots are vertices of X_1.

Example 1.2. Consider the function ...

This may seem like a silly little fact, but the ideas come up often: in the Lefschetz Fixpoint theorem, Jordan decomposition from linear algebra ...

How do we extend this to infinite sets? Let's try using a fixpoint theorem! In this case, we'll be looking for the largest fixpoint of a decreasing function. Note that on infinite sets the trees mentioned earlier may be infinite (consider the example of $x \mapsto x+1$ on $\mathbb{N}$). But these are actually permutations! So maybe the structure looks the same? But how do we define the *permutation set*?

Attempt 1: take the maximal set from

$$\{A \mid f \text{ is a permutation on } A\}$$

§1.1 Useful identities on sets

Lemma 1.3. *For any two families A_α, B_α indexed by $\alpha \in \mathcal{A}$, we have*

$$\bigcup_{\alpha\in\mathcal{A}} A_\alpha \setminus \bigcup_{\alpha\in\mathcal{A}} B_\alpha \subseteq \bigcup_{\alpha\in\mathcal{A}} (A_\alpha \setminus B_\alpha)$$

§1.2 Literature used for this chpater

Chapter 2

Point-set topology

This chapter summaries basic, point-set topology. It is roughly equivalent to Chapter 2 of Munkres, without the metric space material.

§2.1 The location trilemma

Fix a topological space X, a set A an a point x. In terms of set theory, a point has two *locations* with respect to A – either $x \in A$ or $x \notin A$. In topology, we care about a little more. Although any endpoint of a closed interval $[a, b]$ certainly is *in* the interval, we would not say it is *inside* that interval. The difference becomes even starker when looking at sets like $\mathbb{N} \subseteq \mathbb{R}$ or $\mathbb{Q} \subseteq \mathbb{R}$.

The right way – or at least *a right way* – is to say that a point x is *inside* a set A, if apart from $x \in A$, a whole neighbourhood $U \ni x$ is contained in A. Since this readily implies $x \in A$, we may dispense with that requirement in any definition. Metaphorically, this means that x *cannot escape* A or *everything* x *sees is inside of* A.

The notion that $x \notin A$ has a very similar topological analogue. We will say that x *is outside of* A if not only $x \notin A$, but a whole neighbourhood $U \ni x$ is disjoint from A. Another word for that is that x *is separated from* A. This also provides the metaphorical meaning.

With these two notions lifted up from mere set-based combinatoris to topology, we have made a bit of a problem. It could be the case that for a point x and set A neither of the above properties hold. Pictorially speaking, x is then *not far from* A, but also not really *inside of* A – a kind of in-between state. In this case we call x a *boundary point of* A.

Topological properties of operators (such as Int, Cl, Bd etc.) can often be deduced from this simple observation, which we state as a lemma. This was first taught to me by Krzysztof Omiljanowski. I am not aware of any name for the fact, so I made up my own.

Lemma 2.1. *(The location trilemma.) Let* X *be a topological space,* $A \subseteq X$ *and* $x \in X$ *(not neccesarily* $x \in A$*). Then exactly one of the following is true:*

1. *For some neighbourhood* U*,* $x \in U \subseteq A$*. Then* x *is called an* interior point*.*
2. *For some neighbourhood* U*,* $x \in U \subseteq A^c$*. Then* x *is called an* exterior point*. We also say that* x *is* separated from A *by the neighbourhood* U*.*
3. *For all neighbourhoods* U*,* $U \cap A \neq \varnothing$ *and* $U \cap A^c \neq \varnothing$*. Then* x *is called a* boundary point*.*

Proof. Clause (1) implies that $x \in A$ and (2) implies that $x \in A^c$, so they cannot hold at once. Moreover both of them imply that (3) is false. We will show that if both (1) and (2) do not hold, then (3) does.

Pick any neighbourhood $U \ni x$. Since (1) does not hold, $U \nsubseteq A$, so $U \cap A^c \neq \varnothing$. Analogously, $U \cap A \neq \varnothing$. $\square$

The names we gave to the properties of points with respect to A are not random – they correspond exactly to the usual definitions of Int, Bd, Ext.

§2.2 The closure axiomatization

As one should know, a topology on X might as well be defined in terms of which sets are closed rather than which sets are open. In practice, another way of looking at closed sets might also pop up. In many scenarios, there is an operator Cl, which we might call a *closure* operator, of the signature

$$\mathrm{Cl} : \mathcal{P}(X) \to \mathcal{P}(X),$$

which adjoins to a set A some elements that are in a given sense *reachable*, *deducible*, *obtainable* etc. from A.

An example of such an operator would be, for any given topology $\mathcal{T}$ on X, the closure operator of that topology. One might wonder if from an operator one might recover a topology. If we are going to do that, there are a few questions we need to answer.

How do we recover open sets? We can get open sets as complements of closed sets. Then the question is how do we recover closed sets. The key property to use is that a set C is closed iff $\mathrm{Cl}\, C = C$.

Definition 2.2. An operator

$$\mathbf{c} : \mathcal{P}(X) \to \mathcal{P}(X)$$

is said to satisfy the Kuratowski closure axioms if it satisfies to following

TODO: add this to Anki, make a flashcard, whatever

(K1) it preserves $\varnothing$, i.e. $\mathbf{c}(\varnothing) = \varnothing$;

(K2) it is *extensive*, i.e. $A \subseteq \mathbf{c}(A)$ for all A;

(K3) it is *idempotent*, i.e. $\mathbf{c}(\mathbf{c}(A)) = \mathbf{c}(A)$ for all A;

(K4) it distributes over finite sums, i.e. $\mathbf{c}(A \cup B) = \mathbf{c}(A) \cup \mathbf{c}(B)$ for all A, B.

Now we prove that this actually defines a topology.

Lemma 2.3. *Let* $\mathbf{c}$ *be an operator*

$$\mathbf{c} : \mathcal{P}(X) \to \mathcal{P}(X)$$

satisfying the Kuratowski closure axioms. Then, the collection of its fixpoints, i.e.

$$co\mathcal{T} = \{A \mid \mathbf{c}(A) = A\}$$

defines a topology as its family of closed sets.

Proof. From the axiom (K1) we get that $\varnothing \in co\mathcal{T}_{\mathbf{c}}$. We also need $X \in co\mathcal{T}$, which follows from (K2), as no set in X can be larger than all of X. Suppose now that A, B are closed in the sens above. Then we have

$$\mathbf{c}(A \cup B) = \mathbf{c}(A) \cup \mathbf{c}(B) = A \cup B,$$

so $A \cup B$ is closed too. The intersection property is the tricky part.

Take $C_i \in co\mathcal{T}$. Then

$$\bigcap_i C_i \subseteq C_j$$

for all j, so

$$\forall j.\, \mathbf{c}\left(\bigcap_i C_i\right) \subseteq \mathbf{c}(C_j) = C_j,$$

so

$$\mathbf{c}\left(\bigcap_i C_i\right) \subseteq \bigcap_j C_j.$$

The other inclusion follows by extensivity of $\mathbf{c}$, so we actually have the equality

$$\mathbf{c}\left(\bigcap_i C_i\right) = \bigcap_j C_j.$$

□

You may have noticed that (K1), (K2) and (K4) were used in the proof, but not (K3). Then an operator satisying all of the Kuratowski axioms except for (K3) defines a topology via closed sets. On the other hand, the closure operator of that topology definitely has property (K3), so these two are different operators!

Luckily, we can recover the closure operator using lattice theory. First, a definition and a lemma.

Definition 2.4. An operator

$$\mathbf{c} : \mathcal{P}(X) \to \mathcal{P}(X)$$

is called a *Čech closure operator* if it satisfies the Kuratowski axioms (K1), (K2) and (K4).

Lemma 2.5. *Any operator satisfying (K4) is* monotonic*, i.e.*

$$A \subseteq B \Rightarrow \mathbf{c}(A) \subseteq \mathbf{c}(B).$$

for all $A, B \subseteq X$.

Proof. TODO! □

Now that we've seen these, we'll want to recover the actual A. If you think of the operator $\mathbf{c}$ as *enriching* a set, a closed set is one that *is completely enriched.* Then, how do we get $\operatorname{Cl} A$? Well, we start by enriching it

$$A \to \operatorname{Cl} A,$$

but this may not be enough, so we enrich the result and get

$$A \to \operatorname{Cl} A \to \operatorname{Cl}(\operatorname{Cl} A),$$

but this may not be enough, so we enrich the result

$$A \to \operatorname{Cl} A \to \operatorname{Cl}(\operatorname{Cl} A) \to \operatorname{Cl}(\operatorname{Cl}(\operatorname{Cl} A)),$$

and... we're stuck in an infinite loop. In such cases, it helps to take a peek past infinity and consider

$$\bigcup_{k=0}^{n} \operatorname{Cl}^k A.$$

This still does not work.

Example 2.6. You can define a Cech closure operator which fails this. Let

$$X = \mathbb{N} \cup \{\infty\}$$

and

$$\mathbf{c}(A) = \begin{cases} \varnothing & \text{when } A = \varnothing \\ X & \text{when } \infty \in A \\ \{0, 1, \ldots, \sup A + 1\} & \text{when } A \neq \varnothing \text{ and } \infty \notin A \end{cases}$$

You can check that this is a Cech closure operator, but the only closed sets in $\mathbf{co}\mathcal{T}_{\mathbf{c}}$ are $\varnothing$ and X, i.e. this is the indiscrete topology on X.

In reality you need a transfinite number of iterations. However, for a Kuratowski operator this chain stabilizes almost immediately, so we're saved!

Lemma 2.7. *For a Kuratowski closure operator* $\mathbf{c}$*, the closure operator of its generated topology* $\mathcal{T}_{\mathbf{c}}$ *is precisely* $\mathbf{c}$*.*

Proof. Let $C \supseteq A$ be the topological closure of A. Then

$$\mathbf{c}(A) \subseteq \mathbf{c}(C) = C.$$

Since $\mathbf{c}$ is idempotent by (K3), $\mathbf{c}(A)$ is closed in $\mathcal{T}_{\mathbf{c}}$. Since a topological closure is the smallest closed set containing A, we have that

$$C \subseteq \mathbf{c}(A) \Rightarrow C = \mathbf{c}(A)$$

□

Connection with lattice theory. Note that we want to find the least (because topological closure is the smallest closed set) fixpoint (because of the definition of topology) of $\mathbf{c}$ greater than A. This is exactly the setting of the Kleene fixpoint theorem, and what we're doing here is using that.

§2.3 Literature used for this chapter

1. Munkres, chapter 2
2. Wikipedia, Kuratowski Closure Axioms
3. Wikipedia, Preclosure operator

Chapter 3

Metrisable spaces

§3.1 Countability properties

Many countability properties are equivalent for metrisable spaces. In particular

Ankified **Theorem 3.1.** *Let X be a metrisable topological space. The following are equivalent:*

1. *X is separable,*
2. *X is second countable,*
3. *X is Lindelöf.*

Proof ($1 \Leftrightarrow 2$). Let x_n be a countable dense set. We will show that $B(x_n, q)$ for $n \in \mathbb{N}$ and $q \in \mathbb{Q}_+$ is a basis. Take an open set U and arbitrary $x \in U$. Since U is open in a metric space, $B(x, r) \subseteq U$ for some $r > 0$. Take an element x_n of the dense set such that $d(x, x_n) < r/4$. Then

$$x \in B(x_n, q) \subseteq B(x, r) \subseteq U$$

for any $r/4 \leqslant q \leqslant r/2$, so the given balls are indeed a basis.

The other direction is trivial – pick any element is each of the (countably many) basis sets. $\square$

Proof ($1 \Rightarrow 3$). Take an open cover $\mathcal{U}$. By an argument analogous as in the proof above, for all $x \in X$ there is a ball $x \in B(x_n, q) \subseteq U$ for some $U \in \mathcal{U}$. Since there are only countably many such balls, countably many U from $\mathcal{U}$ suffice to cover all of X. $\square$

Remark. The proof of $3 \Rightarrow 1$ is 16.3.

Remark. The presented proof of ($1 \Rightarrow 3$) is *de facto* a composition of the proof of ($1 \Rightarrow 2$) and a proof that *any*, not only metrisable, secound countable space is Lindelöf.

Lemma 3.2. *A second countable topological space X is Lindelöf.*

Proof. Take an open cover $\mathcal{U}$ and a countable basis $\mathcal{B} = \{B_i : i \in \mathbb{N}\}$. For any $x \in U \in \mathcal{U}$, there is a basis set $B_i \in \mathcal{B}$ such that

$$x \in B_i \subseteq U.$$

For each B_i that *appears* in this way we can pick just one $U \in \mathcal{U}$ containing B_i, so there is a countable subcover of $\mathcal{U}$. What might not be clear is why this is a cover. Note that since $\mathcal{B}$ is a basis, for each $x \in X$ there will be a basis set containing it that *appears* in the supposed subcover, so it is indeed a subcover.

In more abstract terms, $\mathcal{U}$ has an inscribed cover which is a subfamily of $\mathcal{B}$. This inscribed cover is countable, and therefore there is a countable subcover of $\mathcal{U}$. □

What about first countability? You might wonder if first countability together with separability are enough to ensure second countability. I have a feeling that this fails, however. In the proof for metrisable spaces we not only use that the balls form a symmetric local basis, but this *system* of local bases is *symmetric*, i.e.

$$x \in B(y,r) \iff d(x,y) < r \iff d(y,x) < r \iff y \in B(x,r).$$

Local connectedness is enough to ensure that a system of symmetric first-countable bases can be built via the chain trick 3.3. Oops, this is not true XD.

There is a counterexample: the Sorgenfrey line $\mathbb{S}$. It has a countable local basis $[x,\, x+1/n)$, countable dense set Q, but no countable basis.

Lemma 3.3 (Finite chain trick)**.** *In a connected space X with a connected/open (CHECK) cover $\mathcal{S}$, every element x can be reached from any other element y through a finite chain of intersecting $S_i \in \mathcal{S}$.*

Chapter 4

Abelian groups

We will now explore some features of abelian groups. These will later be recontextualized into features of modules (for those of you know, an abelian group is essentialy the same as a $\mathbb{Z}$-module). In this chapter, we will derive a structure theorem.

§4.1 Basic properties

§4.2 Torsion

Central to the study of abelian groups is the notion of torsion. We begin with two definitions, one for elements and one for groups.

Definition 4.1. An element of an abelian group is called a **torsion element** if it is of finite order. An element of infinite order is called **torsion-free**.

Definition 4.2. An abelian group is called a **torsion group** if all its elements are torsion. An abelian gropu is called **torsion-free** if all its *nonzero* elements are torsion-free.

One might ask why we define this only for abelian groups. The following crucial fact is the reason – more precisely, the fact that its proof relies on commutativity.

Theorem 4.3. *If α, β are two torsion elements then their sum (and difference) is also a torsion element. Thus, the torsion elements form a subgroup.*

Proof. Let α have order n and β have order m. Then we have

$$nm(\alpha+\beta) = nm\alpha + nm\beta = m(n\alpha) + n(m\beta) = m\cdot 0 + n\cdot 0 = 0.$$

The subgroup claim is proved by noting that the same computation hold as well for difference and that the identity element is of course torsion. □

Since we have a well defined kind of subgroup, we might as well give it a name.

Definition 4.4. Let G be an abelian group. Its **torsion subgroup**, denoted

$$\mathrm{T}(G) \mathrm{or} Tor(G)$$

is the subgroup consisting of all the torsion elements of G.

We may hope that such a nice notion of substructure is also valid for the torsion-free elements of a group. This fails, however, as the set of torsionfree elements is the complement of a subgroup, which will very often fail to be a subgroup. On the level of elements, we have the following fact.

Theorem 4.5. *Let $a \in G$ be torsionfree and $b \in Tor(G)$. Then $a + b$ is torsionfree.*

Proof. We give two proofs. For the first one, suppose $a + b$ is torsion. Then

$$a = (a + b) - b$$

would be torsion (since the torsion elements form a subgroup).

For the second proof, suppose $a + b$ is torsion with order n and that b has order m. Then

$$0 = mn(a + b) = mna + mnb = mna,$$

so a is torsion as well. □

All is not lost though! We cannot define a notion of *torsion-free subgroup*, but there is a simple way of killing the torsion – quotienting!

4.2.1 A weird example of groups

4.2.2 Functors

4.2.3 Tensors

§4.3 Structure theorem

§4.4 A smudge of infinite abelian group theory: the prufer groups

Prufer groups are: divisible, p-torsion, subgroups, infinitely generated

§4.5 Sources for this chapter

1. Ludomir Newelski, Algebra II
2. Ludomir Newelski, Algebra 2R
3. Wikipedia: torsion subgroup, torsion, torsion-free group, Prufer groups

Chapter 5

The Künneth formula

This chapter accomplishes a lofty goal: to calculate the homology groups of a product topological space.

§5.1 Setup

We will first try and discover how this all works for CW-homologies, and then use what we've found to generalize the statements to singular homology. The reason why CW-complexes are good for the job is that given CW-structures on X and Y, we can easily form a CW-structure on $X \times Y$. One should expect to form a simplicial structure on a product of simplicial complexes, but that would require subdividing a product of n- and m-simplices into $(n+m)$-simplices, which, if anything, is unpleasant.

We have a product structure on $X \times Y$ given by products of cells and products of characteristic maps. Let us now try to calculate the cellular boundaries.

One of the cells is larger in dimension: then we get a coefficient of zero, because the map turns out to be constant.

If the above does not hold, then we have to reduce exactly one dimension by one. However, if in the dimension we keep we use a different cell of that dimension, the same problem applies.

Generally, we have the following:

Lemma 5.1. *Suppose we have a continuous map*

$$f : I^k \to I^{k-1},$$

which is equal to a product map

$$f = id_I \times g$$

for some continuous

$$g : I^{k-1} \to I^{k-2}.$$

Then, the degree of

$$\widetilde{f} : \operatorname{Bd} I^k \to I^{k-1} / \operatorname{Bd} I^{k-1}$$

is equal to the degree of

$$\widetilde{g} : \operatorname{Bd} I^{k-1} \to I^{k-2} / \operatorname{Bd} I^{k-2}.$$

Remark. Note that since this is not the same sphere, the degree only makes sense *modulo orientation*.

§5.2 Orientations

Remember that when describing a CW-complex structure, there is an additional degree of freedom – the orientation of each cell. This orientation needs to be taken into account.

For the purposes of this chapter, it is very important to see how an orientation of the interior can be turned into an orientation of the boundary.

One way of doing this would be to just take the *vector to the right and up* in an Euclidean space. However, this leads to an inconsistent choice of orientation on the boundary. A correct way is to pick an inward facing normal at each point on the boundary and then pick an orientation on the boundary which

This can formally be described in homology by noting that the boundary homomorphism is an isomorphism. The determined choice of generator for I^k then projects down to $\operatorname{Bd} I^k$.

§5.3 Literature used for this chapter

1. Allen Hatcher, Algebraic Topology

2. MathOverflow, biography of Hermann Kunneth: https://mathoverflow.net/questions/114215/who-was-hermann-k

3. https://sili-math.github.io/AT2020/Lecture-22.pdf

Chapter 6

Local properties in algebra

§6.1 Localization of modules

§6.2 Local rings

This section is devoted to a class of rings very important in algebraic geometry – local rings. Although the definition is purely algebraic, it corresponds to notions of *looking around a point* on a (affine) variety.

6.2.1 Characterisations of local rings

All material in this section applies to (unital) rings. They need not be commutative for the results to work. All the proofs in this section are my own work.

Definition 6.1 (local ring)**.** A ring R having a unique maximal ideal $\mathfrak{m}$ is called a **local ring** and denoted $(R, \mathfrak{m})$.

It turns out there is another characterisation of such rings, which is of particular usefulness in working with the second point of interest of this chapter – Discrete Valuation Rigs, whose definition we postpone for now.

Theorem 6.2 (local ring characterisation)**.** *A ring R is local iff the set of noninvertible elements forms an ideal. This is the the unique maximal ideal.*

Proof. The first implication is trivial, since an ideal containing a unit is not proper.

Suppose now R is local with maximal ideal $\mathfrak{m}$. Since $\mathfrak{m}$ is proper, it does not contain any invertible element. Take a noninvertible $r \in R$. The principal ideal (r) can be extended (via Zorn's lemma) to a maximal ideal, which must be the unique maximal ideal $\mathfrak{m}$, so $r \in \mathfrak{m}$.

□

There is also an element-wise characterisation of local rings. To see it emerge, let us try a different line of attack for the previous proof.

We know that $\mathfrak{m}$ contains only noninvertible elements, but we still have to prove that it contains all such elements. By way of contradiction, suppose $r \notin \mathfrak{m}$. Now by maximality we have

$$\mathfrak{m} + (r) = R$$

and in particular

$$m + ar = 1$$

for some $a \in R, m \in \mathfrak{m}$. We would like to derive a contradiction, so we would hope that the identity above lets us conclude that r is invertible. The property below is what we need

Theorem 6.3 (local ring addition)**.** *A ring R is local iff it has the following property: for all a, b such that*

$$a + b \in R^\star$$

at least one of a, b is invertible.

Proof. The line of attack above shows how this implies R being local. Once we have this property, we can take any maximal ideal. If there was some noninvertible element it did not contain, we would be able to extend it.

On the other hand, take a local ring. We will show it is impossible to sum two noninvertibles to a unit. We have that

$$(a) + (b) \subseteq \mathfrak{m} + \mathfrak{m} = \mathfrak{m},$$

which contains no units.

□

It should be noted that the property expressed in the previous theorem might be rephrased in two ways: first, take an arbitrary finite sum instead of two elements and second, have the sum be equal to 1 instead of invertible. It should be easy to see that all such characterisations are equivalent.

6.2.2 Examples and generic constructions

6.2.3 Properties of local rings

Krull's intersection, Nakayama's, Kaplansky's theorem

§6.3 Valuations on Fields and Discrete Valuation Rings

§6.4 Literature used for this chapter

1. Piotr Kowalski, Algebraic Curves. Chapter 2.3.

2. nlab page for local rings

3. https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=2933&context=td (for Kaplansky's theorem)

4. Wikipedia pages

Chapter 7

Algebraic Geometry, problemset 5

Problem 4

Since v is a smooth point of C, the ring $\mathcal{O}_v$ is a DVR. A local parameter is defined as any uniformizing parameter of that DVR and having a zero of order 1 means being of valuation 1 in that DVR.

Thus is suffices to prove the following: in a DVR, an element is of valuation 1 iff it is irreducible.

Since a DVR is a local ring, let us denote its maximal ideal by $\mathfrak{m} = (r)$. Note that r must be irreducible, otherwise the ideal $\mathfrak{m}$ would not be maximal. The valuation on this DVR can be described as the usual r-adic valuation or belonging to a power of the maximal ideal (see solution of **Problem 7.**).

Let f be a local parameter, i.e. irreducible. Then, by **Remark 2.37.**

$$(f) = \mathfrak{m} = (r),$$

so f and r are associates, therefore $f = ur$ for some unit u. We have that $u \in \mathcal{O}_v \setminus \mathfrak{m}$, so $v(u) = 0$ and

$$v(f) = v(u) + v(r) = 0 + 1 = 1.$$

On the other hand, let f be of valuation 1. Then, as the valuation is the r-adic valuation we have that

$$f = ra$$

for some $r \nmid a$, so $a \in \mathcal{O}_v \setminus \mathfrak{m}$. This implies a is a unit, so f and r are associated, so f is irreducible since r is.

Problem 5

Problem 5a

Consider a tangent line $L = V(\alpha X + \beta Y + \gamma)$.

If L is tangent, the intersection number $I(0, L \cap C) > 1 > 0$, so $0 \in L$, so $\gamma = 0$. By the same reasoning, F has zero constant term.

It is easy to see that a line is a smooth curve. Therefore, the curve C is tangent to L iff F is of valuation at least two in $\mathcal{O}_0$ (which is a DVR). By **Problem 7** that is equivalent to F being a member of $\mathfrak{m}_v^2$ (where F is reinterpreted as $F + I(L)$ in $K(L)$).

Algebraically, the square of the maximal ideal corresponds to

$$I_L^2(0)\mathcal{O}_0 = \left\{ \frac{G}{H} : G \in I_L^2(0), H \notin I_L(0) \right\}.$$

Thus, F being of valuation at least 2 is equivalent to it being of the form

$$F = \frac{1}{N} \cdot \sum_i G_i H_i$$

as a rational function in $K(L)$ for some $N \in K[L] \setminus I_L(0)$ and $G_i, H_i \in I_L(0)$. This is again equivalent to the identity

$$FN = \sum_i G_i H_i$$

in $K[L]$, which in turn is equivalent to F having the form

$$FN = \sum_i G_i H_i + P(\alpha X + \beta Y)$$

for some polynomials such that $N(0) \neq 0$ and $G_i(0) = H_i(0) = 0$.

After this introduction, we will show

$$T_0 C = V\left(\partial_X F(0) X + \partial_Y F(0) Y\right)$$

via two inclusions, first from left to right.

If both partials are 0, the vanishing set is the whole space $\mathbb{A}^2$, so this inclusion is trivial. Suppose now at least one partial is nonzero. Differentiating both sides of the above identity w.r.t. X we get that

$$\partial_X F \cdot N + F \cdot \partial_X N = \left(\sum_i \partial_X G_i \cdot H_i + G_i \cdot \partial_X H_i \right) + \partial_X P \cdot (\alpha X + \beta Y) + \alpha P.$$

Evaluating both sides at 0 and remembering which polynomials vanish at 0 we obtain

$$N(0)\partial_X F(0) = P(0)\alpha.$$

We can now repeat this for differentiation w.r.t. Y. This implies that the vectors

$$[\partial_X F(0), \partial_Y F(0)]^T, [\alpha, \beta]^T \in \mathbb{K}^2$$

are linearly dependent, so they describe the same line.

Now for the other inclusion. We will lean on the fact that the first partials evaluated at zero are the coefficients of the monomials X, Y in the polynomial F. Suppose first both partials of F are zero. Then we can write F as

$$F = X^2 G(X,Y) + XYH(X,Y) + Y^2 P(X,Y) + 0 \cdot (\alpha X + \beta Y)$$

which is the form we needed. This gives that *any* line is tangent to C at 0, so the tangent space is the whole plane (as is $V(0 \cdot X + 0 \cdot Y)$).

If, on the other hand, some partial is nonzero, we can write F as

$$F = 1 \cdot (\partial_X F(0) \cdot X + \partial_Y F(0) Y) + X^2 G(X,Y) + XYH(X,Y) + Y^2 P(X,Y).$$

This proves that C is indeed tangent to $V(\partial_X F(0) \cdot X + \partial_Y F(0) Y)$.

Problem 5b

From the previous subproblem and the lecture we know that both T_0C and $I_C(0)/I_C(0)^2$ are finite-dimensional K-vector spaces. This allows us to use a theorem of linear algebra which states that a bilinear map such as is given in the problem induces an isomorphism iff for all $P \neq 0$ the function

$$x \mapsto \Phi(x, P)$$

is nonzero (i.e. is nonzero for some x) or, equivalently, that if this function is zero then so is P. Here P should be understood as the representative regular function in $K[C]$ or a polynomial which represents that function (so a representative of representatives).

Suppose then that this function is zero. First consider the case that both partials of F are zero. Then the tangent space is the whole plane and we have that for all $x, y \in K$

$$\partial_X P(0)x + \partial_Y P(0)y = 0,$$

which gives that both partials of P vanish at zero, so we have the form

$$P = X^2P_1 + XYP_2 + Y^2P_3$$

for P as a polynomial, which in particular implies

$$P \in I_C(0)^2$$

as a regular function, so

$$P = 0$$

in $I_C(0)/I_C(0)^2$.

Now consider what happens when F has a nonzero partial. This means that

$$\partial_X P(0)x + \partial_Y P(0)y = 0$$

for all points such that

$$\partial_X F(0)x + \partial_Y F(0)y = 0,$$

which gives that either the gradient of P is zero (in which case we can repeat the reasoning from the previous case) or that the gradients of P and F are linearly dependent. Then there exists a scalar α such that

$$P - \alpha F$$

has zero first partials, so

$$P = \alpha F + X^2P_1 + XYP_2 + Y^2P_3$$

as a polynomial, so

$$P = X^2P_1 + XYP_2 + Y^2P_3 + I(C)$$

as a regular function. As we have done many times, we now conclude that

$$P \in I_C(0)^2,$$

which is what we needed.

Problem 6

Problem 6a

Let

$$a = r^n \frac{\alpha}{\beta}, b = r^m \frac{\gamma}{\delta},$$

with $r \nmid \alpha, \beta, \gamma, \delta$ and wlog $n \geqslant m$. Then

$$a + b = r^m \frac{r^{n-m}\alpha\delta + \beta\gamma}{\beta\delta}.$$

Note that r does not divide the numerator (since r is irreducible and thus prime in a UFD). If $n > m$, it will also not divide the denominator, but if $n = m$ it might. In the first case the valuation is exactly m, while in the second case it might change – but only increase.

Problem 6b

We have

$$ab = r^{n+m} \frac{\alpha\gamma}{\beta\delta}.$$

By virtue of r being prime, we have $r \nmid \alpha\gamma\beta\delta$, so

$$v_r(ab) = n + m.$$

Problem 6c

For every $n \in \mathbb{Z}$ we have

$$v_r(r^n) = n,$$

so the valuation is indeed surjective.

Problem 7

We claim that

$$\mathfrak{m}^n = \{x : v_R(x) \geqslant n\},$$

from which the problem follows immediately. For $n = 0$ the claim follows from the nonnegativity of the valuation.

Let $r \in R$ be the uniformizing parameter. Then we have that

$$\mathfrak{m} = (r)$$

by **remark 2.37.**, so an element of $a \in \mathfrak{m}^n$ is of the form

$$a = \sum_i a_i r^n$$

for some $a_i \in R$, so

$$v_R(a) \geqslant \min(v_r(a_1 r^n), v_r(a_2 r^n), \ldots, v_r(a_j r^n)) \geqslant \min(n, n, \ldots n) = n.$$

This gives us one inclusion of the claim. For the other one, take an element x of valuation no less than n. The definition of valuation implies that for some $m \geqslant n$

$$x = r^m y = (r^{m-n} y) r^n \in \mathfrak{m}^n.$$

This completes the solution.

Problem 8

Take $x, y \in \mathcal{O}_v$. Then we have

$$v(xy) = v(x) + v(y) \geqslant v(y) \geqslant 0$$

and

$$v(x + y) \geqslant \min(v(x), v(y)) \geqslant 0.$$

This implies that $\mathcal{O}_v$ is a subring of L. If we take x, y such that both valuations are positive, the sum has a positive valuation, and if at least y has a positive valuation then the product does as well. This implies that $\mathfrak{m}_v$ is an ideal.

Note that for any valuation we have

$$v(1) = v(1 \cdot 1) = v(1) + v(1),$$

so

$$v(1) = 0.$$

Now take any $x \in \mathcal{O}_v \setminus \mathfrak{m}_v$. This means that $v(x) = 0$. Let y be the multiplicative inverse (in L!) of x. Then

$$0 = v(1) = v(xy) = v(x) + v(y),$$

so $v(y) = 0$ and $y \in \mathcal{O}_v$. If $v(x) > 0$, then $v(y) < 0$ and $y \notin \mathcal{O}_v$. We have just proved that the set of noninvertible elements of $\mathcal{O}_v$ is an ideal. Therefore $(\mathcal{O}_v, \mathfrak{m}_v)$ is a local ring.

Since the valuation is surjective, $\mathfrak{m}_v$ is nonempty and $\mathcal{O}_v$ has nonzero noninvertible elements, so it is not a field. To finish the proof that $\mathcal{O}_v$ is a DVR all we need to do is show that $\mathcal{O}_v$ is a PID.

To do that, take r to be any element of valuation 1. Such an element exists by surjectivity of the valuation.

Claim. Let $t \in \mathcal{O}_v$ and $v(t) = n \geqslant 0$. Then $t = ur^n$ for some unit $u \in \mathcal{O}_v$.

Proof. Let

$$\alpha = \frac{t}{r^n} \in L.$$

Then α has valuation 0, so it actually is an invertible element of $\mathcal{O}_v$.

Now take an ideal I. Note that if I has any element of valuation n, then by the claim above it contains all elements of valuation n as they all are associated with r^n. It will also contain an element (so all elements) of higher valuations by virtue of being closed under multiplication by r.

This lets us conclude that

$$I = (r^n)$$

where n is the smallest valuation achievable by an element of I.

Chapter 8

Approximation properties of smooth functions

Smoothness is a really good property that a function can have. However, a lot of the functions we need aren't smooth, i.e. ReLU, $\max(0, \min(1, x))$, $\chi_{[a,b]}$ and so on. In this chapter, we will show how to effectively approximate such functions with very good smooth lookalikes.

§8.1 Characteristic functions on intervals

Let us make our way towards approximating the interval characteristic function. In the special case of the real line, the story begins with quite a wonderful function – the magnificent

$$e^{-1/x^2}.$$

Taken literally, it's not quite defined at 0. However, as $x \to 0$, $-1/x^2 \to -\infty$, so we can define the value at 0 to be 0. Notice that the convergence of the argument to $-\infty$ is really fast – fast enough the we can retain smoothness at 0. Let us state the properties formally.

Lemma 8.1. *The function*

$$g(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is smooth and satisfies the following properties:

1. $g(0) = 0$,
2. $g(x) \leq 1$,
3. $\lim_{x\to 0} g(x)/P(x) = 0$ *for any* $P \in \mathbb{R}[X]$,
4. $g^{(n)}(0) = 0$ *for all* n.

Proof. Properties (1) and (2) should be obvious. For (3) we perform the change of variables $u := 1/x^2$ to obtain

$$\lim_{x\to 0} \frac{g(x)}{P(x)} = \lim_{u\to\infty} \frac{e^{-u}}{P(\pm 1/\sqrt{u})} = \lim_{u\to\infty} \frac{u^{\deg P/2} e^{-u}}{Q(\pm\sqrt{u})} = 0,$$

where $Q \in \mathbb{R}[X]$ (actually, it's P with the coefficients written backwards). We use (3) to derive an (4). By induction, we have that at all $x \neq 0$

$$g^{(n)} = \frac{g(x)P_n(x)}{Q_n(x)}$$

for some $P, Q \in \mathbb{R}[X]$ and the only root of Q is 0. Indeed,

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{g(x)P_n(x)}{Q_n(x)} = \frac{-\frac{1}{x^2}Q_n(x)P_n(x)g(x) + P_n'(x)Q_n(x)g(x) - P_nQ_n'g(x)}{Q_n^2(x)} = g(x)\frac{P_{n+1}(x)}{Q_{n+1}(x)},$$

where we have expanded the fraction by x^2 to get the desired form. This gives that

$$Q_n = x^{a_n},$$

where

$$a_0 = 0$$
$$a_{n+1} = 2a_n + 2,$$

from which we can derive

$$a_n = 2(2^n - 1).$$

□

The magic of this function is that, since it has all the derivatives at 0 equal to 0, we might as well say it's zero on all negative numbers *and still get a smooth function*! This is very much not the case for more primitive attempts to do this – look at the case of x or more generally x^k and see that it only gives as a function of class C^{k-1}.

This maneover let us discard half of the real line a smooth fashion, and this means we're halfways towards approximating $\chi_{[0,1]}$. The idea will be to use functions like this:

$$g(x)g(1-x).$$

This has a small bump, but since g is very flat near 0, the product of two such functions will not be very large. The solution solution to this problem is to widen the interval for the moment, so that values closer to 1 get multiplied together in the formula above. Let us define

$$h_n = g(x)g(n-x).$$

This should look better, but now

$$\operatorname{supp} h_n = [0, n]$$

where we would much prefer

$$\operatorname{supp} h_n = [0, 1].$$

No worries – we can see that the *shape* is right, so squeezing the function will do.

Lemma 8.2. *The function sequence*

$$f_n(x) = h_n(nx) = g(nx)g(n - nx).$$

converges to $\chi_{[0,1]}$ *pointwise a.e. (outside of* $\{0,1\}$*) and in* L^1*.*

Proof. Let us quantify the idea that f_n *is almost* 1 *on almost all of* $[0,1]$. In terms of h_n, these functions should be almost 1 on almost all of $[0,n]$. We will pick an α (varying with n), for which we will be able to establish a good bound. Suppose

$$\alpha n \leq x \leq (1-\alpha)n.$$

This gives

$$\begin{aligned}
\frac{1}{\alpha^2 n^2} &\geq \frac{1}{x^2} \\
\frac{2}{\alpha^2 n^2} &\geq \frac{1}{x^2} + \frac{1}{(n-x)^2} \\
-\frac{2}{\alpha^2 n^2} &\leq -\frac{1}{x^2} - \frac{1}{(n-x)^2} \\
e^{-\frac{2}{\alpha^2 n^2}} &\leq h_n(x)
\end{aligned}$$

We now want to pick a sequence α such that both

$$\lim_n \alpha = 0$$

and

$$\lim_n \frac{2}{\alpha^2 n^2} = 0.$$

Of course, $\alpha = 1/\sqrt{n}$ does the trick nicely. We can get better convergence estimates for a different choice. □

Another kind of useful function we might like to have are piecewise linear functions. To effectively approximate those, we need just see how to approximate ReLU and ... – all piecewise linear functions are sums of shifted/scaled combinations of these two.

§8.2 Higher dimensions

§8.3 Partitions of unity

§8.4 Approximative identities on $\mathbb{R}$

Now we will try to develop a more general way of approximating functions.

§8.5 Banach algebras

§8.6 Literature used for this chapter

Chapter 9

Differential Geometry, Lecture 1

Lemma 9.1. *Let*

$$f, g : (a, b) \to \mathbb{R}^n$$

be two differentiable functions and $\langle -, - \rangle_B$ *be a bilinear form. Then the derivative of the* B*-dot product is*

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle f, g \rangle_B = \left\langle \frac{\mathrm{d}f}{\mathrm{d}t}, g \right\rangle_B + \left\langle f, \frac{\mathrm{d}g}{\mathrm{d}t} \right\rangle_B .$$

If the form is symmetric and $f = g$*, we have additionally*

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle f, f \rangle_B = 2 \left\langle \frac{\mathrm{d}f}{\mathrm{d}t}, f \right\rangle_B .$$

Proof. Let the matrix of the form be

$$[B] = b_{i,j}.$$

Then

$$\langle f, g \rangle_B = \sum_{i,j} b_{ij} f_i g_j .$$

□

Lemma 9.2. *The distance between two points on the evolvent is given by*

$$(c(t_1), c(t_2)) = \frac{|\kappa_\gamma(t_2) - \kappa_\gamma(t_1)|}{|\kappa_\gamma(t_1)\kappa_\gamma(t_2)|} = |r_1$$

Observation. Curvature is the determinant of the Frenet frame.

Observation. Walking the curve backwards negates the curvatues, but does not change its derivative!

Chapter 10

Differential Geometry, Lecture 2

§10.1 Curves in $\mathbb{R}^3$

Let $\gamma : (a,b) \to \mathbb{R}^3$ be parametrised by arclength and $\gamma'' \neq 0$. Then

$$\gamma' \perp \gamma''$$

and we can define

$$N := \frac{\gamma''}{||\gamma''||}.$$

We need a third vector, ideally so that the three vectors are positively oriented. No worries, we can put down

$$B := \tau \times N.$$

Definition 10.1. We define the curvature in $\mathbb{R}^3$ to be

$$\kappa_\gamma := ||\gamma''||$$

.

Definition 10.2. We define *trójnóg Freneta* to be the ordered orthonormal basis

$$(\tau, N, B).$$

We have an analogue of the Frenet equations

Theorem 10.3 (3D Frenet Equations)**.** *For the* Frenet trójnóg *we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}(T, N, B) = (T, N, B)M$$

Proof. The first row follows from the definition. Let

$$A = (T, N, B).$$

This is an orthonormal matrix for all $t \in (a,b)$, so it traces a curve in SO(3) and we want its derivative. We will show

$$A^{-1}A'$$

is skew-symmetric – this will give us the theorem. We know that

$$AA^T = \mathrm{Id},$$

so let us differentiate both sides and obtain

$$(A')^T A + A^T A' = 0$$

and

$$(A^T A')^T = (A')^T A = -A^T A'.$$

Where we have used the matrix Leibniz rule. □

Definition 10.4. The function τ mentioned in the previous theorem is called **torsion**.

We have another analogue to curves in $\mathbb{R}^2$.

Theorem 10.5 (Fundamental Theorem of Curves is $\mathbb{R}^3$)**.** *For a positive function*

$$\kappa : (a, b) \to \mathbb{R}_+,$$

an arbitrary function

$$\tau : (a, b) \to \mathbb{R}$$

and any positively oriented orthonormal basis of $\mathbb{R}^3$ (the initial Frenet frame), there exists exactly one arclength parametrised curve

$$\gamma : (a, b) \to \mathbb{R}^3$$

which has Frenet frame at $t = 0$ equal to B.

Proof (sketch). As in the 2-dimensional case – put down a differential equation. □

What is torsion? The torsion describes how much a curve escapes a plane in which it starts (the *osculating plane*).

§10.2 Curvature of surfaces

Consider a surface

$$S \subseteq \mathbb{R}^3.$$

We will want to define its curvature in terms of plane curvature. We will work with surfaces given as immersions

$$\Sigma : U \to \mathbb{R}^3,$$

where $U \subseteq \mathbb{R}^2$.

Definition 10.6. Let $p \in U \subseteq \mathbb{R}^2$ and Σ be as above. The **Riemann metric** on TU is the smooth, positive definite bilinear form given at p by

$$\langle v_p, w_p \rangle_\Sigma = \langle \mathrm{d}\Sigma(v_p), \mathrm{d}\Sigma(w_p) \rangle .$$

Notation. We will denote by

$$\partial_x = \mathrm{d}\Sigma(\partial_x)$$
$$\partial_y = \mathrm{d}\Sigma(\partial_y)$$

the push by Σ of the standard vector field on $\mathbb{R}^2$.

Definition 10.7. The **surface normal vector** to the surface Σ at p is

$$n_p = \frac{\partial_x \times \partial_y}{||\partial_x \times \partial_y||}.$$

Now we can start talking about curvature!

Definition 10.8. For $v \in T_p\Sigma$ of length 1, the **curvature**

$$\kappa_p(v)$$

is equal to the plane curvature κ at p of the curve, which is the intersection of Σ and the plane

$$\mathrm{Lin}\{n_p, v\}.$$

(This is a set – the parametrization can be recovered with aid of the inverse function theorem).

Since we have defined the above notion of curvature for $||v|| = 1$, we have a function

$$\kappa : S^1 \to \mathbb{R}$$

10.2.1 Example of surface curvature

Let us isometrically change the coordinate system so that $p = 0$ and take

$$T_p\Sigma = \mathrm{Lin}\{e_1, e_2\}.$$

Problem. This can be done so that in a neighbourhood of 0 the surface looks like the graph of

$$f = \frac{1}{2}\alpha x^2 + \frac{1}{2}\beta y^2 + O(x^3 + y^3).$$

Let $v = (\cos\theta, \sin\theta, 0)$. We have that

$$n_p = (0, 0, 1)$$

and so

$$\gamma_v(t) = f(tv) = \frac{1}{2}\alpha\cos^2\theta + \frac{1}{2}\beta\sin^2\theta + O(t^3) = ct^2 + O(t^3).$$

The curvature of this after going back to the plane is the same as the curve

$$g(t) = ct^2.$$

What is the curvature of $g(t)$ at $t = 0$? After applying a homothety of scale c^{-1} we get a curve $h(t) = t^2$, which has curvature 2 at $t = 0$. By considering osculating circles, the curvature of $g(t)$ is

$$2c = \alpha\cos^2\theta + \beta\sin^2\theta.$$

The extreme values of this are α, β obtained for $v = (1, 0, 0)$ and $v = (0, 1, 0)$. Note that the vectors are orthogonal!

10.2.2 Why will we multiply these?

Picture one – immersing a square into $\mathbb{R}^3$ as a cylinder.

§10.3 Euclidean connection

Take two vector fields X, Y and a surface Σ. We have

$$D_X Y = (X(Y_1), X(Y_2), X(Y_3)).$$

A part of that will be tangent to Σ, and part will be normal. Writing this decomposition down defines

$$D_X Y = \nabla_X Y + \mathbb{I}(X, Y) n_p.$$

Lemma 10.9 (Propeties of ∇). *For any surface Σ, vector fields X, Y, Z and function f we have*

1. $\nabla_{fX} Y = f \nabla_X Y$
2. ∇agrees with $\langle -, - \rangle$, *i.e.*

$$X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle .$$

3. *the connection ∇ is torsion-free, i.e.*

$$\nabla_X Y - \nabla_Y X = [X, Y],$$

4. *the second fundamental form is symmetric and bilinear, i.e.*

$$\mathbb{I}(fX, Y) = f\mathbb{I}(X, Y)$$

$$\mathbb{I}(X, Y) = \mathbb{I}(Y, X).$$

Proof. For the first one,

$$D_{fX} Y = f D_X Y = f \nabla_X Y + f \mathrm{II}(X, Y) n.$$

For the second one

$$D_X(fY) = X(f)Y + f D_X Y = X(f)Y + f \nabla_X Y + f \mathrm{II}(X, Y) n.$$

For (2)

$$X \langle Y, Z \rangle = X \left(\sum_i Y_i Z_i \right) = \sum_i X(Y_i) Z_i + Y_i X(Z_i) = \langle D_X(Y), Z \rangle + \langle Y, D_X(Z) \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle .$$

For (3), we know that

$$[X, Y] f = X(Yf) - Y(Xf)$$

Let us extend X, Y to vector fields $\widetilde{X}, \widetilde{Y}$. Then

$$D_X Y - D_Y X = D_{\widetilde{X}} \widetilde{Y} - D_{\widetilde{Y}} \widetilde{X} = [\widetilde{X}, \widetilde{Y}] = [X, Y].$$

□

This gives some corollaries.

1. The value of $(D_X Y)_p$ depends only on X_p, Y_p and the first derivaties of Y. Therefore the same is true for $\nabla_X Y$ and $\mathbb{I}(X,Y)$.
2. Since $\mathbb{I}$ is symmetric, the value of $\mathbb{I}(X,Y)_p$ depends only on X_p and Y_p. And particular, $\mathbb{I}$ is a bilinear form on $T_p\Sigma$.

Definition 10.10. A **linear connection** is an $\mathbb{R}$-bilinear map

$$\nabla : M \times M \to M$$

such that for any two vector fields and a smooth function $f \in C^\infty(M)$ it holds that

1. $\nabla_{fX} Y = f\nabla_X Y$ and
2. $\nabla_X fY = X(f)Y + f\nabla_X Y$.

Definition 10.11. A linear connection is **consistent with metric** if

$$X \langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle .$$

Definition 10.12. A linear connection ∇ is **torsion-free** if

$$\nabla_X Y - \nabla_Y X = [X,Y].$$

Theorem 10.13 (Fundamental Theorem of Differential Geometry)**.** *Let M be a Riemannian manifold. Then there exists exactly one linear connection that is torsionfree and consistent with metric. It is called the* ***Levi-Civita connection****.*

Definition 10.14. Let M be smooth manifold and TM be its tangent bundle. A **Riemann metric** on M is a smooth, bilinear, symmetric, positive definite form

$$g_p = \langle -, -\rangle_p : T_pM \times T_pM \to \mathbb{R}.$$

Smoothness means that for a chart φ, we have a basis of TM around p called ∂_i and for that basis all the functions

$$g_{ij}(p) = g_p(\partial_i(p), \partial_j(p)).$$

Example. (Pullback of Riemann metric). Let $\Sigma : U \to \mathbb{R}^3$. Then for $p \in U$ we can take

$$g_p : T_pU \times T_pU \to \mathbb{R}$$

by putting down

$$g_p(v,w) = \langle D\Sigma(v), D\Sigma(w)\rangle$$

where $\langle -, -\rangle$ is the $\mathbb{R}^3$ dot product.

Example (hyperbolic plane). Take $\mathbb{H}^2 = \mathbb{R} \times \mathbb{R}_+$. We can give a Riemann Metric by

$$\langle (v_1, v_2), (w_1, w_2) \rangle = \frac{v_1 w_1 + v_2 w_2}{y^2}.$$

Why is this called a *Riemann Metric*? Well, since we can measure tangent vectors, we can measure the *length* of curves!

$$d(\gamma) = \int_0^1 \langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle_{\gamma(t)} \, \mathrm{d}t.$$

Now, we can measure the distance by

$$d_g(x, y) = \inf d(x).$$

Proof of the Fundamental Theorem of Differential Geometry. Existence work by partitions of unity. For uniqueness, we will use

$$\begin{aligned} X \langle Y, Z \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \\ Y \langle Z, X \rangle &= \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_Y X \rangle \\ -Z \langle X, Y \rangle &= - \langle \nabla_Z X, Y \rangle - \langle X, \nabla_Z Y \rangle . \end{aligned}$$

□

§10.4 The shape operator

In th previous lecture we saw how to find the principal curvatures in terms of a Taylor approximations. Now we shall try and see how to find these invariants in a more general way.

Remark. Take a unit tangent vector to a surface Σ at p, called u_p. Now take a curve γ through p such that $\dot{\gamma} = u_p$ and γ lies in the plane generated by u_p and the normal vector n_p. We can compute

$$D_{\dot{\gamma}} \dot{\gamma} = \kappa_\gamma(p) n(p),$$

so we have

$$\mathrm{I\!I}(u_p, u_p) = \kappa_u.$$

Definition 10.15 (Shape operator)**.** The shape operator

$$S : T_p\Sigma \to T_p\Sigma$$

is defined by

$$S v_p = -D_{v_p} n.$$

Lemma 10.16. *For vectors v_p, w_p we have*

$$\langle S v_p, w_p \rangle = \mathrm{I\!I}(v_p, w_p).$$

Proof. Take vector fields V, W which at p are equal to v_p, w_p. Then

$$0 = v_p \langle W, n \rangle = \langle D_V W, n \rangle + \langle Y, D_X n \rangle = \mathrm{I\!I}(v_p, w_p) - \langle w_p, S v_p \rangle .$$

We didn't actually need v_p to be a whole field. □

Definition 10.17 (Gaussian Curvature)**.** The Gaussian curvature at p is defined as

$$\kappa_p = \kappa_1(p) \kappa_2(p).$$

Remark.

$$\det S_p = \kappa_p.$$

For an orthonormal (in the sense of the $\mathbb{R}^3$ dot product) basis v, w of $T_p\Sigma$ we have

$$\det S_p = \det \begin{pmatrix} \mathrm{I\!I}(v,v) & \mathrm{I\!I}(v,w) \\ \mathrm{I\!I}(w,w) & \mathrm{I\!I}(w,w) \end{pmatrix}.$$

Theorem 10.18 (Theorema egregium)**.** *The Gaussian curvature κ_p can be calculated from the Levi-Civita connection.*

Corollary. Take an open Riemannian subset (U, g) of $\mathbb{R}^2$ and two embeddings

$$\Sigma_1, \Sigma_2 : U \to \mathbb{R}^3.$$

Then the curvature κ_p does not depend on the embedding!

Proof. We want to express the determinant of $\mathrm{I\!I}$ using only the connection. By the definition of the Lie bracket,

$$\begin{aligned} 0 &= D_X D_Y Z - D_Y D_X Z - D_{[X,Y]} Z \\ &= D_X(\nabla_X Z + \mathrm{I\!I}(Y,Z)n) - D_Y(\nabla_X Z + \mathrm{I\!I}(X,Z)n) \\ &- \nabla_{[X,Y]} Z - \mathrm{I\!I}\left([X,Y], Z\right) n \\ &= \nabla_X \nabla_y Z + \mathrm{I\!I}(X, \nabla_Y Z)n + (X\mathrm{I\!I}(Y,Z))n - \mathrm{I\!I}(Y,Z)SX \\ &- (\text{ the same for } Y\ ) \\ &- \nabla_{[X,Y]} Z - \mathrm{I\!I}([X,Y], Z)n. \end{aligned}$$

□

Taking the tangent part of this we obtain

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z = \mathrm{I\!I}(Y,Z)SX - \mathrm{I\!I}(X,Z)SY.$$

We denote the value of either side by $R(X,Y)Z$, and call it the **Riemann curvature tensor**. We expand this to

$$\langle R(X,Y)Z, W\rangle = \mathrm{I\!I}...$$

Remarks. This value depends only on X_p, Y_p, Z_p, W_p, and in a linear way! Also R_p depends on g_{ij} and the first and second derivaties of g_{ij}.

§10.5 Tensor bundles

Recall the definition of the **tensor product** (over $\mathbb{R}$). Let V, W be linear spaces. We first use the *multilinear space interpretation.*

$$V \otimes W = \mathrm{Lin}\left\{v \otimes w : v \in V, w \in W\right\}/N,$$

where the subspace N is the linear closure Lin given by the relations

$$\begin{aligned} (\alpha v_1 + \beta v_2 \otimes w) &= \alpha(v_1 \otimes w) + \beta(v_2 \otimes w) \\ (\alpha v_1 + \beta v_2 \otimes w) &= \alpha(v_1 \otimes w) + \beta(v_2 \otimes w) \end{aligned}$$

We know that

$$\dim V \otimes W = \dim V \dim W.$$

In particular, if we have bases e_i, f_j of V and W, the tensors $e_i \otimes f_j$. There is a universal propety, and also

$$Mult(V, W; \mathbb{R}) \cong (V \otimes W)^* \cong V^\star \otimes W^\star.$$

If we have a pure tensor $t = \alpha \otimes \beta$ for $\alpha \in V^\star$, $\beta \in W^*$, wehave a binilear functional

$$(\alpha \otimes \beta)(v, w) = \alpha(v)\beta(w).$$

Excercise.

$$Mult(V_1, V_2, \ldots, V_n; V_{n+1}) \cong Mult(V_1, V_2, \ldots, V_n, V_{n+1}^*; \mathbb{R}) \cong \ldots \cong V_1^* \otimes V_2^* \otimes \ldots \otimes V_{n+1},$$

where all the mappings are natural.

What does it mean that the maps are natural? If we have linear operators $A_i : V_i \to W_i$, then we have a tensor product of these

$$A_1 \otimes A_2 \otimes \ldots A_n : \otimes V_i \to \otimes_{i=1}^n W_i$$

defined by

$$\left(\otimes_{i=1}^n A_i\right)(v_1 \otimes \ldots \otimes v_n) = A_i(v_i) \otimes \ldots \otimes A_n(v_n).$$

In patricular there is a diagram, between $W^* \otimes W^*$.

Definition 10.19 (Linear bundle)**.** A k-dimensional **linear bundle** on a smooth manifold M is a map

$$\pi : E \to M,$$

where E is a manifold, with the properties

1. π is onto,
2. for all $p \in M$, the fiber $E_p := \pi^{-1}[p]$ is a k-dimensional linear space,
3. for all $p \in M$ there is an open $U \ni p$ and a **local trivialization**

$$\varphi : \pi^{-1}[U] \to U \times \mathbb{R}^k$$

such that

$$\varphi_q : E_q \to \{q\} \times \mathbb{R}^k$$

is a linear isomorphism (φ_q is the restriction of φ to E_q).

Examples.

1. The trivial bundle $\varepsilon_M^k := M \times \mathbb{R}^k$.
2. The tangent bundle TM.
3. Let $M \subseteq \mathbb{R}^k$. Then we have

$$T\mathbb{R}^k|_M = TM \times (TM)^\perp.$$

4. Let $M \subseteq N$. There is a quotient bundle

$$TN|_M/TM$$

given by

$$E_p = T_pN/T_pM.$$

This is called the **normal bundle**.

5. For linear bundles E, F there is also a fiber product bundle $E \oplus F$.

6. There is also a tensor a tensor product bundle $E \otimes F$ for any linear bundles E, F.

7. There is also a dual bundle E^* for any linear bundle E.

8. Tautological bundle.

Chapter 11

The Gauss-Bonnet Theorem

Lemma 11.1 (Cartan's Formula)**.** *For any $\omega \in \Omega^1 M$ and vector fields v, w*

$$\mathrm{d}\,\omega(v,w) = v\omega(w) - w\omega(v) - \omega([v,w]).$$

Triangulability. Not all topological manifolds are triangulable. Turns out the for 4-manifolds,

Corollaries On S^2 there are no nonvanishing vector fields. (hairy ball)

Corollary. For any Riemman metric g on the torus T^2

$$\int\limits_{T^2} \kappa \,\mathrm{d}S = 0.$$

(1). Ad absurdum, take a nonvanishing vector field $X \in \Xi(S^2)$, take

$$e_1 = \frac{X}{|X|}$$

and pick e_2 so that $e_1 \perp e_2$ and (e_1, e_2) is positively oriented. Now we have

$$0 < \int\limits_{S} \kappa \,\mathrm{d}S = \int\limits_{S^2} \mathrm{d}\,\omega \,\mathrm{d}S = \int\limits_{\partial S^2} \omega = 0,$$

since $\kappa = 1$. □

(2). On the 2-torus there is a global reper (from the product representation). □

Flat metric on the torus. You can get a flat torus metric from the fundamental polygon (or product).

Proof (Gauss-Bonnet). Let us take an area F bounded by γ and a reper e_1, e_2. Then we have

$$\int\limits_{F} \kappa \mathrm{d}\,S = \int\limits_{F} \mathrm{d}\,\omega = \int\limits_{\partial F} \omega = \int\limits_{0}^{1} \omega(\gamma'(t))\,\mathrm{d}t =$$

$$= \int_0^1 -\left\langle \nabla_{\gamma'(t)} e_1, e_2 \right\rangle \mathrm{d}t = \int_0^1 -\langle \nabla_t \widetilde{e_1}, \widetilde{e_2} \rangle \ \mathrm{d}t,$$

where

$$\widetilde{e_i} = e_i \circ \gamma.$$

□

Lemma 11.2 (Integrating repers)**.** *Let γ be a curve in Σ, e_i and f_i be two orthonormal repers, and f_i be parallel to the curve. Then the value of*

$$\int_0^1 \langle \nabla_t e_1, e_2 \rangle \ \mathrm{d}t$$

is equal to the change of angle between f_1 and e_1. This is a way of measuring how much e_1 rotates around γ.

Proof. We have

$$e_1 = \cos\alpha(t) f_1 + \sin\alpha(t) f_2$$

and

$$e_2 = -\sin\alpha(t) f_1 + \cos\alpha(t) f_2.$$

Then (by properties of connection and the connection being parallel)

$$\langle \nabla_t e_1, e_2 \rangle = \langle \alpha'(t)(-\sin\alpha(t)) f_1 + \cos\alpha(t) f_2, e_2 \rangle = \alpha'(t) \langle e_2, e_2 \rangle = \alpha'(t).$$

So we have

$$\int_0^1 \langle \nabla_t e_1, e_2 \rangle \ \mathrm{d}t = \alpha(1) - \alpha(0).$$

□

Let us take a triangle $\Delta \subseteq \Sigma$ bounded by

$$\gamma : [0, 3] \to U \subseteq \Sigma.$$

Consider three repers:

1. (e_1, e_2) with $e_1 = \lambda \partial x_1$
2. (f_1, f_2) parallel to γ
3. (τ, n) – the Frenet frame of γ.

Because γ need not be smooth at the vertices, the Frenet reper will have jumps.

Chapter 12

Differential Geometry, Problemset 1

Lemma 12.1. *The second derivative of*

$$g \circ f$$

is given by

$$(g'' \circ f) \cdot (f')^2 + (g' \circ f) \cdot f''$$

Proof. By the Chain Rule,

$$(g \circ f)' = f' \cdot (g' \circ f).$$

To take the derivative of this, we use the product rule and chain rule. □

Problem 1

Solved in the lecture.

Problem 2

Solved in the lecture.

Problem 3

By 12.1, if the hypothesis holds for a curve γ, it also holds for all reparametrizations. Therefore, we may assume that γ is parametrized by arclength. Denote the center of the circle as c. We now have

$$\begin{aligned} f(\gamma) &= \langle \gamma - c, \gamma - c \rangle - R^2 \\ \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma) &= 2\langle \dot{\gamma}, \gamma - c \rangle \\ \frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\gamma) &= 2\langle \ddot{\gamma}, \gamma - c \rangle + 2\langle \dot{\gamma}, \dot{\gamma} \rangle \end{aligned}$$

Now, $f(\gamma) = 0$ iff $\gamma(t) \in C(x, R)$. Then, the first derivative is 0 iff $\dot{\gamma}$ is perpendicular to the radius of the circle. Note that the tangent to the circle has precisely this property, so this gives us equivalence of γ being tangent to the circle and the derivative of $f(\gamma)$ being $= 0$.

For the third part, we assume arclength parametrization. Then the equation becomes

$$0 = 1 + \langle \ddot{\gamma}, \gamma - c \rangle,$$

which means that the length of $\ddot{\gamma}$ must be $\frac{1}{R}$ and its direction must be opposite to $\gamma - c$, i.e. it must be directed toward the centre of the circle. This is true for the second derivative of the arclength parametrised circle, so the second derivative of $f(\gamma)$ is zero iff the second derivatives of γ and the arclength parametrization of the circle agree.

Problem 4

An equivalent definition:

Definition 12.2. For an arclength-parametrized curve γ, the circle of best fit to γ at s_0 is the unique tangent circle with the same signed curvature at the point of contact.

Since any tangent circle has the same tangent, is also has the same positive normal ($N(s)$), and the second derivative is the curvature times the normal by definition of curvature **??**.

Problem 6

Assume the curve is parametrized by arclength. We will compute how the distance from the center of the circle changes along the curve. Let the circle have curvature κ. Then the radius of the circle is $1/\kappa$ and its center is the point

$$\gamma(0) + \frac{1}{\kappa} N(0).$$

Then, the vector from the centre to a point on the curve is

$$r(s) := \gamma(s) - \gamma(0) - \frac{1}{\kappa} N(0).$$

This is clearly not changed by translating the whole configuration, so without loss of generality $\gamma(0) = 0$. In what follows, we use the dot product differentiation formula **??**.

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s} ||r(s)||^2 &= \frac{\mathrm{d}}{\mathrm{d}s} \langle r(s), r(s) \rangle \\ &= 2 \langle \dot{r}(s), r(s) \rangle \\ &= 2 \langle \gamma(s), \dot{\gamma}(s)) \rangle - 2 \left\langle \frac{1}{\kappa} N(0), \dot{\gamma}(s) \right\rangle . \end{aligned}$$

At $s = 0$ both terms come out to 0, so we need to compute another derivative to see what is going on. We have

$$\begin{aligned} \frac{\mathrm{d}^2}{\mathrm{d}s^2} ||r(s)||^2 &= 4 \langle \dot{\gamma}(s), \dot{\gamma}(s) \rangle + 4 \langle \gamma(s), \ddot{\gamma}(s) \rangle - 4 \left\langle \frac{1}{\kappa} N(0), \ddot{\gamma}(s) \right\rangle \\ &= 4 + 4 \langle \gamma(s), \ddot{\gamma}(s) \rangle - 4 \frac{\kappa_\gamma(s)}{\kappa} \langle N(0), N(s) \rangle , \end{aligned}$$

where we have used **??** and **??**. At $s = 0$ this comes out to

$$4 \left(1 - \frac{\kappa_\gamma(0)}{\kappa} \right) ,$$

which is positive for $\kappa_\gamma(0) < \kappa$ and negative for $\kappa_\gamma(0) > \kappa$. This concludes the problem: for example, if the curve γ stays inside the circle (even locally!), then we cannot have $\kappa_\gamma(0) < \kappa$, because then the distance would be increasing (by Taylor **??**).

Remark. This still works for negative κ. The only thing we need to check to make sure of that is that the center of the circle is where it is. It gets more tricky for $\kappa = 0$ – in that case the appropriate reformulation of *inside* and *outside* is on one or the other side of the line, and instead of the distance we should consider

$$\langle \gamma(s) - \gamma(0), N(0) \rangle.$$

The derivative of this is

$$2\langle \dot\gamma(s), N(0) \rangle,$$

which equals 0 at $s = 0$. The second derivative is

$$4\langle \ddot\gamma(s), N(0) \rangle = 4\kappa_\gamma(s)\langle N(s), N(0) \rangle$$

by Frenet **??**. At s this is just $4\kappa_\gamma(0)$, and analysing signs as above finishes the problem.

Problem 5

Recall from Problem 6 that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}||r(s)||^2 = 4 + 4\langle \gamma(s), \ddot\gamma(s) \rangle - 4\frac{\kappa_\gamma(s)}{\kappa}\langle N(0), N(s) \rangle$$

and that the first derivative disappears at $s = 0$ for a tangent circle. For a circle of best fit, the curvature $\kappa = \kappa_\gamma(0)$ (see my solution of Problem 4), so the second derivative disappears as well. Computing the derivative of this

$$\frac{\mathrm{d}^3}{\mathrm{d}s^3}||r(s)||^2 = 4\langle \gamma(s), \dddot\gamma(s) \rangle + 4\langle \dot\gamma(s), \ddot\gamma(s) \rangle - 4\frac{\kappa'_\gamma(s)}{\kappa}\langle N(0), N(s) \rangle - 4\frac{\kappa_\gamma(s)}{\kappa}\langle N(0), \dot N(s) \rangle$$

For $s = 0$ this gives

$$\frac{\mathrm{d}^3}{\mathrm{d}s^3}||r(s)||^2 = -4\frac{\kappa'_\gamma(s)}{\kappa}.$$

So, if the derivative is nonzero

Problem 7

I have a shitty calculatory solution, but nothing nice geometrically. It goes like

1. Parametrize the points on the parabola by (p, p^2).
2. Centers of tangent circles lie on $p - 2pt, p^2 + pt$ and have radii $t\sqrt{1 + 4p^2}$.
3. Together with $y = x^2$ this gives an equation for the intersection points. Divided by $(x - p)^2$ the equation is
$$1 + (x + p)^2 + 2t = 0.$$
4. To avoid having two solutions, pick t so that $x = p$ is a solution of this. Then, $x = -3p$ is another solution (same value of the square).

5. Caclulate that in this configuration

$$t = -\frac{1+4p^2}{2}.$$

6. The tangent to the parabola at the chosen point is $(1, -6p)$. The point is $(-3p, 9p^2)$ and the center of the circle is $(2p + 4p^3, -p^2 - 1/2)$. The vector from the circle to the point is then $(5p + 4p^3, -10p^2 - 1/2)$.

7. The dot product between the tangent and the vector from the point to the center is a nonzero polynomial, so for almost all p the circle is not tangent.

Problem 8

Geometrically, the curve rolls into itself, like a spiral. Note that the curvature can also be negative or zero at some points, but increasing. In what follows, we consider a point of self-intersection.

12.0.1 Non-zero curvature

By the **??**, we know that if the curvature at some point is nonzero, then points in the direction of the absolute value of curvature are properly inside the osculating circle. Therefore, there can be no self-intersection at a point of nonzero curvature.

12.0.2 Zero curvature

For a point of zero curvature, we know from the remark to problem 6 that the first derivative of the signed distance from the line (which is the osculating circle at a point of curvature 0) is 0 and the second derivative is strictly positive after the point of zero curvature and strictly negative before (in a small neighbourhood!).

Taking points arbitrarily close to zero, we see that

TO BE CONTINUED!

Problem 9

Chapter 13

Measure Theory Bank of Lemmas

Definition 13.1 (Outer measure)**.** A nonnegative set function $\mu^* : \mathcal{P}(X) \to \mathbb{R}$ is called an **outer measure** when it satisfies the following three properties:

1. it is null on the empty set: $\mu^*(\varnothing) = 0$,
2. it is monotone: $A \subseteq B \implies \mu^*(A) \leqslant \mu^*(B)$,
3. is countably subadditive:
$$\mu^* \left(\bigcup_{j=1}^{\infty} A_j \right) \leqslant \sum_{j=1}^{\infty} \mu^*(A_j)$$

Lemma 13.2 (Generating an outer measure)**.** *Let μ be a measure on the measurable space (X, Σ) and μ^* be the function*
$$\mu^* : \mathcal{P}(X) \to \mathbb{R}$$
defined by
$$\mu^*(A) := \inf \{\mu(B) : B \in \Sigma, A \subseteq B\}.$$
Then μ^* is the unique outer measure which extends μ. Moreover, the infimum may be taken over any σ-ring of sets (AM I SURE OF THIS?) that generates Σ.

Proof. TODO. □

Lemma 13.3 (Generating a σ-algebra)**.** *Fix a space X. For any family of sets $\mathcal{A}$, $\sigma(\mathcal{A})$ can be generated by any of the following sets of operations:*

1. *the empty set, complements and countable unions.*
2. *the empty set, complements, finite unions and increasing countable unions.*

Lemma 13.4. *Let $\mathcal{B}$ be a basis for the topology of X. Then*
$$\sigma(\mathcal{B}) = \operatorname{Bor} X.$$

Lemma 13.5. *The set operations and the measure taking operations are continuous with respect to the symmetric difference pseudometric.*

Lemma 13.6. *A bounded measurable function f on a measure space (X, μ) can be uniformly approximated by simple functions.*

Proof. Let
$$|f| \leqslant [-M, M].$$
We will construct the approximation by considering *bins* of the values of f, i.e. the sets
$$A_k := f^{-1}\Big[[k\varepsilon, (k+1)\varepsilon)\Big]$$
for $k \in \mathbb{Z}$. In such a *bin*, all the values are within an ε of each other. Since f is bounded, all but finitely many of the bins are empty X, so the function
$$\widetilde{f}_\varepsilon := \sum_{A_k \neq \varnothing} (k\varepsilon) \cdot \chi_{A_k}$$
□

Remark. This works equally well for almost everywhere bounded functions, giving almost everywhere uniform convergence.

Lemma 13.7. *Let A, A_1, ..., A_k be measurable sets such that*
$$\forall k.\, \mu(A_i \cap A) \geq (1 - \delta_i)\mu(A).$$
Then
$$\mu(A \cap A_1 \cap A_2 \cap \ldots \cap A_k) \geq \left(1 - \sum \delta_i\right) \mu(A).$$

Proof. Union bound on the sets
$$A \cap A_i^c.$$
□

Remark. This also works for an infinite sequence of sets A_k; we obtain.
$$\mu\left(A \cap \bigcap_k A_k\right) \geqslant \left(1 - \sum_k \delta_k\right) \mu(A).$$

Chapter 14

An introduction to geometric measure theory

In this chapter, we study the links between the topology and geometry of $\mathbb{R}$ and the Lebesgue measure. We first give two examples of how the two structures agree, and one example of how they don't.

Isometries. Consider the group $\operatorname{Isom}\mathbb{R}$ of the isometries of $\mathbb{R}$ with the euclidean metric. One easily shows that this group consists of functions of the form

$$x + a \text{ or } a - x$$

for $a \in \mathbb{R}$. The Lebesgue measure is invariant on transformations $g \in \operatorname{Isom}\mathbb{R}$, i.e.

$$\lambda(gA) = \lambda(A)$$

for all measurable $A \subseteq \mathbb{R}$. A corollary of this is that the Lebesgue measure is invariant w.r.t. the addition operation on $\mathbb{R}$, which gives the reals the structure of a topological group.

Affine transformations. Similarly to the above, the Lebesque measure work well with the action of the affine transformation group $\operatorname{Aff}\mathbb{R}$. Directly from the definition, the group of affine transformations consists of the functions

$$g_{a,b}(x) := ax + b$$

for $r \neq 0$, and the interaction with measure is given by

$$\lambda(g_{a,b}A) = |a| \cdot \lambda(A).$$

Topology. There is a disconnect between the topological (nonempty interior) and measure-theoretic (positive measure) notions of *large* or *non-negligable* – the topological notion is strictly stronger! Indeed, a set with nonempty interior has positive measure, but if we enumerate the rationals as

$$\mathbb{Q} = \{q_1, q_2, \ldots\}$$

the set

$$\mathbb{R} \setminus \bigcup_{n=1}^{\infty} (q_n - \frac{\varepsilon}{2^{n+1}}, q_n + \frac{\varepsilon}{2^{n+1}})$$

has comeasure ε, but is nowhere dense.

However, there does exist a link between the two notions. It is a bit more subtle.

Definition 14.1. Fix a measurable set $A \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is called a **density point** iff

$$\lim_{\delta \to 0^+} \frac{\lambda(A \cap B(x, \delta))}{2\delta} = 1.$$

The 2δ in the numerator is of course $\lambda(B(x, \delta))$.

Definition 14.2. The set of density points of A will be denoted $\phi(A)$.

Note that a density point is by neccesity an accumulation point. The promised link between geometry, measure and topology is provided by the theorem below.

Theorem 14.3 (Lebesgue Density Theorem)**.** *Let $A \subseteq \mathbb{R}$ be a measurable set. Then almost all points of A are density points of A in the sense that*

Ankified

$$\lambda^*(A \setminus \phi(A)) = 0.$$

Remark. Note that the theorem follows trivially for null sets. Also, for a given A, we may as well apply the theorem to A^c to get that almost all points outside of A have density 0.

For the proof of the **Lebesgue Density Theorem**, we will need a tool, which we introduce now and prove later.

Definition 14.4. A family $\mathcal{J}$ of nontrivial closed intervals is called a **Vitali cover** of a set A (not necessarily measurable) if for any given $\varepsilon > 0$ and $x \in A$ there is an interval $J \in \mathcal{J}$ such that

$$\operatorname{diam} J < \varepsilon \wedge x \in J.$$

In particular

$$A \subseteq \bigcup \mathcal{J}.$$

Theorem 14.5 (Vitali Covering Theorem)**.** *If $\mathcal{J}$ is a Vitali cover of A, there exists a (possibly finite) sequence of pariwaise disjoint segments $J_n \in \mathcal{J}$ such that*

Ankified

$$\lambda^* \left(A \setminus \bigcup_n J_n \right) = 0.$$

Why is this theorem useful? Vitali's theorem may not sound very smart on first glance. Its strength lies in the *disjointness* of the cover. If we go about choosing the cover J_n without any guarantees, we can for example choose

$$\bigcup_{q_n \in \mathbb{Q}} \left(q_n - \frac{\varepsilon}{2^{n+1}}, q_n + \frac{\varepsilon}{2^{n+1}} \right)$$

and get stuck! We have only covered a subset of the reals of size ε, but we cannot use any other segment by density of $\mathbb{Q}$.

Corllary. If a set A has a Vitali cover consisting of its subsets (up to measure zero), then A is measurable.

Proof of the Lebesgue Density Theorem. We represent

$$A \setminus \phi(A) = \bigcup_k A_k$$

for

$$A_k = \left\{ a \in A : \liminf_{\delta \to 0^+} \frac{\lambda(A \cap B(a,\delta))}{2\delta} < 1 - \frac{1}{k} \right\}.$$

It suffices to show

$$\lambda^*(A_k) = 0$$

for all k to finish the proof. Since we may represent A as

$$A = \bigcup_{z \in \mathbb{Z}} A \cap [z-1, z+1]$$

and being a density point of A is the same as being a density point of one of the *cutouts* in the sum above, we may assume without loss of generality that $A \subseteq [0,1]$.

By definition of outer measure, we can approximate A_k from above by an open set U such that

$$\lambda^*(A_k) \leqslant \lambda(U) \leqslant \lambda^*(A_k) + \varepsilon.$$

Construct a covering

$$\mathcal{J} = \left\{ [a,b] : [a,b] \subseteq U,\ \lambda\Big(A \cap [a,b]\Big) \leqslant \left(1 - \frac{1}{k}\right) \lambda[a,b] \right\}.$$

It is a Vitali cover of A_k. By Vitali's Theorem we can pick a pairwise disjoint sequence of intervals $J_i \in \mathcal{J}$ for which

$$\lambda^* \left(A_k \setminus \bigcup_i J_i \right) = 0.$$

This gives

$$\begin{aligned}
\lambda^*(A_k) &= \lambda^* \left(A_k \cap \bigcup_i J_i \right) \\
&\leqslant \sum_i \lambda^*(A_k \cap J_i) \\
&\leqslant \sum_i \lambda^*(A \cap J_i) \\
&\leqslant \left(1 - \frac{1}{k}\right) \sum_i \lambda(J_i) \\
&\leqslant \left(1 - \frac{1}{k}\right) \lambda(U) \\
&\leqslant \left(1 - \frac{1}{k}\right) \left(\lambda^*(A_k) + \varepsilon\right).
\end{aligned}$$

The passage from line 2 to 3 may seem trivial, but is in fact crucial. This is the place where we use $A_k \subseteq A$! Otherwise the theorem is quite absurd, even for simple examples like $[0, 1]$. Since $\lambda^*(A_k) \leqslant \lambda(A) < \infty$, we can rearrange this to obtain

$$\lambda^*(A_k) \leqslant (k-1)\varepsilon.$$

Since ε can be picked arbitrarily close to 0, we get

$$\lambda^*(A_k) = 0.$$

□

Proof of the Vitali Covering Theorem. The key to avoiding the *trap* we wrote about after stating the VCT is to choose the segments to be as large as possible – or at least not embarassingly small.

Without loss of generality, A is bounded since we can sum the coverings of $A \cap (n, n+1)$. The sequence of segments we pick is denoted J_n. In that case we may also assume $\bigcup \mathcal{J}$ is bounded. Its prefixes are

$$\begin{aligned} P_n &:= \bigcup_{i<n} J_i \\ \mathcal{J}_n &:= \{J \in \mathcal{J} : J \cap P_n = \varnothing\} \end{aligned}$$

and the *width* of what we can choose is

$$\gamma_n := \sup_{J \in \mathcal{J}_n} \operatorname{diam} J.$$

Note that in particular

$$\begin{aligned} P_1 &= \varnothing, \\ \mathcal{J}_1 &= \mathcal{J} \\ \gamma_1 &\leqslant \operatorname{diam} A < \infty. \end{aligned}$$

At each step, we choose J_n so that

$$\operatorname{diam} J_n \geqslant \frac{1}{2}\gamma_n,$$

or we stop if $\gamma_n = 0$ at some point. If indeed $\gamma_n = 0$, then by the definition of a Vitali cover, each $a \in A \setminus \mathcal{J}_{n-1}$ is an accumulation point of $\mathcal{J}_{n-1}$, but as a finite sum of closed intervals $\mathcal{J}_{n-1}$ is actually closed, so $a \in \mathcal{J}_{n-1}$.

Claim 1. The sequence γ_n converges monotonically to 0.

Proof of Claim 1. Being the supremum of ever decreasing sets, γ_n is decreasing. It is also positive, so the sequence converges and $\lim_n \gamma_n \geqslant 0$. Suppose that $\lim_n \gamma_n = c > 0$. Then in the construction, we would almost always choose disjoint intervals of diameter at least $c/2$. This is impossible, since $\bigcup \mathcal{J}$ was assumed to be bounded, so it has finite measure! □

If finitely many steps of the choice procedure are enough, there is nothing left to The key to proving that the choice procedure is correct will be the **blowup**, which we define for $J = [x-r, x+r]$ as

$$\widetilde{J} := [x - 5r,\, x + 5r]$$

Claim 2. At all steps of the construction

$$A \subseteq \mathcal{J}_n \cup \bigcup_{i \geqslant n} \widetilde{J}_i.$$

Proof of Claim 2. The set $\mathcal{J}_n$ is closed as a union of closed intervals. Therefore, if $a \in A \setminus \mathcal{J}_n$, there is a nondegenerate interval $I \ni a$. Since $\gamma_n \to 0$ by **Claim 1**, I is not considered in the construction of the sequence J_n for almost all n. Let n_0 be the last step where it is considered. Then we must have $I \cap J_{n_0+1} \neq 0$, because that is the step at which I is no longer considered.

We will show that this implies $a \in \widetilde{J}_{n_0+1}$. Let $J_{n_0+1} = [x - r,\, x + r]$ and $y \in I \cap J_{n_0+1}$. Then we have

$$\begin{aligned} d(a, x) &\leqslant d(a, y) + d(y, z) \\ &\leqslant \operatorname{diam} I + r \\ &\leqslant (2 \cdot \operatorname{diam} J_{n_0+1}) + r \\ &= 2 \cdot 2r + r \\ &= 5r, \end{aligned}$$

where the diameter bound comes from the definition of γ_{n_0} and the fact that I is still available at step n_0 of the construction. □

To finish the proof of Vitali's Covering Theorem, we compute that for all n

$$\begin{aligned} \lambda^*\left(A \setminus P_n\right) &\leqslant \lambda^*\left(A \cap \bigcup_{i \geqslant n} \widetilde{J}_i\right) \\ &\leqslant \lambda\left(\bigcup_{i \geqslant n} \widetilde{J}_i\right) \\ &\leqslant \sum_{i \geqslant n} \lambda(\widetilde{J}_i) \\ &= 5 \sum_{i \geqslant n} \lambda(J_i). \end{aligned}$$

These are the tails of the convergent series

$$\sum_{i=1}^{\infty} \lambda(J_i) = \lambda\left(\bigcup_i J_i\right) < \infty,$$

so we get

$$\lambda^*\left(A \setminus \bigcup_i J_i\right) \leqslant \lambda^*\left(A \setminus P_n\right) \to 0.$$

□

Remark. Retracing the argument behind **Claim 2.**, we might prove that for any $\alpha < 1$, if we define γ_n with a coefficient of α instead of $\frac{1}{2}$, the constant used for blowing up intervals can be brought down to

$$1 + \frac{2}{\alpha}.$$

In particular, we can get arbitrarily close to 3.

§14.1 Corollaries and the Lebesgue Differentiation Theorem

Theorem 14.6 (Lebesgue Differentiation Theorem)**.** *Let $f \in L^1(\mathbb{R})$. Then, for almost all x,*

$$\lim_{\delta \to 0^+} \frac{1}{2\delta} \int_{x-\delta}^{x+\delta} f(s)\,\mathrm{d}\lambda(s) = f(x).$$

Proof. For characteristic functions, this is just a restatement of the Lebesgue Density Theorem. □

§14.2 Generalization to metric spaces

The argument in the proof of the VCT was written so that it is easliy generalizable to any metric space with a measure on its Borel sets.

To be more precise, what we need to lift the argument is that

$$\mu\left(B(x, 5r)\right) \leqslant C\mu\left(B(x, r)\right)$$

for some constant C. We can also substitute any constant larger that 3 instead of 5.

Chapter 15

One (Cantor) set to rule them all

§15.1 Ternary Cantor

Let us begin by making a construction. Take the closed interval $C_0 := [0, 1]$ and remove the middle one third of it in such a way that the remaining two interval are closed. The result of this is

$$C_1 := \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

Now, repeat the operation of cutting out the middle third and call the result C_2. We can repeat this *ad inifinitum* and obtain a decresing sequence of sets

$$[0, 1] = C_0 \supset C_1 \supset C_2 \supset \dots.$$

Perhaps surprisingly, there are numbers which are not removed at any step, i.e. the intersection

$$\mathcal{C}_3 := \bigcap_{k=0}^{\infty} C_k$$

is nonempty! It contains 0 and 1. In fact, any number which can be written in base 3 using only 0's and 2's is an element of this intersection. These are in fact all such numbers. We introduce a tool to prove that.

Lemma 15.1. *Let $b \geqslant 2$ be a positional system base and x_0 be a number with k digits after the positional point. Then, the numbers formed by adjoining (perhaps infinitely many) digits to the base b representation of x_0 are all the numbers in the interval*

$$\left[x_0,\, x_0 + b^{-k}\right].$$

If we disallow the infinite extension by the digit $(b-1)$, we get the interval

$$\left[x_0,\, x_0 + b^{-k}\right).$$

Finally, if we allow only finite extensions, we get the b-ary numbers in the second interval.

Lemma 15.2. *A real number $x \in [0, 1]$ is an element of $\mathcal{C}_3$ iff x can be written in base using only the digits 0 and 2.*

Proof (if). We proceed by induction with the induction thesis: *x belongs to $\mathcal{C}_3$ iff x can be written in base 3 so that its first k digits are 0 or 2.* It should be clear that this thesis is equivalent to the lemma statement. For each k, the statement is true by the previous lemma. □

§15.2 Abstract Cantor

The ternary Cantor set has many interesting properties. However, to study it, we will move to a more convenient representation. We can think of the Cantor set as the set of leaves of an infinite binary tree – starting from the root, at each level we choose whether to go right or left, or whether to insert 0 or 2 as the next digit in the base-3 representation of an $x \in \mathcal{C}_3$.

In this way, we can represent the ternary Cantor as

$$\mathcal{C} := \{0,1\}^{\mathbb{N}}.$$

That the map we just described is a bijection follows from

Lemma 15.3. *Let $d_k, \widetilde{d_k}$ be two sequences of base b digits. Then the corresponding real numbers are equal iff d_k and $\widetilde{d}$ agree on some prefix and afterwards one of them is 0 and the other is $(b-1)$.*

Proof. The condition implies equality of numbers by the sum of a geometric series. The other direction follows by looking at the first moment the expansions differ at then bounding the series sum. □

What is missing from this description is the topology. We topologize $\mathcal{C}$ by the metric

$$d(x,y) = \begin{cases} 0 & \text{for } x = y \\ \frac{1}{n} & \text{for } x \neq y, \end{cases}$$

which can also be written succintly as

$$d(x,y) = \frac{1}{n_0(x,y)}$$

with the notation

$$n_0(x,y) := \inf\{n : x_n \neq y_n\}$$

for the first index at which x and y differ. The function d may not look like a metric at first sight, but in fact it has an even better property.

Lemma 15.4. *For the metric d described above we have for all $x, y, z \in \mathcal{C}$*

$$d(x,z) \leqslant \max\{d(x,y), d(y,z)\}.$$

*In particular, d is an **ultrametric**.*

Proof. Recall that $n_0(x,z)$ is the first position at which x and z differ. Then any y has to differ with at least one of y and z at n_0, but might even earlier. This gives

$$n_0(x,z) \geqslant \min\bigl(n_0(x,y), n_0(y,z)\bigr).$$

Since the function $x \mapsto 1/x$ is decreasing, the thesis follows. □

We have established that $(\mathcal{C}, d)$ is a metric space. It is, in fact, homeomorphic with the subspace topology of $\mathcal{C}_3$ inherited from $[0,1]$.

Lemma 15.5. *The function*

$$h_3 : \mathcal{C} \to \mathcal{C}_3$$

defined by

$$h_3(x) := \sum_{k=1}^{\infty} \frac{2x_k}{3^k}$$

is a homeomorphism.

Proof. Bijectivity follows from the number-system lemma 15.3 and 15.2. For continuity, put down $n_0 := n_0(x, y)$ and compute

$$\begin{aligned}|h_3(x) - h_3(y)| &= \sum_{k=1}^{\infty} \frac{2|x_k - y_k|}{3^k} \\ &\leqslant \sum_{k=n_0}^{\infty} \frac{2}{3^k} \\ &= \frac{2}{3^{n_0}} \cdot \frac{3}{2} \\ &= \frac{1}{3^{n_0-1}}.\end{aligned}$$

The continuity of the inverse follows from the bound

$$|h_3(x) - h_3(y)| \geqslant \frac{2}{3^{n_0}}.$$

□

The function h_3 in 15.5 can be understood as a base 3 expansion operator. When we consider a base 2 expansion instead, we lose bijectivity, but we can cover the whole interval.

Lemma 15.6. *The function*

$$h_2 : \mathcal{C} \to [0, 1]$$

given by

$$h_2(x) := \sum_{k=1}^{\infty} \frac{x_k}{2^k}$$

is a continuous surjection.

Proof. Surjectivity follows from number system properties, and continuity is essentialy the same calculation as in the proof of 15.5. □

Theorem 15.7 (The Universal Property of the Cantor Set)**.** *Every metrizable compact space K is a continuous image of $\mathcal{C}$.*

Proof. Considering an element of $\mathcal{C}$ as a binary expansion, we have by 15.6 a surjection

$$h_2 : \{0, 1\}^{\mathbb{N}} \twoheadrightarrow [0, 1].$$

The space K can be embedded into the Hilbert cube by the **Urysohn Metrization Theorem ??**. By compactness of K, the image of the embedding is a compact and thus a closed subset. We also have a surjection

$$h : \{0, 1\}^{\mathbb{N}} \to [0, 1]^{\mathbb{N}}$$

by using the previous surjection and *unweaving* the Cantor set into the product of countably many Cantor sets, i.e. using

$$\mathbb{N} \cong \mathbb{N} \times \mathbb{N} \implies \mathcal{C} = \{0,1\}^{\mathbb{N}} \cong \{0,1\}^{\mathbb{N}\times\mathbb{N}} \cong \left(\{0,1\}^{\mathbb{N}}\right)^{\mathbb{N}} = \mathcal{C}^{\mathbb{N}}.$$

The last step is using the fact that any closed set of $\mathcal{C}$ is a retract of $\mathcal{C}$, which is **??**. □

A warning against generalization. If K is a compact set, it embeds into a *Tichonov Cube*

$$K \to [0,1]^{\Gamma}$$

and we can surject the Tichonov cube with a generalized Cantor set

$$\{0,1\}^{\Gamma},$$

but the universality theorem fails!

§15.3 Topology of the Cantor set

Definition 15.8 (Cantor Cylinder)**.** Let

$$\varphi : \mathbb{N} \multimap \{0,1\}$$

be a partial function with finite domain. Then we define the **cylinder set** with base φ as

$$[\varphi] := \{x \in \{0,1\} : x|_I = \varphi\}.$$

Lemma 15.9. *The sets* $[\varphi]$ *form a base of the topology of* $\{0,1\}^{\mathbb{N}}$*.*

Definition 15.10. A set $A \subseteq \mathcal{C}$ is **determined** by $I \subseteq \mathbb{N}$, which we donte by $A \sim I$ if for all $x \in A$, $y \in \mathcal{C}$ we have

$$x|_I = y|_I \implies y \in A.$$

Equivalently,

$$\pi_I^{-1}\pi_I[A] = A.$$

Lemma 15.11 (Clopen sets in the Cantor set)**.** *A set* $A \subseteq \mathcal{C}$ *is clopen iff* $A \sim I$ *for some finite* $I \subseteq \mathbb{N}$*. In particular, clopen sets can be written as a finite union of disjoint basis clopens* $[\varphi_i]$ *for* φ_i *with finite domain.*

(Direction one). If A is clopen, then

$$A = \bigcup_i [\varphi_i]$$

for some finitely many (by compactness) φ_i with finite domain I_i. Then

$$A \sim \bigcup_i I_i.$$

□

(Other direction). if $A \sim I$, blabla □

Immediately, a lemma follows.

Lemma 15.12 (Cantor set is zerodimensional)**.** *The Cantor set $\mathcal{C}$ is zerodimensional, i.e. it has a base of clopen sets.*

Theorem 15.13 (Topological characterisation of the Cantor set)**.** *If a topological space K is compact, metrizable, zerodimensional with no isolated points, then*

$$K \cong C.$$

§15.4 The group structure

The Cantor set has a natural abelian group structure given by its product structure. We can phrase it even more efficiently when we think of $\mathcal{C}$ as $\mathcal{P}(\mathbb{N})$ – the symmetric difference (or xor for the informatically inclined).

$$A \oplus B := A \Delta B$$

Every element has order two!

Fact. Together with the operation $\oplus$, the Cantor set $\mathcal{C}$ is a compact topological group, i.e. the function

$$(x, y) \mapsto x \oplus y$$

is continuous (in general the second element is inversed, but here every element is its own inverse anyway).

§15.5 Measure

We can define the measure on the Cantor set as a countable product of probability measures:

$$\nu = \bigotimes_{n=1}^{\infty} (\frac{1}{2}(\delta_0 + \delta_1)).$$

But we will do it by hand.

Definition 15.14. Let $A \subseteq \mathcal{C}$ be clopen. Then

$$A \sim \{1, 2, \ldots, n\}$$

for some n. Let

$$A' := \pi_{\{1,2,\ldots,n\}}[A]$$

We define its measure to be

$$\nu(A) := \frac{\#A'}{2^n}.$$

This makes sense with the probabilistic definition.

Theorem 15.15 (Well-definedness of the premeasure)**.** *The function*

$$\nu : \operatorname{Clop} \mathcal{C} \to \mathbb{R}$$

is a well-defined, additive function on the set algebra $\operatorname{Clop} \mathcal{C}$*.*

Proof. Since the Cantor set is compact, ν is automatically downward continuous on the empty set. By Caratheodory's Theorem, ν extends uniquely to a probabilistic measure on

$$\operatorname{Bor}\mathcal{C} = \sigma\left(\operatorname{Clop}\mathcal{C}\right).$$

What now??

$$\mathcal{A} := \left\{B \in \operatorname{Bor}\mathcal{C} : \forall \varepsilon > 0 . \exists A \in \operatorname{Clop} C . \nu(A \Delta B) < \varepsilon\right\}.$$

We prove that this is a σ-algebra.

There is a nice formula for cylinders.

Lemma 15.16 (Measure of a cylinder)**.** *For a partial function*

$$\varphi : \mathbb{N} \multimap \{0, 1\},$$

its cylinder has measure

$$\nu\left[\phi\right] = 2^{-|\operatorname{dom}\varphi|}.$$

The result holds even if $\operatorname{dom}\varphi$ *is infinite, in which case the measure is* 0*.*

Proof. For finite-domain partial functions ϕ, take

$$\operatorname{dom}\varphi =: I \subseteq \{1, 2, \ldots, n\} =: [n]$$

for some n. Then

$$\left|\pi_{[n]}[\varphi]\right| = \frac{2^{n-|I|}}{2^n} = 2^{-|I|}.$$

For infinite-domain functions ϕ, taking a decreasing intersection

$$[\phi] = \bigcap_n \left[\phi|_{[n]}\right]$$

shows that the measure of the intersection is 0. $\square$

Theorem 15.17. *The measure ν is the Haar measure on $\mathcal{C}$, that is, the unique probability measure invariant under group actions*

$$\nu\left(x \oplus B\right) = \nu(B)$$

for all $x \in \mathcal{C}$, $B \subseteq \mathcal{C}$.

Proof. Let us first consider $B = [\varphi]$, and $I = \operatorname{dom}\varphi$. Then

$$\nu\left(x \oplus [\varphi]\right) = \nu\left([x \oplus \varphi]\right) = \nu\left([\varphi]\right).$$

A clopen is a disjoint sum of $[\varphi_i]$ for finitely many φ_i, so additivity on clopens follows. Now, take a superficially different measure

$$\nu_x(B) := \nu\left(x \oplus B\right).$$

Since ν and ν_x agree on clopens, by uniqueness in Caratheodory's Theorem they agree on all sets. $\square$

Note the isomorphism

$$(C, \oplus) \cong (\mathcal{P}(\mathbb{N}), \Delta)$$

of (topological) groups.

§15.6 Normal number theorem

Definition 15.18. Let $A \subseteq \mathcal{C}$. We call A a **tail** set if

$$A \sim \{k : k \geqslant n\}$$

for all n. Equivalently, if $x \in A$ and $x(n) = y(n)$ for almost all n, then $y \in A$.

Example. A naturally occuring example of a tail set is

$$A_\beta := \left\{ x \in \mathcal{C} : \lim_n \frac{x(1) + \ldots x(n)}{n} = \beta \right\}.$$

Theorem 15.19 (Kolmogorov zero-one law for the Cantor set)**.** *A borel tail set $A \subseteq \mathcal{C}$ has measure 0 or 1.*

Proof. Take a basis set $[\varphi]$. We have

$$\nu\left([\varphi] \cap A\right) = \nu\left([\varphi]\right) \cdot \nu(A).$$

From this immediately follows that this work for any $B \in \operatorname{Clop} \mathcal{C}$. Now approximate A by a clopen B so that

$$\nu\left(A \Delta \mathcal{C}\right).$$

To finish the proof, compute

$$\nu(A) \cdot \nu(B) = \nu(A \cap B) \geqslant \nu(A) - \varepsilon \nu(A).$$

□

Returning to the example we have $\nu(A_\beta) \in \{0, 1\}$. We have

$$\nu(A_\beta) = \nu(A_{1-\beta}).$$

Theorem 15.20 (Borel's normal number theorem)**.**

$$\nu\left(A_{\frac{1}{2}}\right) = 1.$$

Remark. According to Billingsley, this theorem was the founding work of modern probability theorem, which is founded on limit theorems.

Proof. Denote for $\alpha < \frac{1}{2}$

$$B_n^\alpha = \left\{ x \in \mathcal{C} : \frac{x_1 + x_2 + \ldots + x_n}{n} \leqslant \alpha \right\}.$$

We claim that there exists a θ such that

$$\nu\left(B_n^\alpha\right) \leqslant \theta^n.$$

Then

$$\nu\left(B_n^\alpha\right) = \frac{c_n}{2^n},$$

where

$$c_n = \sum_{k=1}^{\lfloor \alpha n \rfloor}.$$

□

Chapter 16

Measures on separable, metrisable topological spaces

§16.1 Basic properties

For brevity, we will denote the class of separable and metrisable topological spaces by $\mathcal{SM}$. A lot of the time, it is easier to work with such spaces in a *common box*, i.e. use a universal space in which all of these spaces can be embedded. Luckily, we have such a space – the Hilbert Cube.

Ankified

Theorem 16.1. *Every $\mathcal{SM}$ topological space embeds in the Hilbert Cube.*

Proof. Fix a metric $d \leqslant 1$ and a countable dense subset $x_n \in X$. We define the embedding as

$$f_n(x) := d(x, x_n).$$

This is a product of continuous functions, so it continuous. It is injective, as if $f(x) = f(y)$, then a subsequence of x_n convergent to x is also convergent to y, so $x = y$.

The most difficult fact is that this is open. To see this, take $x \in U \subseteq X$ with U open. We will show that $f[U]$ is open. The neighbourhood U contains some ball $B(x, r)$. We can find an element x_k of the countable dense set such that $d(x, x_k) < r/4$. Then

$$x \in B(x_k, r/2) \subseteq B(x, r) \subseteq U,$$

which implies

$$f(x) \in f\left[B(x_k, r/2)\right] \subseteq f[U].$$

But, by definition of f,

$$f\left[B(x_k, r/2)\right] = f[X] \cap \pi_k^{-1}(-\infty, r/2).$$

Since x was an arbitrary element of U, we have that $f[U]$ is open, so f is a homeomorphism onto its image! □

You may wonder how this relates to the fact that all compact metrisable embed in the Hilbert Cube (as closed sets!). In turns out that compact metrisable spaces are $\mathcal{SM}$. We only need separability, and compactness together with a covering by balls gives us a countable dense subset rather easily.

Lemma 16.2. *If a metrisable topological K is compact, then it is separable.*

Proof. For each n, finitely many balls of radius $1/n$ cover K by compactness. Taking the centers of all such balls over all n yields a countable dense subset. □

The proof above can be trivially extended to totally bounded spaces and Lindelöf spaces. In the second case, we have countably many centers of balls at each step.

Lemma 16.3. *Let K be a metrisable topological space with is either*

1. *compact,*
2. *Lindelöf,*
3. *σ-compact,*
4. *or totally bounded.*

Then K is separable.

We will now investigate for a moment how properties of $\mathcal{SM}$ spaces are reflected in functions on such spaces. Since we care about topology, we restrict our attention to continuous functions. Unfortunately, even continuous functions on an arbitrary $\mathcal{SM}$ space can have an untame structure. Therefore we restrict our attention to bounded functions.

Definition 16.4. The space of bounded, continuous functions from a topological space X to $\mathbb{R}$ is denoted by

$$C_b(X).$$

If we want bounded functions into $\mathbb{C}$, we use the notation

$$C_b(X;\mathbb{C}).$$

This function space has the obvious structure of a linear space, and even an algebra with pointwise addition, scaling and multiplication. This space also has its own topology induced by the supremum norm.

Lemma 16.5. *Let X be an arbitrary.*

$$C_b(X)$$

is a Banach algebra under pointwise operations and the supremum norm.

Proof for compact spaces. Take a Cauchy sequence f_n. For each $x \in X$, $f_n(x)$ is Cauchy, so it converges. Therefore, the sequence of functions coverges pointwise to a limit function f. Suppose the convergence is not uniform. Then for some $\varepsilon > 0$ we can take a sequence x_n such that

$$f_n(x_n) - f(x_n) \geqslant \varepsilon.$$

By the $\varepsilon/3$ trick, f is continous. By compactness of X, x_n has a subsequence convergent to x_0. Since $f_n(x_0) \to f(x_0)$ and $f(x_n) \to f(x_0)$. TODO! □

Proof. The only thing one need to check is that Cauchy sequences actually converge. Let f_n be a Cauchy sequence. For any $x \in X$, $f_n(x)$ is a Cauchy sequence of real numbers, so it converges. Therefore, f_n converges pointwise to a function f. Note that so far we don't know if the convergence is uniform, or even if the function is continuous.

Let N_ε be the point after which the sequence f_n is ε-close. Then for $n, m > N_\varepsilon$ we have

$$|f_n - f_m| \leqslant \varepsilon$$

uniformly on X. Keeping n constant and passing with m to the limit we have

$$|f_n(x) - f(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leqslant \varepsilon.$$

Therefore, f_n converges to f uniformly. In particular, f is continuous and bounded. □

We recall here a useful theorem, whose proof can be found in the literature.

Theorem 16.6 (Stone-Weierstrass)**.** *Let K be a compact, Hausdorff space and let $W \subseteq C_b(K)$ be a subalgebra. If the subalgebra distinguishes points, we have*

$$\overline{W} = C_b(K).$$

Proof. A classic proof due to Lebesgue can be found in Engelking (or prof. Szwarc's functional analysis notes). □

The first and so far only property we investigate for $C_b(X)$ is separability. It turns out that this space is rarely separable, and an exact characterisation can be given in terms of X.

Ankified

Theorem 16.7. *4 Let X be an $\mathcal{SM}$ topological space. Then $C_b(X)$ is separable iff X is compact.*

Proof ($\Leftarrow$). We know by 16.1 that X can be regarded as a subspace of $[0,1]^{\mathbb{N}}$. Since X is compact, it is a closed subset of the cube. By the Stone-Weierstrass Theorem 16.6, the algebra generated by coordinate projections, which consists of finite linear combinations of finite products of coordinate projections, is dense in $C_b(X)$. This subalgebra is in general not countable, however. One small fix is required to find a countable dense subset of $C_b(X)$ – only take rational coefficients in linear combinations. □

Proof ($\Rightarrow$). We'll mirror the proof of the nonseparability of $C_b(\mathbb{R})$ – we will find $\mathfrak{c}$ many balls of radius $\frac{1}{2}$.

Since X is noncompact and metrisable, we have a discrete sequence of elements of X. Call it a_n and let $A = \{x_n : n \in \mathbb{N}\}$. For a subset I of the natural numbers, we define the function $f_I : A \to [0,1]$ by

$$f_I(x_i) := \begin{cases} 1 & i \in I \\ 0 & i \notin I. \end{cases}$$

We can extend all these to function $\widetilde{f}_I : A \to [0,1]$ with the Tietze Extension Theorem. Now, for $I \neq J$, if we look at an element $x_0 \in I\Delta J$ we get

$$\left\| \widetilde{f}_I - \widetilde{f}_J \right\|_\infty \geqslant \left| \widetilde{f}_I(x_0) - \widetilde{f}_J(x_0) \right| = 1.$$

□

Lemma 16.8 (Generating algebras)**.** *Take a finite or countable subset*

$$\mathcal{F} = \{f_1, f_2, \ldots\} \subseteq C_b(X).$$

Then, the subalgebra $\langle \mathcal{F} \rangle$ generated by $\mathcal{F}$ contains functions of the form

$$\sum_{i=1}^{n} \prod_{j=1}^{m_i} a_i j f_{ij}$$

for some $a_{ij} \in \mathbb{R}$ and $f_{ij} \in \mathcal{F}$. We can get a countable dense set by taking all such expressions with rational coefficients.

§16.2 Polish spaces

We now turn to a subclass of $\mathcal{SM}$ spaces which is particularly useful and important.

Definition 16.9. A **Polish space** is an $\mathcal{SM}$ space X, which is completely metrisable.

Please note that this depends on the topology and not on any given metric for the space, as the example below shows.

Example. The space $(0, 1)$ is Polish, since it is homeomorphic to $(0, \infty)$. However, it is definitely not complete with regards to its standard metric! An explicit complete metric can be given by

$$d(x, y) := \left| \operatorname{tg} \frac{x\pi}{2} - \operatorname{tg} \frac{y\pi}{2} \right|,$$

which is the pullback of the complete metric from $(0, \infty)$ by a homeomorphism.

Example. An even weirder example is

$$\mathbb{R} \setminus \mathbb{Q} \cong \mathbb{N}^{\mathbb{N}}.$$

That this is completely metrisable can be seen from the following theorem.

Theorem 16.10 (Alexandroff)**.** *A subspace Y of a Polish space X is itself Polish iff Y is a G_δ subset of X.*

Proof. Let ρ' be a new metric on Y given by

$$\rho'(y_1, y_2) = \rho(y_1, y_2) + \sum_n \min\left(\frac{1}{2^n}, \left| \frac{1}{\rho(y_1, V_n^c)} - \frac{1}{\rho(y_2, V_n^c)} \right| \right)$$

The rest of the details can be found in Kerchis' classical book on Descriptive Set Theory. □

Definition 16.11. In a topological space X, the **Borel subsets** of X are precisely the elements of $\operatorname{Bor} X := \sigma(\tau_X)$.

Definition 16.12. For a topological, but especially $\mathcal{SM}$ or Polish space X, we denote the set of **probability measures** on $\operatorname{Bor} X$ by $\mathbb{P}(X)$.

We need a tool before stating proving the next theorem.

Lemma 16.13. *A closed set in a metric space is G_δ. Conversely, any open set is F_σ.*

Proof. We will use the ε-neighbourhoods of F, i.e.

$$F_\varepsilon := \{x \in X : d(x, F) < \varepsilon\},$$

which are open. Since F is closed, we have

$$F = \bigcap_{n=1}^{\infty} F_{1/n}.$$

We need closedness for the left-to-right inclusion. □

Theorem 16.14 (First Regularity Theorem)**.** *For any $\mathcal{SM}$ (in fact, any metrisable) space X and $\mu \in \mathbb{P}(X)$, the measure μ is **regular**, that is for any $B \in \operatorname{Bor} X$ and $\varepsilon > 0$ there are two sets*
Ankified *$F \subseteq B \subseteq V$, respectively closed and open, such that*

$$\mu\,(V \setminus F) < \varepsilon.$$

Proof. Let $\mathcal{A}$ be the family of all sets with the given propety. We will prove that it is a σ-algebra and that it contains closed sets. By 13.3, we only have to check for complements, finite sums and ascending countable sums.

Closed sets are in $\mathcal{A}$ because they are G_δ, see Lemma 16.13, and because a probability measure is downward continuous.

Closure under complements is inherent in the definition. If $F \subseteq B \subseteq V$, then

$$V^c \subseteq B^c \subseteq F^c,$$

F^c is open, V^c is closed and

$$F^c \setminus V^c = V \setminus F,$$

so the approximation still works.

Finite sums are easy. If $F_i \subseteq B_i \subseteq V_i$ for $1 \leqslant i \leqslant n$ with

$$\nu\,(V_i \setminus F_i) \leqslant \frac{\varepsilon}{n},$$

then

$$\bigcup_{i=1}^{n} F_i \subseteq \bigcup_{i=1}^{n} B_i \subseteq \bigcup_{i=1}^{n} V_i$$

and

$$\bigcup_{i=1}^{n} V_i \setminus \bigcup_{i=1}^{n} F_i \subseteq \bigcup_{i=1}^{n} (V_i \setminus F_i).$$

Passing to measure

$$\nu\left(\bigcup_{i=1}^{n} V_i \setminus \bigcup_{i=1}^{n} F_i\right) \leqslant n \cdot \frac{\varepsilon}{n} = \varepsilon.$$

This works, because we have implicitly used that finite sums of closed sets are closed, and the same for open sets. However, this fails for countable sums in the *closed* part. How do we repair this? Let's do the setup first. For an increasing sequence B_n take approximations $F_n \subseteq B_n \subseteq V_n$ such that

$$\mu\,(V_n \setminus F_n) < \frac{\varepsilon}{2^{n+1}}$$

and let $B = \bigcup_n B_n$. By summing prefixes of F_n, we may assume that F_n is an increasing sequence. Then

$$\bigcup_{n=1}^{\infty} F_n \subseteq \bigcup_{n=1}^{\infty} B_n \subseteq \bigcup_{n=1}^{\infty} V_n$$

and

$$\bigcup_{n=1}^{\infty} V_n \setminus \bigcup_{n=1}^{\infty} F_n \subseteq \bigcup_{n=1}^{\infty} (V_n \setminus F_n)$$

(this is a general identity on sets). We are now facing the closed set problem again. Fortunately, all measures are upward continuous, so we can take a really good approximation by a prefix! Namely,

$$\mu\left(\bigcup_{n=1}^{\infty} F_n \setminus \bigcup_{n=1}^{N} F_n\right) < \frac{\varepsilon}{2},$$

and obtain an approximation

$$\bigcup_{n=1}^{N} F_n \subseteq B \subseteq \bigcup_{n=1}^{\infty} V_n$$

with

$$\mu\left(\bigcup_{n=1}^{\infty} V_n \setminus \bigcup_{n=1}^{N} F_n\right) < \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+1}} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Remark. This proof does not use separability at all, and only uses metrisability to obtain the supplementary lemma 16.13. Similarly, we only need that μ is a probability measure to get downward continuity. Therefore, the proof lifts immediately to σ-finite measures on spaces where closed sets are G_δ.

Remark. This implies that we don't care much about Descriptive Set Theory. For $X \in \mathcal{SM}$ and $\mu \in \mathbb{P}(X)$, we only care about F_σ and G_δ sets. More formally, for any $B \in \operatorname{Bor} X$, any set is up to a set of measure 0 an F_σ from below and G_δ from above.

Remark. An analytical set is an image of a Polish space.

Ankified **Theorem 16.15.** *If X is a Polish space and $\mu \in \mathbb{P}(X)$, then μ is **tight**, i.e. for every $\varepsilon > 0$ there is a compact K such that*

$$\mu(K) > 1 - \varepsilon.$$

Proof. Let $d(-,-)$ be a complete metric on X and let x_n be a countable dense set.

$$X = \bigcup_{n=1}^{\infty} B\left(x_n, \frac{1}{n}\right).$$

By upward continuity of measure, we can take a k_n such that the first k_n balls are large, i.e. larger than

$$1 - \frac{\varepsilon}{2^n}.$$

Now denote

$$K_n := \bigcup_{k=1}^{k_n} \overline{B\left(x_k, \frac{1}{k}\right)},\ K = \bigcap_{n=1}^{\infty} K_n.$$

To see that K is large in μ, use Lemma 13.7. Note that K is closed.

To prove compactness, we will find for a sequence x_k in K a convergent subsequence. We mirror the proof that [0, 1] is sequentially compact. Infinitely many elements of x_k will belong to one of the (finitely many!) balls that make up K_1, and of those inifitely many will land in one of the balls that make up K_2 etc. Therefore, x_n has a Cauchy subsequence. Since X is Polish, this subsequence is convergent in X, but as K is closed, the limit is actually in K. □

Example. Let $X = \mathbb{N}^{\mathbb{N}}$.

Corollary. Every Borel set can be approximated by a compact set.

Remark. This works for some non-Polish spaces, for example $\mathbb{Q}$. This also sometimes *almost fails* for Polish spaces.

Remark. There is an uncountable $X \subseteq \mathbb{R}$ such that no uncountable closed set in contained in X. Note that it can't be Borel. By **List 3.1.**, there is a measure on X that vanishes on points. In particular, all compact subsets are countable and therefore have measure 0. This is the Baire space.

Lemma 16.16. *Let (X, d) be a complete metric space and $F_n \subseteq X$ be a descending sequence of nonempty closed sets. If*

$$\operatorname{diam} F_n \to 0$$

then the intersection $\bigcap_{n=1}^{\infty} F_n$ is nonempty.

Theorem 16.17 (Lusin)**.** *If $X \in \mathcal{SM}$ and $f : X \to \mathbb{R}$ is a Borel function, then for any $\mu \in \mathbb{P}(X)$ and $\varepsilon > 0$ there exists a closed set F of large measure such that f is continuous on F. If X is Polish, the set F can even be compact.*

Characteristic functions. Let $f = \chi_B$ for $B \in \operatorname{Bor} X$. By regularity of measure, we can find a closed F and open U such that $F \subseteq A$ and $\mu(U \setminus F) < \varepsilon$. The function is constant on F and constant on U^c. □

Simple functions. Let $f = \sum_i a_i \chi_{B_i}$. Take good sets F_i for χ_{B_i}. Then f is continuous on

$$F_1 \cap F_2 \cap \ldots \cap F_n.$$

□

Bounded functions. They are uniform limits of simple functions, and the thesis is preserved by uniform limits. Let $f_i \to f$. Then find sets F_i, on which f_i is continuous and

$$\mu(F_i) > 1 - \frac{\varepsilon}{2^i}.$$

□

General functions. For some large M, $f^{-1}[-M, M]$ has measure larger than $\varepsilon/2$. Approximate it from below by a closed set by regularity. □

Function spaces

We have that $L_1(\mu) \cong \ell_1(\kappa)$.

Definition 16.18 (Separable of measures)**.** A measure $\mu \in \mathbb{P}(X)$ is **separable** iff the function space $L_1(X, \mu)$ is separable.

This definition is external in a way. Can we characterise this property is terms of the measure itself?

Lemma 16.19. *Let $(X, \mathcal{B}, \mu)$ be a probabilistic measure space. If there is a countable subcollection $\mathcal{A} \subseteq \mathcal{B}$ such that*

$$\inf_{A \in \mathcal{A}} B \Delta A = 0.$$

Proof. Take rational finite linear combinations of χ_{A_i}. The closure of this set contains. The closure of this set contains all charactestic functions, then all simple functions, and then all integrable functions. □

Theorem 16.20. *If $X \in \mathcal{SM}$ and $\mu \in \mathbb{P}(X)$, then $L^1(\mu)$ is separable.*

Proof. Note that X is second countable. There is a countable basis $\mathcal{B}$ closed under finite unions. Then we can approximate any open U approximated well by a set in $\mathcal{B}$, and any $B \in \operatorname{Bor} X$ can be approximated by open U, and use the triangle inequality for symmetric difference. Then use Lemma 16.19. □

Lemma 16.21. *A continous linear functional is bounded.*

Proof. Take v_n of length 1, such that $f(v_n) \leqslant n$. Then $v_n/n \to 0$, but $f(v_n)/n \not\to 0$. □

For any probability measure $\mu : X \to \mathbb{R}$, we can interpret it as a functional

$$\hat{mu} : C_b(X) \to \mathbb{R}.$$

This functional has norm one, but is the $\hat{}$ operator injective?

Lemma 16.22. *For $X \in \mathcal{SM}$ and $\mu, \nu \in \mathbb{P}(X)$. If $\hat{(}\mu) = \hat{(}\nu)$, then $\mu = \nu$.*

Proof. We prove that the measures agree on open sets, the general thesis follows by regularity. Pick U and an increasing sequence of clopsed sets F_n. By Urysohn's Lemma, there are functions $f_n : X \to [0, 1]$ that are 1 on F_n and 0 on U^c. Use LDCT to finish. □

Are all functionals measure? No, for sign reasons. A functional can send the constant function 1 to -13 or 7. What about functionals of norm 1?

Theorem 16.23 (Riesz representation theorem)**.** *Let K be a compact metrisable (and hence also separable!) space, and let*

$$\varphi : C_b(K) \to \mathbb{R}$$

be a linear functional which is positive and of norm 1. Then there is a probability measure μ such that

$$\varphi(f) = \int_K f(x) \, \mathrm{d}\mu(x).$$

Chapter 17

Boolean algebras

Theorem 17.1. *Let (X, Σ, μ) be a nonatomic Borel probability space and μ be separable (or, more scientifically, of countable type). Then there exists a Borel isomorphism*

$$h : \Sigma/\mathcal{N}(\mu) \to \mathrm{Bor}[0,1]/\mathcal{N}(\lambda)$$

that preserves the measure, i.e.

$$\lambda \circ h = \mu.$$

A measure is of countable type iff there is a countable family $\mathcal{C}_1 \subseteq \Sigma$ such that for all $A \in \Sigma$ and $\varepsilon > 0$ there is a $C \in \mathcal{C}_1$ such that

$$\mu\left(C \Delta A\right) < \varepsilon,$$

i.e. the algebra with the Frechet-Nikodym metric is separable.

Proof. This will be a back-and-forth proof. Denote the algebras by $\mathbb{A}, \mathbb{B}$. We shall endeavour to find the countable dense sets $\mathbb{A}_0, \mathbb{B}_0$ which are isomorphic via an isomorphism h_0. Fix $\mathcal{C}_1, \mathcal{C}_2$ – countable dense subsets of Σ_1, Σ_2. We will express them as

$$\mathbb{A}_0 = \bigcup_{n=1}^{\infty} \mathbb{A}_n$$

and we will have isomorphisms $h_n : \mathbb{A}_n \to \mathbb{B}_n$ that extent one another. When picking a new set, some atoms will split in two and we will extend the isomorphism using the Darboux property of nonatomic measures. $\square$

Corollary. Informally said, there is only one nonatomic probability measure on $\mathcal{SM}$ spaces.

Theorem 17.2. *Let X be $\mathcal{SM}$, $\mu \in \mathbb{P}(X)$. Then*

$$L^p(\mu) \cong L^p[0,1],$$

i.e. they are ***linearly isometric****. In particular, the first space is Banach!*

Proof. We do the standard thing. Put down

$$T(\chi) = T(\chi_B),$$

where $h(\dot{A}) = \dot{B}$. By linearity, we extend T to simple functions, i.e. it is a linear isometry between the normed spaces
$$Simp(\Sigma) \to Simp(\mathrm{Bor}[0,1]).$$
We have a linear isometry between dense subspaces, so we have a dense isometry between whole spaces. Explicitly
$$\left\| \sum c_i \chi_{A_i} \right\|_p^p = \sum |c_i|^p \, \mu(A_i).$$
□

Remark. For Polish spaces, we even have a whole isomorphisma, not only between algebras!

Questions we yet have to answer.

1. There is an $\mathcal{SM}$ space X and $\mu \in \mathbb{P}(X)$ such that $\mu(K) = 0$ for all compact $K \subseteq X$. This is the Bernstein set.
2. The existance of an $\mathcal{SM}$ space such that all probability measures are purely atomic.
3. Interesting measures on nonseparable metric spaces.
4. Does λ extend to all subsets of $[0,1]$.

§17.1 Recap of set theory.

An ordinal number is a set well-ordered by the

Lemma 17.3 (Ordinal number ordering). *Any two well-ordered sets are either isomorphic or one embeds as an initial segment of the other.*

The set ω is the only inifite well-ordered set without a maximal element whose all initial segments are finite.

Theorem 17.4 (Bernstein Set Theorem). *There exists a set $Z \subseteq [0,\, 1]$ such that for all uncountable compact sets K*
$$Z \cap K \neq \varnothing \neq Z^c \cap K.$$

Proof. There are $\mathfrak{c}$-many such sets K and they all have an embedded Cantor set (every uncountable Polish space does), so they are of size $\mathfrak{c}$. Now for $\alpha < \mathfrak{c}$ we can take
$$x_\alpha, y_\alpha \in K_\alpha \setminus \{x_\beta, y_\beta : \beta < \alpha\}$$
and
$$Z := \{x_\alpha\}.$$
□

Lemma 17.5 (Measure of the Bernstein Set). *We have*
$$\lambda^*(Z) = \lambda^*(Z^c) = 1.$$

Proof. If $\lambda^*(Z) < 1$, then there is a $B \in \mathrm{Bor}[0,\, 1]$ such that $\lambda(B) < 1$, so $B^c \subseteq Z^c$ contains a compact set of positive measure, contradiction. □

Theorem 17.6 (Non-tight measure)**.** *The Lebesgue measure restricted to the Bernstein set is not tight.*

Proof. The Borel substs of Z are restriction of Borel subsets. The

$$\mu(B \cap Z) = \lambda(B)$$

defines a measure. The Bernstein set has only countable compact subsets, and on them

$$\mu(K) = 0.$$

□

Remark. The Bernstein set is even more nonmeasurable. In fact, for all nonatomic measures

$$\nu^*(Z) = \nu^*(Z^c) = 1.$$

We not turn to the Continuum Hypothesis.

Theorem 17.7 (Lusin set)**.** *Suppose the Continuum Hypothesis. Then there exists a **Lusin set**, i.e. an uncountable $L \subseteq [0,\,1]$ such that $L \cap P$ for all closed nowhere dense sets P.*

Proof. Let $\{F_\alpha : \alpha < \omega_1\}$ be a list of all closed nowhere dense sets. Then we define

$$L = \{x_\alpha : \alpha < \omega_1\},$$

where

$$x_\alpha \in [0,\,1] \setminus \left(\bigcup_{\beta<\alpha} F_\beta \cup \{x_\beta : \beta < \alpha\} \right) \neq \varnothing.$$

□

The consequences of that for measure theory.

Theorem 17.8. *Let L be a Lusin set. Then*

1. *Every nonatomic $\mu \in \mathbb{P}[0,1]$ has $\mu^*(L) = 0$.*
2. *Every $\nu \in \mathbb{P}(L)$ is nonatomic.*

Proof of (1). Let $\mu \in \mathbb{P}[0,1]$ be nonatomic. Then there is a sequence of nowhere dense $F_n \subseteq [0,1]$ such that

$$\mu\left(\bigcup F_n\right) = 1.$$

Dually, $\mu(G) = 0$ for a dense G_δ set G. This follows from regularity – take a measurable hull of a countable dense set. □

Proof of (2). Ad absurdum, a measure $\nu \in \mathbb{P}(L)$ extends to a measure on $\mathbb{P}[0,\,1]$ which is still nonatomic, contraqdicting (1). □

Remark. This gives that measure and category are quite orthogonal with what is understood as *small.*

Theorem 17.9. *If there exists a Lusin set $L \subseteq [0,\,1]$ then there are sets $E_n \subseteq [0,\,1]$ such that λ does not admit an extension to a measure on*

$$\sigma\left(\mathrm{Bor}[0,\,1], E_1, \ldots\right)$$

Chapter 18

Measures on Topological Spaces, Problemset 1

Problem 4

Extension 1

We show that the set can be the graph of a function! Let Z be a borel set of positive measure and define

$$T_Z = \{x : \lambda(Z_x) > 0\}.$$

Then T_Z is a measurable set by Fubini's Theorem. We can pick a compact subset T'_Z. A compact set of positive measure has at least $\mathfrak{c}$ elements, and there are as many borel sets. Then, enumerate borel sets of $\mathbb{R}^2$.

Problem 5

Set of undefined density at 0

TODO

Set of density t at 0

Presented in class by **Michał Baran**. Fix $t \in (0, 1)$.

The set we will construct will be symmetric around 0. We will find a sequence b_n such that with

$$A_n = \left(\frac{1}{n} - b_n, \frac{1}{n}\right)$$

we will have for all n

$$\frac{t}{n} = \lambda\left(\bigcup_{k=n}^{\infty} A_n\right) = \sum_{k=n}^{\infty} b_k,$$

so

$$b_n = \sum_{k=n}^{\infty} b_k - \sum_{k=n+1}^{\infty} b_k = \frac{t}{n(n+1)}.$$

Consider

$$A := \bigcup_{k=1}^{\infty} A_k \cup -A_k.$$

We will bound the fraction

$$\frac{\lambda\left(A \cap (-\delta, \delta)\right)}{2\delta} = \frac{\lambda\left(A \cap (0, \delta)\right)}{\delta}$$

from above and below. For $\delta \in (1/(n+1), 1/n]$ we have

$$\bigcup_{k=n+1} A_k \subseteq A \cap (0, \delta) \subseteq \bigcup_{k=n} A_k,$$

passing to measure

$$\frac{t}{n+1} \leqslant \lambda\left(A \cap (0, \delta)\right) \leqslant \frac{t}{n}.$$

When divided by δ, we get the result by the squeeze theorem.

Remark. The solution would work equally well if instead of $a_n = 1/n$ we used a sequence that converges to 0 monotonically and satisfies

$$\frac{a_n - a_{n+1}}{a_n} \to 0.$$

Problem 9

Presented in class by **dr Arturo Martinez Celiz**.

Wlog, everything happens within (0, 1). Following the hint, take a countable sequence A_i such that the set $B := \bigcup_i A_i$ has maximal measure.

By this choice, for any $C \in \mathcal{A}$, we have

$$\lambda\Big((C \cup B)\, \Delta B\Big) = 0,$$

so that

$$\phi(C \cup B) = \phi(B)$$

and

$$C \subseteq \phi(C \cup B) = \phi(B).$$

Since C was arbitrary

$$\bigcup \mathcal{A} \subseteq \phi(B) \implies B \subseteq \bigcup \mathcal{A} \phi(B).$$

Since $\lambda(B) = \lambda\left(\phi(B)\right)$ we know that the sum of $\mathcal{A}$ is measurable.

Problem 10

Presented in class by **Szymon Smolarek**.

We take a cover of **regular sets**, i.e. a family for which there exists a constant C such that

$$\operatorname{diam}^2 A \leqslant C\lambda_2(A).$$

It can be proven that if such a family is a Vitali cover of a set $A \subseteq \mathbb{R}^2$, an analogue of the VCT holds.

The family of all triangles does not satisfy the regularity condition – think of keeping one segment constant and bringing the third verted evert closer to the segment. To deal with this, we subdivide the family $\mathcal{T}$ into subfamilies

$$\mathcal{T}_n := \left\{ T \in \mathcal{T} : \operatorname{diam}^2 T \leqslant n\lambda_2(A) \right\}.$$

Reducing to a given subfamily, we can cover each triangle T by arbitrarily small traingles similar to T contained within T. This gives us a regular Vitali cover $\widetilde{T}_n$ of $\bigcup \mathcal{T}_n$.

Problem 11

Stated in class by **Szymon Smolarek**.

Theorem 18.1 (Steinhaus theorem for the Cantor Set)**.** *For any measurable set A, the set*

$$A \oplus A$$

contains an open neighbourhood of 0.

Theorem 18.2 (Vitali Covering Theorem for the Cantor Set)**.** *If a family of clopens $\mathcal{J} \subseteq \operatorname{Clop} \mathcal{C}$ is a Vitali cover of A, then there is a sequence $J_n \in \mathcal{J}$ such that*

$$\nu^* \left(A \setminus \bigcup_n J_n \right) = 0.$$

Theorem 18.3 (Lebesgue Density Theorem for the Cantor Set)**.** *Let $A \subseteq \mathcal{C}$. An element $a \in A$ is a **density point** of A if*

$$\lim_{n\to\infty} \frac{\nu\left(A \cap [a|_{[n]}]\right)}{2^{-n}} = 1.$$

If A is measurable, then almost all points of A are density points of A.

The proofs are quite the same, as $\mathcal{C}$ is a topological group and the measure ν is its Haar measure.

Problem 12

Hint. Use Baire's theorem.

Chapter 19

Measures on Topological Spaces, Problemset 2

Problem 1

Such an a exists by compactness of A and continuity of metric. If we have two a_1, a_2 such that

$$\rho(x, a_1) = \rho(x, a_2)$$

then a_1, a_2 must agree and disagree with x at all places, so in fact $a_1 = a_2$, thus r_A is well-defined. For any $a \in A$, $d(a, A) = 0 = d(a, a)$, so r_A is a retraction. What remains to be shown is continuity.

Let x, y agree up to $n_0(x, y)$. Then $r_A(x)$ and $r_A(y)$ also agree up to $n_0(x, y)$ – if they differed earlier, we could use $r_A(x)$ instead of $r_A(y)$ and get a closer point a in the definition! So we have

$$n_0(x, y) \leqslant n_0\left(r_A(x), r_A(y)\right)$$

and

$$d(x, y) \geqslant d\left(r_A(x), r_A(y)\right).$$

Remark. The metric $d(x, y) = 1/n_0(x, y)$, i.e. the first moment where x and y differ, won't work, because it can't tell apart points from which x differs at the same position!

Problem 2

Problem 3

Any $A, B \in \operatorname{Clop} \mathcal{C}$ can be written as disjoint sums of the basis cylinders $[\varphi]$ by 15.11. Since the condition distributes over disjoint sums, we will prove the statement for $A = [\varphi]$ and $B = [\psi]$ with

$$|\operatorname{dom} \varphi|, |\operatorname{dom} \psi| < \infty.$$

Let $I = \operatorname{dom} \varphi$, $J = \operatorname{dom} \psi$ be the disjoint(!) domains of φ, ψ. There is a function τ on $I \cup J$ such that

$$\tau|_I = \phi,\ \tau|_J = \psi.$$

For such a function,

$$[\varphi] \cap [\psi] = [\tau] .$$

Now take an n such that $I \cup J \subseteq \{1, 2, \ldots, n\}$ and denote the last set as $[n]$. By 15.16 we compute

$$\begin{aligned} \nu\,[\varphi] &= 2^{-|I|} \\ \nu\,[\psi] &= 2^{-|J|} \\ \nu\,[\tau] &= 2^{-|I \cup J|}, \end{aligned}$$

and $|I \cup J| = |I| + |J|$ finishes the proof.

The general case

Now take arbitrary $A, B \in \operatorname{Bor} \mathcal{C}$ such that $A \sim I$, $B \sim J$. Approximate A, B by clopens A', B' to within an ε, i.e. so that

$$\nu(A \Delta A'),\ \nu(B \Delta B') < \varepsilon.$$

We cannot use the clopen statement we just proved since a priori A' and B' could be determined by sets with nonempty intersection. We can, however, improve the approximation with

$$\widetilde{A} := \pi_I^{-1} \pi_I A'.$$

The set $\widetilde{A}$ is still a clopen – since A' was determined by a finite set K, $\widetilde{A}$ is determined by $K \cap I$. Additionally we have

$$\widetilde{A} \Delta A \subseteq A' \Delta A,$$

so we have improved the approximation! Now, do the same for B' and use the statement for clopens to finish up the solution.

Warning! The reasoning below does not work! (For tail sets, for example)

We can approximate A, B by decreasing sequences of clopens by putting down

$$A_n := \pi_{[n]}^{-1} \pi_{[n]} A$$

and the same for B_n. We also approximate their intersection by decreasing clopens in the same way, i.e.

$$C_n := \pi_{[n]}^{-1} \pi_{[n]} (A \cap B).$$

For these approximations

$$C_n = A_n \cap B_n,$$

so by the first subproblem

$$\nu(C_n) = \nu(A_n \cap B_n) = \nu(A_n) \cdot \nu(B_n).$$

Since the measure ν is probabilistic, and hence continuous, by passing to the limit $n \to \infty$ we get what we need.

The general case, by Dominik

We use the $\pi - \lambda$ Lemma **??** twice. Fix disjoint I and J. First, we make $A \sim I$ a clopen. The family of Borel sets B determined by J such that

$$\nu(A \cap B) = \nu(A) \cdot \nu(B)$$

is a λ-system containing all clopens determined by J, so it contains all Borel sets betermined by J.

For the second step, now take an arbitrary Borel $A \sim I$. By the previous step, the family of Borel sets $B \sim J$ that work contains all clopens determined by J and is a λ-system, so it contains all Borel sets.

The general case, shown by dr Celiz

Fix $A \sim I$, wlog $J = I^c$. Now take an open $U \supseteq A$.

Claim. Let A be Borel, $A \sim I$ and $a \in A$ and an open set $U \supseteq A$. Then there is a finite set $F \subseteq I$ such that $[a|_F] \subseteq U$.

After the Claim. Once the Claim is proved, we can prove the thesis for arbitrary open sets, and thus for any open sets by approximation.

Problem 4

Presented in class by **Szymon Smolarek**.

We estimate the complement of this set, which we denote by

$$B(\varepsilon) = A(\varepsilon)^c.$$

Consider a sequence of partial functions

$$\varphi_k : [kn,\, (k+1)n) \multimap\!\!\!\rightarrow \{0, 1\}$$

given by $\varphi_k(j) := \varepsilon(j - kn)$, thinking of ε as a partial function $\varepsilon : [0, n) \multimap\!\!\!\rightarrow \{0, 1\}$. Then

$$B(\varepsilon) \subseteq \bigcap_{k=1}^{\infty} [\varphi_k]^c,$$

so by the previous problem and 15.16 we have

$$\nu(B(\varepsilon)) \leqslant \nu\left(\bigcap_{k=1}^{\infty} [\varphi]^c\right) \leqslant \prod_{k=1}^{\infty} \nu\left([\varphi_k]^c\right) = 0,$$

since $\nu\left([\varphi_k]^c\right) = 1 - 1/2^n$.

Problem 6

Any clopen $C \in \operatorname{Clop}\mathcal{C}$ is a disjoint sum of basis cylinders by 15.11. Since $\oplus$ is a group operation, the function

$$l_x(y) = x \oplus y$$

is bijective, so on the level of sets l_x distributes over disjoint sums. We check the property for a cylinder $[\varphi]$. This is easy, since

$$\nu\Big(x \oplus [\varphi]\Big) = \nu[x \oplus \varphi] = 2^{-|\operatorname{dom}\varphi|} = \nu[\varphi]$$

by 15.16. Now consider the family of sets

$$\mathcal{A} := \Big\{A : \forall x \in \mathcal{C}.\, \nu(A) = \nu(x \oplus A)\Big\}.$$

We will show that this is a σ-algebra. Since we have already shown that it contains all the clopens, which form a basis of the topology on $\mathcal{C}$, it will automatically be equal to $\operatorname{Bor}\mathcal{C}$ by 13.4.

A σ-algebra can be generated by complements and countable sums (see 13.3). As mentioned before, l_x respects these operations, so

$$\nu(x \oplus A^c) = \nu\Big((x \oplus A)^c\Big) = 1 - \nu(x \oplus A) = 1 - \nu(A) = \nu(A^c)$$

and

$$\nu\left(x \oplus \bigcup_i A_i\right) = \nu\left(\bigcup_i x \oplus A_i\right) = \sum_i \nu(x \oplus A_i) = \sum_i \nu(A_i) = \nu\left(\bigcup_i A_i\right).$$

Problem 7

The identification is

$$A \mapsto \chi_A,\ x \mapsto \{n : x_n = 1\}.$$

One easily checks that these two are mutually inverse. Addition modulo 2 comes out to 1 iff exactly one of the summands is 1, and this corresponds exactly to belonging to the symmetric difference.

Problem 8

A filter cannot contain both A and A^c, since then it would contain $A \cap A^c = \varnothing$. Thus, a filter containing for all A either A or A^c is maximal.

For the other direction, suppose neither A nor A^c is in a filter $\mathcal{F}$. We define its *extension* by A as

$$\mathcal{F}_A = \{A' \cap F : A \subseteq A', F \in \mathcal{F}\}.$$

We check that this is a filter.

1. If $\varnothing \in \mathcal{F}_A$, $\mathcal{F}$ contains a set disjoint with A, so by the superset property it contains A^c.
2. Let $A_1 \cap F_1, A_2 \cap F_2 \in \mathcal{F}_A$. Then

$$A \subseteq A_1 \cap A_2, F_1 \cap F_2 \in \mathcal{F},$$

so $(A_1 \cap F_1) \cap (A_2 \cap F_2) = (A_1 \cap A_2) \cap (F_1 \cap F_2) \in \mathcal{F}_A$.

3. Let $B \supseteq A' \cap F$. Then

$$B = B \cup (A' \cap F) = (B \cup A') \cap (B \cup F)$$

and $A \subseteq A' \cup B$, $F \subseteq B \cup F$, so $B \in \mathcal{F}_A$.

Of course, $A \in \mathcal{F}_A \setminus \mathcal{F}$, so $\mathcal{F}$ was not maximal in the first place.

Remark. One can check that $\mathcal{F}_A$ is the minimal filter containing $\mathcal{F}$ and A.

Remark. A filter is free iff it contains the Frechet filter.

Problem 9

Non-measurability

Take an $F \subseteq \mathcal{C}$ corresponding to $\mathcal{F}$ is the sense of problem 7. Assume F is measurable. By the Zero-One Law, either $\nu(F) = 1$ or $\nu(F) = 0$.

Take the set of complements of F, i.e. $F \oplus \mathbb{1}$. We have

$$F \cap (F \oplus \mathbb{1}) = \varnothing,\ F \cup (F \oplus \mathbb{1}) = \mathcal{C}.$$

We now have $\nu(F) = 1/2$, which is a contradiction.

Remark. The only principal ultrafilters are generated by singletons, so they are definitely measurable.

Outer/inner measure

Approximate the filter F by a tail Borel set.

Problem 10

By the $\pi - \lambda$ Lemma, it is enough to check this for dyadic intervals, i.e. the intervals

$$\left[\frac{k}{2^l}, \frac{k+1}{2^l}\right].$$

By 15.1, modulo a point this corresponds to the cylinder specifying k in the binary number system. This has size $1/2^l$ by 15.16, which agrees with the Lebesgue measure.

Problem 11

Just do what's in the hint :). For measurability of f, it is enough to check that $f^{-1}[C]$ is measurable for a basis cylinder $C = [\varphi]$. But

$$f^{-1}[C] = A_\varphi,$$

which is measurable. Also, the identity

$$f[\mu] = \nu$$

contains all finite cylinders, so also all clopens, so also all Borel sets, because the family

$$\left\{\mu(f^{-1}B) = \nu(B)\right\}$$

is a λ-system.

Chapter 20

Measures on Topological Spaces, Problemset 3

Problem 1

By transfinite induction, each $B \in \operatorname{Bor} Y$ is of the form $\widetilde{B} \cap Y$ for some $\widetilde{B} \in \operatorname{Bor} X$.

The axiom $\nu(\varnothing) = 0$ is immediate from $\varnothing = \varnothing \cap Y$.

For countable additivity, take a sequence of pairwise disjoint sets $B_n \in \operatorname{Bor} Y$. By an earlier observation, we may represent them as $Y \cap \widetilde{B}_n$ for a sequence $\widetilde{B}_n \in \operatorname{Bor} X$. By Problem 2, we may in fact assume

$$\mu(\widetilde{B}_n) = \mu^*(Y \cap B_n)$$

if we take $\widetilde{B}_n$ to be measurable hulls. If B_n, B_m give disjoint sets in Y, then we have

$$\widetilde{B}_n \cap \widetilde{B}_m \subseteq (\widetilde{B}_n \setminus Y) \cup (\widetilde{B}_m \setminus Y)$$

so when we pass to outer measure we see that $\mu(\widetilde{B}_n \cap \widetilde{B}_m) = 0$. Note that $B_n \cap Y = \widetilde{B}_n \cap Y$. We have

$$0 \leqslant \mu^* \left(\bigcup_{n=1}^{\infty} \widetilde{B}_n \setminus Y \right) \leqslant \sum_{n=1}^{\infty} \mu^*(\widetilde{B}_n \setminus Y) = 0,$$

so

$$\begin{aligned}
\mu^* \left(\bigcup_{n=1}^{\infty} \widetilde{B}_n \cap Y \right) &= \mu^* \left(\bigcup_{n=1}^{\infty} \widetilde{B}_n \right) \\
&= \mu \left(\bigcup_{n=1}^{\infty} \widetilde{B}_n \right) \\
&= \sum_{n=1}^{\infty} \mu \left(\widetilde{B}_n \right) \\
&= \sum_{n=1}^{\infty} \mu^* \left(\widetilde{B}_n \cap Y \right).
\end{aligned}$$

Problem 2

By definition of outer measure (as an infimum), we can choose measurable sets $H_n \supseteq Z$ such that

$$\mu^*(Z) \leqslant \mu(H_n) < \mu^*(Z) + \frac{1}{n}.$$

Take

$$H = \bigcap_{n=1}^{\infty} H_n.$$

This H contains Z, so we have

$$\mu(H) = \mu^*(Z)$$

by squeezing. By regularity of borel measures (see 16.14), we can take a G_δ upper approximation of H with the same measure.

For the second part, take two measurable hulls H_1 and H_2. Since $Z \subseteq H_1 \cap H_2$, we have

$$H_1 \Delta H_2 \subseteq (H_1 \setminus Z) \cup (H_2 \setminus Z),$$

and the RHS has (outer) measure 0, so the LHS does as well.

Observation. Hulls work well with set unions, i.e. for a countable union of Z_i with hulls H_i, the union of H_i is a hull for that union. Intersections and complements are more problematic.

Problem 3

By a *base* I understand a basis for the topology, in particular $\sigma(\mathcal{U}) = \operatorname{Bor} X$. First, we will show that $\mu = \nu$ on open sets. Take an open set U. It can be represented as

$$U = \bigcup_{n=1}^{\infty} U_n = \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{k} U_n$$

for some $U_n \in \mathcal{U}$. Since $\mathcal{U}$ is closed under finite sums, the inner sums are also in $\mathcal{U}$, so μ and ν agree on them. But since the outer sums on RHS are increasing, $\mu(U) = \nu(V)$. Applying the $\pi - \lambda$ Lemma shows that μ and ν agree on all Borel sets.

Alternatively, since these are probability measures, they also agree on closed sets (by complements), so the regularity property 16.14 does the job.

Cardinality

Since X has at least two distinct points x_1, x_2, we have at least $\mathfrak{c}$ measures, as witnessed by

$$p\delta_{x_1} + (1-p)\delta_{x_2}.$$

On the other hand, we have

Lemma 20.1. *If X is $\mathcal{SM}$, then X is second countable.*

Proof. By 16.1, $X \hookrightarrow [0, 1]^{\mathbb{N}}$ which is second countable, so X is second countable as well. □

Since the values of a probability measure μ are determined by its values on a countable basis, we know that there are at most as many measures as functions in $[0, 1]^{\mathbb{N}}$. The cardinality of the Hilbert Cube is

$$\mathfrak{c}^{\aleph_0} = \left(2^{\aleph_0}\right)^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0} = \mathfrak{c},$$

which gives the upper bound.

Problem 4

Let $\mathcal{F}$ be the family of functions which have this property. It is of course closed under uniform limits, but because of the regularity property of Borel measures (see 16.14) and Egorov's Theorem it is also closed under almost uniform limits, so also under pointwise limits.

Problem 5

Think of ω^ω as an infinite product. There is a number n such that

$$A_1 := \mu\left(\pi_0^{-1}\left[k,\, n\right]\right) > 1 - \frac{\varepsilon}{2}.$$

The intervals denote finite subsets of ω. Analogously, define

$$A_k := \mu\left(\pi_k^{-1}\left[0,\, n_k\right]\right) > 1 - \frac{\varepsilon}{2^k}.$$

By upward continuity of μ, for each k such an n_k exists. Then

$$\mu\left(\bigcap_{k=1}^{\infty} A_k\right) > 1 - \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = 1 - \varepsilon.$$

On the other hand,

$$K := \bigcap_{k=1}^{\infty} A_k = \prod_{k=1}^{\infty} [0, n_k],$$

which is a product of compact sets, so compact by Tychonoff.

Problem 6

Forward

We will first state and prove useful propeties of the Baire space ω^ω.

Lemma 20.2. *Every Polish space is a continuous image of the Baire Space ω^ω.*

An idea that may pop into your head is to pick a countable dense subset $x_n \in X$ and define

$$f : \omega^\omega \to X$$

via

$$f(s) = \lim_n x_{s_n}.$$

This is sort of the right idea, but runs into the problem that there are nonconvergent sequences, to its only a partial function. By analogy with the Cantor set, the Baire space retracts onto any of its closed subspaces. This does not save us, since $\operatorname{dom} f$ is not closed – the convergence of the sequence depends only on a tail set of indices, and the metric of ω^ω is defined via prefixes. There is a solution.

Proof. Define the set $D \subseteq \omega^\omega$ as

$$D := \left\{ s \in \omega^\omega : d(x_{s_n}, \lim_n x_{s_n}) < \frac{1}{n} \right\}.$$

This set is closed, so there is a retraction $r : \omega^\omega \to D$. The function is uniformly continuous on D, so $f \circ r$ surjects X. $\square$

We now proceed to try and surject the Souslin scheme result with the Baire Space. Assume the result is nonempty. For a sequence $\sigma \in \omega^{<\omega}$ denote by $[\sigma]$ the cylinder of sequences beginning with σ. These are clopen sets. Moreover, the sets $[n]$ for $n \in \omega$ form a partition of ω^ω into disjoint open sets. For each n, take a function

$$f_n : \omega^\omega \to F_n$$

and *merge* them using the disjoint sets S_n. That is, we put down $g_1 = f_n$ on S_n. Now, we can do this for prefixes of any arbitrary (finite) length k, giving us a function g_k. The key insight is that *surjects the Souslin scheme up to level* k. More concretely

$$g_k(\sigma) \in F_{\sigma|k}.$$

If we could take a pointwise limit $g = \lim_{k=1} g_k$, we would have $g(\sigma) \in$

$$g(\sigma) \in \bigcap_{k=1}^{\infty} F_{\sigma|k}.$$

We run into three problems

1. we may happen upon an empty set $F_{\sigma|k}$ somewhere in the scheme,
2. the limit might not exist,
3. the infinite intersection may be empty or contain more than one point, in which case g may not be a surjection.

Removing empty sets. If some set in the Souslin scheme is empty, we replace its subtree with on of its siblings. This does not change the result of the whole operation. If all siblings are empty, we travel up a level and treat the parent as though it was empty. Since the set we are trying to surject is nonempty, at some point we will be able to use a nonempty sibling. This fixes the first problem. Turns out the other two have a rather elegant solution, which I stole from Kechris.

Ensuring nonemptiness. If we have that $\operatorname{diam} F_\sigma < 1/l$, where l is the length of σ, then each intersection of $F_{\sigma|k}$ is a singleton. We can easily do this, since X can be covered by finitely many closed balls of radius $1/n$ for any n, since it is separable. Now insert a new level of the Souslin scheme tree in between two existing ones, where we take an F_σ and subdivide it into

$$F_\sigma \cap B\left(x_i, \frac{1}{|\sigma|}\right),$$

where $\{x_i\}$ is a countable dense set.

Mopping up. Take the initial Souslin scheme, insert the levels needed to have diameters tending to zero along every path and remove empty sets. Not that removing empty sets does not move any set downward in the tree, so the diameter bound is maintained. Then do the g_k construction. Because of the diameter bound, the convergence is now uniform, so we get a continuous function.

Formalisms. There are some issues with the constructions I used. They have to be done level-by-level to work, and an induction principle is needed!

Problem 9

The pushforward operator $f[-]$ is a covariant functor from the category of measurable spaces and Borel maps into an appropriate category (even the forgetful Set will suffice, though we could take sth like measure algebras). The function f has a Borel section s, so the operator $f[-]$ has a section $s[-]$, and in particular it must be surjective.

Chapter 21

Munkres

§21.1 The Countability Axioms

21.1.1 Definitions

Definition 21.1 (Basis at a point)**.** Let $x \in X$. A collection $\mathcal{B}$ of neighbourhoods of x is a **basis at** x iff every open neighbourhood $U \ni x$ contains a member of $\mathcal{B}$, i.e. there is a $B \in \mathcal{B}$ s. that $x \in B \subseteq U$.

Definition 21.2 (First countability)**.** A space X is said to be **first-countable** if it has a countable basis $\mathcal{B}_x$ at each of its points $x \in X$.

Definition 21.3 (Second countability)**.** A space X is said to be **second-countable** if it has a countable basis.

Definition 21.4 (Dense subset)**.** A subset $D \subseteq X$ is said to be **dense** if

$$\operatorname{Cl} D = X.$$

Definition 21.5 (Lindelöf property)**.** A space X is called **Lindelöf** if every open cover of X admits a countable subcover.

.

21.1.2 Lemmas

Lemma 21.6 (Picking a decreasing basis at point)**.** *Let $\mathcal{B} = \{B_1, B_2, \ldots\}$ be a basis for the topology of X at a point $x \in X$. Then the sequence*

$$B_1, B_1 \cap B_2, B_1 \cap B_2 \cap B_3, \ldots$$

is a decreasing inscribed basis at X.

Proof. All of these sets contain x. Since any neighbourhood of $U \ni x$ contains some B_k, it will also contain the set $B_1 \cap B_2 \cap \ldots \cap B_k$. □

21.1.3 Theorems

Theorem 21.7 (Continuity in first-countable spaces)**.** *If $f : X \to Y$ is continuous, then for all sequences $x_n \to x$ we have $f(x_n) \to f(x)$. If additionally X is first-countable, then then converse is also true.*

Proof. For the forward implication, pick any $U \in N_{f(x)}$. Then $V := f^{-1}[U]$ is open, so, for all but finitely many indices n, $x_n \in V$. Then for the same indices $f(x_n) \in U$.

For the backward, let B_n be a countable basis of X at x and suppose without loss of generality that B_n is decreasing. Pick again an open neighbourhood $U \in N_{f(x)}$. If $f^{-1}[U]$ is not open, then choose a sequence $x_n \in B_n \setminus f^{-1}[U]$. This means that $f(x_n) \notin U$ for all n, but $f(x_n) \to f(x)$, a contradiction. □

Theorem 21.8 (Closures in first-countable spaces)**.** *If $x_n \in A$ and $x_n \to x$, we can conclude that $x \in \operatorname{Cl} A$. The converse holds in a first-countable space.*

Proof. For the forward implication, every neighbourhood $U \ni x$ contains a point of A, and even infinitely many of them, so $x \in \operatorname{Cl} A$.

For the backward implication, suppose $x \in \operatorname{Cl} A$ and $\mathcal{B} = \{B_1, B_2, \ldots\}$ is a countable basis at x, without loss of generality decreasing. Then every set B_n contains a point of A, so we can for a sequence $x_n \in A \cap B_n$. Since $B_i \searrow$, for any neighbourhood $U \ni x$, $B_N \subseteq U$ for some N, and so U contains all x_n for $n \geqslant N$. Therefore $x_n \to x$ is a sequence of elements of A convergent to x. □

Theorem 21.9 (First- and second-countability are hereditary)**.** *Let X be second- or first-countable space. Then, if Y is subspace of X, Y is also, respectively, second- or first-countable.*

Proof. Let B_k be a basis for X. Then $B_k \cap Y$ is a basis for Y. Indeed, any open set $U \subseteq Y$ is of the form $V \cap Y$ for some open $V \subseteq X$. Then $V \supseteq B_k$ for some k, to $U = V \cap Y \supseteq B_k \cap Y$. An analogous argument holds for a basis at $y \in Y$. □

21.1.4 Problems

Problem 21.1 (P30.18 from [Mun14])**.** Let G be a first-countable topological group. Show that if G is either separable or Lindelöf, then G is also second-countable.

Solution (for separable spaces). Let $D = \{d_1, d_2, \ldots\}$ be the countable dense subset and $\mathcal{B} = \{B_1, B_2 \ldots\}$ a basis at e. We will use the dense subset to *spread* the local basis over the group. Concretely, we claim that the countable basis can be given by

$$d_m B_n,$$

i.e. left translations of the basis at e. Pick any open set $U \subseteq G$ and its element $g \in U \subseteq G$. The left-transalted set $g^{-1}U$ contains $e = g^{-1}g$, so we can pick a neighbourhood $B_n \subseteq g^{-1}U$ or, equivalently, $g \in gB_n \subseteq U$. At this point it might be true that for some d_m we will have $g \in d_m B_n \subseteq U$. However, by continuity of the group operations we can shrink B_n to B_k such that $B_k \cdot B_k \subseteq B_m$. Then, pick a $d_m \in gB_k$. We can now compute that

$$\begin{aligned} d_m &\in gB_k \\ g^{-1}d_m &\in B_k \\ e \in g^{-1}d_m B_k &\subseteq B_k B_k \subseteq B_m \\ g \in d_m B_k &\subseteq gB_m \subseteq U, \end{aligned}$$

which completes the proof. □

Solution (for Lindelöf spaces). Let $\mathcal{B} = \{B_1, B_2, \ldots\}$ be a countable basis at e. For a given B_k, consider the cover

$$\mathcal{A}_k = \{gB_k : g \in G\} .$$

By the Lindelöf property, we can pick a countable subset $G_k \subseteq G$ such that the family $\{gB_k : g \in G_k\}$is a subcover of $\mathcal{A}_{\|}$. Now do this for all $k \in \mathbb{N}$ and, for convenince, take

$$G_0 := \bigcup_{k=1}^{\infty} G_k.$$

We claim that sets of the form

$$g_0 B_k$$

for $g_0 \in G_0$ and $B_k \in \mathcal{B}$ form a countable basis. Pick any open set and its element $g \in U \subseteq G$. We repeat the solution of the previous subproblem. To be explicit, we can find an m such that $gB_m \subseteq U$. Now pick a B_k with the property that $B_k \cdot B_k \subseteq B_m$. Since the family $\{gB_k : g \in G_0\}$, we can find some g_0 s. that $g \in g_0 B_k \subseteq gB_m \subseteq U$. □

Remark. The set G_0 is actually a dense set for the topology of G.

Remark. This is one of the situations in which the relationships between countability axioms can be reversed, that is a subset of weaker axioms implies second-countability.

21.1.5 Chapter review questions

1. What is the significance of first-countability for closures?

   Answer: In a first-countable space closures are determined by sequences. That is, $x \in \operatorname{Cl} A$iff there is a sequence $x_n \in A$ s. that $x_n \to x$. In general only the backward implication holds. See here for a proof.

2. What is the significance of first-countability for continuous function?

   Answer: If X is first-countable, then a function $f : X \to Y$ is continuous iff for all sequences $x_n \to x$ we have $f(x_n) \to f(x)$. In general, only the forward implication holds. Prove this.

3. What examples does Munkres give for a first-countable space that is not second-countable?

   Answer: $\mathbb{R}^\omega$ is first-countable as it is a metric space, but is not second-countable, because it has a discrete set $\{0, 1\}^\omega$ of size $\mathfrak{c}$.

4. How large can a discrete subspace of a second-countable space be?

   Answer: At most countable. A discrete subspace can be injected into the basis, so it has a smaller size.

5. What are the relations between countability propeties?

   Answer: Second-countability implies all other properties. See here for a proof. Not even a conjuction of the three others implies second-countability, as shown by the Sorgenfrey line $\mathbb{S}$.

6. What is the significance of the Sorgenfrey line $\mathbb{S}$ for countability axioms? Which countability properties does it have?

7. What is the significance of the Sorgenfrey plane $\mathbb{S}^2$ for countability axioms? Which countability properties does it have?

   Answer: The Sorgenfrey plane is first-countable, separable, but neither second-countable nor Lindelöf. Thus it is a product of Lindelöf spaces which is not Lindelöf, establishing that Lindelöfness is not productive. The proof uses the fact that the antidiagonal is a closed discrete subspace.

8. What is the significance of the ordered square I_o^2 for countability axioms? Which countability properties does it have?

9. How can you make it easier to check the Lindelöf property?

10. Under what conditions do the relations between countability axioms reverse, i.e. what can you use to conclude that a space is second-countable from the other axioms?

    Answer: (1) metrisable spaces are all first-countable, but either the Lindelöf property or separability imply second-countability (proof) (2) a first-countable topological group satisfying either the Lindelöf property or separability is second-countable (proof).

21.1.6 Questions about property preservation

1. How does first-countability behave under taking subspaces?

   Answer: First-countability is preserved under taking subspaces. See here.

2. How does first-countability behave under products?

3. How does first-countability behave under coproducts?

   Answer: Since first-countability only concerns what happens around a point, it is preserved under arbitrary coproducts.

4. How does first-countability behave under surjections?

5. How does first-countability behave under quotients?

6. How does second-countability behave under taking subspaces? **Answer:** Second-countability is preserved under taking subspaces. See here.

7. How does second-countability behave under products?

8. How does second-countability behave under coproducts?

9. How does second-countability behave under surjections?

10. How does second-countability behave under quotients?

11. How does being Lindelöf behave under taking subspaces?

12. How does being Lindelöf behave under products?

13. How does being Lindelöf behave under coproducts?

14. How does being Lindelöf behave under surjections?

15. How does being Lindelöf behave under quotients?

    Answer: Since the Lindelöf property is surjective, it is also preserved under quotients.

16. How does (topological) separability behave under taking subspaces?

    Answer: Separability is not preserved under taking subspaces, not even closed ones. This is shown by the antidiagonal subspace Δ^- of the Sorgenfrey line.

17. How does separability behave under products?

18. How does separability behave under coproducts?

19. How does separability behave under surjections?

20. How does separability behave under quotients?

    Answer: Since separability is surjective, it is preserved under quotients.

Chapter 22

Boolean algebras

To study Boolean algebras it is best to see them in action. There are two most basic uses for Boolean algebras:

1. the algebra of truth values $\{0, 1\}$ of (classical, propositional) logic.
2. the algebraic structure of $\mathcal{P}$ with set-theoretic operations.

Definition 22.1 (Propositional formulas)**.** Fix a set Var. We call its members propositional variables. Then a **(propositional) formula** over built using the variables from Var and connectives $\wedge, \vee, \neg, \rightarrow, \leftrightarrow$.

§22.1 Axioms of Boolean algebras

We will now define a Boolean algebra in earnest as an algebraic structure. All the axioms are taken from Wikipedia.

Definition 22.2 (Axioms of Boolean algebras)**.** A Boolean algebra is an algebraic structure $\mathcal{B} = (B, 0, 1, \wedge, \vee, \neq)$ satisfying the following axioms

$a \vee (b \vee c) = (a \vee b) \vee c$	$a \wedge (b \wedge c) = (a \wedge b) \wedge c$	associativity
$a \vee b = b \vee a$	$a \wedge b = b \wedge a$	commutativity
$a \vee (a \wedge b) = a$	$a \wedge (a \vee b) = a$	absorption
$a \vee 0 = a$	$a \wedge 0 = a$	identity
$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$	$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$	distributivity
$a \vee (\neg a) = 1$	$a \wedge (\neg a) = 0$	complement

Terminology. It is a good practice to use different terms for Boolean algebras and logical terms, and even use different symbols if possible. Thus the algebraic operations are called **bottom**, **top**, **meet**, **join**, **complement** and not false, true, disjunction etc.

Theorem 22.3 (NAND axiomatisation)**.** *An algebraic structure $\mathcal{B} = (B, |)$ with the following axiom*

$$((a|b)|c)|(a|((a|c)|a)) = c$$

is a Boolean algebra, when $|$ is interpreted as the NAND operator.

Source: https://en.wikipedia.org/wiki/Minimal_axioms_for_Boolean_algebra.

Observe that for the examples we've seen so far – the truth value algebra $\mathbb{B}$ and powerset algebras $\mathcal{P}(X)$ have exactly the same identities. That is, an equation like

$$x \vee z \wedge x \wedge z = x$$

holds in all Boolean algebras iff it holds in the Boolean algebra $\mathbb{B}$. One direction is trivial since $\mathbb{B} = \mathcal{P}(\{x_0\})$. The other direction is not much difficult, since we can check identities on sets element-wise. Of course, this does not mean that one specific algebra can't have its own identities – we're talking about *laws of Boolean algebra.* A natural question to ask is whether this extends to all abstract Boolean algebras. And in fact it does! That is, we have the following theorem

Theorem 22.4. *Let $f(\overline{x}), g(\overline{x})$ be two terms of Boolean connectives. Then the sentence*

$$\forall \overline{x}. f(\overline{x}) = g(\overline{x})$$

holds in all Boolean algebras iff it holds in the truth value algebra $\mathbb{B}$.

We will prove this theorem in logical terms. First, we restate it.

Observation. In a Boolean algebra $\mathcal{B}$

$$b_1 = b_2$$

is equivalent to

$$b_1 \leftrightarrow b_2 = 1.$$

Definition 22.5. We write $\mathcal{B}, v \models \varphi$ if

$$[[\varphi]]_v = 1$$

and $\mathcal{B} \models \varphi$ if $\mathcal{B}, v \models \varphi$ regardless of the valuation v.

So, to prove our theorem we actually want to prove that $\mathbb{B} \models (f \leftrightarrow g)$ iff for all algebras $\mathcal{B} \models (f \leftrightarrow g)$. It turns out that this is true for an arbitrary formula! For a proof, see the section on filters in Boolean algebras.

§22.2 Filters in Boolean algebras

Consider the Boolean algebra $\mathcal{P}(\mathbb{N})$. Its members are sets of various cardinality, and there is a very sharp divide between two kinds of cardinalities – finite and infinite. However, an inifite set may not be *that large*. Consider the set of powers of 2. One might be tempted to say that although it is infinite, it is quite small for an infinite sets, as it has a lot of gaps.

Within the sets of infinite cardinality there is a collection of especially large sets – those that contain all but finitely many numbers. Since this means they have finite complement, we call them **cofinite**. Observe that for any two cofinite sets A, B we have

$$(A \cap B)^c = A^c \cup B^c,$$

so $A \cap B$ is also cofinite, which is not at all true if we replace the word *cofinite* with *infinite*. This motivates the following definition.

Definition 22.6 (Filter)**.** Fix a Boolean algebra $\mathcal{B}$. A collection $F \subseteq \mathcal{B}$ of subsets of X is called a **filter** if the two following conditions are met

(1) if $a \in F$ and $a \leqslant b$, then $b \in F$,

(2) if $a, b \in F$ then $a \wedge b \in F$.

In the case $\mathcal{B} = \mathcal{P}(X)$ we refer to F as a *filter on* X.

This definition is actually satisfied by $F = \mathcal{P}(X)$ in the powerset algebra $\mathcal{P}(X)$, so our notion of *family of very large sets* also contains the empty set, which is not ideal.

Definition 22.7 (Proper filter)**.** Fix a Boolean algebra $\mathcal{B}$. A filter F is **proper** if $0 \notin F$. Equivalently, $F \neq \mathcal{B}$.

Example 22.8 (Fréchet filter)**.** Fix an infinite set X. Then the family $\acute{F}$ of all cofinite subsets of X is a proper filter in $\mathcal{P}(X)$, this has already been verified in the text at the beginning of the section. We call this filter the *Fréchet filter* on X.

Example 22.9 (Conull filter)**.** Fix a complete measure space (X, Σ, μ). The family of all conull members of Σ forms a filter on X.

There is also a second intuition behind the idea of a filter, coming more directly from logic. One can think of a liter as a *state of knowledge*, i.e. all the facts derivable from a set of base facts. The following example is one of the ways of making this notion precise. (I learned this intuition from section 2.5 of [SU06])

Example 22.10 (Filters of propositional formulas)**.** The set of propositional formulas on a set Var, quotiented by equivalence of formulas, is a Boolean algebra with operations given by logical connectives. For a formula φ we can define a filter

$$F_\varphi := \{\psi \mid \varphi \to \psi \text{ is a tautology}\}.$$

This filter is proper iff φ is satisfiable. In fact, for any (possible infinte set) Γ of propositional formulas we can define a filter

$$F_\Gamma := \{\psi \mid \varphi_1 \wedge \varphi_2 \ldots \wedge \varphi_n \to \psi \text{ is a tautology for some } \varphi_i \in \Gamma\}.$$

This, too, is a proper filter iff φ is satisfiable, but the proof of this requires compactness of propositional logic.

Lemma 22.11 (Filter extension)**.** *Suppose that $a, \neg a \notin F$. Then there exists a filter $F_a \supseteq F$ extending F.*

Proof. TODO. □

Definition 22.12 (Ultrafilter)**.** A proper filter F is an **ultrafilter** if it satisfies any of the following equivalent conditions:

1. $a \in F$ or $\neg a \in F$,

2. either $a \in F$ or $\neg a \in F$,

3. if $a \vee b \in F$ then $a \in F$ or $b \in F$,

4. $a \vee b \in F$ iff $a \in F$ or $b \in F$.

What about meets? Any filter satisfies $a \wedge b \in F$ iff both $a \in F$ and $b \in F$.

Proof of equivalence. A proper filter cannot contain 0, which establishes (1) $\Leftrightarrow$ (2). Order properties establish (3) $\Leftrightarrow$ (4).

For the downward implication, see that since $a \vee b \in F$ then $F \not\ni \neg(a \vee b) = \neg a \wedge \neg b$, so by the observation above $\neg a \notin F$ or $\neg b \notin F$, which implies $a \in F$ or $b \in F$.

For the upward implication, $F \ni 1 = a \vee \neg a$, so at least one of $a, \neg a$ is in F. $\square$

Terminology. By the extension lemma 22.11, a filter satisfying (1) or (2) is a maximal proper filter. The term *prime filter* is best ascribed to a filter satisfying either of the conditions (3) or (4). The reason we introduce these different words is twofold: firstly, the notions are not equivalent in slightly different algebras (e.g. the Heyting algebras of intuitionistic propositional logic). Secondly, we want to tease the relationship with ring theory.

Problem 22.1 (P2.13 from [SU06])**.** Suppose $\mathcal{B}, v \not\models \varphi$. Then there exists a prime filter $F \subseteq B$ with $[[\neg\varphi]]_v \in F$. Secondly, the binary valuation given by $w(p) = 1$ iff $v(p) \in F$ satisfies $[[\varphi]]_w = 0$.

Solution. Let $f = [[\varphi]]_v$. Since $\mathcal{B}, v \not\models \varphi$ we know that $f \neq 1$, so $\neg f \neq 0$, so the principal filter $F_{\neg f}$ is proper, therefore it can be extended to a prime filter F.

Taking the valuation w from the statement, we turn to computing $[[\varphi]]_w$. By induction on subformulas, $[[\psi]]_w$ is either 0 or 1 for any propositional formula ψ (we can also just note that $\{0, 1\}$ is a subalgebra of $\mathcal{B}$ and $\mathcal{B}$ is nontrivial since $\mathcal{B}, v \not\models \varphi$).

Also by structural induction on ψ we prove that $[[\psi]]_w = 1$ iff $[[\psi]]_v \in F$. For variables this follows from the definition. For negation, this follows from the filter F being prime, so also maximal – F contains either $[[\psi]]_v$ or $[[\neg\psi]]_v = \neg [[\psi]]_v$. The inductive step for disjunction follows from property (4) is 22.12, and for conjunction from the remark after 22.12. Finally, since $\neg f \in F$, we have that $[[\neg\varphi]] \in F$, so $[[\neg\varphi]]_w = 1$ and $[[\varphi]]_w = 0$. $\square$

Problem 22.2 (P2.14 from [SU06])**.** Let $\mathcal{B}_0$ denote any nondegenerate Boolean algebra, $\mathbb{B}$ be the algebra of truth values and φ be a propositional formula. TFAE:

(i) $\mathbb{B} \models \varphi$,

(ii) $\mathcal{B}_0 \models \varphi$,

(iii) $\mathcal{B} \models \varphi$ for all Boolean algebras $\mathcal{B}$.

Solution. The lower upward implication is trivial. The second upward implication follows from that fact that $\mathbb{B}$ embeds into $\mathcal{B}_0$ as a subalgebra. The interesting implication is $(i) \Rightarrow (iii)$. It follows immediately from Stone's Theorem **??**, since equalities of sets can be checked element-wise.

For a proof without Stone's Theorem, use the previous problem. That is, suppose $\mathcal{B} \not\models \varphi$ for some Boolean algebra φ, so $\mathcal{B}, v \not\models \varphi$ for some valuation $v : \mathsf{Var} \to \mathcal{B}$. By the previous problem there exists a binary valuation $w : \mathsf{Var} \to \{0_\mathcal{B}, 1_\mathcal{B}\}$ which gives $[[\varphi]]_w = 0_\mathcal{B}$. Since the subalgebra $\{0_\mathcal{B}, 1_\mathcal{B}\}$ is isomorphic to $\mathbb{B}$, we get $\mathbb{B}, \widetilde{w} \not\models \varphi$, so $\mathbb{B} \not\models \varphi$. $\square$

Chapter 23

Algebraic semantics of intuitionistic logic

This chapter will introduce intuitionistic logic and two ways to give it semantics: Heyting algebras and Kripke models.

Problem 23.1 (Problem 2.1. from [SU06])**.** At least one of the numbers $e+\pi, e\pi$ is transcendental.

Solution. If both these numbers were algebraic, then so would be the two roots of the polynomial

$$X^2 - (e + \pi)X + e\pi = (X - e)(X - \pi),$$

which would imply that both e and π are algebraic numbers. It is, however, well-known that they are transcendental. $\square$

Definition 23.1 (Lattice as a poset)**.** A partially ordered set $(A, \leqslant)$ is called a **lattice** if for every pair $a, b \in A$ the bounds $\sup\{a, b\}$ and $\inf\{a, b\}$ can be defined. These are then called **join** and **meet**, respectively, and are denoted $a \sqcup b$ and $a \sqcap b$.

Definition 23.2 (Lattice, algebraic)**.** A **lattice** is an algebraic structure $(A, \sqcap, \sqcup)$ satisfying the following four pairs of dual axioms:

$a \sqcup (b \sqcup c) = (a \sqcup b) \sqcup c$	$a \sqcap (b \sqcap c) = (a \sqcap b) \sqcap c$	associativity
$a \sqcup b = b \sqcup a$	$a \sqcap b = b \sqcap a$	commutativity
$a \sqcup (a \sqcap b) = a$	$a \sqcap (a \sqcup b) = a$	absorption
$a \sqcup a = a$	$a \sqcap a = a$	idempotency

Problem 23.2 (Problem 2.6. from [SU06])**.** Prove that a lattice defined algebraically as in 23.2 can be given a poset structure with the ordering given by

$$a \leqslant b \Leftrightarrow a \sqcup b = b.$$

Show that this is a lattice in the sense of 23.1 and that joins and meets in this lattice are given by $\sqcup, \sqcap$.

Solution. First, we check the conditions for order. The relation is reflexive since $a \sqcup a = a$. For transitivity, suppose $a \leqslant b$ and $b \leqslant c$, so that we have

$$a \sqcup b = b,$$
$$b \sqcup c = c.$$

Then we use the associativity axiom to compute

$$a \sqcup c = a \sqcup (b \sqcup c) = (a \sqcup b) \sqcup c = b \sqcup c = c,$$

so $a \leqslant c$. Now we prove that $\leqslant$ is antisymmetric. Suppose $a \leqslant b$ and $b \leqslant a$. By commutativity we have

$$a = b \sqcup a = a \sqcup b = b.$$

Let us now prove that $a \sqcup b = \sup\{a, b\}$. By idempotency and associativity we have

$$a \sqcup (a \sqcup b) = (a \sqcup a) \sqcup b = a \sqcup b,$$

so $a \sqcup b \geqslant a$, and of course the same for b, so $a \sqcup b$ is an upper bound of $\{a, b\}$. Take any other upper bound $c \geqslant \{a, b\}$. Then by associativity

$$(a \sqcup b) \sqcup c = a \sqcup (b \sqcup c) = a \sqcup c = c,$$

so $a \sqcup b \leqslant c$ and $a \sqcup b$ is the least upper bound. For the infimum use Lemma 23.3 and retread the proof for sup. □

Remark. A dual (with $\sqcup$ and $\sqcap$ switched) proof would work if we defined the order as

$$a \leqslant b \Leftrightarrow a \sqcap b = a.$$

In face, we have the following.

Lemma 23.3 (Dual definitions of order in algebraic lattices)**.** *In an algebraic lattice* $(A, \sqcap, \sqcup)$ *the conditions*

$$a \sqcup b = b$$

and

$$a \sqcap b = a$$

are equivalent, and both define an order relation.

Proof. The *order* part of the statement is the previous problem. For equivalency, suppose now that $a \sqcap b = a$. Then

$$a \sqcup b = (a \sqcap b) \sqcup b = b$$

by the absorption axiom. The upward implication is dual. □

Lemma 23.4 (Monotnicity proerties of lattices)**.** *The lattice operations* $\sqcap, \sqcup$ *are increasing in each of their arguments.*

Proof. TODO. □

Chapter 24

Random things

Theorem 24.1 (Duality law for Boolean algebra)**.** *Every identity of boolean algebra is has a dual*
`Ankified` *identity, i.e. one obtained by interchanging* 0 *with* 1 *and* $\wedge$ *with* $\vee$*.*

Proof. Negation! □

Appendix A

Some values of Tor and Ext

$\mathrm{Tor}_{\mathbb{Z}}$	$\mathbb{Z}$	$\mathbb{Z}_m$	G
$\mathbb{Z}$	0	0	0
$\mathbb{Z}_n$	0	$\mathbb{Z}_{\gcd(m,n)}$	$\ker(G \xrightarrow{\cdot n} G)$

Bibliography

[Mun14] James Munkres. *Topology.* Pearson Education Ltd, 2014.

[SU06] Morten Heine Sørensen and Pawel Urzyczyn. *Lectures on the Curry-Howard isomorphism.* Elsevier, Amsterdam; Oxford, 2006.