Descriptive Set Theory

January 10, 2025

Contents

0.1 Introduction: Perfect Subsets of the Real Line

Descriptive set theory nowadays is understood as the study of definable subsets of Polish Spaces. Many of its problems and techniques arose out of efforts to answer basic questions about the real numbers. A prominent example is the *Continuum Hypothesis* (CH):

> **Continuum Hypothesis (Cantor, 1890s)**
> If $A \subseteq \mathbb{R}$ is uncountable, then there exists a bijection between A and $\mathbb{R}$. That is, is every uncountable subset of $\mathbb{R}$ is of the same cardinality as $\mathbb{R}$.

Early approaches tried to show that CH holds for a number of sets with an easy topological structure.

For closed sets, the situation is less clear. Given a set $A \subseteq \mathbb{R}$, we call $x \in \mathbb{R}$ a **limit point** of A if

$$\forall \epsilon > 0\, \exists z \in A\, [z \neq x \,\&\, z \in U_\varepsilon(x)], \tag{1}$$

where $U_\varepsilon(x)$ denotes the standard ε-neighborhood of x in $\mathbb{R}$

Definition 0.1.1. *A non-empty set $P \subseteq \mathbb{R}$ is* ***perfect*** *if it is closed and every point of P is a limit point.*

In other words, a perfect set is a closed set that has no isolated points. We can also deduce that for a perfect set P, every neighborhood of a point $p \in P$ contains infinitely many points from P.

Cantor set

Obviously, $\mathbb{R}$ itself is perfect, as is any closed interval in $\mathbb{R}$. There are totally disconnected perfect sets, such as the **middle-third Cantor set** in $[0, 1]$

> **Theorem (Cantor, 1884)**
> A perfect subset of $\mathbb{R}$ has the same cardinality as $\mathbb{R}$.

> **Hint**
>
> - Argue it suffices to construct an injection from $2^{\mathbb{N}}$ (the set of all infinite binary sequences) into the perfect set.
> - Start with any point x in the perfect set and open neighborhood $U_\varepsilon(x)$. Use the perfect set property to find two points x_0, x_1 distinct from x and each other in $U_\varepsilon(x)$.
> - These points will 'guide' the mapping: All sequences in $2^{\mathbb{N}}$ starting with 0 will be mapped to a point close to x_0, while all sequences starting with 1 will be a mapped to a point close to x_1.
> - Now iterate with x_0 and x_1 in place of x.

Proof. Let $P \subseteq \mathbb{R}$ be perfect. We construct an injection from the set $2^{\mathbb{N}}$ of all infinite binary sequences into P. An infinite binary sequence $\xi = \xi_0\xi_1\xi_2\ldots$ can be identified with a real number $\in [0, 1]$ via the mapping

$$\xi \mapsto \sum_{i \geq 0} \xi_i 2^{-i-1}. \tag{2}$$

Note that this mapping is onto. It follows that the cardinality of P is at least as large as the cardinality of $[0,1]$. The Schröder-Bernstein Theorem (for a proof see e.g. **?**) implies that $|P| = |\mathbb{R}|$.

To construct the desired injection, choose $x \in P$ and let $\varepsilon_0 = 1 = 2^0$. Since P is perfect, $P \cap U_{\varepsilon_0}(x)$ is infinite. Let $x_0 \neq x_1$ be two points in $P \cap U_{\varepsilon_0}(x)$, distinct from x. Let ε_1 be such that $\varepsilon_1 \leq 1/2$, $U_{\varepsilon_1}(x_0), U_{\varepsilon_1}(x_1) \subseteq U_{\varepsilon_0}(x)$, and $\overline{U_{\varepsilon_1}(x_0)} \cap \overline{U_{\varepsilon_1}(x_1)} = \emptyset$, where $\overline{U}$ denotes the closure of U.

We can iterate this procedure recursively with smaller and smaller diameters, using the fact that P is perfect. This gives rise to a so-called **Cantor scheme**, a family of open balls (U_σ) satisfying certain *nesting conditions*. Here the index σ is a finite binary sequence, also called a *string*. A Cantor scheme is defined by the following properties.

1. $\operatorname{diam}(U_\sigma) \leq 2^{-|\sigma|}$, where $|\sigma|$ denotes the length of σ.

2. If τ is a proper extension of σ, then $U_\tau \subset U_\sigma$.

3. If τ and σ are *incompatible* (i.e. neither extends the other), then

$$U_\tau \cap U_\sigma = \emptyset.$$

4. The center of each U_σ, call it x_σ, is in P.

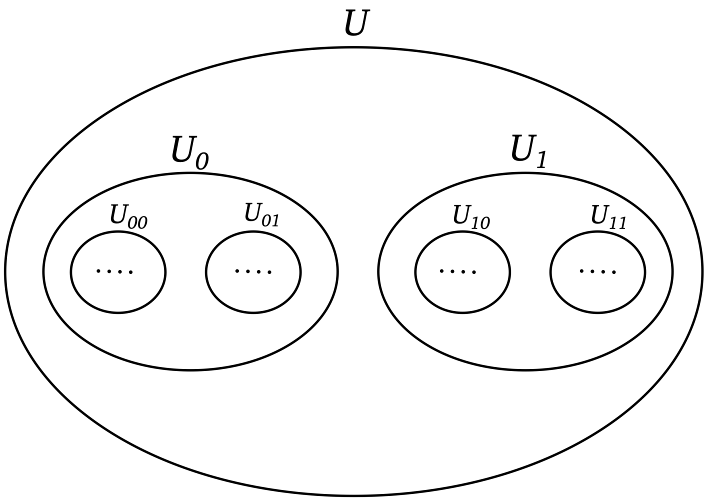

Figure 1: Nested structure of a Cantor scheme

Let ξ be an infinite binary sequence. Given $n \geq 0$, we denote by $\xi\restriction_n$ the string formed by the first n bits of ξ, i.e.

$$\xi\restriction_n = \xi_0\xi_1 \ldots \xi_{n-1}. \tag{3}$$

The finite initial segments give rise to a sequence $x_{\xi\restriction_n}$ of centers. By properties (1.) and (2.), this is a Cauchy sequence. By (4.), the sequence lies in P. Since P is closed, the limit x_ξ is in P. By (3.), the mapping $\xi \mapsto x_\xi$ is well-defined and injective.

$\square$

Thus, to show that a set of reals has the same cardinality as $\mathbb{R}$, it suffices to show the set contains a perfect subset. The next theorem establishes that the Continuum Hypothesis holds for all closed subsets of $\mathbb{R}$.

Cantor-Bendixson Theorem
Every uncountable closed subset of $\mathbb{R}$ contains a perfect subset.

Hint
Consider the set of **condensation points**, i.e. the set of all points for which any open neighborhood has *uncountable* intersection with the given closed set.

Proof. Let $C \subseteq \mathbb{R}$ be uncountable and closed. We say $z \in \mathbb{R}$ is a *condensation point* of C if

$$\forall \varepsilon > 0 \, [U_\varepsilon(z) \cap C \text{ uncountable}]. \tag{4}$$

Let D be the set of all condensation points of C. Note that $D \subseteq C$, since every condensation point is clearly a limit point and C is closed.

Furthermore, we claim that D is perfect. Clearly D is closed. Suppose $z \in D$ and $\varepsilon > 0$. Then $U_\varepsilon(z) \cap C$ is uncountable. We would like to conclude that $U_\varepsilon(z) \cap D$ is uncountable, too, since this would mean in particular that $U_\varepsilon(z) \cap D$ is infinite. The conclusion holds if $C \setminus D$ is countable.

To show that $C \setminus D$ is countable, assume that $y \in C \setminus D$. Then, for some $\delta > 0$, $U_\delta(y) \cap C$ is countable. We can find and interval $I(y) \subseteq U_\delta(y)$ that contains y and has rational endpoints. There are at most countably many intervals with rational endpoints and hence for each $y \in C \setminus D$ there are at most countably many choices for $I(y)$. Thus, we have

$$C \setminus D \subseteq \bigcup_{y \in C \setminus D} I_y \cap C. \tag{5}$$

The right hand side is a countable union of countable sets, hence countable.

$\square$

We will later encounter an alternative (more constructive) proof that gives additional information about the complexity of the closed set C. For now we conclude with the fact we were aiming to prove in this lecture.

Corollary 0.1.0.1. *Every closed subset of $\mathbb{R}$ is either countable or of the cardinality of the continuum.*

The results of this lecture give us a blueprint on how to verify the Continuum Hypothesis for a given family $\mathcal{F}$ of sets (of reals):

A family $\mathcal{F}$ of sets (of reals) has the **perfect set property** if every set in $\mathcal{F}$ is either countable or has a perfect subset.

Question
Which families of sets have the perfect set property?

0.2 Ordinals and cardinals

0.2.1 Ordinals

It will be important for us to extend the usual counting process beyond the natural numbers. To give an example, let us return for a moment to perfect subsets of the reals. To show that every uncountable closed subset of $\mathbb{R}$ contains a perfect subset, we considered the *condensation points* of the set. There is another, more gradual way, to arrive at a perfect subset. When Cantor studied convergence of Fourier series, he introduced the **derivative** of a set:

A' = { x $\in A$: x is a limit point of $A\}$(6) We can iterate the derivative and consider $A', A'', A''', \dots$. This yields a descending sequence of sets

A$' \supseteq A'' \supseteq A''' \supseteq \dots \supseteq A^{(n)} \supseteq \dots$ (7)

As the sequence is nested, we can take a "limit":

A$^{\infty} = \bigcap_n A^{(n)}$(8) But the process does not necessarily stop here. A^{∞} may have isolated points again, so that $A^{\infty} \supsetneq (A^{\infty})'$.

Let us introduce ω as a new number to be used in place of ∞ above. We can continue the counting process:

1,2,3, $\dots, \omega, \omega+1, \omega+2, \dots, \omega+\omega, \omega+\omega+1, \dots, \omega+\omega+\omega, \dots, \omega\cdot\omega, \dots$ (9)

We can then define, for example, $A^{\omega+1} := (A^{\omega})'$. As intuitively clear from above, the new transfinite numbers come with a natural ordering, so we can also put $A^{\omega+\omega} := \bigcap_{\alpha<\omega+\omega} A^{\alpha}$

Another way to count into the transfinite is to reorder the natural numbers and first enumerate all powers of two, followed by all powers of three and so on:

$$1, 2, 4, 8, \dots 3, 9, 27, \dots 5, 25, 125, \dots$$

This still leaves an infinite reservoir of numbers like $0, 6, 10, \dots$.

A number of questions arises:

- Can this process continued indefinitely?
- Is there a *unifying* principle behind the various ways to count into the transfinite?
- Can we define operations like $+$ and $\cdot$ on these infinite numbers independent of the way we represent these numbers, and without leading to contradictions?

These questions can be addressed by developing the theory of *ordinal numbers* via the concept of a *well-founded order.*

Orders and well-orders

Definition 0.2.1. *A (**reflexive** or **non-strict**) **partial order** on a set A is a binary relation $\leq$ on A such that for all $a, b, c \in A$,*

- $a \leq a$ *(reflexive)*
- $a \leq b \ \wedge\ b \leq a \quad\Rightarrow\quad a = b$ *(anti-symmetric)*
- $a \leq b \ \wedge\ b \leq c \quad\Rightarrow\quad a \leq c$ *(transitive).*

*A **linear** (or **total**) **order** additionally satisfies for all $a, b \in A$,*

- $a \leq b \ \vee\ b \leq a$ *(connected).*

With any reflexive partial order $\leq$ we can associate an *irreflexive* one by letting

$$a < b \quad :\Longleftrightarrow \quad a \leq b \ \wedge\ a \neq b. \tag{10}$$

Likewise, we can obtain a reflexive order from an irreflexive one by defining

$$a \leq b \quad :\Longleftrightarrow \quad a < b \ \vee\ a = b. \tag{11}$$

In light of this, we will usually just speak of partial or linear orders, without further specifying whether it is reflexive or irreflexive.

Example 0.2.1. • *The usual orders on $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are linear orders.*

- *The relation*

$$f \leq g \quad :\Longleftrightarrow \quad \forall x \in \mathbb{R} \;\; f(x) \leq g(x)(12)$$

  is a partial order on real valued functions on $\mathbb{R}$ but not a linear order.
- *The subset relation $\subseteq$ is a partial order on the power set of any set A, but it is a linear order only when $A = \emptyset$.*

We can enumerate the natural numbers one after another, but for the other standard ordered number domains this is not possible: We cannot find a place to begin counting (as in the case of $\mathbb{Z}$) or there is no "next bigger" element (as in the case of $\mathbb{Q}$ or $\mathbb{R}$).

To enumerate these domains in the form $\{a_0, a_1, \dots, a_n, a_{n+1}, \dots\}$ we have to reorder them in a way that

- we can start with a smallest element,

and once we have a arrived at an element a,

- we know with which element we continue the enumeration (i.e. there is an *immediate successor* to a),
- we can continue the enumeration even if we have already enumerated infinitely many elements before a (but elements of the domain still remain).

These requirements can be combined into a single property: every non-empty subset (i.e. the elements not enumerated yet) has a least element (to be enumerated next).

Definition 0.2.2. *A linear order $(A, <)$ is a **well-order** if*

$$\forall Z \; (\emptyset \neq Z \subseteq A \;\Rightarrow\; \exists x \in Z \; \forall y \in Z \; x \leq y)(13)$$

Orders themselves can be compared using *embeddings*.

Definition 0.2.3. *An **embedding** of a partial order $(A, <_A)$ into another partial order $(B, <_B)$ is a mapping $f : A \to B$ such that for all $x, y \in A$*

$$x <_A y \quad \Longleftrightarrow \quad f(x) <_B f(y).(14)$$

*Two orders are **isomorphic** if there exists a bijective embedding of one into the other.*

Of course every order is isomorphic to itself (*automorphic*) via the identity. But many orders allow automorphisms other than the identity (e.g. $\mathbb{Z}$ with $z \mapsto z + 1$), or are isomorphic to a proper subset (for example, $\mathbb{R}$ and $(0, 1)$). As we will see, well-orders are very rigid in this regard.

We start with a simple observation.

Proposition 0.2.1. *Let $(A, <)$ be a well-order and assume $f : A \to A$ is a self-embedding. Then for all $x \in A$, $x \leq f(x)$.*

Proof. If the set $\{x \in A \colon f(x) < x\}$ is non-empty, it has a minimal element z. But since f is increasing, this would imply $f(f(z)) < f(z)$, contradicting the minimality of z.

□

We immediately obtain

Corollary 0.2.0.1. *The only automorphism of a well-order is the identity.*

An **initial segment** of an order $(A, <)$ is given by all elements that are smaller than a given element b. We denote this initial segment by $A\,|_b$.

Corollary 0.2.0.2. *No well-order is isomorphic to an initial segment of itself.*

Proof. Suppose $f : A \to A\,|_b$ is an isomorphism. Then $\operatorname{ran}(f) = A\,|_b$ and $f(x) < b$ for all $x \in A$. In particular, $f(b) < b$, contradicting Lemma

□

Ordinal numbers

Cantor defined ordinal numbers (or ordinals) as **isomorphism classes of well-orders**. Later, *von Neumann* suggested a system of representatives particularly suitable for set theoretic considerations. The idea is to define the order $<$ through the $\in$-relation on a set. For example,

$\{ \emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$(15) represents a 3-element well-order.

We will impose some conditions on the sets we allow as ordinals. Given a set A, we use $\in_A$ to denote the $\in$-relation *on* A:

$\in_A = \{(x, y) \colon x, y \in A \ \wedge \ x \in y\}$(16)

Definition 0.2.4. *A set A is **transitive** if*

$$\forall x \in A \ \ x \subseteq A \tag{trans}$$

In other words, transitive sets cannot "hide" elements in subsets.

Definition 0.2.5. *A set A is an **ordinal** if it is transitive and well-ordered by $\in_A$*

It is customary to use *lower case Greek letters* $\alpha, \beta, \gamma, \ldots$ to denote ordinals.

If we exclude certain pathological sets from the beginning, we can further simplify this definition.

Definition 0.2.6. *A set A is **well-founded** if every non-empty subset has a $\in$-minimal element:*

$$\forall B \subseteq A \, (B \neq \emptyset \ \Rightarrow \ \exists y \in B \, \forall z \in B \ \ z \notin y) \tag{17}$$

Sets which contain themselves ($A \in A$) are not well-founded− $\{A\}$ would be a subset without a $\in$-minimal element. Similarly, well-founded sets cannot have cycles like $a \in b \in c \in a$.

Proposition 0.2.2. *Assume every set is well-founded. A set α is an **ordinal** if and only if it is transitive and linearly ordered by $\in_\alpha$.*

If we write out the formulas in full, we see Definition 0.2.5 is much simpler than the original one. Most notably, in Definition 0.2.5 we only use only **bounded quantifiers** (of the form $\forall y \in a$), whereas in the original form we have to quantify over arbitrary subsets of a. This is an important difference whose impact will become clear later on.

We can now develop the theory of ordinals based on this definition.

Proposition 0.2.3. *Any element of an ordinal is an ordinal.*

Proof. Any subset of a linear order is again a linear order under the induced order relation. It remains to show that $(b, \in_b)$ is transitive. Let α be an ordinal, and assume $b \in \alpha$. Let $x \in c \in b$. We claim $x \in b$. Since α is transitive, $b \subseteq \alpha$ and hence $c \in \alpha$. By transitivity of α again, $c \in \alpha$. Thus $x, b \in \alpha$, and since $\in_\alpha$ linearly orders α, we must have

x $\in b \ \vee \ x = b \ \vee \ b \in x$.(18) If $x = b$, we get $b \in c \in b$, contradicting well-foundedness. Similar for $b \in x$. Therefore, $x \in b$.

∎

The well-ordering of ordinals

Proposition 0.2.3 suggests we can order ordinals by letting

$\alpha < \beta \ :\iff \ \alpha \in \beta.$(19) By Proposition 0.2.3, an ordinal then contains precisely the ordinals smaller than it:

$\alpha = \{\beta : \beta < \alpha\}.$(20)

$\in$ defines a partial order on all ordinals: As all sets are well-founded, irreflexivity holds, and since ordinals are transitive sets, $<$ is a transitive relation.

Proposition 0.2.4. *For any ordinals α, β,*

$\alpha < \beta \iff \alpha \subset \beta.$(21)

Proof. The $\Rightarrow$ direction follows directly from the transitivity of ordinals. For $\Leftarrow$, we show something more general, namely that any transitive proper subset of an ordinal is itself an ordinal and is an element of the superset ordinal:

$\mathrm{trans}(a) \ \wedge \ a \subset \beta \quad \Rightarrow \quad \mathrm{Ord}(a) \ \wedge \ a \in \beta.$(22) If $a \subset \beta$, a is linearly ordered by $\in$ (as a subset of β). Further, if a is transitive, a is an ordinal.

It remains to show $a \in \beta$. Since a is a proper subset of β, by well-foundedness there exists a $\in$-minimal element of $\gamma \in \beta \setminus a$. We claim $a = \gamma$. By $\in$-minimality of γ, every element of γ cannot be in $\beta \setminus a$ and therefore has to be in a. Hence $\gamma \subseteq a$. On the other hand, if $x \in a$, then, by assumption $x \in \beta$, and since $\in$ linearly orders β,

$\mathrm{x} \in \gamma \ \vee \ x = \gamma \ \vee \ \gamma \in x.$(23) The latter two are impossible due to $\gamma \notin a$. Hence $x \in \gamma$ and therefore $a \subseteq \gamma$, yielding $a = \gamma$.

□

Theorem (well-ordering of ordinals)
The ordinal numbers are well-ordered by $<$.

Hint
Most properties follow directly from well-foundedness and the fact that ordinals are transitive as sets.

To show that ordinals are linearly ordered by $<$, look at the intersection of two ordinals and try to apply Proposition 0.2.4.

Proof. We first show $<$ is a linear order. Irreflexivity follows from well-foundedness of $\in$. Transitivity of $<$ follow from the transitivity of ordinals as sets. To show

$\alpha < \beta \ \vee \ \alpha = \beta \ \vee \ \beta < \alpha,$(24) observe that the intersection of two ordinals is an ordinal, the *minimum* of the two ordinals. Let $\gamma = \alpha \cap \beta$. Then $\gamma \subseteq \alpha$, so by Proposition 0.2.4, $\gamma \in \alpha$ or $\gamma = \alpha$ and similarly $\gamma \in \beta$ or $\gamma = \beta$. But in the case $\gamma \in \alpha, \gamma \in \beta$ we would have $\gamma \in \alpha \cap \beta = \gamma$, contradicting well-foundedness.

Finally, if A is a non-empty set of ordinals, the well-ordering condition on $<$ spells out as

$\exists \alpha \in A \forall \beta \in A \ \beta \notin \alpha.$(25) But this holds since we assume all sets are well-founded.

□

Basic properties of ordinals

Using the results obtained so far. we can now deduce some basic facts about the structure of ordinals:

- $0 = \emptyset$ is the *smallest ordinal.*
- Every ordinal α has an **immediate successor** under the ordering $<$:

  $$\alpha' = \alpha + 1 = \alpha \cup \{\alpha\}. \tag{26}$$

  Clearly $\alpha < \alpha + 1$. If $\alpha < \beta$, then by Proposition 0.2.4, $\alpha \subset \beta$ and $\alpha \in \beta$. Hence $\alpha + 1 \subseteq \beta$ and therefore $\alpha + 1 \leq \beta$.
- The *finite ordinals* are exactly the **natural numbers** ("set theoretic version"):

  $$0 = \emptyset, \quad 1 = 0 + 1 = \emptyset \cup \{\emptyset\} = \{\emptyset\}, \quad 2 = 1 + 1 = \{\emptyset, \{\emptyset\}\}, \dots \tag{27}$$
- The set of all natural numbers is transitive and well-ordered by $\in$ and thus itself an ordinal, the first infinite ordinal ω:

  $\omega = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \dots\}$(28)
- ω is also the first instance of a *limit ordinal*: A **successor ordinal** is any ordinal of the form $\alpha + 1$. Any ordinal λ that is not a successor is called a **limit ordinal**. Being limit is equivalent to the following property:

  $\lambda \neq 0 \ \wedge \ \forall \alpha < \lambda \ (\alpha + 1 < \lambda)$.(29) This shows immediately that ω is limit.
- More generally, if A is a set of ordinals, $\sup A = \bigcup_{\alpha \in A} \alpha$ is an ordinal and is the least upper bound for A.
- The first limit ordinal ω is followed by a number of successor ordinals as well as their limits as limit ordinals:

  $$\omega, \omega + 1, \omega + 2, \dots \omega + \omega, \omega + \omega + 1, \omega + \omega + 2, \dots \omega + \omega + \omega,$$
  $$\omega + \omega + \omega + 1, \omega + \omega + \omega + 2, \dots \omega \cdot \omega, \omega \cdot \omega + 1, \dots, \omega^{\omega} \dots \omega^{\omega^{\omega}} \dots$$

All of the ordinals listed here are still countable (as sets). The supremum of the set of all countable ordinals is denoted by ω_1, the **first uncountable ordinal**. After ω_1, we have again successors, limits, and so on.

Metamathematical issues

Is there a *set* Ord of *all* ordinals? If so, it would be well-ordered by $\in$ and also transitive (since, by Proposition 0.2.3, every element of an ordinal is an ordinal) and therefore an ordinal. But then Ord +1 would be an ordinal not contained in Ord (by well-foundedness), contradiction.

This is know as the **Anomaly of Burali-Forti**. It tells us that somehow the collection of all ordinals is *too big* to form a set. It also warns us that if we handle the intuitive concept of a set too carelessly, it might lead to contradictions and inconsistencies.

Later on we will develop an axiomatic approach to sets which aims to exclude antinomies like this. In this framework, we will be able to formally show that Ord is not a set. It forms what we will call a **proper class**.

Representing well-orders as ordinals

We introduced ordinals with the goal to have a specific representation for any well-order.

> **Theorem (representation theorem for well-orders)**
> Any well-ordered set $(A, <)$ is order-isomorphic to a unique ordinal α. The isomorphism is unique.

The ordinal α is called the **order type** of $(A, <)$. We will delay the proof of this theorem for a while, until we learn how to extend *induction* and *recursion* into the transfinite.

0.3 Polish spaces

0.3.1 Polish Spaces

The proofs in the introduction section are quite general, that is, they make little use of specific properties of $\mathbb{R}$. If we scan the arguments carefully, we see that we can replace $\mathbb{R}$ by any metric space that is **complete and contains a countable basis of the topology**.

Review of some concepts from topology

Basis Let $(X, \mathcal{O})$ be a topological space. A family $\mathcal{B} \subseteq \mathcal{O}$ of subsets if X is a **basis** for the topology if every open set from $\mathcal{O}$ is the union of elements of $\mathcal{B}$. For example, the open intervals with rational endpoints form a basis of the standard topology of $\mathbb{R}$. A family $\mathcal{S} \subseteq \mathcal{O}$ is a **subbasis** if the set of finite intersections of sets in $\mathcal{S}$ is a basis for the topology.

Finally, if $\mathcal{S}$ is any family of subsets of X, the **topology generated by** $\mathcal{S}$ is the smallest topology on X containing $\mathcal{S}$. It consists of all unions of finite intersections of sets in $\mathcal{S} \cup \{X, \emptyset\}$.

Density A set $D \subset X$ is **dense** if for any open $U \neq \emptyset$ there exists $z \in D \cap U$. If a topological space $(X, \mathcal{O})$ has a countable dense subset, the space is called **separable**.

Products If $(X_i)_{i \in I}$ is a family of topological spaces, one defines the **product topology** on $\Pi_{i \in I} X_i$ to be the topology generated by the sets $\pi_i^{-1}(U)$, where $i \in I$, $U \subseteq X_i$ is open, and $\pi_i : \Pi_{i \in I} X_i \to X_i$ is the ith projection.

Now suppose (X, d) is a metric space. With each point $x \in X$ and every $\varepsilon > 0$ we associate an **ε-neighborhood** or **ε-ball**

$$U_\varepsilon(x) = \{y \in X : d(x, y) < \varepsilon\}. \tag{30}$$

The topology generated by the ε-neighborhoods is called the *topology of the metric space* (X, d). If this topology agrees with a given topology $\mathcal{O}$ on X, we say the metric d is **compatible** with the topology $\mathcal{O}$. If for a topological space $(X, \mathcal{O})$ there exists a compatible metric, $(X, \mathcal{O})$ is called **metrizable**.

If a topological space $(X, \mathcal{O})$ is separable and metrizable, then the balls with center in a countable dense subset D and rational radius form a *countable base of the topology.*

Polish spaces – the basics

Definition 0.3.1. *A **Polish space** is a separable topological space X for which exists a compatible metric d such that (X, d) is a complete metric space.*

There may be many different compatible metrics that make X complete. If X is already given as a complete metric space with countable dense subset, then we call X a *Polish metric space.*

The standard example is, of course, $\mathbb{R}$, the set of real numbers. One can obtain other Polish spaces using the following basic observations. (We leave the proof as an exercise.)

Proposition 0.3.1. *1. A closed subset of a Polish space is Polish.*

2. The product of a countable (in particular, finite) sequence of Polish spaces is Polish.

3. Any topological space homeomorphic to a Polish space is Polish.

We conclude that $\mathbb{R}^n$, $\mathbb{C}$, $\mathbb{C}^n$, the unit interval $[0, 1]$, the unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, and the infinite dimensional spaces $\mathbb{R}^{\mathbb{N}}$ and $[0, 1]^{\mathbb{N}}$ (the *Hilbert cube*) are Polish spaces.

Any countable set with the **discrete topology** is Polish, by means of the **discrete metric** $d(x, y) = 1 \Leftrightarrow x \neq y$.

Some subsets of Polish spaces are Polish but not closed.

Exercise
By choosing a suitable metric, show that $(0,1)$, the open unit interval, is a Polish space.

We will later characterize all subsets of Polish spaces that are Polish themselves.

Product spaces

In a certain sense, the most important Polish spaces are of the form $A^{\mathbb{N}}$, where A is a countable set carrying the discrete topology. The standard cases are

> $2^{\mathbb{N}}$, the Cantor space and $\mathbb{N}^{\mathbb{N}}$, the Baire space.

We will, for now, denote elements from $A^{\mathbb{N}}$ by lower case greek letters from the beginning of the alphabet. The n-th term of α we denote by either $\alpha(n)$ or α_n, whichever is more convenient.

We endow A with the discrete topology. The product topology on these spaces has a convenient characterization. Given a set A, let $2^{<\mathbb{N}}[A]$ be the sets of all finite sequences over A. Given $\sigma, \tau \in A^{<\mathbb{N}}$, we write $\sigma \subseteq \tau$ to indicate that σ is an initial segment of τ. $\subset$ means the initial segment is proper. This notation extends naturally to hold between elements of $2^{<\mathbb{N}}[A]$ and $A^{\mathbb{N}}$, $\sigma \subset \alpha$ meaning that σ is a finite initial segment of α.

A basis for the product topology on $A^{\mathbb{N}}$ is given by the **cylinder sets**

$$N_\sigma = \{\alpha \in A^{\mathbb{N}} \colon \sigma \subset \alpha\}, \tag{31}$$

that is, the set of all infinite sequences extending σ. The complement of a cylinder is a union of cylinders and hence open. Therefore, each set N_σ is clopen.

A compatible metric is given by

$$d(\alpha,\beta) = \begin{cases} 2^{-N} & \text{where } N \text{ is least such that } \alpha_N \neq \beta_N \\ 0 & \text{if } \alpha = \beta. \end{cases} \tag{32}$$

The representation of the topology via cylinders (which are characterized by finite objects) allows for a combinatorial treatment of many questions and will be essential later on.

Proposition 0.3.2 (Topological properties of). $A^{\mathbb{N}}$
Let A be a countable set, equipped with the discrete topology. Suppose $A^{\mathbb{N}}$ is equipped with the product topology. Then the following hold.

1. *$A^{\mathbb{N}}$ is Polish.*
2. *$A^{\mathbb{N}}$ is zero-dimensional, i.e. it has a basis of clopen sets.*
3. *$A^{\mathbb{N}}$ is compact if and only if A is finite.*

Via the mapping

$$\alpha \mapsto \sum_{i=0}^{\infty} \frac{2\alpha_i}{3^{i+1}}, \tag{33}$$

$2^{\mathbb{N}}$ is homeomorphic to the middle-third Cantor set in $\mathbb{R}$, whereas the **continued fraction** mapping

$$\beta \mapsto \beta_0 + \cfrac{1}{\beta_1 + \cfrac{1}{\beta_2 + \cfrac{1}{\beta_3 + \dots}}} \tag{34}$$

provides a homeomorphism between $\times(\mathbb{N} \setminus \{0\})^{\mathbb{N}}$ and the irrational real numbers.

The universal role played by the discrete product spaces is manifested in the following results.

Theorem 0.3.1. *Every uncountable Polish space contains a homeomorphic embedding of Cantor space* $2^{\mathbb{N}}$.

The proof is similar to the proof of Theorem (Cantor, 1884). Note that the proof actually constructs an embedding of $2^{\mathbb{N}}$. The continuity of the mapping is straightforward.

In a similar way we can adapt the proof of Cantor-Bendixson Theorem to show that the *perfect subset property* holds for closed subsets of Polish spaces.

Theorem 0.3.2 (Cantor-Bendixson Theorem for Polish spaces)**.** *Every uncountable closed subset of a Polish space contains a perfect subset.*

Finally, we can characterize Polish spaces as continuous images of Baire space.

Theorem 0.3.3. *Every Polish space* X *is the continuous image of* $\mathbb{N}^{\mathbb{N}}$.

Proof. Let d be a compatible metric on X, and let $D = \{x_i \colon i \in \mathbb{N}\}$ be a countable dense subset of X. Every point in X is the limit of a sequence in D. Define a mapping $g : \mathbb{N}^{\mathbb{N}} \to X$ by putting

$$\alpha = \alpha(0)\,\alpha(1)\,\alpha(2)\cdots \mapsto \lim_n x_{\alpha(n)}. \tag{35}$$

The problem is, of course, that the limit on the right hand side not necessarily exists. We have to proceed more carefully. Given $\alpha \in \mathbb{N}$, we put $y_0^\alpha = x_{\alpha(0)}$ and define iteratively

$$y_{n+1}^\alpha = \begin{cases} x_{\alpha(n+1)} & \text{if } d(y_n^\alpha, x_{\alpha(n+1)}) < 2^{-n}, \\ y_n^\alpha & \text{otherwise .} \end{cases} \tag{36}$$

The resulting sequence (y_n^α) is clearly Cauchy in X, and hence converges to some point $y^\alpha \in X$, by completeness. We define

$$f(\alpha) = y^\alpha. \tag{37}$$

f is continuous, since if α and β agree up to length N (that is, their distance is at most 2^{-N} with respect to the above metric), then the sequences (y_n^α) and (y_n^β) will agree up to index N, and all further terms are within 2^{-N} of y_N^α and y_N^β, respectively.

Finally, since D is dense in X, f is a surjection.

∎

0.3.2 Excursion: The Urysohn Space

Recall that a mapping $f : X \to Y$ between two metric spaces (X, d_X) and (Y, d_Y) is an **isometry** if

$$d_Y(f(x), f(y)) = d_X(x, y) \quad \text{for all } x, y \in X, \tag{38}$$

that is, an isometry is a mapping that preserves distances. The function f is also called an *isometric embedding* of X into Y. X and Y are *isometric* if there exists a bijective isometry between them.

Universal spaces

Theorem 0.3.4. *There exists a Polish metric space $\mathbb{U}$ such that every Polish metric space isometrically embeds into $\mathbb{U}$.*

A concrete example of such a space is $\mathcal{C}[0, 1]$.

Exercise

Show that the set $\mathcal{C}[0, 1]$ of all continuous, real-valued functions on $[0, 1]$ with the metric

$$d(f, g) = \sup\{|f(x) - g(x)| \colon x \in [0, 1]\} \tag{39}$$

contains an isomorphic copy of any Polish metric space.

But this example is not quite what we have in mind here. There exists another space with a stronger, more "intrinsic" universality property. This space was first constructed by Pavel Urysohn in 1927 **?**.

The construction features an **amalgamation principle** that has surfaced in other areas like model theory or graph theory.

Extensions of finite isometries and Urysohn universality

Suppose X is a Polish metric space. Let $D = \{x_1, x_2, \dots\}$ be a countable, dense subset. We first observe that, to isometrically embed X into another Polish space, it is sufficient to embed D.

Lemma 0.3.5. *If Y is Polish, then any isometric embedding e of D into Y extends to an isometric embedding e^* of X into Y.*

Proof. Given $z \in X$, let (x_{i_n}) be a sequence in D converging to z. Since (x_{i_n}) converges, it is Cauchy.

e is an isometry, and thus $y_n := e(x_{i_n})$ is Cauchy, and since Y is Polish, (y_n) converges to some $y \in Y$. Put $e^*(z) = y$.

To see that this mapping is well-defined, let (x_{j_n}) be another sequence with $x_{j_n} \to z$. Then $d(x_{i_n}, x_{j_n}) \to 0$, and hence $d(e(x_{i_n}), e(x_{j_n}) = d(y_n, e(x_{j_n})) \to 0$, implying $e(x_{j_n}) \to y$.

Furthermore, suppose $w = \lim x_{k_n}$ is another point in X. Then (since a metric is a continuous mapping from $Y \times Y \to \mathbb{R}$)

$$d(e^*(z), e^*(w)) = \lim d(e(x_{i_n}), e(x_{k_n})) = \lim d(x_{i_n}, x_{k_n}) = d(z, w). \tag{40}$$

Thus e^* is an isometry.

□

In order to embed D, we can now exploit the inductive structure of $\mathbb{N}$ and reduce the task to extending finite isometries.

Suppose we have constructed an isometry e between $F_N = \{x_1, \dots, x_N\} \subset D$ and a space Y. We would like to extend the isometry to include x_{N+1}. For this we have to find an element $y \in Y$ such that for all $i \leq N$

$$d_Y(y, e(x_i)) = d_X(x_{N+1}, x_i). \tag{41}$$

This extension property gives rise to the following definition.

Definition 0.3.2. *A Polish metric space* (Y, d_Y) *is* ***Urysohn universal*** *if for every finite subspace* $F \subset Y$ *and any extension* $F^* = F \cup \{x^*\}$ *with metric a* d^* *such that*

$$d^*|_{F\times F} = d_Y|_{F\times F}, \tag{42}$$

there exists a point $u \in Y$ *such that*

$$d_Y(u, x) = d^*(x^*, x) \quad \text{for all } x \in F. \tag{43}$$

As outlined above, the extension property of Urysohn universal spaces implies the desired isometric embedding property.

Proposition 0.3.3. *Let* U *be a Urysohn universal Polish metric space. For any Polish metric space* (X, d) *there exists an isometric embedding from* X *into* U*.*

But the extension property also implies a strong intrinsic extension property for the Urysohn space itself.

Proposition 0.3.4. *Let* U *be a Urysohn universal Polish metric space. Every isometry between finite subsets of* $\mathbb{U}$ *extends to an isometry of* U *onto itself.*

The proof applies the Back-and-forth method that you may know from the rationals: every order-isomorphism between finite subsets of $\mathbb{Q}$ extends to an automorphism of $(\mathbb{Q}, <)$.

This property (which can be formulated for structures in general) is also known as **homogeneity**. It plays an important role, for example, in model theory **?** and in the topological dynamics of automorphism groups of countable structures **?**.

Exercise
Show that any two Urysohn universal spaces are isometric.

We will prove the existence of this unique Polish space, which we denote by $\mathbb{U}$, in the following sections.

Constructing the Urysohn space – a first approximation

We first give a construction of a space that has the extension property, but is not Polish. After that we will take additional steps to turn it into a Polish space.

The crucial idea is to observe that if X is a metric space and $x \in X$, then the mapping $f_x : X \to \mathbb{R}^{\geq 0}$ given by

$$f_x(y) = d_X(x, y) \tag{44}$$

is 1-Lipschitz. Recall that a function g between metric spaces X and Y is L**-Lipschitz**, $L > 0$ if for every $x, y \in X$,

$$d(g(x), g(y)) \leq L\, d(x, y). \tag{45}$$

Let $\mathrm{Lip}_1(X)$ be the set of 1-Lipschitz mappings from X to $\mathbb{R}$. We endow $\mathrm{Lip}_1(X)$ with the supremum metric

$$d(f, g) = \sup\{|f(x) - g(x)| \colon x \in X\}. \tag{46}$$

If $\mathrm{diam}(X) \leq \mathrm{d}$ and f, g are 1-Lipschitz, then $d(f, g)$ is indeed finite. However, we will later need that the resulting space is also bounded. Let $\mathrm{Lip}_1^{\mathrm{d}}(X)$ be the space of all 1-Lipschitz functions from X to $[0, \mathrm{d}]$.

Clearly, $\mathrm{diam}(\mathrm{Lip}_1^{\mathrm{d}}(X)) \leq \mathrm{d}$.

With this metric, the mapping $x \mapsto f_x(y) = d(x, y)$ becomes an isometry: We have

$$d(f_x, f_z) = \sup\{|d(x,y) - d(z,y)| \colon y \in X\}. \tag{47}$$

By the reverse triangle inequality, this is always $\leq d(x,z)$. On the other hand, setting $y = z$ yields $d(f_x, f_z) \geq d(x,z)$. This embedding of X into $\mathrm{Lip}_1^{\mathrm{d}}(X)$ is called the **Kuratowski embedding**.

We use this fact as follows: If $X^* = X \sqcup \{x^*\}$ and d^* is an extension of d_X, then f_{x^*} is an element of $\mathrm{Lip}_1^{\mathrm{d}}(X)$, and as above, for any $x \in X$

$$d(f_{x^*}, f_x) = d^*(x^*, x). \tag{48}$$

Hence $\mathrm{Lip}_1^{\mathrm{d}}(X)$ has an extension property of the kind we are looking for.

Iterative construction: Let X_0 be any non-empty Polish space with finite diameter $\mathrm{d} > 0$. Given X_n, let $\mathrm{d}(n) = \mathrm{diam}(X_n)$ and set $X_{n+1} = \mathrm{Lip}_1^{2\mathrm{d}(n)}(X_n)$. Finally, put $X_\infty = \bigcup_n X_n$. Note that X_∞ inherits a well-defined metric d from the X_n, which embed isometrically into it.

We wan to verify that X_∞ has the extension property needed to be Urysohn universal. Let F be a finite subset of X_∞. There exists N such that $F \subset X_N$. Suppose $F^* = F \sqcup \{x^*\}$ and d^* is an extension of d to F^*. Let $\mathrm{d}^* = \mathrm{diam}(F^*)$. Note that $\mathrm{diam}(X_n) = 2^n \mathrm{d}$. Choose M so that $M \geq N$ and $\mathrm{diam}(X_M) \geq \mathrm{d}^*$. The next lemma ensures that we can find $f \in X_{M+1}$ such that $f(x) = d^*(x^*, x)$ for all $x \in F$.

Lemma 0.3.6 (McShane-Whitney)**.** *Let X be a metric space with $\mathrm{diam}(X) \leq \mathrm{d}$, $A \subseteq X$, and $f \in \mathrm{Lip}_1^{\mathrm{d}}(A)$, then f can be extended to a 1-Lipschitz function f^* on all of X such that*

$$f^*|_A = f \qquad \text{and} \qquad f^* \in \mathrm{Lip}_1^{2\mathrm{d}}(X). \tag{49}$$

Proof. For each $a \in A$ define $f_a : X \to \mathbb{R}$ as

$$f_a(x) = f(a) + d(a,x). \tag{50}$$

Then f_a is 1-Lipschitz, by the reverse triangle inequality. Let

$$f^*(x) = \inf\{f_a(x) \colon a \in A\}. \tag{51}$$

Then $f^*(a) = f(a)$ for all $a \in A$. Let $x, y \in X$ and $\varepsilon > 0$. Wlog assume $f^*(y) \geq f^*(x)$. Pick $a \in A$ so that $f_a(x) \leq f^*(x) + \varepsilon$. Then

$$\begin{aligned} |f^*(x) - f^*(y)| = f^*(y) - f^*(x) &\leq f^*(y) - f_a(x) + \varepsilon \\ &\leq f_a(y) - f_a(x) + \varepsilon \leq d(x,y) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we have $|f^*(x) - f^*(y)| \leq d(x,y)$.

Finally, we have $f(a) \leq f_a(x) \leq f(a) + \mathrm{d}$ and thus $0 \leq f^*(x) \leq f_a(x) \leq 2\mathrm{d}$.

□

Finishing the construction

The set X_∞ we constructed has two deficiencies with respect to our goal of constructing a Urysohn universal space: X_∞ is not necessarily separable, and X_∞ is not necessarily complete.

To make X_∞ separable, we observe that if X is compact, then the set $\mathrm{Lip}_1^{\mathrm{d}}(X)$ is closed in $\mathcal{C}(X)$ (the set of all real-valued continuous functions on X), bounded, and equicontinuous. By the **Arzelà-Ascoli Theorem**, $\mathrm{Lip}_1^{\mathrm{d}}(X)$ is compact. Every compact metric space is separable: For every $\varepsilon > 0$, there exists a finite covering of the space with sets of diam $< \varepsilon$. Letting ε traverse all positive rationals and picking a point from each set in an ε-covering yields a countable dense subset. Hence if we start with X_0 compact, each X_n will be compact, too. A countable union of separable spaces is separable, thus X_∞ is separable.

To obtain a complete space, we can pass from X_∞ to its *completion* $\overline{X_\infty}$. First note that if a metric space X is separable, so is its completion $\overline{X}$. However, we also have to ensure that $\overline{X_\infty}$ retains the universality property of X_∞.

Lemma 0.3.7. *If a complete metric space (Y, d) admits a dense Urysohn universal subspace $\mathcal{U}$, then Y is Urysohn universal.*

Proof. We follow **?**. Let $F = \{x_1, \ldots, x_n\} \subset Y$ and assume $F^* = F \sqcup \{x^*\}$ is an extension with metric d^*.

We first note that Y is **approximately universal**. This means that for any $\varepsilon > 0$, there exists a point $y^* \in Y$ such that

$$|d(y*, x) - d^*(x^*, x)| < \varepsilon \quad \text{for all } x \in F. \tag{$*$}$$

This can be seen as follows. Since $\mathcal{U}$ is dense in Y, we can find a finite set $F_\varepsilon = \{z_1, \ldots, z_n\} \subset \mathcal{U}$ such that

$$d(x_i, z_i) < \varepsilon \quad \text{for } 1 \leq i \leq n. \tag{52}$$

Now use the Urysohn universality of $\mathcal{U}$ for the set $G^* = \{z_1, \ldots, z_n\} \sqcup \{x^*\}$ with the metric

$$d^{**}(x^*, z_i) = d^*(x^*, x_i) \qquad (i = 1, \ldots, n) \tag{53}$$

to find $z \in \mathcal{U}$ with

$$d(z, z_i) = d^{**}(x^*, z_i) = d^*(x^*, x_i) \qquad (i = 1, \ldots, n) \tag{54}$$

Then, by the reverse triangle inequality,

$$|d(z, x_i) - d^*(x^*, z_i)| = |d(z, x_i) - d(z, z_i)| \leq d(z_i, x_i) = \varepsilon, \tag{55}$$

as required.

We use this approximate universality to construct a Cauchy sequence (y_k) in Y of 'approximate' extension points that satisfy $(*)$ for smaller and smaller ε.

Let $0 < \delta = \max\{d^*(x^*, x_i) \colon 1 \leq i \leq n\}$. The formal requirements for the sequence (y_i) are as follows.

1. $|d(y_k, x_i) - d^*(x^*, x_i)| \leq 2^{-k}\delta$.
2. $d(y_{k+1}, y_k) \leq 2^{-k}\delta$.

The sequence necessarily converges in Y and the limit point must be a true extension point, due to (1.)

Suppose we have already constructed $y_1, \ldots, y_k$ satisfying (1.), (2.). Add an (abstract) point y^*_{k+1} to $F_k = F \cup \{y_1, \ldots, y_k\}$. Let $F^*_{k+1} = F_k \sqcup \{y^*_{k+1}\}$.

We want to use approximate universality on F^*_{k+1}. To this end we have to define a metric e^* on F^*_{k+1} that has the following properties

$$\begin{array}{rl} (i) & e^*|_{F_k} = d|_{F_k} \\ (ii) & e^*(y^*_{k+1}, x_i) = d^*(x^*, x_i) \quad (1 \leq i \leq n) \\ (iii) & e^*(y^*_{k+1}, y_k) = 2^{-k-1}\delta \end{array}$$

Indeed such a metric exists: The condition (i) already defines a metric on the set F_k. The conditions (i)-(iii) also define a metric on $F \cup \{y_k, y^*_{k+1}\}$ – the only thing to check for this is the triangle inequality for y_k, y^*_{k+1}:

$$|e^*(x_i, y_k) - e^*(y^*_{k+1}, x_i)| = |d(x_i, y_k) - d^*(x^*, x_i)| \leq 2^{-k}\delta = e^*(y_k, y^*_{k+1}), \tag{56}$$

by (1.). These metrics agree on the set

$$F_k \cap (F \cup \{y_k, y^*_{k+1}\}) = F \cup \{y_k\}. \tag{57}$$

Therefore, we can "merge" them to a metric on all of F^*_{k+1} by letting

$$e^*(y^*_{k+1}, y_j) = \inf\{e^*(y^*_{k+1}, z) + e^*(z, y_j) \colon z \in \{y_1, \dots, y_{k-1}\}\}. \tag{58}$$

Now choose $\varepsilon < 2^{-k-1}\delta$ and apply approximate universality to F^*_{k+1}. This yields a point $y_{k+1} \in Y$ such that

$$|d(y_{k+1}, z) - e^*(y^*_{k+1}, z)| < 2^{-k-1}\delta \tag{59}$$

for all $z \in F_k$. By definition of e^*, we have

$$|d(y_{k+1}, x_i) - d^*(y^*_{k+1}, z)| < 2^{-k-1}\delta \tag{60}$$

for $1 \leq i \leq n$, and (iii) yields

$$d(y_{k+1}, y_k) < e^*(y^*_{k+1}, y_k) + \varepsilon \leq 2^{-k-1}\delta + 2^{-k-1}\delta = 2^{-k}\delta \tag{61}$$

as required.

□

0.3.3 Trees

Let A be a set. Recall that the set of all finite sequences over A is denoted by $2^{<\mathbb{N}}[A]$, while $A^{\mathbb{N}}$ denotes the set of all infinite sequences over A. Given $\alpha \in A^{\mathbb{N}}$, $n \in \mathbb{N}$, $\alpha\,|_n$ denotes the initial segment of α of length n.

Definition 0.3.3. *A* ***tree*** *on A is a set $T \subseteq 2^{<\mathbb{N}}[A]$ that is* ***closed under prefixes****, that is*

$$\forall \sigma, \tau\ [\tau \in T \,\&\, \sigma \subseteq \tau \ \Rightarrow\ \sigma \in T] \tag{62}$$

We call the elements of T **nodes**.

A sequence $\alpha \in A^{\mathbb{N}}$ is an **infinite path through** or **infinite branch of** T if for all n, $\alpha\,|_n = (\alpha_0, \alpha_1, \dots, \alpha_{n-1}) \in T$. We denote the set of infinite paths through T by $[T]$.

An important criterion for a tree to have infinite paths is the following.

Theorem 0.3.8 (Königs Lemma)**.** *Any tree T with infinitely many nodes that is* ***finite branching*** *(i.e. each node has at most finitely many immediate extensions) has an infinite path.*

Proof. We construct an infinite path inductively.

Let T_σ denote the tree "above" σ, i.e. $T_\sigma = \{\tau \in 2^{<\mathbb{N}}[A] \colon \sigma^\frown\tau \in T\}$. If T is finite branching, by the **pigeonhole principle**, at least one of the sets T_σ for $|\sigma| = 1$ must be infinite. Pick such a σ and let $\alpha\,|_1 = \sigma$.

Repeat the argument for $T = T_\sigma$ and continue inductively. This yields a sequence $\alpha \in [T]$.

□

If $[T] = \emptyset$, we call T **well-founded**. The motivation behind this is that T is well-founded if and only if the **inverse prefix** relation

$$\sigma \preceq \tau \quad :\Leftrightarrow \quad \sigma \supseteq \tau \tag{63}$$

is well-founded, i.e. it does not have an infinite descending chain.

If $T \neq \emptyset$ is well-founded, we can assign T an ordinal number, its **rank** $\rho(T)$.

- If σ is a **terminal node**, i.e. σ has no extensions in T, then let $\rho_T(\sigma) = 0$.
- If σ is not terminal, and $\rho_T(\tau)$ has been defined for all $\tau \supset \sigma$, we set $\rho_T(\sigma) = \sup\{\rho_T(\tau) + 1 \colon \tau \in T, \tau \supset \sigma\}$.
- Finally, set $\rho(T) = \sup\{\rho_T(\sigma) + 1 \colon \sigma \in T\} = \rho_T(\varnothing) + 1$, where $\varnothing$ denotes the empty string.

Orderings on trees

Now suppose A is linearly ordered by a relation $\leq_A$. The **lexicographical ordering** $\leq_{\text{lex}}$ of $2^{<\mathbb{N}}[A]$ is defined as

$$\sigma \leq_{\text{lex}} \tau \quad :\Leftrightarrow \sigma = \tau \ \text{ or } \ \exists i < |\sigma|, |\tau|\ (\sigma_i <_A \tau_i \text{ and } \forall j < i\ \sigma_j = \tau_j)$$

This ordering extends to $A^{\mathbb{N}}$ in a natural way.

Proposition 0.3.5. *If $\leq_A$ is a well-ordering of A and T is a tree on A with $[T] \neq \emptyset$, then $[T]$ has a $\leq_{\text{lex}}$-minimal element, the* ***leftmost branch****.*

Proof. We **prune** the tree T by deleting any node that is not on an infinite branch. This yields a subtree $T' \subseteq T$ with $[T'] = [T]$.

Let $T'_n = \{\sigma \in T' \colon |\sigma| = n\}$. Since $\leq_A$ is a well-ordering on A, T'_1 must have a $\leq_{\text{lex}}$-least element. Denote it by $\alpha\,|_1$. Since T' is pruned, $\alpha\,|_1$ must have an extension in T, and we can repeat the argument to obtain $\alpha\,|_2$.

Continuing inductively, we define an infinite path α through T', and it is straightforward to check that α is a $\leq_{\text{lex}}$-minimal element of $[T']$ and hence of $[T]$.

□

We can combine the $\leq_{\text{lex}}$-ordering with the inverse prefix order to obtain a linear ordering of $2^{<\mathbb{N}}[A]$. This ordering has the nice property that if A is well-ordered and T is well-founded, then the ordering restricted to T is a well-ordering.

Definition 0.3.4. *The **Kleene-Brouwer ordering** $\leq_{\text{KB}}$ of $2^{<\mathbb{N}}[A]$ is defined as follows.*

$$\sigma \leq_{\text{KB}} \tau \quad :\Leftrightarrow \quad \sigma \supseteq \tau \quad \text{or} \quad \sigma \leq_{\text{lex}} \tau.$$

This means σ is smaller than τ if it is a proper extension of τ or "to the left" of τ.

We now have

Proposition 0.3.6. *Assume $(A, \leq)$ is a well-ordered set. Then for any tree T on A,*

$$T \text{ is well-founded} \quad \Leftrightarrow \quad \leq_{\text{KB}} \text{ restricted to } T \text{ is a well-ordering.}$$

Proof. Suppose T is not well-founded. Let $\alpha \in [T]$. Then $\alpha\mid_0, \alpha\mid_1, \dots$ is an infinite descending sequence with respect to $\leq_{\text{KB}}$.

Conversely, suppose $\sigma_0 >_{\text{KB}} \sigma_1 >_{\text{KB}} \dots$ is an infinite descending sequence in T. By the definition of $>_{\text{KB}}$, this implies $\sigma_1(0) \geq_A \sigma_2(0) \geq_A \dots$ as a sequence in A. Since A is well-ordered, this sequence must eventually be constant, say $\sigma_n(0) = a_0$ for all $n \geq n_0$.

Since the σ_n are descending, by the definition of $\leq_{\text{KB}}$ it follows that $|\sigma_n| \geq 2$ for $n > n_0$. Hence we can consider the sequence $\sigma_{n_0+1}(1) \geq_A \sigma_{n_0+2}(1) \geq_A \dots$ in A. Again, this must be constant $= a_1$ eventually. Inductively, we obtain a sequence $\alpha = (a_0, a_1, a_2, \dots) \in [T]$, that is, T is not well-founded. $\square$

> **Caution**
> The order type of a well-founded tree under $\leq_{\text{KB}}$ usually is not equal to its rank $\rho(T)$.

Coding trees

We can also define an ordering on $2^{<\mathbb{N}}[A]$ via an injective mapping from $2^{<\mathbb{N}}[A]$ to some linearly ordered set A. We will use this repeatedly for the case $A = \mathbb{N}$ and $A = \{0, 1\}$.

For $A = \mathbb{N}$, we can use the standard coding mapping

$$\pi : (a_0, a_1, \dots, a_n) \mapsto p_0^{a_0+1} p_1^{a_1} \cdots p_n^{a_n},$$

where p_k is the kth prime number. This embeds $\mathbb{N}^{<\mathbb{N}}$ into $\mathbb{N}$, and we can well-order $\mathbb{N}^{<\mathbb{N}}$ by letting $\sigma < \tau$ if and only if $\pi(\sigma) < \pi(\tau)$.

For $A = \{0, 1\}$ we set

$$\pi : (b_0, b_1, \dots, b_{n-1}) \mapsto (2^n - 1) + \sum_{i=0}^{n-1} b_i 2^i.$$

These two mappings allows us henceforth to see **trees as subsets of the natural numbers**. This will be an important component in exploring the relation between topological and arithmetical complexity.

Trees and closed sets

Let A be a set with the discrete topology. Consider $A^{\mathbb{N}}$ with the product topology (and compatible metric) defined in Lecture 2.

Proposition 0.3.7. *A set $F \subseteq A^{\mathbb{N}}$ is closed if and only if there exists a tree T on A such that $F = [T]$.*

Proof. Suppose F is closed. Let

$$T_F = \{\sigma \in 2^{<\mathbb{N}}[A] \colon \sigma \subset \alpha \text{ for some } \alpha \in F\}.$$

Then clearly $F \subset [T_F]$. Suppose $\alpha \in [T_F]$. This means for any n, $\alpha\mid_n \in T_F$, which implies that there exists $\beta_n \in F$ such that $\alpha_n \subset \beta_n$. The sequence (β_n) converges to α, and since F is closed, $\alpha \in F$.

For the other direction, suppose $F = [T]$. Let $\alpha \in A^{\mathbb{N}} \setminus F$. Then there exists an n such that $\alpha \mid_n \notin T$. Since a tree is closed under prefixes, no extension of $\alpha \mid_n$ can be in T. This implies $N_{\alpha|_n} \subseteq A^{\mathbb{N}} \setminus F$, and hence $A^{\mathbb{N}} \setminus F$ is open.

$\square$

Continuous mappings on product spaces

Let $f : A^{\mathbb{N}} \to A^{\mathbb{N}}$ be continuous. We define a mapping $\phi : 2^{<\mathbb{N}}[A] \to 2^{<\mathbb{N}}[A]$ by setting

$$\phi(\sigma) = \text{ the longest } \tau \text{ with } |\tau| \leq |\sigma| \text{ such that } N_\sigma \subseteq f^{-1}(N_\tau).$$

This mapping has the following properties:

1. It is **monotone**, i.e. $\sigma \subseteq \tau$ implies $\phi(\sigma) \subseteq \phi(\tau)$.
2. For any $\alpha \in A^{\mathbb{N}}$ we have $\lim_n |\phi(\alpha \mid_n)| = \infty$. This follows directly from the continuity of f: For any neighborhood N_τ of $f(\alpha)$ there exists a neighborhood N_σ of α such that $f(N_\sigma) \subseteq N_\tau$. But τ has to be of the form $\tau = f(\alpha) \mid_m$, and σ of the form $\alpha \mid_n$. Hence for any m there must exist an n such that $\phi(\alpha \mid_n) \supseteq f(\alpha) \mid_m$.

On the other hand, if a function $\phi : 2^{<\mathbb{N}}[A] \to 2^{<\mathbb{N}}[A]$ satisfies (1.) and (2.), it induces a function $\phi^* : A^{\mathbb{N}} \to A^{\mathbb{N}}$ by letting

$$\phi^*(\alpha) = \lim_n \phi(\alpha \mid_n) = \text{ the unique sequence extending all } \phi(\alpha \mid_n).$$

This ϕ^* is indeed continuous: The preimage of N_τ under ϕ^* is given by

$$(\phi^*)^{-1}(N_\tau) = \bigcup \{N_\sigma \colon \phi(\sigma) \supseteq \tau\},$$

which is an open set.

We have shown the following.

Proposition 0.3.8. *A mapping $f : A^{\mathbb{N}} \to A^{\mathbb{N}}$ is continuous if and only if there exists a mapping ϕ satisfying (1) and (2) such that $f = \phi^*$.*

Note that we can completely describe a topological concept, continuity, through a relation between finite strings.

0.4 The Borel hierarchy

0.4.1 Borel Sets

The Borel sets in a topological space are the σ-algebra generated by the open sets. That means one can build up the Borel sets from the open sets by iterating the operations of complementation and countable union. This generates sets that are more and more complicated, which is reflected in the **Borel hierarchy**. The complexity is reflected on the logical side by the number of quantifier changes needed to define the set. There is a close connection between the arithmetical hierarchy in computability and the Borel hierarchy.

Definition 0.4.1. *Let X be a set. A σ-**algebra** $\mathcal{S}$ on X is a collection of subsets of X such that $\mathcal{S}$ is closed under complements and countable unions, that is*

- *if $A \in \mathcal{S}$, then $X \setminus A \in \mathcal{S}$, and*
- *if $(A_n)_{n\in\mathbb{N}}$ is a sequence of sets in $\mathcal{S}$, then $\bigcup_n A_n \in \mathcal{S}$,*

If the enveloping space X is clear, we use $\neg A$ to denote the complement of A in X.
It is easy to derive that a σ-algebra is also closed under the following set-theoretic operations:

- *countable intersections* – we have $\bigcap A_n = \neg \bigcup_n \neg A_n$.
- *differences* – we have $A \setminus B = A \cap \neg B$.
- *Symmetric differences* – we have $A \triangle B = (A \cap \neg B) \cup (\neg A \cap B)$.

Definition 0.4.2. *Let $(X, \mathcal{O})$ be a topological space. The collection of **Borel sets** in X is the smallest σ-algebra containing the open sets in $\mathcal{O}$.*

Of course, one has to make sure that this collection actually exists. For this, note that the intersection of any collection of σ-algebras is again a σ-algebra, so the Borel sets are just the intersection of all σ-algebras containing $\mathcal{O}$. (Note the the full power set of X is such a σ-algebras, so we are not taking an empty intersection.)

The definition of Borel sets is rather *"external"*. It does not give us much of an idea what Borel sets *look like*. One can arrive at the family of Borel sets also through a construction from *"within"*. This reveals more structure and gives rise to the **Borel hierarchy**.

The Borel hierarchy

To generate the Borel sets, we start with the open sets. By closing under complements, we obtain the closed sets. We also have to close under countable unions. The open sets are already closed under this operation, but the closed sets are not.

Countable unions of closed sets are classically known as F_σ sets. Their complements, i.e. countable intersections of open sets, are the G_δ sets.

We can continue this way and form the $F_{\sigma\delta}$ sets – countable intersections of F_σ sets – the $G_{\delta\sigma}$ sets – countable unions of G_δ sets – and so on.

The $\sigma\delta$-notation soon becomes rather impractical, and hence we replace it by something more convenient, and much more suggestive, as we will see later.

To make the hierarchy that we are introducing well-behaved, we focus on **metrizable spaces**.

Definition 0.4.3. *Let X be a metrizable topological space. We inductively define the following collection of subsets of X.*

$$
\begin{aligned}
\mathbf{\Sigma}^0_1(X) &= \{U \colon U \subseteq X \text{ open } \}\\
\mathbf{\Pi}^0_n(X) &= \{\neg A \colon A \in \mathbf{\Sigma}^0_n(X)\} = \neg \mathbf{\Sigma}^0_n(X)\\
\mathbf{\Sigma}^0_{n+1}(X) &= \{\bigcup_k A_k \colon A_k \in \mathbf{\Pi}^0_n(X)\}
\end{aligned}
$$

Hence the open sets are precisely the sets in $\mathbf{\Sigma}^0_1$, the closed sets are the sets in $\mathbf{\Pi}^0_1$, the F_σ sets from the class $\mathbf{\Sigma}^0_2$ etc. If it is clear what the underlying space X is, we drop the reference to it and simply write $\mathbf{\Sigma}^0_n$ and $\mathbf{\Pi}^0_n$. Besides, we will say that a set $A \subseteq X$ *is* (or *is not*) $\mathbf{\Sigma}^0_n$ or $\mathbf{\Pi}^0_n$, respectively.

> **Question:** Does the collection of all $\mathbf{\Sigma}^0_n$ and $\mathbf{\Pi}^0_n$ exhaust the Borel sets of X?

We will see that the answer is no. We have to extend our inductive construction into the transfinite and consider classes $\mathbf{\Sigma}^0_\xi$, where ξ is a countable infinite ordinal.

The Borel sets of finite order

We fix a Polish space X. We want to establish the basic relationships between the different classes $\mathbf{\Sigma}^0_n$ and $\mathbf{\Pi}^0_m$ for X.

It follows from the definitions that $\mathbf{\Pi}^0_n \subseteq \mathbf{\Sigma}^0_{n+1}$ and $\mathbf{\Sigma}^0_n \subseteq \mathbf{\Pi}_{n+1}$.

Lemma 0.4.1. *In any metric space (X, d), every closed set is a G_δ set.*

Proof. Let $F \subset X$ be closed. For $n \geq 0$, put

$$F_n = \bigcup_{x \in F} U_{2^{-n}}(x). \tag{64}$$

Each F_n is open, and $F \subseteq \bigcup_{n\in\mathbb{N}} F_n$.

Moreover, if $x \in \bigcup_{n\in\mathbb{N}} F_n$, then there exists a sequence (x_n) such that for all n, $x_n \in F$ and $x \in U_{2^{-n}}(x_n)$. It follows that $x_n \to x$, and since F is closed, $x \in F$. Thus

$$F = \bigcup_{n\in\mathbb{N}} F_n, \tag{65}$$

which is G_δ.

□

Corollary 0.4.1.1. $\mathbf{\Sigma}^0_1 \subseteq \mathbf{\Sigma}^0_2$ *and* $\mathbf{\Pi}^0_1 \subseteq \mathbf{\Pi}^0_2$.

The second statement follows by passing to complements: If F is closed,

$$F = \neg\neg F = \neg \bigcup F_n = \bigcup \neg F_n,$$

where the F_n are closed.

There are also sets that can be both $\mathbf{\Sigma}^0_2$ and $\mathbf{\Pi}^0_2$, but neither $\mathbf{\Sigma}^0_1$ nor $\mathbf{\Pi}^0_1$. For example, consider the half-open interval $[0, 1)$.

$$[0,1) = \bigcup_n [0, 1 - 1/n] = \bigcap_m (-1/n, 1).$$

Therefore, it makes sense to define the **hybrid classes**:

$$\mathbf{\Delta}^0_n = \mathbf{\Sigma}^0_n \cap \mathbf{\Pi}^0_n.$$

Using induction, we can extend the inclusions in a straightforward way to higher n.

$$\begin{tikzcd}[column sep=small]
 & \mathbf{\Sigma}^0_1 \arrow[dr, phantom, sloped, "\subseteq"] & & \mathbf{\Sigma}^0_2 \arrow[dr, phantom, sloped, "\subseteq"] & & \mathbf{\Sigma}^0_3 \arrow[dr, phantom, sloped, "\subseteq"] & \\
\mathbf{\Delta}^0_1 \arrow[ur, phantom, sloped, "\subseteq"] \arrow[dr, phantom, sloped, "\subseteq"] & & \mathbf{\Delta}^0_2 \arrow[ur, phantom, sloped, "\subseteq"] \arrow[dr, phantom, sloped, "\subseteq"] & & \mathbf{\Delta}^0_3 \arrow[ur, phantom, sloped, "\subseteq"] \arrow[dr, phantom, sloped, "\subseteq"] & & {} \\
 & \mathbf{\Pi}^0_1 \arrow[ur, phantom, sloped, "\subseteq"] & & \mathbf{\Pi}^0_2 \arrow[ur, phantom, sloped, "\subseteq"] & & \mathbf{\Pi}^0_3 \arrow[ur, phantom, sloped, "\subseteq"] &
\end{tikzcd}$$

Theorem 0.4.2 (Weak Hierarchy Theorem)**.**

Are the inclusions are proper?

If the space is discrete, every open set is closed and vice versa, and hence the whole hierarchy collapses.

Any countable set is $\mathbf{\Sigma}^0_2$, since a singleton set is closed, and a countable set is a countable union of singletons. In a perfect Polish space, we can find countable sets that are neither open nor closed. The complements of such sets then provide examples of $\mathbf{\Pi}^0_2$ sets that are neither open nor closed, showing that the first two levels of the Borel hierarchy are proper for perfect Polish spaces.

Using the concept of **Baire category**, we will later show that the rationals $\mathbb{Q}$ are $\mathbf{\Sigma}^0_2$ but not $\mathbf{\Pi}^0_2$, thereby separating $\mathbf{\Sigma}^0_2$ and $\mathbf{\Pi}^0_2$.

It is much harder to find specific examples for the higher levels, e.g. a $\mathbf{\Sigma}^0_5$ set that is not $\mathbf{\Sigma}^0_4$. This separation will be much facilitated by the introduction of a **definability framework** for the Borel sets. Therefore, we defer the proof of the strong hierarchy theorem for a while.

Examples of Borel sets – continuity points of functions

Theorem 0.4.3 (Young). *Let $f : X \to Y$ be a mapping between metric spaces. Then*

$$C_f = \{x \colon f \text{ is continuous at } x\}$$

is a $\mathbf{\Pi}^0_2$ (i.e. G_δ) set.

Proof. The function f is continuous at a if and only if for any $\varepsilon > 0$,

$$\exists \delta > 0\, \forall x, y\, [x, y \in U_\delta(a) \;\Rightarrow\; d(f(x), f(y)) < \varepsilon]. \tag{$*$}$$

Given $\varepsilon > 0$, let

$$C_\varepsilon = \{a \colon \; (*) \text{ holds at } a \text{ for } \varepsilon\}.$$

We claim that C_ε is open. Suppose $a \in C_\varepsilon$. Choose a suitable δ that witnesses that $a \in C_\varepsilon$. We show $U_\delta(a) \subseteq C_\varepsilon$. Let $b \in U_\delta(a)$. Choose δ^* so that $U_{\delta^*}(b) \subseteq U_\delta(a)$. Then

$$x, y \in U_{\delta^*}(b) \;\Rightarrow\; x, y \in U_\delta(a) \;\Rightarrow\; d(f(x), f(y)) < \varepsilon.$$

Notice further that $\varepsilon > \varepsilon^*$ implies $C_\varepsilon \supseteq C_{\varepsilon^*}$. Hence we can represent C_f as

$$C_f = \bigcap_{n \in \mathbb{N}} C_{1/n},$$

a countable intersection of open sets.

$\square$

Here is a nice application of Young's theorem.

The function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 0 & x \text{ irrational}, \\ 1 & x = 0, \\ 1/q & x = p/q,\ p \in \mathbb{Z},\ q \in \mathbb{Z}^{>0},\ p, q \text{ relatively prime} \end{cases}$$

is a function that is continuous at every irrational, discontinuous at every rational number. How about the other way around – **discontinous at exactly the irrationals?** As noted above, the rationals are a $\mathbf{\Sigma}^0_2$ set that is not $\mathbf{\Pi}^0_2$. Hence **such a function cannot exist**.

We finish this lecture by showing that Young's Theorem can be reversed.

Theorem 0.4.4. *Given a $\mathbf{\Pi}^0_2$ subset A of a perfect Polish space X, there exists a mapping $f : X \to \mathbb{R}$ such that f is continuous at every point in A, and discontinuous at every other point, i.e. $C_f = A$.*

Proof. Fix a countable dense subset $D \subseteq X$. We first deal with the easier case that A is open. Let

$$f(x) = \begin{cases} 0 & x \in A \text{ or } x \in \neg\, \overline{A} \cap D, \\ 1 & \text{otherwise}. \end{cases}$$

It is clear that f is continuous on A. Now assume $x \notin A$. If $x \notin \overline{A}$, then there exists $U_\varepsilon(x) \subseteq \neg\, \overline{A}$. Any $U_{\varepsilon^*}(x) \subseteq U_\varepsilon(x)$ contains points from both D and $\neg\, D$, so it is clear that f is not continuous at x. Finally, let $x \in \overline{A} \setminus A$. Then $f(x) = 1$, but points of A are arbitrarily close, where f takes value 0.

Now we extend this approach to general $\mathbf{\Pi}^0_2$ sets. Suppose

$$A = \bigcap_n G_n, \qquad G_n \text{ open.}$$

By replacing G_n with $G_n^* = G_1 \cap \cdots \cap G_n$, we can assume that

$$X = G_0 \supseteq G_1 \supseteq G_2 \supseteq G_3 \supseteq \ldots$$

The idea is to define f_n as above for each G_n and then "amalgamate" the f_n in a suitable way. Assume for each n, $f_n : X \to \mathbb{R}$ is defined as above such that $C_{f_n} = G_n$. Let (b_n) be a sequence of positive real numbers such that for all n,

$$b_n > \sum_{k>n} b_k,$$

for example, $b_n = 1/n!$. We now form the series

$$f(x) = \sum_n b_n f_n(x).$$

Since $|f_n(x)| \leq 1$, $|f(x)| \leq \sum_n b_n < \infty$. Furthermore, (f_n) converges uniformly to f, for

$$|f(x) - f_n(x)| \leq \sum_{k>n} b_k < b_n,$$

and the last bound is independent of x and converges to 0.

It follows by uniform convergence that if each f_n is continuous at x, f is continuous on x, too. Hence f is continuous on A.

Now assume $x \notin A$. Then there exist n such that $x \in G_n \setminus G_{n+1}$. Hence

$$f_0(x) = \cdots = f_n(x) = 0.$$

Again, we distinguish two cases.

First, assume $x \notin \overline{G_{n+1}}$. Then there exists $\delta > 0$ such that $U_\delta(x) \subseteq \neg\, G_{n+1}$. This also implies $U_\delta(x) \subseteq \neg\, G_k$ for any $k \geq n+1$. Besides, since G_n is open, we can chose δ sufficiently small so that $U_\delta(x) \subseteq G_n$. For $y \in \neg\, D \cap U_\delta(x)$ we have $f_k(y) = 1$ for all $k \geq n+1$, and hence $f(y) = \sum_{k>n} b_k f_k(y) > 0$. On the other hand, if $y \in D \cap U_\delta(x)$, then $f_k(y) = 0$ for all $k \geq n+1$, and also $f_0(y) = \cdots = f_n(y) = 0$, since $y \in G_n$, and thus $f(y) = 0$. Hence there are points arbitrarily close to x whose f-values differ by a constant lower bound, which implies f is not continuous in x.

Finally, suppose $x \in \overline{G_{n+1}}$. Then $f_{n+1}(x) = 1$ and hence $f(x) \geq b_{n+1} > 0$. On the other hand, for any $y \in G_{n+1}$, $f(y) \leq \sum_{k>n+1} b_k < b_{n+1} = f(x)$. That is, there are points arbitrarily close to x whose f-value differs from $f(x)$ by a constant lower bound. Hence f is discontinuous at x in this case, too.

$\square$

0.4.2 Subspaces of Polish Spaces

Closed subsets of Polish spaces (with the subspace topology) are Polish (Proposition 0.3.1).

What about open subsets like $(0,1) \subset \mathbb{R}$? It is clear from this example that we have to find a different compatible metric.

Proposition 0.4.1. *Any open subset of a Polish space X is Polish.*

Proof. Let $U \subset X$ be open, where we assume that $U \neq X$. Consider the set

$$F = \{(t,x) \in \mathbb{R} \times X \colon\; t \cdot d(x, X \setminus U) = 1\}. \tag{66}$$

The mapping $x \mapsto d(x, X \setminus U)$ from X to $\mathbb{R}$ is continuous. Therefore, F is closed and thus Polish.

If we restrict the projection $\pi_2 : \mathbb{R} \times X \to X$ to F, we obtain a homeomorphism between F and U. As homeomorphic images of Polish spaces are Polish (Proposition 0.3.1), U is Polish.

□

Exercise

Give an alternative proof of the preceding result by considering the following:

Let d be a compatible metric on X and define

$$\overline{d}(x,y) = \frac{d(x,y)}{1 + d(x,y)}. \tag{67}$$

Show that this is a metric that induces the same topology.

Now let

$$d_U(x,y) = \overline{d}(x,y) + \left| \frac{1}{\overline{d}(x, X \setminus U)} - \frac{1}{\overline{d}(y, X \setminus U)} \right|. \tag{68}$$

Verfiy that this is a metric on U compatible with the subspace topology with respect to which U is complete.

Proposition 0.4.2. *Let X be a Polish space, and suppose (Y_n) is a sequence of Polish subspaces of X. Then $\bigcap_n Y_n$ is a Polish subspace of X.*

Proof. Consider the mapping $f : X \to X^{\mathbb{N}}$ given by $x \mapsto (x,x,x,\dots)$. The restriction of f to $\bigcap_n Y_n$ is a homeomorphism between $\bigcap_n Y_n$ and the diagonal $\Delta \subseteq \prod_n Y_n$,

$$\Delta = \{(x,x,x,\dots) \colon\; x \in Y_n \text{ for all } n \in \mathbb{N}\}. \tag{69}$$

Δ is closed in the product space $\prod_n Y_n$ and hence Polish, and this property pushes over to $\bigcap_n Y_n$ (see Proposition 0.3.1).

□

Hence every G_δ subset of a Polish space is Polish. This is as far as we can get.

Theorem 0.4.5 (Mazurkiewicz)**.** *A subset of a Polish space is Polish if and only if it is G_δ.*

We have already established the "if" direction of this result. For the other direction, we need a lemma that is interesting in its own right.

Lemma 0.4.6 (Kuratowski extension lemma)**.** *Suppose X, Y are Polish spaces, $A \subseteq X$, and $f : A \to Y$ continuous. Then there exists a G_δ set G with $A \subseteq G \subseteq \overline{A}$ and a continuous extension $g : G \to Y$ of f.*

Compare this with the last lecture, where we showed that the points of continuity of a function is always a G_δ set (Theorem **??**).

Proof. We can adapt the *ε-oscillation set* C_ε used in the proof of Theorem **??** to the domain A:

$$C_\varepsilon^A = \{x \in X \colon \exists \delta > 0\, \forall a, b \in A\, [a, b \in U_\delta(x) \Rightarrow d(f(a), f(b)) < \varepsilon]\}. \tag{$*$}$$

As before, C_ε^A is open and hence

$$G = \overline{A} \cap \bigcap_n C_{1/n}^A \tag{70}$$

is G_δ and since f is continuous, $A \subseteq G \subseteq \overline{A}$.

To extend f to G, let $x \in G$. Since $x \in \overline{A}$, there exists (a_n) in A with $x = \lim_n a_n$. As $x \in \bigcap_n C_{1/n}^A$, $(f(a_n))$ is Cauchy. Y is complete, so there exists $y = \lim_n f(a_n) \in Y$. It is straightforward to verify that y is independent of the choice of (a_n) and agrees with $f(x)$ for $x \in A$. Hence we can put

$$g(x) = y, \tag{71}$$

which yields the desired continuous extension.

□

Now assume $Y \subset X$ is Polish but not G_δ. Then, by the previous lemma, the identity mapping $\mathrm{id} : Y \to Y$ has a proper continuous extension $g : G \to Y$ to a G_δ set G with $Y \subsetneq G \subseteq \overline{Y}$. Let $x \in G \setminus Y$. Y is dense in G, so there exists (y_n) in Y with $x = \lim_n y_n$. By continuity

$$x = \lim_n y_n = \lim_n g(y_n) = g(x) \in Y, \tag{72}$$

contradiction. This completes the proof of Theorem **??**.

Borel set as clopen sets

More complicated Borel sets in Polish spaces are not Polish anymore in the subspace topology, as we just saw. But what if we are allowed to change the topology? In the process, we would like to "preserve" as much as possible of the original space. It turns out we can change the topology so that a given Borel set becomes clopen while inducing the same family of Borel sets overall.

We start with closed sets.

Lemma 0.4.7. *If X is a Polish space with topology $\mathcal{O}$, and $F \subseteq X$ is closed, then there exists a finer topology $\mathcal{O}' \supseteq \mathcal{O}$ such that $\mathcal{O}$ and $\mathcal{O}'$ give rise to the same class of Borel sets in X, and F is clopen with respect to $\mathcal{O}'$.*

Proof. By Proposition 0.3.1 and Proposition **??**, respectively, F and $X \setminus F$ are Polish spaces with compatible metrics d_F and $d_{X\setminus F}$, respectively. Wlog $d_F, d_{X\setminus F} < 1$. We form the *disjoint union of the spaces F and $X \setminus F$*: This is the set $X = F \sqcup X \setminus F$ with the following topology, $\mathcal{O}'$. $U \subseteq F \sqcup X \setminus F$ is in $\mathcal{O}'$ if and only if $U \cap F$ is open (in F) and $U \cap X \setminus F$ is open (in $X \setminus F$).

The disjoint union is Polish, as witnessed by the following metric.

$$d_\sqcup(x, y) = \begin{cases} d_F(x, y) & \text{if } x, y \in F, \\ d_{X\setminus F}(x, y) & \text{if } x, y \in X \setminus F, \\ 2 & \text{otherwise.} \end{cases} \tag{73}$$

It is straightforward to check that d is compatible with $\mathcal{O}'$. Furthermore, let (x_n) be Cauchy in $(X, d_\sqcup)$. Then the x_n are completely in F or in $X \setminus F$ from some point on, and hence (x_n) converges.

Under the disjoint union topology, F is is clopen. Moreover, an open set in this topology is a disjoint union of an open set in $X \setminus F$, which also open the original topology $\mathcal{O}$, and an intersection of an open set from $\mathcal{O}$ with F. Such sets are Borel in $(X, \mathcal{O})$, hence $(X, \mathcal{O})$ and $(X, \mathcal{O}')$ have the same Borel sets.

□

Theorem 0.4.8. *Let X be a Polish space with topology $\mathcal{O}$, and suppose $B \subseteq X$ is Borel. Then there exists a finer Polish topology $\mathcal{O}' \supseteq \mathcal{O}$ such that $\mathcal{O}$ and $\mathcal{O}'$ give rise to the same class of Borel sets in X, and B is clopen with respect to $\mathcal{O}'$.*

Proof. Let $\mathcal{S}$ be the family of all subsets A of X for which a finer topology exists that has the same Borel sets as $\mathcal{O}$ and in which A is clopen.

We will show that $\mathcal{S}$ is a σ-algebra, which by the previous Lemma contains the closed sets. Hence $\mathcal{S}$ must contain all Borel sets, and we are done.

$\mathcal{S}$ is clearly closed under complements, since the complement of a clopen set is clopen in any topology.

So assume now that $\{A_n\}$ is a countable family of sets in $\mathcal{S}$. Let $\mathcal{O}_n$ be a topology on X that makes A_n clopen and does not introduce new Borel sets.

Let $\mathcal{O}_\infty$ be the topology generated by $\bigcup_n \mathcal{O}_n$. Then $\bigcup_n A_n$ is open in $(X, \mathcal{O}_\infty)$, and we can apply Lemma 0.4.7. For this to work, however, we have to show that $(X, \mathcal{O}_\infty)$ is Polish and does not introduce any new Borel sets.

We know that the product space $\prod(X, \mathcal{O}_n)$ is Polish. Consider the mapping $\phi : X \to \prod_n X$

$$x \mapsto (x, x, x, \dots). \tag{74}$$

Observe that ϕ is a continuous mapping between $(X, \mathcal{O}_\infty)$ and $\prod_n X$. The preimage of a basic open set $U_1 \times U_2 \times \cdots \times U_n \times X \times X \times \cdots$ under ϕ is just the intersection of the U_i. Furthermore, ϕ is clearly one-to-one, and the inverse mapping between $\phi(X)$ and X is continuous, too.

If we can show that $\phi(X)$ is closed in $\prod_n X$, we know it is Polish as a closed subset of a Polish space, and since $(X, \mathcal{O}_\infty)$ is homeomorphic to $\phi(X)$, we can conclude it is Polish.

To see that $\phi(X)$ is closed in $\prod_n X$, let $(y_1, y_2, y_3, \dots) \in \neg\,\phi(X)$. Then there exist $i < j$ such that $y_i \neq y_j$. Since $(X, \mathcal{O})$ is Polish, we can pick U, V open, disjoint such that $y_i \in U$, $y_j \in V$. Since each $\mathcal{O}_n$ refines $\mathcal{O}$, U is open in $\mathcal{O}_i$, and V is open in $\mathcal{O}_j$. Therefore,

$$X_1 \times X_2 \times \cdots \times X_{i-1} \times U \times X_{i+1} \times \cdots \times X_{j_1} \times V \times X_{j+1} \times X_{j+2} \times \cdots \tag{75}$$

where $X_k = X$ for $k \neq i, j$, is an open neighborhood of $(y_1, y_2, y_3, \dots)$ completely contained in $\neg\,\phi(X)$.

Finally, too see that the Borel sets of $(X, \mathcal{O}_\infty)$ are the same as the ones of $(X, \mathcal{O})$, for each n, let $\{U_i^{(n)}\}_{i \in \mathbb{N}}$ be a basis for $\mathcal{O}_n$. By assumption, all sets in $\mathcal{O}_n$ are Borel sets of $(X, \mathcal{O})$. The set $\{U_i^{(n)}\}_{i,n \in \mathbb{N}}$ is a subbasis for $\mathcal{O}_\infty$. This means that any open set in $(X, \mathcal{O}_\infty)$ is a countable union of finite intersections of the $U_i^{(n)}$. Since every $U_i^{(n)}$ is Borel in $(X, \mathcal{O})$, this means that any open set in $\mathcal{O}_\infty$ is Borel in $(X, \mathcal{O})$. Since the Borel sets are closed under complementation and countable unions, this in turn implies that every Borel set of $(X, \mathcal{O}_\infty)$ is already Borel in $(X, \mathcal{O})$.

□

Corollary 0.4.8.1 (Perfect subset property for Borel sets; Alexandroff, Hausdorff)**.** *In a Polish space, every uncountable Borel set has a perfect subset.*

Proof. Let $(X, \mathcal{O})$ be Polish, and assume $B \subseteq X$ is Borel. We can choose a finer topology $\mathcal{O}' \supseteq \mathcal{O}$ so that B becomes clopen, but the Borel sets stay the same. By Theorem **??**, B is Polish with respect to the subspace topology $\mathcal{O}'|_B$

Suppose B is uncountable. By Theorem **??** there exists a continuous injection f from $2^{\mathbb{N}}$ (with respect to the standard topology) into $(B, \mathcal{O}'|_B)$.

Since $\mathcal{O}'$ is finer than $\mathcal{O}$, f is continuous with respect to $\mathcal{O}$, too. Since $2^{\mathbb{N}}$ is compact, $f(2^{\mathbb{N}})$ is closed with respect to $\mathcal{O}$. Finally, $f(2^{\mathbb{N}})$ has no isolated points with respect to $\mathcal{O}'$, which then also holds for the coarser topology $\mathcal{O}$.

Therefore, B has a perfect subset.

□

0.4.3 Measure and Baire Category

At the end of the previous section, we saw that Borel sets are well-behaved in the sense that they possess the **perfect subset property**. Two other important regularity properties are **measurability** and the **Baire property**, which we will introduce in this section.

Filters and Ideals

The most common measure of size is, of course, cardinality. In the presence of uncountable sets (like in a perfect Polish space), the usual division is between countable and uncountable sets. The smallness of the countable sets is reflected, in particular, by two properties: A subset of a countable set is countable, and countable unions of countable set are countable. These characteristics are shared with other notions of smallness, two of which we will encounter in this lecture.

Definition 0.4.4. *A non-empty family $\mathcal{I} \subseteq \mathcal{P}(X)$ of subsets of a given set X is an **ideal** if*

$$\text{(I1)} \qquad A \in \mathcal{I} \text{ and } B \subseteq A \text{ implies } B \in \mathcal{I},$$
$$\text{(I2)} \qquad A, B \in \mathcal{I} \text{ implies } A \cup B \in \mathcal{I}.$$

If we have closure even under *countable unions*, we speak of a **σ-ideal**. For example, while the countable sets in $\mathbb{R}$ form a σ-ideal, the finite subsets only form an ideal.

Another example of ideals are the so-called **principal ideals**. These are ideals of the form

$$\langle Z \rangle = \{A \colon A \subseteq Z\} \tag{76}$$

for a fixed $Z \subseteq X$.

The **dual notion** to an ideal is that of a **filter**. It reflects that the sets in a filter share some **largeness property**.

Definition 0.4.5. *A non-empty family $\mathcal{F} \subseteq \mathcal{P}(X)$ of subsets of a given set X is a **filter** if*

$$\text{(F1)} \qquad A \in \mathcal{F} \text{ and } B \supseteq A \text{ implies } B \in \mathcal{F},$$
$$\text{(F2)} \qquad A, B \in \mathcal{F} \text{ implies } A \cap B \in \mathcal{F}.$$

Again, closure under countable intersections yields **σ-filters**.

If $\mathcal{I}$ is a (σ-) ideal, then $\mathcal{F} = \{\neg A \colon A \in \mathcal{I}\}$ is a (σ-) filter. Hence the co-finite subsets of $\mathbb{R}$ form a filter, and the co-countable subsets form a σ-filter.

Note that the complement of a (σ-) ideal (in $\mathcal{P}(X)$) is not necessarily a (σ-) filter. This is true, however, for a special class of ideals/filters.

Definition 0.4.6. *A non-empty family $\mathcal{I} \subseteq \mathcal{P}(X)$ is a **prime ideal** if it is an ideal for which*

> *for every $A \in X$, either $A \in \mathcal{I}$ or $\neg A \in \mathcal{I}$ (but not both).*

*An **ultrafilter** is a filter whose complement in $\mathcal{P}(X)$ is a prime ideal.*

In light of the small-/largeness motivation, prime ideals and ultrafilters provide a *complete* separation of X: Each set is either small or large.

Measures

Coarsely speaking, a measure assigns a size to a set in a way that reflects our basic geometric intuition about sizes: The size of the union of disjoint objects is the sum of their sizes. The question whether this can be done in a consistent way for *all* subsets of a given space is of fundamental importance and has motivated many questions in set theory.

The formally, a measure μ on X is a $[0, \infty]$-valued function defined on subsets of X that satisfies

$$\begin{aligned}
&\text{(M1)} \qquad \mu(\emptyset) = 0 \\
&\text{(M2)} \qquad \mu\Big(\bigcup_n A_n\Big) = \sum_n \mu(A_n), \\
&\qquad\qquad \text{whenever the } A_n \text{ are pairwise disjoint.}
\end{aligned}$$

The question is, of course, which subsets of X can be assigned a measure. The condition (M2) suggests that this family is closed under countable unions. Furthermore, if $A \subseteq X$, then the equation $\mu(X) = \mu(A) + \mu(\neg A)$ suggests that $\neg A$ should be measurable, too. In other words, the sets who are assigned a measure form a σ-algebra.

Definition 0.4.7. *A **measurable space** is a pair $(X, \mathcal{S})$, where X is a set and $\mathcal{S}$ is a σ-algebra on X. A **measure** on a measurable space $(X, \mathcal{S})$ is a function $\mu : \mathcal{S} \to [0, \infty]$ that satisfies (M1) and (M2) for any pairwise disjoint family $\{A_n\}$ in $\mathcal{S}$. If μ is a measure on $(X, \mathcal{S})$, then the triple $(X, \mathcal{S}, \mu)$ is called a **measure space**.*

If we want the measure μ to reflect also some other basic intuition about geometric sizes, this often puts restrictions on the σ-algebra of measurable sets. For example, in $\mathbb{R}$ the measure of an interval should be its *length.* We will see later that, if we assume the Axiom of Choice, it is impossible to assign every subset of $\mathbb{R}$ a measure, so that (M1) and (M2) are satisfied, and the measure of an interval is its length.

To have some control over what the σ-algebra of measurable sets should be, one can construct a measure more carefully, start with a measure on basic objects such as intervals or balls, and then extend it to larger classes of sets by approximation.

An essential component in this extension process is the concept of an **outer measure**.

Definition 0.4.8. *An **outer measure** on a set X is a function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ such that*

$$\begin{aligned}
&\text{(O1)} \qquad \mu^*(\emptyset) = 0, \\
&\text{(O2)} \qquad A \subseteq B \textit{ implies } \mu^*(A) \leq \mu^*(B), \\
&\text{(O3)} \qquad \mu^*\Big(\bigcup_n A_n\Big) \leq \sum_n \mu^*(A_n), \\
&\qquad\qquad \textit{for any countable family } \{A_n\} \textit{ in } X.
\end{aligned}$$

An outer measure hence weakens the conditions of **additivity** (M2) to **subadditivity** (O3). This makes it possible to have non-trivial outer measures that are defined on *all* subsets of X.

The usefulness of outer measures lies in the fact that they can always be restricted to subset of $\mathcal{P}(X)$ on which they behave as measures.

Definition 0.4.9. *Let μ^* be an outer measure on X. A set $A \subseteq X$ is μ^*-**measurable** if*

$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \setminus A) \quad \textit{for all } B \subseteq X. \tag{77}$$

This definition is justified rather by its consequences than by its intuitive appeal. Regarding the latter, suffice it to say here that outer measures may be rather far from being even *finitely* additive. The definition singles out those sets that split all other sets correctly, with regard to measure.

Proposition 0.4.3. *The class of μ^*-measurable sets forms a σ-algebra $\mathcal{M}$, and the restriction of μ^* to $\mathcal{M}$ is a measure.*

A proof can be found in any standard book on measure theory, for instance **?** or **?**.

The size of the σ-algebra of measurable sets depends, of course, on the outer measure μ^*. If μ^* is behaving rather pathetically, we cannot expect $\mathcal{M}$ to contain many sets.

Lebesgue measure

A standard way to obtain "nice" outer measures is to start with a well-behaved function defined on a certain class of sets, and then approximate. The paradigm for this approach is the construction of **Lebesgue measure** on $\mathbb{R}$.

Definition 0.4.10. *The **Lebesgue outer measure** λ^* of a set $A \subseteq \mathbb{R}$ is defined as*

$$\lambda^*(A) = \inf\left\{\sum_n |b_n - a_n| \colon A \subseteq \bigcup_n (a_n, b_n)\right\}. \tag{78}$$

> **Exercise**
> Show that λ^* indeed defines an outer measure.

We call the λ^*-measurable sets **Lebesgue measurable**.
The following two facts are also standard **?**.

Proposition 0.4.4. *If $I \subseteq \mathbb{R}$ is an interval, then $\lambda^*(I)$ is equal to the length of I (possibly infinite).*

Proposition 0.4.5. *Any interval $I \subseteq \mathbb{R}$ is Lebesgue measurable.*

Corollary 0.4.8.2. *Any Borel set in $\mathbb{R}$ is Lebesgue measurable*

Proof. This follows from Proposition 0.4.3, Proposition 0.4.5 and the fact that any open set in $\mathbb{R}$ is a countable union of intervals.

□

The construction of Lebesgue measure can be generalized and extended to other metric spaces, for example through the concept of **Hausdorff measures**.

All these measures are **Borel measures**, in the sense that the Borel sets are measurable. However, there are measurable sets that are not Borel sets. The reason for this lies in the presence of **nullsets**, which are measure theoretically "easy" (since they do not contribute any measure at all), but can be topologically quite complicated.

Nullsets

Let μ^* be an outer measure on X. If $\mu^*(A) = 0$, then A is called a μ^***-nullset**.

Proposition 0.4.6. *Any μ^*-nullset is μ^*-measurable.*

Proof. Suppose $\mu^*(A) = 0$. Let $B \subseteq X$. Then, since μ^* is subadditive and monotone,

$$\mu^*(B) \leq \mu^*(B \cap A) + \mu^*(B \cap \neg A) = \mu^*(B \cap \neg A) \leq \mu^*(B), \tag{79}$$

and therefore $\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap \neg A)$.

□

The next result confirms the intuition that nullsets are a notion of smallness.

Proposition 0.4.7. *The μ^*-nullsets form a σ-ideal.*

Proof. (I1) follows directly from monotonicity (O2). Countable additivity follows immediately from subadditivity (O3).

□

In case of Lebesgue measure, we can use Proposition Proposition 0.4.6 to further describe the Lebesgue measurable subsets of $\mathbb{R}$.

Proposition 0.4.8. *A set $A \subseteq \mathbb{R}$ is Lebesgue measurable if and only if it is the difference of a $\mathbf{\Pi}^0_2$ set and a Lebesgue nullset.*

Proof. We first assume $\lambda^*(A) < \infty$. Let $G_n \subseteq \mathbb{R}$ be an open set such that $G_n \supseteq A$ and $\lambda^*(G_n) \leq \lambda^*(A) + 1/n$. The existence of such a G_n follows from the definition of λ^*, and the fact that every open set is the disjoint union of open intervals. Then $G = \bigcap_n G_n$ is $\mathbf{\Pi}^0_2$, $A \subseteq G$, and for all n,

$$\lambda^*(A) \leq \lambda^*(G) \leq \lambda^*(A) + 1/n \tag{80}$$

hence $\lambda^*(A) = \lambda^*(G)$. Hence for $N = G \setminus A$, since A is measurable,

$$\lambda^*(N) = \lambda^*(G) - \lambda^*(A) = 0 \quad \text{and} \quad A = G \setminus N. \tag{81}$$

If $\lambda^*(A) = \infty$, we set $A_m = A \cap [m, m+1)$ for $m \in \mathbb{Z}$. By monotonicity, each $\lambda^*(A_m)$ is finite. For each $m \in \mathbb{Z}$, $n \in \mathbb{N}$, pick $G_n^{(m)}$ open such that $\lambda^*(G_n^{(m)}) \leq \lambda^*(A) + 1/2^{n+2|m|+1}$. Then, with

$$\bigcap_{n\in\mathbb{N}} \bigcup_{m\in\mathbb{Z}} G_n^{(m)}, \tag{82}$$

$N = G \setminus A$ is the desired set.

For the other direction, note that the measurable sets form a σ-algebra which contains both the Borel sets and the nullsets. Hence any set that is the difference of a Borel set and a nullset is measurable, too.

□

> **Exercise**
> Show that each Lebesgue measurable set can be written as a disjoint union of a $\mathbf{\Sigma}^0_2$ set and a nullset.

Hence if a set is measurable, it differs from a (rather simple) Borel set only by a nullset.

We also obtain the following characterization of the σ-algebra of Lebesgue measurable sets.

Proposition 0.4.9. *The σ-algebra of Lebesgue measurable sets in $\mathbb{R}$ is the smallest σ-algebra containing the open sets and the nullsets.*

As mentioned before, there are Lebesgue measurable sets that are not Borel sets. We will eventually encounter such sets. The question which sets exactly are Lebesgue measurable was one of the major questions that drove the development of descriptive set theory, just like the question which uncountable sets have perfect subsets.

Baire category

The basic paradigm for smallness here is of topological nature. A set is small if it does not look anything like an open set, not even under closure. In the following, let X be a Polish space.

Definition 0.4.11. *A set $A \subseteq X$ is **nowhere dense** if its complement contains an open, dense set.*

Being nowhere dense means for any open set $U \subseteq X$ we can find a non-empty open subset $V \subseteq U$ such that $V \subseteq \neg A$. In other words, a nowhere dense set is "full of holes".

Examples of nowhere dense sets are all finite, or more generally, all discrete subsets of a perfect Polish space, i.e. sets all whose points are isolated. There are non-discrete nowhere dense sets, such as $\{0\} \cup \{1/n : n \in \mathbb{N}\}$ in $\mathbb{R}$, even uncountable ones, such as the middle-third Cantor set.

The nowhere dense sets form an ideal, but not a σ-ideal: Every singleton set is nowhere dense, but there are countable sets that are not, such as the rationals $\mathbb{Q}$ in $\mathbb{R}$.

To obtain a σ-ideal, we close the nowhere dense sets under countable unions.

Definition 0.4.12. *A set $A \subseteq X$ is **meager** or of **first category** if it is the countable union of nowhere dense sets. Non-meager sets are also called sets of **second category**. Complements of meager sets are called **comeager** or **residual**.*

The meager subsets of X form a σ-ideal. Examples of meager sets are all countable sets, but there are uncountable ones (Cantor set).

The concept of Baire category is often used in existence proofs: To show that a set with a certain property exists, one shows that the set of points *not having the property* is meager. A famous example is Banach's proof of the existence of continuous, nowhere differentiable functions. For this to work, of course, we have to ensure that the complements of meager sets are non-empty.

Theorem 0.4.9 (Baire Category Theorem)**.** *For any Polish space X, the following statements hold.*

- ***(a)*** *For every meager set $M \subseteq X$, the complement $\neg M$ is dense in X.*
- ***(b)*** *No non-empty open set is meager.*
- ***(c)*** *If $\{D_n\}$ is a countable family of open, dense sets, then $\bigcap_n D_n$ is dense.*

Proof. (a) Assume $M = \bigcup_n N_n$, where each N_n is nowhere dense. Then $\neg M = \bigcap D_n$, where each D_n contains a dense, open set. Let $U \subseteq X$ be open.

We construct a point $x \in U \cap \neg M$ by induction. We can find an open ball B_1 of radius < 1 such that $\overline{B_1} \subseteq U \cap D_1$, since D_1 contains a dense open set. In the next step, we use the same property of D_2 to find an open ball B_2 of radius $< 1/2$ whose closure is completely contained in $B_1 \cap D_2$. Continuing inductively, we obtain a nested sequence of balls B_n of radius $< 1/n$ such that $\overline{B_n} \subseteq B_{n-1} \cap D_n$.

Let x_n be the center of B_n. Then (x_n) is a Cauchy sequence, so $x = \lim_n x_n$ exists in X. Since for any n, all but finitely many x_i are in B_n, we have $x \in \overline{B_n}$ for all n. Therefore, by construction

$$x \in \bigcap_n \overline{B_n} = \bigcap_n B_n \subseteq U \cap \bigcap_n D_n \subseteq U. \tag{83}$$

(b) follows immediately from (a), the proof of (c) is exactly the same as that for (a). In fact, the three statements are equivalent.

□

Any topological space that satisfies the three equivalent conditions (a)-(c) is called a **Baire space** (not to be confused with *the* Baire space $\mathbb{N}^{\mathbb{N}}$ – the latter is, of course, a Baire space, too).

As an application, we determine the exact location of $\mathbb{Q}$ in the Borel hierarchy of $\mathbb{R}$.

Corollary 0.4.9.1. $\mathbb{Q}$ *is not a* $\mathbf{\Pi}^0_2$ *set, hence a true* $\mathbf{\Sigma}^0_2$ *set.*

Proof. Note that $\mathbb{R}$ cannot be meager, by (b). Since $\mathbb{Q}$ is meager, $\mathbb{R} \setminus \mathbb{Q}$ cannot be meager either. If $\mathbb{Q}$ were a $\mathbf{\Pi}^0_2$ set, it would be the intersection of open, dense sets and hence its complement $\mathbb{R} \setminus \mathbb{Q}$ would be meager.

□

The Baire property

We have seen that the measurable sets are precisely the ones that differ from a $\mathbf{\Pi}^0_2$ set by a nullset. We can introduce a similar concept for Baire category.

Definition 0.4.13. *A set $B \subseteq X$ has the* ***Baire property*** *if there exists an open set G and a meager set M such that*

$$B \triangle G = M, \tag{84}$$

where $\triangle$ denotes the symmetric difference between two sets: $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

Exercise

Show that $\triangle$ is commutative, associative, and satisfies the distributive law

$$A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C). \tag{85}$$

In the above definition, we can replace open sets by closed sets.

Lemma 0.4.10. *A set B has the Baire property if and only if it can be represented in the form $B = F \triangle M$, where F is closed and M is meager.*

Proof. Suppose $B = G \triangle M$, G open and M meager.

Then $N = \overline{G} \setminus G$ is nowhere dense and closed. Furthermore, $Q = M \triangle N$ is meager (it is the union of two meager sets). We easily verify that $G = \overline{G} \triangle N$, and therefore

$$B = G \triangle M = (\overline{G} \triangle N) \triangle M = \overline{G} \triangle (N \triangle M) = \overline{G} \triangle Q, \tag{86}$$

as desired.

The converse direction is similar, using the interior instead of the closure.

□

Proposition 0.4.10. *The sets having the Baire property form a σ-algebra.*

Proof. To show closure under complement, note that $\neg (A \triangle B) = \neg A \triangle B$. Therefore, if $B = G \triangle M$ with G open and M meager, we have $\neg B = \neg G \triangle M$, and we can use Lemma **??**.

Now assume $B = \bigcup B_i$, and for each i there exist open G_i and meager M_i such that $B_i = G_i \triangle M_i$.

Let $G = \bigcup G_i$ and $M = \bigcup M_i$. Then G is open and M is meager (since the meager sets for a σ-ideal).

We easily check that

$$G \setminus M \subseteq B \subseteq G \cup M. \tag{87}$$

This implies $B \triangle G \subseteq M$ and hence $B \triangle G$ is meager.

Since

$$B = G \triangle (B \triangle G), \tag{88}$$

we conclude that B has the Baire property.

□

Corollary 0.4.10.1. *The σ-algebra of sets having the Baire property is the smallest σ-algebra containing all open and all meager sets.*

Exercise

Show that B has the Baire property if and only if it can be represented as a G_δ set plus a meager set.

As in the case of measure, there exist non-Borel sets with the Baire property, and using the Axiom of Choice one can show that there exists set that do not have the Baire property.

We conclude this lecture with a note on the relationship between measure and category. From the results so far it seems that they behave quite similarly. This might lead to the conjecture that maybe they more or less coincide. This is not so, in fact, they are quite orthogonal to each other, as the next result shows.

Proposition 0.4.11. *The real numbers can be partitioned into two subsets, one a Lebesgue nullset and the other one meager.*

Proof. Let (G_n) be a sequence of open sets witnessing that $\mathbb{Q}$ is a nullset, i.e. each G_n is a union of disjoint open intervals that covers $\mathbb{Q}$ and whose total length does not exceed 2^{-n}. Then $G = \bigcap_n G_n$ is a nullset, but at the same time it is an intersection of open dense sets, thus comeager, hence its complement is meager.

□

0.4.4 The Axiom of Choice

In the previous lectures, a number of **regularity principles** for sets of real numbers emerged:

- **(PS)** the **perfect subset property**,
- **(LM)** **Lebesgue measurability**,
- **(BP)** the **Baire property**.

We have seen that the Borel sets in $\mathbb{R}$ have all these properties. In this lecture we will show how to construct counterexamples for each of these principles. The proofs make essential use if the **Axiom of Choice**:

> **(AC)** Every set $\mathcal{X}$ of non-empty sets has a choice-function.

A **choice function** for $\mathcal{X}$ is a function f that assigns every set $Y \in \mathcal{X}$ an element $y \in Y$.

One of the most famous applications of the Axiom of Choice is Vitali's construction of a non-Lebesgue measurable set.

Theorem 0.4.11 (Vitali)**.** *There exists a set $A \subseteq \mathbb{R}$ that is not Lebesgue measurable.*

Proof. Put

$$x \sim y \text{ if and only if } x - y \in \mathbb{Q}.$$

It is straightforward to check that this is an equivalence relation on $\mathbb{R}$. Using a choice function on the equivalence classes of $\sim$ intersected with the unit interval $[0,1]$, we pick from each equivalence class a representative from $[0,1]$, and collect them in a set S.

If we let, for $r \in \mathbb{Q}$,

$$S_r = \{s + r \colon s \in S\},$$

then

$$S_r \cap S_t \quad \text{for } r \neq t. \tag{89}$$

Suppose S is measurable. Then so is each S_r, and $\lambda(S_r) = \lambda(S)$.

If $\lambda(S) = 0$, then $\lambda(\mathbb{R}) = 0$, which is impossible. On the other hand, if $\lambda(S) > 0$, then, by countable additivity,

$$2 = \lambda([0,2]) \geq \lambda\left(\bigcup_{r \in \mathbb{Q} \cap [0,1]} S_r\right) = \sum_{r \in \mathbb{Q} \cap [0,1]} \lambda(S) = \infty,$$

contradiction.

□

The Axiom of Choice is equivalent to a number of other principles. We will use the **Well-ordering Principle**:

> **(WO)** Every set X can be well-ordered.

This means that one can define a binary relation $<$ on X so that every non-empty subset of X has a $<$-minimal element.

We use (WO) to construct a set $B \subseteq \mathbb{R}$ such neither B nor $\mathbb{R} \setminus B$ contains a perfect subset. Such sets are called **Bernstein sets**.

Theorem 0.4.12. *There exists a Bernstein set.*

Proof. Let $\mathcal{P}$ be the set of perfect subsets of $\mathbb{R}$. We can well-order this set, say

$$\mathcal{P} = \{P_\xi \colon \xi < 2^{\aleph_0}\}.$$

Note that every perfect subset corresponds to Cantor-Scheme, which can be coded by a real number (see **??**). Therefore, there are at most $2^{\aleph_0}$-many perfect subsets of $\mathbb{R}$, and it is not hard to see that there are exactly $2^{\aleph_0}$-many.

Furthermore, we assume each P_ξ is well-ordered.

Pick $a_0 \neq b_0$ from P_0. Assume we have chosen $\xi < 2^{\aleph_0}$, and $\{a_\beta \colon \beta < \xi\}$ and $\{b_\beta \colon \beta < \xi\}$ so that

$$a_\beta, b_\beta \in P_\beta \quad \text{and} \quad \text{all } a_\beta, b_\gamma \text{ pairwise distinct,}$$

we can choose $a_\xi, b_\xi \in P_\xi$ to be the first two elements of $P_\xi \setminus \bigcup_{\gamma<\xi}\{a_\gamma, b_\gamma\}$. This is possible since a perfect subset of $\mathbb{R}$ has cardinality $2^{\aleph_0}$, and $\xi < 2^{\aleph_0}$.

Put

$$A = \{a_\xi \colon \xi < 2^{\aleph_0}\} \qquad B = \{b_\xi \colon \xi < 2^{\aleph_0}\}.$$

Neither A nor B has a perfect subset by construction, and since $A \subseteq \mathbb{R} \setminus B$, B is a Bernstein set.

□

Proposition 0.4.12. *A Bernstein set does not have the Baire property.*

Proof. Assume for a contradiction a Bernstein set B has the Baire property. By an exercise in the previous chapter, we can write $B = M \cup G$, where M is meager and G is G_δ.

At least one of B, $\mathbb{R} \setminus B$ is not meager. Wlog assume B is not meager. (If not, obtain the representation "meager $\cup$ G_δ" above for $\mathbb{R} \setminus B$ and proceed analogously.) Then $G \subseteq B$ must be non-meager, too, and hence is an uncountable G_δ set. By Theorem **??**, G is Polish and hence must contain a perfect subset, contradiction.

□

Exercise

Show that a Bernstein set is not Lebesgue measurable.

The existence of arbitrary choice functions appears to be a rather strong assumption. It has consequences that seem paradoxical in the sense that they conflict with basic intuitions we have about objects and they behavior with respect to size or other characteristics. Arguably the most famous example is the **Banach-Tarski Paradox**, which uses the Axiom of Choice to partition a ball in $\mathbb{R}^3$ into finitely many pieces, and then, using rigid transformations (i.e.\ rotations and translations), to assemble them into two balls of the original size.

On the other hand, the Axiom of Choice implies or is even equivalent to many principles that are applied throughout many areas of mathematics, such as the existence of bases of vector spaces, Zorn's Lemma, Tychonoff's Theorem on the compactness of product spaces, the Hahn-Banach Theorem, or the Prime Ideal Theorem.

For some applications, however, a weaker form of the Axiom of Choice is sufficient.

The **Axiom of Countable Choice**:

> **(AC$_\omega$)** Every *countable* family $\mathcal{X}$ of non-empty sets has a choice-function.

Stronger than Countable Choice, but still weaker than the full Axiom of Choice is **Axiom of Dependent Choice**:

> **(DC)** If E is a binary relation on a non-empty set A, and if for every $a \in A$ there exists $b \in A$ such that $a\,E\,b$, then there exists a function $f : \mathbb{N} \to A$ such that for all $n \in \mathbb{N}$, $f(n)\;E\;f(n+1)$.

A seminal result by **?** showed that DC is no longer sufficient to prove the existence of non-regular sets in the above sense. He constructed (though under a large cardinal assumption) a model of ZF+DC in which every set of real numbers is Lebesgue measurable, has the Baire property, and has the perfect subset property.

0.4.5 Coding Borel Sets

In this chapter, we take a further look at Borel subsets of $\mathbb{N}^{\mathbb{N}}$. As is common in this setting, we call the elements of $\mathbb{N}^{\mathbb{N}}$ **reals**. This is motivated by the fact that, via the continued fration expansion, $\mathbb{N}^{\mathbb{N}}$ is homeomorphic to the set of irrational real numbers. Suppose $U \subseteq \mathbb{N}^{\mathbb{N}}$ is open. Then there exists a set $W \subseteq \mathbb{N}^{<\mathbb{N}}$ such that

$$U = \bigcup_{\sigma \in W} N_\sigma. \tag{90}$$

Using a standard (effective) coding procedure, we can identify a finite sequence of natural numbers with a natural number, and thus can see W as a subset of $\mathbb{N}$.

If we provide a Turing machine with oracle W, we can semi-effectively test for membership in U as follows. Assume we want to determine whether some $\alpha \in \mathbb{N}^{\mathbb{N}}$ is in U. Write α on another oracle tape, and start scanning the W oracle. If we retrieve a σ that coincides with an initial segment of α, we know $\alpha \in U$. On the other hand, if $\alpha \in U$, then we will eventually find some $\alpha \restriction_n$ in W. If $\alpha \notin U$, then the search will run forever. In other words, given W, U is **semi-decidable**, or, extending terminology from subsets of $\mathbb{N}$ to subsets of $\mathbb{N}^{\mathbb{N}}$, U is **recusively enumerable** relative to W.

Similarly, we can identify a closed set F with the code for the tree

$$T_F = \{\alpha \restriction_n : \alpha \in F, \, n \in \mathbb{N}\}. \tag{91}$$

Then determining whether $\alpha \in F$ is **co-r.e.** in (the code of) T_F. If $\alpha \notin F$ we will learn so after a finite amount of time.

These simple observations suggest the following general approach to Borel sets.

In this lecture we will fully develop this correspondence. Later, we will see that it even extends beyond the finite level.

Some notation for reals, strings, and numbers

We fix a computable bijection $\pi : \mathbb{N} \to \mathbb{N}^{<\mathbb{N}}$. In general, we will often use string and their images under π interchangeably, that is, for example, if $A \subset \mathbb{N}$, we will write $\sigma \in A$ to denote $\pi(\sigma) \in A$. We will also freely identify infinite binary sequences with the set of natural numbers they represent as their characteristic function.

Furthermore, let $\langle ., . \rangle$ be the standard coding function for pairs,

$$\langle x, y \rangle = \frac{(x+y)(x+y+1)}{2} + y. \tag{92}$$

Finally, let us define the following operation on elements of Baire (or Cantor) space: Given $\beta \in \mathbb{N}^{\mathbb{N}}$,

- let β' be the real defined by $\beta'(n) = \beta(n+1)$. (*We cut the first entry.*)
- for $m \geq 0$, let $(\beta)_m$ be the **m-th column** of β, $(\beta)_m(n) = \beta(\langle m, n \rangle)$.

Borel codes of finite order

Borel codes are defined inductively.

Definition 0.4.14. *Let $\gamma \in \mathbb{N}^{\mathbb{N}}$.*

- *Suppose $\gamma \in \mathbb{N}^{\mathbb{N}}$ is such that $\gamma(0) = 1$ and $\gamma' \in \mathbb{N}^{\mathbb{N}}$. γ **is a $\mathbf{\Sigma}^0_1$ code** for the open set*
$$U = \bigcup_{\gamma'(\sigma)=0} N_\sigma$$
- *If γ is such that $\gamma(0) = 2$ and γ' is a $\mathbf{\Sigma}^0_n$ code for $A \subseteq \mathbb{N}^{\mathbb{N}}$, we say γ **is a $\mathbf{\Pi}^0_n$ code** for $\neg A$.*
- *If γ is such that $\gamma(0) = 3$ and for each m, $(\gamma')_m$ is a $\mathbf{\Pi}^0_n$ code of a set A_m, we say γ **is a $\mathbf{\Sigma}^0_{n+1}$ code** for $\bigcup_n A_n$.*

The first position in each code indicates the kind of set it codes – an open set, a complement, or a union.

Note that the definition of Borel code actually assigns codes to **representations of sets**. A Borel set can have (and has) multiple codes, just as it has multiple representations. We can, for example, represent an open set by different sets W of initial segments.

Moreover, every $\mathbf{\Sigma}^0_1$ set is also $\mathbf{\Sigma}^0_2$, and thus a set has codes which reflect the "more complicated" definition of the $\mathbf{\Sigma}^0_1$ set as a union of closed sets. It is useful to keep this distinction between a Borel set and its Borel representation in mind.

The following is a straightforward induction.

Proposition 0.4.13. *Every $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$) set has a $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$) Borel code, and every $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$) code represents a $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$) set.*

Computing with Borel codes

Suppose γ is a **computable** code for an F_σ set B. We may assume γ is of the form $(3, \gamma')$, with each column $(\gamma')_m$ being of the form $(2, 1, (\alpha)_m)$, coding a closed set F_m.

With this, we can express membership in B as follows:

$$\begin{aligned} \beta \in B \quad &\Leftrightarrow \quad \exists m \, [\beta \text{ is in the set coded by } (\gamma')_m] \\ &\Leftrightarrow \quad \exists m \forall n \, [\beta \restriction_n \text{ is not in the set coded by } (\alpha)_m]. \\ &\Leftrightarrow \quad \exists m \forall n \, [(\alpha)_m(\beta \restriction_n) \neq 0]. \end{aligned}$$

Note that, since we assume γ to be computable, the **inner predicate** $R(m, \sigma)$ given by

$$R(m, \sigma) :\Longleftrightarrow (\alpha)_m(\sigma) \neq 0 \tag{93}$$

is **decidable**, that is, it can be decided by a Turing machine.

Hence any $\mathbf{\Sigma}^0_2$ set B with a computable code can be represented in the following form:

> There exists a decidable predicate $R(m, \sigma)$ such that

$$\beta \in B \quad \Leftrightarrow \quad \exists m \, \forall n \, \neg R(m, \beta \restriction_n).$$

Conversely, if $R(m, \sigma)$ is a (decidable) predicate, let

$$F_m = \{\beta \colon \forall n \, R(m, \beta \restriction_n)\}. \tag{94}$$

We claim that F_m is closed: Define a tree T_m by letting

$$\sigma \in T_m :\Longleftrightarrow \forall \tau \subseteq \sigma \, R(m, \tau). \tag{95}$$

Then $[T_m] = F_m$. Moreover,

$$\beta \in \bigcup_m F_m \iff \exists m \forall n \, R(m, \beta \restriction_n) \tag{96}$$

Thus, there seems to be a close connection between F_σ sets with computable Borel codes and sets definable by Σ^0_2 formulas over computable predicates. Given that we introduced the notation $\mathbf{\Sigma}^0_2$ for F_σ sets earlier, this is perhaps not very surprising.

In this analysis, there seems to be nothing specific about the F_σ used in the example. Indeed, it can be extended to Borel sets of finite order, which we will do next.

We will next introduce the **lightface** Borel hierarchy and show that it corresponds to Borel sets of finite order with recursive codes. Using **relativization**, we then obtain a complete characterization of Borel sets of finite order: *They are precisely those sets definable by arithmetical formulas, relative to a real parameter.*

But before we do that, we observe a basic fact about how we can compute with codes.

Lemma 0.4.13. *Suppose γ is a Borel code of finite order representing a set $B \subseteq \mathbb{N}^{\mathbb{N}}$. Suppose further $C \subseteq \mathbb{N}^{\mathbb{N}}$ is clopen and both C and its complement have computable $\mathbf{\Sigma}^0_1$ codes. We can, uniformly in γ, compute Borel codes for $B \cap C$ and $B \cup C$ of the same Borel complexity as γ.*

Lemma 0.4.14. *Suppose γ is a Borel code of finite order representing a set $B \subseteq \mathbb{N}^{\mathbb{N}}$. Then can, uniformly in γ and k, compute Borel codes of the same Borel complexity as γ for the set*

$$B'_k = \{\delta \colon (k, \delta) \in B\} \tag{97}$$

We leave the proofs as an exercise. Proceed by induction on the Borel complexity of γ.

The effective Borel hierarchy

Definition 0.4.15 (The Lightface Hierarchy)**.** *A set $A \subseteq \mathbb{N}^{\mathbb{N}}$ is*

- *(lightface) Σ^0_1 if there exists a computable predicate $R(\sigma)$ such that*
$$\alpha \in A \iff \exists k\, R(\alpha\,|_k),$$
- *(lightface) Π^0_n if $\neg A$ is Σ^0_n,*
- *(lightface) Σ^0_{n+1} if there exists a Π^0_n set P such that*
$$\alpha \in A \iff \exists n\, (n, \alpha) \in P.$$

The following result is at the heart of the effective theory.

Proposition 0.4.14. *Let $A \subseteq \mathbb{N}^{\mathbb{N}}$. Then*

> *A is (lightface) Σ^0_n (Π^0_n) iff A has a computable $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$) code.*

Proof. ($\Rightarrow$) We proceed by induction on the Borel complexity.

Suppose A is Σ^0_1. Let R be computable such that $A = \{\alpha \colon \exists n\, R(\alpha\,|_n)\}$. Let

$$W = \{\sigma \in \mathbb{N}^{<\mathbb{N}} \colon R(\sigma)\}. \tag{98}$$

We have $\alpha \in A$ if and only if $\alpha \in \bigcup_{\sigma \in W} N_\sigma$. Since R is decidable, W is computable and $\gamma \in \mathbb{N}^{\mathbb{N}}$ given by

$$\gamma(n) = \begin{cases} 1 & n = 0, \\ 0 & n \geq 1 \,\&\, \pi(n-1) \in W, \\ 17 & n \geq 1 \,\&\, \pi(n-1) \notin W, \end{cases} \tag{99}$$

is a computable $\mathbf{\Sigma}^0_1$ code for A.

If A is Π^0_n, then $A = \neg B$ for some Σ^0_n B. By inductive hypothesis, B has a computable $\mathbf{\Sigma}^0_n$ code γ. Then $(2, \gamma)$ is a computable $\mathbf{\Pi}^0_n$ code for $\neg A$.

Finally, assume that A is Σ^0_{n+1}. Let P be Π^0_n such that $\alpha \in A \iff \exists n\, (n, \alpha) \in P$.

By inductive hypothesis, P has a computable $\mathbf{\Pi}^0_n$ code γ. If we let $P_m = \{\beta \colon (m, \beta) \in P\}$, then $A = \bigcup P_m$. Thus, it suffices to show that we can uniformly obtain codes for P_m. This follows from Lemma **??**.

($\Leftarrow$) We proceed by induction on the complexity of the code γ.

If γ is of the form $(1, \alpha)$, with α coding an open set U. Then

$$\alpha \in U \iff \exists n\, \alpha(|_n) = 0. \tag{100}$$

Since γ is assumed to be computable, the computable relation

$$R(\sigma) :\iff \alpha(\sigma) = 0 \tag{101}$$

witnesses that U is Π^0_1.

If $\gamma = (2, \alpha)$ is a $\mathbf{\Pi}^0_n$ code, then α is a $\mathbf{\Sigma}^0_n$ code. By inductive hypothesis, the set coded by α is Σ^0_n, so by definition of the effective hierarchy and the Borel codes, γ codes a Π^0_n set.

Finally, assume $\gamma = (3, \alpha)$ is a $\mathbf{\Sigma}^0_{n+1}$ code for a set B. Then each $(\alpha)_m$ is a $\mathbf{\Pi}^0_n$ code for a set A_m.

Lemma 0.4.15. *If (α_m) is a uniformly computable sequence of $\mathbf{\Pi}^0_n$ codes for sets A_m, respectively, then there exists a $\mathbf{\Pi}^0_n$ code α for the set*

$$A = \{(m, \beta) \colon \beta \in A_m\} \tag{102}$$

Proof. Similar to Lemma **??**

□

By inductive hypothesis, the set A as in the Lemma is Π^0_n and we have

$$\beta \in B \iff \exists m(m, \beta) \in A, \tag{103}$$

which implies B is Σ^0_{n+1}.

□

Relativization

Using relativized computations via oracles, we can define a relativized version of the effective Borel hierarchy. This way we can capture *all* Borel sets of finite order, not just the ones with computable codes.

Definition 0.4.16. *Let $\gamma \in \mathbb{N}^{\mathbb{N}}$. A set $A \subseteq \mathbb{N}^{\mathbb{N}}$ is*

- ***(a)*** $\Sigma^0_1(\gamma)$ *if there exists a predicate $R(x)$ computable in γ such that*
$$\alpha \in A \iff \exists n\, R(\alpha\,|_n),$$
- ***(b)*** $\Pi^0_n(\gamma)$ *if $\neg A$ is $\Sigma^0_n(\gamma)$,*
- ***(c)*** $\Sigma^0_{n+1}(\gamma)$ *if there exists a $\Pi^0_n(\gamma)$ set P such that*
$$\alpha \in A \iff \exists n\, [(n, \alpha) \in P].$$

A straightforward relativization gives the following analogue of Proposition 0.4.14.

Proposition 0.4.15. *Let $A \subseteq \mathbb{N}^{\mathbb{N}}$ and $\gamma \in \mathbb{N}^{\mathbb{N}}$. Then*

> *A is $\Sigma^0_n(\gamma)$ ($\Pi^0_n(\gamma)$) if and only if A has a $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$) code computable in γ.*

We can now present the **fundamental theorem of effective descriptive set theory**.

Theorem 0.4.16. *A set $A \subseteq \mathbb{N}^{\mathbb{N}}$ is $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$) if and only if it is $\Sigma^0_n(\gamma)$ ($\Pi^0_n(\gamma)$) for some $\gamma \in \mathbb{N}^{\mathbb{N}}$.*

Proof. If A is $\mathbf{\Sigma}^0_n$, then by Proposition **??** it has a $\mathbf{\Sigma}^0_n$-code γ, and by Proposition 0.4.15, A is $\Sigma^0_n(\gamma)$. The other direction follows immediately from Proposition 0.4.15.

The argument for $\mathbf{\Pi}^0_n$ is completely analogous.

□

Definability in Arithmetic

One of the fundamental insights of computability theory is the close relation between computability and definability in arithmetic. The recursively enumerable subsets of $\mathbb{N}$ are precisely the sets Σ_1-definable over the standard model of arithmetic, $(\mathbb{N}, +, \cdot, 0, 1)$, and **Post's Theorem** uses this result to establish a rigid connection between levels of arithmetical complexity and computational complexity.

As indicated above, we can use this relation to give a characterization of the Borel sets of finite order in terms of definability. Since we are dealing with subsets of $\mathbb{N}^{\mathbb{N}}$, that is, with sets of functions on $\mathbb{N}$ rather than just functions on $\mathbb{N}$, we will work in the framework of **second order arithmetic**.

The **language of second order arithmetic** has two kinds of variables: **number variables** $x, y, z, \dots$ (and sometimes k, l, m, n if they are not used as metavariables), to be interpreted as elements of $\mathbb{N}$, and **function variables** $\alpha, \beta, \gamma, \dots$, intended to range over functions from $\mathbb{N}$ into $\mathbb{N}$, i.e. elements of Baire space, i.e. reals. The non-logical symbols are the binary function symbols $+, \cdot$, the binary relation symbol $<$, the **application function** symbol Ap, and the constants $\underline{0}, \underline{1}$. **Numerical terms** are defined in usual way using $+, \cdot, \underline{0}, \underline{1}$, and involve only number variables. **Atomic formulas** are $t_1 = t_2$, $t_1 < t_2$, and $\mathrm{Ap}(\alpha, t_1) = t_2$, where t_1, t_2 are numerical terms.

The **standard model of second order arithmetic** is the structure

$$\mathcal{A}^2 = (\mathbb{N}, \mathbb{N}^{\mathbb{N}}, \mathrm{Ap}, +, \cdot, <, 0, 1), \tag{104}$$

where $+$ and $\cdot$ are the usual operations on natural numbers, $<$ is the standard ordering of $\mathbb{N}$. The two domains are connected by the binary operation $\mathrm{Ap} : \mathbb{N}^{\mathbb{N}} \times \mathbb{N} \to \mathbb{N}$, defined as

$$\mathrm{Ap}(\alpha, x) = \alpha(x). \tag{105}$$

A relation $R \subseteq \mathbb{N}^m \times (\mathbb{N}^{\mathbb{N}})^n$ is **definable over** $\mathcal{A}^2$ if there exists a formula φ of second order arithmetic such that for any $x_1, \ldots, x_m \in \mathbb{N}$ and $\alpha_1, \ldots \alpha_n \in \mathbb{N}^{\mathbb{N}}$,

$$R(x_1, \ldots, x_m, \alpha_1, \ldots \alpha_n) \quad \text{iff} \quad \mathcal{A}^2 \models \varphi[x_1, \ldots, x_m, \alpha_1, \ldots \alpha_n]. \tag{106}$$

Theorem 0.4.17. *A set $A \subseteq \mathbb{N}^{\mathbb{N}}$ is Σ^0_n (Π^0_n) if and only if it is definable over $\mathcal{A}^2$ by a Σ^0_n (Π^0_n) formula.*

Here, Σ^0_n (Π^0_n) formula means that we can **only quantify over number variables**, as opposed to Σ^1_n (Π^1_n) formulas, where we can also quantify over function variables.

The proof is a straightforward extension of the standard argument for subsets of $\mathbb{N}$.

To formulate the fundamental Theorem 0.4.16 in terms of definability, we need the concept of **relative definability**. We add a new constant function symbol $\underline{\gamma}$ to the language. Given a function γ, a relation is **definable in** γ if it is definable over the structure

$$\mathcal{A}^2(\gamma) = (\mathbb{N}, \mathbb{N}^{\mathbb{N}}, \mathrm{Ap}, +, \cdot, <, 0, 1, \gamma), \tag{107}$$

where the symbol $\underline{\gamma}$ is interpreted as γ.

Then the following holds.

Theorem 0.4.18. *A set $A \subseteq \mathbb{N}^{\mathbb{N}}$ is $\boldsymbol{\Sigma}^0_n$ $(\boldsymbol{\Pi}^0_n)$ if and only if it is definable in γ by a Σ^0_n (Π^0_n) formula, for some $\gamma \in \mathbb{N}^{\mathbb{N}}$.*

This theorem facilitates the description of Borel sets considerably. As an example, consider the set

$$A = \{\alpha \colon \alpha \text{ eventually constant}\}. \tag{108}$$

We have

$$\alpha \in A \iff \exists n \forall m [m \geq n \Rightarrow \alpha(n) = \alpha(m)] \tag{109}$$

The right hand side is a Σ^0_2-formula. Hence the set A is Σ^0_2.

0.4.6 The Structure of Borel Sets

In this chapter, we further investigate the structure of Borel sets. We will use the results of the previous lecture to derive various closure properties and other structural results. As an application, we see that the Borel hierarchy is indeed proper.

Notation

Before we go on, we have to address some notational issues. So far we have used notation quite liberally, especially when it came to product sets. We will continue to do so, but we want to put this on a firmer footing.

Using coding, we can identify any product space $\mathbb{N}^m \times (\mathbb{N}^\mathbb{N})^n$ with $\mathbb{N}^\mathbb{N}$. One way to do this is to fix, for each $n \geq 1$, an effective homeomorphism $\theta_n : (\mathbb{N}^\mathbb{N})^n \to \mathbb{N}^\mathbb{N}$ and map

$$(k_1, \dots, k_m, \alpha_1, \dots, \alpha_n) \mapsto (k_1, \dots, k_m, \theta_n(\alpha_1, \dots, \alpha_n)). \tag{110}$$

Here $(k_1, \dots, k_m, \theta_n(\alpha_1, \dots, \alpha_n))$ is just a suggestive way of writing the concatenation

$$\langle k_1 \rangle^\frown \cdots ^\frown \langle k_m \rangle^\frown \theta_n(\alpha_1, \dots, \alpha_n). \tag{111}$$

We have already used this notation in the previous lecture. In the following, we will continue to switch freely between product sets and their coded counterparts, as subsets of $\mathbb{N}^\mathbb{N}$.

Another notation identifies sets and relations. We will identify sets $A \subseteq \mathbb{N}^m \times (\mathbb{N}^\mathbb{N})^n$ with the relation they induce and write $A(k_1, \dots, k_m, \alpha_1, \dots, \alpha_n)$ instead of $(k_1, \dots, k_m, \alpha_1, \dots, \alpha_n) \in A$. Conversely, we will identify relations with the set they induce.

Normal forms

Theorem Theorem **??** tells us that a set $A \subseteq \mathbb{N}^\mathbb{N}$ is $\boldsymbol{\Sigma}^0_n$ if and only if it is definable by a Σ^0_n formulas over $\mathcal{A}^2$, relative to some parameter. That means that there exists a **bounded formula** $\phi(x_1, \dots, x_n, \alpha, \underline{\gamma})$ (i.e. all quantifiers are bounded) such that

$$A(\alpha) \iff \exists x_1 \ \dots \ \mathsf{Q} x_n \ \phi(x_1, \dots, x_n, \alpha, \gamma) \text{ holds (in the standard model).} \tag{112}$$

Here γ is the parameter, and Q is "$\exists$" if n is odd, and "$\forall$" if n is even.

Similarly, $A \subseteq \mathbb{N}^\mathbb{N}$ is $\boldsymbol{\Pi}^0_n$ if and only if it is definable as

$$A(\alpha) \iff \forall x_1 \ \dots \ \mathsf{Q} x_n \ \phi(x_1, \dots, x_n, \alpha, \gamma) \text{ holds (in the standard model).} \tag{113}$$

where $\phi(x_1, \dots, x_n, \alpha, \underline{\gamma})$ is bounded, and Q is "$\forall$" if n is odd, and "$\exists$" if n is even.

What do sets defined by bounded formulas look like? An atomic formula (without parameters) either contains no function variable at all, or it is of the form $\alpha(t_1) = t_2$. This implies that the truth of an atomic formula is determined by *finitely many positions* in α. This remains true if we consider logical combinations of atomic formulas, or even bounded quantification. Hence a bounded formula defines an open subset of $\mathbb{N}^\mathbb{N}$.

On the other hand, the reals for which a bounded formula does not hold are definable by a bounded formula, too, since the negation of a bounded formula is again a bounded formula. We conclude that **bounded formulas define clopen subsets of** $\mathbb{N}^\mathbb{N}$. On the other hand, if we have $\boldsymbol{\Sigma}^0_1$-code for a set A and its complement, we can decide the relation $A(\alpha \restriction_n)$ computably in the code.

Hence we can formulate the Normal Form above as follows. $A \subseteq \mathbb{N}^\mathbb{N}$ is $\boldsymbol{\Sigma}^0_n$ if and only if there exists a clopen set $R \subseteq \mathbb{N}^n \times \mathbb{N}^\mathbb{N}$

$$A(\alpha) \iff \exists x_1 \ \dots \ \mathsf{Q} x_n \ R(x_1, \dots, x_n, \alpha), \tag{114}$$

and similarly for $\boldsymbol{\Pi}^0_n$ sets.

Closure properties

We can use the Normal Form to derive several closure properties of $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$).

If $P \subseteq \mathbb{N} \times \mathbb{N}^{\mathbb{N}}$, we define the **projection of** P **along** $\mathbb{N}$, $\exists^{\mathbb{N}} P$, as

$$\exists^{\mathbb{N}} P = \{\alpha \colon \exists n\, P(n, \alpha)\}. \tag{115}$$

We already encountered this operation in the definition of the effective Borel hierarchy (Definition **??**). The dual operation is

$$\forall^{\mathbb{N}} P = \{\alpha \colon \forall n\, P(n, \alpha)\}. \tag{116}$$

Proposition 0.4.16. *For each $n \geq 1$, $\mathbf{\Sigma}^0_n$ is closed under $\exists^{\mathbb{N}}$, and $\mathbf{\Pi}^0_n$ is closed under $\forall^{\mathbb{N}}$.*

Proof. We prove the result for Σ^0_n (lightface). The boldface case follows by relativization, and the proof for $\mathbf{\Pi}^0_n$ is completely dual.

Let $\phi(x_1, \dots, x_n, z, \alpha)$ be a bounded formula such that

$$A(z, \alpha) \iff \exists x_1 \, \dots \, \mathsf{Q} x_n \; \phi(x_1, \dots, x_n, z, \alpha) \text{ holds.} \tag{117}$$

Then

$$\exists^{\mathbb{N}} A(\alpha) \iff \exists x_0 \exists x_1 \, \dots \, \mathsf{Q} x_n \; \phi(x_1, \dots, x_n, x_0, \alpha) \tag{118}$$

We can collect two existential number quantifiers into one by using the pairing function $\langle ., . \rangle$, or rather, its inverses, which we will denote by $(.)_0$ and $(.)_1$. (Recall that the pairing function is definable by a bounded formula.) Then

$$\exists^{\mathbb{N}} A(\alpha) \iff \exists z_1 \, \dots \, \mathsf{Q} z_n \; \phi((z_1)_1, \dots, z_n, (z_1)_0, \alpha), \tag{119}$$

as desired.

□

One can use similar applications of coding and quantifier manipulation to prove a number of other closure properties, Often they follow also directly from the topological definitions, but it is good to have several techniques at hand.

Proposition 0.4.17. *For all $n \geq 1$,*

- ***(a)** $\mathbf{\Sigma}^0_n$ is closed under countable unions and finite intersections.*
- ***(b)** $\mathbf{\Pi}^0_n$ is closed under finite unions and countable intersections.*
- ***(c)** $\mathbf{\Delta}^0_n$ is closed under finite unions, finite intersections, and complements.*

Proof. One can prove this by induction along the hierarchy. To obtain the closure under finite unions and intersections, one can use the following logical equivalences.

$$\begin{aligned} \exists x\, P(x) \;\wedge\; \exists y\, R(y) &\iff \exists x \exists y\, (P(x) \;\wedge\; R(y)) \\ \forall x\, P(x) \;\vee\; \forall y\, R(y) &\iff \forall x \forall y\, (P(x) \;\vee\; R(y)) \end{aligned}$$

□

Given $P \subseteq \mathbb{N} \times \mathbb{N}^{\mathbb{N}}$, the **bounded projection** along $\mathbb{N}$ is defined as

$$\exists^{\leq} P = \{(n, \alpha) \colon \exists m \leq n\, P(m, \alpha)\}. \tag{120}$$

and the dual is

$$\forall^{\leq} P = \{(n, \alpha) \colon \forall m \leq n\, P(m, \alpha)\}. \tag{121}$$

Proposition 0.4.18. *For all $n \geq 1$, $\mathbf{\Sigma}^0_n$, $\mathbf{\Pi}^0_n$, and $\mathbf{\Delta}^0_n$ are closed under $\exists^{\leq}$ and $\forall^{\leq}$.*

Proof. In this case we use the computable coding function $\pi : \mathbb{N} \to \mathbb{N}^{<\mathbb{N}}$. We have the following equivalence, which immediately implies the closure properties for $\mathbf{\Sigma}^0_n$ and $\mathbf{\Pi}^0_n$, respectively, and hence also for $\mathbf{\Delta}^0_n$.

$$\begin{aligned} \forall m \le n\, \exists k\, P(m,k) &\iff \exists k\, \forall m \le n\, P(m, \pi(k)_m) \\ \exists m \le n\, \forall k\, P(m,k) &\iff \forall k\, \exists m \le n\, P(m, \pi(k)_m) \end{aligned}$$

□

Finally, the levels of the Borel hierarchy are closed under continuous preimages.

Proposition 0.4.19. *For all $n \ge 1$, for any $A \subseteq \mathbb{N}^{\mathbb{N}}$, and for any continuous $f : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$, if A is $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$, $\mathbf{\Delta}^0_n$) then $f^{-1}(A)$ is $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$, $\mathbf{\Delta}^0_n$).*

Proof. This follows easily by induction on n, since open and closed sets are closed under continuous preimages.

However, we can also argue via definability, since by Proposition 0.3.8 one can represent a continuous function through a monotone mapping ψ from finite strings to finite strings. We have

$$f^{-1}(A) = \{\alpha \colon A(f(\alpha))\}. \tag{122}$$

Let R be clopen such that

$$A(\alpha) \iff \exists x_1 \ldots \mathsf{Q} x_n\, R(x_1, \ldots, x_n, \alpha). \tag{123}$$

Since clopen predicates depend only on a finite initial segment of α, we can substitute $f(\alpha)$ for α. The resulting formula defines $f^{-1}(A)$, and is equivalent to a Σ^0_n-formula relative to a parameter coding the mapping ψ.

□

Universal sets

Let Γ be a family of subsets defined in various Polish spaces. Of course we have in mind the classes $\mathbf{\Sigma}^0_n$ or $\mathbf{\Pi}^0_n$, but the concept of a **universal set** can be defined quite generally.

Definition 0.4.17. *Let Y be a set. A set $U \subseteq X \times Y$ is **Y-universal for** Γ if $U \in \Gamma$, and for every set A in Γ, there exists a $y \in Y$ such that*

$$A = \{x \colon (x,y) \in U\}. \tag{124}$$

A universal set for Γ can be thought of as a **parametrization** of Γ, the second component providing a **code** or **parameter** for each set in Γ.

A well-known example of a universal set is the **generalized halting problem**,

$$K_0 = \{(x,e) \colon \text{ the } e\text{-th Turing machine halts on input } x\}. \tag{125}$$

In the sense of the above definition, K_0 is $\mathbb{N}$-universal for the family of recursively enumerable sets.

Proposition 0.4.20. *For any $n \ge 1$, there exists a set $U \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}}$ that is $\mathbb{N}^{\mathbb{N}}$-universal for $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$).*

Proof. We can use the Borel codes defined in the previous lecture.

First of all, notice that for each $n \ge 1$, the set of all $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$)-codes is homeomorphic to $\mathbb{N}^{\mathbb{N}}$. This follows easily from the definition of the Borel codes. Hence, if we fix n, every $\gamma \in \mathbb{N}^{\mathbb{N}}$ represents a $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$)-code of a $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$) set, and every such set in turn has a code $\gamma \in \mathbb{N}^{\mathbb{N}}$.

For fixed n, we let

$$U_n = \{(\alpha, \gamma) \colon \gamma \in \mathbb{N}^{\mathbb{N}} \text{ and } \alpha \text{ is in the } \mathbf{\Sigma}^0_n\ (\mathbf{\Pi}^0_n) \text{ set coded by } \gamma\}. \tag{126}$$

It follows easily from Theorem 0.4.16 that U_n is $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$), too, and it is clear from the definition of U that it parametrizes $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$).

□

The result can be generalized to hold for arbitrary Polish spaces X, i.e. for any $n \geq 1$, there exists a set $U \subseteq \mathbb{N}^{\mathbb{N}} \times X$ that is $\mathbb{N}^{\mathbb{N}}$-universal for $\mathbf{\Sigma}^0_n(X)$ ($\mathbf{\Pi}^0_n(X)$). To achieve this, one has to define Borel codes for X. This can be done by fixing a countable basis (V_n) of the topology of X, and assign a sequence $\gamma \in \mathbb{N}^{\mathbb{N}}$ the open set

$$U_\gamma = \bigcup_{n\in\mathbb{N}} V_{\gamma(n)}. \tag{127}$$

The definition of codes for higher levels is then similar to Definition Definition **??**.

As in the case of the halting problem, we can use the existence of universal sets to show that the levels of the Borel hierarchy are proper. The crucial point is that we can use universal sets to **diagonalize**.

Theorem 0.4.19. *For any $n \geq 1$, $\mathbf{\Sigma}^0_n \neq \mathbf{\Pi}^0_n$.*

Proof. Let U be an $\mathbb{N}^{\mathbb{N}}$-universal set for $\mathbf{\Sigma}^0_n$. Put

$$D = \{\alpha \colon (\alpha, \alpha) \in U\}. \tag{128}$$

Since U is $\mathbf{\Sigma}^0_n$, D is $\mathbf{\Sigma}^0_n$, too. Then $\neg D$ is $\mathbf{\Pi}^0_n$, but cannot be $\mathbf{\Sigma}^0_n$, for then there would exist β such that

$$\neg D = \{\alpha \colon (\alpha, \beta) \in U\}, \tag{129}$$

and thus

$$\beta \in D \iff (\beta, \beta) \in U \iff \beta \in \neg D, \tag{130}$$

a contradiction.

□

The diagonal set D can obviously be defined for any universal set U, and hence the same proof yields a $\mathbf{\Pi}^0_n$ set that is not $\mathbf{\Sigma}^0_n$.

Corollary 0.4.19.1. *For any $n \geq 1$,*

$$\mathbf{\Delta}^0_n \subsetneq \mathbf{\Sigma}^0_n \subsetneq \mathbf{\Delta}^0_{n+1}$$
$$\mathbf{\Delta}^0_n \subsetneq \mathbf{\Pi}^0_n \subsetneq \mathbf{\Delta}^0_{n+1}.$$

Proof. Since $\mathbf{\Sigma}^0_n \nsubseteq \mathbf{\Pi}^0_n$ and $\mathbf{\Pi}^0_n \nsubseteq \mathbf{\Sigma}^0_n$, $\mathbf{\Delta}^0_n \subsetneq \mathbf{\Sigma}^0_n, \mathbf{\Pi}^0_n$. On the other hand if $\mathbf{\Sigma}^0_n = \mathbf{\Delta}^0_{n+1}$, then $\mathbf{\Sigma}^0_n$ would be closed under complements, and hence $\mathbf{\Sigma}^0_n = \mathbf{\Pi}^0_n$, contradicting Theorem **??**.

□

Borel sets of transfinite order

We saw that the **Borel sets of finite order**

$$\mathrm{Borel}_\omega = \bigcup_{n<\omega} \mathbf{\Sigma}^0_n \tag{131}$$

form a proper hierarchy. This fact also implies that Borel_ω does not exhaust all Borel sets.

Proposition 0.4.21. *There exists a Borel set B that is not $\mathbf{\Sigma}^0_n$ for any $n \in \mathbb{N}$.*

Proof. For every $n \in \mathbb{N}$, pick a set B_n in $\mathbf{\Pi}^0_n \setminus \mathbf{\Sigma}^0_n$. Put

$$B = \bigcup_{n \in \mathbb{N}} \{(n, \alpha) \colon \alpha \in B_n\}. \tag{132}$$

Each of the sets in the union is Borel and hence B is Borel. If B were of finite order, it would be $\mathbf{\Sigma}^0_k$ for some $k \geq 1$. Since each $\mathbf{\Sigma}^0_n$ is closed under finite intersections, it follows that for all $m \geq 1$,

$$B \cap N_{\langle m \rangle} \tag{133}$$

is $\mathbf{\Sigma}^0_k$. But $B \cap N_{\langle m \rangle}$ is homeomorphic to B_m, hence B_m in $\mathbf{\Sigma}^0_k$ for all $m \geq 1$, contradiction. □

We can extend the Borel hierarchy to arbitrary ordinals.

Definition 0.4.18. *Let X be a Polish space. Given an ordinal ξ, we define*

$$\begin{aligned}
&\mathbf{\Sigma}^0_\xi(X) = \{\bigcup_k A_k \colon A_k \in \mathbf{\Pi}^0_{\zeta_k}(X),\ \zeta_k < \xi\},\\
&\mathbf{\Pi}^0_\xi(X) = \{\neg A \colon A \in \mathbf{\Sigma}^0_\xi(X)\} = \neg\, \mathbf{\Sigma}^0_\xi(X),\\
&\mathbf{\Delta}^0_\xi(X) = \mathbf{\Sigma}^0_\xi(X) \cap \mathbf{\Pi}^0_\xi(X).
\end{aligned}$$

It actually suffices to consider ordinals up to ω_1, the first uncountable ordinal.

Proposition 0.4.22. *For every Borel set B there exists $\xi < \omega_1$ such that $B \in \mathbf{\Sigma}^0_\xi$.*

Proof. If B is open, this is clear. It is also clear if B is the complement of a Borel for which the statement has been verified.

Assume finally that

$$B = \bigcup_n B_n, \quad \text{where each } B_n \text{ is Borel,} \tag{134}$$

and assume the statement holds for each B_n. For each n, let ξ_n be a countable ordinal such that

$$B_n \in \mathbf{\Pi}^0_{\xi_n}. \tag{135}$$

Then

$$B \in \mathbf{\Sigma}^0_\xi, \quad \text{where } \xi = \sup\{\xi_n + 1 \colon n \in \mathbb{N}\}. \tag{136}$$

Since each ξ_n is countable, ξ is countable. □

Borel sets of infinite order have the **same closure properties as their counterparts of finite order**. The proofs, however have to proceed by induction using the topological properties of $\mathbf{\Sigma}^0_\xi$ and $\mathbf{\Pi}^0_\xi$, since the characterization via definability in arithmetic is no longer available – the arithmetical hierarchy reaches only to ω.

Similarly, the Hierarchy Theorem (Theorem **??**) extends to the transfinite levels. As the finite levels, this follows from the existence of universal sets for each level, which we now prove for the full hierarchy.

Proposition 0.4.23. *For each $\xi < \omega_1$, there exists a $\mathbb{N}^\mathbb{N}$-universal set for $\mathbf{\Sigma}^0_\xi$ ($\mathbf{\Pi}^0_\xi$).*

Proof. If U is $\mathbb{N}^\mathbb{N}$-universal for $\mathbf{\Sigma}^0_\xi$, then

$$\neg U = \{(\alpha, \gamma) \colon (\alpha, \gamma) \notin U\} \tag{137}$$

is $\mathbb{N}^\mathbb{N}$-universal for $\mathbf{\Pi}^0_\xi$, since for any $\mathbf{\Pi}^0_\xi$ set A, $B = \neg A$ is $\mathbf{\Sigma}^0_\xi$ and hence there exists a γ such that

$$B = \{\beta\colon (\beta, \gamma) \in U\} \tag{138}$$

and hence

$$A = \{\alpha\colon (\alpha, \gamma) \notin U\}. \tag{139}$$

It remains to show that each $\mathbf{\Sigma}^0_\xi$ has an $\mathbb{N}^\mathbb{N}$-universal set. By induction hypothesis, for every $\eta < \xi$ exists a $\mathbb{N}^\mathbb{N}$-universal set U_η for $\mathbf{\Pi}^0_\eta$. Since ξ is countable, we can pick a monotone sequence of ordinals (ξ_n) such that $\xi = \sup\{\xi_n + 1\colon n < \omega\}$. Define

$$U_\xi = \{(\alpha, \gamma)\colon \exists n(\alpha, (\gamma)_n) \in U_{\xi_n}\}, \tag{140}$$

where $(\gamma)_n$ denotes the nth column of γ.

It is straightforward to check that U_ξ is $\mathbb{N}^\mathbb{N}$-universal for $\mathbf{\Sigma}^0_\xi$. (Note that any set A in $\mathbf{\Sigma}^0_\xi$ can be represented as $\bigcup_n A_n$ with $A_n \in \mathbf{\Pi}^0_{\xi_n}$, since $(\xi_n + 1)$ is cofinal in ξ.)

□

The construction of the universal $\mathbf{\Sigma}^0_\xi$ set bears some resemblance to the construction of a $\mathbf{\Sigma}^0_{n+1}$ code. It is indeed possible to formally define Borel codes for *all* Borel sets.

Definition 0.4.19. *Let $\gamma \in \mathbb{N}^\mathbb{N}$.*

- *Suppose $\gamma \in \mathbb{N}^\mathbb{N}$ is such that $\gamma(0) = 1$ and $\gamma' \in \mathbb{N}^\mathbb{N}$. γ is a Borel code for the open set*
$$U = \bigcup_{\gamma'(\sigma)=0} N_\sigma$$
- *If γ is such that $\gamma(0) = 2$ and γ' is a Borel code for $A \subseteq \mathbb{N}^\mathbb{N}$, we say γ is a Borel code for $\neg A$.*
- *If γ is such that $\gamma(0) = 3$ and for each m, γ'_m is a Borel code of a set A_m, we say γ is a Borel code for $\bigcup_n A_n$.*

Any Borel code induces a well-founded tree (given by the coding nodes 1, 2,and 3). We can also consider Borel sets with computable codes. But there is no more straightforward connection with effective definability. It is possible to do this, but it requires a careful development of what it means to take effective unions along countable ordinals. We will return to it later.

Looking further ahead, one can show that **the set of all Borel codes is not Borel** (exercise – use a diagonalization argument as in the proof of Theorem **??**). At the heart of this lies the fact that we cannot, in a Borel way, describe whether an arbitrary tree over $\mathbb{N}$ is well-founded or not. This will soon be a central topic when we turn our investigation to analytic and co-analytic sets.

0.4.7 Continuous Images of Borel sets

In 1916, Nikolai Lusin asked his student Mikhail Souslin to study a paper by Henri Lebesgue. Souslin found a number of errors, including a lemma that asserted that the projection of a Borel is again Borel. In this lecture we will study the behavior of Borel sets under continuous functions. We will see that on the one hand every Borel set is the continuous image of a closed set, but that on the other hand continuous images of Borel sets are not always Borel.

This gives rise to a new family of sets, the **analytic** sets, which form a proper superclass of the Borel sets with interesting properties.

Borel sets as continuous images of closed sets

We have seen in Theorem **??** that every Polish space is the continuous image of Baire space $\mathbb{N}^{\mathbb{N}}$. As we will see now, we can strengthen this result.

Theorem 0.4.20 (Lusin and Souslin)**.** *Let X be a Polish space. Then there exists a closed subset $F \subseteq \mathbb{N}^{\mathbb{N}}$ and a continuous bijection $f : F \to X$ that can be extended to a continuous surjection $g : \mathbb{N}^{\mathbb{N}} \to X$.*

We have seen (Theorem **??**) that every uncountable Polish space contains a homeomorphic embedding of Cantor space. This was achieved by means of a **Cantor scheme**. To prove Theorem **??**, we take up this idea again and adapt it to the Baire space.

Definition 0.4.20. *A **Lusin scheme** on a set X is a family $(F_\sigma)_{\sigma \in \mathbb{N}^{<\mathbb{N}}}$ of subsets of X such that*

- ***(i)*** *$\sigma \subseteq \tau$ implies $F_\sigma \supseteq F_\tau$,*
- ***(ii)*** *for all $\tau \in \mathbb{N}^{<\mathbb{N}}$, $i \neq j \in \mathbb{N}$, $F_{\tau\frown\langle i\rangle} \cap F_{\tau\frown\langle j\rangle} = \emptyset$.*

If it has the additional property that

- ***(iii)*** $\operatorname{diam}(F_{\alpha|_n}) \to 0$ *for $n \to \infty$,*

then we can, similarly to a Cantor scheme, define the set

$$D = \{\alpha \in \mathbb{N}^{\mathbb{N}} \colon \bigcap_{n \in \mathbb{N}} F_{\alpha|_n} \neq \emptyset\}$$

*and an **associated map** $f : D \to X$ by*

$$\{f(\alpha)\} = \bigcap_{n \in \mathbb{N}} F_{\alpha|_n}.$$

Properties (i)-(iii) ensure that f is continuous and injective.

To prove the theorem we devise a Lusin scheme on X such that D will be closed, and f will be a surjection, too. This is ensured by the following additional properties.

- **(a)** $F_\emptyset = X$,
- **(b)** Each F_τ is $\mathbf{\Sigma}^0_[2]$,
- **(c)** For each τ, $\operatorname{diam}(F_\sigma) \leq 1/2^{|\sigma|}$,
- **(d)** $F_\tau = \bigcup_{i \in \mathbb{N}} F_{\tau\frown\langle i\rangle} = \bigcup_{i \in \mathbb{N}} \overline{F_{\tau\frown\langle i\rangle}}$.

For this we have to show that every $\mathbf{\Sigma}^0_[2]$ set $F \subseteq X$ can be written, for given $\varepsilon > 0$, as $F = \bigcup_{i \in \mathbb{N}} F_i$, where the F_i are pairwise disjoint $\mathbf{\Sigma}^0_[2]$ sets of diameter $< \varepsilon$ so that $\overline{F_i} \subseteq F$:

Let $F = \bigcup_{i \in \mathbb{N}} C_i$, where C_i is closed, and $C_i \subseteq C_{i+1}$. Then $F = \bigcup_{i \in \mathbb{N}} (C_{i+1} \setminus C_i)$.

Let (U_n) be a covering of X with open sets of diameter $< \varepsilon$. Put $D^{(i)}_n = U_n \cap (C_{i+1} \setminus C_i)$. Then $D^{(i)}_n$ is $\mathbf{\Delta}^0_2$. Now let $E^{(i)}_n = D^{(i)}_n \setminus (D^{(i)}_1 \cup \cdots \cup D^{(i)}_{n-1})$.

Then $C_{i+1} \setminus C_i = \bigcup_{n \in \mathbb{N}} E^{(i)}_n$ where the $E^{(i)}_j$ are $\mathbf{\Sigma}^0_[2]$ sets of diameter $< \varepsilon$. Therefore,

$$F = \bigcup_{i,n \in \mathbb{N}} E^{(i)}_n \quad \text{and} \quad \overline{E^{(i)}_n} \subseteq \overline{C_{i+1} \setminus C_i} \subseteq C_{i+1} \subseteq F.$$

The mapping f associated with this Lusin scheme is surjective due to (a) and (d). Furthermore, the domain D of f is closed: Suppose $\alpha_n \in D$, $\alpha_n \to \alpha$. Then $f(\alpha_n)$ is Cauchy, since for $\varepsilon > 0$, there exists N with $\operatorname{diam}(F_{\alpha|_N}) < \varepsilon$ and n_0 such that $\alpha_n \mid_N = \alpha \mid_N$ for all $n \geq n_0$, so that $d(f(\alpha_n), f(\alpha_m)) < \varepsilon$ whenever $n, m \geq n_0$. Since X is Polish $f(\alpha_n) \to y$ for some $y \in X$.

By (d) we have $y \in \bigcap_n \overline{F_{\alpha|_n}} = \bigcap_n F_{\alpha|_n}$, hence $\alpha \in D$ and $f(\alpha) = y$.

It remains to show that we can extend f to a continuous surjection $g : \mathbb{N}^{\mathbb{N}} \to X$. Say a closed subset C of a topological space Y is a **retract** of Y if there exists a continuous surjection $g : Y \to C$ such that $g \mid_C = \mathrm{id}$.

Lemma 0.4.21. *Every non-empty closed subset of $\mathbb{N}^{\mathbb{N}}$ is a retract of $\mathbb{N}^{\mathbb{N}}$.*

Proof. Let $C \subseteq \mathbb{N}^{\mathbb{N}}$ be closed, and let T be a pruned tree such that $[T] = C$. We define a monotone mapping $\phi : \mathbb{N}^{<\mathbb{N}} \to T$ such that $\phi(\sigma) = \sigma$ for all $\sigma \in T$. Then the induced (continuous) mapping $\phi^* : \mathbb{N}^{\mathbb{N}} \to C$ is the desired retract.

Define ϕ by induction. Let $\phi(\varnothing) = \varnothing$. Given $\phi(\tau)$, let

$$\phi(\tau^\frown\langle m\rangle) = \begin{cases} \tau^\frown\langle m\rangle & \text{if } \tau^\frown\langle m\rangle \in T, \\ \text{any } \phi(\tau)^\frown\langle k\rangle \in T & \text{otherwise.} \end{cases} \tag{141}$$

Note that k must exist since T is pruned.

□

If we combine the retract function with f, we then obtain the desired surjection $\mathbb{N}^{\mathbb{N}} \to X$. This concludes the proof of Theorem **??**.

Refining the topology as in Theorem **??**, we can extend the result from Polish spaces to Borel sets.

Corollary 0.4.21.1 (Lusin and Souslin)**.** *For every Borel subset B of a Polish space X there exists a closed set $F \subseteq \mathbb{N}^{\mathbb{N}}$ and a continuous bijection $f : F \to B$. Furthermore, f can be extended to a continuous surjection $g : \mathbb{N}^{\mathbb{N}} \to B$.*

Proof. Enlarge the topology $\mathcal{O}$ of X to a topology $\mathcal{O}_B$ for which B is clopen. By Theorem **??**, $(B, \mathcal{O}_B \mid_B)$ is a Polish space. By the previous theorem, there exists a closed set $F \subset \mathbb{N}^{\mathbb{N}}$ and a continuous bijection $f : \mathbb{N}^{\mathbb{N}} \to (B, \mathcal{O}_B \mid_B)$. Since $\mathcal{O} \subseteq \mathcal{O}_B$, $f : F \to B$ is continuous for $\mathcal{O}$, too.

□

This theorem can be reversed in the following sense.

Theorem 0.4.22 (Lusin and Souslin)**.** *Suppose X, Y are Polish and $f : X \to Y$ is continuous. If $A \subseteq X$ is Borel and $f \mid_A$ is injective, then $f(A)$ is Borel.*

For a proof (which uses facts about analytic sets), see **?** (II.15.1).

Images of Borel sets under arbitrary continuous functions

As announced in the introduction, Borel sets are *not* closed under *arbitrary* continuous mappings.

Theorem 0.4.23 (Souslin)**.** *The Borel sets are not closed under continuous images.*

Proof. Let $U \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}}$ be $\mathbb{N}^{\mathbb{N}}$-universal for $\mathbf{\Pi}^0_[1](\mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}})$. Define

$$F := \{(\alpha, \beta) \colon \exists \gamma \ (\alpha, \gamma, \beta) \in U\}.$$

We claim that this set is $\mathbb{N}^{\mathbb{N}}$-*universal for the set of all continuous images of closed subsets of* $\mathbb{N}^{\mathbb{N}}$:

On the one hand F is a projection of a closed set, and projections are continuous. This implies that all the sets $F_\beta = \{\alpha \colon (\alpha, \beta) \in F\}$ are continuous images of a closed set.

On the other hand, if $f : C \to \mathbb{N}^{\mathbb{N}}$ is continuous with $C \subseteq \mathbb{N}^{\mathbb{N}}$ closed (possibly empty) and $f(C) = A$, then

$$\alpha \in A \iff \exists \gamma \ (\gamma, \alpha) \in \mathrm{Graph}(f) \iff \exists \gamma \ (\alpha, \gamma) \in \mathrm{Graph}(f^{-1}).$$

Since f is continuous, $\mathrm{Graph}(f)$ and hence also $\mathrm{Graph}(f^{-1})$ are closed subsets of $\mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}}$. Thus, by the universality of U, there exists β such that

$$\mathrm{Graph}(f^{-1}) = U_\beta = \{(\alpha, \gamma) \colon (\alpha, \gamma, \beta) \in U\}, \tag{142}$$

and hence

$$A = F_\beta. \tag{143}$$

F cannot be Borel: Otherwise $D_F = \{\alpha \colon (\alpha, \alpha) \notin F\}$ were Borel. By Corollary **??**, every Borel set is the image of a closed set under a continuous mapping. This implies that $D_F = F_\beta$. But then

$$\beta \in D_F \iff \beta \in F_\beta \iff (\beta, \beta) \in F \iff \beta \notin D_F, \tag{144}$$

contradiction.

□

0.5 Projective sets

0.5.1 Analytic Sets

Definition 0.5.1. *A subset A of a Polish space X is* ***analytic*** *if it is empty or there exists a continuous function $f : \mathbb{N}^{\mathbb{N}} \to X$ such that $f(\mathbb{N}^{\mathbb{N}}) = A$.*

We will later see that the analytic sets correspond to the sets definable by means of Σ^1_1 formulas, that is formulas in the language of second order arithmetic that have **one existential function quantifier**.

Therefore, we will denote the analytic subsets of X also by

$$\boldsymbol{\Sigma}^1_1(X). \tag{145}$$

Here are some simple properties of analytic sets.

Proposition 0.5.1.
- ***(i)*** *Every Borel set is analytic.*
- ***(ii)*** *A continuous image of analytic set is analytic.*
- ***(iii)*** *Countable unions of analytic sets are analytic.*

Proof. (i): This follows directly from Corollary **??**.

(ii): The composition of continuous mappings is continuous.

(iii): Let A_n be analytic and $f_n : \mathbb{N}^{\mathbb{N}} \to X$ such that $f_n(\mathbb{N}^{\mathbb{N}}) = A_n$. Define $f : \mathbb{N}^{\mathbb{N}} \to X$ by

$$f(m, \alpha) = f_n(\alpha). \tag{146}$$

Then f is continuous and $f(\mathbb{N}^{\mathbb{N}}) = \bigcup_n A_n$.

□

We can use our previous results about Borel sets to give various equivalent characterizations of analytic sets.

Proposition 0.5.2. *For a subset A of a Polish space X, the following are equivalent.*

- ***(i)*** *A is analytic,*
- ***(ii)*** *A is empty or there exists a Polish space Y and a continuous $f : Y \to X$ such that $f(Y) = A$,*
- ***(iii)*** *A is empty or there exists a Polish space Y, a Borel set $B \subseteq Y$ and a continuous $f : Y \to X$ such that $f(B) = A$.*
- ***(iv)*** *A is the* ***projection*** *of a closed set $F \subseteq \mathbb{N}^{\mathbb{N}} \times X$ along $\mathbb{N}^{\mathbb{N}}$,*
- ***(v)*** *A is the projection of a $\boldsymbol{\Pi}^0_2$ set $G \subseteq 2^{\mathbb{N}} \times X$ along $2^{\mathbb{N}}$,*
- ***(vi)*** *A is the projection of a Borel set $B \subseteq X \times Y$ along Y, for some Polish space Y.*

Proof. (i) ⇔ (ii): Follows from Theorem **??** and Proposition 0.5.1 (ii).

(ii) ⇔ (iii): Follows from Corollary Corollary **??** and Proposition 0.5.1 (ii).

(i) ⇒ (iv): Let $f : \mathbb{N}^{\mathbb{N}} \to X$ be continuous, $f(\mathbb{N}^{\mathbb{N}}) = A$. Then

$$x \in A \iff \exists \alpha \, (\alpha, x) \in \mathrm{Graph}(f), \tag{147}$$

hence A is the projection of the closed set $\mathrm{Graph}(f)$ along $\mathbb{N}^{\mathbb{N}}$.

(iv) ⇒ (iii): Clear, since projections are continuous.

(iv) ⇒ (v): $\mathbb{N}^{\mathbb{N}}$ is homeomorphic to a $\boldsymbol{\Pi}^0_2$ subset of $2^{\mathbb{N}}$. (Exercise!)

(v) ⇒ (vi), (vi) ⇒ (iii): Obvious.

□

The Lusin Separation Theorem

In a course on computability theory one learns that there are **effectively inseparable** disjoint computably enumerable sets. i.e. disjoint c.e. sets $W, Z \subseteq \mathbb{N}$ for which no recursive set A exists with $W \subseteq A$ and $A \cap Z = \emptyset$.

In contrast to this, disjoint analytic sets can always be separated by a Borel set – they are **Borel separable**.

Theorem 0.5.1 (Lusin)**.** *Let $A, B \subseteq X$ be disjoint analytic sets. Then there exists a Borel $C \subseteq X$ such that*

$$A \subseteq C \quad \textit{and} \quad B \cap C = \emptyset, \tag{148}$$

Proof. Let $f : \mathbb{N}^{\mathbb{N}} \to A$ and $g : \mathbb{N}^{\mathbb{N}} \to B$ be continuous surjections.

We argue by contradiction. The key idea is: if A and B are Borel inseparable, then, for some $i, j \in \mathbb{N}$, $A_{\langle i \rangle} = f(N_{\langle i \rangle})$ and $B_{\langle j \rangle} = g(N_{\langle j \rangle})$ are Borel inseparable.

This follows from the observation

(∗) if the sets $R_{m,n}$ separate the sets P_m, Q_n (for each m, n), then $R = \bigcup_m \bigcap_n R_{m,n}$ separates the sets $P = \bigcup_m P_m$, $Q = \bigcup_n Q_n$.

So, by using (∗) repeatedly, we can construct sequences $\alpha, \beta \in \mathbb{N}^{\mathbb{N}}$ such that for all n, $A_{\alpha|_n}$ and $B_{\beta|_n}$ are Borel inseparable, where

$$A_\sigma = f(N_\sigma) \quad \text{and} \quad B_\sigma = g(N_\sigma). \tag{149}$$

Then we have $f(\alpha) \in A$ and $g(\beta) \in B$, and since A and B are disjoint, $f(\alpha) \neq g(\beta)$. Let U, V be disjoint open sets such that $f(\alpha) \in U$, $g(\beta) \in V$. Since f and g are continuous, there exists N such that $f(N_{\alpha|_N}) \subseteq U$, $g(N_{\beta|_N}) \subseteq V$, hence U separates $A_{\alpha|_N}$ and $B_{\beta|_N}$, contradiction.

□

The Separation Theorem yields a nice characterization of the Borel sets.

Theorem 0.5.2 (Souslin)**.** *If a set A and its complement $\neg A$ are both analytic, then A is Borel.*

Proof. In Theorem **??**, chose $A = A$ and $B = \neg A$.

□

It follows from Theorem **??** and the Theorem **??** that the analytic sets are not closed under complements.

Sets whose complement is analytic are called **co-analytic**. Analogous to the levels of the Borel hierarchy, the co-analytic subsets of a Polish space X are denoted by

$$\mathbf{\Pi}^1_1(X). \tag{150}$$

If we define, again analogy to the Borel hierarchy,

$$\mathbf{\Delta}^1_1(X) = \mathbf{\Sigma}^1_{(1)}(X) \cap \mathbf{\Pi}^1_1(X), \tag{151}$$

then Souslin's Theorem states that

$$\text{Borel}(X) = \mathbf{\Delta}^1_1(X). \tag{152}$$

The Souslin operation

Souslin schemes give an alternative presentation of analytic sets which will be useful later.

Definition 0.5.2. *A **Souslin scheme** on a set X is a family $P = (P_\sigma)_{\sigma \in \mathbb{N}^{<\mathbb{N}}}$ of subsets of X indexed by $\mathbb{N}^{<\mathbb{N}}$.*

*The **Souslin operation** $\mathcal{A}$ for a Souslin scheme is given by*

$$\mathcal{A}P = \bigcup_{\alpha \in \mathbb{N}^{\mathbb{N}}} \bigcap_{n \in \mathbb{N}} P_{\alpha|_n}. \tag{153}$$

This means

$$x \in \mathcal{A}P \iff \exists \alpha \in \mathbb{N}^{\mathbb{N}}\, \forall n \in \mathbb{N}\; x \in P_{\alpha|_n}. \tag{**}$$

The analytic sets are precisely the sets that can be obtained by Souslin operations on closed sets. If a Γ is a family of subsets of a set X, we let

$$\mathcal{A}\Gamma = \{\mathcal{A}P\colon\ P = (P_\sigma) \text{ is a Souslin scheme with } P_\sigma \in \Gamma \text{ for all } \sigma\}. \tag{154}$$

Theorem 0.5.3.

$$\mathbf{\Sigma}^1_1(X) \ = \ \mathcal{A}\mathbf{\Pi}^0_1(X). \tag{155}$$

Proof. Suppose $f : \mathbb{N}^{\mathbb{N}} \to X$ is continuous with $f(\mathbb{N}^{\mathbb{N}}) = A$. Then

$$x \in A \iff \exists\alpha \in \mathbb{N}^{\mathbb{N}}\ \forall n \in \mathbb{N}\ x \in \overline{f(N_{\alpha|n})}.$$

Hence if we let $P_\sigma = \overline{f(N_\sigma)}$, then

$$A = \mathcal{A}\,P,$$

for the Souslin scheme $P = (P_\sigma)$.

To see that any set A in $\mathcal{A}\mathbf{\Pi}^0_1(X)$ is analytic, consider $(**)$. If the P_σ are closed, the condition

$$(\alpha, x) \in F \iff \forall n \in \mathbb{N}\ x \in P_{\alpha|n} \tag{156}$$

defines a closed subset of $\mathbb{N}^{\mathbb{N}} \times X$ such that A is the projection of F along $\mathbb{N}^{\mathbb{N}}$.

□

Note that the Souslin scheme (P_σ) used in the previous proof has the additional property that

$$\sigma \subseteq \tau \quad \Rightarrow \quad P_\sigma \supseteq P_\tau. \tag{157}$$

Such Souslin schemes are called **regular**. By replacing any Souslin scheme P_σ with

$$Q_\sigma = \bigcap_{\tau \subseteq \sigma} P_\tau, \tag{158}$$

we obtain a regular Souslin scheme $Q = (Q_\sigma)$ with $\mathcal{A}\,Q = \mathcal{A}\,P$. Moreover, if the P_σ are from a class Γ, and Γ is closed under finite intersections, then the Q_σ are also from Γ. In particular, **any analytic set can be obtained from a regular Souslin scheme of closed sets** via the Souslin operation.

0.5.2 Regularity Properties of Analytic sets

In this lecture we verify that the analytic sets are Lebesgue measurable (LM) and have the Baire property (BP). Since both properties are closed under complements, they also hold for the class of **co-analytic sets $\mathbf{\Pi}^1_1$**.

The analytic sets also have the perfect subset property (PS).

> **Exercise**
> Show that if $A \subseteq \mathbb{N}^\mathbb{N}$ is analytic and uncountable, then it contains a perfect subset.
> (*Hint: Since A is analytic, there exists a continuous mapping $f : \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ such that $A = f(\mathbb{N}^\mathbb{N})$. Construct an embedding of $2^\mathbb{N}$ into A. Show that we can find two disjoint open sets U_0, U_1 whose intersection with $A = f(\mathbb{N}^\mathbb{N})$ is uncountable. The preimages of the U_i are disjoint open subsets with uncountable images. Show that this process can be continued and defines in the limit an injection of $2^\mathbb{N}$ into A.*)

For Borel sets, one proves (LM) and (BP) by showing that the class of sets having (LM) (or (BP), respectively) forms a σ-algebra and contains the open sets. For the analytic sets, this method is no longer available. We can, however, prove a similar property with respect to the Souslin operation $\mathcal{A}$, which can be seen as an extension of basic set theoretic operations into the uncountable.

More specifically, we will show the following.

- The Souslin operation $\mathcal{A}$ is **idempotent**, i.e. $\mathcal{A}\mathcal{A}\,\Gamma = \mathcal{A}\Gamma$. This implies that the analytic sets are closed under $\mathcal{A}$.
- The family of sets with (LM) (or (BP), respectively), is closed under the Souslin operation. Since the closed sets have both properties, and the Souslin operator is clearly monotone on classes, this yields the desired regularity results.

Idempotence of the Souslin operation

Theorem 0.5.4. *For every family Γ of subsets of a set X,*

$$\mathcal{A}\mathcal{A}\Gamma = \mathcal{A}\Gamma.$$

Proof. We clearly have $\Gamma \subseteq \mathcal{A}\Gamma$, so that we only need to prove $\mathcal{A}\mathcal{A}\Gamma \subseteq \mathcal{A}\Gamma$.

Suppose $A = \mathcal{A}P$ with $P_\sigma \in \mathcal{A}\Gamma$, that is, $P_\sigma = \mathcal{A}Q_{\sigma,\tau}$ with $Q_{\sigma,\tau} \in \Gamma$. Then

$$\begin{aligned} z \in A &\iff \exists\alpha\,\forall m\,(z \in P_{\alpha|_m}) \\ &\iff \exists\alpha\,\forall m\,\exists\beta\,\forall n\,(z \in Q_{\alpha|_m,\beta|_n}) \\ &\iff \exists\alpha\,\exists\beta\,\forall m\,\forall n\,(z \in Q_{\alpha|_m,(\beta)_m|_n}), \end{aligned}$$

where $(\beta)_m$ denotes the m-th column of β.

Now we contract the two function quantifiers to a single one, using a (computable) homeomorphism $\mathbb{N}^\mathbb{N} \times \mathbb{N}^\mathbb{N}$, and the two universal number quantifiers into a single one using the paring function $\langle .,. \rangle$. Then A can be characterized as

$$z \in A \iff \exists\gamma\,\forall k(z \in R_{\gamma|_k})$$

where $R_\sigma = Q_{\varphi(\sigma),\psi(\sigma)} \in \Gamma$ for suitable coding functions φ, ψ. We leave an explicit definition of these coding functions as an exercise.

$\square$

Corollary 0.5.4.1.

$$\mathcal{A}\mathbf{\Sigma}^1_1 = \mathbf{\Sigma}^1_1. \tag{159}$$

Lebesgue measurability of analytic sets

We start with a lemma that essentially says that we can envelop any set with a smallest (up to measure 0) measurable set.

Lemma 0.5.5. *For every set $A \subseteq \mathbb{R}$ there exists a set $B \subseteq \mathbb{R}$ so that*

- ***(i)*** *$A \subseteq B$ and B is Lebesgue measurable,*
- ***(ii)*** *if B' is such that $A \subseteq B' \subseteq B$ and is Lebesgue measurable, then $B \setminus B'$ has measure 0.*

Proof. Suppose first that $\lambda^*(A) < \infty$. For every $n \geq 0$, there exists an open set $O_n \supseteq A$ with $\lambda^*(O_n) = \lambda(O_n) < \lambda^*(A) + 1/n$. Then $B = \bigcap_n O_n$ is measurable, and $\lambda(B) = \lambda^*(A)$. Furthermore, if $A \subseteq B' \subseteq B$, then $\lambda^*(A) \leq \lambda^*(B') \leq \lambda^*(B)$.

If B' is also measurable, then

$$\lambda^*(B) = \lambda^*(B \cap B') + \lambda^*(B \setminus B') = \lambda^*(B') + \lambda^*(B \setminus B'), \tag{160}$$

hence $\lambda^*(B \setminus B') = 0$.

If $\lambda^*(A) = \infty$, let $A_n = A \cap [m, m+1)$ for $m \in \mathbb{Z}$. Then $\lambda^*(A_m) \leq 1$, and we can choose $B_m \supseteq A_m$ measurable such that $\lambda^*(B_m) = \lambda^*(A_m)$. Then $B = \bigcup_{m \in \mathbb{Z}} B_m$ has the desired property.

□

We now apply the lemma to show that Lebesgue measurability is closed under the Souslin operation. The basic idea is to approximate the local "branches" of the Souslin operation on a Souslin scheme by measurable sets from outside, in the sense of the lemma. It turns out that the total error we make by this approximation is negligible, and hence the overall result of the Souslin operation differs from a measurable set only by a nullset and hence is measurable.

Proposition 0.5.3. *The class* **LM** *of all Lebesgue measurable sets $\subseteq \mathbb{R}$ is closed under the Souslin operation, that is,*

$$\mathcal{A}\,\mathbf{LM} \subseteq \mathbf{LM}. \tag{161}$$

Proof. Let $A = (A_\sigma)$ be a Souslin scheme with each A_σ measurable. We can assume that (A_σ) is regular. For each $\sigma \in \mathbb{N}^{<\mathbb{N}}$ we let

$$A^\sigma = \bigcup_{\alpha \supset \sigma} \bigcap_{n \in \mathbb{N}} A_{\alpha|_n} \subseteq A_\sigma.$$

Note that $A^\varnothing = \mathcal{A}\,A$.

By the previous lemma, there exist measurable sets $B^\sigma \supseteq A^\sigma$ so that for every measurable $B \supseteq A^\sigma$, $B^\sigma \setminus B$ is null.

By replacing B^σ with $B^\sigma \cap A_\sigma$, we can further assume $B^\sigma \subseteq A_\sigma$, and also that (B^σ) is a regular Souslin scheme.

Now let $C_\sigma = B^\sigma \setminus \bigcup_{n \in \mathbb{N}} B^{\sigma^\frown \langle n \rangle}$. Each C_σ is a nullset, by the choice of the B^σ and the fact that $A^\sigma = \bigcup_{n \in \mathbb{N}} A^{\sigma^\frown \langle n \rangle} \subseteq \bigcup_{n \in \mathbb{N}} B^{\sigma^\frown \langle n \rangle}$. Hence $C = \bigcup_\sigma C_\sigma$ is a nullset, too.

It remains to show that

$$B^\varnothing \setminus C \ \subseteq \ A^\varnothing = \mathcal{A}\,A,$$

for this implies $B^\varnothing \setminus A^\varnothing \subseteq C$ is null, which in turn implies that $A^\varnothing$ is Lebesgue measurable (since it differs from a measurable set by a nullset).

So let $x \in B^\varnothing \setminus C$. Since $x \notin C_\varnothing$, there is an $\alpha(0)$ with $x \in B^{\langle \alpha(0) \rangle}$.

Given $\alpha\,|_n$ with $x \in B^{\alpha|_n}$, we can choose $\alpha(n)$ so that $x \in B^{\alpha|_{n+1}}$. This is possible because $x \notin C_{\alpha|_n}$. This way we construct $\alpha \in \mathbb{N}^\mathbb{N}$ with

$$x \in \bigcap_n B^{\alpha|_n} \subseteq \bigcap_n A_{\alpha|_n} \subseteq A^\varnothing.$$

□

Corollary 0.5.5.1. *Every analytic set is Lebesgue measurable.*

Proof. By the idempotence of $\mathcal{A}$, $\mathcal{A}\mathbf{\Sigma}^1_1 = \mathcal{A}\mathcal{A}\mathbf{\Pi}^0_1 = \mathcal{A}\mathbf{\Pi}^0_1 = \mathbf{\Sigma}^1_1$. On the other hand, we have $\mathcal{A}\mathbf{\Pi}^0_1 \subseteq \mathcal{A}\mathbf{LM} = \mathbf{LM}$, since the Souslin operation is monotone on classes. This yields $\mathbf{\Sigma}^1_1 \subseteq \mathbf{LM}$.

□

Universally measurable sets

The previous proof is general enough to work for other kinds of measures on arbitrary Polish spaces.

Given a Polish space X, a **Borel measure** on X is a countably additive set function μ defined on a σ-algebra of the Borel sets in X. A set is **μ-measurable** if it can be represented as a union of a Borel set and a μ-nullset. A measure μ is **σ-finite** if $X = \bigcup_n X_n$, where X_n is μ-measurable with $\mu(X_n) < \infty$. Lebesgue measure is σ-finite Borel measure on the Polish space $\mathbb{R}$.

A set $A \subseteq X$ is **universally measurable** if it is μ-measurable for every σ-finite Borel measure on X.

Theorem 0.5.6 (Lusin). *In a Polish space, every analytic is universally measurable.*

Baire property of analytic sets

Inspecting the proof of Proposition **??**, we see that it works for the Baire property as well (with "measure 0" replaced by "meager", of course), provided we can prove a Baire category version of Lemma 0.5.5.

Lemma 0.5.7. *Let X be a Polish space. For every set $A \subseteq X$ there exists a set $B \subseteq X$ so that*

- ***(i)*** *$A \subseteq B$ and B has the Baire property,*
- ***(ii)*** *if $Z \subseteq B \setminus A$ and Z has the Baire property, then Z is meager.*

Proof. Let $U_1, U_2, \dots$ be an enumeration of countable base of the topology for X. Given $A \subseteq \mathbb{R}$ set

$$A^* := \{x \in \mathbb{R} \colon \forall i\ (x \in U_i \Rightarrow U_i \cap A \text{ not meager})\}.$$

Note that A^* is closed: If $x \notin A^*$, then there exists i with $x \in U_i$ & $U_i \cap A$ null. If $y \in U_i$, then $y \notin A^*$, since $U_i \cap A$ is null. Hence $U_i \subseteq \neg A^*$.

We have

$$A \setminus A^* = \bigcup \{A \cap U_i \colon A \cap U_i \text{ meager}\},$$

which is a countable union of meager sets and hence meager.

If we let $B = A \cup A^* = A^* \cup (A \setminus A^*)$, then B is a union of a meager set and a closed set and hence has the Baire property.

Now assume $B' \supseteq A$ has the Baire property. Then $C = B \setminus B'$ has the Baire property, too. Suppose C is not meager, then $U_i \setminus C$ is meager for some i, and hence also $U_i \cap A \subseteq (U_i \setminus C)$. Besides, $U_i \cap C \neq \emptyset$, for otherwise $U_i \subseteq U_i \setminus C$ would be meager. Thus there exists $x \in U_i$ with $x \notin A^*$, which by definition of A^* implies that $U_i \cap A$ is not meager, a contradiction.

□

By adapting the proof of Proposition **??**, we obtain the Baire category version and hence can deduce that analytic sets have the Baire property.

Proposition 0.5.4. *In any Polish space X, the class* $\mathbf{BP}$ *of all sets $\subseteq X$ with the Baire property is closed under the Souslin operation, i.e.*

$$\mathcal{A}\,\mathbf{BP} \subseteq \mathbf{BP}. \tag{162}$$

0.5.3 The Projective Hierarchy

We saw in the previous chapters that analytic sets are projections of closed sets and hence can be written as

$$x \in A \iff \exists \alpha \in \mathbb{N}^{\mathbb{N}}\, F(\alpha, x), \tag{163}$$

where $F \subseteq \mathbb{N}^{\mathbb{N}} \times X$ is closed. It follows that co-analytic sets can be written in the form

$$x \in A \iff \forall \alpha \in \mathbb{N}^{\mathbb{N}}\, U(\alpha, x),$$

for some open $U \subseteq \mathbb{N}^{\mathbb{N}} \times X$.

Using quantifier manipulations that allow to switch number and function quantifiers,

$$\begin{aligned} \forall m\, \exists \alpha\; P(m, \alpha) &\iff \exists \beta\, \forall m\, P(m, (\beta)_m) \\ \exists m\, \forall \alpha\; P(m, \alpha) &\iff \forall \beta\; \exists m\; P(m, (\beta)_m), \end{aligned}$$

we obtain that *both* the analytic sets and the co-analytic sets are closed under countable unions and intersections.

We have seen (Proposition 0.5.1) that the analytic sets are closed under continuous images. Taking continuous images of co-analytic sets, however, leads out of the co-analytic sets.

Using continuous images (or rather, the special case of **projections**), we define the **projective hierarchy**. Recall our notation $\exists^{\mathbb{N}}$ for projection along $\mathbb{N}$, with $\forall^{\mathbb{N}}$ its dual. We denote by $\exists^{\mathbb{N}^{\mathbb{N}}}$ and $\forall^{\mathbb{N}^{\mathbb{N}}}$ projection along $\mathbb{N}^{\mathbb{N}}$ and its dual, respectively.

$$\begin{aligned} \mathbf{\Sigma}^1_1(X) &= \exists^{\mathbb{N}^{\mathbb{N}}}\, \mathbf{\Pi}^0_1(X) \\ \mathbf{\Pi}^1_n(X) &= \neg\, \mathbf{\Sigma}^1_n(X) \\ \mathbf{\Sigma}^1_{n+1}(X) &= \exists^{\mathbb{N}^{\mathbb{N}}} \mathbf{\Pi}^1_n(X) \\ \mathbf{\Delta}^1_n(X) &= \mathbf{\Sigma}^1_n(X) \cap \mathbf{\Pi}^1_n(X) \end{aligned}$$

Hence a set $P \subseteq X$ is

$$\begin{array}{lll r} \mathbf{\Sigma}^1_1 & \text{iff} & P(x) \Leftrightarrow \exists \alpha\; F(\alpha, x) & \text{for a closed set } F \subseteq \mathbb{N}^{\mathbb{N}} \times X, \\ \mathbf{\Pi}^1_1 & \text{iff} & P(x) \Leftrightarrow \forall \alpha\; G(\alpha, x) & \text{for an open set } G \subseteq \mathbb{N}^{\mathbb{N}} \times X, \\ \mathbf{\Sigma}^1_2 & \text{iff} & P(x) \Leftrightarrow \exists \alpha \forall \beta\; G(\alpha, \beta, x) & \text{for an open set } G \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \times X, \\ \mathbf{\Pi}^1_2 & \text{iff} & P(x) \Leftrightarrow \forall \alpha \exists \beta\; F(\alpha, \beta, x) & \text{for a closed set } F \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \times X, \\ & & \vdots & \end{array}$$

These characterizations clearly indicate a relation between being projective and being definable in second order arithmetic using function quantifiers.

The effective projective hierarchy

We have seen that the Borel sets of finite order correspond to the sets definable (from parameters) by formulas using only number quantifiers (**arithmetical formulas**). A similar relation holds between projective sets and sets definable by formulas using both number and function quantifiers. In fact, the way we defined the projective hierarchy makes this easy to see.

Historically, however, the topological approach and the definability approach happened separately, the former devised by the Russian school of Souslin, Lusin, and others, while the effective approach was pursued by Kleene. Kleene named the sets in his effective hierarchy *analytical* sets, which to this day is a source of much confusion.

Definition 0.5.3 (Kleene)**.** • *A set $A \subseteq \mathbb{N}^{\mathbb{N}}$ is (lightface) Σ^1_1 if there exists a computable relation $R(\sigma, \tau)$ such that*

$$\alpha \in A \iff \exists \beta \forall n \, R(\alpha\,|_n, \beta\,|_n) \tag{164}$$

- *$A \subseteq \mathbb{N}^{\mathbb{N}}$ is (lightface) Π^1_n if $\neg A$ is Σ^1_n.*
- *$A \subseteq \mathbb{N}^{\mathbb{N}}$ is (lightface) Σ^1_{n+1} if it is $\exists^{\mathbb{N}^{\mathbb{N}}}\Pi^1_{n+1}$, that is, if it a projection of a Π^1_{n+1} relation along $\mathbb{N}^{\mathbb{N}}$.*
- *A set that is Σ^1_n and Π^1_n at the same time is called Δ^1_n.*

As before, we can relativize this hierarchy with respect to a parameter $\gamma \in \mathbb{N}^{\mathbb{N}}$, by requiring R to be computable only relative to γ. This gives rise to classes $\Sigma^1_n(\gamma)$, $\Pi^1_n(\gamma)$, and $\Delta^1_n(\gamma)$. Then the Theorem 0.4.16 can be extended as follows:

$$\boldsymbol{\Sigma}^1_n = \bigcup_{\gamma \in \mathbb{N}^{\mathbb{N}}} \Sigma^1_n(\gamma) \qquad \boldsymbol{\Pi}^1_n = \bigcup_{\gamma \in \mathbb{N}^{\mathbb{N}}} \Pi^1_n(\gamma) \tag{165}$$

To complete the connection with definability, we also have following analogue to Theorem **??**.

Theorem 0.5.8. *A set $A \subseteq \mathbb{N}^{\mathbb{N}}$ is $\boldsymbol{\Sigma}^1_n$ ($\boldsymbol{\Pi}^1_n$) if and only if it is definable, relative to some $\gamma \in \mathbb{N}^{\mathbb{N}}$, by a Σ^1_n (Π^1_n) formula in second order arithmetic.*

Examples of projective sets

Here are a few examples of projective sets that occur naturally in mathematics.

Analytic sets:

- $\{K \subseteq X \colon K$ compact and uncountable$\}$ is a $\boldsymbol{\Sigma}^1_1$ subset of the space $K(X)$ of compact subsets of X.
- $\{f \in \mathcal{C}[0,1] \colon f$ continuously differentiable on $[0,1]\}$ is a $\boldsymbol{\Sigma}^1_1$ subset of $\mathcal{C}[0,1]$.

Co-analytic sets:

- $\{f \in \mathcal{C}[0,1] \colon f$ differentiable on $[0,1]\}$ is a $\boldsymbol{\Pi}^1_1$ subset of $\mathcal{C}[0,1]$.
- $\{f \in \mathcal{C}[0,1] \colon f$ nowhere differentiable on $[0,1]\}$ is a $\boldsymbol{\Pi}^1_1$ subset of $\mathcal{C}[0,1]$.
- WF $= \{\alpha \in 2^{\mathbb{N}} \colon \alpha$ codes a well -founded tree on $\mathbb{N}\}$ is a $\boldsymbol{\Pi}^1_1$ subset of the space Tr of trees, which can be seen as a closed subspace of $2^{\mathbb{N}^{<\mathbb{N}}}$, and hence is Polish. As we will see, the set WF is a prototypical $\boldsymbol{\Pi}^1_1$ set.

Higher levels:

- $\{f \in \mathcal{C}[0,1] \colon f$ satisfies the Mean Value Theorem $[0,1]\}$ is a $\boldsymbol{\Pi}^1_2$ subset of $\mathcal{C}[0,1]$.

(Here f satisfies the Mean Value Theorem if for all $a < b \in [0,1]$ there exists c with $a < c < b$ such that $f'(c)$ exists and $f(b) - f(a) = f'(c)(b-a)$.)

Some structural properties of the projective hierarchy

The quantifier manipulations mentioned above yield the following closure properties.

Proposition 0.5.5. • ***(i)*** *The classes $\boldsymbol{\Sigma}^1_n$ are closed under continuous preimages, countable intersections and unions, and continuous images (in particular, $\exists^{\mathbb{N}^{\mathbb{N}}}$).*

- ***(ii)*** *The classes $\boldsymbol{\Pi}^1_n$ are closed under continuous preimages, countable intersections and unions, and co-projections $\forall^{\mathbb{N}^{\mathbb{N}}}$.*
- ***(iii)*** *The classes $\boldsymbol{\Delta}^1_n$ are closed under continuous preimages, complements, countable intersections and unions. (In particular, they form a σ-algebra.)*

To show that the hierarchy is proper, we need the existence of universal sets.

Proposition 0.5.6. *For every Polish space X, there is a $\mathbb{N}^{\mathbb{N}}$-universal set for $\boldsymbol{\Sigma}^1_n$ and for $\boldsymbol{\Pi}^1_n$.*

Proof. By induction on n. We have seen (Theorem **??**) that there exists a $\mathbb{N}^{\mathbb{N}}$-universal set for $\mathbf{\Sigma}^1_1$. Now note that if $U \in \mathbf{\Sigma}^1_n(\mathbb{N}^{\mathbb{N}} \times X)$ is $\mathbb{N}^{\mathbb{N}}$-universal for $\mathbf{\Sigma}^1_n(X)$, then $\neg U$ is $\mathbb{N}^{\mathbb{N}}$-universal for $\mathbf{\Pi}^1_n(X)$, and if $U \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \times X$ is $\mathbb{N}^{\mathbb{N}}$-universal for $\mathbf{\Pi}^1_n(\mathbb{N}^{\mathbb{N}} \times X)$, then

$$V = \{(\alpha, z) \colon \exists\beta\, (\alpha, \beta, z) \in U\}$$

is $\mathbb{N}^{\mathbb{N}}$-universal for $\mathbf{\Sigma}^1_{n+1}$.

□

Corollary 0.5.8.1. *For every $n \geq 1$, $\mathbf{\Sigma}^1_n \not\subseteq \mathbf{\Pi}^1_n$ and $\mathbf{\Pi}^1_n \not\subseteq \mathbf{\Sigma}^1_n$. Morover,*

$$\mathbf{\Sigma}^1_n \subsetneq \mathbf{\Delta}^1_{n+1} \subsetneq \mathbf{\Sigma}^1_{n+1}$$
$$\mathbf{\Pi}^1_n \subsetneq \mathbf{\Delta}^1_{n+1} \subsetneq \mathbf{\Pi}^1_{n+1}$$

The proof is similar to the proofs of Theorem **??** and Corollary 0.4.19.1.

Regularity properties of projective sets

We have seen (Corollary 0.5.5.1 and Proposition **??**) that all analytic sets are Lebesgue measurable and have the Baire property. Since these properties are closed under complements, it follows that the same holds for co-analytic ($\mathbf{\Pi}^1_1$) sets. Analytic sets also have the perfect-set property, but if you worked out the exercise, you will see that the proof does not carry over to complements of analytic sets. Can we find a different proof?

Similarly, it does not seem impossible to extend the regularity properties (LM) and (BP) to higher levels of the projective hierarchy. We will soon see that there are metamathematical limits that prevent us from doing so.

Without explicitly mentioning it, up to now we have been working in ZF, Zermelo-Fraenkel set theory, plus a weak form of Choice ($\mathsf{AC}_\omega(\mathbb{N}^{\mathbb{N}})$). If we add the full Axiom of Choice (AC), we saw that the regularity properties do not extend to all sets. Solovay's model of $\mathsf{ZF} + \mathsf{DC}$ shows that the use of a strong version of Choice is necessary for this.

On the other hand, the proofs gave us no direct indication how 'complex' the non-regular sets we constructed are. We will study a model of ZF in which exists a $\mathbf{\Delta}^1_2$ set which neither is Lebesgue measurable nor has the Baire property. This, together with the Solovay model, shows we cannot settle in ZF alone the question of whether the projective sets are measurable or have the Baire property. We would have to add additional axioms.

A key feature in the construction of a non-measurable $\mathbf{\Delta}^1_2$ set is the use of the well-ordering principle rather than the Axiom of Choice.

Proposition 0.5.7. *Suppose $<_W \subseteq \mathbb{R} \times \mathbb{R}$ is a well-ordering of $\mathbb{R}$ of order-type ω_1, then the set*

$$A = \{(x, y) \colon x <_W y\} \tag{166}$$

neither is Lebesgue measurable nor has the Baire property.

Lebesgue measure here refers to the product measure $\lambda\times\lambda$, which is the unique translation invariant measure defined on the Borel σ-algebra generated by the rectangles $I \times J$, where I and J are open intervals, and $(\lambda \times \lambda)(I \times J) = \lambda(I)\lambda(J)$.

Proof. Since $<_W$ is of order type ω_1, for every $y \in \mathbb{R}$, the set $A_y = \{x \colon x <_W y\}$ is countable, and hence of Lebesgue measure zero.

Fubini's Theorem implies that if $A \subseteq \mathbb{R}^2$ is measurable, then

$$(\lambda \times \lambda)(A) = \int \lambda(A_y) d\lambda(y) = 0.$$

So if A is measurable, then $(\lambda \times \lambda)(A) = 0$. The complement of A is $\neg A = \{(x, y) \colon x \geq_W y\}$. As above, for any $x \in \mathbb{R}$, $(\neg A)_x = \{y \colon x \geq_W y\}$ is countable, and hence $\lambda(\neg A)_x = 0$ for all x.

Again, by Fubini's Theorem, $(\lambda \times \lambda)(\neg A) = 0$, and thus $(\lambda \times \lambda)(\mathbb{R}) = (\lambda \times \lambda)(A \cup \neg A) = (\lambda \times \lambda)(A) + (\lambda \times \lambda)(\neg A) = 0$, a contradiction.

We can apply a similar reasoning for Baire category, using the Lemma below. The sections A_y and $\neg A_x$ are countable, and hence meager.

$\square$

The following lemma provides a Baire category analogue to Fubini's Theorem.

Lemma 0.5.9. *Let $A \subseteq \mathbb{R}^2$ have the property of Baire. Then A is meager if and only if $A_x = \{y \colon (x, y) \in A\}$ is meager for all x except a meager set.*

For a proof see **?**.

Therefore, if the Continuum hypothesis (CH) holds in a model and we can well-order $\mathbb{R}$ (or $\mathbb{N}^{\mathbb{N}}$, $2^{\mathbb{N}}$) within a certain complexity (as a subset of $\mathbb{R}^2$), we can find a non-regular set of the same complexity. The question now becomes how (hard it is) to define a well-ordering of $\mathbb{R}$, and of course if CH holds.

0.5.4 Co-Analytic sets

We will see that, in many ways, $\mathbf{\Pi}^1_1$ sets form the frontier between classical descriptive set theory and metamathematics. This chapter can be seen as the start of our transition to metamathematics. We will detail the distinguished role well-founded relations play in the analysis of $\mathbf{\Pi}^1_1$ sets.

Normal forms

Analytic sets are projections of closed sets. Closed sets are in $\mathbb{N}^\mathbb{N} \times \mathbb{N}^\mathbb{N}$ are infinite paths through **trees on** $\mathbb{N} \times \mathbb{N}$, i.e. two-dimensional trees.

Definition 0.5.4. *A set $T \subseteq \mathbb{N}^{<\mathbb{N}} \times \mathbb{N}^{<\mathbb{N}}$ is a* ***two-dimensional tree*** *if*

- ***(i)*** *$(\sigma, \tau) \in T$ implies $|\sigma| = |\tau|$ and*
- ***(ii)*** *$(\sigma, \tau) \in T$ implies $(\sigma\,|_n, \tau\,|_n) \in T$ for all $n \leq |\sigma|$.*

An ***infinite branch*** *of T is a pair $(\alpha, \beta) \in \mathbb{N}^\mathbb{N} \times \mathbb{N}^\mathbb{N}$ so that*

$$\forall n \in \mathbb{N}\ (\alpha\,|_n, \beta\,|_n) \in T.$$

As in the one-dimensional case, we use $[T]$ to denote the set of all infinite paths through T. It follows that $A \subseteq \mathbb{N}^\mathbb{N}$ is analytic if and only if there exists a two-dimensional tree T on $\mathbb{N} \times \mathbb{N}$ such that

$$\begin{aligned} \alpha \in A &\iff \exists \beta\, (\alpha, \beta) \in [T] \\ &\iff \exists \beta\, \forall n\, (\alpha\,|_n, \beta\,|_n) \in T. \end{aligned}$$

Another way to write this is to put, for given T and $\alpha \in \mathbb{N}^\mathbb{N}$,

$$T(\alpha) = \{\tau \colon (\alpha\,|_{|\tau|}, \tau) \in T\}. \tag{167}$$

Then we have, with T witnessing that A is analytic,

$$\alpha \in A \iff T(\alpha) \text{ has an infinite path } \iff T(\alpha) \text{ is not well-founded.} \tag{168}$$

We obtain the following normal form for co-analytic sets.

Proposition 0.5.8 (Normal form for co-analytic sets)**.** *A set $A \subseteq \mathbb{N}^\mathbb{N}$ is $\mathbf{\Pi}^1_1$ if and only if there exists a two-dimensional tree T such that*

$$\alpha \in A \iff T(\alpha) \textit{ is well-founded.} \tag{169}$$

If A is (lightface) Π^1_1, then there exists a computable such T, and the mapping $\alpha \mapsto T(\alpha)$ is computable, as a mapping between reals and trees (which can be coded by reals). This relativizes, i.e. for a $\Pi^1_1(\gamma)$ set, the mapping $\alpha \mapsto T(\alpha)$ is computable in γ. Since computable mappings are continuous, we obtain that the in the above proposition, the mapping $\alpha \mapsto T(\alpha)$ is continuous.

$\mathbf{\Pi}^1_1$-complete sets

How does one show that a specific set is *not* Borel? A related question is: Given a definition of a set in second order arithmetic, how can we tell that there is not an easier definition (in the sense that it uses less quantifier changes, no function quantifiers etc.)? The notion of **completeness** for classes in Polish spaces provides a general method to answer such questions.

Definition 0.5.5. *Let X, Y be Polish spaces. We say a set $A \subseteq X$ is* ***Wadge reducible*** *to $B \subseteq Y$, written $A \leq_{\mathrm{W}} B$, if there exists a continuous function $f : X \to Y$ such that*

$$x \in A \iff f(x) \in B. \tag{170}$$

The important fact about Wadge reducibility is that it preserves classes closed under continuous preimages.

Proposition 0.5.9. *Let Γ be a family of subsets in Polish spaces (such as the classes of the Borel or projective hierarchy). If Γ is closed under continuous preimages, then $A \leq_{\mathrm{W}} B$ and $B \in \Gamma$ implies $A \in \Gamma$.*

Proof. If $A \leq_{\mathrm{W}} B$ via f, then $A = f^{-1}(B)$.

□

Definition 0.5.6. *A set $A \subseteq X$ is Γ-**complete** if $A \in \Gamma$ and for all $B \in \Gamma$, $B \leq_{\mathrm{W}} A$.*

Γ-complete sets can be seen as the most complicated members of Γ. In particular, for the $\mathbf{\Sigma}/\mathbf{\Pi}$ classes complete sets cannot be members of the dual class. For instance, a $\mathbf{\Pi}^1_1$-complete set cannot be $\mathbf{\Sigma}^1_1$, since this would mean it is Borel, and hence every $\mathbf{\Pi}^1_1$ set would be Borel, which we have seen is not true.

If $A \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}}$ is $\mathbb{N}^{\mathbb{N}}$-universal for some class Γ in the Borel or projective hierarchy, then the set

$$\{\alpha \oplus \beta \colon (\alpha, \beta) \in A\} \tag{171}$$

is Γ-complete, where $\oplus$ here denotes the pairing function for reals

$$\alpha \oplus \beta(n) = \begin{cases} \alpha(k) & n = 2k, \\ \beta(k) & n = 2k+1. \end{cases} \tag{172}$$

Since $\oplus$ is continuous, and $B \in \Gamma$ if and only if $B = A_\gamma$ for some $\gamma \in \mathbb{N}^{\mathbb{N}}$, we have in that case that $B \leq_{\mathrm{W}} A$ via the mapping

$$f(\beta) = \gamma \oplus \beta. \tag{173}$$

It follows that complete sets exist for all levels of the Borel and projective hierarchy. However, the universal sets they are based on are rather abstract objects. Complete sets are most useful when we can show that a *specific property* implies completeness. We will encounter next an important example for the class of co-analytic sets.

Well-founded relations and well-orderings

Given a real in $\beta \in \mathbb{N}^{\mathbb{N}}$, we can associate with β a binary relation E_β on $\mathbb{N}$:

$$E_\beta(m, n) :\iff \beta(\langle m, n\rangle) = 0. \tag{174}$$

A binary relation E on a set X is **well-founded** if every non-empty $Y \subseteq X$ has an E-minimal element y_0, that is, there is no $z \in Y$ with $E(z, y)$.

Let

$$\mathrm{WF} = \{\beta \in \mathbb{N}^{\mathbb{N}} \colon E_\beta \text{ is well-founded}\}. \tag{175}$$

Then

$$\beta \in \mathrm{WF} \iff \forall \gamma \in \mathbb{N}^{\mathbb{N}}\, \exists n\, \forall m\, [\gamma(n) E_\beta \gamma(m)], \tag{176}$$

and hence WF is Π^1_1.

A closely related set is

$$\mathrm{WOrd} = \{\beta \in \mathbb{N}^{\mathbb{N}} \colon E_\beta \text{ is a well-ordering}\}. \tag{177}$$

Then

$$\beta \in \mathrm{WOrd} \iff \beta \in \mathrm{WF} \text{ and } E_\beta \text{ is a linear ordering}. \tag{178}$$

Coding a linear order is easily seen to be Π^0_1, hence WOrd is Π^1_1, too.

Theorem 0.5.10. *The sets* WF *and* WOrd *are $\mathbf{\Pi}^1_1$-complete.*

Proof. We have seen in the chapter on Trees that a tree has an infinite path if and only if the inverse prefix ordering is ill-founded. Trees can be coded as reals, and hence Proposition 0.5.8 yields immediately that WF is $\mathbf{\Pi}^1_1$-complete.

For WOrd we use the Kleene-Brouwer ordering and refer to Proposition **??**.

□

The theorem lets us gain further insights in the structure of co-analytic sets. If $\alpha \in \mathbb{N}^{\mathbb{N}}$ codes a well-ordering on $\mathbb{N}$, let

$$\|\alpha\| = \text{ order type of well-ordering coded by } \alpha.$$

It is clear that $\|\alpha\| < \omega_1$. For a fixed ordinal $\xi < \omega_1$, we let

$$\mathrm{WOrd}_\xi = \{\alpha \in \mathrm{WOrd} \colon \|\alpha\| \leq \xi\}.$$

Lemma 0.5.11. *For any $\xi < \omega_1$, the set* WOrd_ξ *is Borel.*

Proof. Let $\alpha \in \mathbb{N}^{\mathbb{N}}$. We say $m \in \mathbb{N}$ is in the **domain** of E_α, $m \in \mathrm{dom}(E_\alpha)$, if

$$\exists n\, [mE_\alpha n \ \vee \ nE_\alpha m].$$

It is clear from the definition of E_α that $\mathrm{dom}(E_\alpha)$ is Borel. For $\xi < \omega_1$, let

$$B_\xi = \{(\alpha, n) \colon E_\alpha|_{\{m \colon mE_\alpha n\}} \text{ is a well-ordering of order type } \leq \xi\}$$

We show by transfinite induction that every B_ξ is Borel. Suppose B_ζ is Borel for all $\zeta < \xi$. Then, since ξ is countable, $\bigcup_{\zeta<\xi} B_\zeta$ is Borel, too. But

$$(\alpha, n) \in B_\xi \iff \forall m\, [mE_\alpha n \Rightarrow (\alpha, m) \in \bigcup_{\zeta<\xi} B_\zeta],$$

and from this it follows that B_ξ is Borel. Finally, note that

$$\alpha \in \mathrm{WOrd}_\xi \iff \forall n\, [n \in \mathrm{dom}(E_\alpha) \Rightarrow (\alpha, n) \in B_\xi],$$

which implies that WOrd_ξ is Borel.

□

Corollary 0.5.11.1. *Every $\mathbf{\Pi}^1_1$ set is a union of $\aleph_1$ many Borel sets.*

Proof. Since WOrd is $\mathbf{\Pi}^1_1$-complete, every co-analytic set A is the preimage of WOrd for some continuous function f. We have

$$\mathrm{WOrd} = \bigcup_{\xi<\omega_1} \mathrm{WOrd}_\xi,$$

and hence

$$A = \bigcup_{\xi<\omega_1} f^{-1}(\mathrm{WOrd}_\xi).$$

Since continuous preimages of Borel sets are Borel, the result follows.

□

If we work instead with the set

$$C_\xi = \{\alpha \colon \alpha \in \mathrm{WOrd}_\xi \text{ or } \exists n \in \mathrm{dom}(E_\alpha)\},$$
$E_\alpha|_{\{m \colon mE_\alpha n\}}$ is a well-ordering of order type ξ

then we get that $\mathrm{WOrd} = \bigcap_{\xi<\omega_1} C_\xi$, and hence

Corollary 0.5.11.2. *Every $\mathbf{\Pi}^1_1$ set can be obtained as a union or intersection of $\aleph_1$-many Borel sets. Consequently, the same holds for every $\mathbf{\Sigma}^1_1$ set.*

The previous results allow us to solve the cardinality problem of co-analytic sets at least partially.

Corollary 0.5.11.3. *Every* $\mathbf{\Pi}^1_1$ *set is either countable, of cardinality* $\aleph_1$*, or of cardinality* $2^{\aleph_0}$*.*

We conclude the chapter with another application of Lemma **??** – a useful tool for analyzing $\mathbf{\Sigma}^1_1$ sets:

Theorem 0.5.12 (-bounding)**.** $\mathbf{\Sigma}^1_1$
For every analytic $A \subseteq \mathrm{WOrd}$ *there exists an ordinal* $\nu < \omega_1$ *such that*

$$\forall x \in A \quad \| x \| < \nu. \tag{179}$$

Proof. If such a ν did not exist, then

$$\alpha \in \mathrm{WOrd} \iff \exists \nu \, [\alpha \in A \ \wedge \ \mathrm{WOrd}_\nu]. \tag{180}$$

The right-hand side is $\mathbf{\Sigma}^1_1$, and hence WOrd would be $\mathbf{\Sigma}^1_1$, contradiction.

□

An analogous statement holds for WF, with respect to the rank function ρ of a well-founded relation.

0.6 Axiomatic set theory

0.6.1 The Axioms of Set Theory

In the previous chapters, we have repeatedly brought up a metamathematical context, such as the use of the Axiom of Choice, or Solovay's model in which every set of reals is measurable. But we have not really distinguished between results in a formal theory and in the metatheory, mostly because we did not really establish a formal theory to begin with. We have treated descriptive set theory like most other mathematical theories in that we defined our basic objects (Polish spaces, Borel sets, etc.) and then started proving facts about them "the usual way", as we would prove facts about commutative rings or locally compact topological spaces. But in order to make better sense of the metamathematical issues, we have to "pause" and talk a bit about the axioms of set theory.

Comprehension and Russell Antinomy

To develop set theory formally as a theory of first order logic, we first need to fix the **language**. We want to consider the notion of a *set* as foundational (with the intention to develop everything else from it), we provide only **one binary relation symbol**, $\in$. We will denote this as the **language of set theory**, $\mathcal{L}_{\in}$.

How can we axiomatize the concept of set?

In his famous 1895 paper "*Beiträge zu Begründung der transfiniten Mengenlehre*" **?**, Georg Cantor writes

> Unter einer "*Menge*" verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die "*Elemente*" von M genannt werden) zu einem Ganzen.

This can be translated approximately as: *A set is any collection of certain distinct objects of our intuition or our thought into a whole.*

We can try to make this more precise as follows:

> For every property $P(x)$ exists a set M of all objects x having property P: $M = \{x : M(x)\}$.

This can be formalized in the language of set theory as an **axiom scheme**: For every $\mathcal{L}_{\in}$-formula $\varphi(x)$,

> (**Comprehension**)$_{\varphi}$ $\exists y \, \forall x \; (x \in y \leftrightarrow \varphi(x))$

This axiom, however, is **contradictory**.

Russell's antinomy (1903)
If we choose $\varphi(x) = \neg x \in x$, then the Comprehension axiom yields the existence of a set r with

$$\forall x (x \in r \leftrightarrow x \notin x).$$

In particular, for $x = r$:

$$r \in r \leftrightarrow r \notin r,$$

contradiction!

We obtain similar contradictions if we choose as $\varphi(x)$ the formula

$x = x$	antinomy of the set of all sets (Cantor)
x is a cardinal	Cantor, around 1899, published 1932
x is an ordinal	antinomy of Burali-Forti

These antinomies are, however, not as direct as Russell's and require some further development of the theory in order to derive a contradiction.

Regarding the existence of sets, we have to distinguish between

- **classes**, which we will denote by capital letters $A, B, \ldots$ (for some specific, important classes we will also use boldface) and
- **sets**, denoted in this context by lower case letters $a, b, \ldots, x, y \ldots$.

An arbitrary property $\varphi(x)$ will define a corresponding class

$$A = \{x \colon \varphi(x)\} \tag{181}$$

As we saw above, classes are not necessarily sets: some are "*too large*" to be a set, as for example the **class of all sets**,

$$\mathbf{V} = \{x \colon x = x\} \tag{182}$$

Classes that are not sets are called **proper classes**.

You should keep in mind that, in the formal theory ZF, we do not have variables for classes, so the definition above is *informal*. Any class variable, as well as expressions like $a \in A$, should be seen as *abbreviations* for a formal expression using the underlying formula. There are formal systems (such as **Bernays-Gödel set theory**) that use classes explicitly, but they are used less frequently.

The Axioms of ZFC

We begin with the **Axiom of Extensionality**, which is essential for the equality relation between sets.

> (**Extensionality**) $\quad \forall x(x \in a \leftrightarrow x \in b) \to a = b.$

Consequently, two sets coincide if they have the same elements.

The basic idea of the **Zermelo-Fraenkel** axiom system ZF is that we avoid introducing sets that are "too large" (and hence would lead to contradictions) by allowing new sets only if they can be **"generated" from a given set** by a number of fixed, well-behaved operations.

So let us postulate that at least one set exists:

> (**Set Existence**) $\quad \exists x(x = x)$

This axiom is not strictly necessary, as the existence of a set also follows from other axioms in ZF (or usually even from the underlying axioms of first-order logic). But it is good to have it as a starting point here for emphasis.

We have seen we cannot use full comprehension for sets. In its place we introduce the scheme of

> (**Separation**)$_\varphi$ $\quad \forall a \exists y \forall x(x \in y \leftrightarrow x \in a \wedge \varphi(x, \ldots))$

By Extensionality, the set y is unique.

Separation allows us to select from any class $\{x \colon \varphi(x, \ldots)\}$ those elements that are in a given set a and collect them in a **set**

$$\{x \in a \colon \varphi(x, \ldots)\}$$

Next, we have

> (**Pairing**) $\quad \exists y \forall x(x \in y \leftrightarrow x = a \vee x = b).$

This axioms allows forming *pairs of sets*, specifically

$$\{a,b\} := \{x \colon x = a \ \vee \ x = b\}$$
$$\{a\} := \{x \colon x = a\} \qquad \text{singleton set}$$
$$(a,b) := \{\{a\},\{a,b\}\} \qquad \text{ordered pair}$$

While for the *pair set* $\{a,b\} = \{b,a\}$ the order is not important, we have for the *ordered pair*

$$(a,b) = (c,d) \leftrightarrow a = c \wedge b = d.$$

Hence, we can introduce **binary relations** as classes of ordered pairs

$$\mathrm{Rel}(R) :\leftrightarrow \forall u \in R\, \exists x, y\ (u = (x,y)).$$

As usual, by identifying **functions** with their graphs, we can introduce functions as a special kind of relation:

$$F : A \to B :\leftrightarrow \forall x \in A\ \exists! y \in B\ (x,y) \in F.$$

We write $\mathrm{Fun}(a)$ to denote that fact that a is a function.
Further elementary axioms:

(**Union**) $\forall a \exists y \forall z (z \in y \leftrightarrow \exists x \in a\ z \in x)$

(**Replacement**)$_{\mathrm{F}}$

$$\forall a ((F : a \to \mathbf{V}) \ \to \ \exists u \forall y (y \in u \leftrightarrow \exists x \in a\ y = F(x)))$$

(**Power Set**) $\forall a \exists y \forall z (z \in y \leftrightarrow z \subseteq a)$

It follows that, for a given set a, the following classes are sets:

$$\bigcup a = \bigcup_{x \in a} x := \{z \colon \exists x \in a\ z \in x\} \qquad \text{union}$$
$$F[a] = \{F(x) | x \in a\} := \{y \colon \exists x \in a\ y = F(x)\} \qquad \text{image set}$$
$$\mathcal{P}(a) := \{x \colon x \subseteq a\} \qquad \text{power set}$$

(**Infinity**) $\exists x (\emptyset \in x \wedge \forall y (y \in x \to y \cup \{y\} \in x))$

The axiom of Infinity is a “pure” set existence axiom, that is, it does not depend on another set already existing. It therefore renders the axiom of Set Existence above redundant. It also implies the existence of the set $\mathbb{N}$ of natural numbers (along with the operation $+$), which we will address in more detail below.

Using $\mathbb{N}$, we can introduce the other basic number sets:

- $\mathbb{Z} = (\mathbb{N} \times \mathbb{N}) / \sim_{\mathbb{Z}}$, where $(x,y) \sim_{\mathbb{Z}} (u,v) :\Leftrightarrow x + v = y + u$. Multiplication on $\mathbb{Z}$ can be defined inductively (see below).
- $\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z}) / \sim_{\mathbb{Q}}$, where $(x,y) \sim_{\mathbb{Q}} (u,v) :\Leftrightarrow xv = yu$.
- We can extend the linear order $<$ of $\mathbb{N}$ to $\mathbb{Z}$ and then to $\mathbb{Q}$ in the usual way. Then we can introduce the **real numbers** as the set of **Dedekind cuts**:

$$\mathbb{R} = \{x \in \mathcal{P}(\mathbb{Q}) \colon x \neq \emptyset \ \wedge \ x \neq \mathbb{Q} \ \wedge \ \forall z \in x \forall y \in \mathbb{Q}\, y < z \to y \in x\}. \tag{183}$$

The final axiom of ZF is

(**Foundation**) $\forall a\, (a \neq \emptyset \to \exists x \in a\, \forall y \in x\, y \notin a)$.

Foundation rules out, for example, that a set can be an element of itself. More precisely, the axiom states that $\in$-relation is **well-founded** on any set.

We can also formalize the **Axiom of Choice**:

(**Choice**) $\forall a (\forall x \in a\; x \neq \emptyset \quad \to \quad \exists f (\mathrm{Fun}(f) \;\wedge\; \mathrm{dom}(f) = a \;\wedge\; \forall x \in a\; f(x) \in x))$

We denote the axiom system $\mathsf{ZF} + \mathsf{AC}$ as ZFC – **Zermelo-Fraenkel with Choice**.

0.6.2 Recursion and the Von-Neumann Hierarchy

Transfinite induction

While the class Ord of all ordinals is not a set, it is still transitive and well-ordered by $\in$. Regarding the associated order $\leq$, every *set* of ordinals a has a **supremum** $\bigcup a = \bigcup_{\xi \in a} \xi$ and (if $a \neq \emptyset$) an **infimum** $\bigcap a = \bigcap_{\xi \in a} \xi$, which is the *smallest element* of a. Such a smallest element exists actually for every (non-empty) *class* A (since if $\xi \in A$, we only need to find the infimum of the *set* of ordinals $\leq \xi$.) This allows us to prove properties about *all* ordinals by **induction**.

Proposition 0.6.1 (Induction for ordinals, I)**.** *For every property φ,*

$$\forall \alpha \, [\forall \xi < \alpha \, \varphi(\xi) \to \varphi(\alpha)] \to \forall \alpha \, \varphi(\alpha).$$

We have repeatedly used induction already for ordinals $< \omega_1$, the first uncountable ordinal.

To prove this principle simply observe that if $\forall \alpha \, \varphi(\alpha)$ failed there would have to be a *smallest* α with $\neg\varphi(\alpha)$, contradicting the induction hypothesis.

Since every ordinal is either 0, a successor, or a limit ordinal, we have the following variant of induction.

Proposition 0.6.2 (Induction for ordinals, II)**.** *For every property φ, if*

- ***(i)*** $\varphi(0)$*,*
- ***(ii)*** $\forall \alpha(\varphi(\alpha) \to \varphi(\alpha + 1))$*, and*
- ***(iii)*** $(\forall \xi < \lambda \, \varphi(\xi)) \to \varphi(\lambda)$ *for all limit* λ*,*

then $\forall \alpha \, \varphi(\alpha)$*.*

(i) and (ii) coincide with the usual induction scheme for natural numbers. To cover *all* ordinals we need to add (iii).

Ordinal recursion

The induction principle can be used to define functions by **recursion**. For example, **addition** on the natural numbers is given by

$$\begin{aligned} x + 0 &= x \\ x + (y + 1) &= (x + y) + 1. \end{aligned}$$

In the case of ordinals, we have to consider the limit case, too.

Theorem 0.6.1 (Recursion on ordinals)**.** *If $G : \mathrm{Ord} \times \mathrm{V} \longrightarrow \mathrm{V}$ is a function and a is a set, then there exists a unique function $F : \mathrm{Ord} \longrightarrow \mathrm{V}$ such that for all $\alpha \in \mathrm{Ord}$,*

$$F(\alpha) = G(\alpha, F\,|_{\alpha}) \tag{184}$$

Proof. The uniqueness of the function F follows by induction.

To show the existence of F, we define the following:

- Call h *tame* if
$$\exists \alpha \, (h : \alpha \to \mathrm{V} \wedge \forall \xi \in \alpha \; h(\xi) = G(\xi, h\,|_{\alpha}))$$
- Say h is *compatible* with g if
$$\forall x \in \mathrm{Dom}(h) \cap \mathrm{Dom}(g) \; h(x) = g(x)$$

It follows by induction that any two tame functions are compatible.

This lets us define the desired F as

$$F := \bigcup \{h\colon h \text{ tame}\} \tag{185}$$

Then F is a function (otherwise there would be two incompatible tame functions), its domain is transitive, and satisfies the recursion condition (since it is the union of tame functions).

It remains to show that F is defined on all of Ord. If $D = \mathrm{Dom}(F) \neq \mathrm{Ord}$, then we would have $D = \alpha$ for some ordinal α. In particular B is a set therefore $F = f$ is a set, for some tame f. This f could be extended to a tame $h = f \cup \{(\alpha, G(\alpha, f\restriction_\alpha))\}$, contradiction.

□

Note that we defined F **explicitly** as a *union* of all partial solution to the recursion equation.

As with induction, we have the following variant of the recursion principle.

Proposition 0.6.3 (Recursion on ordinals, variant)**.** *If $G, H : \mathrm{Ord} \times \mathrm{V} \longrightarrow \mathrm{V}$ are functions and a is a set, then there exists a unique function $F : \mathrm{Ord} \longrightarrow \mathrm{V}$ such that*

$$\begin{aligned} F(0) &= a \\ F(\alpha+1) &= G(\alpha, F(\alpha)) \\ F(\lambda) &= H(\lambda, \{F(\xi) \colon \xi < \lambda\}) \quad \textit{for } \mathrm{Lim}(\lambda). \end{aligned}$$

We can establish a similar principle for a well-ordering $<$ on a class A. In case of a proper class, though, we have to require that for every $a \in A$, the class of all **predecessors** of a

$$S(a, <) := \{x \in A \colon x < a\},$$

is a set (if A is a set, this follows automatically by *Separation*). If this is the case, the recursion principle yields a function $F : A \to \mathrm{V}$ such that

$$F(a) = G(a, F\restriction_{S(a,<)}).$$

Recursion for well-founded relations

More generally, we can define induction and recursion on **well-founded** relations. We already encountered those in a previous chapter.

Definition 0.6.1. *A relation R on a class A is **well-founded** if it satisfies the **minimality condition***

$$(\mathrm{Min}_R) \qquad \emptyset \neq b \subset A \to \exists x \in b\, \forall y \in b\, (\neg yRx)$$

*and the **set condition***

$$\forall x \in A\ S(a, R) := \{x \colon xRa\} \textit{ is a set}$$

If A is a set, the minimality condition is again automatically satisfied by *Separation*.

The set condition allows for taking the R**-transitive closure** of a set $a \in A$: the smallest superset $\mathrm{TC}_R(a)$ of a that is R*-transitive*:

$$\forall x \in \mathrm{TC}_R(a)\ S(x, R) \subseteq \mathrm{TC}_R(a) \tag{186}$$

This is done by recursion over the natural numbers. The following is an important example.

Example 0.6.1 (Transitive closure of a set)**.** *By the axiom of Foundation, $\in$ is a well-founded relation on* V*. (The set condition is satisfied since $S(a, \in) = a$.)*

*We can form the **transitive closure**, the smallest transitive superset, of a set a as*

$$\mathrm{TC}(a) := a \cup \bigcup a \cup \bigcup\bigcup a \ldots = \bigcup_{n<\omega} U^n(a), \qquad \textit{where}$$
$$U^0(a) = a, U^{n+1}(a) = \bigcup U^n(a).$$

This is an example of definition by recursion along $\mathbb{N}$.

We can use the existence of TC_R as a set to strengthen the minimality condition to *subclasses*, similar to the case of the well-ordering of Ord:

Lemma 0.6.2. *For every non-empty class $B \subseteq A$, there exists $x \in B$ such that*

$$\forall y \in B \, \neg yRx \tag{187}$$

To prove this lemma, simply pick any $x \in B$, take its transitive R-closure, and intersect it with B:

$$C = \mathrm{TC}_R(x) \cap B. \tag{188}$$

C is a set, and by the minimality condition (Min_R) has an R-minimal element a. a has to be minimal for B, too, since otherwise there exists $b \in B$ with bRa. Since $a \in \mathrm{TC}_R(x)$, $b \in \mathrm{TC}_R(x)$, and therefore $b \in C$, contradicting the minimality of a.

The lemma implies a corresponding **induction principle for well-founded relations**:

$$(\mathrm{Ind}_R) \qquad \forall x \in A[\forall y(yRx \to \varphi(y)) \to \varphi(x)] \to \forall x \in A \, \varphi(x)).$$

This in turn yields the following.

Theorem 0.6.3 (Recursion principle for well-founded relations)**.** *Let R be a well-founded relation on a class A. The for every function $G : A \times \mathrm{V} \longrightarrow \mathrm{V}$ exists a unique function $F : A \to \mathrm{V}$ such that*

$$F(a) = G(a, F \restriction_{\{x|xRa\}}) \text{ for all } a \in A.$$

The Von-Neumann hierarchy

Is there a way to systematically build the V, the universe of all sets, "from below"?

We start with the empty set:

$$V_0 = \emptyset \tag{189}$$

Given V_α, the *Power Set* axiom requires the set of all subsets to exist, so we set

$$V_{\alpha+1} = \mathcal{P}(V_\alpha). \tag{190}$$

Finally, at limit stages we simply collect all sets we have obtained so far:

$$V_\lambda = \bigcup_{\xi<\lambda} V_\xi \quad \text{for limit } \lambda \tag{191}$$

What we really are doing here is to construct a function $V : \mathrm{Ord} \to \mathrm{V}$ by ordinal recursion. Think $V_\alpha = V(\alpha)$.

Remarkably, if we assume the axiom of *Foundation*, we reach *all* sets this way.

Theorem 0.6.4. *For every set x there exists an ordinal α with $x \in V_\alpha$, that is,*

$$\mathrm{V} = \bigcup_{\alpha \in \mathrm{Ord}} V_\alpha \tag{192}$$

Proof. Let C be the class of all sets not in any V_α. Since $\in$ is well-founded, if C is non-empty, it has a $\in$-minimal element x. This implies that for all $z \in x$, $z \in \bigcup_{\alpha\in\mathrm{Ord}} V_\alpha$. Define a function h by mapping each $z \in x$ to the *least* α so that $z \in V_\alpha$. Since x is a set, $h[x]$ is a set of ordinals, by *Replacement*. This set or ordinals has a supremum, say γ. Then $x \subseteq V_\gamma$ and therefore,

$$x \in \mathcal{P}(V_\gamma) = V_{\gamma+1}. \tag{193}$$

Hence C must be empty, and the theorem follows.

□

We can now split the question of "how large" V is into two sub-questions:

- How "**long**" is V, that is, how many ordinals are there? Axioms for **large cardinals** attempt to extend this "length" as far as possible.
- How "**wide**" is V, that is, how large is the power set of a set? A rather "slim" universe is given by the **constructible sets**, which we will encounter soon.

0.6.3 Models of Set Theory

You may have noticed that, when introducing the axioms of ZFC, we never *really* answered the question "*What is a set?*". Instead, we developed a formal theory of axioms for a binary relation that somehow describe "*how sets work*", that is, how we can obtain sets from given ones using well-known operations like power set and union.

We have then seen how we can develop a lot of standard mathematical *objects* (like $\mathbb{N}$, $\mathbb{R}$) and *techniques* (like induction and definition by recursion) **inside** this formal system. In fact, most of mathematics can be developed formally inside this system. Almost all proofs you find in any standard math book are proofs that can be formalized in ZFC. It is very tedious to do this for us humans, but there is little doubt it can be done, and in fact, looking at the recent work on **proof assistants** (like Coq or Lean), many parts of mathematics have been formalized (albeit not directly in ZFC).

This expressiveness gives ZFC its foundational importance, but it is also the cause for much confusion for someone who first studies set theory.

From a pedagogical point of view, in what follows it is helpful to assume a "**Platonist**" perspective of mathematics, and set theory in particular, namely that *sets and the relations between exist independently (and outside) of the* ZFC *axioms.* The set of real numbers exists, and our development of $\mathbb{R}$ inside ZFC is just a formal way to describe them. From this perspective, the axioms of ZF (AC is a little different) are just obvious truths about sets, just like the **Peano axioms** are obvious truths about natural numbers.

Among other things, this perspective allows us to treat ZF just like any other mathematical theory, like *group theory* or the theory of *algebraically closed fields.* In particular, we can think about **models of set theory** the way we would think about models of group theory, in the sense of model theory.

A **model** would simply be a set M together with a binary relation E on S such that

$$(M, E) \models \mathsf{ZF}, \tag{194}$$

that is, all axioms hold when interpreted in (M, E). Note that we use "*set*" in this context not in the formal sense, but in the "meta"-sense (the Platonist world of sets).

Working in the meta-theory ("*that what is mathematically true*"), we know by Gödel's completeness theorem that

> if ZF is consistent, then it has a model.

This model should be seen as a **set-theoretic universe**: Its elements can be seen as sets, and the interpretation E of the $\in$-symbol will tell us how these sets are connected via the element-relation.

Note that E does not have to be the *actual* element relation on a set (of sets), but just some binary relation so that the axioms are satisfied.

In the meta-world, there are, of course, sets other than M, but that does not matter here, since al we are interested in is giving *some* universe in which our axioms hold. (Timothy Chow has suggested that set theory should rather be called "*universe theory.* He is right in the sense that what axiomatic set theory does is to define such *universes of sets*, rather than what a set is.)

In the meta-theory, we can then follow the usual techniques to show provability or non-provability results.

If we want to prove that CH is consistent with ZF (assuming ZF is consistent), we need to find a model in which both hold.

One difficulty in working with models of set theory is that they can look very different depending on whether you look at a model "**from the inside**" or "**from the outside**".

To illustrate this, assume ZF is consistent. Then, by the **Löwenheim-Skolem theorem**, there exists a **countable model** for ZF. Yet it is a theorem of ZF that *there exists an uncountable set.* This is often referred to as **Skolem's paradox**, although it is not really an antinomy.

If we break this down a bit, we see that the apparent paradox is really just a matter of perspective (*inside* or *outside*). Assume (M, E) is a countable model of ZF. Then there exists $x \in M$ such that there is no injection from x to the natural numbers. Since M is countable, x can have at most countably many elements. So why is this *not* a contradiction? We should really read the statement above as

> there is no injection **in** M from x to M**'s version** of the natural numbers.

In other words, even though x is countable *from the outside*, x appears uncountable *inside* M since a mapping witnessing its countability does not exist in M.

This is a first warning sign that models of ZF can behave in very unexpected ways. For another example, recall the axiom of *Foundation* asserts that the $\in$-relation is well-founded. But again, this means only "from the inside".

Proposition 0.6.4. *If* ZF *is consistent, than there exists a model* (M, E) *of* ZF *such that* (M, E) *is ill-founded.*

Proof. Introduce new constant symbols c_n $(n \in \mathbb{N})$ and add the formulas $\varphi_n \equiv c_{n+1} \in c_n$ to the axioms of ZF. It is not hard to show, using the compactness theorem, that $\mathsf{ZF} + \bigcup_n \varphi_n$ has a model (M^*, E^*), for which the set $\{c_n : n \in \mathbb{N}\}$ is ill-founded.

□

Since, as mentioned above, the model (M^*, E^*) satisfies *Foundation*, the set $\{c_n : n \in \mathbb{N}\}$ is actually *not in* the model (and neither can be any other set with an infinite descending $\in$-chain).

Mostowski collapse

If we restrict ourselves to models on which the E-relation is *actually* well-founded (i.e. from the outside), then interestingly these models look in way "*natural*": They can be assumed to be the $\in$-relation on a set. Such models are also called **standard**.

Given a set theoretic structure (M, E) (not necessarily a model of ZF), for each $x \in M$ let

$$\operatorname{ext}_E(x) = \{y \in X : y \, E \, x\}$$

If E behaves "set-like", then it will respect the axiom of *Extensionality*, i.e. two sets are identical if and only if they have the same elements. Therefore we say that E is **extensional** if

$$x, z \in X, \; x \neq z \quad \text{implies} \quad \operatorname{ext}_E(x) \neq \operatorname{ext}_E(z).$$

Furthermore, as stated above, we want to exclude infinite descending E-chains. We say that E is **well-founded** if

> every non-empty set $Y \subseteq X$ has an E-minimal element.

Theorem 0.6.5 (Mostowski collapse)**.** *If E is an extensional and well-founded relation on a set X, then there exists a transitive set S and a bijection $\pi : X \to S$ such that*

$$x \, E \, y \iff \pi(x) \in \pi(y) \quad \textit{for all } x, y \in X.$$

Moreover, S and π are unique.

Proof. We construct π and $S = \operatorname{im}(\pi)$ by recursion on E, which is possible since it is well-founded.

For each $x \in X$, let
$$\pi(x) = \{\pi(y) : y \, E \, x\},$$
and set $S = \operatorname{im}(\pi)$.

The injectivity of π follows from the extensionality of π by induction along E: Suppose we have shown
$$\forall z \; (zEx \to \forall y \in X(\pi(z) = \pi(y) \to z = y)).$$

and we have to show that it holds for x. Assume $\pi(x) = \pi(y)$ for some $y \in X$. Then

$$\begin{aligned} cEx \Rightarrow &\quad \pi(c) \in \pi(x) = \pi(y) \\ \Rightarrow &\quad \pi(c) = \pi(z) \qquad \text{for some } zEy \\ \Rightarrow &\quad c = z \qquad \text{(by ind. hyp., since } cEx) \\ \Rightarrow &\quad cEy. \end{aligned}$$

Similarly, we get $cEy \Rightarrow cEx$, hence $x = y$ as desired due to the extensionality of E. Finally we have

$$\begin{aligned} \pi(x) \in \pi(y) \Rightarrow &\quad \pi(x) = \pi(c) \qquad \text{for some } cEy \\ \Rightarrow &\quad x = c \qquad \text{(since } \pi \text{ is injective)} \\ \Rightarrow &\quad xEy \end{aligned}$$

Thus π is an isomorphism.

To see the uniqueness of π and S, assume ρ, T are such that the statement of the theorem is satisfied. Then $\pi \circ \rho^{-1}$ is an isomorphism between $(T, \in)$ and $(S, \in)$. Now apply the following lemma.

□

Lemma 0.6.6. *Suppose X, Y are sets, and θ is an isomorphism between $(X, \in)$ and $(Y, \in)$. Then $X = Y$ and $\theta(x) = x$ for all $x \in X$.*

Proof. By induction on the well-founded relation $\in$. Assume that $\theta(z) = z$ for all $z \in x$ and let $y = \theta(x)$.

We have $x \subseteq y$ because if $z \in x$, then $z = \theta(z) \in \theta(x) = y$.

We also have $y \subseteq x$: Let $t \in y$. Since $y \in Y$, there is $z \in X$ with $\theta(z) = t$. Since $\theta(z) \in y$ and $y = \theta(x)$, we have $z \in x$, and thus $t = \theta(z) = z \in x$.

Hence $x = y$, and this also implies $\theta(x) = x$.

□

0.6.4 Large Cardinals

Inaccessible cardinals

The cardinality of V_α grows rather fast relative to α. For example,

$$|V_{\omega+\alpha}| = \beth_\alpha \tag{195}$$

where the **beth function** $\beth_\alpha$ is defined as

$$\begin{aligned} \beth_0 &= \aleph_0 \\ \beth_{\alpha+1} &= 2^{\beth_\alpha} \\ \beth_\lambda &= \sup\{\beth_\alpha \colon \alpha < \lambda\} \qquad \lambda \text{ limit} \end{aligned}$$

This presents difficulties for the axiom of *Replacement* to hold in a V_α, since we could define a function on a set of sufficiently high cardinality that maps to sets in V_α whose ranks are cofinal in α (and the image would not be an element of V_α).

The existence of **inaccessible cardinals** ensures that the von-Neumann hierarchy is "long enough" for α to eventually "catch up" with the cardinality of V_α.

Recall the enumeration of all cardinals by means of the **$\aleph$-sequence**:

$$\aleph_0 = \omega, \quad \aleph_{\alpha+1} = \aleph_\alpha^+, \quad \aleph_\lambda = \sup\{\aleph_\xi \colon \xi < \lambda\} \text{ for limit } \lambda.$$

Here κ^+ is the least cardinal $> \kappa$. Some cardinals are limits of short sequences of cardinals – for example,

$$\aleph_\omega = \lim_n \aleph_n$$

is uncountable, but a limit of a countable sequence of smaller cardinals. Generally, cardinals who are a limit of a sequence of cardinals of length smaller than their cardinality are called **singular**. Non-singular cardinals are called **regular**:

$$\mathrm{reg}(\kappa) :\iff \forall \alpha < \kappa\, \forall f\ (f : \alpha \to \kappa \ \to\ \sup_{\xi<\alpha} f(\xi) < \kappa).$$

In other words, a regular cardinal κ cannot be reached by less then κ-many steps. The first example of a regular cardinal is $\aleph_0$.

> **Exercise**
> Show that all **successor cardinals**, i.e. cardinals of the form $\aleph_{\alpha+1}$ are regular. (Use the Axiom of Choice.)

On the other hand,

$$\aleph_\omega, \aleph_{\omega+\omega}, \aleph_{\aleph_\omega}, \aleph_{\aleph_{\aleph_\omega}}, \ldots$$

are singular and this suggests the question:

> Are there regular cardinals of the form $\aleph_\lambda$ with λ limit?

This is captured by the notion of **inaccessibility**.

Definition 0.6.2 (Hausdorff 1908, Tarski, Zermelo 1930)**.** *An uncountable cardinal $\kappa > \omega$ is*

- ***weakly inaccessible*** *if*

$$\begin{aligned} &\mathrm{reg}(\kappa) \ \wedge\ \exists\lambda(\lim(\lambda) \ \wedge\ \kappa = \aleph_\lambda) \\ &\quad (\Leftrightarrow \mathrm{reg}(\kappa) \ \wedge\ \forall \alpha < \kappa\ \alpha^+ < \kappa) \end{aligned}$$

- ***(strongly) inaccessible*** *if*

$$\operatorname{reg}(\kappa) \ \wedge \ \forall \alpha < \kappa \ \ 2^{\alpha} < \kappa$$

Under the **Generalized Continuum Hypothesis**,

(GCH) $\forall \alpha \ \ 2^{\aleph_\alpha} = \aleph_\alpha^+$

weakly and strongly inaccessible cardinals coincide.
If $\kappa > \omega$ is inaccessible, then $\kappa = \aleph_\kappa$. Moreover, we have

Proposition 0.6.5. *If κ is strongly inaccessible, $|V_\kappa| = \kappa$.*

Proof. It suffices to show that $|V_\alpha| < \kappa$ for all $\alpha < \kappa$. This follows by a straightforward induction, using the fact that κ is strongly inaccessible.

□

This in turn implies we can bound the cardinality of elements of V_κ.

Proposition 0.6.6. *Suppose κ is strongly inaccessible and $x \subset V_\kappa$. Then*

$$x \in V_\kappa \ \Leftrightarrow \ |x| < \kappa. \tag{196}$$

Proof. ($\Rightarrow$) $x \in V_\kappa$ implies $|x| < |V_\kappa|$. Apply Proposition **??**.
($\Leftarrow$) Since $x \subseteq V_\kappa$, each $y \in x$ has rank $< \kappa$. Since $|x| < \kappa$, by regularity of κ,

$$\operatorname{rank}(x) = \sup\{\operatorname{rank}(y) + 1 \colon y \in x\} < \kappa \tag{197}$$

which implies $x \in V_\kappa$.

□

We have already seen that for limit $\alpha > \omega$, V_α is a model of all ZFC axioms except *Replacement.*

Theorem 0.6.7. *If κ is strongly inaccessible, then $V_\kappa \models \mathsf{ZFC}$.*

Proof. We verify that V_κ satisfies the axiom of *Replacement.* Suppose $x \in V_\kappa$ and $f : x \to V_\kappa$ is a function. Then $f[x] \subseteq V_\kappa$, and by Proposition 0.6.6, $|f[x]| \leq |x| < \kappa$. Applying the other direction of Proposition 0.6.6 to $f[x]$, we obtain $f[x] \in V_\kappa$, as desired.

□

Suppose an inaccessible cardinal exists, and let κ be the least inaccessible.
It is not hard to verify that

$$V_\kappa \models \mathsf{ZFC} + \text{"there does not exist an inaccessible cardinal"}. \tag{198}$$

(You verify that being a inaccessible cardinal is absolute for V_κ.) Therefore, the existence of an inaccessible cardinal is not provable from ZFC. This fact also follows from Gödel's second incompleteness theorem.

Measurability

We have seen that (assuming the Axiom of Choice) there subsets of $\mathbb{R}$ that are not Lebesgue measurable. Inspecting the proof, we see that we only use the following properties of Lebesgue measure:

- σ-additivity,
- translation invariance ($\lambda(A) = \lambda(A + r)$),
- $\lambda(A) > 0$ for some A.

For spaces without an additive structures, instead of translation invariance, we can consider a **non-triviality condition**:

$$m(\{x\}) = 0 \quad \text{for all } x \tag{199}$$

The **generalized measure problem** asks whether there exists a set M together with a measure function

$$m : \mathcal{P}(M) \to [0, \infty),$$

so that the following conditions are met:

- (**M1**) $m(M) = 1$
- (**M2**) $\forall x \in M\ m(\{x\}) = 0$
- (**M3**) if $(A_i)_{i<\omega}$ is a countable sequence of disjoint sets $\subseteq M$, then

$$m\left(\bigcup_{i<\omega} A_i\right) = \sum_{i<\omega} m(A_i)$$

The structure of the set M does not play any role here, so we can replace it by a cardinal κ outright. One can also consider strengthening σ-additivity to κ**-additivity**:

> If $\gamma < \kappa$ and $(A_\xi)_{\xi<\lambda}$ is a sequence of disjoint subsets of κ, then

$$m(\bigcup_{\xi<\gamma} A_\xi) = \sum_{\xi<\gamma} m(A_\xi).$$

A transfinite sum $\sum_{\xi<\gamma}$ is given as the supremum of all sums over finite subsequences:

$$\sum_{\xi<\gamma} m(A_\xi) = \sup\left\{\sum_{\xi\in F} m(A_\xi) \colon F \subseteq \gamma \text{ finite}\right\}. \tag{200}$$

Hence, ω_1-additive is the same as σ-additive.

Theorem 0.6.8 (Banach)**.** *If κ is the least cardinal for which a measure satisfying (M1)-(M3) exists, then any such measure on κ is already κ-additive.*

Proof. Suppose m is a measure on κ that is not κ-additive. Then, for some $\gamma < \kappa$, there exists a sequence $(A_\xi)_{\xi<\gamma}$ of disjoint subsets of κ so that

$$m(\bigcup_{\xi<\gamma} A_\xi) \neq \sum_{\xi<\gamma} m(A_\xi). \tag{201}$$

Since a measure is always σ-additive, $\gamma > \omega$ has to hold, and there can be at most countably many A_ξ with $m(A_\xi) > 0$.

We can drop those A_ξ, and by the σ-additivity of m for the remaining ξ it has to hold that $m(A_\xi) = 0$ while $m\left(\bigcup_{\xi<\gamma} A_\xi\right) = r > 0$.

By putting

$$\overline{m}(X) = \frac{m(\bigcup_{\xi\in X} A_\xi)}{r}$$

we obtain a measure on $\gamma < \kappa$, contradicting the minimality of κ.

□

Measurable cardinals If m is a measure on κ, the **associated ideal**

$$\mathcal{I}_m = \{x \subseteq \kappa \colon m(x) = 0\} \tag{202}$$

is a σ-ideal, or, complementing the notion of ω_1-additivity, a ω_1**-complete ideal**.

> **Exercise**
> Show that $\mathcal{I}_m$ is not principal.

The corresponding filter

$$\mathcal{F}_m = \{x \subseteq \kappa \colon m(x) = 1\} \tag{203}$$

is then ω_1-complete, too.

A measure m is **two-valued** if it only assumes the values 0 and 1. In this case the corresponding filter $\mathcal{F}_m$ is an **ultrafilter** (and $\mathcal{I}_m$ is a **prime ideal**).

Conversely, if U is ω_1-complete, non-principal ultrafilter on κ, we can define a two-valued measure $m : \mathcal{P}(\kappa) \to \{0, 1\}$ on κ by letting

$$m(x) = \begin{cases} 1 & \text{if } x \in U, \\ 0 & \text{otherwise.} \end{cases} \tag{204}$$

Definition 0.6.3. *Let κ be an uncountable cardinal.*

- *κ is **real-valued measurable** if there exists a κ-additive measure on κ.*
- *κ is **measurable** if there exists a κ-additive, two-valued measure on κ, or, equivalently, if there exists a κ-complete, non-principal ultrafilter on κ.*

In the following, we will see that measurability implies inaccessibility.

Lemma 0.6.9. *If U is a κ-complete, non-principal ultrafilter on κ, then every $X \in U$ has cardinality κ.*

Proof. Since U is non-principal, no *singleton* set $\{x\}$ can be in U (for this would imply $\kappa \setminus \{x\} \notin U$ and therefore no subset of it would be in U either, contradicting the non-principality of U).

If $X \in U$ and $|X| < \kappa$, then X is the union of $< \kappa$ many singletons. Since $\neg U$ is a κ-complete prime ideal, this implies $X \in \neg U$, contradiction.

□

Proposition 0.6.7. *If κ is measurable, then it is regular.*

Proof. If κ were singular, it would be the union of $< \kappa$-many sets of cardinality $< \kappa$. Applying Lemma 0.6.9 leads to a contradiction.

□

Theorem 0.6.10. *A measurable cardinal is (strongly) inaccessible.*

Proof. By Proposition 0.6.7, any measurable cardinal is regular. Assume for a contradiction there exists $\gamma < \kappa$ with $2^\gamma > \kappa$. As $2^\gamma > \kappa$, there exists a set S of functions $f : \gamma \to \{0, 1\}$ with $|S| = \kappa$. Let U be a κ-complete, non-principal ultrafilter on S.

For $\alpha < \gamma, i \in \{0, 1\}$, let

$$X_{\alpha,i} = \{f \in S \colon f(\alpha) = i\} \tag{205}$$

and let $g(\alpha) = i$ if and only if $X_{\alpha,i} \in U$. Since U is an ultrafilter, g is well-defined on γ.

Since $\gamma < \kappa$ and U is κ-complete,

$$X = \bigcap_{\alpha<\gamma} X_{\alpha,g(\alpha)} \tag{206}$$

is in U. But $|X| \leq 1$, since the only function possibly in X is g. This contradicts Lemma 0.6.9. $\square$

Exercise
Show that every real-valued measurable cardinal is weakly inaccessible.

Proposition 0.6.8. *If κ is real-valued measurable, then κ is measurable or $\kappa \leq 2^{\aleph_0}$.*

Thus, if κ is real-valued measurable but not measurable, then the continuum $2^{\aleph_0}$ has to be very large.

Partition properties

Another concept of largeness is related to the existence of large **homogeneous sets** for partitions.

For given set S and $n \in \mathbb{N}$, let

$$[S]^n := \{X \subseteq S\colon\ |X| = n\}$$

be the set of all n-element subsets of S. For cardinals κ, λ, we define

$$\kappa \to (\lambda)^n_k$$

to mean that any partition $F : [S]^n \to \{1, \dots, k\}$ mit $|S| = \kappa$ has an F**-homogeneous subset** of cardinality λ, that is, a set H, $|H| = \lambda$, such that

$$F|_{[H]^n} \equiv \text{ constant.}$$

Ramsey's theorem (1929/39) says that for any $n, k \in \mathbb{N}$,

$$\aleph_0 \to (\aleph_0)^n_k.$$

Do there exist uncountable cardinals with similar properties?

A cardinal κ is **weakly compact** if it is uncountable and $\kappa \to (\kappa)^2_2$ holds.

Exercise
Show that for any cardinal κ, $2^\kappa \nrightarrow (\kappa^+)^2_2$, and use this to infer that any weakly compact cardinal is inaccessible.
(Thus the existence of weakly compact cardinals cannot be established in ZFC.)

Measurable cardinals have even stronger homogeneity properties. Let $[S]^{<\omega}$ be the set of all finite subsets of S. If $F : [S]^{<\omega} \to I$ is a partition, then $H \subseteq S$ is F**-homogenenous** if

$$F|_{[H]^n} \equiv \text{ constant}$$

for all $n \in \mathbb{N}$.

Theorem 0.6.11 (Rowbottom)**.** *Let κ be a measurable cardinal and let $F : [\kappa]^{<\omega} \to \lambda$ a partition of $[\kappa]^{<\omega}$ into $\lambda < \kappa$ pieces. Then there exists an F-homogeneous set $H \subseteq \kappa$ with $|H| = \kappa$.*

In general, any cardinal that satisfies the statement of the theorem is called **Ramsey**.

To prove Theorem **??**, we introduce **normal ultrafilters**.

Definition 0.6.4. *Given a sequence of sets $(A_\xi)_{\xi<\gamma}$, the **diagonal intersection** is given as*

$$\Delta_{\xi<\gamma} A_\xi = \{\alpha < \gamma\colon \alpha \in \bigcap_{\xi<\alpha} A_\xi\}. \tag{207}$$

*A filter F on a cardinal κ is **normal** if for any κ-sequence $(A_\xi)_{\xi<\kappa}$, $A_\xi \in F$, the diagonal intersection $\Delta_{\xi<\kappa} A_\xi$ is in F.*

Let us assume as a convention that a filter on a cardinal κ always contains the end-segments $\{\xi \colon \alpha \leq \xi < \kappa\}$.

Exercise
Show that a normal filter on κ is κ-complete.

Exercise
Show that if there is a normal filter over κ, then κ is uncountable and regular.

Exercise
Show that if κ is measurable, then there is a normal ultrafilter on κ.

Proof. (Proof of Theorem **??**)

Let U be a normal filter over κ. We show that for every n, for any $g : [\kappa]^n \to \gamma$ with $\gamma < \kappa$, there is a set $H_n \in U$ such that $g_n \restriction_{[H_n]^n} \equiv$ const. The intersection of the H_n is again in U and satisfies the statement of the the theorem.

We proceed by induction. The case $n = 1$ follows from the κ-completeness of U. Now assume $g : [\kappa]^{n+1} \to \gamma$, with $\gamma < \kappa$.

For each $S \in [\kappa]^n$, define $g_s : \kappa \to \gamma$ by

$$g_S(\alpha) = \begin{cases} g(S \cup \{\alpha\}) & \text{if } \max S < \alpha \\ 0 & \text{otherwise} \end{cases} \tag{208}$$

By κ-completeness of U, g_S is constant on a set $Y_S \in U$, say

$$g_S \restriction_{Y_S} \equiv \delta_S < \gamma. \tag{209}$$

We now define a function $h : [\kappa]^n \to \gamma$ by letting

$$h(S) = \delta_S. \tag{210}$$

By induction hypothesis, h is constant on a set $Z \subseteq \kappa$ in U (and hence of size κ), say $h \restriction_{[Z]^n} \equiv \delta < \kappa$.

For each $\alpha < \kappa$, let

$$Y_\alpha = \bigcap \{Y_S \colon \max S \leq \alpha\} \tag{211}$$

By κ-completeness, $Y_\alpha \in U$, and by normality

$$H = Z \cap \Delta_{\alpha<\kappa} Y_\alpha \in U \tag{212}$$

By Lemma 0.6.9, H has cardinality κ.

We claim that g is constant on $[H]^{n+1}$: Let $T \in [H]^{n+1}$. Write T as $S \cup \{\alpha\}$ with $\max S < \alpha$. Then

$$\begin{aligned} \alpha \in H \Rightarrow & \quad \alpha \in \Delta_{\gamma<\kappa} Y_\gamma \\ \Rightarrow & \quad \alpha \in \bigcap_{\beta<\alpha} Y_\beta \\ \Rightarrow & \quad \alpha \in Y_{\max S} \\ \Rightarrow & \quad \alpha \in Y_S \\ \Rightarrow & \quad g_S(\alpha) = \delta_S \end{aligned}$$

On the other hand, $S \subseteq H$ implies $S \subseteq Z$ and hence by definition of Z, $h(S) = \delta_S = \delta$, so $g(T) = g_S(\alpha) = \delta_S = \delta$.

$\square$

0.7 Constructibility

0.7.1 The Constructible Universe

A set X is **(first-order) definable** in a set Y (from parameters) if there exists a first-order formula $\varphi(x_0, x_1, \ldots, x_n)$ in the language of set theory (i.e. only using the binary relation symbol $\in$) such that for some $a_1, \ldots, a_n \in Y$,

$$X = \{y \in Y : (Y, \in) \models \varphi[y, a_1, \ldots, a_n]\}. \tag{213}$$

The constructible universe is built as a cumulative hierarchy of sets along the ordinals. In each successor step, instead of adding all subsets of the current set, only the **definable** ones are added. Formally, L is defined as follows. Given a set Y, let

$$\mathcal{P}_{\mathrm{Def}}(Y) = \{X \subseteq Y : X \text{ is definable in } Y \text{ from parameters}\}, \tag{214}$$

where the underlying language is the language of set theory. Now put

$$\begin{aligned} L_0 &= \emptyset \\ L_{\alpha+1} &= \mathcal{P}_{\mathrm{Def}}(L_\alpha) \\ L_\lambda &= \bigcup_{\alpha<\lambda} L_\alpha \quad (\lambda \text{ limit ordinal}) \end{aligned}$$

Finally, let

$$L = \bigcup_{\alpha \in \mathrm{Ord}} L_\alpha. \tag{215}$$

Basic properties of L

The first Proposition tells us that the L_ξ are set-theoretically nice structures and linearly ordered by the $\subseteq$-relation.

Proposition 0.7.1. *For each ordinal ξ:*

- ***(1)*** *L_ξ is transitive.*
- ***(2)*** *For $\alpha < \xi$, $L_\alpha \subseteq L_\xi$.*
- ***(3)*** *For $\alpha < \xi$, $L_\alpha \in L_\xi$.*
- ***(4)*** *If $\xi \geq \omega$, then $|L_{\xi+1}| = |L_\xi|$.*

Proof. We show the first two statements statements simultaneously by induction.

They are clear for $\xi = 0$ and ξ limit, so assume $\xi = \alpha + 1$.

Suppose $x \in L_\alpha$. Consider the formula $\varphi(x_0) \equiv x_0 \in x$ (here x is a parameter). φ defines the set

$$x' = \{a \in L_\alpha : L_\alpha \models \varphi[a]\} = \{a \in L_\alpha : a \in x\}. \tag{216}$$

By induction hypothesis, L_α is transitive, and hence $a \in x$ implies $a \in L_\alpha$, and hence $x' = x$, so $x \in L_{\alpha+1}$. This yields $L_\alpha \subseteq L_\xi$. Now if $x \in L_\xi$, then $x \subseteq L_\alpha$, and hence $x \subseteq L_\xi$. Thus L_ξ is transitive.

For the third statement, note that the formula $\varphi(x_0) \equiv x_0 = x_0$ defines L_α in L_α, and hence $L_\alpha \in L_{\alpha+1}$.

For (4), notice that there are only countably many formulas.

□

Next, we show that L contains all ordinals and that ξ "shows up" exactly after ξ steps.

Proposition 0.7.2. *For any ξ,*

- ***(1)*** *$\xi \subseteq L_\xi$,*
- ***(2)*** *$L_\xi \cap \mathrm{Ord} = \xi$.*

Proof. Clearly, (1) follows from (2).

To show (2), one again proceeds by induction. Again, the statement is clear for 0 and limit ordinals, so assume $\xi = \alpha + 1$ and $L_\alpha \cap \mathrm{Ord} = \alpha$.

We need to show $L_{\alpha+1} \cap \mathrm{Ord} = \alpha + 1 = \alpha \cup \{\alpha\}$. Since $L_\alpha \subseteq L_{\alpha+1}$, we have $\alpha \subseteq L_{\alpha+1} \cap \mathrm{Ord}$. On the other hand, since $L_{\alpha+1} \subseteq \mathcal{P}(L_\alpha)$, we have $L_{\alpha+1} \cap \mathrm{Ord} \subseteq \alpha + 1$. It thus remains to show that $\alpha \in L_{\alpha+1}$.

We observed before that the statement

> $\varphi_{\mathrm{Ord}}(x)$: *$x$ is transitive and linearly ordered by $\in$*

can be formalized by a Δ_0 formula, which are absolute for transitive classes.

Thus,

$$\alpha = \{a \in L_\alpha \colon L_\alpha \models \varphi_{\mathrm{Ord}}[a]\},$$

and hence we can conclude $\alpha \in L_{\alpha+1}$.

□

Defining definability

We want to study L as a model of ZF. In order to do that, we need a better understanding of $\mathcal{P}_{\mathrm{Def}}$. Our definition, so far, is “from the outside” (i.e. in the meta-theory). This will make it hard to understand how formulas behave relative to L, in particular, whether we can apply any of the absoluteness results we obtained. We will therefore have to develop (or at least, convince ourselves how we can develop) definability **inside** ZF. We can then combine this with some powerful results about **cumulative hierarchies** (such as V and L) to prove that L is a model not only of ZF, but also of CH and AC.

Coding set theoretic formulas We follow the standard procedure of **Gödelization** and assign every set theoretic formula φ a **Gödel number** $\ulcorner \varphi \urcorner$.

We fix the set of variables as $\{v_n : n \in \omega\}$ and put

$$\ulcorner v_n \urcorner = (1, n). \tag{217}$$

It will be helpful to extend the language by adding, for each set x, a formal constant, which we will denote by $\underline{x}$. We code this constant by

$$\ulcorner \underline{x} \urcorner = (2, x) \tag{218}$$

This allows us, for a given structure $(a, \in)$, to express statements about the elements of a by formulas using the corresponding constants $\underline{x}$ for $x \in a$. If $\underline{x} \in a$ is interpreted in $(a, \in)$ by itself, we call this the **canonical interpretation**.

The Gödelization of formulas now follows the usual, recursion pattern:

$$\begin{aligned}
\ulcorner x = y \urcorner &= (3, (\ulcorner x \urcorner, \ulcorner y \urcorner)) \\
\ulcorner x \in y \urcorner &= (4, (\ulcorner x \urcorner, \ulcorner y \urcorner)) \\
\ulcorner \neg\varphi \urcorner &= (5, \ulcorner \varphi \urcorner) \\
\ulcorner \varphi \vee \psi \urcorner &= (6, (\ulcorner \varphi \urcorner, \ulcorner \psi \urcorner)) \\
\ulcorner \exists v_n \varphi \urcorner &= (7, (n, \ulcorner \varphi \urcorner))
\end{aligned}$$

Definability of syntactic notions We can now express various syntactic statements about formulas, such as "v is a variable", as a set theoretic formula via the corresponding codes:

$$\mathrm{Vbl}(a) \;:\Leftrightarrow\; \exists y \in a \exists x \in y\, (a = (1, x) \wedge x = \omega) \tag{219}$$

where of course we substitute a suitable Δ_0 formula for "$x \in \omega$".

Proposition 0.7.3. *The following relations are Δ_1-definable over* ZF*.*

> $\mathrm{Fml}^n(e)$: *e is the code of a formula φ whose free variables are among $v_0, \ldots, v_{n-1}$;*
> $\mathrm{Fml}^n_a(e)$: *e is the code of a formula φ whose free variables are among $v_0, \ldots, v_{n-1}$, and whose constants are of the form $\underline{a_i}$ with $a_i \in a$*

Definability of the satisfaction relation Let a be a set. If we replace a formula φ by its code $\ulcorner\varphi\urcorner$, we have to express the fact that φ^a holds now as a statement about validity in the structure $(a, \in)$, using the code. That is, we have to **formalize** the notion of truth. This can be done using the recursive definition of truth. This way we obtain a Δ_1-definable predicate $\mathrm{Sat}(a, e)$ expressing

> $\mathrm{Sat}(a, e)$: e is the code of a formula $\varphi(\underline{a_0}, \ldots, \underline{a_n})$ that does not contain any free variables, and φ is true in $(a, \in)$ under the canonical interpretation.

In place of $\mathrm{Sat}(a, e)$, we will also write $(a, \in) \models e$. For any single formula, this formalization of truth then agrees with the validity of the corresponding relativization:

Theorem 0.7.1. *Let $\varphi(v_0, \ldots, v_{n-1})$ be a set theoretic formula, and assume $a_0, \ldots, a_{n-1} \in a$. Then it is provable in ZF that*

$$\varphi^a(a_0, \ldots, a_{n-1}) \;\leftrightarrow\; \mathrm{Sat}(a, \ulcorner\varphi(\underline{a_0}, \ldots, \underline{a_{n-1}})\urcorner). \tag{220}$$

(Keep in mind, however, that the general satisfaction relation, over all formulas, is not formalizable in ZF.)

The Theorem is proved via induction over the structure of φ. For atomic φ, both sides express the same fact, since we use the canonical interpretation of constants. The definition of relativization ensures that for the inductive cases, both sides behave identically wit respect to the corresponding subformulas.

Definability of definability We can now formalize the notion of definability we used informally above in the definition of L:

> $\mathcal{P}_{\mathrm{Def}}(a) = \{x \subseteq a \colon \exists e\, (\mathrm{Fml}^1_a(e) \;\wedge\; x = \{z \in a \colon (a, \in) \models e(z)\})\}$

Here, $e(z)$ is defined so that for a formula $\varphi(v_0)$ with code e,

> $e(z)$ is the code of the formula we obtain when we substitute $\underline{z}$ for v_0 in φ:

$$\ulcorner\varphi(v_0)\urcorner(z) = \ulcorner\varphi(\underline{z})\urcorner$$

With little effort, one can read off the complexity of the mapping $a \mapsto \mathcal{P}_{\mathrm{Def}}(a)$.

Theorem 0.7.2. *The relation $b = \mathcal{P}_{\mathrm{Def}}(a)$ is Δ_1-definable.*

Proof. (*Sketch*) Taking into account the complexity of the various sub-formulas, we see that the mapping $a \mapsto \mathcal{P}_{\mathrm{Def}}(a)$ is Σ_1-definable.

The graph of a Σ_1-definable function f (with domain V) is Δ_1, since the complement is given as

$$f(x) \neq y \;\Leftrightarrow\; \exists z(f(x) = z \,\wedge\, y \neq z). \tag{221}$$

Thus, $b = \mathcal{P}_{\mathrm{Def}}(a)$ is Δ_1.

□

This *complexity bound is of central importance* to the applications of L, and we will return to it soon (Proposition **??**).

Cumulative hierarchies

Many facts about L hold more generally for **cumulative hierarchies**.

Definition 0.7.1. *A sequence $(M_\alpha)_{\alpha\in\mathrm{Ord}}$ of sets is a **cumulative hierarchy** if*

- ***(H1)*** *each M_α is transitive,*
- ***(H2)*** *$\alpha < \beta$ implies $M_\alpha \subseteq M_\beta$,*
- ***(H3)*** *For limit λ,*

$$M_\lambda = \bigcup_{\alpha<\lambda} M_\alpha$$

The von-Neumann universe V ($M_\alpha = V_\alpha$) and Gödel's L ($M_\alpha = L_\alpha$) are the most important examples of cumulative hierarchies.

Definition 0.7.2. *A function $F : \mathrm{Ord} \to \mathrm{Ord}$ is **normal** if*

$$\alpha < \beta \;\Rightarrow\; F(\alpha) < F(\beta) \qquad \textit{(strictly increasing)}$$
$$\lambda \textit{ limit } \Rightarrow F(\lambda) = \bigcup_{\alpha<\lambda} F(\alpha) \qquad \textit{(continuous)}$$

The images of normal functions are called **clubs** (short for **closed, unbounded**).

> **Exercise**
> Show that a normal function has arbitrarily large fixed points, that is,
>
> $$\forall\alpha\exists\beta\;(\beta > \alpha \,\wedge\, \beta = F(\beta)) \tag{222}$$

Theorem 0.7.3 (Reflection for cumulative hierarchies)**.** *Let $(M_\alpha)_{\alpha\in\mathrm{Ord}}$ be a cumulative hierarchy and let $M = \bigcup_\alpha M_\alpha$.*

For every set theoretic formula $\varphi(v_0, \ldots, v_{n-1})$ there exists a normal function F such that

$$F(\alpha) = \alpha \;\;\to\;\; \forall a_0, \ldots, a_{n-1} \in M_\alpha\; (\varphi^{M_\alpha}(\vec{a}) \leftrightarrow \varphi^M(\vec{a})) \tag{223}$$

Proof. Proceed by induction on the formula structure. We focus on the case $\varphi \equiv \exists y\psi$. The other cases are straightforward due to the definition of relativization.

By induction hypothesis, there exists a normal function G such that

$$G(\alpha) = \alpha \;\;\to\;\; \forall\vec{a}, b \in M_\alpha\; (\psi^{M_\alpha}(\vec{a}, b) \leftrightarrow \psi^M(\vec{a}, b)) \tag{224}$$

We define a function H by

$$H(\alpha) = \text{ least } \beta > \alpha \text{ with } \forall\vec{a} \in M_\alpha\; (\exists b \in M\; \psi^M(\vec{a}, b) \;\to\; \exists b \in M_\beta\; \psi^M(\vec{a}, b)) \tag{225}$$

We use H to define, via transfinite recursion, another normal function F:

$$\begin{aligned} F(0) &= 0 \\ F(\alpha+1) &= H(F(\alpha)) \\ F(\lambda) &= \bigcup_{\alpha<\lambda} F(\alpha) \quad \text{for } \lambda \text{ limit} \end{aligned}$$

The composition $F^* = F \circ G$ is again normal, and its fixed points are simultaneously fixed points of F and G. It is now straightforward to check that F^* has the desired property.

□

Corollary 0.7.3.1 (Scott-Scarpellini Theorem)**.** *If $(M_\alpha)_{\alpha\in\mathrm{Ord}}$ is a cumulative hierarchy, $M = \bigcup_\alpha M_\alpha$, and $\varphi(\vec{x})$ is a set-theoretic formula, then*

$$\forall\alpha\exists\beta>\alpha\ \forall a_0,\ldots,a_{n-1} \in M_\beta\ (\varphi^{M_\beta}(\vec{a}) \leftrightarrow \varphi^M(\vec{a})) \tag{226}$$

By taking conjunctions, it is possible to generalize the reflection theorem to **finite sets of formulas**. Again, it is not possible (unless ZF is inconsistent) to extend this to arbitrary sets of formulas (or we could produce, in ZF, a *set model* of ZF, contradicting the *second incompleteness theorem*).

Corollary 0.7.3.2. ZF *is not finitely axiomatizable.*

Inner models

Definition 0.7.3. *A class M is an **inner model of** ZF if*

- *M is transitive,*
- *M contains all ordinals,*
- *σ^M holds for all axioms σ of* ZF.

Theorem 0.7.4 (Characterization of inner models)**.** *A class M is an inner model of* ZF *if and only if there exists a sequence $(M_\alpha)_{\alpha\in\mathrm{Ord}}$ such that for all $\alpha, \beta, \lambda \in$ Ord,*

- ***(I1)*** *$M = \bigcup_{\alpha\in\mathrm{Ord}} M_\alpha$ is a cumulative hierarchy,*
- ***(I2)*** *$\mathcal{P}_{\mathrm{Def}}(M_\alpha) \subseteq M_{\alpha+1} \subseteq \mathcal{P}(M_\alpha)$*

Proof. ($\Rightarrow$) Suppose M is an inner model of ZF. Let

$$M_\alpha = V_\alpha^M = V_\alpha \cap M. \tag{227}$$

This defines a cumulative hierarchy, so (I1) is satisfied. For (I2), first note that (Power Set)^M if and only if $\forall x \in M(\mathcal{P}(x) \cap M \in M)$. Now we can use the absoluteness of $\mathcal{P}_{\mathrm{Def}}$ (Theorem **??**) and the fact that M satisfies the axiom of *Separation* to conclude $\mathcal{P}_{\mathrm{Def}}(M_\alpha) \subseteq M_{\alpha+1}$.

($\Leftarrow$) *Extensionality* and *Foundation* hold in all transitive classes. *Set Existence* is satisfied in any cumulative hierarchy (since $\emptyset \in M$).

Union: By absoluteness, (Union)^M if and only if $\forall x \in M\ \bigcup x \in M$. The latter holds in M by (I2) and the fact that $y = \bigcup x$ is definable.

Pairing: Similar to *Union.*

Separation: Suppose $a, b_1, \ldots, b_n \in M$ and $\varphi(v_0, v_1, \ldots, v_n)$ is a formula. We have to argue that the set

$$z = \{x \in a\colon \varphi^M(x, b_1, \ldots, b_n)\}$$

is in M. By the reflection theorem for cumulative hierarchies, there exists α such that $a, b_1, \ldots, b_n \in M_\alpha$ and for all $x \in M_\alpha$,

$$\varphi^M(x, b_1, \ldots, b_n) \ \leftrightarrow\ \varphi^{M_\alpha}(x, b_1, \ldots, b_n).$$

This implies $z \in \mathcal{P}_{\mathrm{Def}}(M_\alpha)$ and hence by (I2), $z \in M_{\alpha+1}$. By (I1), $z \in M$.

Power Set: Suppose $a \in M$, say $a \in M_\alpha$. The set $z = \mathcal{P}(a) \cap M$ has a Δ_0-definition over M_α: the formula "$x \subseteq a$". Therefore, by (I2), $z \in M_{\alpha+1}$ and hence $z \in M$. z is the power set of a relative to M since, by absoluteness of $\subseteq$,

$$(z = \mathcal{P}(a))^M \iff \forall x \in M(x \in z \iff x \subseteq a) \iff z = \mathcal{P}(a) \cap M$$

Replacement: Assume a function F on M is defined by a formula $\varphi(x, y, \vec{a})$ ($\vec{a} \in M$ being parameters), that is

$$\forall x, y \in M\ (\varphi^M(x, y, \vec{a}) \wedge \varphi^M(x, z, \vec{a})\ \rightarrow\ y = z)$$

Let b be a set. By reflection, there exists an α such that $\vec{a}, b \in M_\alpha$ and the following two formulas hold:

$$\forall x, y, z \in M_\alpha\ (\varphi^M(x, y, \vec{a})) \leftrightarrow\ \varphi^{M_\alpha}(x, y, \vec{a}))$$
$$\forall x \in M_\alpha\ (\exists y \in M \varphi^M(x, y, \vec{a})\ \leftrightarrow\ \exists y \in M_\alpha \varphi^{M_\alpha}(x, y, \vec{a}))$$

Since $b \subseteq M_\alpha$ (transitivity), this implies

$$\forall x \in b\ (\exists y \in M \varphi^M(x, y, \vec{a})\ \leftrightarrow\ \exists y \in M_\alpha \varphi^M(x, y, \vec{a}))$$

and therefore

$$\{y : \exists x \in b\, \varphi^M(x, y, \vec{a})\} = \{y : \exists x \in b\, \varphi^{M_\alpha}(x, y, \vec{a})\}$$

The left side defines the image of F in M, which, by the right side, is in $\mathcal{P}_{\mathrm{Def}}(M_\alpha)$, and thus, by (I2), in $M_{\alpha+1}$.

Infinity: "$x = \omega$" is Δ_0, and since by (I2), $L_\omega \subseteq M_\omega$, we have that $\omega \in M_{\omega+1}$ and that this witnesses the axiom of *Infinity*.

□

We see that V and L lie at the extreme ends of the spectrum of inner models.

Corollary 0.7.4.1. *L is an inner model of* ZF*.*

0.7.2 The Axiom of Constructibility

We can add to ZF the axiom that all sets are constructible, i.e.

> $(\mathsf{V}=\mathsf{L})$ $\quad \forall x \exists y\,(y \text{ is an ordinal } \wedge\ x \in L_y)$.

This axiom is usually denoted by $\mathsf{V}=\mathsf{L}$. We may be tempted to think that L is then trivially a model of $\mathsf{ZF}+\mathsf{V}=\mathsf{L}$. But this is not at all clear, since this has to hold **relative to** L, i.e. $(\mathsf{V}=\mathsf{L})^L$.

This means that

$$\forall x \in L\, \exists y \in L\,(y \text{ is an ordinal } \wedge\ (x \in L_y)^L).$$

To verify this, we need to make sure that *inside* L, L "means the same as" L. This is, of course, an absoluteness property, and we therefore revisit the complexity of the formulas defining the constructible universe.

We have seen that the map $a \mapsto \mathcal{P}_{\mathrm{Def}}(a)$ is Σ_1. This important implications for the map $\alpha \mapsto L_\alpha$.

Proposition 0.7.4. *The map $\alpha \mapsto L_\alpha$ is Δ_1.*

Proof. We first show that the mapping is Σ_1. The mapping is obtained by ordinal recursion over the function $a \mapsto \mathcal{P}_{\mathrm{Def}}(a)$.

In general, if a function $G : \mathrm{V} \to \mathrm{V}$ is Σ_1 and $F : \mathrm{Ord} \to \mathrm{V}$ is obtained by recursion from G, i.e. $F(\alpha) = G(F\restriction_\alpha)$, then F is also Σ_1. This is because

$$\begin{aligned} y = F(\alpha)\ \leftrightarrow\ \alpha \in \mathrm{Ord} \wedge \exists f\,(f \text{ function } \wedge \mathrm{dom}(f) = \alpha \\ \wedge \forall \beta < \alpha (f(\beta) = G(f\restriction_\beta) \wedge y = G(f)). \end{aligned}$$

Applying some of the various prefix transformations for Σ_1-formulas, and using that being an ordinal, being an function, being the domain of a function, etc., are all Δ_0 properties, the above formula can be shown to be Σ_1, too.

In our case, G is a function that applies either $\mathcal{P}_{\mathrm{Def}}$ or $\bigcup$ (both at most Δ_1), depending on whether the input is a function defined on a successor ordinal or a limit ordinal (or applies the identity if neither is the case). Fortunately, this case distinction is also Δ_0, and hence we obtain that $G : \alpha \mapsto L_\alpha$ is Σ_1.

Finally, as in Theorem **??**, observe that if G is a Σ_1 function with a Δ_1 domain (Ord), then G is actually Δ_1, since we have

$$G(x) \neq y\ \Leftrightarrow\ \exists z (G(x) = z\ \wedge\ y \neq z) \tag{228}$$

so the complement of the graph of G is Σ_1-definable, too.

□

Corollary 0.7.4.2.
- *(1) The relations $x = L_\alpha$ and $x \in L_\alpha$ are Δ_1.*
- *(2) The predicate $x \in L$ is Σ_1.*
- *(3) The axiom $\mathsf{V}=\mathsf{L}$ is Π_2.*

We can relativize the definition of L to other classes M. If M is is an inner model, then the development of L can be done *relative to* M. Since M is a ZF-model, it has to contain all the sets L_α^M (as we developed definability and proved facts about it *inside* ZF). As M is transitive, the mapping $G : \alpha \to L_\alpha$ is absolute for M and we obtain, for all α,

$$L_\alpha^M = L_\alpha. \tag{229}$$

Theorem 0.7.5.
- *(1) If M is any transitive proper class model of ZF, then $L = L^M \subseteq M$.*
- *(2) L is a model of $\mathsf{ZF}+\mathsf{V}=\mathsf{L}$.*

Proof. (1) follows immediately from the fact that for such M, $L_\alpha^M = L_\alpha$.

(2) We have

$$\begin{aligned}(\mathsf{V}=\mathsf{L})^L &\leftrightarrow \forall x \in L \exists y \in L\,(y \text{ is an ordinal } \wedge\ x \in L_y)^L \\ &\leftrightarrow \forall x \in L \exists \alpha\,(x \in L_\alpha)^L && (\mathrm{Ord} \subset L \text{ and absolute}) \\ &\leftrightarrow \forall x \in L \exists \alpha\,(x \in L_\alpha) && (\text{by } (1))\end{aligned}$$

The last statement is true since $L = \bigcup_\alpha L_\alpha$.

□

Constructibility and the Axiom of Choice

Every well-ordering on a transitive set X can be extended to a well-ordering of $\mathcal{P}_{\mathrm{Def}}(X)$.

Note that every element of $\mathcal{P}_{\mathrm{Def}}(X)$ is determined by a pair $(\psi, \vec{a})$, where ψ is a set-theoretic formula, and $\vec{a} = (a_1, \dots, a_n) \in X^{<\omega}$ is a finite sequence of parameters.

For each $z \in \mathcal{P}_{\mathrm{Def}}(X)$ there may exist more than one such pair (i.e.\ z can have more than one definition), but by well-ordering the pairs $(\psi, \vec{a})$, we can assign each $z \in \mathcal{P}_{\mathrm{Def}}(X)$ its **least** definition, and subsequently order $\mathcal{P}_{\mathrm{Def}}(X)$ by comparing least definitions. Elements already in X will form an initial segment.

Such an order on the pairs $(\psi, \vec{a})$ can be obtained in a **definable way**: First use the order on X to order $X^{<\omega}$ length-lexicographically, order the formulas through their Gödel numbers, and finally put

$$(\psi, \vec{a}) < (\varphi, \vec{b}) \quad \text{iff} \quad \psi < \varphi \text{ or } (\psi < \varphi \text{ and } \vec{a} < \vec{b}). \tag{230}$$

Based on this, we can order all levels of L so that the following hold:

- **(1)** $<_L|_{V_\omega}$ is a standard well-ordering of V_ω (as for example given by a canonical isomorphism $(V_\omega, \in) \leftrightarrow (\mathbb{N}, E)$, see **?**)
- **(2)** $<_L|_{L_{\alpha+1}}$ is the order on $\mathcal{P}_{\mathrm{Def}}(L_\alpha)$ induced by $<_L |L_\alpha$.
- **(3)** $<_L|_{L_\lambda} = \bigcup_{\alpha<\lambda} <_L|_{L_\alpha}$ for a limit ordinal $\lambda > \omega$.

It is straightforward to verify that this is indeed a well-ordering on L. Moreover, the relation $<_L$ is Δ_1. (To verify this, we have to spell out all the details of the above definition. This is a little involved, so we skip this here and refer to **?**.)

Theorem 0.7.6. $\mathsf{V}=\mathsf{L}$ *implies* AC

Since L is a model of $\mathsf{ZF}+\mathsf{V}=\mathsf{L}$, we obtain

Corollary 0.7.6.1. *If* ZF *is consistent, then* $\mathsf{ZF}+\mathsf{AC}(=\mathsf{ZFC})$ *is consistent, too.*

Condensation and the Continuum Hypothesis

We now show that $\mathsf{V}=\mathsf{L}$ implies the Continuum Hypothesis. The proof works by showing that under $\mathsf{V}=\mathsf{L}$, every subset of a cardinal κ will be constructed by stage κ^+. This is made possible by a "**condensation**" argument: If any subset x of κ is in L, then it must show up at some stage L_λ. κ and x generate an elementary substructure M of L_λ or cardinality κ. If we could show that this M **itself must be an** L_β, we can use the fact that the cardinality of the L_α behaves "tamely" along the ordinals, as evidenced by the following.

Proposition 0.7.5. *For all* $\alpha \geq \omega$, $|L_\alpha| = |\alpha|$.

Proof. We know that $\alpha \subseteq L_\alpha$. Hence $|\alpha| \leq |L_\alpha|$. To show $|\alpha| \geq |L_\alpha|$, we work by induction on α.

If $\alpha = \beta + 1$, then by Proposition **??**(4), $|L_\alpha| = |L_\beta| = |\beta| \leq |\alpha|$.

If α is limit, then L_α is a union of $|\alpha|$ many sets of cardinality $\leq |\alpha|$ (by inductive hypothesis), hence of cardinality $\leq |\alpha|$.

□

But why would an elementary substructure of an L_λ have to be itself an L_β? This is where the absoluteness of the construction of L strikes yet again!

Lemma 0.7.7 (Condensation lemma)**.** *There is a finite set T of axioms of ZF – Power Set so that if M is a transitive set with $M \models T + \mathsf{V} = \mathsf{L}$, then $M = L_\lambda$ for some limit ordinal λ.*

Proof. Let the axioms of T be *Pairing*, *Union*, *Set Existence*, together with all (instances of) axioms of ZF used to prove that all the theorems leading up to the fact that for all α, L_α exists and that $\alpha \mapsto L_\alpha$ is Δ_1 (and hence absolute). (We have proved only finitely meany theorems so far so we only needed finitely many axioms!)

Suppose for a transitive set M, $M \models T + \mathsf{V} = \mathsf{L}$. Let λ be the least ordinal not in M. We must have that $\mathrm{Ord}^M = \lambda$, by absoluteness of ordinal. Moreover, λ must be a limit ordinal since for each $\alpha \in M$, $\alpha \cup \{\alpha\}$ is in M since M satisfies *Pairing* and *Union.*

We have that

$$M \models \forall x \exists \alpha \in \mathrm{Ord}(x \in L_\alpha), \tag{231}$$

thus

$$\forall x \in M \exists \alpha < \lambda (x \in L_\alpha^M). \tag{232}$$

By absoluteness of $\alpha \mapsto L_\alpha$, we have $L_\alpha^M = L_\alpha$ and therefore

$$M \subseteq \bigcup_{\alpha \in M} L_\alpha = \bigcup_{\alpha < \lambda} L_\alpha = L_\lambda. \tag{233}$$

On the other hand, for each $\alpha < \lambda$, L_α^M exists in M (since T is strong enough to prove this), and by absoluteness

$$L_\lambda = \bigcup_{\alpha < \lambda} L_\alpha = \bigcup_{\alpha \in M} L_\alpha^M \subseteq M. \tag{234}$$

□

We now put condensation to use as described above.

Lemma 0.7.8. *Suppose $V = L$. If κ is a cardinal and $x \subseteq \kappa$, then $x \in L_{\kappa^+}$.*

Proof. Since we assume $\mathsf{V} = \mathsf{L}$, there exists limit $\lambda > \kappa$ such that $x \in L_\lambda$ and such that $L_\lambda \models T + \mathsf{V} = \mathsf{L}$, where T is as in the **condensation lemma**. Such a λ exists by the **reflection theorem** (Theorem 0.7.3). Let $X = \kappa \cup \{x\}$. By choice of λ, $X \subseteq L_\lambda$.

By the Löwenheim-Skolem Theorem, there exists an **elementary substructure** $N \preceq L_\lambda$ such that

$$X \subseteq N \subseteq L_\lambda \quad \text{and} \quad |N| = |X|. \tag{$*$}$$

N is not necessarily transitive, but since it is well-founded we can take its **Mostowski collapse** (Theorem **??**) and obtain a **transitive** set M together with an **isomorphism** $\pi : (N, \in) \to (M, \in)$.

Since κ is contained in both M and N, and is already transitive, it is straightforward to show via induction that $\pi(\alpha) = \alpha$ for all $\alpha \in \kappa$. Since $x \subseteq \kappa$, this also yields $\pi(x) = x$. This implies in turn that $x \in M$.

As $(M, \in)$ is isomorphic to $(N, \in)$ and $N \preceq L_\lambda$, M satisfies the same sentences as $(L_\lambda, \in)$. In particular, $M \models T + \mathsf{V} = \mathsf{L}$. By the **condensation lemma**, $M = L_\beta$ for some β.

This implies, by Proposition **??**,

$$|\beta| = |L_\beta| = |M| = |N| = |X| = \kappa < \kappa^+ \leq \lambda. \tag{235}$$

Since $x \in L_\beta$ and $\beta < \kappa^+$, it follows that $x \in L_{\kappa^+}$, as desired.

□

Theorem 0.7.9 (Gödel)**.** *If* $\mathsf{V} = \mathsf{L}$*, then for all cardinals* κ*,* $2^\kappa = \kappa^+$*.*

Proof. If $\mathsf{V} = \mathsf{L}$, then by Lemma **??**, $\mathcal{P}(\kappa) \subseteq L_{\kappa^+}$. With Proposition **??**, we obtain

$$2^\kappa = |\mathcal{P}(\kappa)| \leq |L_{\kappa^+}| = \kappa^+. \tag{236}$$

□

Corollary 0.7.9.1. *If* ZF *is consistent, so is* $\mathsf{ZF} + \mathsf{AC} + \mathsf{GCH}$.

0.7.3 Constructible Reals

In this lecture we transfer the results about L to the projective hierarchy. The main idea is to relate sets of reals that are defined by set theoretic formulas to sets defined in second order arithmetic.

The set of constructible reals

What is the complexity of the set $\mathbb{N}^{\mathbb{N}} \cap L$? In particular, is it in the projective hierarchy? The set of all constructible reals is defined by a Σ_1 formula over set theory:

$$\varphi(x_0) \;\equiv\; \exists y\,[y \text{ is an ordinal } \;\wedge\; x_0 \in L_y \;\wedge\; x_0 \text{ is a set of natural numbers }]. \tag{237}$$

We would like to replace this formula by an "equivalent" one in the language of second order arithmetic. In particular, we would like to replace the quantifier $\exists y$ by a quantifier over the reals.

The key for doing this is Lemma **??**: every constructible real shows up at a countable stage of L. Hence if $\alpha \in L \cap \mathbb{N}^{\mathbb{N}}$, there exists a countable ξ such that $x \in L_\xi$. Since $|\xi| = |L_\xi|$, L_ξ is countable, too. Hence we can hope to replace L_ξ by something like "*there exists a real that codes a model that looks like L_ξ*".

The **condensation lemma** (Lemma 0.7.7) allows us to do this. Let $\varphi_{\mathsf{V=L}}$ be the conjunction of the axioms in T and the axiom $\mathsf{V} = \mathsf{L}$.

Recall that any real $\alpha \in \mathbb{N}^{\mathbb{N}}$ codes a set theoretic structure

$$(\omega, E_\alpha) \qquad \text{where } E_\alpha = \{\langle m,n\rangle \colon \alpha(\langle m,n\rangle) = 0\}.$$

Unfortunately, the condensation lemma only holds for **transitive sets** (and (ω, E_α) may look very different from a transitive set model), so simply requiring $(\omega, E_\beta) \models \varphi_{\mathsf{V=L}}$ is not enough. But we know from Theorem **??** (**Mostowski collapse**) that if E_β is well-founded and extensional, we can map it isomorphically to a transitive set S on which we interpret E_β as $\in$. By the condensation lemma, this S must then be an L_ξ.

So, for reals, we can formulate membership in L now as follows:

$$\begin{aligned}\alpha \in L \cap \mathbb{N}^{\mathbb{N}} \iff \exists \beta \exists m\, [&E_\beta \text{ is extensional and well-founded} \\ &\wedge\; (\omega, E_\beta) \models \varphi_{\mathsf{V=L}} \;\wedge\; \pi_\beta(m) = \alpha],\end{aligned} \tag{$*$}$$

where π_β is the Isomorphism of the Mostowski collapse of E_β.

It remains to show that the notions occurring inside the square brackets are definable in second order arithmetic.

Proposition 0.7.6.
- ***(a)*** *For any $n \in \mathbb{N}$, the following set is Σ^0_n:*

$$\{(m,\sigma,\gamma) \in \mathbb{N} \times \mathbb{N}^{<\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \colon m = \ulcorner\varphi\urcorner \;\wedge\; \varphi \textit{ is } \Sigma_n \;\wedge\; (\omega, E_\gamma) \models \varphi[\sigma]\}$$

- ***(b)*** *If $\alpha \in \mathbb{N}^{\mathbb{N}}$ and E_α is well-founded and extensional, then the following set is arithmetic in α:*

$$\{(m,\gamma) \in \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \colon \pi_\alpha(m) = \gamma\}$$

Proof. (a) can be established similar to showing that Sat-predicate of Theorem 0.7.1 is Δ_1-definable. One does this first for Σ_1 formulas and then uses induction. Using Gödelization, one carefully defines all syntactical notions using arithmetic formulas. Then, one uses the recursive definition of truth to establish the definability of the satisfaction relation.

Since we work with relations over $\mathbb{N}$ now instead of arbitrary sets, it is not that easy anymore to keep quantifiers bounded. But since we are only interested in the complexity of $\models$ for Σ_n-formulas, this helps us bound the overall complexity at Σ^0_n

(c) By analyzing the recursive definition and using the definition of $\mathbb{N}$ in ZF, one first shows that the set

$$\{(m,p) \in \mathbb{N} \times \mathbb{N} \colon \pi_\alpha(m) = p\} \tag{238}$$

is arithmetic in α.

Let $\psi(v_0, v_1, v_2)$ be the formula $\langle v_0, v_1 \rangle \in v_2$. Then

$$\pi_\alpha(m) = \gamma \iff \\ \forall p, q\ (\gamma(p) = q \leftrightarrow \exists i, j\ (\pi_\alpha(i) = p \wedge \pi_\alpha(j) = q \wedge (\omega, E_\alpha) \models \psi[i, j, m]))$$

Now apply the previous observation and (a).

□

Finally, note that

$$E_\beta \text{ is extensional} \iff \forall m, n\, [\forall k(kE_\beta m \leftrightarrow kE_\beta n) \rightarrow m = n]. \tag{239}$$

Hence this is arithmetical. And we have already seen that coding a well-founded relation over $\mathbb{N}$ is Π^1_1.

Now we know the complexity of all parts of $(*)$ and can put everything together.

Theorem 0.7.10. *The set $L \cap \mathbb{N}^\mathbb{N}$ is Σ^1_2.*

In a similar way we can show that the relation $\alpha <_L \beta$ over $(L \cap \mathbb{N}^\mathbb{N})^2$ is Σ^1_2 (using that $<_L$ is Δ_1-definable).

Exercise

Recall that given $\alpha \in \mathbb{N}^\mathbb{N}, n \in \mathbb{N}$, $(\alpha)_n$ denotes the n-th column of α.

Show that the following relation R over $(L \cap \mathbb{N}^\mathbb{N})^2$ is Σ^1_2.

$$(\alpha, \beta) \in R :\iff \{(\alpha)_n \colon n \in \mathbb{N}\} = \{\gamma \colon \gamma <_L \beta\} \tag{240}$$

In other words, α codes the (countable) $<_L$-initial segment restricted to $\mathbb{N}^\mathbb{N}$.

If $\mathsf{V} = \mathsf{L}$, then the set is actually Δ^1_2, since then

$$\alpha <_L \beta \iff \alpha \neq \beta \wedge \neg(\beta <_L \alpha). \tag{241}$$

Finally, since $\mathsf{V} = \mathsf{L}$ implies CH, we can use Proposition 0.5.7 to show the existence of non-measurable sets under $\mathsf{V} = \mathsf{L}$.

Corollary 0.7.10.1. *If $\mathsf{V} = \mathsf{L}$, then there exists a Δ^1_2 set that is not Lebesgue-measurable and does not have the Baire property.*

An uncountable Π^1_1 set without a perfect subset

We now show that under the assumption $\mathsf{V} = \mathsf{L}$, the **perfect set property** fails at level Π^1_1.

We start with constructing an example at the Σ^1_2 level.

Recall that if $\alpha \in \mathbb{N}^\mathbb{N}$ codes a well-ordering on $\mathbb{N}$, then

$$\|\alpha\| = \text{ order type of well-ordering coded by } \alpha.$$

Proposition 0.7.7. *If $\mathsf{V} = \mathsf{L}$, there exists an uncountable Σ^1_2 set in $\mathbb{N}^\mathbb{N}$ without a perfect subset.*

Proof. Let $A \subseteq \mathbb{N}^{\mathbb{N}}$ be given by

$$x \in A \iff x \in \text{WOrd} \wedge \forall y <_L x\, (\| y \| \neq \| x \|).$$

In other words, A collects the $<_L$-least code for every ordinal $< \omega_1$.

Clearly A is uncountable, since it has a representative for every countable ordinal and hence of cardinality ω_1.

Moreover, A is Σ^1_2: Let R be the Σ^1_2-relation of the exercise above. Then

$$x \in A \iff x \in \text{WOrd} \wedge \exists z\, (R(z,x) \wedge \forall n\, (\| (z)_n \| \neq \| x \|).$$

The relation $\| (z)_n \| \neq \| x \|$ Π^1_1, hence A is Σ^1_2.

Finally, we see that A **does not have an uncountable $\boldsymbol{\Sigma}^1_1$ subset** (hence, since all perfect sets are closed, no perfect subset): By $\boldsymbol{\Sigma}^1_1$-boundedness (Theorem 0.5.12), for any $\boldsymbol{\Sigma}^1_1$ subset $X \subseteq A$ the set $\{\| x \| : x \in X\}$ bounded by an ordinal $\gamma < \omega_1$, hence countable.

□

It is possible to get this example down to Π^1_1 using the powerful technique of **uniformization**.

Definition 0.7.4. *Suppose $A \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}}$. We say $A^* \subseteq A$* ***uniformizes*** *A if*

$$\forall x\, [\exists y\, A(x,y) \to \exists! y\, A^*(x,y)] \tag{242}$$

A pointclass Γ has the ***uniformization property*** *if*

$$A \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \wedge A \in \Gamma \quad \Rightarrow \quad \exists A^* \in \Gamma\, (A^* \textit{ uniformizes } A).$$

Theorem 0.7.11 (Kondo)**.** *$\boldsymbol{\Pi}^1_1$ has the uniformization property.*

Theorem 0.7.12. *If $\mathsf{V} = \mathsf{L}$, then there exists an uncountable $\boldsymbol{\Pi}^1_1$ set without a perfect subset.*

Proof. Let A be the Σ^1_2 set from the proof of Proposition 0.7.7. $A \subseteq \mathbb{N}^{\mathbb{N}}$ is the projection of a Π^1_1 set $B \subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}}$. If we apply uniformization to B, we obtain a uniformizing set B^* whose projection is still A.

B^* is uncountable, but does not contain a perfect subset: If $P \subset B^*$ were such a subset, then P would be (the graph of) a function and uncountable, and the projection $\exists^{\mathbb{N}^{\mathbb{N}}} P$ would be an uncountable $\boldsymbol{\Sigma}^1_1$ subset of A, contradiction.

□

0.7.4 Shoenfield Absoluteness

Tree representations of $\mathbf{\Sigma}^1_2$ sets

Analytic sets are projections of closed sets. Closed sets are in $\mathbb{N}^\mathbb{N}$ are infinite paths through trees on $\mathbb{N}$.

We call a set $A \subseteq \mathbb{N}^\mathbb{N}$ **Y-Souslin** if A is the projection $\exists^{Y^\mathbb{N}}[T]$ of some $[T]$, where T is a tree on $\mathbb{N} \times Y$, that is

$$A = \exists^{Y^\mathbb{N}}[T] = \{\alpha \colon \exists y \in Y^\mathbb{N}\, (\alpha, y) \in [T]\}. \tag{243}$$

Theorem 0.7.13 (Shoenfield, 1961)**.** *Every $\mathbf{\Sigma}^1_2$ set is ω_1-Souslin. In particular, if A is Σ^1_2 then there is a tree $T \in L$ on $\mathbb{N} \times \omega_1$ such that $A = \exists^{(\omega_1)^\mathbb{N}}[T]$.*

Proof. Assume first A is Π^1_1. There is a recursive tree T on $\mathbb{N} \times \mathbb{N}$ (and hence, in L, since "*being recursive*" is definable) such that

$$\alpha \in A \iff T(\alpha) \text{ is well-founded.} \tag{244}$$

Hence, $\alpha \in A$ if and only if there exists an order preserving map $\pi : T(\alpha) \to \omega_1$. We recast this now in terms of getting an infinite branch through a tree.

Let $\{\sigma_i \colon i \in \mathbb{N}\}$ be a recursive enumeration of $\mathbb{N}^{<\mathbb{N}}$. We may assume for this enumeration that $|\sigma_i| \le i$. We define a tree $\widetilde{T}$ on $\mathbb{N} \times \omega_1$ by

$$\widetilde{T} = \{(\sigma, \tau) :\ \forall i, j < |\sigma|\, [\sigma_i \supset \sigma_j \ \wedge\ (\sigma \restriction_{|\sigma_i|}, \sigma_i) \in T\ \to\ \tau(i) < \tau(j)]\}. \tag{245}$$

The tree $\widetilde{T}$ is in L, since it is definable from T and ω_1. Furthermore, if $\alpha \in A$, then the existence of an order-preserving map $\pi : T(\alpha) \to \omega_1$ implies that there is an infinite path (α, η) through $\widetilde{T}$.

Conversely, if such a path (α, η) exists, then there is an order preserving map $\pi : T(\alpha) \to \omega_1$. Hence we have

$$\alpha \in A\ \leftrightarrow\ \exists \eta \in (\omega_1)^\mathbb{N}\ (\alpha, \eta) \in [\widetilde{T}]\ \leftrightarrow\ \alpha \in \exists^{(\omega_1)^\mathbb{N}}[\widetilde{T}], \tag{246}$$

so A is of the desired form.

Next, we extend the representation to Σ^1_2.

If A is Σ^1_2, then there is a Π^1_1 set $B \subseteq \mathbb{N}^\mathbb{N} \times \mathbb{N}^\mathbb{N}$ such that $A = \exists^{\mathbb{N}^\mathbb{N}} B$. Since $B \in \Pi^1_1$, we can employ the tree representation of Π^1_1 to obtain a tree T over $\mathbb{N} \times \mathbb{N} \times \omega_1$ such that $B = \exists^{(\omega_1)^\mathbb{N}}[T]$.

Now we recast T as a tree T' over $\mathbb{N} \times \omega_1$ such that $\exists^{(\omega_1)^\mathbb{N}}[T'] = \exists^{(\omega_1)^\mathbb{N}} B$. This is done by using a bijection between $\mathbb{N} \times \omega_1$ and ω_1.

This way we can cast the $\mathbb{N} \times \omega_1$ component of T into a single ω_1 component, and thus transform the tree T into a tree T' over $\mathbb{N} \times \omega_1$ such that $\exists^{(\omega_1)^\mathbb{N}}[T'] = \exists^{(\omega_1)^\mathbb{N}}[B]$.

□

$\mathbf{\Sigma}^1_2$ sets as unions of Borel sets

We can use Shoenfield's tree representation to extend Corollary **??** to Σ^1_2 sets.

Theorem 0.7.14 (Sierpinski, 1925)**.** *Every Σ^1_2 set is a union of $\aleph_1$-many Borel sets.*

Sierpinski's original proof used AC. The following proof does not make use of choice.

Proof. Let $A \subseteq \mathbb{N}^\mathbb{N}$ be Σ^1_2. By Theorem 0.7.13 there exists a tree T on $\mathbb{N} \times \omega_1$ such that $A = \exists^{(\omega_1)^\mathbb{N}}[T]$. For any $\xi < \omega_1$ let

$$T^\xi = \{(\sigma, \eta) \in T \colon \forall i \le |\eta|\ \eta(i) < \xi\}. \tag{247}$$

Since the cofinality of ω_1 is greater than ω (this can be proved without using AC), every $d : \omega \to \omega_1$ has its range included in some $\xi < \omega_1$. Thus we have

$$A = \bigcup_{\xi<\omega_1} \exists^{(\omega_1)^{\mathbb{N}}}[T^\xi]. \tag{248}$$

For all $\xi < \omega_1$, the set $\exists^{(\omega_1)^{\mathbb{N}}}[T^\xi]$ is Σ^1_1, because the tree T^ξ is a tree on a product of countable sets and hence is isomorphic to a tree on $\mathbb{N} \times \mathbb{N}$. By Corollary 0.5.11.2, each Σ^1_1 set is the union of $\aleph_1$ many Borel sets, from which the result follows.

□

As for co-analytic sets, an immediate consequence of this theorem is (using the perfect set property of Borel sets):

Corollary 0.7.14.1. *Every $\mathbf{\Sigma}^1_2$ set has cardinality at most $\aleph_1$ or has a perfect subset and hence cardinality $2^{\aleph_0}$.*

Absoluteness of Σ^1_2 relations

Shoenfield used the tree representation of $\mathbf{\Sigma}^1_2$ sets to establish an important absoluteness result for Σ^1_2 sets of reals.

Suppose $A \subseteq \mathbb{N}^{\mathbb{N}}$ is Σ^1_2. Then, by the Kleene Normal Form there exists a bounded formula of second order arithmetic $\varphi(v_0, v_1, v_2)$ such that

$$\alpha \in A \iff \exists\beta_0\, \forall\beta_1\, \exists m\, \varphi(\alpha\mid_m, \beta_0\mid_m, \beta_1\mid_m). \tag{249}$$

Let M be an inner model of ZF. Arithmetical formulas can be interpreted in ZF and we can also relativize them. This allows us to introduce a relativized version of A by identifying, as usual, a set with the predicate that defines it:

$$A^M(\alpha) :\iff (\exists\beta_0 \in M \cap \mathbb{N}^{\mathbb{N}})\, (\forall\beta_1 \in M \cap \mathbb{N}^{\mathbb{N}})\, (\exists m)\, \varphi(\alpha\mid_m, \beta_0\mid_m, \beta_1\mid_m) \tag{250}$$

Note that we do not have to relativize the inner natural number quantifier, since $\mathbb{N}$ is absolute for inner models, and also not the formula φ, since a bounded arithmetic formula translates into a bounded set-theoretic formula (with only natural number quantifiers) and is therefore absolute for M.

We can then say that A is **absolute for** M if for any $\alpha \in M$,

$$A^M(\alpha) \iff A(\alpha). \tag{251}$$

Absoluteness can be extended and relativized in a straightforward manner to predicates analytical in some $\gamma \in \mathbb{N}^{\mathbb{N}} \cap M$.

All arithmetic predicates are absolute, since all quantifiers are natural number quantifiers. Shoenfield absoluteness extends this absoluteness to Σ^1_2 and Π^1_2 predicates.

Theorem 0.7.15 (Shoenfield absoluteness)**.** *Every $\Sigma^1_2(\gamma)$ predicate and every $\Pi^1_2(\gamma)$ predicate is absolute for all inner models M of ZFC such that $\gamma \in M$. In particular, all Σ^1_2 and Π^1_2 relations are absolute for L.*

Proof. We show the theorem for Σ^1_2 predicates. For the relativized version, one uses the **relative constructible universe** $L[\gamma]$, see **?** or **?**.

Let A be a Σ^1_2 relation. For simplicity, we assume that A is unary. Fix a tree representation of A as a projection of a Π^1_1 set. So, let T be a recursive tree on $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ such that

$$A(\alpha) \iff \exists\beta\, T(\alpha, \beta) \text{ is well-founded.} \tag{252}$$

Note that T is in M (since it is recursive and hence definable).

Now assume $\alpha \in M$ and $A^M(\alpha)$. Hence there is a $\beta \in M$ such that $T(\alpha, \beta)$ is well-founded in M. This is equivalent to the fact that in M there exists an order preserving mapping $\pi : T(\alpha, \beta) \to \mathbf{Ord}$.

Since M is an inner model and T is absolute, the mapping exists also in V. Hence $T(\alpha, \beta)$ is well-founded in V and thus $A(\alpha)$.

For the converse assume that $\alpha \in M$ and $A(\alpha)$. Now we use the alternative tree representation of A given by Theorem 0.7.13. Let $U \in L \subseteq M$ be a tree on $\mathbb{N} \times \omega_1$ such that $A = \exists^{(\omega_1)^{\mathbb{N}}} U$.

As before, let

$$U(\alpha) = \{(\alpha\,|_n, \tau) \in U \colon n \in \mathbb{N}, \tau \in (\omega_1)^n, \} \tag{253}$$

Then, for any $\alpha \in \mathbb{N}^{\mathbb{N}}$,

$$\begin{aligned} A(\alpha) &\iff \exists \lambda \in (\omega_1)^{\mathbb{N}}\ (\alpha, \lambda) \text{ infinite path through } U. \\ &\iff U(\alpha) \text{ not well-founded.} \end{aligned}$$

This means that there exists no order preserving map $U(\alpha) \to \omega_1$. But then such a map cannot exist in M either. Thus, $U(\alpha)$ is a tree in M which is ill-founded in the sense of M. Thus, by Shoenfield's Representation Theorem relativized to M, $A^M(\alpha)$.

Absoluteness for Π^1_2 follows by employing the same reasoning, using that the complement is Σ^1_2.

$\square$

By analyzing the proof one sees that it actually suffices that M is a transitive $\in$-model of a certain finite collection of axioms ZF such that $\omega_1 \subseteq M$.

The result is the best possible with respect to the analytical hierarchy, since the statement

$$\exists \alpha\ [\alpha \notin L] \tag{254}$$

is Σ^1_3, but cannot be absolute for $M = L$.

Shoenfield's absoluteness theorem also holds for sentences rather than predicates, with a similar proof. This means a Σ^1_2 statement is true in L if and only if it holds in V. Many results of classical analysis are Σ^1_2 statements. The Shoenfield absoluteness theorem says that if they can be established under $\mathsf{V} = \mathsf{L}$, they can be established in ZF alone.

On the negative side, as we will soon see, Shoenfield absoluteness also puts strong limits on the use of forcing to establish independence results in analysis.

0.8 Reference

0.8.1 Bibliography

Bibliography

W. Ackermann. Die Widerspruchsfreiheit der allgemeinen Mengenlehre. *Mathematische Annalen*, 114 (1):305–315, 1937.

G. Cantor. Beiträge zur Begründung der transfiniten Mengenlehre. *Mathematische Annalen*, 46(4): 481–512, 1895.

M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1999.

P. R. Halmos. *Measure Theory*. Van Nostrand, 1950.

T. Jech. *Set Theory*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003.

A. Kanamori. *The Higher Infinite*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, second edition, 2003.

A. S. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.

A. S. Kechris, V. G. Pestov, and S. Todorcevic. Fraı
ssé Limits, Ramsey Theory, and topological dynamics of automorphism groups. *Geometric And Functional Analysis*, 15(1):106–189, 2005.

D. Macpherson. A survey of homogeneous structures. *Discrete Mathematics*, 311(15):1599–1634, 2011.

H. L. Royden and P. Fitzpatrick. *Real analysis*, volume 32. Macmillan New York, 1988.

R. Solovay. A model of set-theory in which every set of reals is Lebesgue measurable. *The Annals of Mathematics*, 1970.

P. Urysohn. Sur un espace métrique universel. I, II. *Bull. Sci. Math.*, 51:1–38, 1927.