Lecture 10: The Structure of Borel Sets

In this lecture 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

$$(k_1, \dots, k_m, \alpha_1, \dots, \alpha_n) \mapsto (k_1, \dots, k_m, \theta_n(\alpha_1, \dots, \alpha_n)).$$

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

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

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 9.9 tells us that a set $A \subseteq \mathbb{N}^\mathbb{N}$ is $\mathbf{\Sigma}^0_n$ 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* $\varphi(x_1, \dots, x_n, \alpha, \underline{\gamma})$ (i.e. all quantifiers are bounded) such that

$$A(\alpha) \quad \Longleftrightarrow \quad \exists x_1 \dots \mathsf{Q} x_n \, \varphi(x_1, \dots, x_n, \alpha, \gamma) \text{ holds (in the standard model).}$$

Here γ is the parameter, and Q is "∃" if n is odd, and "∀" if n is even.

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

$$A(\alpha) \quad \Leftrightarrow \quad \forall x_1 \ldots \mathsf{Q} x_n\, \varphi(x_1, \ldots, x_n, \alpha, \gamma) \text{ holds (in the standard model).}$$

where $\varphi(x_1, \ldots, x_n, \alpha, \underline{\gamma})$ 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 $\mathbf{\Sigma}^0_1$-code for a set A and its complement, we can decide the relation $A(\alpha\restriction_n)$ recursively in the code.

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

$$A(\alpha) \quad \Leftrightarrow \quad \exists x_1 \ldots \mathsf{Q} x_n\, R(x_1, \ldots, x_n, \alpha),$$

and similarly for $\mathbf{\Pi}^0_n$ 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

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

The dual operation is

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

Proposition 10.1: *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 $\varphi(x_1, \ldots, x_n, z, \alpha)$ be a bounded formula such that

$$A(z,\alpha) \quad \Longleftrightarrow \quad \exists x_1 \ldots \mathsf{Q} x_n \; \varphi(x_1, \ldots, x_n, z, \alpha) \text{ holds.}$$

Then

$$\exists^{\mathbb{N}} A(\alpha) \quad \Longleftrightarrow \quad \exists x_0 \exists x_1 \ldots \mathsf{Q} x_n \; \varphi(x_1, \ldots, x_n, x_0, \alpha)$$

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

$$\exists^{\mathbb{N}} A(\alpha) \quad \Longleftrightarrow \quad \exists z_1 \ldots \mathsf{Q} z_n \; \varphi((z_1)_1, \ldots, z_n, (z_1)_0, \alpha),$$

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 10.2:

(a) For all $n \geq 1$, $\mathbf{\Sigma}^0_n$ is closed under countable unions and finite intersections.

(b) For all $n \geq 1$, $\mathbf{\Pi}^0_n$ is closed under finite unions and countable intersections.

(c) For all $n \geq 1$, $\mathbf{\Delta}^0_n$ 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.

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

□

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

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

and the dual is

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

Proposition 10.3: *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 coding function $\pi : \mathbb{N}^{<\mathbb{N}} \to \mathbb{N}$. We can define a *partial inverse*

$$(k)_i = \begin{cases} \sigma_i & \text{if } k = \pi(\sigma) \text{ for a finite sequence } \sigma = (\sigma_0, \dots, \sigma_{n-1}) \\ & \quad \text{and } i < n \\ 0 & \text{otherwise.} \end{cases}$$

Using this decoding, we have the following equivalence, which immediately imply the closure properties for $\mathbf{\Sigma}^0_n$ and $\mathbf{\Pi}^0_n$, respectively, and hence also for $\mathbf{\Delta}^0_n$.

$$\begin{aligned} \forall m \leq n \, \exists k \, P(m,k) \quad &\Longleftrightarrow \quad \exists k \, \forall m \leq n \, P(m,(k)_m) \\ \exists m \leq n \, \forall k \, P(m,k) \quad &\Longleftrightarrow \quad \forall k \, \exists m \leq n \, P(m,(k)_m) \end{aligned}$$

□

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

Proposition 10.4: *For all $n \geq 1$, 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 4.7 one can represent a continuous function through a monotone mapping ψ from finite strings to finite strings. We have

$$f^{-1}(A) = \{\alpha : A(f(\alpha))\}.$$

Let R be clopen such that

$$A(\alpha) \quad \Longleftrightarrow \quad \exists x_1 \dots \mathsf{Q} x_n \, R(x_1, \dots, x_n, \alpha).$$

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

$$A = \{x : (x, y) \in U\}.$$

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*,

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

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

We have seen in the previous lecture that there is a strong connection between r.e. sets and Σ^0_1 sets. The relation is based on the fact that each Σ^0_1 set in $\mathbb{N}^{\mathbb{N}}$ has a *code* that is r.e. We can use the code to obtain a universal set for $\mathbf{\Sigma}^0_1$.

Proposition 10.6: *For any $n \geq 1$, 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 $n \geq 1$, the set of all $\mathbf{\Sigma}^0_1$-codes or $\mathbf{\Pi}^0_1$-codes is homeomorphic to $\mathbb{N}^{\mathbb{N}}$. This follows easily from the definition of the Borel codes. Hence, if we fix n, every $\gamma \in \mathbb{N}$ 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 $\gamma \in \mathbb{N}$.

Now we let, for fixed n,

$$U = \{(\alpha, \gamma) : \ \alpha \text{ is in the } \Sigma^0_n \ (\mathbf{\Pi}^0_n) \text{ set coded by } \gamma\}.$$

It follows easily from Theorem 9.9 that U is $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$), too, and it is clear from the definition of U that it parametrizes $\mathbf{\Sigma}^0_n$ ($\mathbf{\Pi}^0_n$). $\square$

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

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

The definition of codes for higher levels is then similar to Definition 9.1.

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 10.7: *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

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

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

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

and thus

$$\beta \in D \quad \Leftrightarrow \quad (\beta, \beta) \in U \quad \Leftrightarrow \quad \beta \in \neg D,$$

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 10.8: *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 $\mathbf{\Sigma}^0_n \subsetneq \mathbf{\Pi}^0_n$ and $\mathbf{\Pi}^0_n \subsetneq \mathbf{\Sigma}^0_n$, $\mathbf{\Delta}^0_n \subsetneq \mathbf{\Sigma}^0_n, \mathbf{\Pi}^0_n$. 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 10.7. □

Borel sets of transfinite order

We saw that the *Borel sets of finite order*

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

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

Proposition 10.9: *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

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

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 $m \geq 1$,

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

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. $\square$

We can extend the Borel hierarchy to arbitrary ordinals.

Definition 10.10: Let X be a Polish space. Given an ordinal ξ, we define

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

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

Proposition 10.11: *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

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

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

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

Then

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

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 10.7 extends to the transfinite levels. For the finite levels, this followed from the existence of universal sets for each level.

Proposition 10.12: *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

$$\neg U = \{(\alpha,\gamma) : (\alpha,\gamma) \notin U\}$$

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

$$B = \{\beta : (\beta,\gamma) \in U\}$$

and hence

$$A = \{\alpha : (\alpha,\gamma) \notin U\}.$$

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 $\xi = \sup\{\xi_n + 1 : n < \omega\}$. Define

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

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.) $\square$

This general proof of existence of universal sets does not use Borel codes, since those were defined only for Borel sets of finite order. The proof of Proposition 10.12 provides an idea how we could extend the definition of a code to transfinite orders: Take unions of codes along a cofinal sequence. However, we would like to this in an effective way, and it is not clear how to do this for infinite ordinals in general.

We will later return to this question, when we introduce *computable ordinals*.