# Combinatorial and topological games

*Baire spaces and applications in classical mathematics*

José Iovino

Department of Mathematics
The University of Texas at San Antonio

For the UT San Antonio Mathematics student research seminar
and the joint University of Toronto–UT San Antonio Mathematics research group

© 2026 José Iovino. All rights reserved.
These notes are a work in progress. Please do not post or redistribute copies;
link instead to the current version at `utsa-quantum-math.github.io`.
Comments are welcome at `jose.iovino@utsa.edu`.

Contents

Introduction 2

**Part I. Topological games and Baire spaces** 7
1. Baire category 8
2. Infinite games 10
3. The Banach–Mazur game 12
4. The strong Choquet game 17
5. The converse of Choquet's theorem 22
6. Summary: from complete metrizability to Baire category 28
7. Notes and historical remarks 29

**Part II. Applications of Banach–Mazur and Choquet games** 31
8. Applications to Diophantine approximation 33
9. Applications to homogeneous dynamics 38
10. Applications to topological groups 42
11. Applications to infinite-dimensional Ramsey theory 45
12. Applications to Banach space theory 49
13. Applications to ultrafilters 53

**Part III. Applications of other topological games** 60
14. Applications to set-theoretic topology 61
15. Applications to descriptive set theory 63

**Part IV. Applications of non-topological games** 68
16. Applications to the probabilistic method 68
17. Applications to fixed point theory 71

**Part V. Ehrenfeucht–Fraïssé games and Fraïssé limits** 74
18. Fraïssé limits 75
19. Elementary equivalence and the game 96
20. The Ehrenfeucht–Fraïssé theorem 98
21. Linear orders, parity and connectivity 102
22. Back-and-forth systems and ultrapowers 107
23. Extensions to other logics 109
24. Notes and historical remarks 111

**Part VI. Banach–Mazur games in computational complexity** 113
25. Category inside complexity classes 114
26. Strings, machines and complexity classes 114
27. The game 118
28. The main result 120
29. Solving the problem with the game 120
Exercises 125

**Part VII. The Scottish Book, von Neumann, Ulam, Texas, Toronto, and the triumph of games, logic, topology and combinatorics** 126
30. The marble tables of the Scottish Café 126

31. Problem 43 128
32. Von Neumann's invitation and the bridge to America 129
33. War, and the survival of the book 129
34. Ulam after Lwów 130
35. The 1957 Los Alamos mimeograph 130
36. Problem 153: a live goose 130
37. The road to Denton: Dan Mauldin and North Texas 131
38. The New Scottish Book 131
39. Set-theoretic topology and Ramsey spaces at Toronto 131
40. The triumph of games: combinatorics, logic and the real-valued setting 132
41. Epilogue: A Scottish Book of our own 133
Further reading 133
Appendix A. Real-valued structures 133
References 135

## Introduction

In 1935 Stanisław Mazur entered a problem in the Scottish Book (see Figure 1). The book was a notebook kept at the Scottish Café, in the beautiful city of Lwów (see Figure 2). At the café's tables the mathematicians of Lwów did a good part of their thinking. The problem concerns a game in which two players take turns choosing intervals of the real line, each inside the previous one. Mazur asked when the first player can force the intersection of the intervals to meet a given set, and when the second player can keep the intersection disjoint from it. Banach found the answer but never published a proof. The first published proof is Oxtoby's, in a paper of 1957 whose title calls the game by the name it has kept: the Banach–Mazur game.

A game is, at first sight, a curious way of asking a question about a set. Mazur's game turned out to be a second name for a notion that topologists already had. In Part I we show that the second player can force the intersection of the moves to lie inside a set exactly when the set is large in the sense of Baire category (see Theorem 3.5). Category and the game are thus one notion under two modes of presentation: the first given by countable unions of nowhere dense sets, the second by what a player can enforce. The second presentation has proved the more fertile, and these notes are an attempt to show why.

The reason begins with the shape of mathematical statements. Most statements worth proving alternate quantifiers: for every $\varepsilon$ there is a $\delta$, for every point there is a neighborhood, for every finite configuration there is a larger one with some property. A statement of this shape is a contest between two people, a skeptic who chooses the instances of "for every" and a defender who answers each of them with an instance of "there is". The statement is true exactly when the defender has a winning strategy, a rule that meets every challenge in time. Logicians have made this observation into a semantics, in the dialogue games of Paul Lorenzen and in the game-theoretical semantics of Jaakko Hintikka. Every student of analysis applies the observation, usually without saying so, when trying to defeat an imaginary opponent who chooses $\varepsilon$. A statement with finitely many quantifiers is a game of finitely many moves. Mazur's game goes further: the players alternate forever, and the property at stake, whether the intersection of the moves lies in a given set, is decided only at the end of an infinite play. Seen in this way, games are not a device that mathematics borrowed from recreation. Games

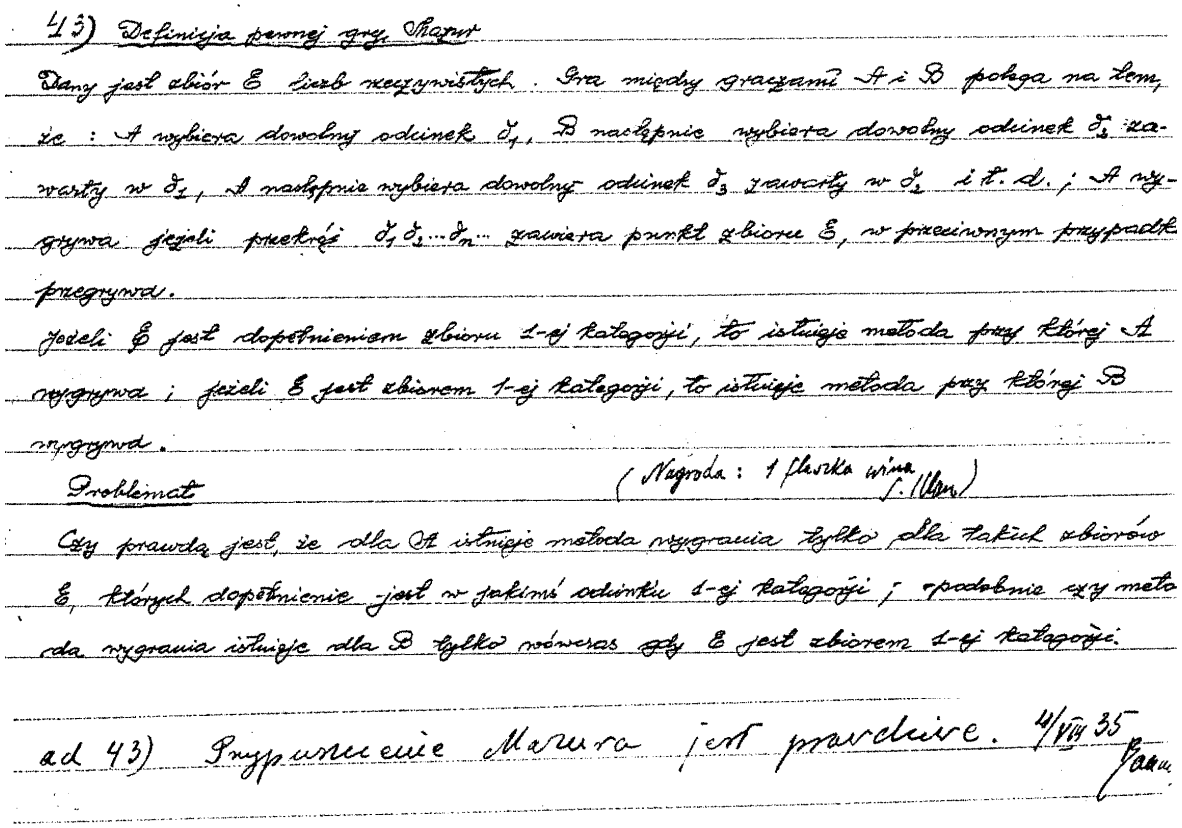



FIGURE 1. Mazur's Problem 43 in the Scottish Book, with the prize he offered and his signature, and below it, from the opposite page, Banach's signed note of 4 August 1935 that Mazur's conjecture is true (courtesy of the Banach family, via the Stefan Banach Portal, Adam Mickiewicz University, Poznań, `kielich.amu.edu.pl/Stefan_Banach`).

are the form in which mathematical assertions already come, made explicit and allowed to run without end.

Two features of games recur behind all the variations of the rules, and between them they account for much of what follows. The first is that a winning strategy is built one move at a time: a global question, whether a set is large, becomes a local one, what to play next. The second is that strategies can be *combined*: a player can pursue several goals at once by taking turns between them, one round for one goal and the next round for another. This second feature explains why the sets for which Player II has a winning strategy survive countable intersections. Comeager sets, whose complements are countable unions of nowhere dense sets, survive such intersections too, as do sets of full measure. Largeness in the sense of dimension, by contrast, does not survive countable intersections: two sets of full Hausdorff dimension can be disjoint. The badly approximable numbers, the real numbers $x$ for which there is $c > 0$ with $|x - p/q| \geq c/q^2$ for all integers $p$ and $q \geq 1$, form one such set, and their complement another. For each value of the shrinking factor that Schmidt's game prescribes to the second player, the winning sets of the game have both virtues: they have full dimension, and they are closed under countable intersections. We show both facts in Section 8, where we also tell the history of the problem.

Once a notion has become a game, the rules of the game can be changed, and each change asks a new question. Over the last sixty years the sizes of the moves have been prescribed, the players have been allowed to play vectors, points or open covers instead of intervals, and the second player has been limited in how much she may compute. Each change has found its problem in some unexpected corner of mathematics. In number theory, the game of Wolfgang Schmidt measures how badly a real number can be approximated by rationals; played on spaces of lattices, the same game describes the bounded orbits of flows that Dani, and later Kleinbock and Margulis, studied. In analysis, a game decides when a separately continuous function of two variables is jointly continuous, another game describes the Asplund spaces of functional analysis, and Gowers's dichotomy for Banach spaces, which settled a problem of Banach, rests on a game played with sequences of vectors. In infinite combinatorics, the Kastanas game characterizes the sets that are Ramsey in the sense of the Ellentuck topology. In topology, the Menger game turns a covering property into a contest, and Menger's conjecture becomes the assertion that the game is determined, an assertion that Fremlin and Miller refuted.

A game raises a question that a theorem does not raise: must one of the players have a winning strategy? For games of finite length the answer is yes, by an argument that goes back to Zermelo's study of chess in 1913. For infinite games the question is deep, and the answer depends on the set that decides the winner. Martin proved in 1975 that a game is determined whenever that set is Borel, and the theorem organizes the descriptive set theory of Part III, beginning with the Wadge hierarchy, which ranks the Borel sets by a game. At the other extreme, a nonprincipal ultrafilter on the set $\omega$ of natural numbers defines a game in which neither player has a winning strategy (see Theorem 13.1). Such ultrafilters exist by the axiom of choice, so the assertion that every game of this kind is determined contradicts the axiom. Mycielski and Steinhaus, starting from a variant of Mazur's game that Ulam had entered in the Scottish Book, proposed in 1962 to take determinacy as an axiom instead of choice. Through such games, the questions asked over coffee in Lwów reached the axioms of set theory.

Games also prove impossibilities, and they replace luck by method. In the Erdős–Selfridge theorem of Part IV, the expected outcome of random play becomes a potential function that guides a deterministic strategy, an early instance of what computer scientists call derandomization. In the same part, the fact that the game of Hex cannot end in a draw turns out, as Gale showed in 1979, to be equivalent to the Brouwer fixed point theorem. In Part VI, the second player of Mazur's game may perform only a limited amount of computation. The result, due to Mehlhorn and Lutz, is a notion of Baire category inside complexity classes, in which the languages that a computer decides in polynomial time form a small class among those that it decides in exponential time, although ordinary category finds the two classes equally small.

Logic, finally, has a game of its own, found in Warsaw, the other great center of Polish mathematics between the wars, and in that game the circle closes. If statements are games, two structures that satisfy the same statements should be indistinguishable by play. The Ehrenfeucht–Fraïssé game of Part V makes this precise: for structures for a finite relational vocabulary, the second player survives $n$ rounds exactly when the two structures satisfy the same sentences with at most $n$ nested quantifiers. With this game we prove that no first-order sentence expresses the parity of a finite linear order, or the connectivity of a finite graph. Allowed to run forever, the game becomes a method of construction. The method yields the rational order, the random graph, the algebraic closure of a finite field and the field of

algebraic numbers, each as the unique countable structure that contains every piece of a given kind and in which every isomorphism between pieces extends to a symmetry of the whole.

Games are played by people, and these notes keep the people in view. Each story begins with the problem that a game solved and with the mathematicians who posed and solved the problem. The last part, Part VII, follows the Scottish Book itself: how the notebook survived the war, how its problems traveled to Los Alamos, Denton and Toronto, and how the games that it started came back to settle a problem of Banach.

The theory of the Banach–Mazur game and of its strengthening, the strong Choquet game, occupies Part I, together with what the two games reveal about Baire spaces and complete metrizability. That part is self-contained and can be read as a short course on its own. Once Part I has been read, the order is the reader's to choose: most sections of Parts II–IV stand on their own, as does Part VI, and at the opening of Part II we indicate which of its sections lean on which. These sections share roughly the same plan (the problem, its history, the game, the main result, and how the game solves the problem), so a reader who knows one of them knows how to read the others. The part on logic is built differently, as a chapter of a textbook, with the history gathered at the end; like Part IV, that part draws on Part I only for the language of games of Section 2. Its section on Fraïssé limits needs neither games nor logic, and is meant as an entrance to logic for readers who come from algebra or analysis.

Since every game in Part II is the original one with a few rules changed, we record the original here, in the form in which we will keep modifying it. Mazur's question was about a set $E$ of real numbers at which the first player aims. We write $S = \mathbb{R} \setminus E$ for the set at which the second player aims, and we give the payoff to her (see the conventions below). In Part I we prove the theorem in the box below for an arbitrary topological space in place of the real line (see Theorem 3.5).

**The Banach–Mazur game.** Fix a set $S \subseteq \mathbb{R}$. Player I chooses a closed interval $U_0$ of positive length; Player II chooses a closed interval $V_0 \subseteq U_0$ of positive length; Player I chooses $U_1 \subseteq V_0$; and so on, forever. Player II wins if every point of $\bigcap_{k=0}^{\infty} V_k$ belongs to $S$. Theorem (Banach; Oxtoby, 1957): Player II has a winning strategy if and only if the complement $\mathbb{R} \setminus S$ is meager (see Definition 1.1), that is, if and only if the set $S$ is "large" in the sense of Baire category.

*Conventions.* We follow the usage of Part I throughout: Player I moves first, and, in games of Banach–Mazur type, Player II answers inside Player I's move. Under the same usage, the payoff set belongs to Player II, whom we call "she". In most of our games she is the one trying to produce an object with a given property, and Player I, whom we call "he", is the one trying to stop her. In the Maker–Breaker games of Section 16 the roles are reversed: Maker, who builds, is Player I. Where an author used other names for the players ($\beta$ and $\alpha$, Black and White, ONE and TWO, Maker and Breaker, Spoiler and Duplicator), we say which of them is our Player I. Among the exercises at the end of the sections, the more demanding ones are marked $(*)$.



(a) (b)

FIGURE 2. The Scottish Café: (a) Plac Akademicki in Lwów in 1925, with the café on the ground floor of the turreted corner building on the right, today Prospekt Shevchenka 27, photographed by Marek Münz (National Library of Poland, Polona, public domain); and (b) the interior of the café in the 1930s, by an unknown photographer (public domain).

Part

# I

# Topological games and Baire spaces

A set of real numbers may be large in more than one sense: in measure, in dimension, or in category. Of these, the third is the one most naturally described by a game. A set of the first category is a countable union of pieces that are thin everywhere. A player who answers each open set with a smaller one can step clear of one such piece at every move. By dodging the pieces one after another, the player keeps the intersection of the moves outside the whole union. That only sets of the first category can be dodged in this way is the deeper half of the theorem of Banach and Oxtoby, which we prove in Section 3. In this part we turn the idea of dodging into a theory of *topological games*: infinite games in which two players alternately choose open sets (in one of them Player I also chooses points). In each of these games a topological property of the underlying space is characterized by which player has a winning strategy.

**Roadmap**

Two theorems carry this part, and each of them turns a property of a space into a fact about a game. The first is that a non-empty space is a Baire space exactly when the first player has no winning strategy in the Banach–Mazur game (see Theorem 3.3). That the space is Baire and that Player I is powerless are one fact, stated once in the language of topology and once in the language of play. The second is that a non-empty metrizable space is completely metrizable exactly when the second player has a winning strategy in the strong Choquet game (see Theorem 4.4).

In Section 1 we recall category and Baire spaces, and in Section 2 we say what a game is. The first theorem is proved in Section 3. In Section 4 the moves of Player I acquire points; there we prove the easy half of the second theorem and derive from it the Baire category theorem (see Corollary 4.5). In the same section we extend the easy half to Čech-complete spaces (see Subsection 4.1). The hard half, which we prove in Section 5, is that a winning strategy for Player II produces a complete metric that induces the topology of the space. The proof uses three classical tools that are worth knowing for their own sake: Alexandrov's theorem on countable intersections of open sets, A. H. Stone's theorem on locally finite refinements, and Kőnig's lemma. We close

with a single chain of implications that gathers all these results (see Section 6), and with notes on their history (see Section 7). A reader who knows Baire category may begin with Section 2. The later parts draw on this one unequally: the applications in Part II rest on the Banach–Mazur and strong Choquet games of Sections 3 and 4, and Part VI rests on the Banach–Mazur game alone. By contrast, Parts III–V need little more than the language of Section 2. If the part has a moral, the moral is that a strategy can stand in for a construction: the complete metric of Section 5 is read off from what Player II knows how to do.

**Prerequisites.** We assume a first course in metric spaces and the basic vocabulary of point-set topology: open and closed sets, closure and interior, dense sets, continuity, and Cauchy sequences and completeness. We also use the following notions:

(1) A topological space $X$ is *metrizable* if there is a metric on $X$ whose open sets are exactly the open sets of $X$; such a metric is called *compatible*. The space is *completely metrizable* if the space has a compatible metric that is complete. Completeness is a property of the metric, whereas complete metrizability is a property of the topology. The set of irrational numbers, for example, is not complete in the usual metric of $\mathbb{R}$, yet the set carries a complete compatible metric (see Section 5).
(2) A *$G_\delta$ set* is a countable intersection of open sets.
(3) Every metric space $(X, d)$ embeds isometrically as a dense subspace of a complete metric space, the *completion* of $(X, d)$, which is unique up to an isometry that fixes $X$. For instance, the completion of $\mathbb{Q}$ is $\mathbb{R}$.
(4) *Zorn's lemma*: if every chain in a non-empty partially ordered set has an upper bound, then the set has a maximal element. We use the lemma to find maximal families of pairwise disjoint open sets.

## 1. Baire category

Baire category measures the size of a set by topology alone: a set is small when the set is a countable union of sets that are thin everywhere. In this section we recall the notions of category. We then show that a space is a Baire space exactly when no non-empty open subset of the space is small in this sense, that is, of the first category (see Theorem 1.3 below). For a subset $E$ of a topological space we denote the closure of $E$ by $\overline{E}$, and the interior of $E$ by $E^\circ$.

**Definition 1.1.** Let $X$ be a topological space, and let $E \subseteq X$ be given.

(1) The set $E$ is *nowhere dense* in $X$ if $(\overline{E})^\circ = \emptyset$.
(2) The set $E$ is of the *first category* in $X$ (or *meager*) if $E$ is the union of a countable family of nowhere dense sets. A set that is not of the first category is of the *second category* (or *non-meager*).
(3) The set $E$ is *comeager* (or *residual*) in $X$ if the complement $X \setminus E$ is of the first category.
(4) The space $X$ is called a *Baire space* if the intersection of every countable family of dense open subsets of $X$ is dense in $X$.
(5) The set $E$ has the *Baire property* if there is an open set $O$ such that the symmetric difference $E \triangle O$ is of the first category; that is, if $E$ is an open set up to an error of the first category.

**Remark 1.2.** The notions in Definition 1.1 classify sets by their topological "size"; they form a qualitative analogue of measure theory.

(1) A nowhere dense set is thin everywhere. Equivalently (see Exercise 1 at the end of this section), a set $E$ is nowhere dense exactly when every non-empty open set contains a non-empty open set disjoint from $E$. Every finite subset of $\mathbb{R}$ is nowhere dense, and so is the Cantor ternary set (see Figure 3).

(2) A set of the first category is topologically negligible, and a countable union of such sets is again of the first category, just as a countable union of Lebesgue-null sets is null. The rationals $\mathbb{Q} = \bigcup_{q\in\mathbb{Q}}\{q\}$ form a set of the first category in $\mathbb{R}$. The analogy has limits: the line $\mathbb{R}$ is the union of a set of the first category and a set of Lebesgue measure zero [Oxt80, Chapter 1]. Hence a set that is small in one sense may be large in the other.

(3) A comeager set contains "most" points of the space. A property is said to hold of the *generic* point of a topological space $X$ if the set of points that have it is comeager. In $\mathbb{R}$, the set of irrational numbers is comeager, so the generic real number is irrational.

(4) In Theorem 1.3 below we show that a space is a Baire space exactly when no non-empty open subset of the space is negligible. This *local* condition is stronger than asking that the space itself be of the second category in itself. Consider the disjoint union $X = (\mathbb{Q}\times\{0\}) \cup (\mathbb{R}\times\{1\}) \subseteq \mathbb{R}^2$ (see Figure 4). The open subset $\mathbb{R}\times\{1\}$ of $X$, a copy of $\mathbb{R}$, is of the second category in itself by the Baire category theorem (see Corollary 4.5). If a subset of an open set $U \subseteq X$ is nowhere dense in $X$, then the subset is also nowhere dense in $U$. Consequently, the set $\mathbb{R}\times\{1\}$ is of the second category in $X$, and therefore so is $X$. The space $X$, however, is not a Baire space, since the open subset $\mathbb{Q}\times\{0\}$ of $X$ is of the first category.

(5) Every open set $O$ has the Baire property, because $O \triangle O = \emptyset$. Every closed set $F$ has the Baire property as well: the set $F \triangle F^\circ = F \setminus F^\circ$ is closed and has empty interior, hence nowhere dense. The sets with the Baire property form a $\sigma$-algebra (see [Oxt80]), and therefore every Borel set has the Baire property.




FIGURE 3. The first stages of the Cantor ternary set $C = \bigcap_{n=0}^{\infty} C_n$, where $C_0 = [0,1]$ and $C_{n+1}$ is obtained from $C_n$ by removing the open middle third of each of its intervals.




FIGURE 4. The space $X = (\mathbb{Q}\times\{0\}) \cup (\mathbb{R}\times\{1\})$ of Remark 1.2(4), whose two lines are open subsets of $X$.

**Theorem 1.3.** *A topological space $X$ is a Baire space if and only if no non-empty open subset of $X$ is of the first category.*

*Proof.* The idea is to pass to complements: a closed set is nowhere dense exactly when the complement of the set is a dense open set.

Suppose first that $X$ is a Baire space, and let $G \subseteq X$ be open and non-empty. If $G = \bigcup_{n=1}^{\infty} E_n$ with each $E_n$ nowhere dense, then each set $V_n = X \setminus \overline{E_n}$ is open and dense in $X$. Since $X$ is a Baire space, the intersection $\bigcap_{n=1}^{\infty} V_n$ is dense in $X$, and the intersection therefore meets the non-empty open set $G$. However, we have

$$G \cap \bigcap_{n=1}^{\infty} V_n = G \setminus \bigcup_{n=1}^{\infty} \overline{E_n} = \emptyset,$$

a contradiction, which shows that $G$ is of the second category.

Conversely, suppose that no non-empty open subset of $X$ is of the first category. Let $(V_n)_{n\geq 1}$ be a sequence of dense open subsets of $X$, and let $O \subseteq X$ be open and non-empty; we show that $O$ meets $\bigcap_{n=1}^{\infty} V_n$. For each $n$, the set $E_n = X \setminus V_n$ is closed, and $E_n$ has empty interior because $V_n$ is dense; thus each $E_n$ is nowhere dense. If $O$ missed $\bigcap_{n=1}^{\infty} V_n$, the set $O$ would be contained in $\bigcup_{n=1}^{\infty} E_n$, and would therefore be a non-empty open set of the first category, contrary to our hypothesis. Hence the open set $O$ meets $\bigcap_{n=1}^{\infty} V_n$. Since $O$ was arbitrary, the intersection $\bigcap_{n=1}^{\infty} V_n$ is dense, which proves that $X$ is a Baire space. ☺

**Exercises.**

(1) Let $X$ be a topological space, and let $E \subseteq X$ be given. Show that $E$ is nowhere dense in the sense of Definition 1.1(1) if and only if every non-empty open subset of $X$ contains a non-empty open set disjoint from $E$ (see Remark 1.2(1)). Show also that if $E$ is nowhere dense, then so is the closure of $E$.

(2) Show that the union of two nowhere dense subsets of a topological space is nowhere dense. Show by an example in $\mathbb{R}$ that the union of countably many nowhere dense sets need not be nowhere dense.

(3) Let $X$ be a topological space. Show that $X$ is a Baire space if and only if every comeager subset of $X$ is dense in $X$. Deduce that in a non-empty Baire space no comeager set is of the first category.

(4) Let $X$ be a Baire space, and let $U \subseteq X$ be a non-empty open set. Show that $U$, with the subspace topology, is a Baire space. (Hint: show that a subset of $U$ that is nowhere dense in $U$ is nowhere dense in $X$, and use Theorem 1.3.)

(5) (∗) Let $X$ be a Baire space, and let $D \subseteq X$ be a dense $G_\delta$ set. Show that $D$, with the subspace topology, is a Baire space. (Hint: a dense open subset of $D$ has the form $W \cap D$ with $W$ open and dense in $X$.)

## 2. Infinite games

In this section we fix the formal framework for all the games of these notes, from the Banach–Mazur game to the Ehrenfeucht–Fraïssé game. The framework consists of trees of positions, plays, payoff sets, and strategies.

A definition of a game must settle three matters. The first is the state of play: what is known when a move is made, namely the finite sequence of moves so far, and which moves are allowed at that point. The second is the outcome: an infinite play is won by Player II exactly when the play belongs to a given set. The third is a method of play: a rule that

assigns a move to every position at which a player is to move. Trees of positions settle the first matter, payoff sets of plays the second, and strategies the third.

**Definition 2.1.** Let $A$ be a non-empty set. As usual, we write $\omega = \{0, 1, 2, \dots\}$ for the set of natural numbers. We denote by $A^{<\omega}$ the collection of all finite sequences $s = (s_0, \dots, s_{n-1})$ of elements of $A$, with $n \in \omega$, and by $A^\omega$ the set of all infinite sequences $(a_n)_{n\in\omega}$ of elements of $A$. The length $n$ of $s$ is denoted by $|s|$; the empty sequence $\langle\rangle$ has length 0. If $s, t \in A^{<\omega}$ satisfy $|t| \le |s|$ and $t_i = s_i$ for $0 \le i < |t|$, we write $t \subseteq s$ and say that $s$ *extends* $t$. When sequences are regarded as functions defined on initial segments of $\omega$, this relation is literally inclusion. For $s \in A^{<\omega}$ and $a \in A$, we write $s^\frown a = (s_0, \dots, s_{|s|-1}, a)$ for the extension of $s$ by $a$.

A non-empty subset $T \subseteq A^{<\omega}$ is called a *tree* if $t \in T$ whenever $s \in T$ and $t \subseteq s$. For each $s \in T$, we write

$$\mathrm{Succ}(s) = \{a \in A : s^\frown a \in T\}.$$

We assume throughout that $\mathrm{Succ}(s) \neq \emptyset$ for every $s \in T$. An infinite sequence $x = (a_n)_{n=0}^\infty \in A^\omega$ is a *play* in $T$ if $(a_0, \dots, a_{n-1}) \in T$ for all $n \ge 1$. The set of all such plays is denoted by $[T]$.

**Definition 2.2.** An *infinite game* is a pair $(T, W)$, where $T \subseteq A^{<\omega}$ is a tree on a non-empty set $A$ and $W \subseteq [T]$ is a set of plays, the *payoff set*. Two players, I and II, alternately choose elements of $A$: Player I chooses $a_{2n}$, and Player II chooses $a_{2n+1}$, for $n \in \omega$, subject to the rule that $(a_0, \dots, a_k) \in T$ for every $k \in \omega$. Player II *wins* the resulting play $x = (a_n)_{n\in\omega} \in [T]$ if $x \in W$; otherwise, Player I wins. The elements of $T$ are called *positions*.

A *strategy* for Player I is a function $\sigma$, defined on the positions of even length, such that $\sigma(s) \in \mathrm{Succ}(s)$ for every such position $s$. A play or position is *consistent* with $\sigma$ if $a_{2n} = \sigma(a_0, \dots, a_{2n-1})$ for every $n \in \omega$ for which $a_{2n}$ is defined. We say that $\sigma$ is a *winning strategy* for Player I if every play consistent with $\sigma$ lies in $[T] \setminus W$.

A strategy $\tau$ for Player II and the notion of a winning strategy for Player II are defined analogously.



**Remarks 2.3.**

(1) In these notes, the payoff set $W$ is the set of plays won by Player II. Many texts, among them Kechris's book [Kec95], use the opposite convention, in which the payoff set is Player I's. Ours suits the games below, in which Player II is the player who tries to build something. The object to be built is a point common to all the moves, or an intersection that lies inside a given set. As in the Introduction, we call Player I "he" and Player II "she".

(2) The two players need not choose from the same kind of object. If Player I chooses from a set $A_\mathrm{I}$ and Player II from a set $A_\mathrm{II}$, we take $A = A_\mathrm{I} \cup A_\mathrm{II}$ and let the tree $T$ record which moves are legal.

(3) At most one player has a winning strategy. Indeed, if $\sigma$ and $\tau$ were winning strategies for Players I and II, respectively, the unique play consistent with both would lie both outside and inside $W$. That this play exists and is unique is the content of Exercise 2 at the end of this section.

(4) A game is called *determined* if one of the two players has a winning strategy. Not every game is determined: there are Baire spaces $X$ in which neither player has a winning strategy in the Banach–Mazur game $\mathrm{BM}(X)$ of Section 3 (see the remark after Theorem 3.3). The axiom of choice, moreover, produces undetermined games in which the players choose finite binary strings (see Section 13). On the other side stands Martin's theorem (see Section 15): every game in which the players choose

natural numbers and whose payoff set is a Borel subset of the space $\omega^\omega$ is determined. Here the space $\omega^\omega$ is the product of countably many copies of the discrete space $\omega$.

(5) A strategy sees the whole position, not only the last move. The move that a strategy $\sigma$ for Player I prescribes at a position $(a_0, \dots, a_{2n-1})$, with $n \in \omega$, may depend on everything played so far (see Figure 5). This memory can be indispensable (see Exercise 5 at the end of this section).

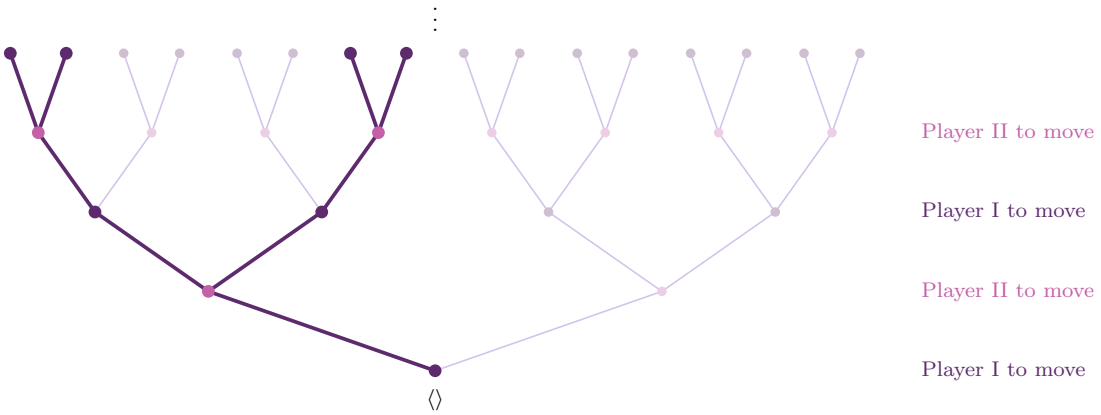



FIGURE 5. A game tree with two moves available at each position and the root $\langle\rangle$ at the bottom, with the positions consistent with a strategy $\sigma$ for Player I drawn dark.

**Exercises.**

(1) Let $A = \{0, 1\}$ and $T = A^{<\omega}$, and write each play as $(a_n)_{n\in\omega}$. For each of the following payoff sets $W \subseteq [T]$, decide which player has a winning strategy in $(T, W)$ and describe one: (a) the plays containing at least one 1; (b) the plays with $a_n = 0$ for every $n \in \omega$; (c) the plays with $a_1 = a_0$; (d) the plays in which Player I plays 1 infinitely often.

(2) Let $(T, W)$ be an infinite game, and let $\sigma$ and $\tau$ be strategies for Player I and Player II, respectively. Show that there is exactly one play consistent with both $\sigma$ and $\tau$, a fact used in Remarks 2.3(3).

(3) Let $T$ be a tree, and let $W \subseteq W' \subseteq [T]$ be given. Show that a winning strategy for Player II in $(T, W)$ is winning in $(T, W')$, and that a winning strategy for Player I in $(T, W')$ is winning in $(T, W)$.

(4) Let $(T, W)$ be an infinite game, and suppose that there is a natural number $N$ such that whether a play lies in $W$ depends only on the first $N$ moves of the play. Show that the game $(T, W)$ is determined. (Hint: decide who wins from each position of length $N$, and work backwards to $\langle\rangle$.)

(5) $(*)$ Let $A = \{0, 1\}$ and $T = A^{<\omega}$, and let $W$ be the set of plays $(a_n)_{n\in\omega}$ with $a_{2n+1} = a_0$ for every $n \in \omega$. Show that Player II has a winning strategy, but that no strategy for her whose move depends only on the last move of Player I is winning.

## 3. The Banach–Mazur game

In the Banach–Mazur game, two players take turns shrinking an open set, and the second player wants the sets to keep a point in common. In this section we define the game and show

that a non-empty space is a Baire space exactly when the first player has no winning strategy (see Theorem 3.3 below). We then let the players aim at a given set and prove the theorem of Banach and Oxtoby (see Theorem 3.5 below). This theorem asserts that the second player can force the intersection of the moves to lie inside the set exactly when the complement of the set is of the first category.

**Definition 3.1.** Let $X$ be a non-empty topological space, and let $\mathcal{O}^+(X)$ be the collection of all non-empty open subsets of $X$. The *Banach–Mazur game* on $X$, denoted by $\mathrm{BM}(X)$, is the game $(T, W)$ where:

(1) The set of moves is $A = \mathcal{O}^+(X)$.
(2) The tree $T$ consists of all finite sequences $(U_0, V_0, U_1, V_1, \dots)$ of sets in $\mathcal{O}^+(X)$ with $U_0 \supseteq V_0 \supseteq U_1 \supseteq V_1 \supseteq \cdots$.
(3) The payoff set is $W = \{(U_0, V_0, U_1, \dots) \in [T] : \bigcap_{n=0}^{\infty} U_n \neq \emptyset\}$.

Thus Player I plays the sets $U_n$, Player II plays the sets $V_n$, and Player II wins if the sets played have a point in common.

**Example 3.2.** (1) In $\mathrm{BM}(\mathbb{Q})$, Player I has a winning strategy. Fix an enumeration $q_0, q_1, \dots$ of $\mathbb{Q}$, and let Player I play $U_0 = \mathbb{Q} \setminus \{q_0\}$ and $U_n = V_{n-1} \setminus \{q_n\}$ for $n \geq 1$. These sets are open because the points of $\mathbb{Q}$ are closed, and they are non-empty because $\mathbb{Q}$ has no isolated points. Since every $q_n$ is missing from $U_n$, the intersection $\bigcap_n U_n$ is empty (see Figure 6).

(2) In $\mathrm{BM}(\mathbb{R})$, Player II has a winning strategy: in answer to $U_n$, she plays a bounded open interval $V_n$ with $\overline{V_n} \subseteq U_n$. The compact sets $\overline{V_n}$ are nested and non-empty, hence they have a common point, and that point lies in every $U_n$ because $\overline{V_n} \subseteq U_n$.

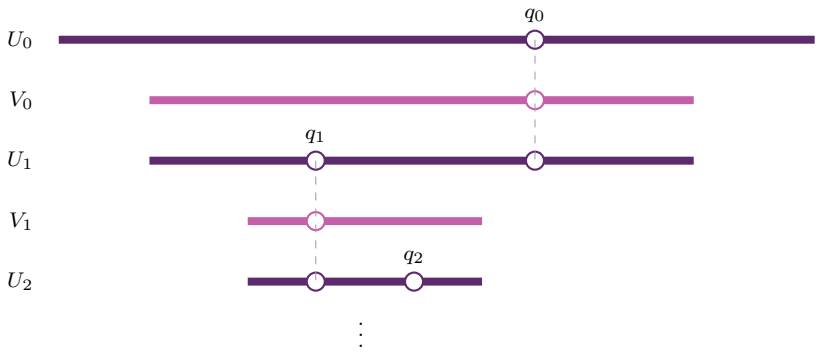



FIGURE 6. The winning strategy of Player I in $\mathrm{BM}(\mathbb{Q})$ of Example 3.2(1), which removes at step $n$ the rational $q_n$ of a fixed enumeration of $\mathbb{Q}$.

**Theorem 3.3.** *Let $X$ be a non-empty topological space.*

(1) *Player I has a winning strategy in* $\mathrm{BM}(X)$ *if and only if $X$ contains a non-empty open set of the first category.*
(2) *The space $X$ is a Baire space if and only if Player I has no winning strategy in* $\mathrm{BM}(X)$.

*Proof.* By Theorem 1.3, part (2) is an immediate consequence of part (1), which we now prove. The idea is that a nowhere dense set can be avoided in a single move. If a non-empty open set is a countable union of nowhere dense sets, Player I opens with this set and avoids

one of the pieces at each later move. Conversely, the plays of a winning strategy for Player I, laid side by side, show that his first move is of the first category.

Suppose first that $G \subseteq X$ is open, non-empty, and of the first category. Then there are nowhere dense sets $E_0, E_1, \ldots$ with $G = \bigcup_{n=0}^{\infty} E_n$. Player I opens with $U_0 = G$; if Player II responds with $V_0 \subseteq U_0$, the set $V_0 \setminus \overline{E_0}$ is open, and the set is non-empty because $\overline{E_0}$ has empty interior. Player I may therefore choose a non-empty open set $U_1 \subseteq V_0 \setminus \overline{E_0}$. In general, for $n \geq 1$, given $V_{n-1}$, Player I chooses an open set

$$U_n \subseteq V_{n-1} \setminus \overline{E_{n-1}}, \qquad U_n \neq \emptyset.$$

This defines a strategy $\sigma$ for Player I. Every play consistent with $\sigma$ satisfies $U_n \cap E_{n-1} = \emptyset$ for all $n \geq 1$, whence

$$\bigcap_{n=0}^{\infty} U_n \subseteq G \setminus \bigcup_{n=0}^{\infty} E_n = \emptyset.$$

Thus the strategy $\sigma$ is winning for Player I.

Conversely, suppose that Player I has a winning strategy $\sigma$, and let $U_0 = \sigma(\langle\rangle)$. We show that the non-empty open set $U_0$ is of the first category. For a position $p$ of odd length $2n+1$ consistent with $\sigma$, we shall write $U^p$ for the last entry of $p$, the set $U_n$ played by Player I.

We build, for each $n \geq 0$, a family $\mathcal{P}_n$ of positions of length $2n+1$ consistent with $\sigma$, such that:

(a) the sets $U^p$, $p \in \mathcal{P}_n$, are pairwise disjoint;
(b) their union $D_n = \bigcup_{p \in \mathcal{P}_n} U^p$ is dense in $U_0$;
(c) every $p \in \mathcal{P}_{n+1}$ extends some $q \in \mathcal{P}_n$.

The family $\mathcal{P}_0$ consists of the single position $(U_0)$. Suppose that $\mathcal{P}_n$ has been built, and let $q \in \mathcal{P}_n$ be given. Each non-empty open $V \subseteq U^q$ is a legal move for Player II at $q$, and leads to the position $q^\frown V^\frown \sigma(q^\frown V)$ of length $2n+3$. The families of such positions whose last entries are pairwise disjoint are closed under unions of chains. By Zorn's lemma, we may therefore choose a maximal family $\mathcal{P}(q)$ among them. The union of the last entries of the positions in $\mathcal{P}(q)$ is dense in $U^q$: otherwise some non-empty open $O \subseteq U^q$ would miss all these last entries. In that case the position $q^\frown O^\frown \sigma(q^\frown O)$, whose last entry lies in $O$, could be added to $\mathcal{P}(q)$, contradicting maximality. Let $\mathcal{P}_{n+1}$ be the union $\bigcup_{q \in \mathcal{P}_n} \mathcal{P}(q)$, a family for which property (c) holds by construction. Property (a) holds because last entries of positions in the same $\mathcal{P}(q)$ are disjoint, and those in $\mathcal{P}(q)$ and $\mathcal{P}(q')$ with $q \neq q'$ lie in the disjoint sets $U^q$ and $U^{q'}$. For (b), we argue by induction on $n$; the case $n = 0$ is trivial, since $D_0 = U_0$. By the inductive hypothesis, every non-empty open set $O \subseteq U_0$ meets $D_n$, hence meets $U^q$ for some $q \in \mathcal{P}_n$. The union of the last entries of the positions in $\mathcal{P}(q)$ is dense in $U^q$ and therefore meets the non-empty open set $O \cap U^q$. As this union is contained in $D_{n+1}$, the set $O$ meets $D_{n+1}$. The first two levels are shown in Figure 7.

Each $D_n$ is open and dense in $U_0$; hence the set $U_0 \setminus D_n$ is nowhere dense in $X$. Suppose, toward a contradiction, that $U_0$ is of the second category. Then the first-category set $\bigcup_n (U_0 \setminus D_n)$ does not contain $U_0$, and there is a point $x \in \bigcap_n D_n$. By (a), for each $n$ there is exactly one $p_n \in \mathcal{P}_n$ with $x \in U^{p_n}$. Property (c) provides a position $q \in \mathcal{P}_n$ extended by $p_{n+1}$; since the moves along a play are nested, we have $x \in U^{p_{n+1}} \subseteq U^q$, and property (a) forces $q = p_n$. Thus the positions $p_0 \subseteq p_1 \subseteq \cdots$ fit together into a single play consistent with $\sigma$, and the point $x$ lies in every set $U^{p_n}$ that Player I plays in this play. This contradicts the assumption that $\sigma$ is winning. Hence the opening move $U_0$ is a non-empty open set of the first category, and the proof is complete. ☺

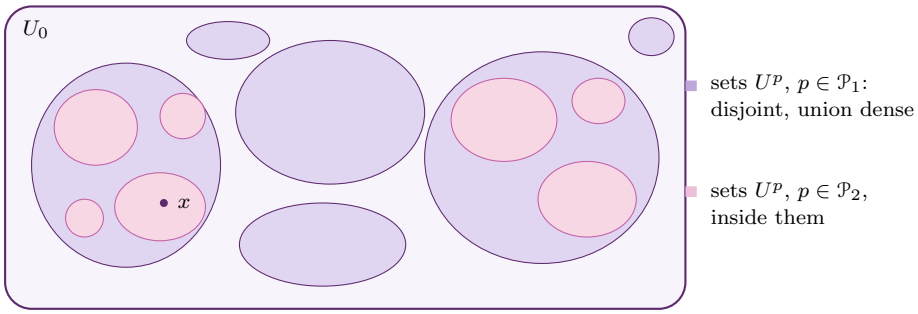



FIGURE 7. The first two levels of the construction in the proof of Theorem 3.3, inside the opening move $U_0$ of a winning strategy $\sigma$ of Player I.

**Remark 3.4.** When a non-empty topological space $X$ is a Baire space, part (2) of Theorem 3.3 rules out a winning strategy for Player I in $\mathrm{BM}(X)$. Player II, however, need not have one either: there are Baire spaces in which neither player has a winning strategy (see [Tel87]). Such spaces occur already among the subspaces of the real line. In 1908 Bernstein constructed subsets of $\mathbb{R}$ that, like their complements, meet every uncountable closed subset of $\mathbb{R}$. Every set $B \subseteq \mathbb{R}$ with this property, regarded as a subspace of $\mathbb{R}$, is a Baire space on which neither player of $\mathrm{BM}(B)$ has a winning strategy.

In Mazur's original game, and in most applications in Part II, the players aim at a given set. Let $X$ be a non-empty topological space, and let $S \subseteq X$ be given. The game $\mathrm{BM}(X, S)$ has the same moves as $\mathrm{BM}(X)$, but Player II wins the play $(U_0, V_0, U_1, \dots)$ if

$$\bigcap_{n=0}^{\infty} U_n \subseteq S,$$

that is, if no point outside $S$ survives. Since the moves are nested, the moves of the two players have the same intersection, $\bigcap_n U_n = \bigcap_n V_n$. Unlike in $\mathrm{BM}(X)$, Player II also wins when the intersection is empty; Player I wins only if he produces a point outside $S$.

**Theorem 3.5 (Banach; Oxtoby).** *Let $X$ be a non-empty topological space, and let $S \subseteq X$ be given.*

(1) *Player II has a winning strategy in $\mathrm{BM}(X, S)$ if and only if $X \setminus S$ is of the first category in $X$.*
(2) *If $X$ is a complete metric space, then Player I has a winning strategy in $\mathrm{BM}(X, S)$ if and only if there is a non-empty open set $U$ such that $U \cap S$ is of the first category in $X$.*

*Proof.* Both parts rest on the two ideas of the proof of Theorem 3.3, of which the first is that a nowhere dense set can be avoided in a single move. The second is that a winning strategy, unfolded into maximal families of positions, exhibits a set of the first category. (1) Suppose first that $X \setminus S = \bigcup_n E_n$ with each $E_n$ nowhere dense. In answer to $U_n$, Player II plays a non-empty open set $V_n \subseteq U_n \setminus \overline{E_n}$. A point of $\bigcap_n V_n$ lies outside every $E_n$, hence in $S$, so this strategy is winning.

Conversely, let $\tau$ be a winning strategy for Player II. We repeat the construction in the proof of Theorem 3.3, now with the moves of Player II. For a position $p$ of non-zero even length consistent with $\tau$, we write $V^p$ for the last entry of $p$. As in the proof of Theorem 3.3, Zorn's lemma provides families $\mathcal{P}_n$, for $n \in \omega$, of positions of length $2n + 2$ consistent with

$\tau$ such that every $p \in \mathcal{P}_{n+1}$ extends some $q \in \mathcal{P}_n$. Within each family $\mathcal{P}_n$, the sets $V^p$ are pairwise disjoint, and their union $D_n$ is dense in $X$. For $\mathcal{P}_0$ we take a maximal family of positions $(U, \tau(U))$ whose last entries are pairwise disjoint. The union of the last entries of the positions in $\mathcal{P}_0$ is dense, because for every non-empty open set $O$ the answer $\tau(O)$ lies in $O$. At a position $q \in \mathcal{P}_n$, every non-empty open $U \subseteq V^q$ is a legal move for Player I, and the answer of $\tau$ lies in $U$, so the maximality argument goes through unchanged. Now let $x \in \bigcap_n D_n$ be given; exactly as before, the positions $p_n \in \mathcal{P}_n$ with $x \in V^{p_n}$ fit together into a single play consistent with $\tau$, and every set played in this play contains $x$. Since $\tau$ is winning, the point $x$ lies in $S$. Thus the intersection $\bigcap_n D_n$ is contained in $S$, whence $X \setminus S \subseteq \bigcup_n (X \setminus D_n)$. Each $X \setminus D_n$ is closed with empty interior, hence nowhere dense, and therefore the complement $X \setminus S$ is of the first category.

(2) Let $d$ be the metric of $X$, which is complete by hypothesis. Suppose first that $U$ is a non-empty open set and that $U \cap S = \bigcup_n E_n$ with each $E_n$ nowhere dense. Player I plays $U_0 = U$ and, in answer to $V_{n-1}$ for each $n \geq 1$, a non-empty open set $U_n$ with $\overline{U_n} \subseteq V_{n-1} \setminus \overline{E_{n-1}}$ and with diameter $\operatorname{diam} U_n = \sup\{d(x,y) : x, y \in U_n\}$ less than $2^{-n}$. The closed sets $\overline{U_n}$ are nested and their diameters tend to 0, so by completeness they have a point $x$ in common. The point $x$ lies in every $U_n$, because $x \in \overline{U_{n+1}} \subseteq V_n \subseteq U_n$, and in particular in $U = U_0$; the point lies in no $E_n$, because $\overline{U_{n+1}}$ misses $\overline{E_n}$. Hence Player I has produced a point of $U \setminus S$, namely $x$, and he has won.

Conversely, let $\sigma$ be a winning strategy for Player I, and write $U_0 = \sigma(\langle\rangle)$. We show that $U_0 \cap S$ is of the first category. Once more we follow the construction in the proof of Theorem 3.3, with one change: at level $n$, the families $\mathcal{P}(q)$ use only moves $V$ of Player II with $\operatorname{diam} V < 2^{-n}$. Every non-empty open set contains such a set $V$ (a small open ball), and the maximality argument shows, as before, that the unions $D_n$ are dense in $U_0$. Every point $x \in \bigcap_n D_n$ again picks out a play consistent with $\sigma$ whose intersection contains $x$. The diameters of the moves tend to 0, so this intersection is exactly $\{x\}$. Since $\sigma$ is winning, the intersection contains a point outside $S$, and this point can only be $x$. Thus the intersection $\bigcap_n D_n$ lies in $U_0 \setminus S$, and the set $U_0 \cap S$ is contained in $\bigcup_n (U_0 \setminus D_n)$, which is of the first category because each $U_0 \setminus D_n$ is nowhere dense. Hence Player I's first move $U_0$ is the required non-empty open set. ☺



**Remark 3.6.** By part (1) of Theorem 3.5, Player II has a winning strategy in $\mathrm{BM}(\mathbb{R}, S)$ when $S$ is the set of irrational numbers. One such strategy removes the rationals one at a time, just as Player I did in Example 3.2(1). Part (1) needs no hypothesis on the space. In part (2), the "only if" direction uses the metric but not the completeness of the metric, which is needed only in the "if" direction, where Player I must produce a point. Without completeness the "if" direction fails, as the space $X = \mathbb{Q}$ with the usual metric shows for $S = \emptyset$: the set $U \cap S$ is then empty for every non-empty open set $U$. Player II nevertheless has a winning strategy in $\mathrm{BM}(\mathbb{Q}, \emptyset)$ by part (1), because $\mathbb{Q}$ is of the first category in itself.

**Exercises.**

(1) Let $C$ be the Cantor ternary set (see Figure 3). For each of the sets $S$ below, decide which player has a winning strategy in $\mathrm{BM}(\mathbb{R}, S)$, and describe one explicitly: (a) the irrational numbers; (b) the rational numbers; (c) the set $C$; (d) the set $\mathbb{R} \setminus C$. Check your answers against Theorem 3.5.

(2) Let $X$ be a non-empty compact Hausdorff space. Show that Player II has a winning strategy in $\mathrm{BM}(X)$. (Hint: such a space is regular, so every point of a non-empty open set $U$ has an open neighborhood whose closure lies in $U$.)

(3) Let $X$ be a non-empty topological space, and let $S_0, S_1, \dots \subseteq X$ be given. Suppose that Player II has a winning strategy $\tau_k$ in $\mathrm{BM}(X, S_k)$ for every $k \in \omega$. Show directly that she has a winning strategy in $\mathrm{BM}(X, \bigcap_k S_k)$, and then give a second proof using Theorem 3.5. (Hint for the direct proof: split the set $\omega$ of rounds into pairwise disjoint infinite sets $R_k$, for $k \in \omega$, and let her follow $\tau_k$ in the rounds that belong to $R_k$.)

(4) Let $X$ be a Baire space, and let $U \subseteq X$ be a non-empty open set. Use Theorem 3.3 to show that $U$ is a Baire space, by turning a winning strategy for Player I in $\mathrm{BM}(U)$ into one in $\mathrm{BM}(X)$.

(5) $(*)$ Let $X$ be a non-empty complete metric space, and let $S \subseteq X$ have the Baire property. Show that the game $\mathrm{BM}(X, S)$ is determined. (Hint: if $U \cap S$ is of the first category for no non-empty open set $U$, show that every open set $O$ with $S \triangle O$ of the first category is dense.)

## 4. The strong Choquet game

The Banach–Mazur game detects category, but not completeness: a metrizable space for which Player II has a winning strategy in the Banach–Mazur game need not be completely metrizable (see Remark 4.6 below). In this section we let Player I name a point with each of his moves, and we show that the resulting game characterizes complete metrizability among metrizable spaces (see Theorem 4.4 below). Here we prove the easy half of the theorem and derive from it the Baire category theorem (see Corollary 4.5 below); the hard half is proved in Section 5. At the end of the section we leave metric spaces and turn to Čech-completeness, a form of completeness that makes sense in every Tychonov space. In Subsection 4.1 we show that Player II has a winning strategy in the new game on every non-empty Čech-complete space, and that the converse fails.

**Definition 4.1.** Let $X$ be a non-empty topological space, and recall that $\mathcal{O}^+(X)$ is the collection of all non-empty open subsets of $X$. The *strong Choquet game* on $X$, denoted by $\mathrm{Ch}(X)$, is played as follows. Player I chooses a pair $(x_0, U_0)$ with $U_0 \in \mathcal{O}^+(X)$ and $x_0 \in U_0$. Player II responds with $V_0 \in \mathcal{O}^+(X)$ such that $x_0 \in V_0 \subseteq U_0$. At step $n \geq 1$, Player I chooses $(x_n, U_n)$ such that $x_n \in U_n \subseteq V_{n-1}$, and Player II responds with $V_n \in \mathcal{O}^+(X)$ such that $x_n \in V_n \subseteq U_n$. Player II wins the play if $\bigcap_{n=0}^{\infty} U_n \neq \emptyset$.

In the format of Definition 2.2, Player I chooses from $X \times \mathcal{O}^+(X)$ and Player II from $\mathcal{O}^+(X)$. If the points $x_n$ are omitted, what remains is exactly the Banach–Mazur game $\mathrm{BM}(X)$, which in this context is also called the *Choquet game*. Choquet called the two players $\beta$ (the first) and $\alpha$ (the second). Following him, we call a non-empty topological space $X$ *$\alpha$-favorable* if Player II has a winning strategy in $\mathrm{BM}(X)$, and *strongly $\alpha$-favorable* if she has a winning strategy in $\mathrm{Ch}(X)$. Kechris [Kec95] calls these spaces *Choquet spaces* and *strong Choquet spaces*, respectively.

**Lemma 4.2.** *Every strongly $\alpha$-favorable space is $\alpha$-favorable.*

*Proof.* The idea is that Player II can supply the missing points herself. Let $X$ be a strongly $\alpha$-favorable space, and let $\tau$ be a winning strategy for Player II in $\mathrm{Ch}(X)$. In $\mathrm{BM}(X)$, when Player I plays $U_n$, Player II picks a point $x_n \in U_n$ and answers with the set that $\tau$ assigns to the position $(x_0, U_0), V_0, \dots, V_{n-1}, (x_n, U_n)$ of $\mathrm{Ch}(X)$. The first move $(x_0, U_0)$ is legal, and for $n \geq 1$ we have $x_n \in U_n \subseteq V_{n-1}$, so every subsequent move $(x_n, U_n)$ is legal in $\mathrm{Ch}(X)$. Every play of $\mathrm{BM}(X)$ that follows this rule therefore arises, by omitting the points, from a play of $\mathrm{Ch}(X)$ consistent with $\tau$. Because $\tau$ is winning, the sets $U_n$ of every such play have a common point. Hence this rule is a winning strategy for Player II in $\mathrm{BM}(X)$. ☺

**Example 4.3.** Player I wins Ch($\mathbb{Q}$): he follows the strategy of Example 3.2(1), adding any point of each set he plays.

**Theorem 4.4 (Choquet).** *A non-empty metrizable space is completely metrizable if and only if it is strongly $\alpha$-favorable.*

*Proof.* Let $X$ be a metrizable space, and suppose first that $X$ is completely metrizable, with a compatible complete metric $d$. We define a strategy $\tau$ for Player II whose answers are balls with radii tending to 0. The points of Player I then form a Cauchy sequence, and the limit of the sequence is a common point of the play. When Player I plays $(x_n, U_n)$, Player II chooses $r_n$ with $0 < r_n < 2^{-n}$ such that the closed ball $B[x_n, r_n] = \{y \in X : d(x_n, y) \leq r_n\}$ is contained in $U_n$, and plays the open ball $V_n = B(x_n, r_n) = \{y \in X : d(x_n, y) < r_n\}$.

Let $(x_0, U_0), V_0, (x_1, U_1), V_1, \ldots$ be a play consistent with $\tau$. For $m > n$ we have $x_m \in U_m \subseteq V_n = B(x_n, r_n)$, whence $d(x_m, x_n) < 2^{-n}$. Thus the points $x_n$ form a Cauchy sequence, which converges to some $x \in X$ because $d$ is complete. For each $n$, the points $x_m$ with $m \geq n$ all lie in the closed ball $B[x_n, r_n]$; so does the limit $x$, and the ball $B[x_n, r_n]$ is contained in $U_n$. Consequently, the limit $x$ lies in every $U_n$, and the strategy $\tau$ is winning for Player II (see Figure 8).

The converse, that a strongly $\alpha$-favorable metrizable space is completely metrizable, is proved in Section 5.

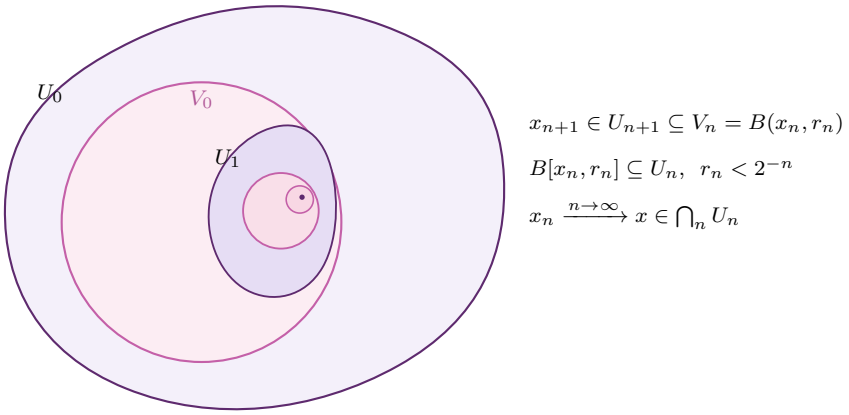



FIGURE 8. The strategy of Player II in the easy half of Theorem 4.4, which answers each point $x_n$ of Player I with the open ball $B(x_n, r_n)$.

**Corollary 4.5 (Baire category theorem).** *Every completely metrizable space is a Baire space.*

*Proof.* Let $X$ be completely metrizable; since the empty space is trivially a Baire space, we may assume that $X$ is non-empty. By Theorem 4.4 and Lemma 4.2, Player II has a winning strategy in BM($X$). Since at most one player has a winning strategy (see Remarks 2.3(3)), Player I has none, and by Theorem 3.3 the space $X$ is a Baire space.

**Remark 4.6.** The points $x_n$ named by Player I cannot be dropped from the strong Choquet game in Theorem 4.4: an $\alpha$-favorable metrizable space need not be completely metrizable. Let $X = \{(a, b) \in \mathbb{R}^2 : b > 0\} \cup (\mathbb{Q} \times \{0\})$, with the topology inherited from $\mathbb{R}^2$. The upper half plane $H = \{(a, b) \in \mathbb{R}^2 : b > 0\}$ is open and dense in $X$. Player II wins BM($X$) by

always answering with an open disk whose closure lies in $U_n \cap H$ and whose radius is less than $2^{-n}$. The closures of these disks are nested and compact, and their common point lies in every $U_n$. The space $X$, however, is not completely metrizable. Indeed, a closed subspace of a completely metrizable space is completely metrizable, since a Cauchy sequence in the subspace converges in the space and the limit lies in the subspace. By Corollary 4.5, such a subspace is therefore a Baire space. The closed subspace $\mathbb{Q} \times \{0\}$ of $X$ is nevertheless not a Baire space, being a copy of $\mathbb{Q}$ and hence a countable union of nowhere dense singletons. By the hard half of Theorem 4.4 (proved in Section 5), it follows that Player II has no winning strategy in $\mathrm{Ch}(X)$. A direct argument, independent of Section 5, leads to the same conclusion. The argument rests on the observation that in $\mathrm{Ch}(X)$ Player I can choose his points on the horizontal axis, and that Player II must answer with sets that contain them. The strategy by which Player I then wins is shown in Figure 9.

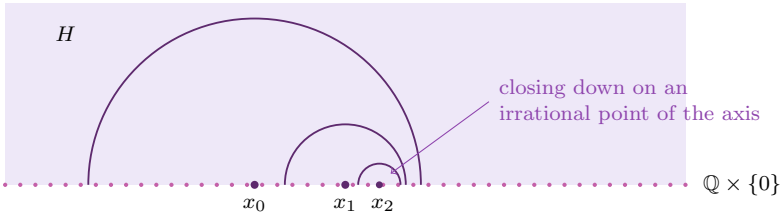



FIGURE 9. The winning strategy of Player I in $\mathrm{Ch}(X)$ for the space $X$ of Remark 4.6, by balls about rational points of the horizontal axis that close down on a point of the axis that is not rational.

4.1. **Čech-complete spaces.** Because Player I names points, the strong Choquet game detects complete metrizability among metrizable spaces, which the Banach–Mazur game cannot do (see Remark 4.6). Completeness, however, is a metric notion; the notion has a purely topological counterpart, to which the easy half of Theorem 4.4 extends.

To make this precise, recall that a *Tychonov space* is a Hausdorff space in which every closed set and every point outside it are separated by a continuous real-valued function. Recall also that a *compactification* of a Tychonov space $X$ is a compact Hausdorff space that contains $X$ as a dense subspace. The Stone–Čech compactification $\beta X$ is the largest compactification of $X$, in the sense that every compactification of $X$ is the image of $\beta X$ under a continuous map that fixes the points of $X$.

**Definition 4.7 (Čech, 1937).** A Tychonov space $X$ is *Čech-complete* if $X$ is a $G_\delta$ subset of the Stone–Čech compactification $\beta X$.

A Tychonov space is Čech-complete if and only if the space is a $G_\delta$ subset of some compactification, and the space is then a $G_\delta$ subset of every compactification (see Engelking's book [Eng89, Section 3.9]). In particular, every compact Hausdorff space is Čech-complete, and so is every locally compact Hausdorff space, since a non-compact one is open in its one-point compactification. For metrizable spaces nothing new arises: a metrizable space is Čech-complete exactly when the space is completely metrizable (Engelking proves this theorem of Čech in his book [Eng89, Theorem 4.3.26]). By Theorem 4.4, Čech-completeness and strong $\alpha$-favorability therefore coincide on non-empty metrizable spaces. Outside metrizable spaces only one implication survives.

**Proposition 4.8.** *Every non-empty Čech-complete space is strongly $\alpha$-favorable.*

*Proof.* The idea is to play the game in $\beta X$, where compactness supplies a common point of the play, while the $G_\delta$ condition keeps that point in $X$. Let $X$ be Čech-complete, and let $G_0, G_1, \dots$ be open subsets of $\beta X$ with $X = \bigcap_n G_n$; closures below are taken in $\beta X$. Along with each answer $V_n$, Player II records an open set $W_n \subseteq \beta X$ with $V_n = W_n \cap X$, and she chooses the two sets as follows. When Player I plays $(x_n, U_n)$, she takes an open set $\widetilde{U}_n \subseteq \beta X$ with $\widetilde{U}_n \cap X = U_n$. For $n \geq 1$ she may assume that $\widetilde{U}_n \subseteq W_{n-1}$: since $U_n \subseteq V_{n-1} \subseteq W_{n-1}$, the set $\widetilde{U}_n \cap W_{n-1}$ has the same intersection $U_n$ with $X$. The point $x_n$ lies in the open set $\widetilde{U}_n \cap G_n$, and the compact Hausdorff space $\beta X$ is regular; hence there is an open set $W_n$ with $x_n \in W_n$ and $\overline{W_n} \subseteq \widetilde{U}_n \cap G_n$. Player II answers with $V_n = W_n \cap X$, a legal move because $x_n \in V_n \subseteq \widetilde{U}_n \cap X = U_n$.

In every play consistent with this strategy, the inclusions $\overline{W_{n+1}} \subseteq \widetilde{U}_{n+1} \subseteq W_n$ hold for every $n$, and the sets $\overline{W_n}$ thus form a decreasing sequence of compact sets, which are non-empty because $x_n \in \overline{W_n}$. By compactness, their intersection contains a point $y$ (see Figure 10). Since each $\overline{W_n}$ lies in $\widetilde{U}_n \cap G_n$, the point $y$ lies in $\bigcap_n G_n = X$. Consequently, this point lies in $\widetilde{U}_n \cap X = U_n$ for every $n$, whence the play has the common point $y$ and the strategy is winning. ☺

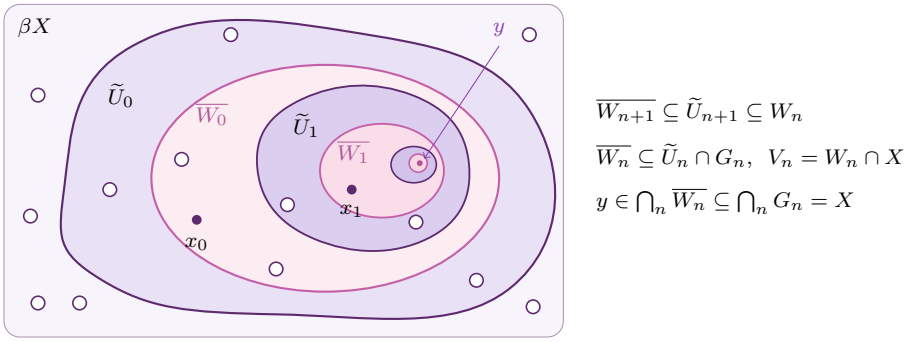



FIGURE 10. The proof of Proposition 4.8, drawn in $\beta X$, where $X = \bigcap_n G_n$ for open subsets $G_0, G_1, \dots$ of $\beta X$ and the white dots are points of $\beta X$ that are not in $X$.

As in the proof of Corollary 4.5, it follows that every Čech-complete space is a Baire space (see Figure 11). The converse of Proposition 4.8 fails, even among paracompact spaces: the Michael line, obtained from the real line by isolating the irrational points, is strongly $\alpha$-favorable without being Čech-complete (see Example 4.9 below). Here a Hausdorff space is *paracompact* if every open cover of the space has a locally finite open refinement that covers the space; locally finite families and refinements are defined before Lemma 5.3 below.

**Example 4.9 (The Michael line).** Let $M$ be the real line with the topology whose open sets are the unions of a set open in the usual topology with a set of irrational numbers. Every irrational number is then an isolated point, while every rational number keeps its usual neighborhoods (see Figure 12). The space $M$ is Hausdorff and paracompact (see Engelking's book [Eng89, Example 5.1.32]), hence normal, and in particular Tychonov.

(1) *The space $M$ is strongly $\alpha$-favorable.* When Player I plays $(x_n, U_n)$ with $x_n$ irrational, Player II answers with $V_n = \{x_n\}$; every later move equals $\{x_n\}$, and the point $x_n$ lies in every set of the play. When $x_n$ is rational, the set $U_n$ contains a usual open interval about $x_n$. In that case Player II answers with an interval $V_n = (x_n - r_n, x_n + r_n)$,

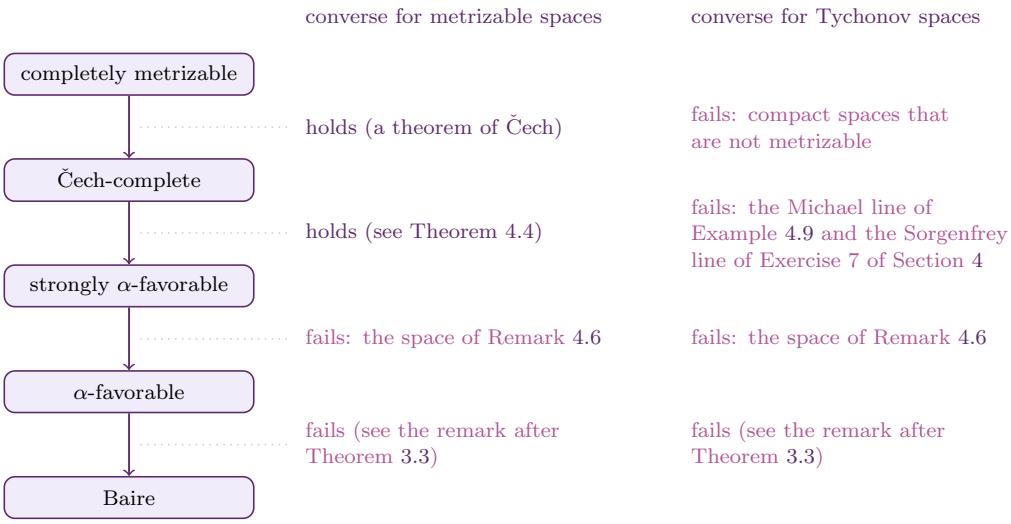



FIGURE 11. Completeness properties of a Tychonov space, from the strongest to the weakest, with whether the converse of each arrow holds among metrizable spaces and among all Tychonov spaces.

where $r_n > 0$ is so small that the closed interval $[x_n - r_n, x_n + r_n]$ lies in $U_n$. If Player I plays only rational points, these closed intervals are nested, since each lies in the previous answer, and compact in the usual topology of $\mathbb{R}$; they therefore have a common point, which lies in every $U_n$.

(2) *The space $M$ is not Čech-complete.* The set $\mathbb{Q}$ is closed in $M$, because the complement of $\mathbb{Q}$ consists of isolated points, and the subspace topology of $\mathbb{Q}$ is the usual one. If $M$ were Čech-complete, then the closed subspace $\mathbb{Q}$ of $M$ would be Čech-complete as well, since closed subspaces of Čech-complete spaces are Čech-complete (see Engelking's book [Eng89, Theorem 3.9.6]). The subspace $\mathbb{Q}$ would therefore be a Baire space; yet the space $\mathbb{Q}$, the union of countably many nowhere dense singletons, is not a Baire space.

The Michael line separates the two properties in one more respect: unlike Čech-completeness, strong $\alpha$-favorability does not pass to closed subspaces, since Player I wins $\mathrm{Ch}(\mathbb{Q})$ (see the example after Lemma 4.2).

Among paracompact spaces, Čech-completeness also has an exact characterization by a game of another kind, played not on the space but on what the Stone–Čech compactification adds to it. The characterization asserts that a paracompact space $X$ is Čech-complete if and only if Player II has a winning strategy in the Menger game of Section 14 below, played on the remainder $\beta X \setminus X$. Aurichi and Bella [AB15, Corollary 3.3] deduce this equivalence from a theorem of Telgársky [Tel84].

**Exercises.**

(1) Let $X$ be a non-empty compact Hausdorff space. Show directly, without Proposition 4.8, that Player II has a winning strategy in $\mathrm{Ch}(X)$, that is, that $X$ is strongly $\alpha$-favorable.

(2) Let $X$ be a non-empty completely metrizable space with no isolated points. Show that the space $X$ is uncountable. (Hint: each singleton is closed with empty interior; use Corollary 4.5.)

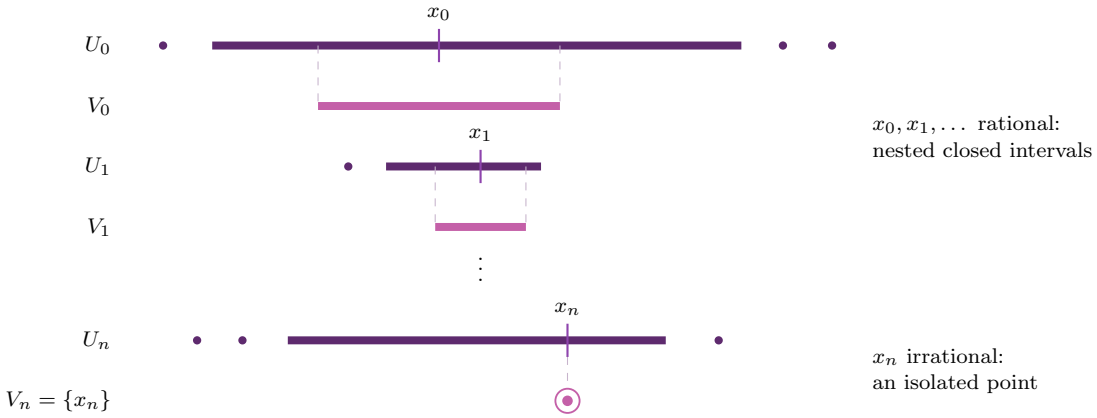



FIGURE 12. The answers of Player II in the strong Choquet game on the Michael line $M$ of Example 4.9, to rational points (top) and to an irrational point (bottom), with the isolated irrational points drawn as small dots.

(3) The interval $(0, 1)$ is not complete in the usual metric. Show directly that Player II has a winning strategy in $\mathrm{Ch}((0, 1))$. (Hint: imitate the proof of Theorem 4.4, choosing $r_n < 2^{-n}$ so that the closed interval $[x_n - r_n, x_n + r_n]$ lies in $U_n$.)

(4) Let $X = \{(a, b) \in \mathbb{R}^2 : b > 0\} \cup (\mathbb{Q} \times \{0\})$ be the space of Remark 4.6. Write out in full the winning strategy for Player I in $\mathrm{Ch}(X)$ indicated in Figure 9, and check that the strategy is winning. (Hint: enumerate $\mathbb{Q} \times \{0\}$ as $p_0, p_1, \ldots$, and at step $n$ choose $U_n$ inside a closed disk $D_n$ of radius less than $2^{-n}$ that misses $p_n$ and, for $n \geq 1$, lies in $D_{n-1}$.)

(5) $(*)$ Let $P = \mathbb{R} \setminus \mathbb{Q}$ be the space of irrational numbers, with the usual metric, and enumerate $\mathbb{Q}$ as $q_0, q_1, \ldots$. Show directly that Player II has a winning strategy in $\mathrm{Ch}(P)$. (Hint: answer $(x_n, U_n)$ with $V_n = (x_n - r_n, x_n + r_n) \cap P$, where the closed interval $[x_n - r_n, x_n + r_n]$ misses $q_n$ and the points of this interval that lie in $P$ lie in $U_n$.)

(6) Recall that a topological space is *countably compact* if every countable open cover has a finite subcover, or, equivalently, if every decreasing sequence of non-empty closed sets has a non-empty intersection. Show that every non-empty countably compact regular space is strongly $\alpha$-favorable (this extends Exercise 1). (Hint: answer $(x_n, U_n)$ with an open set $V_n$ such that $x_n \in V_n$ and $\overline{V_n} \subseteq U_n$.)

(7) $(*)$ Let $S$ be the Sorgenfrey line, the set $\mathbb{R}$ with the topology whose basic open sets are the intervals $[a, b)$ with $a < b$ real. Show that $S$ is strongly $\alpha$-favorable. (Hint: answer $(x_n, U_n)$ with $V_n = [x_n, x_n + r_n)$, where $r_n > 0$ and $[x_n, x_n + 2r_n) \subseteq U_n$. The points $x_n$ then form a non-decreasing sequence, bounded above by $x_0 + r_0$, and the limit of this sequence lies in every $U_n$.) The space $S$ is, after the Michael line, a second strongly $\alpha$-favorable space that is not Čech-complete. Indeed, by the paragraph after Theorem 10.1 below, Čech-complete spaces are strongly Baire (a property defined in Section 10), whereas $S$ is not (see Exercise 4 of Section 10).

## 5. The converse of Choquet's theorem

In this section we prove the hard half of Theorem 4.4: every non-empty metrizable space on which Player II has a winning strategy in the strong Choquet game is completely metrizable.

The proof is long, so we first explain where a complete metric can come from and what the strategy has to do with such a metric.

**The completion.** Let $X$ be a non-empty metrizable space, and let $\tau$ be a winning strategy for Player II in the strong Choquet game $\mathrm{Ch}(X)$. We fix a compatible metric $d$ on $X$, and we denote by $\widetilde{X}$ the completion of $(X, d)$. If $d$ is complete, then $\widetilde{X} = X$; in general the completion contains new points, and the position of $X$ inside the completion decides whether $X$ is completely metrizable. Two subspaces of $\mathbb{R}$ illustrate the role of this position. The irrationals form the $G_\delta$ set $\bigcap_{q\in\mathbb{Q}}(\mathbb{R} \setminus \{q\})$, and they are completely metrizable: continued fractions make them homeomorphic to the space of sequences of positive integers, which is complete in its usual metric. The rationals do not form a $G_\delta$ subset of $\mathbb{R}$, for otherwise the rationals and the irrationals would be two disjoint dense $G_\delta$ subsets of $\mathbb{R}$. By the Baire category theorem (see Corollary 4.5), however, the intersection of two dense $G_\delta$ subsets of $\mathbb{R}$ is dense, and in particular non-empty. In Lemma 5.2 below we show that every $G_\delta$ subset of a complete metric space is completely metrizable. Our goal is therefore to find open sets $G_1, G_2, \dots$ of $\widetilde{X}$ with $X = \bigcap_n G_n$. In the proof of Proposition 4.8, a winning strategy came from a representation of this kind, that of a Čech-complete space $X$ as a $G_\delta$ subset of $\beta X$. Here the roles are exchanged: the strategy must produce the representation.

**The open sets of the strategy.** The strategy $\tau$ of Player II is a machine that answers each point with an open set that contains the point. At the first stage, Player I may play any point $x$ of $X$ (with the open set $X$), and the strategy answers with an open set $V \ni x$. These sets cover $X$; once each of them is enlarged slightly into $\widetilde{X}$, their union is an open subset $G_1$ of $\widetilde{X}$ that contains $X$. At the second stage, Player I continues each of these positions with each point of the set just played, and so on. The open set $G_n$ is roughly the union of the (enlarged) sets played by Player II at stage $n$ over all positions, and $G_n$ therefore contains $X$. Suppose now that a point $y$ of the completion lies in every $G_n$. If the sets containing $y$ at the various stages all came from *one* play, then the play would have a common point, because Player II wins the play. The sets will be chosen with diameters tending to 0, so this common point would be $y$; in particular, the point $y$ would belong to $X$.

**The difficulty.** At each stage, however, the point $y$ may lie in sets that come from many different positions. A choice of one position at each stage need not form a single play, which requires that the position chosen at stage $n + 1$ extend the one chosen at stage $n$. If at every stage only *finitely* many positions are involved, then a compactness argument, Kőnig's lemma (see Lemma 5.4 below), makes the choices fit together. The finiteness is supplied by a theorem of A. H. Stone (see Lemma 5.3 below): every open cover of a metric space has an open refinement, still covering the space, in which each point has a neighborhood that meets only finitely many members. The two ingredients combine as shown in Figure 13.

5.1. **A preliminary lemma and three classical tools.** In this subsection we prove a lemma on extending open sets and the three classical tools of the proof: Alexandrov's theorem, A. H. Stone's theorem and Kőnig's lemma.

Throughout this subsection we fix a metric space $(X, d)$ and a metric space $(\widetilde{X}, \tilde{d})$ that contains $X$ as a dense subset, with $\tilde{d}$ extending $d$. Closures and diameters of subsets of $\widetilde{X}$ are taken in $\widetilde{X}$. In the preliminary lemma we extend each open set of $X$ to an open set of $\widetilde{X}$ with the same points of $X$ and the same diameter (see Figure 14). For an open set $V \subseteq X$, we write $\widetilde{V} = \widetilde{X} \setminus \overline{X \setminus V}$.

**Lemma 5.1.** *Let $(X, d)$ and $(\widetilde{X}, \tilde{d})$ be as fixed at the beginning of this subsection, and let $V$ and $V'$ be open subsets of $X$. Then the sets $\widetilde{V}$ and $\widetilde{V'}$ satisfy the following:*

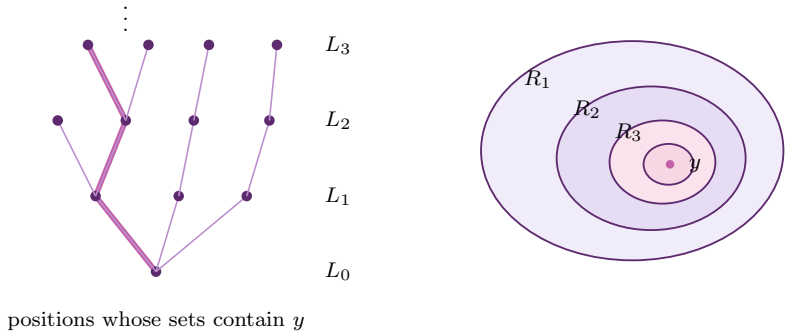



FIGURE 13. The end of the proof of the hard half of Theorem 4.4: the finite levels $L_n$ of pairs whose sets contain $y$, with a branch given by Kőnig's lemma in pink (left), and the sets $R_n$ along that branch shrinking to $y$ (right).

(1) *the set* $\widetilde{V}$ *is open in* $\widetilde{X}$*, and* $\widetilde{V} \cap X = V$*;*
(2) *if* $V \subseteq V'$*, then* $\widetilde{V} \subseteq \widetilde{V'}$*;*
(3) *the diameters satisfy* $\operatorname{diam} \widetilde{V} = \operatorname{diam} V$*.*

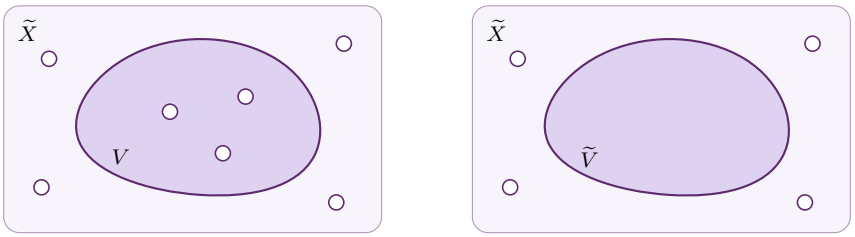



FIGURE 14. The extension $\widetilde{V}$ in $\widetilde{X}$ (right) of an open set $V$ of $X$ (left), as in Lemma 5.1, with the points of $\widetilde{X}$ that are not in $X$ drawn as white dots.

*Proof.* (1) The set $\widetilde{V}$ is the complement of a closed set, hence open. Let $x \in V$ be given. Since $V$ is open in $X$, there is $\varepsilon > 0$ such that the open ball $B(x, \varepsilon)$ of $X$ lies in $V$. The ball of radius $\varepsilon$ about $x$ in $\widetilde{X}$ meets $X$ exactly in $B(x, \varepsilon)$, and this ball therefore misses $X \setminus V$, which gives $x \in \widetilde{V}$. If $x \in X \setminus V$, then $x \in \overline{X \setminus V}$, and $x \notin \widetilde{V}$. Thus $\widetilde{V} \cap X = V$.

(2) If $V \subseteq V'$, then $X \setminus V' \subseteq X \setminus V$, and the complements of the closures reverse the inclusion.

(3) We first show that $\widetilde{V} \subseteq \overline{V}$. Let $y \in \widetilde{V}$ and $\varepsilon > 0$ be given. The points of $\widetilde{V}$ at distance less than $\varepsilon$ from $y$ form a non-empty open set, which meets the dense set $X$; by (1), every point of $\widetilde{V} \cap X$ lies in $V$. Since $\varepsilon$ is arbitrary, every ball about $y$ meets $V$, that is, $y \in \overline{V}$. It follows that $\operatorname{diam} \widetilde{V} \leq \operatorname{diam} \overline{V} = \operatorname{diam} V$. The reverse inequality holds because $V \subseteq \widetilde{V}$. ☺

The first tool is Alexandrov's theorem, announced at the beginning of this section; it reduces our problem to finding a $G_\delta$ set.

**Lemma 5.2 (P. S. Alexandrov).** *Every* $G_\delta$ *subset of a complete metric space is completely metrizable.*

*Proof.* Let $(Y, \rho_0)$ be a complete metric space, and let $G = \bigcap_{n \geq 1} G_n$ with each $G_n$ open in $Y$. The idea is to add to $\rho_0$ terms that blow up near the complements of the sets $G_n$,

in order that no Cauchy sequence may approach those complements. We may assume that each closed set $F_n = Y \setminus G_n$ is non-empty, discarding the indices with $G_n = Y$ (if no index remains, then $G = Y$ and there is nothing to prove). For $x \in G$, the distance $\rho_0(x, F_n) = \inf\{\rho_0(x, z) : z \in F_n\}$ is positive, because $F_n$ is closed and $x \notin F_n$. The function $f_n$ given on $G$ by $f_n(x) = 1/\rho_0(x, F_n)$ is continuous, and for $x, y \in G$ we define

$$\rho(x, y) = \rho_0(x, y) + \sum_{n \geq 1} 2^{-n} \min\{1,\, |f_n(x) - f_n(y)|\}.$$

Each summand of $\rho$ is a pseudometric on $G$, the series converges, and the first summand $\rho_0$ is a metric; the sum $\rho$ is therefore a metric on $G$.

The metric $\rho$ is compatible with the topology of $G$. Indeed, convergence for $\rho$ implies convergence for $\rho_0$, because $\rho \geq \rho_0$. Conversely, let $x \in G$ be given, and suppose that a sequence $(x_k)_{k \geq 1}$ in $G$ converges to $x$ with respect to $\rho_0$. Each term of the series then tends to 0, by the continuity of $f_n$, and the $n$th term never exceeds $2^{-n}$; dominated convergence for series yields $\rho(x_k, x) \xrightarrow{k \to \infty} 0$.

The metric $\rho$ is complete. Let $(x_k)_{k \geq 1}$ be a $\rho$-Cauchy sequence in $G$. Since $\rho \geq \rho_0$, the sequence is $\rho_0$-Cauchy, and it converges to some $y \in Y$ because $\rho_0$ is complete. For each $n$, the sequence $(f_n(x_k))_k$ is Cauchy, because $\min\{1, |f_n(x_k) - f_n(x_l)|\} \leq 2^n \rho(x_k, x_l)$; the sequence is therefore bounded by some constant $c_n > 0$. Thus $\rho_0(x_k, F_n) \geq 1/c_n$ for every $k$, and in the limit $\rho_0(y, F_n) \geq 1/c_n > 0$. It follows that $y \in G_n$ for every $n$, that is, $y \in G$, and $\rho(x_k, y) \xrightarrow{k \to \infty} 0$ by the compatibility just proved. Consequently, the space $G$ carries the complete compatible metric $\rho$ and is completely metrizable. ☺

The second tool supplies the finiteness. A family $\{V_i : i \in I\}$ of subsets of a topological space $Y$ is *locally finite* if every point of $Y$ has a neighborhood that meets $V_i$ for only finitely many indices $i \in I$. A family $\mathcal{V}$ *refines* a family $\mathcal{U}$, and is then called a *refinement* of $\mathcal{U}$, if every member of $\mathcal{V}$ is contained in some member of $\mathcal{U}$.

**Lemma 5.3 (A. H. Stone).** *Every open cover of a metric space has a locally finite open refinement that covers the space.*

The idea of the proof is to cover the space by small balls, assigning each ball to the first member of the cover that contains the center of the ball. A ball is used only when its assigned member contains it with room to spare, and small balls are no longer used around points that larger balls already cover.

*Proof.* Let $(Y, \rho)$ be a metric space, and let $\{U_s : s \in \Lambda\}$ be an open cover of $Y$; we write $B(c, r)$ for the open ball of radius $r$ about $c$. By the well-ordering theorem we may assume that $\Lambda$ is well ordered. For $s \in \Lambda$ and $n \geq 1$, let $V_{s,n}$ be the union of the balls $B(c, 2^{-n})$ over all centers $c \in Y$ such that:

(i) the index $s$ is the least one with $c \in U_s$;
(ii) the center $c$ lies in no set $V_{t,j}$ with $t \in \Lambda$ and $j < n$;
(iii) the ball $B(c, 3 \cdot 2^{-n})$ is contained in $U_s$.

The definition is a recursion on $n$, because (ii) involves only the sets with smaller second index. Each $V_{s,n}$ is open, and $V_{s,n} \subseteq U_s$ by (iii).

*The sets $V_{s,n}$ cover $Y$.* Let $c \in Y$ be given, let $s$ be the least index with $c \in U_s$, and let $n$ be so large that $B(c, 3 \cdot 2^{-n}) \subseteq U_s$. Either the point $c$ lies in some $V_{t,j}$ with $j < n$, or it satisfies (i)–(iii), and then $c \in B(c, 2^{-n}) \subseteq V_{s,n}$.

*The family is locally finite.* Let $c \in Y$ be given, and let $s$ and $n$ be indices with $c \in V_{s,n}$; since $V_{s,n}$ is open, there is $j \geq 1$ with $B(c, 2^{-j}) \subseteq V_{s,n}$. We show that the ball $B(c, 2^{-n-j})$

meets no set $V_{t,i}$ with $i \geq n+j$, and at most one set $V_{t,i}$ for each $i < n+j$; the ball then meets at most $n+j-1$ sets of the family.

Suppose first that $i \geq n+j$, and let $c'$ be a center used for $V_{t,i}$. Since $n < i$, condition (ii) gives $c' \notin V_{s,n}$, so $\rho(c,c') \geq 2^{-j}$. Every point of $B(c', 2^{-i})$ is therefore at distance at least $2^{-j} - 2^{-i} \geq 2^{-j} - 2^{-n-j} \geq 2^{-n-j}$ from $c$, and the ball $B(c, 2^{-n-j})$ misses $V_{t,i}$.

Suppose next that $i < n+j$, and that $p \in V_{t,i}$ and $q \in V_{t',i}$ with $t < t'$. There are centers $c_1$ for $V_{t,i}$ and $c_2$ for $V_{t',i}$ with $\rho(p,c_1) < 2^{-i}$ and $\rho(q,c_2) < 2^{-i}$. Condition (iii) places the ball $B(c_1, 3\cdot 2^{-i})$ inside $U_t$, and condition (i) keeps $c_2$ out of $U_t$, because $t'$ is the least index with $c_2 \in U_{t'}$ and $t < t'$. Hence $\rho(c_1,c_2) \geq 3\cdot 2^{-i}$ and

$$\rho(p,q) \geq \rho(c_1,c_2) - \rho(p,c_1) - \rho(q,c_2) > 2^{-i} \geq 2\cdot 2^{-n-j}.$$

Two points of $B(c, 2^{-n-j})$ are at distance less than $2\cdot 2^{-n-j}$, so the ball cannot meet both $V_{t,i}$ and $V_{t',i}$. The sets $V_{s,n}$ therefore form a locally finite open refinement of $\{U_s : s \in \Lambda\}$ that covers $Y$. ☺

The third tool turns finitely many choices at each stage into one coherent choice.

**Lemma 5.4 (D. Kőnig).** *Let $L_0, L_1, \ldots$ be non-empty finite sets, and for each $n \geq 0$ let $\pi_n \colon L_{n+1} \to L_n$ be a function. Then there are $\ell_n \in L_n$, for $n \geq 0$, with $\pi_n(\ell_{n+1}) = \ell_n$ for every $n$.*

*Proof.* The idea is to follow, level by level, the elements that are reached from every later level. For $m \geq n$ we write $\pi_{n,m} \colon L_m \to L_n$ for the composite $\pi_n \circ \cdots \circ \pi_{m-1}$ (the identity when $m = n$). We call $\ell \in L_n$ *good* if for every $m \geq n$ some element of $L_m$ is mapped to $\ell$ by $\pi_{n,m}$. If $\ell = \pi_{n,m}(z)$ for some $z \in L_m$, then $\ell = \pi_{n,k}(\pi_{k,m}(z))$ for every $k$ with $n \leq k \leq m$. Hence an element of $L_n$ is good as soon as it lies in the image of $\pi_{n,m}$ for infinitely many $m$.

Some element of $L_0$ is good: every $L_m$ is non-empty, so each $m$ supplies an element of $L_0$ in the image of $\pi_{0,m}$. Since $L_0$ is finite, one element is supplied by infinitely many $m$. If $\ell \in L_n$ is good, then some element of the finite set $\pi_n^{-1}(\ell)$ is good: for each $m > n$, an element of $L_m$ mapped to $\ell$ passes through some element of $\pi_n^{-1}(\ell)$. Of the finitely many elements of $\pi_n^{-1}(\ell)$, one serves infinitely many $m$. We may therefore choose recursively a good element $\ell_0 \in L_0$ and, for each $n$, a good element $\ell_{n+1} \in \pi_n^{-1}(\ell_n)$. The resulting sequence satisfies $\pi_n(\ell_{n+1}) = \ell_n$ for every $n$, as required. ☺

### 5.2. **The proof.**

*Proof of the converse of Theorem 4.4.* Let $X$ be a non-empty metrizable space, and let $\tau$ be a winning strategy for Player II in $\mathrm{Ch}(X)$. Let $d$ be a compatible metric on $X$, and let $(\widetilde{X}, \tilde{d})$ be the completion of $(X,d)$, in which $X$ is dense. Following the plan described at the beginning of this section, we build from $\tau$ open sets $G_1, G_2, \ldots$ of $\widetilde{X}$, each containing $X$. We use Kőnig's lemma to show that the intersection of these open sets is exactly $X$. As in Lemma 5.1, for an open set $V \subseteq X$ we define $\widetilde{V} = \widetilde{X} \setminus \overline{X \setminus V}$, the closure being taken in $\widetilde{X}$. We write $B(x,r)$ for the open ball of radius $r$ about $x$ in $X$. For $n \geq 0$, a *position of stage* $n$ is a position of length $2n$ of $\mathrm{Ch}(X)$ consistent with $\tau$. A position of stage $n$ ends with the move $V_{n-1}$ of Player II when $n \geq 1$, and it is the empty position when $n = 0$.

*Step 1: the families $\mathcal{S}_n$.* By recursion on $n \geq 0$ we construct families $\mathcal{S}_n$ of pairs $(R,p)$, where $R$ is a non-empty open subset of $\widetilde{X}$ and $p$ is a position of stage $n$. The union of the sets $R$ with $(R,p) \in \mathcal{S}_n$ is denoted by $G_n$. We require the following:

(P1) if $n \geq 1$ and $(R,p) \in \mathcal{S}_n$, then $R \subseteq \widetilde{V}$, where $V$ is the last move of $p$, and $\operatorname{diam} R \leq 2^{-n}$;

(P2) the set $G_n$ contains $X$;
(P3) the family of sets $R$, $(R,p) \in \mathcal{S}_n$, is locally finite in $G_n$;
(P4) each $(R,p) \in \mathcal{S}_{n+1}$ has a designated *parent* $(R',p') \in \mathcal{S}_n$ such that $R \subseteq R'$ and $p$ extends $p'$ by two moves: Player I plays $(x, R' \cap X)$ for some point $x \in R' \cap X$, and Player II answers as $\tau$ prescribes.

Let $\mathcal{S}_0 = \{(\widetilde{X}, \langle\rangle)\}$; conditions (P2) and (P3) hold trivially, and (P1) is vacuous for $n = 0$.

Suppose that $\mathcal{S}_n$ has been constructed. Let $(R',p') \in \mathcal{S}_n$ and $x \in R' \cap X$ be given. The move $(x, R' \cap X)$ is legal for Player I at $p'$. Indeed, for $n = 0$ it is a first move, and for $n \geq 1$ condition (P1) and Lemma 5.1(1) give $R' \cap X \subseteq \widetilde{V} \cap X = V$, where $V$ is the last move of $p'$. Let $V_x$ be the answer of $\tau$, which satisfies $x \in V_x \subseteq R' \cap X$, and let $q_x$ be the resulting position of stage $n+1$. (The sets $V_x$, the positions $q_x$ and the sets $Q_x$ and $O_x$ below depend also on the pair $(R',p')$, which we suppress from the notation.) We define

$$Q_x = V_x \cap B(x, 2^{-n-2}) \qquad \text{and} \qquad O_x = R' \cap \widetilde{Q_x}.$$

The set $O_x$ is open in $\widetilde{X}$ and contained in $R'$. It contains $x$, because $x \in R'$ and $x \in Q_x \subseteq \widetilde{Q_x}$ by Lemma 5.1(1); it lies in $\widetilde{V_x}$, because $Q_x \subseteq V_x$ and Lemma 5.1(2) applies. Finally, the diameter of $O_x$ is at most $\operatorname{diam} \widetilde{Q_x} = \operatorname{diam} Q_x \leq 2^{-n-1}$, by Lemma 5.1(3) and because $Q_x$ lies in a ball of radius $2^{-n-2}$. By (P2), every point $x$ of $X$ lies in some $R'$ with $(R',p') \in \mathcal{S}_n$, and then $x \in O_x$. Hence the sets $O_x$, over all pairs $(R',p') \in \mathcal{S}_n$ and all $x \in R' \cap X$, cover $X$. Their union $H$ is open in $\widetilde{X}$, and the sets $O_x$ form an open cover of the metric space $H$. By Lemma 5.3, this cover has a locally finite open refinement $\mathcal{R}$ that covers $H$; we discard its empty members. For each $R \in \mathcal{R}$ we choose a set $O_x$ that contains $R$, coming from a pair $(R',p')$ and a point $x$, and we place $(R, q_x)$ in $\mathcal{S}_{n+1}$, with designated parent $(R',p')$.

We now verify that conditions (P1)–(P4) hold for $n+1$. Each pair $(R,q_x)$ in $\mathcal{S}_{n+1}$ satisfies (P1), because $R \subseteq O_x \subseteq \widetilde{V_x}$, where $V_x$ is the last move of $q_x$, and $\operatorname{diam} R \leq \operatorname{diam} O_x \leq 2^{-n-1}$. The union $G_{n+1}$ equals $H$, which contains $X$, as (P2) demands, and condition (P3) holds because the family $\mathcal{R}$ is locally finite in $H = G_{n+1}$. Finally, the pair $(R,q_x)$ and its designated parent $(R',p')$ satisfy (P4), since $R \subseteq O_x \subseteq R'$ and $q_x$ extends $p'$ by the moves $(x, R' \cap X)$ and $V_x$.

*Step 2: a $G_\delta$ set containing $X$.* The set $G = \bigcap_{n \geq 1} G_n$ is a $G_\delta$ subset of $\widetilde{X}$, and $X \subseteq G$ by (P2). It remains to be shown that $G \subseteq X$.

*Step 3: every point of $G$ is the common point of a play.* Let $y \in G$ be given. For $n \geq 0$, let $L_n$ be the set of pairs $(R,p) \in \mathcal{S}_n$ with $y \in R$. Each $L_n$ is non-empty, because $y \in G_n$ (and $L_0$ consists of the single pair $(\widetilde{X}, \langle\rangle)$). Each $L_n$ is also finite, by (P3): since $y \in G_n$, some neighborhood of $y$ meets only finitely many of the sets $R$ with $(R,p) \in \mathcal{S}_n$, among them every one that contains $y$. By (P4), the parent $(R',p')$ of a member $(R,p)$ of $L_{n+1}$ satisfies $y \in R \subseteq R'$, so it lies in $L_n$; let $\pi_n \colon L_{n+1} \to L_n$ be the parent map. By Lemma 5.4, there are $(R_n, p_n) \in L_n$ such that $(R_n, p_n)$ is the parent of $(R_{n+1}, p_{n+1})$ for every $n$.

By (P4), for each $n \geq 0$ the position $p_{n+1}$ extends $p_n$ by a move $(x_n, R_n \cap X)$ of Player I, for some point $x_n \in R_n \cap X$, and by the answer of $\tau$. Hence the positions $p_n$ fit together into a single play of $\mathrm{Ch}(X)$, consistent with $\tau$, in which Player I plays the sets $R_n \cap X$. Since $\tau$ is winning, there is a point $z \in \bigcap_n (R_n \cap X)$. For every $n \geq 1$, the points $y$ and $z$ both lie in $R_n$, which has diameter at most $2^{-n}$ by (P1). Hence $\tilde{d}(y,z) = 0$, and $y = z \in X$ (see Figure 13).

*Step 4: conclusion.* By Steps 2 and 3, we have $X = G$, a $G_\delta$ subset of the complete metric space $\widetilde{X}$. Since $\tilde{d}$ extends $d$, the topology that $X$ inherits from $\widetilde{X}$ is its own. Alexandrov's theorem (see Lemma 5.2) now shows that the space $X$ is completely metrizable. ☺

**Remark 5.5.** The points played by Player I are essential to the proof. In Step 1, Player I plays every point $x$ of every set $R' \cap X$, and Player II must answer with a set that contains $x$, whence each $G_n$ contains all of $X$. Without the points, the converse of Theorem 4.4 fails (see Remark 4.6). For separable spaces, Kechris gives a different proof [Kec95, Section 8].

**Exercises.**

(1) Let $X$ be the space $\mathbb{R} \setminus \{0\}$ with the usual metric, whose completion is $\widetilde{X} = \mathbb{R}$. Compute the set $\widetilde{V} = \widetilde{X} \setminus \overline{X \setminus V}$, with the closure taken in $\widetilde{X}$, for $V = (-1, 0) \cup (0, 1)$ and for $V = (0, 1)$, and check the three statements of Lemma 5.1 in both cases.

(2) Show that the open cover $\{(-n, n) : n \geq 1\}$ of $\mathbb{R}$ is not locally finite, and find a locally finite open refinement of it that covers $\mathbb{R}$, as promised by Lemma 5.3.

(3) Apply the proof of Lemma 5.2 to $G = (0, 1)$, a single open subset of $Y = \mathbb{R}$ with the usual metric $\rho_0$, and write down the metric $\rho$ explicitly. Show that the sequence $(1/k)_{k\geq 2}$ is Cauchy for the usual metric but not for $\rho$.

(4) Show that the conclusion of Lemma 5.4 can fail if the sets $L_n$ are allowed to be infinite. (Hint: let $L_0$ have one element, and for $n \geq 1$ let $L_n$ consist of the pairs $(i, n)$ of integers with $i \geq n$. Let $\pi_0$ be constant, and define $\pi_n$ for $n \geq 1$ by $\pi_n(i, n + 1) = (i, n)$.)

(5) $(*)$ Let $(X, d)$ be a metric space whose topology is completely metrizable, possibly by a metric other than $d$. Use Theorem 4.4 and the proof of its converse in Section 5, with the metric $d$, to show that $X$ is a $G_\delta$ subset of the completion of $(X, d)$. Deduce that $\mathbb{Q}$ is not completely metrizable.

## 6. Summary: from complete metrizability to Baire category

The main results of this part fit into one chain of implications, which runs from complete metrizability, through the two games, down to Baire category.

For a non-empty metrizable space $X$, with $\mathrm{Ch}(X)$ the strong Choquet game and $\mathrm{BM}(X)$ the Banach–Mazur game on $X$, the chain is the following:

$X$ is completely metrizable

$\Updownarrow$ Theorem 4.4

Player II has a winning strategy in $\mathrm{Ch}(X)$

$\Downarrow$ Lemma 4.2

Player II has a winning strategy in $\mathrm{BM}(X)$

$\Downarrow$ Remarks 2.3(3)

Player I has no winning strategy in $\mathrm{BM}(X)$

$\Updownarrow$ Theorem 3.3

$X$ is a Baire space

$\Updownarrow$ Theorem 1.3

no non-empty open subset of $X$ is of the first category.

Followed from top to bottom, the chain yields the Baire category theorem (see Corollary 4.5). The implication given by Lemma 4.2 cannot be reversed, as the space of Remark 4.6 shows. This space, the open upper half plane together with the rational points of the horizontal axis, is metrizable, and Player II wins its Banach–Mazur game. The space is not, however, completely metrizable, and Player II therefore has no winning strategy in its strong Choquet game. The games $\mathrm{BM}(X, S)$ with a target set $S \subseteq X$, in which Player II wins when the

intersection of the moves lies in $S$, measure the size of sets inside $X$. By Theorem 3.5, Player II has a winning strategy in $\mathrm{BM}(X, S)$ exactly when $S$ is comeager in $X$. When $X$ is a complete metric space, the same theorem shows that Player I has one exactly when $U \cap S$ is of the first category in $X$ for some non-empty open set $U$.

## 7. Notes and historical remarks

7.1. **Baire category.** The characterization of Baire spaces in Theorem 1.3 is an elementary reformulation of their definition and carries no attribution. It is not to be confused with the *Banach category theorem* [Ban30], which is deeper: in any topological space, the union of an arbitrary family of open sets of the first category is again of the first category. The title of Oxtoby's paper [Oxt57] refers to this theorem.

7.2. **The Banach–Mazur and strong Choquet games.** As told in the Introduction, the Banach–Mazur game is Problem 43 of the Scottish Book, posed by S. Mazur in 1935 [Mau81] (see also Section 31). In Mazur's formulation, two players alternately choose nested intervals of the real line, and the first player wins if their intersection meets a given set $E$. Mazur observed two facts: the first player wins if $E$ is comeager in some interval, and the second player wins if $E$ is of the first category. He conjectured the converses. On 4 August 1935 Banach recorded in the Scottish Book that Mazur's conjecture is true. Banach's proof was never published. The first published proof is Oxtoby's [Oxt57]. Oxtoby showed that in an arbitrary topological space the second player has a winning strategy exactly when $E$ is of the first category. The other converse, that the first player wins exactly when $E$ is comeager in some non-empty open set, requires further hypotheses on the space; it holds, for instance, in complete metric spaces. Together, Oxtoby's theorem and this converse become the two parts of Theorem 3.5 under the following dictionary:

(1) Mazur's first player is our Player I, and Mazur's second player is our Player II.
(2) Mazur's set $E$, at which the first player aims, is the complement $X \setminus S$ of the set $S$ at which Player II aims in the game $\mathrm{BM}(X, S)$ on a space $X$.
(3) "The intersection meets $E$" becomes "the intersection is not contained in $S$", so Mazur's first player wins exactly when our Player I wins.

Whether the moves are open or closed intervals makes no difference to the proofs. The game $\mathrm{BM}(X)$, in which Player II needs only a non-empty intersection, differs from $\mathrm{BM}(X, X)$, which Player II wins trivially. The proof of Theorem 3.3 is Oxtoby's argument adapted to $\mathrm{BM}(X)$. Telgársky's survey [Tel87], written for the fiftieth anniversary of the Banach–Mazur game, covers its later history.

The strong Choquet game goes back to Choquet's work on Baire spaces [Cho58]. Choquet expounds the game, together with Theorem 4.4, in his *Lectures on Analysis* [Cho69], where the players are called $\beta$ and $\alpha$. Kechris states Theorem 4.4 for separable metrizable spaces, where it characterizes the Polish spaces, that is, the separable completely metrizable spaces [Kec95, Section 8]. The proof in Section 5, for arbitrary metrizable spaces, rests on three classical results. The first, Lemma 5.2, is due to Alexandrov [Ale24]. The second, Lemma 5.3, is A. H. Stone's theorem that metric spaces are paracompact [Sto48]; the short proof given here is M. E. Rudin's [Rud69]. The third, Lemma 5.4, is Kőnig's infinity lemma [Kőn27], first stated for graphs.

7.3. **Čech-complete spaces.** Čech introduced the spaces now named after him in 1937 [Čec37], in the paper in which he constructed the Stone–Čech compactification independently of M. H. Stone. He called them topologically complete, and he proved

that a metrizable space is topologically complete exactly when it is completely metrizable. The result of Proposition 4.8, that every Čech-complete space is strongly $\alpha$-favorable, is folklore. In his survey [Tel87], Telgársky also treats the many variants of the strong Choquet game and the completeness properties that they describe. Michael introduced the space of Example 4.9 in 1963, as a paracompact space whose product with the irrationals is not normal [Mic63].

Part
# II

# Applications of Banach–Mazur and Choquet games

In Part I we showed that the Banach–Mazur game measures size: a set is large in the sense of Baire category exactly when Player II can steer the play into the set (see Theorem 3.5). In this part that measure is employed outside topology. In one form or another, each problem in this part asks for a proof that a set is large, or that an object exists when no formula for the object is in sight. In the last section the question is turned around, and the game explains why an object that the axiom of choice provides cannot be written down. In each case the answer is a game obtained by changing a rule or two, and with each new rule the game measures something new. The rules so changed are those of the Banach–Mazur game of Section 3 or of a game of Choquet type in which the players shrink sets (see Section 4). Through all these changes the two features of games described in the Introduction persist. The first is that a winning strategy is built one move at a time; the second, that a player can pursue several goals at once by taking turns between them.

**Roadmap**

The sections run from the concrete to the set-theoretic. The part opens in number theory with Schmidt's game, in which the sizes of the moves are prescribed. This game shows that the badly approximable numbers form a set that is large in dimension and survives countable intersections (see Section 8). Transported to the space of lattices, the same game becomes a game about bounded orbits of flows (see Section 9). A more demanding winning condition yields a class of spaces, the strongly Baire spaces, on which a separately continuous group multiplication is always jointly continuous (see Section 10). In the Ellentuck topology, which is finer than the topology that the infinite subsets of the natural numbers inherit from the Cantor space, a game of Banach–Mazur type characterizes the families of infinite sets that are Ramsey in a hereditary sense (see Section 11). In Banach spaces one game answers arbitrary sets with relatively weak*-open ones, and another answers block subspaces with vectors; the first detects Asplund spaces, and the second drives Gowers's dichotomy (see Section 12). The part ends with the nonprincipal ultrafilters on the natural numbers (see Section 13).

The sections can be read in any order, with a few dependencies. The section on homogeneous dynamics builds on the proof in Section 8, and the sections on Banach spaces and on ultrafilters both draw on the Ramsey theory of Section 11. The Ramsey theory in turn borrows the definition of an ultrafilter from the beginning of Section 13. The sections also differ in the background expected of the reader. The sections on Diophantine approximation and ultrafilters need little beyond Part I. In the section on ultrafilters we give complete proofs of the main points; in the section on Diophantine approximation we prove in full that **Bad**(1) is winning and that winning sets survive countable intersections, and we sketch the rest. The sections on homogeneous dynamics and Banach spaces reach into Lie groups and functional analysis; each of them opens with a short glossary. The heart of the section on homogeneous dynamics is a computation with $2 \times 2$ matrices that we do by hand. In the sections on topological groups and Ramsey theory we sketch the main arguments rather than prove every step. Subsections marked *optional* may be skipped on a first reading. The games of this part are set side by side in Table 1.

| **Section** | **Moves** | **Player II wins if** | **Main result** |
|---|---|---|---|
| 8. Diophantine approximation | closed balls whose radii shrink by two fixed factors, one for each player | the limit point lies in the target set | Theorem 8.1 |
| 9. Homogeneous dynamics | Schmidt's game on an expanding leaf of a flow on a space of lattices | the orbit of the limit point is bounded | Theorem 9.1 |
| 10. Topological groups | open sets, as in the Banach–Mazur game | the intersection is non-empty, and every sequence of points of a fixed dense set, one in each of her moves, clusters | Theorem 10.1 |
| 11. Ramsey theory | Ellentuck basic sets; Player II adds one element at each move | the set produced lies in a given family of infinite sets | Theorem 11.1 |
| 12. Banach spaces | arbitrary sets against relatively weak$^*$-open sets; block subspaces against vectors | the sets shrink to a point in norm, or have empty intersection; the sequence of vectors lies in a given set of sequences | Theorems 12.1 and 12.2 |
| 13. Ultrafilters | finite binary strings; sets in a given ultrafilter against points or finite sets | the sequence produced lies in the target set; the set produced lies in the ultrafilter | Theorem 13.1 |

Table 1. The games of Part II, in each of which Player I moves first and the payoff set belongs to Player II.

## 8. Applications to Diophantine approximation

**The problem.** Every real number can be approximated by fractions, but some numbers are much harder to approximate than others. Let $n$ be a positive integer, and for $t \in \mathbb{R}$ let $\|t\|$ be the distance from $t$ to the nearest integer. A vector $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ is *badly approximable* if there is a constant $c = c(\mathbf{x}) > 0$ such that

$$\max_{1 \le i \le n} \|q x_i\| \;\ge\; c\, q^{-1/n} \qquad \text{for every integer } q \ge 1.$$

For $n = 1$, writing $x$ for $x_1$, the condition becomes $|x - p/q| \ge c/q^2$ for every fraction $p/q$ with $p \in \mathbb{Z}$ and $q \ge 1$. Let $\mathbf{Bad}(n)$ be the set of badly approximable vectors of $\mathbb{R}^n$. By Dirichlet's theorem (1842), every $\mathbf{x} \in \mathbb{R}^n$ satisfies $\max_i \|q x_i\| < q^{-1/n}$ for infinitely many integers $q \ge 1$. Badly approximable vectors are those for which this bound cannot be improved by more than a constant factor: they are the most stubborn vectors there are. How big is $\mathbf{Bad}(n)$? In particular, is the intersection of countably many translates of $\mathbf{Bad}(n)$ still big?

8.1. **History of the problem and of the solution using topological games.** For $n = 1$ there is a classical description: an irrational number is badly approximable exactly when the partial quotients in the continued fraction expansion of the number are bounded. As Hurwitz showed in 1891, the golden ratio $(1 + \sqrt{5})/2$ is as badly approximable as a number can be. The continued fraction expansion $1 + 1/(1 + 1/(1 + \cdots))$ of the golden ratio has every partial quotient equal to 1, the smallest possible value. Badly approximable numbers are at once rare and common. Borel showed in 1909, by an argument that Bernstein completed in 1912, that $\mathbf{Bad}(1)$ has Lebesgue measure zero, so a number chosen at random is almost surely *not* badly approximable. Jarník proved in 1928 that $\mathbf{Bad}(1)$ has Hausdorff dimension 1, the largest possible. The set $\mathbf{Bad}(1)$ is thus tiny for measure and huge for dimension.

Neither notion helps with intersections: measure sees only that $\mathbf{Bad}(1)$ is a null set, and two sets of full Hausdorff dimension can be disjoint. The set $\mathbf{Bad}(1)$ and the complement of $\mathbf{Bad}(1)$ are an example of such a pair, since the complement has full measure and hence full dimension. A question as innocent as whether there is a number $x$ such that $x + \sqrt{p}$ is badly approximable for every prime $p$ was therefore out of reach.

Wolfgang Schmidt solved the problem in 1966 [Sch66]. To do so he invented a new version of the Banach–Mazur game, now called Schmidt's game. He proved that the winning sets of the game have full Hausdorff dimension, that countable intersections remain winning when the sets share a winning parameter, and that $\mathbf{Bad}(1)$ is winning. In 1969 he extended the last result to $\mathbf{Bad}(n)$, for every $n$ [Sch69].

The game had more to give. For real weights $i, j > 0$ with $i + j = 1$, a vector $(x, y) \in \mathbb{R}^2$ is badly approximable with weights $(i, j)$ if there is $c > 0$ with $\max(\|qx\|^{1/i}, \|qy\|^{1/j}) \ge c/q$ for every integer $q \ge 1$. Let $\mathbf{Bad}(i, j)$ be the set of vectors badly approximable with weights $(i, j)$. In 1983 Schmidt conjectured that $\mathbf{Bad}(\frac{1}{3}, \frac{2}{3})$ and $\mathbf{Bad}(\frac{2}{3}, \frac{1}{3})$ have a point in common. He observed that a counterexample to his conjecture would prove the Littlewood conjecture, which asserts that $\liminf_{q \to \infty} q\, \|qx\|\, \|qy\| = 0$ for all real numbers $x$ and $y$. Badziahin, Pollington and Velani proved Schmidt's conjecture in 2011 [BPV11]. In 2016 An showed that the sets $\mathbf{Bad}(i, j)$ are all winning for Schmidt's game, with one and the same parameter for every pair of weights [An16]. The intersection property of winning sets then yields Schmidt's conjecture in a stronger form, namely that every countable intersection of sets $\mathbf{Bad}(i, j)$ has full dimension.

*Recognition.* Schmidt received the Cole Prize in Number Theory in 1972. He addressed the International Congress of Mathematicians three times: in Nice in 1970, in Vancouver in 1974 (a plenary address) and in Warsaw in 1983. The paper of Badziahin, Pollington and Velani that settled Schmidt's conjecture appeared in the *Annals of Mathematics*.

8.2. **The game.** Schmidt's game is the Banach–Mazur game of Section 3 with one new rule: the sizes of the moves are prescribed in advance. Let $(X, d)$ be a complete metric space and let $S \subseteq X$ be given; fix also two numbers $\alpha, \beta \in (0,1)$, which serve as the shrinking factors of Player II and Player I, respectively. The game is played with closed balls $B[x,r] = \{y \in X : d(x,y) \le r\}$ of prescribed radii, where $x \in X$ and $r > 0$. (In a general metric space a ball, as a set, need not determine the center and radius of the ball, and Schmidt formulates the game with the pairs $(x,r)$. The distinction disappears in $\mathbb{R}^n$, the only space used below.) Schmidt called the players Black and White: Black, who moves first, is our Player I and White is our Player II. In Choquet's terminology our Player I is $\beta$ and our Player II is $\alpha$; in Choquet's games, as in Schmidt's, the letter $\alpha$ thus belongs to Player II. The resemblance ends there: here $\alpha$ and $\beta$ are shrinking factors, and the $\alpha$-winning sets defined below have nothing to do with the $\alpha$-favorable spaces of Section 4.

(1) Player I chooses a closed ball $U_0 = B[x_0, r_0]$ with any radius $r_0 > 0$ he likes.
(2) Player II chooses a closed ball $V_0 \subseteq U_0$ of radius $\alpha r_0$.
(3) Player I chooses a closed ball $U_1 \subseteq V_0$ of radius $\beta\alpha r_0$.
(4) The players continue in this way, and the play produces nested balls $U_0 \supseteq V_0 \supseteq U_1 \supseteq V_1 \supseteq \cdots$, where for every $k \ge 0$ the ball $U_k$ has radius $r_k = (\alpha\beta)^k r_0$ and the ball $V_k$ has radius $\alpha r_k$.

Since the radii tend to 0 and the space is complete, the balls close down on a single point $x_\infty$ (as in the proof of Theorem 4.4). We call this game, in which Player II wins if $x_\infty \in S$, the $(\alpha,\beta)$-*game for* $S$. The set $S$ is $\alpha$-*winning* if Player II has a winning strategy in the $(\alpha,\beta)$-game for $S$ for every $\beta \in (0,1)$. The order of the quantifiers matters. Player II declares her handicap $\alpha$ first, promising to shrink by the factor $\alpha$ every time, and only then does Player I pick $\beta$, as aggressive as he likes. Her strategy may depend on $\beta$, but the number $\alpha$ is fixed once and for all. This order of quantifiers is what lets $\alpha$-winning sets survive countable intersections (see the proof of Theorem 8.1(2) below). On the real line the balls are just closed intervals, and we speak of lengths instead of radii (see Figure 15).

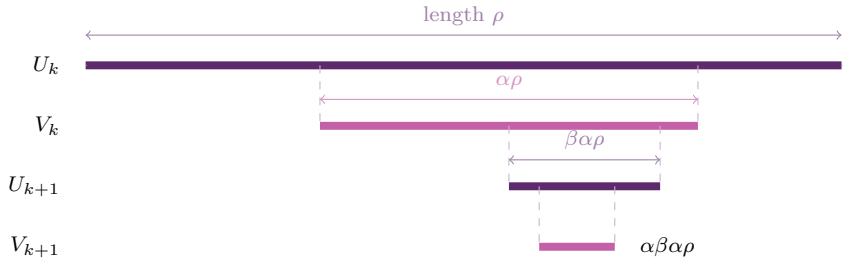



FIGURE 15. Rounds $k$ and $k+1$ of Schmidt's $(\alpha,\beta)$-game on the line, drawn with $\alpha = \frac{1}{2}$ and $\beta = \frac{2}{5}$.

Because the sizes of the moves are prescribed, neither player can make the balls arbitrarily small in a single move. This small change of rule lets the game detect Hausdorff dimension (see the proof of Theorem 8.1(1) below). The game also parts company with category. The

set **Bad**(1) is $\frac{1}{2}$-winning (see Theorem 8.1(3) below), and yet **Bad**(1) is meager. Indeed, for each $c > 0$, the numbers $x$ with $|x - p/q| \geq c/q^2$ for every fraction $p/q$ form a closed set that contains no rational number, hence a nowhere dense set. The set **Bad**(1) is the union of these nowhere dense sets over $c = 1, \frac{1}{2}, \frac{1}{3}, \ldots$.

8.3. **The main result of this section.** Schmidt proved three things about $\alpha$-winning sets: they are large, they survive countable intersections, and the badly approximable vectors form one of them.

**Theorem 8.1 (Schmidt, 1966 and 1969).** *Let $n$ be a positive integer.*

(1) *If $0 < \alpha < 1$ and $S \subseteq \mathbb{R}^n$ is $\alpha$-winning, then the intersection of $S$ with every non-empty open set $U \subseteq \mathbb{R}^n$ has Hausdorff dimension $n$.*
(2) *For every $\alpha \in (0,1)$, the intersection of countably many $\alpha$-winning subsets of $\mathbb{R}^n$ is $\alpha$-winning.*
(3) *The set **Bad**(1) is $\frac{1}{2}$-winning, and there is $\alpha \in (0,1)$ such that **Bad**$(n)$ is $\alpha$-winning.*

Together, the three parts answer both questions of the problem. Let $\alpha \in (0,1)$ be a number for which **Bad**$(n)$ is $\alpha$-winning, as in part (3). A translate of an $\alpha$-winning set is $\alpha$-winning (see Exercise 2 at the end of this section). By part (2), the set $\bigcap_k (\mathbf{Bad}(n) + \mathbf{t}_k)$ is therefore $\alpha$-winning for every sequence $(\mathbf{t}_k)_{k\geq 1}$ of vectors of $\mathbb{R}^n$, and by part (1) this intersection has full dimension in every non-empty open set. In particular, every non-empty open interval contains a set of Hausdorff dimension 1 of numbers $x$ such that $x + \sqrt{p}$ is badly approximable for every prime $p$.

8.4. **Solving the problem with the game.** We go through the three parts of Theorem 8.1, starting with the one that is pure number theory. For simplicity we prove that **Bad**(1) is $\alpha$-winning for every $\alpha < \frac{1}{2}$; Schmidt's sharper bookkeeping also handles $\alpha = \frac{1}{2}$ [Sch66].

**Badly approximable numbers: the plan.** Fix $\alpha \in (0, \frac{1}{2})$ and $\beta \in (0,1)$, and consider the $(\alpha,\beta)$-game for **Bad**(1) on the real line, in which Player I opens with an interval $U_0$ of length $\rho_0$. We write $R$ for the number $1/(\alpha\beta)$, which is greater than 1. Every round shrinks the intervals by the factor $R$, and for every $k \geq 0$ the interval $U_k$ has length $\rho_k = \rho_0 R^{-k}$. We will find a constant $c > 0$ and a strategy for Player II such that the final point $x$ satisfies

$$\left| x - \frac{p}{q} \right| \;\geq\; \frac{c}{q^2} \qquad \text{for every fraction } p/q \text{ with } p \in \mathbb{Z} \text{ and } q \geq 1.$$

It suffices to consider fractions in lowest terms: if $p/q = p'/q'$ with $q \leq q'$, the inequality for $p/q$ implies the one for $p'/q'$, because $c/q^2 \geq c/q'^2$. From now on every fraction is in lowest terms. We call the interval $\left(\frac{p}{q} - \frac{c}{q^2}, \frac{p}{q} + \frac{c}{q^2}\right)$ the *no-go zone* of $p/q$. Player II's job is to keep the final point out of every no-go zone. There are infinitely many zones, but she has infinitely many moves, and the proof consists in giving each move a manageable amount of work.

**Idea 1: fractions with small denominators keep their distance.** If $p/q \neq p'/q'$, then

$$\left| \frac{p}{q} - \frac{p'}{q'} \right| = \frac{|pq' - p'q|}{qq'} \;\geq\; \frac{1}{qq'},$$

because the numerator $pq' - p'q$ is a non-zero integer. Two fractions whose denominators are both small therefore cannot be close together, and this little inequality drives the argument.

**Idea 2: sort the fractions into generations.** Let $A = 1/(3\rho_0)$. We say that $p/q$ belongs to *generation* 0 if $q^2 < A$, and to *generation* $k \geq 1$ if

$$AR^{k-1} \;\leq\; q^2 \;<\; AR^k.$$

Every fraction belongs to exactly one generation, and generation $k$ is the one that Player II will deal with in round $k$. Grouping the denominators by size in this way keeps the fractions of a single generation far apart. Two different fractions $p/q$ and $p'/q'$ of generation $k$ satisfy $qq' < AR^k$, and by Idea 1 they therefore lie at distance more than $1/(AR^k) = 3\rho_0 R^{-k} = 3\rho_k$. If we fatten Player I's interval $U_k$ by $\rho_k$ on each side, we obtain an interval of length $3\rho_k$, which can contain *at most one* fraction of generation $k$. This fraction, if there is one, is the *dangerous fraction* of round $k$.

**Idea 3: the no-go zones are small.** A fraction $p/q$ of generation $k \geq 1$ has $q^2 \geq AR^{k-1}$, so the no-go zone of $p/q$ has length

$$\frac{2c}{q^2} \;\leq\; \frac{2c}{AR^{k-1}} \;=\; 6cR\,\rho_k.$$

We choose $c > 0$ so small that $6cR \leq 1 - 2\alpha$ and $2c \leq (1-2\alpha)\rho_0$ (the second condition takes care of generation 0, whose zones have length at most $2c$). Then every zone of generation $k$ has length at most $(1-2\alpha)\rho_k$, which is less than $\rho_k$. In particular, a zone of generation $k$ that reaches into $U_k$ has the center of the zone within $\rho_k/2$ of $U_k$. Such a zone therefore belongs to a fraction in the fattened interval, that is, to the dangerous fraction.

**The strategy.** In round $k$, Player II looks at Player I's interval $U_k$ and at the no-go zone of the dangerous fraction of round $k$, if there is one. Removing the zone from $U_k$ leaves at most two closed intervals, of total length at least $\rho_k - (1-2\alpha)\rho_k = 2\alpha\rho_k$, and one of them therefore has length at least $\alpha\rho_k$. Player II plays her interval $V_k$ inside that one; if there is no dangerous fraction, she plays any interval of the prescribed length inside $U_k$. These two instructions constitute the whole strategy.

**Proof that the strategy wins.** Let $p/q$ be a fraction, and let $k$ be the generation of $p/q$. The final point $x$ lies in $V_k \subseteq U_k$. If the zone of $p/q$ reaches into $U_k$, the fraction is the dangerous one of round $k$; the interval $V_k$ avoids the zone of $p/q$, and so does the point $x$. Otherwise the zone misses $U_k$ altogether, and with it the point $x$. Either way the inequality $|x - p/q| \geq c/q^2$ holds, whence $x \in \mathbf{Bad}(1)$. The constant $c$ depends on $\alpha$, $\beta$ and $\rho_0$, and this dependence is harmless, since the definition of $\mathbf{Bad}(1)$ allows each number a constant of its own.

The game has done exactly what we promised in the Introduction: it has turned a global question into a local one. The condition "$|x - p/q| \geq c/q^2$ for all fractions" involves infinitely many fractions at once; the game splits it into infinitely many small jobs, each of them easy (dodge one interval).

**Higher dimensions.** For $\mathbf{Bad}(n)$ the same plan works, with Idea 1 replaced by the *simplex lemma*. This lemma asserts that if rational points $\mathbf{p}_0/q_0, \ldots, \mathbf{p}_n/q_n$ of $\mathbb{R}^n$ (with $\mathbf{p}_i \in \mathbb{Z}^n$ and $q_i \geq 1$) do not lie on a common hyperplane, then the simplex they span has volume at least $1/(n!\,q_0 \cdots q_n)$. Indeed, the volume is $1/n!$ times the absolute value of the determinant of the $(n+1) \times (n+1)$ matrix whose $i$-th row is $(1, \mathbf{p}_i/q_i)$. Multiplying the $i$-th row by $q_i$ produces an integer matrix, and a non-zero determinant of this kind is therefore at least $1/(q_0 \cdots q_n)$ in absolute value, exactly as in Idea 1. The generations are chosen as in Idea 2, but by the size of $q^{1+1/n}$ instead of $q^2$, since the no-go zone of $\mathbf{p}/q$ now has radius $c\,q^{-1-1/n}$ in the maximum norm. With the generations so chosen, the simplex lemma implies that all the rational points of one generation that lie near Player I's ball of the corresponding round

sit on a single hyperplane. In each round Player II therefore has to dodge a thin slab around one hyperplane instead of one interval.

**Full dimension.** Suppose that Player II has a winning strategy in the $(\alpha, \beta)$-game for a set $S \subseteq \mathbb{R}^n$. Inside Player II's ball $V_k$, of radius $\alpha r_k$, Player I has room for roughly $\beta^{-n}$ disjoint balls of radius $\beta\alpha r_k$, and he may play any of them. Let him try them all, while Player II keeps following her strategy. The result is a tree of plays: every ball has about $\beta^{-n}$ children, and the radii shrink by the factor $\alpha\beta$ from one level of the tree to the next. The final points of all these plays form a Cantor set, and every point of this Cantor set lies in $S$, because Player II won every one of those plays. Since the children of each ball can be chosen well separated, a Cantor set built this way has Hausdorff dimension at least roughly

$$\frac{\log(\beta^{-n})}{\log(1/(\alpha\beta))} = \frac{n\log(1/\beta)}{\log(1/\alpha)+\log(1/\beta)}$$

(see Schmidt [Sch66] for the precise bound), and this quantity tends to $n$ as $\beta$ tends to 0. Since an $\alpha$-winning set is winning for *every* $\beta$, letting $\beta$ tend to 0 shows that such a set has dimension $n$. If Player I opens with a ball inside a non-empty open set $U$, the same construction yields the local version: the set $S \cap U$ has dimension $n$.

**Countable intersections.** Let $S_1, S_2, \ldots$ be $\alpha$-winning subsets of $\mathbb{R}^n$. Player II now shares her time among them. She splits the rounds into infinitely many infinite arithmetic progressions; for instance, the $k$-th progression may consist of the rounds $2^{k-1}(2m+1)-1$ with $m \geq 0$. In the rounds of the $k$-th progression she plays her strategy for $S_k$, ignoring everything else. Everything that happens between two consecutive rounds of the $k$-th progression consists of her moves for the other sets and of Player I's moves. From the point of view of $S_k$, this whole stretch of play looks like one very big shrinking move by Player I. If the progression has difference $d_k$, the game for $S_k$ is thus the $(\alpha, \beta_k)$-game with $\beta_k = \beta(\alpha\beta)^{d_k-1}$. Since $S_k$ is $\alpha$-winning, she has a winning strategy in this game. The final point, which is the same for every $k$, therefore lies in every $S_k$. Her time-sharing strategy is a fugue: each voice enters at a time of its own and follows a line of its own, and in the end all of them sound together. This time-sharing is the second feature of games described in the Introduction, and it is exactly what the problem needed. Hausdorff dimension has no such device, and measure, which has one, sees only that $\mathbf{Bad}(n)$ is null.

**Exercises.**

(1) Let $n$ be a positive integer, let $0 < \alpha < 1$ be given, and let $S \subseteq \mathbb{R}^n$ be $\alpha$-winning. Show that $S$ meets every non-empty open subset of $\mathbb{R}^n$, that is, that $S$ is dense in $\mathbb{R}^n$.

(2) Let $n$ be a positive integer, let $\alpha \in (0,1)$ be given, let $S \subseteq \mathbb{R}^n$ be $\alpha$-winning, and let a real number $\lambda \neq 0$ and a vector $\mathbf{t} \in \mathbb{R}^n$ be given. Show that the set $\lambda S + \mathbf{t}$ is $\alpha$-winning. (Hint: the map $\mathbf{x} \mapsto \lambda\mathbf{x} + \mathbf{t}$ carries $B[\mathbf{x}, r]$ onto $B[\lambda\mathbf{x}+\mathbf{t}, |\lambda| r]$; transport the strategy.)

(3) Let $\alpha \in (0,1)$ and $\beta \in (0, \frac{1}{2})$ be given, and let $S \subseteq \mathbb{R}$ be countable. Show, without using Theorem 8.1, that Player I has a winning strategy in the $(\alpha, \beta)$-game for $S$. Conclude that $\mathbb{Q}$ is not $\alpha$-winning for any $\alpha \in (0,1)$. (Hint: since $2\beta < 1$, a closed interval contains two disjoint closed subintervals of $\beta$ times the length of the interval.)

(4) Let $0 < \alpha < \frac{1}{2}$ be given. Show directly that the set of irrational numbers is $\alpha$-winning, and more generally that the complement of every countable subset of $\mathbb{R}$ is $\alpha$-winning.

(5) $(*)$ Let $n$ be a positive integer, let $0 < \alpha' < \alpha < 1$ be given, and let $S \subseteq \mathbb{R}^n$ be $\alpha$-winning. Show that $S$ is $\alpha'$-winning. (Hint: given $\beta' \in (0,1)$, let $\beta = \alpha'\beta'/\alpha$,

whence $\alpha\beta = \alpha'\beta'$. In the $(\alpha', \beta')$-game, Player II answers Player I's ball of radius $r_k$ with a ball of radius $\alpha' r_k$ inside the ball of radius $\alpha r_k$ prescribed by her strategy for the $(\alpha, \beta)$-game.)

## 9. Applications to homogeneous dynamics

**The problem.** A *lattice* in $\mathbb{R}^m$, for a positive integer $m$, is a set of the form $g\mathbb{Z}^m$ with $g$ an invertible real $m \times m$ matrix, and the lattice $g\mathbb{Z}^m$ is *unimodular* if $\det g = \pm 1$. The unimodular lattices form a space, which can be identified with the quotient $\mathrm{SL}(m,\mathbb{R})/\mathrm{SL}(m,\mathbb{Z})$ of the group of real $m \times m$ matrices of determinant 1 by the subgroup of integer matrices. Flows on this space move lattices around by stretching some directions and squeezing others. An example is the flow

$$g_t = \mathrm{diag}\big(e^t, \ldots, e^t, e^{-nt}\big), \qquad t \geq 0,$$

on the unimodular lattices in $\mathbb{R}^{n+1}$, where $n$ is a positive integer (here $m = n+1$) and each $g_t$ carries a lattice $\Lambda$ to the lattice $g_t\Lambda$. For a flow of this kind almost every lattice has a dense orbit $\{g_t\Lambda : t \geq 0\}$, which therefore wanders all over the space. By contrast, the lattices whose orbits stay in a bounded region form a set of measure zero. How big is this set of measure zero?

In this section we show that the game of Section 8 is a game about lattices being stretched and squeezed, so a reader of Section 8 already knows the main idea. The words "Lie group", "lattice" and "horospherical subgroup" will appear, but the heart of the matter is a computation with $2 \times 2$ matrices acting on the plane, which we do by hand below. The Lie group language is just the coordinate system that lets the same computation run in higher dimensions.

**A small glossary.** The reader may keep the example $G = \mathrm{SL}(m,\mathbb{R})$, $\Gamma = \mathrm{SL}(m,\mathbb{Z})$ in mind throughout.

- A *Lie group* is a group that is also a smooth manifold on which multiplication and inversion are smooth, such as $\mathrm{SL}(m,\mathbb{R})$.
- A *lattice* $\Gamma$ in a Lie group $G$ is a discrete subgroup such that the quotient $G/\Gamma$ has finite volume. The word is used in two senses here: a lattice in $\mathbb{R}^m$ is a grid of points, and a lattice in $G$ is a subgroup of matrices. Each such grid is itself a lattice, in the sense just defined, in the group $\mathbb{R}^m$. The two senses are linked by the identification of the coset $g\,\mathrm{SL}(m,\mathbb{Z})$ with the grid $g\mathbb{Z}^m$, under which $\mathrm{SL}(m,\mathbb{R})/\mathrm{SL}(m,\mathbb{Z})$ becomes the space of unimodular grids.
- "Almost every" and "measure zero" refer to the natural measure on $G/\Gamma$ induced by the Haar measure of $G$, which the flow preserves.
- A set of points of $G/\Gamma$ is *bounded* if the closure of the set is compact; for grids, boundedness is made concrete by Mahler's criterion in the history below.
- A flow $g_t$ is *quasi-unipotent* if conjugation by $g_t$ stretches no direction of $G$ exponentially. Equivalently, all the eigenvalues of the linear maps that this conjugation induces on the tangent space of $G$ at the identity (the Lie algebra of $G$) have absolute value 1. For $G = \mathrm{SL}(m,\mathbb{R})$ the flow is quasi-unipotent exactly when all the eigenvalues of the matrices $g_t$ have absolute value 1.
- A lattice $\Gamma$ in $G$ is *irreducible* if no subgroup of finite index in $\Gamma$ is a product of lattices in two non-trivial factors of $G$. In a simple group such as $\mathrm{SL}(m,\mathbb{R})$, every lattice is irreducible.

Other terms (nilpotent, semisimple, mixing) appear only where we describe the general case, and a reader may pass over them.

9.1. **History of the problem and of the solution using topological games.** The space of lattices is not compact: one can squash a lattice to create a very short non-zero vector, and then squash further. "Going to infinity" in this space is usually called "going into the cusp". A lattice deep in the cusp is a very thin needle-like grid with a vector of length nearly zero. By Mahler's compactness criterion (1946), squashing a lattice in this way is the only way to escape to infinity. The criterion states that a set of unimodular lattices is bounded exactly when the lengths of their shortest non-zero vectors are bounded away from zero.

The link with the previous section was found by Dani. He showed in 1985 [Dan85] that badly approximable vectors correspond exactly to bounded orbits of the flow above. In 1986 he used Schmidt's games to show that bounded orbits are abundant for several families of flows [Dan86]. In his plenary address at the International Congress of Mathematicians in Kyoto in 1990 [Mar91], Margulis conjectured a general abundance theorem. In the semisimple setting of Theorem 9.1(2), the conjecture asserts that for every one-parameter flow that is not quasi-unipotent (roughly, every flow that really stretches something), the set of points with bounded orbits is *thick*. A set is thick when the set meets every non-empty open set in a set of full Hausdorff dimension.

Kleinbock and Margulis proved the conjecture in 1996 [KM96]. They worked on the expanding leaves of the flow, but their argument rests on the mixing of the flow, and the argument produces sets of full dimension rather than winning strategies. Kleinbock and Weiss introduced *modified Schmidt games*, adapted to flows that stretch different directions at different rates. In 2013 they proved that sets of bounded orbits are winning for these games [KW13], a result that adds the countable intersection property to thickness. In 2018 An, Beresnevich and Velani turned to the sets $\mathbf{Bad}(i,j)$ of Section 8, of vectors badly approximable with weights $(i,j)$. They established that on every nondegenerate planar curve the points of $\mathbf{Bad}(i,j)$ form a winning subset of the curve for Schmidt's game [ABV18]. Here a nondegenerate planar curve is a twice continuously differentiable curve in the plane whose curvature vanishes nowhere.

*Recognition.* Margulis received the Fields Medal in 1978, the Wolf Prize in 2005 and the Abel Prize in 2020. Dani received the Shanti Swarup Bhatnagar Prize in 1990. He was an invited speaker at the International Congress of Mathematicians in Zürich in 1994. The dynamical approach to Diophantine approximation that grew out of this work led Kleinbock and Margulis to a proof of the Baker–Sprindžuk conjectures, published in the *Annals of Mathematics* in 1998 [KM98]. That proof uses non-divergence estimates, not games.

9.2. **The game.** The game is Schmidt's game of Section 8, played not on the whole space but on an *expanding leaf* of a flow $g_t$ on $G/\Gamma$. The expanding horospherical subgroup $H$ consists of the elements $h$ of $G$ with $g_{-t}hg_t \xrightarrow{t\to\infty} e$, where $e$ is the identity of $G$; these are the directions that the flow stretches. Distances in these directions grow exponentially, and a tiny ball soon becomes a ball of unit size. This exponential growth lets Player II deal with the problem one scale at a time (see the dictionary between scales and times below). The group $H$ is a nilpotent Lie group with a natural metric. The expanding leaf through a point $z$ of $G/\Gamma$ is the orbit $Hz$, the image of $H$ under the map $h \mapsto hz$, which need not be injective. The players play Schmidt's game with closed balls in $H$, and Player II wins if the final point $h$ is such that $hz$ has a bounded forward orbit. For the diagonal flow above, the subgroup $H$ consists of the matrices $\left(\begin{smallmatrix} I_n & \mathbf{y} \\ 0 & 1 \end{smallmatrix}\right)$, and the leaf through $\mathbb{Z}^{n+1}$ is the set of lattices

$\Lambda_{\mathbf{y}} = \left(\begin{smallmatrix} I_n & \mathbf{y} \\ 0 & 1 \end{smallmatrix}\right) \mathbb{Z}^{n+1}$ with $\mathbf{y} \in \mathbb{R}^n$. The game on this leaf is therefore literally Schmidt's game on $\mathbb{R}^n$.

9.3. **The main result of this section.** The main result has two parts: a dictionary between numbers and orbits, due to Dani, and the theorem of Kleinbock and Margulis.

**Theorem 9.1 (Dani, 1985; Kleinbock–Margulis, 1996).**

(1) Dani's correspondence. *Let $n$ be a positive integer, let $g_t = \mathrm{diag}(e^t, \dots, e^t, e^{-nt})$ for $t \geq 0$, and for $\mathbf{x} \in \mathbb{R}^n$ let $\Lambda_{\mathbf{x}} = \left(\begin{smallmatrix} I_n & \mathbf{x} \\ 0 & 1 \end{smallmatrix}\right) \mathbb{Z}^{n+1}$. A vector $\mathbf{x} \in \mathbb{R}^n$ lies in the set $\mathbf{Bad}(n)$ of badly approximable vectors if and only if the forward orbit $\{g_t \Lambda_{\mathbf{x}} : t \geq 0\}$ is bounded in the space of unimodular lattices in $\mathbb{R}^{n+1}$.*

(2) Thickness of bounded orbits. *For every non-quasi-unipotent one-parameter subgroup of a connected semisimple Lie group $G$ without compact factors, and every irreducible lattice $\Gamma$ in $G$, the set of points of $G/\Gamma$ with bounded orbits is thick.*

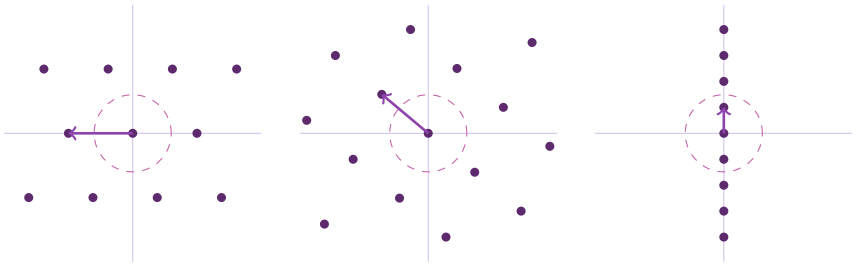
By part (1), Schmidt's theorem of Section 8 is a theorem about one particular flow. Indeed, the vectors $\mathbf{x}$ whose lattices $\Lambda_{\mathbf{x}}$ have bounded forward orbits form the set $\mathbf{Bad}(n)$, which is $\alpha$-winning for some $\alpha \in (0, 1)$. Part (2) extends the abundance of bounded orbits to every flow that stretches. For that part the game serves as a plan rather than as a proof, as we explain in the paragraph on the general case below.

9.4. **Solving the problem with the game.** We first follow the whole mechanism in the simplest case $n = 1$, where every step can be computed by hand, and then explain what changes in general.

**Lattices in the plane.** For $n = 1$ we have $g_t = \mathrm{diag}(e^t, e^{-t})$, and for $x \in \mathbb{R}$ we write $u_x = \left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$; then

$$\Lambda_x = u_x \mathbb{Z}^2 = \{(p + qx,\ q) : p, q \in \mathbb{Z}\}, \qquad g_t \Lambda_x = \{(e^t(p + qx),\ e^{-t} q) : p, q \in \mathbb{Z}\}.$$

The flow stretches the first coordinate and squeezes the second, and the *product* of the two coordinates of a vector, namely $q(p + qx)$, does not change along the orbit. The whole argument rests on this invariant product (see Figure 16).



(a) $x = (\sqrt{5} - 1)/2$, $t = 0$ (b) $x = (\sqrt{5} - 1)/2$, $t = 1.6$ (c) $x = 1/2$, $t = 1.6$

FIGURE 16. The grid $g_t \Lambda_x$ near the origin for the badly approximable number $x = (\sqrt{5} - 1)/2$ in (a) and (b) and for the rational number $x = 1/2$ in (c), with a shortest non-zero vector drawn as an arrow and the circle of radius 0.6 about the origin drawn dashed.

**Dani's correspondence, by hand.** Suppose first that $x \in \mathbf{Bad}(1)$: there is a constant $c > 0$ with $|p + qx| \geq c/|q|$ for all integers $p$ and $q$ with $q \neq 0$. Let $(e^t(p + qx), e^{-t}q)$ be a non-zero vector of $g_t\Lambda_x$, with $t \geq 0$. If $q \neq 0$, the product of the coordinates of this vector has absolute value $|q|\,|p + qx| \geq c$, so the two coordinates cannot both be smaller than $\sqrt{c}$, and the vector has length at least $\sqrt{c}$. If $q = 0$, the vector is $(e^t p, 0)$ with $p \neq 0$, of length at least $e^t \geq 1$. Hence no lattice on the forward orbit has a non-zero vector shorter than $\min(1, \sqrt{c})$, and by Mahler's criterion the orbit is bounded.

Conversely, if $x \notin \mathbf{Bad}(1)$, then for every $\varepsilon \in (0, 1)$ there are integers $p$ and $q \geq 1$ with $|p + qx| < \varepsilon/q$. Let $t$ be the time with $e^{-t}q = \sqrt{\varepsilon}$, which is positive because $q \geq 1 > \sqrt{\varepsilon}$. At this time the vector coming from $(p, q)$ has second coordinate $\sqrt{\varepsilon}$ and first coordinate of absolute value $e^t|p + qx| < (q/\sqrt{\varepsilon})(\varepsilon/q) = \sqrt{\varepsilon}$, and the vector is therefore shorter than $\sqrt{2\varepsilon}$. Since $\varepsilon$ is arbitrary, the orbit makes arbitrarily deep excursions into the cusp, and the orbit is unbounded by Mahler's criterion.

A good rational approximation $-p/q$ of $x$ is thus *the same thing* as an excursion of the orbit into the cusp at a time $t \approx \log q$.

**The dictionary between scales and times.** A one-line computation gives

$$g_t\, u_x\, g_{-t} = u_{e^{2t}x},$$

so for every unimodular lattice $\Lambda$ the flow carries the expanding leaf $\{u_x\Lambda : x \in \mathbb{R}\}$ onto the leaf through $g_t\Lambda$, stretching the leaf by the factor $e^{2t}$. An interval of length $\rho$ in the leaf is stretched to length about 1 at the time $t$ with $e^{2t} \approx 1/\rho$. We now return to the proof in Section 8 that $\mathbf{Bad}(1)$ is $\alpha$-winning. In the proof every round shrinks the intervals by the factor $R = 1/(\alpha\beta)$, and the fractions are sorted into generations by the size of their denominators. In round $k$, Player I's interval has length $\rho_k = \rho_0 R^{-k}$, where $\rho_0$ is the length of his first interval. The fractions $p/q$ that Player II worries about in round $k$ are those of generation $k$, with $q^2$ within a constant factor of $R^k$. By Dani's correspondence as worked out above, these fractions are, up to constant factors, the ones that send the orbit into the cusp at the time $t$ with $e^{2t} \approx R^k \approx 1/\rho_k$. In other words, *round $k$ of the game corresponds to the time at which the current interval has been stretched to unit size. In this round Player II dodges the points whose orbits visit the cusp at this time.* Read in this dictionary, Schmidt's proof is a proof about orbits, and the formula $g_t u_x g_{-t} = u_{e^{2t}x}$ translates each step of one proof into the corresponding step of the other.

**The general case.** Kleinbock and Margulis follow the same plan for an arbitrary flow that stretches, working on an expanding leaf $Hz$. At each scale they look at a ball in the leaf and wait (in their heads) until the flow has stretched the ball to unit size. They then determine the points of the ball that are deep in the cusp at this time. For the diagonal flow on the plane this bad set was a single short interval, the no-go zone of one fraction. In general there is no single dangerous fraction. Here lies the real work: using the mixing of the flow, Kleinbock and Margulis [KM96] show that the bad set occupies only a small proportion of the ball. Removing the bad set at every scale leaves a Cantor-like set whose dimension is close to that of the leaf. The points of this Cantor-like set never go far into the cusp and therefore have bounded orbits. The set of points with bounded orbits thus has full dimension on every expanding leaf. This argument does not produce a winning strategy, since a set that occupies a small proportion of a ball may still meet every ball of the prescribed size inside the larger ball. The winning property came later, with the modified games of Kleinbock and Weiss (see the history above).

**From leaves to the whole space.** Roughly, near any point the space looks like a product of a piece of an expanding leaf with the directions that the flow does not stretch. Whether

an orbit stays bounded depends only mildly on those other directions. Full dimension on every expanding leaf therefore yields full dimension in the whole space, that is, thickness (see Kleinbock and Margulis [KM96]).

**Exercises.**

(1) Let $x \in \mathbb{R}$ and $t \geq 0$ be given, and let $g_t$, $u_x$ and $\Lambda_x$ be as in the case $n = 1$ above. Check by multiplying matrices that $g_t u_x g_{-t} = u_{e^{2t}x}$. Show that $\Lambda_x$ is unimodular, and that the product of the two coordinates of a vector of $g_t\Lambda_x$ coming from $(p, q)$ does not depend on $t$.

(2) Let $x \in \mathbb{R}$ be given. Show that $\Lambda_{x+1} = \Lambda_x$. Explain why this is consistent with Theorem 9.1(1), by checking directly that $x \in \mathbf{Bad}(1)$ if and only if $x + 1 \in \mathbf{Bad}(1)$.

(3) Let $x = a/b$, with integers $a$ and $b \geq 1$. Find a non-zero vector of $g_t\Lambda_x$ of length $be^{-t}$, and deduce that the orbit $\{g_t\Lambda_x : t \geq 0\}$ is unbounded. Check that this agrees with Theorem 9.1(1), by showing that no rational number is badly approximable.

(4) (∗) Let $\varphi = (1+\sqrt{5})/2$ and $\varphi' = (1-\sqrt{5})/2$. Show that $(p+q\varphi)(p+q\varphi') = p^2+pq-q^2$ for all integers $p, q$, and deduce that $|q|\,|p+q\varphi| \geq 1/(1+\sqrt{5})$ whenever $q \neq 0$. Conclude that, for every $t \geq 0$, every non-zero vector of $g_t\Lambda_\varphi$ has length at least $(1+\sqrt{5})^{-1/2}$.

## 10. Applications to topological groups

**The problem.** A *semitopological group* is a group with a topology in which multiplication is *separately* continuous: for every element $a$ of the group, the translations $x \mapsto xa$ and $x \mapsto ax$ are continuous. A semitopological group is a *topological group* if multiplication is jointly continuous and inversion is continuous. Separate continuity is much weaker than joint continuity (the function $f(x, y) = xy/(x^2 + y^2)$, with $f(0, 0) = 0$, is continuous in each variable separately but not jointly continuous at the origin; see Figure 17). When does separate continuity of multiplication force a semitopological group to be a topological group? Behind this lies a more general question: when is a separately continuous real-valued function on a product $X \times Y$ of topological spaces jointly continuous at many points?

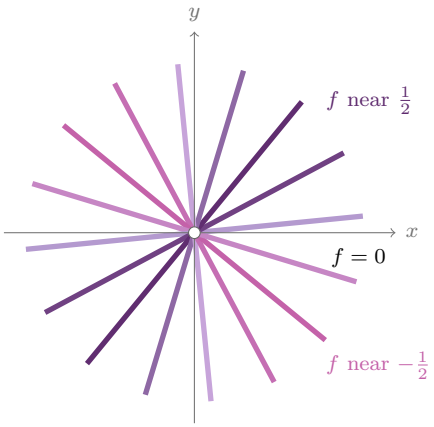



Figure 17. The function $f(x, y) = xy/(x^2 + y^2)$, with $f(0, 0) = 0$, which equals $\frac{1}{2}\sin 2\theta$ on the ray from the origin at angle $\theta$ to the positive $x$-axis, with its value shown by color.

We show in this section that the answer is a matter of completeness. The right amount of completeness, as we also show, is measured by a game of Banach–Mazur type in which Player II must meet a more demanding winning condition.

10.1. **History of the problem and of the solution using topological games.** Baire proved in his thesis of 1899 that a separately continuous function on $\mathbb{R}^2$ is jointly continuous at the points of a dense $G_\delta$ set. In 1957 Ellis proved that every locally compact Hausdorff semitopological group is a topological group [Ell57]. In 1974 Namioka proved a general joint continuity theorem for functions on $X \times K$, with $X$ complete in a suitable sense and $K$ compact, from which he deduced Ellis's theorem [Nam74].

The game-theoretic solution came in three steps. In 1981 Christensen replaced the completeness hypothesis on $X$ in Namioka's theorem by the existence of a winning strategy in a topological game [Chr81]. In 1983 Saint-Raymond showed that the weaker hypothesis that the opposing player has no winning strategy suffices [SR83]. He deduced that the conclusion of Namioka's theorem holds whenever $X$ is a metrizable Baire space. The third step concerns groups. Bouziad proved in 1996 [Bou96] that every Čech-analytic Baire semitopological group is a topological group; the Čech-analytic spaces form a class that contains the Čech-complete spaces of Subsection 4.1. In 2001 Kenderov, Kortezov and Moors introduced a game that describes a class of spaces, the strongly Baire spaces, on which every semitopological group is a topological group [KKM01].

10.2. **The game.** Let $X$ be a topological space, and let $D \subseteq X$ be a dense set. The game $\mathrm{SB}(D)$ (for "strongly Baire") is due to Kenderov, Kortezov and Moors, who call the players $\beta$ and $\alpha$; here $\beta$ becomes Player I and $\alpha$ becomes Player II.

(1) Player I chooses a non-empty open set $U_1 \subseteq X$.
(2) Player II chooses a non-empty open set $V_1 \subseteq U_1$.
(3) In general, for $k \geq 1$, Player I chooses a non-empty open $U_{k+1} \subseteq V_k$ and Player II a non-empty open $V_{k+1} \subseteq U_{k+1}$.
(4) Player II wins if $\bigcap_k V_k \neq \emptyset$ and every sequence $(a_k)_{k\geq 1}$ with $a_k \in V_k \cap D$ has a cluster point in $X$.

A regular space is *strongly Baire* if, for some dense $D$, Player I has no winning strategy in $\mathrm{SB}(D)$. This game is the Banach–Mazur game of Section 3 with a more demanding winning condition for Player II. The games of Christensen and Saint-Raymond are variants of $\mathrm{SB}(D)$ in which Player II also picks points as she goes.

The motivating examples are the Čech-complete spaces of Subsection 4.1, on which Player II wins $\mathrm{SB}(X)$ with the strategy of the proof of Proposition 4.8, stripped of the points. Only the condition on cluster points needs a new argument. To supply the argument, let $X$ be Čech-complete, say with $X = \bigcap_k G_k$ for sets $G_k$ open in a compactification $bX$. For each move $U_k$ of Player I, let $\widetilde{U}_k$ be an open subset of $bX$ with $\widetilde{U}_k \cap X = U_k$. Player II wins with $D = X$ by choosing $V_k$ whose closure in $bX$ lies inside $G_k \cap \widetilde{U}_k$; such a set exists because $bX$ is regular. The closures form a decreasing sequence of non-empty compact sets. Their intersection lies in $\bigcap_k G_k = X$, and the intersection equals $\bigcap_k V_k$ because the closure of $V_{k+1}$ meets $X$ inside $U_{k+1} \subseteq V_k$. Finally, a sequence with $a_k \in V_k$ has a cluster point in $bX$. Since $a_j \in V_k$ for $j \geq k$, this cluster point lies in the closure of every $V_k$, hence in $\bigcap_k V_k$.

10.3. **The main result of this section.** The main result goes from functions to groups. The result comprises Namioka's joint continuity theorem and Ellis's theorem on locally compact groups, together with the theorem of Kenderov, Kortezov and Moors, which contains Ellis's.

**Theorem 10.1 (Ellis, 1957; Namioka, 1974; Kenderov–Kortezov–Moors, 2001).**

(1) Namioka's theorem. *Let $X$ be a Čech-complete space, let $K$ be a compact space, and let $f\colon X \times K \to \mathbb{R}$ be separately continuous. Then there is a dense $G_\delta$ set $A \subseteq X$ such that $f$ is jointly continuous at every point of $A \times K$.*

(2) Ellis's theorem. *Every locally compact Hausdorff semitopological group is a topological group.*

(3) *Every semitopological group whose underlying space is strongly Baire is a topological group.*

By the argument of the preceding subsection, every Čech-complete space is strongly Baire (a Čech-complete space is Tychonov by definition, hence regular). In particular, every locally compact Hausdorff space and every complete metric space is strongly Baire; hence part (3) contains part (2). Part (1) has a useful reformulation (see Exercise 5 below). Let $X$, $K$ and $f$ be as in part (1), and let $C(K)$ be the space of continuous real-valued functions on $K$. The map $x' \mapsto f(x', \cdot)$ from $X$ to $C(K)$ is continuous for the topology of pointwise convergence. For $x \in X$, the function $f$ is jointly continuous at every point of $\{x\} \times K$ exactly when this map is continuous at $x$ for the topology of the supremum norm. Namioka's theorem is therefore about upgrading pointwise convergence to uniform convergence, at least at many points.

10.4. **Solving the problem with the game: a sketch.** The arguments of Christensen, of Saint-Raymond and of Kenderov, Kortezov and Moors share a pattern worth learning, of which we give only a sketch. For the details, see the original papers [Chr81, SR83, KKM01]. Suppose, for instance, that a semitopological group with identity $e$ has a strongly Baire underlying space $X$, and that multiplication is not jointly continuous at $(e, e)$. Then there is a neighborhood $W$ of $e$ such that, for all neighborhoods $O_1$ and $O_2$ of $e$, there are $u \in O_1$ and $v \in O_2$ with $uv \notin W$. Player I uses these pairs to build a strategy as follows. At each move he chooses an open set inside Player II's last move that records a new pair of witnesses. Each new pair is chosen in such a way that separate continuity at any cluster point of the witnesses would be violated. Since $X$ is strongly Baire, there is a dense set $D \subseteq X$ such that Player I has no winning strategy in $\mathrm{SB}(D)$. Some play of $\mathrm{SB}(D)$ consistent with Player I's strategy is therefore won by Player II. This play provides a cluster point of the witnesses, and separate continuity at this point leads to a contradiction. (The witnesses must be taken from the dense set $D$, in order that the winning condition of the game, which only speaks of sequences from $D$, may apply to them. Arranging that the witnesses lie in $D$ is the technical heart of the proofs, and we omit it.) In short, the failure of joint continuity is turned into a strategy for Player I, and the topological hypothesis says exactly that such strategies can be beaten. Continuity of inversion is handled separately. For locally compact groups the argument yields Ellis's theorem. Namioka's original deduction of Ellis's theorem, which uses translations to spread a single point of joint continuity over the whole group, follows the same idea.

**Exercises.**

(1) Let $f(x, y) = xy/(x^2 + y^2)$ for $(x, y) \neq (0, 0)$, and $f(0, 0) = 0$. Show that $f$ is continuous in each variable separately, that $f$ is not jointly continuous at $(0, 0)$, and that $f$ is jointly continuous at every other point of $\mathbb{R}^2$.

(2) Let $X$ be a strongly Baire space, and show that $X$ is a Baire space. (Hint: by Theorem 3.3, a non-empty space $X$ is a Baire space if and only if Player I has no winning strategy in the Banach–Mazur game $\mathrm{BM}(X)$. Moreover, a winning strategy for Player I in $\mathrm{BM}(X)$ is also winning in $\mathrm{SB}(D)$ for every dense set $D$.) Deduce that $\mathbb{Q}$ is not strongly Baire.

(3) Let $X$ be a complete metric space. Show directly that Player II has a winning strategy in $\mathrm{SB}(X)$, by answering $U_k$ with a non-empty open set $V_k$ such that $\overline{V_k} \subseteq U_k$ and $\operatorname{diam} V_k < 2^{-k}$.

(4) Let $\mathbb{R}$ carry the Sorgenfrey topology, whose basic open sets are the intervals $[a, b)$ with $a < b$ real. Show that addition is jointly continuous, that the translations are homeomorphisms, and that $x \mapsto -x$ is not continuous. Assuming that the Sorgenfrey line is regular, deduce from Theorem 10.1(3) that the Sorgenfrey line is not strongly Baire.

(5) $(*)$ Let $X$ be a topological space, let $K$ be a compact space, let $f\colon X \times K \to \mathbb{R}$ be separately continuous, and let $x \in X$ be given. Show that $f$ is jointly continuous at every point of $\{x\} \times K$ if and only if the map $x' \mapsto f(x', \cdot)$ from $X$ to $C(K)$ is continuous at $x$ for the supremum norm.

## 11. Applications to infinite-dimensional Ramsey theory

**The problem.** Ramsey's theorem (1930) says that, for every positive integer $n$, whenever the $n$-element subsets of $\omega = \{0, 1, 2, \dots\}$ are colored with two colors, there is an infinite $H \subseteq \omega$ all of whose $n$-element subsets have the same color. We now color the *infinite* subsets instead. For $H \subseteq \omega$, let $[H]^\omega$ be the set of infinite subsets of $H$. Call $\mathcal{A} \subseteq [\omega]^\omega$ *Ramsey* if there is an infinite $H \subseteq \omega$ such that $[H]^\omega \subseteq \mathcal{A}$ or $[H]^\omega \cap \mathcal{A} = \emptyset$. Using the axiom of choice, one can build sets that are not Ramsey. For instance, let $\mathcal{U}$ be a nonprincipal ultrafilter on $\omega$ (see the definitions at the beginning of Section 13). The set of all $X = \{x_0 < x_1 < \cdots\}$ in $[\omega]^\omega$ for which the union of the intervals $[x_{2k}, x_{2k+1}) = \{m \in \omega : x_{2k} \le m < x_{2k+1}\}$, for $k \in \omega$, belongs to $\mathcal{U}$ is then a set that is not Ramsey (see Exercise 6 below). Which sets are Ramsey?

There is a Rothko hiding in every Pollock: however wildly two colors are splashed on the pairs of natural numbers, there is an infinite set all of whose pairs bear a single color. In this section we show that the picture survives the passage from pairs to infinite sets whenever the coloring is regular enough. More precisely, the sets that are Ramsey in a hereditary sense are exactly those with the Baire property in a suitable topology. Moreover, every analytic set, one that is empty or a continuous image of the space of all sequences of natural numbers, is Ramsey in this hereditary sense. The tool is a game of Banach–Mazur type.

11.1. **History of the problem and of the solution using topological games.** Ramsey's theorem appeared in 1930 [Ram30]. In 1952 Erdős and Rado showed, using the axiom of choice, that the infinite-dimensional analogue of Ramsey's theorem fails [ER52]. Nash-Williams proved in 1965 that clopen sets are Ramsey, for the topology that $[\omega]^\omega$ inherits from the Cantor space $2^\omega$ when every subset of $\omega$ is identified with the characteristic function of the subset [NW65]. Galvin and Prikry extended this result to Borel sets [GP73], and Silver extended the result to analytic sets [Sil70], by a proof that used metamathematical methods. Silver's paper appeared in 1970, three years before that of Galvin and Prikry. Mathias developed the theory further, with applications to forcing, in his work on happy families [Mat77].

Erik Ellentuck solved the problem in 1974 [Ell74]. He introduced a new topology on $[\omega]^\omega$, finer than the usual one. The sets that are Ramsey in a hereditary sense, he proved, are exactly those with the Baire property in this topology. Silver's theorem follows from Ellentuck's in a few lines. In 1983 Kastanas [Kas83] characterized the same sets by the determinacy of a game, which is a Banach–Mazur game in the Ellentuck topology. In Section 12 we show that Gowers's game is the analogue of the Kastanas game for block sequences in Banach spaces.

11.2. **The game.** The game is played in a new topology whose open sets look strange at first, so we begin with the idea behind them. The usual topology on $[\omega]^\omega$ is inherited from the Cantor space $2^\omega$ (see the history above). In this topology two infinite sets are close if they begin in the same way, that is, if they have the same elements below some large integer. Ellentuck's open sets ask for more: they fix the beginning $s$ *and* restrict the rest of the set to come from a given reservoir $M$. Here the beginning $s$ is a finite subset of $\omega$, and the reservoir $M$ is an infinite subset of $\omega$ with $\max s < \min M$, a condition that holds vacuously when $s = \emptyset$. Let

$$[s, M] = \{X \in [\omega]^\omega : s \text{ is an initial segment of } X \text{ and } X \subseteq s \cup M\},$$

where $s$ is an *initial segment* of $X$ if $s \subseteq X$ and every element of $X$ that is smaller than some element of $s$ lies in $s$. Think of $s$ as a stem that has already been decided and of $M$ as the pool of candidates for the rest. The sets $[s, M]$ form a basis of a topology, the *Ellentuck topology* (see Exercise 1 below). A set $\mathcal{A} \subseteq [\omega]^\omega$ is *completely Ramsey* if for every $[s, M]$ there is an infinite $N \subseteq M$ with $[s, N] \subseteq \mathcal{A}$ or $[s, N] \cap \mathcal{A} = \emptyset$.

Let $\mathcal{A} \subseteq [\omega]^\omega$ be given, let $s \subseteq \omega$ be finite, and let $N \subseteq \omega$ be infinite with $\max s < \min N$. The Kastanas game $K(\mathcal{A}, s, N)$ is played as follows.

(1) Player I chooses an infinite $M_0 \subseteq N$.
(2) Player II chooses $n_0 \in M_0$ and an infinite $N_0 \subseteq M_0$ with $n_0 < \min N_0$.
(3) In round $k \geq 1$, Player I chooses an infinite $M_k \subseteq N_{k-1}$, and Player II chooses $n_k \in M_k$ and an infinite $N_k \subseteq M_k$ with $n_k < \min N_k$.
(4) Player II wins if $s \cup \{n_0, n_1, \dots\} \in \mathcal{A}$.

In Kastanas's paper the payoff set is Player I's. The switch does not matter, since the main result below (see Theorem 11.1) concerns determinacy, and since the set $\mathcal{A}$ is completely Ramsey exactly when the complement $[\omega]^\omega \setminus \mathcal{A}$ is (see Exercise 4 below). The moves are basic open sets in disguise: Player I's move $M_k$ stands for $[s \cup \{n_0, \dots, n_{k-1}\}, M_k]$, and Player II answers with the smaller basic open set $[s \cup \{n_0, \dots, n_k\}, N_k]$. The Kastanas game is thus a Banach–Mazur game in the Ellentuck topology, in which Player II must commit to one more element at each move (see Figure 18).

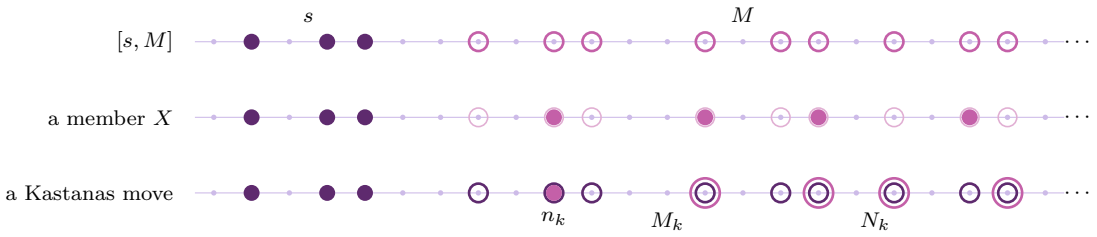



FIGURE 18. The Ellentuck topology on $[\omega]^\omega$: a basic open set $[s, M]$ (top), a member of $[s, M]$ (middle), and round $k$ of the Kastanas game played above the stem $s$ (bottom).

11.3. **The main result of this section.** The main result is Ellentuck's theorem, in the game form due to Kastanas.

**Theorem 11.1 (Ellentuck, 1974; Kastanas, 1983).** *For $\mathcal{A} \subseteq [\omega]^\omega$ the following are equivalent:*

(1) *the set $\mathcal{A}$ is completely Ramsey;*
(2) *the set $\mathcal{A}$ has the Baire property in the Ellentuck topology;*
(3) *for every finite $s \subseteq \omega$ and every infinite $N \subseteq \omega$ with $\max s < \min N$, the Kastanas game $K(\mathcal{A}, s, N)$ is determined.*

In any topological space, the sets with the Baire property are closed under the Souslin operation (see [Kec95]). This operation applies to families of sets $F_u$ indexed by the non-empty finite sequences $u$ of natural numbers. To such a family the operation assigns the union, over all sequences $x = (x_0, x_1, \dots)$ of natural numbers, of the sets $\bigcap_{n\geq 1} F_{(x_0,\dots,x_{n-1})}$. In $[\omega]^\omega$ with the usual topology, as in every separable completely metrizable space, the analytic sets are exactly the sets obtained from closed sets by this operation (see [Kec95]). Every closed set of the usual topology is closed in the finer Ellentuck topology. Hence every analytic set, and in particular every Borel set, has the Baire property in the Ellentuck topology, and is therefore completely Ramsey by Theorem 11.1. Since completely Ramsey sets are Ramsey (see Exercise 4 below), the argument recovers in one stroke the Galvin–Prikry theorem and Silver's theorem. The former asserts that Borel sets are Ramsey, the latter that analytic sets are.

11.4. **Solving the problem with the game.** The proofs of the Galvin–Prikry theorem and of Ellentuck's theorem rest on one device, the *combinatorial forcing* of Galvin and Prikry [GP73]. Fix $\mathcal{A} \subseteq [\omega]^\omega$. For a finite $s \subseteq \omega$ and an infinite $M \subseteq \omega$ with $\max s < \min M$, we say that $M$ *accepts* $s$ if $[s, M] \subseteq \mathcal{A}$, and that $M$ *rejects* $s$ if no infinite $N \subseteq M$ accepts $s$. In plain words, the set $M$ accepts the stem $s$ if every completion of $s$ by infinitely many elements of $M$ lies in $\mathcal{A}$. Likewise, the set $M$ rejects $s$ if no thinning out of $M$ produces a set that accepts $s$. A diagonal argument produces, inside any infinite $M$, an infinite $N$ that *decides* every finite $s \subseteq N$. That is, for each such $s$, the set of elements of $N$ above $\max s$ accepts or rejects $s$. Ellentuck's insight was that these diagonal arguments are exactly the proof that $[\omega]^\omega$, with the Ellentuck topology, is a Baire space. He saw, moreover, that the nowhere dense sets of the Ellentuck topology are exactly the sets $\mathcal{B} \subseteq [\omega]^\omega$ such that every $[s, M]$ contains some $[s, N]$ disjoint from $\mathcal{B}$ [Ell74]. The Kastanas game turns the diagonalization into a play. A winning strategy for Player II in $K(\mathcal{A}, s, N)$, run against all possible moves of Player I and glued together by a diagonal fusion, produces an infinite $M \subseteq N$ with $[s, M] \subseteq \mathcal{A}$. A winning strategy for Player I, in turn, produces an infinite $M \subseteq N$ with $[s, M] \cap \mathcal{A} = \emptyset$ [Kas83]. The property of being completely Ramsey is thus to the Kastanas game what the Baire property is to the Banach–Mazur game.

11.5. **The notation $[\omega]^\omega$ and the partition calculus.** The games of this section are played on the set $[\omega]^\omega$ of infinite subsets of $\omega$, with the Ellentuck topology, and the notation for this set differs from the notation of Section 2. The notation comes from the partition calculus, as we now explain. For a set $S$ and a cardinal $\kappa$, one writes $[S]^\kappa = \{X \subseteq S : |X| = \kappa\}$ and $[S]^{<\kappa} = \{X \subseteq S : |X| < \kappa\}$. Thus the symbols $[\omega]^\omega$ and $[\omega]^{<\omega}$ denote the families of infinite and of finite subsets of $\omega$. These families of subsets are to be distinguished from the spaces $A^\omega$ and $A^{<\omega}$ of infinite and of finite sequences of elements of a set $A$ (see Section 2). The notation $[S]^\kappa$ is that of the partition calculus of P. Erdős and R. Rado [ER56], in which one

writes $\kappa \to (\lambda)^r_\mu$ for cardinals $\kappa$, $\lambda$, $\mu$ and $r$. This relation means that, whenever $|S| = \kappa$ and $[S]^r$ is partitioned into $\mu$ pieces, there is a set $H \subseteq S$ with $|H| = \lambda$ such that $[H]^r$ lies inside a single piece. Such a set $H$ is called *homogeneous*. Ramsey's theorem becomes $\omega \to (\omega)^r_k$ for all positive integers $r$ and $k$. For the infinite exponent the relation fails: using the axiom of choice, Erdős and Rado partitioned $[\omega]^\omega$ into two pieces with no infinite homogeneous set [ER52]. Positive results recover $\omega \to (\omega)^\omega_2$ for partitions that are regular enough. Galvin and Prikry obtained the relation for Borel partitions [GP73], Silver for analytic ones [Sil70], and Ellentuck for partitions with the Baire property in the Ellentuck topology [Ell74]. In this section, the Kastanas game is the bridge to these theorems. The partition calculus and the school of infinite combinatorics that Erdős and his collaborators built around the calculus are set out in the monograph of Erdős, Hajnal, Máté and Rado [EHMR84].

**Exercises.**

(1) Show that the sets $[s, M]$ form a basis of a topology on $[\omega]^\omega$. (Hint: if $[s, M] \cap [t, N]$ is non-empty, then one of $s$ and $t$, say $s$, is an initial segment of the other. In that case the intersection is $[t, P]$, where $P$ is the set of elements of $M \cap N$ greater than $\max t$.)

(2) Show that the Ellentuck topology is finer than the usual (Cantor space) topology: for every positive integer $n$ and every $t \subseteq \{0, \dots, n-1\}$, the set of $X \in [\omega]^\omega$ with $X \cap \{0, \dots, n-1\} = t$ is a basic set $[s, M]$. Show also that, if $M \subseteq \omega$ is infinite and $\omega \setminus M$ is infinite, the set $[\emptyset, M]$ has empty interior in the usual topology.

(3) Let $\mathcal{A}$ be the set of $X \in [\omega]^\omega$ containing infinitely many even numbers, and let $s$ and $N$ be as in the definition of the Kastanas game. Show that Player II has a winning strategy in $K(\mathcal{A}, s, N)$ if $N$ contains only finitely many odd numbers, and that Player I has one otherwise. Find in each case an infinite $N' \subseteq N$ with $[s, N'] \subseteq \mathcal{A}$ or $[s, N'] \cap \mathcal{A} = \emptyset$.

(4) Show that a set $\mathcal{A} \subseteq [\omega]^\omega$ is completely Ramsey if and only if $[\omega]^\omega \setminus \mathcal{A}$ is, and that every completely Ramsey set is Ramsey.

(5) $(*)$ Color the two-element subsets of $\omega$ red or blue, and let $\mathcal{A}$ be the set of $X \in [\omega]^\omega$ whose two least elements form a red pair. Show that $\mathcal{A}$ is completely Ramsey. (Hint: treat separately the cases in which the stem $s$ of a basic set $[s, M]$ has at least two elements, one element, or none; in the last case use Ramsey's theorem.)

(6) Let $\mathcal{U}$ be a nonprincipal ultrafilter on $\omega$, viewed as a subset of $[\omega]^\omega$. Show that $\mathcal{U}$ itself is Ramsey, by finding an infinite $H \notin \mathcal{U}$ and checking that $[H]^\omega \cap \mathcal{U} = \emptyset$. For $X = \{x_0 < x_1 < \cdots\}$ in $[\omega]^\omega$, let $Y_X$ be the union of the intervals $[x_{2k}, x_{2k+1}) = \{m \in \omega : x_{2k} \le m < x_{2k+1}\}$ for $k \in \omega$, and let $\mathcal{A}$ be the set of $X$ with $Y_X \in \mathcal{U}$. Show that $Y_X$ and $Y_{X \setminus \{x_0\}}$ are disjoint and that their union is $\{m \in \omega : m \ge x_0\}$. Deduce that $\mathcal{A}$ is not Ramsey.

## 12. Applications to Banach space theory

**The problems.**
(1) *Differentiability of convex functions.* In calculus, a convex function on an open interval is differentiable at all but countably many points. In infinite dimensions things are much less tidy. A Banach space $X$ is an *Asplund space* if every continuous convex function on an open convex subset of $X$ is Fréchet differentiable on a dense $G_\delta$ subset of the domain of the function. Fréchet differentiability is defined in the glossary below. Which spaces are Asplund spaces?
(2) *The unconditional basic sequence problem.* Does every infinite-dimensional Banach space contain an unconditional basic sequence? Here an unconditional basic sequence is a sequence $(e_n)_{n\geq 1}$ that is a basis (see the glossary below) of the closed subspace spanned by $e_1, e_2, \ldots$. Moreover, every convergent expansion $\sum_n a_n e_n$ is required to keep converging when the signs of the terms are changed or the terms are rearranged.

In this section we bring two games to bear on the geometry of Banach spaces. The fragmentability game of Kenderov and Moors detects the Asplund property, and Gowers's game drives his dichotomy theorem, which explains the answer to the unconditional basic sequence problem. Functional analysis lies outside the background we assume, so we begin with a short glossary.

**A small glossary.** Let $X$ be a Banach space.

- The *dual* $X^*$ is the Banach space of bounded linear functionals on $X$, and the closed unit ball of $X^*$ is denoted by $B_{X^*}$. The *weak* topology* on $X^*$ is the topology of pointwise convergence on the elements of $X$.
- Let $f$ be a function on an open subset $O$ of $X$, and let $x \in O$ be given. The function $f$ is *Gâteaux differentiable* at $x$ if there is $\varphi \in X^*$ such that, for each fixed $h \in X$, we have $f(x+th) = f(x) + t\varphi(h) + o(t)$ as the real number $t$ tends to 0. The function $f$ is *Fréchet differentiable* at $x$ if there is $\varphi \in X^*$ with $f(x+h) = f(x) + \varphi(h) + o(\|h\|)$ as $\|h\| \to 0$.
- A *basis* of $X$ is a sequence $(e_n)_{n\geq 1}$ in $X$ such that every $x \in X$ has a unique expansion $x = \sum_n a_n e_n$ with scalars $a_n$. A *basic sequence* is a sequence that is a basis of the closed linear span of the sequence. For a constant $C \geq 1$, a basic sequence $(e_n)_{n\geq 1}$ is *$C$-unconditional* if $\|\sum_n \varepsilon_n a_n e_n\| \leq C\|\sum_n a_n e_n\|$ for all finitely supported sequences of scalars $(a_n)$ and all signs $\varepsilon_n = \pm 1$. A basic sequence is *unconditional* if the sequence is $C$-unconditional for some $C$.
- The space $X$ is the *direct sum* of closed subspaces $Y_1$ and $Y_2$ if $Y_1 \cap Y_2 = \{0\}$ and $Y_1 + Y_2 = X$.
- Let $K$ be a set with a topology $\mathcal{T}$ and a metric $d$. The set $K$ is *fragmented* by $d$ if every non-empty subset of $K$ has non-empty relatively $\mathcal{T}$-open subsets of arbitrarily small $d$-diameter.

12.1. **History of the problems and their solutions using topological games.** Mazur proved in 1933 that continuous convex functions on separable Banach spaces are Gâteaux differentiable on dense $G_\delta$ sets. Asplund introduced the spaces now named after him in 1968, in a paper in *Acta Mathematica* [Asp68]. The characterization problem was solved in 1975 by Namioka and Phelps [NP75] and by Stegall [Ste75]. The property behind the Asplund spaces is fragmentability. Kenderov and Moors described fragmentability by a game, first in 1996 and then, in the form used below, in a paper published in 1999 [KM99].

The unconditional basic sequence problem was one of the central open problems of Banach space theory for decades. The problem was solved, negatively, by Gowers and Maurey in 1991 (published in 1993) [GM93]. They built a *hereditarily indecomposable* space: an infinite-dimensional Banach space so tangled that no infinite-dimensional closed subspace of the space is the direct sum of two infinite-dimensional closed subspaces. Gowers then showed that such strange spaces are no accident. In papers of 1996 and 2002 he proved that every infinite-dimensional Banach space contains an infinite-dimensional subspace that either has an unconditional basis or is hereditarily indecomposable [Gow96, Gow02]. His proof of this dichotomy rests on an infinite game played with subspaces and vectors. The second paper, which contains the Ramsey theorem behind the game, appeared in the *Annals of Mathematics*.

*Recognition.* Gowers received the Fields Medal in 1998 for research connecting functional analysis and combinatorics; this research includes the solution of the unconditional basic sequence problem and the dichotomy theorem. He was an invited speaker at the International Congress of Mathematicians in Zürich in 1994, and he received the Prize of the European Mathematical Society in 1996.

12.2. **The games.**

12.2.1. *The fragmentability game.* Let $K$ be a set with two topologies $\mathcal{T}_1$ and $\mathcal{T}_2$. The case to keep in mind is the dual ball $B_{X^*}$ of a Banach space $X$, with the weak$^*$ topology as $\mathcal{T}_1$ and the norm topology as $\mathcal{T}_2$. The game is played as follows.

(1) Player I chooses a non-empty subset $U_0 \subseteq K$.
(2) Player II chooses a non-empty relatively $\mathcal{T}_1$-open subset $V_0 \subseteq U_0$.
(3) In round $k \geq 1$, Player I chooses a non-empty subset $U_k \subseteq V_{k-1}$, and Player II a non-empty relatively $\mathcal{T}_1$-open subset $V_k \subseteq U_k$.
(4) Player II wins if $\bigcap_k U_k$ is empty, or if $\bigcap_k U_k$ is a single point $x$ and every $\mathcal{T}_2$-neighborhood of $x$ contains some $U_k$.

This is a Banach–Mazur game with an asymmetry: Player I may play any non-empty subset of the previous move, while Player II must play relatively $\mathcal{T}_1$-open ones. She wins if the sets close in on at most one point and, when they close in on a point, they do so in the topology $\mathcal{T}_2$.

12.2.2. *Gowers's game.* Let $X$ be a Banach space with a basis $(e_n)_{n\geq 1}$. A *block vector* is a non-zero finite linear combination of the $e_n$. A *block subspace* is the closed span of a sequence of block vectors $x_1, x_2, \dots$ with *successive supports*: every basis vector used in $x_k$ comes before every basis vector used in $x_{k+1}$. A sequence of block vectors with successive supports is a *block sequence*, and a vector is *normalized* if the vector has norm 1. Let $\Sigma$ be a set of sequences of normalized block vectors. In round $k \geq 1$, Player I chooses an infinite-dimensional block subspace $Y_k$, and Player II chooses a normalized block vector $x_k \in Y_k$; she wins if $(x_1, x_2, \dots) \in \Sigma$. Player I's subspaces play the role of open sets. In the Banach–Mazur game a small neighborhood limits where the other player may look, and a block subspace does the same. Player I may choose his subspace to live as far out along the basis as he likes, thereby forcing Player II's next vector to be supported beyond any finite stage. He can also restrict the kind of vectors available to her.

12.3. **The main results of this section.** Differentiability is tied to the dual space for a simple reason: the derivative of a function on $X$ at a point is a linear functional, that is, an element of $X^*$. Differentiability is therefore a question about how the candidate derivatives sit inside $X^*$. For a convex function $f$ on an open convex subset $O$ of $X$ and a point $x \in O$, the

candidate derivatives are the *subgradients*, the functionals $\varphi \in X^*$ with $f(y) \geq f(x)+\varphi(y-x)$ for all $y \in O$. Roughly, Fréchet differentiability at a point means that subgradients at nearby points are close to the derivative *in norm*, and the norm is what fragmentability controls. In the first result below we collect several characterizations of Asplund spaces, one of which is exactly what the game of Kenderov and Moors detects; the second result is Gowers's dichotomy.

**Theorem 12.1 (Asplund, 1968; Namioka–Phelps, 1975; Stegall, 1975).** *For a Banach space $X$ the following are equivalent:*

(1) *the space $X$ is an Asplund space;*
(2) *every separable subspace of $X$ has a separable dual;*
(3) *the dual $X^*$ has the Radon–Nikodým property;*
(4) *the dual ball $B_{X^*}$, with the weak* topology, is fragmented by the norm.*

The Radon–Nikodým property in (3) is a property from the theory of vector-valued measures, which we do not use below. Condition (4), in the terms of the glossary at the beginning of this section, says that every non-empty subset of $B_{X^*}$ has non-empty relatively weak*-open subsets of arbitrarily small norm diameter. Kenderov and Moors [KM99] proved that the fragmentability game detects a general form of this property. Let $K$ be a set with two topologies $\mathcal{T}_1$ and $\mathcal{T}_2$, regard $K$ as a space with the topology $\mathcal{T}_1$, and consider the fragmentability game on $K$. Their theorem says that Player II has a winning strategy in this game if and only if the space $K$ is fragmented by a metric on $K$ whose topology contains $\mathcal{T}_2$. In particular, by condition (4), Player II wins the game on $B_{X^*}$ with the weak* and norm topologies whenever $X$ is an Asplund space.

**Theorem 12.2 (Gowers's dichotomy, 1996 and 2002).** *Every infinite-dimensional Banach space contains an infinite-dimensional subspace that either has an unconditional basis or is hereditarily indecomposable.*

The two alternatives exclude each other, because a hereditarily indecomposable space contains no unconditional basic sequence [GM93]. Indeed, consider the closed spans of the even-numbered and of the odd-numbered terms of such a sequence. These two spans would be infinite-dimensional closed subspaces whose direct sum is the closed span of the whole sequence. Every infinite-dimensional Banach space, however large, thus contains a subspace that is either perfectly civilized or thoroughly rebellious; there is no third option.

### 12.4. **Solving the problems with the games.**

12.4.1. *Asplund spaces.* Problem (1) was solved in 1975 without games, by Namioka and Phelps and by Stegall. What the game of Kenderov and Moors adds is a new way to establish fragmentability, and with fragmentability the Asplund property. Instead of producing a fragmenting metric, one describes how Player II should play. We give the easy half of their theorem.

Let $X$ be a Banach space, and suppose that $B_{X^*}$ is fragmented by the norm. Consider the fragmentability game on $B_{X^*}$, with the weak* topology as $\mathcal{T}_1$ and the norm topology as $\mathcal{T}_2$. Then Player II has a simple strategy: she answers each move $U_k$ with a relatively weak*-open subset $V_k \subseteq U_k$ of norm diameter less than $2^{-k}$. Since $U_{k+1} \subseteq V_k$ for every $k$, the intersection $\bigcap_k U_k$ has norm diameter 0, and the intersection is either empty or a single point $x$. In the second case, the set $U_{k+1}$ lies in the norm ball of radius $2^{-k}$ about $x$, whence

every norm neighborhood of $x$ contains some $U_k$, and Player II wins. Conversely, Kenderov and Moors build a fragmenting metric out of a winning strategy for Player II; roughly, two points are close when the strategy takes a long time to separate them. The advantage of the game is that a strategy is a local object, built one move at a time, while a fragmenting metric is a global one. Moreover, describing how to play is often much easier than writing down a metric.

12.4.2. *Gowers's dichotomy.* Let $X$ be a Banach space with a basis. Let $\Sigma$ be a set of sequences of normalized block vectors that is analytic in the product of countably many copies of $X$, and fix a sequence of positive tolerances. Wiggling the members of $\Sigma$ within these tolerances produces a slight enlargement of $\Sigma$. Gowers's infinite Ramsey theorem [Gow02] concerns Gowers's game played below a block subspace $Y$ of $X$, in which Player I is restricted to block subspaces of $Y$. The theorem asserts that, for some such $Y$, either no block sequence of $Y$ lies in $\Sigma$, or Player II has a winning strategy in this game for that enlargement. The theorem is the counterpart, for block sequences, of the Galvin–Prikry theorem and Ellentuck's theorem on Ramsey sets (see Section 11), and Gowers's game plays the part of the Kastanas game. An exact Ramsey statement is false for vectors; what survives is the game, played up to the given tolerances. Gowers's proof uses a "combinatorial forcing" argument, in which subspaces are said to accept or reject finite block sequences.

To obtain the dichotomy, one first passes to a subspace with a basis; every infinite-dimensional Banach space has such a subspace. If this space contains an unconditional basic sequence, the closed span of the sequence gives the first alternative. Otherwise, Gowers uses the scalar-weighted formulation of his infinite Ramsey theorem: the analytic payoff requires finite consecutive segments whose sums have arbitrarily large norm relative to their alternating sums [Gow96, Gow02]. The winning strategy is obtained on a further block subspace, with tolerances chosen small enough to preserve the estimates needed in the argument. From this strategy Gowers extracts a hereditarily indecomposable subspace, none of whose infinite-dimensional closed subspaces is the direct sum of two infinite-dimensional closed subspaces. The game locates this subspace; that such spaces exist at all is the theorem of Gowers and Maurey.

**Exercises.**

(1) Let $c_0$ be the space of real sequences $(a_n)_{n\geq 1}$ converging to 0, with the supremum norm. For $1 \leq p < \infty$, let $\ell_p$ be the space of real sequences $(a_n)_{n\geq 1}$ for which the norm $\|(a_n)\|_p = (\sum_n |a_n|^p)^{1/p}$ is finite. Show that the standard unit vector bases of these spaces are 1-unconditional.

(2) Let $f(x,y) = x^3y/(x^4+y^2)$ for $(x,y) \neq (0,0)$, and $f(0,0) = 0$. Show that $f$ is Gâteaux differentiable at the origin, with $\varphi = 0$, but not Fréchet differentiable there. (Hint: look at $f(x,x^2)$.)

(3) Let $(K,d)$ be a metric space, and let $\mathfrak{T}_1$ and $\mathfrak{T}_2$ both be the topology of $d$. Show that Player II has a winning strategy in the fragmentability game, by answering $U_k$ with the intersection of $U_k$ and a ball of radius $2^{-k-1}$ about a point of $U_k$.

(4) Let $X$ be a Banach space with a 1-unconditional basis $(e_n)_{n\geq 1}$, and let $x_1, x_2, \ldots$ be block vectors with successive supports. Show that $\|\sum_k \varepsilon_k b_k x_k\| \leq \|\sum_k b_k x_k\|$ for all finite sequences of scalars $b_k$ and signs $\varepsilon_k = \pm 1$.

(5) $(*)$ Let $X$ be a Banach space with a basis $(e_n)_{n\geq 1}$, and consider Gowers's game in $X$. Show that Player II has a winning strategy when $\Sigma$ is the set of all sequences of normalized block vectors with successive supports. Show also that Player I has a

winning strategy when $\Sigma$ is the set of sequences whose first term is $e_1/\|e_1\|$. (Hint: the coefficient functionals of a basis are continuous.)

## 13. Applications to ultrafilters

In the last section of this part the Banach–Mazur game is turned on the axiom of choice itself, through the ultrafilters on the natural numbers. We begin with the vocabulary needed to state the problem.

**Definitions.** A *filter* on $\omega$ is a non-empty family $\mathcal{F}$ of subsets of $\omega$ that does not contain $\emptyset$, is closed under finite intersections, and contains all supersets of members of $\mathcal{F}$. An *ultrafilter* is a filter $\mathcal{U}$ such that for every $x \subseteq \omega$ either $x$ or $\omega \setminus x$ belongs to $\mathcal{U}$. For each $n \in \omega$, the ultrafilter of all subsets of $\omega$ containing $n$ is called *principal*; every other ultrafilter is *nonprincipal*. A nonprincipal ultrafilter contains every cofinite set (see Exercise 1 at the end of this section). Such an ultrafilter $\mathcal{U}$ is a *P-point* if, for every partition of $\omega$ into pieces that do not belong to $\mathcal{U}$, some element of $\mathcal{U}$ meets every piece in a finite set. The ultrafilter $\mathcal{U}$ is *selective* if, for every such partition, some element of $\mathcal{U}$ meets every piece in at most one point. The set of all ultrafilters on $\omega$, with the topology whose basic open sets are the sets of ultrafilters containing a given subset of $\omega$, is a compact space, denoted by $\beta\omega$. When each $n \in \omega$ is identified with the principal ultrafilter of the sets containing $n$, the subspace $\omega^* = \beta\omega \setminus \omega$ consists of the nonprincipal ultrafilters; these spaces appear only in passing.

**The problem.** Identify every subset of $\omega$ with the characteristic function of the subset. The power set $\mathcal{P}(\omega)$ then becomes the Cantor space $2^\omega$ of infinite binary sequences, with the product topology, and a filter on $\omega$ becomes a set of reals. (In set theory the elements of $2^\omega$ are called reals, and the subsets of $2^\omega$ sets of reals.) The axiom of choice provides nonprincipal ultrafilters on $\omega$ in abundance (there are $2^{2^{\aleph_0}}$ of them, the points of $\omega^*$), and yet nobody has ever exhibited one. The problem is to explain this state of affairs: to determine how irregular a nonprincipal ultrafilter must be as a set of reals and how much choice the existence of such an ultrafilter requires. A second task is to determine which ultrafilters are tame enough to carry the Ramsey theory of Section 11.

A principal ultrafilter is a point of $\omega$ in disguise: it collects the sets that contain the point. The two-valued measure that the ultrafilter defines (measure 1 for its members and 0 for all other subsets of $\omega$) is the Dirac measure at the point. A nonprincipal ultrafilter is a point at infinity that passes judgment on every subset of $\omega$, consistently and irrevocably, and always against the finite ones. Seen through the Banach–Mazur game on the Cantor space, played with finite binary strings (Ulam's game below), such a judge is invisible to topology. Indeed, in Theorem 13.1(1) below we show that in the game defined by a nonprincipal ultrafilter neither player has a winning strategy. Seen through Galvin's games, the well-mannered judges are the selective ultrafilters and the P-points, and in Theorem 13.1(4) below we show that they are exactly the judges that cannot be outplayed.

13.1. **History of the problem and of the solution using topological games.** Ultrafilters came into mathematics through two doors, the algebra of sets and general topology. Tarski proved in 1930, using the axiom of choice, that every filter of subsets of a set extends to an ultrafilter [Tar30]; Stone made the ultrafilters of a Boolean algebra into the points of a compact space in 1936 and 1937 [Sto36]; and Cartan introduced filters and ultrafilters in 1937 to describe convergence in general topology [Car37]. In 1938 Sierpiński observed that a nonprincipal ultrafilter on $\omega$, regarded as a set of reals, is not Lebesgue measurable [Sie38].

The symmetry behind his argument, the flip $x \mapsto \omega \setminus x$, also shows that such an ultrafilter does not have the Baire property (we prove both facts in Theorem 13.1(2) below). In 1962 Mycielski and Steinhaus proposed the Axiom of Determinacy, by which every game in which the players choose natural numbers is determined [MS62]. The axiom makes every set of reals Lebesgue measurable and gives it the Baire property, and it therefore rules out nonprincipal ultrafilters on $\omega$ altogether (see Theorem 13.1(3) below). Feferman showed in 1965, by forcing, that the existence of a definable nonprincipal ultrafilter on $\omega$ cannot be proved in ZFC (Zermelo–Fraenkel set theory with the axiom of choice) [Fef65]. Shelah showed in 1984 that ZF (the same theory without the axiom of choice) together with the axiom of dependent choices is consistent with the statement that every set of reals has the Baire property, relative to ZF alone [She84]. Some choice beyond dependent choice is therefore needed to produce a single nonprincipal ultrafilter on $\omega$.

Meanwhile the ultrafilters themselves were being sorted. Walter Rudin proved in 1956 that under the continuum hypothesis there are P-points, and deduced that $\omega^*$ is not homogeneous, that is, that some point of $\omega^*$ cannot be carried to some other point by a homeomorphism of $\omega^*$ [Rud56]. Frolík proved in 1967 that $\omega^*$ is not homogeneous in ZFC [Fro67]. Booth and Kunen studied the selective ultrafilters [Boo70], and Kunen proved that a nonprincipal ultrafilter on $\omega$ is selective exactly when, for every coloring of the two-element subsets of $\omega$ in two colors, the ultrafilter contains a set all of whose two-element subsets have the same color. Mathias showed in 1977 that selective ultrafilters carry the Ramsey theory of analytic sets [Mat77]. Kunen proved in 1976 that it is consistent with ZFC that there are no selective ultrafilters [Kun76]. Shelah proved shortly afterwards that it is consistent that there are no P-points at all (see the account of Wimmers [Wim82]).

The topological side of the problem was settled by Jalali-Naini in his thesis of 1976 [JN76] and, independently, by Talagrand in 1980 [Tal80]. Both characterized the meager filters by a combinatorial condition on partitions of $\omega$ into intervals. Talagrand's motivation came from pointwise compact sets of measurable functions, a subject closely tied to the theory of Rosenthal compacta, begun by Rosenthal and developed by Bourgain, Fremlin and Talagrand. The games enter twice. The irregularity of a nonprincipal ultrafilter follows from a strategy-stealing argument in the Banach–Mazur game on the Cantor space, which rests on the same symmetry $x \mapsto \omega \setminus x$ as Sierpiński's proof (we give the argument in the proof of Theorem 13.1(1) below); and Galvin and McKenzie characterized the selective ultrafilters and the P-points by the games described below (see the book of Bartoszyński and Judah [BJ95]).

*Recognition.* Talagrand received the Abel Prize in 2024, and Shelah the Wolf Prize in 2001.

13.2. **The games.**

**Ulam's game.** Let $Z \subseteq 2^\omega$ be given. In the game $\mathrm{Ul}(Z)$, Player I and Player II alternately choose non-empty finite binary strings, forever. Player I chooses $s_0$, Player II answers with $t_0$, Player I continues with $s_1$, and so on. Player II wins if the concatenation

$$x = s_0{}^\frown t_0{}^\frown s_1{}^\frown t_1{}^\frown \cdots$$

belongs to $Z$, where $u{}^\frown v$ denotes the finite binary string $u$ followed by the binary sequence $v$, finite or infinite. For a finite binary string $u$, let $N_u = \{x \in 2^\omega : u \subseteq x\}$ be the basic clopen set of the sequences that begin with $u$. Each move produces a position, the concatenation of the moves made so far, and the basic clopen set of that position lies inside the basic clopen set of the previous one. The game $\mathrm{Ul}(Z)$ is therefore the Banach–Mazur game $\mathrm{BM}(2^\omega, Z)$ of Section 3, with the moves restricted to basic clopen sets. We name the game after Ulam, who proposed, right after Problem 43 of the Scottish Book, a variant in which each move is a single

binary digit (see Section 31). Since the Cantor space is a compact metric space, both parts of the theorem of Banach and Oxtoby apply (see Theorem 3.5). The proof of the theorem works equally well with moves restricted to basic open sets. In the present setting, the first part asserts that Player II has a winning strategy in $\mathrm{Ul}(Z)$ if and only if $Z$ is comeager. By the second part, Player I has one if and only if $Z \cap N_u$ is meager for some finite binary string $u$.

**Galvin's games.** Let $\mathcal{U}$ be an ultrafilter on $\omega$. In the *selective game* $S(\mathcal{U})$, at each turn $k \in \omega$, Player I chooses a set $A_k \in \mathcal{U}$ and Player II responds with an element $n_k \in A_k$; Player II wins if $\{n_0, n_1, \dots\} \in \mathcal{U}$. In the *P-point game* $P(\mathcal{U})$, Player I again chooses sets $A_k \in \mathcal{U}$, but Player II answers each $A_k$ with a finite set $F_k \subseteq A_k$ instead of a single element, and she wins if $\bigcup_k F_k \in \mathcal{U}$. Player I plays the skeptic: at each turn he names a set that $\mathcal{U}$ approves of. His challenge to Player II is to assemble, one element (or one finite batch) at a time, a set that $\mathcal{U}$ approves of as well.

The tame ultrafilters of the main result are the P-points and the selective ultrafilters defined at the beginning of this section. Every selective ultrafilter is a P-point, since meeting each piece of a partition in at most one point is a special case of meeting the piece in a finite set. The counterpart of this inclusion for the games is the content of Exercise 4 below.

13.3. **The main result of this section.** In the main result we gather the behavior of nonprincipal ultrafilters in the two kinds of games, together with the consequences for category, measure and determinacy: in Ulam's game they are invisible, and in Galvin's games the tame ones are exactly those that cannot be outplayed.

**Theorem 13.1 (Sierpiński, 1938; Mycielski–Steinhaus, 1962; Galvin–McKenzie).** *Let $\mathcal{U}$ be a nonprincipal ultrafilter on $\omega$, regarded as a subset of $2^\omega$. Then the following hold.*

(1) *The game $\mathrm{Ul}(\mathcal{U})$ is not determined.*
(2) *The set $\mathcal{U}$ is neither meager nor comeager, does not have the Baire property, and is not measurable for the coin-tossing measure on $2^\omega$, the product of countably many copies of the measure on $\{0,1\}$ that gives measure $\frac{1}{2}$ to each of $0$ and $1$.*
(3) *If the game $\mathrm{Ul}(Z)$ is determined for every $Z \subseteq 2^\omega$, then every subset of $2^\omega$ has the Baire property, and there is no nonprincipal ultrafilter on $\omega$.*
(4) *The ultrafilter $\mathcal{U}$ is selective if and only if Player I has no winning strategy in $S(\mathcal{U})$, and $\mathcal{U}$ is a P-point if and only if Player I has no winning strategy in $P(\mathcal{U})$.*

By part (2), it is no accident that nobody has written down a nonprincipal ultrafilter on $\omega$. Analytic subsets of $2^\omega$ (the empty set and the continuous images of the space $\omega^\omega$ of all sequences of natural numbers) and coanalytic ones (their complements) have the Baire property (see [Kec95]). Consequently, no nonprincipal ultrafilter is analytic or coanalytic. The proofs of (1) and (2), given below, use only two properties of $\mathcal{U}$:

- membership in $\mathcal{U}$ does not depend on finitely many digits, because $\mathcal{U}$ is nonprincipal;
- for every $x \subseteq \omega$, the complement $\bar{x} = \omega \setminus x$ belongs to $\mathcal{U}$ if and only if $x$ does not, because $\mathcal{U}$ is an ultrafilter.

13.4. **Solving the problem with the games.**

**Stealing a strategy.** The idea is to let Player I play two games at once against a strategy of Player II, the second always the mirror image of the first. The purpose of this arrangement is to force the strategy to place both a set and the complement of the set in $\mathcal{U}$. For a finite

binary string $s$, we write $\bar{s}$ for the string obtained by flipping every digit of $s$. Suppose that $\tau$ is a winning strategy for Player II in $\mathrm{Ul}(\mathcal{U})$. Player I opens the first game with the string 0, and the strategy $\tau$ answers $t_0$. He opens the second game with $\overline{0^\frown t_0}$ and receives the answer $u_0$. Back in the first game, he plays $\bar{u}_0$ and receives $t_1$; in the second he plays $\bar{t}_1$ and receives $u_1$; and so on (Figure 19). By construction, each block of either play is the flip of the corresponding stretch of the other. Consequently, the play $x$ of the first game and the play $y$ of the second satisfy $y = \bar{x}$. Both plays are won by $\tau$, whence $x \in \mathcal{U}$ and $\bar{x} \in \mathcal{U}$, which is impossible. Hence Player II has no winning strategy.

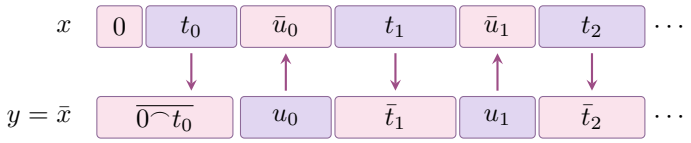



FIGURE 19. Stealing Player II's strategy $\tau$ in $\mathrm{Ul}(\mathcal{U})$, with the play $x$ of the first game on top, the play $y = \bar{x}$ of the second below, the answers of $\tau$ in lavender and the moves of Player I in blush.

Suppose now that Player I has a winning strategy $\sigma$, and let $s$ be the first move of $\sigma$. From the position $s$ on, the strategy $\sigma$ answers the moves of Player II and forces $x \notin \mathcal{U}$, that is, $\bar{x} \in \mathcal{U}$. Flipping every move, we obtain a strategy that, from the position $\bar{s}$, answers the opponent and forces the play into $\mathcal{U}$. Player II can transplant this strategy into $\mathrm{Ul}(\mathcal{U})$. She makes her first move at least as long as $s$, so that at her next turn the position is longer than $s$. She then pretends that the position is $\bar{s}$ followed by a single move of her opponent, namely the digits of the actual position beyond the length of $s$. From then on she answers as the flipped strategy does. The resulting play differs in finitely many digits from a play that the flipped strategy wins, and membership in $\mathcal{U}$ does not depend on finitely many digits. The transplanted strategy is therefore a winning strategy for Player II in $\mathrm{Ul}(\mathcal{U})$, which we have just seen to be impossible. This contradiction completes the proof of part (1).

**Baire property and measure.** If $Z$ is comeager, then Player II has a winning strategy in $\mathrm{Ul}(Z)$. Her strategy is to play, for each $k \in \omega$, her $k$-th move into the $k$-th of countably many dense open sets whose intersection lies in $Z$. If $Z$ is meager, then Player I has a winning strategy of the same kind. By (1), the set $\mathcal{U}$ is therefore neither meager nor comeager. We turn next to sets $Z \subseteq 2^\omega$ that are invariant under finite changes, that is, that contain every sequence differing in finitely many digits from one of their members. Such a set, if the set has the Baire property, is meager or comeager. Indeed, suppose that $Z$ is not meager. Since $Z$ has the Baire property, the set $Z$ differs by a meager set from an open set, which must be non-empty; hence the difference $N_s \setminus Z$ is meager for some string $s$. For every string $t$ of the same length as $s$, the homeomorphism of $2^\omega$ that replaces the initial segment $s$ by $t$ carries $N_s$ onto $N_t$ and preserves $Z$. It follows that, for every such string $t$, the set $N_t \setminus Z$ is meager as well. The finitely many sets $N_t$ cover $2^\omega$, whence the complement of $Z$ is meager. As $\mathcal{U}$ is invariant under finite changes and is neither meager nor comeager, the set $\mathcal{U}$ does not have the Baire property. Measure fares no better: the flip $x \mapsto \bar{x}$ preserves the coin-tossing measure $\lambda$ and carries $\mathcal{U}$ onto the complement of $\mathcal{U}$. If $\mathcal{U}$ were measurable, we would thus have $\lambda(\mathcal{U}) = 1 - \lambda(\mathcal{U}) = \frac{1}{2}$. Kolmogorov's zero–one law, applied to a measurable set that is invariant under finite changes, gives instead $\lambda(\mathcal{U}) \in \{0, 1\}$. Consequently, the set $\mathcal{U}$ is not measurable, and the proof of part (2) is complete.

**Determinacy kills ultrafilters.** We show that, if all of Ulam's games are determined, then every subset of $2^\omega$ has the Baire property; part (2) then leaves no room for a nonprincipal

ultrafilter. Assume that every game $\mathrm{Ul}(Z)$ is determined, and let $Z \subseteq 2^\omega$ and a finite binary string $s$ be given. Applying determinacy to the set $\{z \in 2^\omega : s^\frown z \in Z\}$ and the theorem of Banach and Oxtoby (see Theorem 3.5), we find a string $s' \supseteq s$ such that $Z$ is comeager in $N_{s'}$ or meager in $N_{s'}$. Let $O$ be the union of the sets $N_u$, for finite binary strings $u$, in which $Z$ is comeager, and let $O'$ be the union of those in which $Z$ is meager. The union $O \cup O'$ is then a dense open set, whose complement $2^\omega \setminus (O \cup O')$ is nowhere dense. The symmetric difference $Z \triangle O$ is covered by the sets $O \setminus Z$ and $O' \cap Z$ together with this complement, all of which are meager. Since $Z$ was arbitrary, every subset of $2^\omega$ has the Baire property, and by (2) there is no nonprincipal ultrafilter on $\omega$. The same argument applies to the sets of a pointclass, a family of sets of reals that is closed under continuous preimages (such as the class of analytic sets). In this form the argument leads to the remarks on definable ultrafilters at the end of this section.

**Selective ultrafilters.** The proof of part (4) uses a theorem of Kunen, recorded in the history above, which we quote without proof.

**Lemma 13.2 (Kunen).** *If $\mathcal{U}$ is a selective ultrafilter on $\omega$, then for every coloring of the two-element subsets of $\omega$ in two colors there is a set $H \in \mathcal{U}$ all of whose two-element subsets have the same color.*

The idea of the first half of part (4) is to color each pair according to whether the strategy of Player I still allows the larger element of the pair after the smaller one. By Lemma 13.2, some set in $\mathcal{U}$ is then homogeneous, that is, all pairs from this set have the same color. Player II can play such a set in full. Suppose that $\mathcal{U}$ is selective, and let $\sigma$ be a strategy for Player I in $S(\mathcal{U})$. For a finite set $s \subseteq \omega$, listed increasingly, let $A_s \in \mathcal{U}$ be the move of $\sigma$ after Player II has played the elements of $s$, with $A_s = \omega$ if these moves are not legal. For each $n \in \omega$, let $B_n$ be the intersection $\bigcap\{A_s : s \subseteq \{0, \dots, n-1\}\}$. As a finite intersection of members of $\mathcal{U}$, the set $B_n$ belongs to $\mathcal{U}$. We color a pair $m < n$ pink if $n \in B_{m+1}$, and purple otherwise. By Lemma 13.2 there is a homogeneous set $H \in \mathcal{U}$; replacing $H$ by $H \cap A_\emptyset$, which is still homogeneous and belongs to $\mathcal{U}$, we may assume that $H \subseteq A_\emptyset$. The color of $H$ is not purple: for $m = \min H$, the sets $H \setminus \{m\}$ and $B_{m+1}$ both belong to $\mathcal{U}$ and therefore meet. Any common point $n$ of these two sets yields a pink pair $\{m, n\}$ inside $H$. The color of $H$ is thus pink. We list $H$ increasingly as $h_0 < h_1 < \cdots$. Since $H \subseteq A_\emptyset$ and every pair in $H$ is pink, we have $h_0 \in A_\emptyset$, and $h_{k+1} \in B_{h_k+1} \subseteq A_{\{h_0,\dots,h_k\}}$ for every $k$. Hence Player II may answer $\sigma$ with $h_0, h_1, h_2, \dots$ in turn, and the set she produces is $H \in \mathcal{U}$. Consequently, the strategy $\sigma$ is not winning.

Conversely, suppose that $\mathcal{U}$ is not selective. Fix a partition $\{P_j : j \in J\}$ of $\omega$, indexed by a set $J$, into pieces outside $\mathcal{U}$ such that no element of $\mathcal{U}$ meets every piece in at most one point. Player I wins by always removing the pieces already visited. When Player II has played elements of $P_{j_0}, \dots, P_{j_{k-1}}$, Player I plays $\omega \setminus (P_{j_0} \cup \cdots \cup P_{j_{k-1}})$, which belongs to $\mathcal{U}$ because a finite union of sets outside an ultrafilter lies outside the ultrafilter. The set produced by Player II then meets each piece in at most one point, and the set is not in $\mathcal{U}$. For the P-point game we quote the corresponding characterization, whose proof uses finite choices and the P-point property rather than the homogeneous-set argument above. For the details we refer to the book of Bartoszyński and Judah [BJ95]. This completes the discussion of part (4).

Selectivity is exactly what the Ramsey theory of Section 11 needs. By the theorem of Mathias recalled in the history above, if $\mathcal{U}$ is selective, then every analytic $\mathcal{A} \subseteq [\omega]^\omega$ has a homogeneous set in $\mathcal{U}$: there is $H \in \mathcal{U}$ with $[H]^\omega \subseteq \mathcal{A}$ or $[H]^\omega \cap \mathcal{A} = \emptyset$ [Mat77]. In the other direction, every nonprincipal ultrafilter, selective or not, yields a subset of $[\omega]^\omega$ that is not Ramsey. For this fact we refer to the problem at the beginning of Section 11 and to Exercise 6 at the end of Section 11.

13.5. **Further remarks (optional).**

**Definable ultrafilters.** The proof of part (3) of Theorem 13.1 works one pointclass at a time. The *projective* sets are the sets of reals obtained from the Borel sets by finitely many applications of the operations of taking continuous images and complements. *Projective determinacy* is the statement that every game in which the players choose natural numbers and whose payoff set is projective is determined. By the theorem of Martin and Steel mentioned in Section 15 below, this statement is a consequence of the existence of infinitely many Woodin cardinals. Under projective determinacy, every projective set has the Baire property, so no nonprincipal ultrafilter is projective either. Without an extra hypothesis of this kind, the bound recorded after the main result, that no nonprincipal ultrafilter is analytic or coanalytic, cannot be extended to the next class of projective sets. This next class is that of the $\Delta^1_2$ sets, which we now define. A set of reals is $\Sigma^1_2$ if the set is a continuous image of a coanalytic subset of $\omega^\omega$, and $\Delta^1_2$ if both the set and the complement of the set are $\Sigma^1_2$. In Gödel's constructible universe, the canonical well-ordering of the reals is $\Sigma^1_2$, and a transfinite recursion along this well-ordering produces a $\Sigma^1_2$ nonprincipal ultrafilter $\mathcal{U}$. The ultrafilter $\mathcal{U}$ is even $\Delta^1_2$, since $x \notin \mathcal{U}$ if and only if $\omega \setminus x \in \mathcal{U}$.

**Meager filters.** The game argument shows that a nonprincipal ultrafilter is not meager. In the following theorem we characterize the meager filters, and thereby isolate the part of the ultrafilter property that is responsible for nonmeagerness.

**Theorem 13.3 (Jalali-Naini, 1976; Talagrand, 1980).** *A filter $\mathcal{F}$ on $\omega$ that contains every cofinite set is meager, as a subset of $2^\omega$, if and only if there is a partition of $\omega$ into consecutive finite intervals $I_0 < I_1 < \cdots$ (every element of $I_n$ being smaller than every element of $I_{n+1}$, for $n \in \omega$) such that every element of $\mathcal{F}$ meets all but finitely many of the $I_n$.*

*Proof.* One direction is a direct check. For the other, the idea is to build, from a sequence of closed nowhere dense sets covering $\mathcal{F}$, a single set $e \subseteq \omega$ together with a partition of $\omega$ into intervals. The points of the $k$-th of these closed sets are to disagree with $e$ on every later block of the partition, and these blocks are the intervals we need. If a partition as in the theorem exists, then the filter is contained in the union over $m \in \omega$ of the sets $\{x : x \cap I_n \neq \emptyset \text{ for all } n \geq m\}$. These sets are closed, and nowhere dense, because any finite string can be extended by zeros across a whole interval $I_n$. Suppose now that $\mathcal{F}$ is meager, and let $C_0 \subseteq C_1 \subseteq \cdots$ be closed nowhere dense sets whose union contains $\mathcal{F}$. We choose $0 = n_0 < n_1 < \cdots$ and binary strings $w_k$ of length $n_{k+1} - n_k$ such that $N_{u^\frown w_k} \cap C_k = \emptyset$ for every string $u$ of length $n_k$. This is possible because there are only finitely many such $u$, and we may extend a candidate for $w_k$ once for each of them, each time avoiding $C_k$. Let $I_k = [n_k, n_{k+1})$ and let $e \subseteq \omega$ be the set whose characteristic function is $w_0{}^\frown w_1{}^\frown \cdots$. Since $C_m \subseteq C_k$ for $k \geq m$, every point of $C_m$ disagrees with $e$ somewhere on $I_k$ for each $k \geq m$. Now let $x \in \mathcal{F}$ be given; the set $x \cup e$ then belongs to $\mathcal{F}$ and therefore lies in $C_m$ for some $m$. For every $k \geq m$, the sets $x \cup e$ and $e$ consequently differ somewhere in $I_k$, which means that $x$ meets $I_k$. Hence the intervals $I_k$ witness the condition of the theorem. ☺

A nonprincipal ultrafilter $\mathcal{U}$ fails the condition of Theorem 13.3 for every partition: the union of the even intervals and the union of the odd intervals are complementary. One of these two unions therefore belongs to $\mathcal{U}$, and this union misses infinitely many intervals. This is a second proof that $\mathcal{U}$ is not meager. By Theorem 13.3, nonmeagerness also survives finite-to-one images. Let $\mathcal{F}$ be a filter on $\omega$ that contains every cofinite set, let $f : \omega \to \omega$ be finite-to-one, and let $f(\mathcal{F}) = \{y \subseteq \omega : f^{-1}(y) \in \mathcal{F}\}$ be the image filter. If $f(\mathcal{F})$ is meager, witnessed by a partition $\{I_n\}$, then every $x \in \mathcal{F}$ meets $f^{-1}(I_n)$ for all but finitely many $n$,

because the image $f[x] = \{f(m) : m \in x\}$ belongs to $f(\mathcal{F})$. The sets $f^{-1}(I_n)$ are finite and pairwise disjoint, and infinitely many of them are non-empty. Hence any partition of $\omega$ into consecutive intervals each of which contains a non-empty set of the form $f^{-1}(I_n)$ witnesses that $\mathcal{F}$ is meager.

**Points that nobody can point at.** The space $\omega^*$ is the Stone space of the Boolean algebra $\mathcal{P}(\omega)/\mathrm{fin}$, the quotient of $\mathcal{P}(\omega)$ by the ideal of finite sets. The points of this Stone space, the ultrafilters of the algebra, correspond to the nonprincipal ultrafilters on $\omega$. Pospíšil showed in 1937 that there are $2^{2^{\aleph_0}}$ of them, using an independent family of $2^{\aleph_0}$ subsets of $\omega$ in the sense of Fichtenholz, Kantorovich and Hausdorff. Independence here means that every intersection of finitely many members of the family and of the complements of finitely many other members is infinite. Hence every choice, for each member, of either the member or the complement of the member extends to a nonprincipal ultrafilter, and distinct choices give distinct ultrafilters. A P-point is exactly a point $p \in \omega^*$ such that every $G_\delta$ subset of $\omega^*$ containing $p$ is a neighborhood of $p$. If $D \subseteq \omega^*$ is countable and $p$ lies in the closure of $D$ but not in $D$, then the complement $\omega^* \setminus D$ is a $G_\delta$ set containing $p$ without being a neighborhood of $p$. No homeomorphism of $\omega^*$ can therefore move a P-point to such a point. Points of this kind exist in ZFC, and P-points exist under the continuum hypothesis; this is Rudin's proof that $\omega^*$ is not homogeneous [Rud56]. In 1978 Kunen proved in ZFC that there are *weak P-points*, points of $\omega^*$ that are not limits of any countable subset of the rest of $\omega^*$ [Kun80]. Weak P-points give a proof that $\omega^*$ is not homogeneous without the continuum hypothesis. All of these points are produced by choice, and by the remarks on definable ultrafilters above, none of them can be produced by a definition of low complexity. The large-cardinal picture is sharper still, and involves $L(\mathbb{R})$, the smallest transitive model of ZF that contains all the ordinals and all the reals. Todorčević proved that, in the presence of large cardinals, the selective ultrafilters are exactly the filters that are generic over $L(\mathbb{R})$ for the partial order $\mathcal{P}(\omega)/\mathrm{fin}$ (see Farah [Far98]). A selective ultrafilter $\mathcal{U}$ is thus a generic object for the very Boolean algebra whose Stone space $\mathcal{U}$ inhabits. Every statement about $\mathcal{U}$ that holds in $L(\mathbb{R})[\mathcal{U}]$, the smallest such model that also contains $\mathcal{U}$, is therefore forced by one of the members of $\mathcal{U}$ itself. In the same way, every true statement about a Cohen real is forced by a finite initial segment of that real.

**Exercises.**

(1) Let $\mathcal{U}$ be an ultrafilter on $\omega$ that contains a finite set. Show that $\mathcal{U}$ is principal. Deduce that every nonprincipal ultrafilter $\mathcal{V}$ on $\omega$ contains every cofinite set, and that if $x \in \mathcal{V}$ and $y \subseteq \omega$ differs from $x$ in finitely many points, then $y \in \mathcal{V}$.

(2) For each of the following sets $Z \subseteq 2^\omega$, decide which player has a winning strategy in $\mathrm{Ul}(Z)$, and check the answer against Theorem 3.5: (a) the sequences with infinitely many terms equal to 1; (b) the sequences that are eventually 0; (c) the sequences $x$ with $x(n) = 1$ for infinitely many even $n$; (d) the sequences $x$ with $x(0) = 0$.

(3) Let $\mathcal{F}$ be the filter of cofinite subsets of $\omega$. Show that $\mathcal{F}$ is meager in $2^\omega$, first directly and then as a consequence of Theorem 13.3.

(4) Let $\mathcal{U}$ be a nonprincipal ultrafilter on $\omega$ such that Player I has no winning strategy in $S(\mathcal{U})$. Show, without using Theorem 13.1, that Player I has no winning strategy in $P(\mathcal{U})$ either. (Hint: Player II may play one-element sets in $P(\mathcal{U})$.)

(5) $(*)$ Let $\mathcal{U}$ be a nonprincipal ultrafilter on $\omega$ with the property in the conclusion of Lemma 13.2. Show that $\mathcal{U}$ is selective. (Hint: given a partition of $\omega$ into pieces outside $\mathcal{U}$, color a pair according to whether the two elements of the pair lie in the same piece.)

Part

# III

# Applications of other topological games

The last section of Part II ended with a game that is not determined, the game defined by a nonprincipal ultrafilter on $\omega$ (see Theorem 13.1). In this part the question of determinacy, whether one of the players has a winning strategy, becomes the theme. The players no longer shrink a set: in the Menger game they choose open covers and finite subfamilies of them. In the games of descriptive set theory, the players call out natural numbers, one at a time. What survives from Part I is the framework of Section 2 (see Definition 2.2), in which a game is a tree of positions together with a set of plays won by Player II.

**Roadmap**

A property of a space can be turned into a statement about one player of a game, and a stronger property into a statement about the other player. When this happens, the two properties coincide on a space exactly when the game on the space is determined. In Section 14 we follow this idea through Menger's conjecture: the Menger property and, for separable metrizable spaces, $\sigma$-compactness are equivalent to statements about the two players of the Menger game (see Theorem 14.1). The theorem of Fremlin and Miller shows that the game can be undetermined on a set of reals. In Section 15 determinacy changes roles, and becomes a tool rather than a question. There the games are played with natural numbers, and Martin's theorem says that every game with a Borel payoff set is determined (see Theorem 15.1). Wadge's comparison of two Borel sets then follows from the determinacy of a single game. In this game a winning strategy for either player yields a continuous reduction: of the first set to the second, or of the second to the complement of the first. The two sections are independent of each other.

## 14. Applications to set-theoretic topology

**The problem.** Compact sets are the "small" sets of topology, and countable unions of compact sets ($\sigma$-compact sets) are almost as tame. In 1924 Karl Menger proposed a covering property that looked like a clean characterization of $\sigma$-compactness. A topological space $X$ has the *Menger property* if for every sequence $\mathcal{U}_1, \mathcal{U}_2, \dots$ of open covers of $X$ there are finite sets $\mathcal{F}_n \subseteq \mathcal{U}_n$ such that $\bigcup_{n\geq 1} \mathcal{F}_n$ covers $X$. Every $\sigma$-compact space has the Menger property: if $X = \bigcup_{n\geq 1} K_n$ with each $K_n$ compact, it suffices to take for $\mathcal{F}_n$ a finite subset of $\mathcal{U}_n$ that covers $K_n$. Menger conjectured the converse for metric spaces. The problem is to decide Menger's conjecture, in particular for sets of real numbers.

14.1. **History of the problem and of the solution using topological games.** Menger introduced the property in 1924 [Men24], in a different but equivalent form, while studying a basis property of metric spaces. He conjectured that, among metric spaces, the property characterizes $\sigma$-compactness. Hurewicz gave the property the form above in a paper of 1926 [Hur26]. In the same paper he proved a theorem that is now stated as a characterization of the property by an infinite game. The game form is due to later authors, notably Telgársky and Scheepers. Hurewicz also introduced a stronger property, for which he made the analogous conjecture. Under the continuum hypothesis Menger's conjecture fails: a Luzin set, an uncountable set of reals meeting every meager set in a countable set, has the Menger property but is not $\sigma$-compact. In 1984 Telgársky characterized $\sigma$-compactness of metrizable spaces by the same game [Tel84]. Menger's conjecture thereby became the statement that the Menger game is determined on every set of reals.

In 1988 David Fremlin and Arnold Miller solved the problem [FM88]: they proved, without any hypothesis beyond ZFC (Zermelo–Fraenkel set theory with the axiom of choice), that there is a set of reals with the Menger property that is not $\sigma$-compact. The closely related Rothberger property (1938), in which a single member of each cover is chosen, has a parallel story in terms of games. In 1978 Galvin introduced the point-open game [Gal78], in which one player chooses points and the other chooses open sets containing them (see the next subsection). In 1994 Pawlikowski proved that the point-open game on a set of reals is undetermined if and only if the set is uncountable and has the Rothberger property [Paw94]. A set of reals has *strong measure zero* if, for every sequence of positive lengths, the set can be covered by a sequence of intervals with those lengths. Borel conjectured in 1919 that every set of reals of strong measure zero is countable; Laver proved in 1976 that the conjecture is consistent [Lav76]. Since the Rothberger property implies strong measure zero, every set of reals with the Rothberger property is countable in the model of set theory that Laver constructed. By the characterization of 1994, the point-open game is determined in Laver's model on every set of reals. Under the continuum hypothesis, on the other hand, there are Luzin sets, and by a theorem of Recław [Rec94] the point-open game on every Luzin set is undetermined.

14.2. **The games.** The *Menger game* on a topological space $X$ is played in rounds $n = 1, 2, \dots$ (see Figure 20). In round $n$, Player I chooses an open cover $\mathcal{U}_n$ of $X$, and Player II chooses a finite set $\mathcal{F}_n \subseteq \mathcal{U}_n$. Player II wins if $\bigcup_n \mathcal{F}_n$ covers $X$. In the *Rothberger game* Player II is more frugal still: she chooses a single member of each cover, and she wins if the chosen sets cover $X$. The *point-open game* reverses the winning condition. In round $n$, Player I chooses a point $x_n \in X$ and Player II an open set $U_n \ni x_n$. As in the Menger and Rothberger games, Player II still chooses the open sets, but she now wins if the sets $U_n$ do

not cover $X$. The point-open game is dual to the Rothberger game [Gal78]: Player I has a winning strategy in either of the two games exactly when Player II has one in the other. In the literature Player I is called ONE and Player II is called TWO.

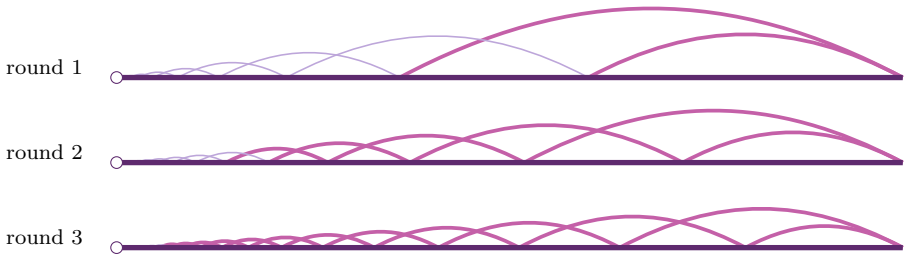



FIGURE 20. Three rounds of the Menger game on the half-open interval $X = (0, 4]$, with the covers of Player I drawn as arcs and the intervals that Player II keeps in pink.

14.3. **The main result of this section.** Hurewicz and Telgársky turned the Menger property and $\sigma$-compactness into statements about the two players of the Menger game. Fremlin and Miller proved that the gap between these two statements is real.

**Theorem 14.1 (Hurewicz, 1926; Telgársky, 1984; Fremlin–Miller, 1988).**

(1) *A topological space $X$ has the Menger property if and only if Player I has no winning strategy in the Menger game on $X$.*
(2) *A separable metrizable space $X$ is $\sigma$-compact if and only if Player II has a winning strategy in the Menger game on $X$.*
(3) *There is a set of reals with the Menger property that is not $\sigma$-compact. Consequently, the Menger game on this set is not determined, and Menger's conjecture is false.*

14.4. **Solving the problem with the games.**

**The Menger property.** One direction of part (1) of Theorem 14.1 is immediate. Suppose that Player I has no winning strategy, and let open covers $\mathcal{U}_1, \mathcal{U}_2, \ldots$ of $X$ be given. The strategy "play $\mathcal{U}_n$ in round $n$" is then not winning: some play against this strategy is won by Player II, and her moves witness the Menger property. The converse is the substantial direction, and we give only an outline. Applied to a constant sequence of covers, the Menger property shows that $X$ is Lindelöf (see Exercise 2 at the end of this section). In other words, every open cover of $X$ has a countable subcover. Player II may therefore restrict herself, at each position, to a fixed countable subcover of the cover that Player I's strategy plays. Only countably many finite positions can then arise, and the strategy plays only countably many covers. The Menger property, applied to a suitable arrangement of these countably many covers, yields a play that defeats the strategy. We omit the details, which go back to Hurewicz [Hur26].

**$\sigma$-compactness.** If $X = \bigcup_{n\geq 1} K_n$ with $K_n$ compact and increasing, Player II has a winning strategy: in round $n$, she chooses a finite subfamily of $\mathcal{U}_n$ that covers $K_n$. Conversely, Telgársky unravels a winning strategy of Player II, position by position, into countably many compact sets that cover $X$ [Tel84].

**Menger's conjecture.** Among separable metrizable spaces, consider those that have the Menger property but are not $\sigma$-compact. By parts (1) and (2) of Theorem 14.1, these are exactly the separable metrizable spaces on which neither player has a winning strategy in the Menger game. Fremlin and Miller obtain such a set of reals from combinatorial properties of families of functions from $\omega$ to $\omega$ under eventual domination, in an argument that splits into cases [FM88]. A function $g$ *eventually dominates* a function $f$ if $f(n) \leq g(n)$ for all but finitely many $n \in \omega$. A family of functions from $\omega$ to $\omega$ is *dominating* if every such function is eventually dominated by a member of the family. The link with the Menger property is a later theorem of Hurewicz, from 1927, about the space $\omega^\omega$ of these functions with the product topology (see Section 15). This theorem asserts that a subset of $\omega^\omega$ has the Menger property exactly when no continuous image of the subset in $\omega^\omega$ is a dominating family. Seen as a statement about games, Menger's conjecture is attractive and false for one and the same reason: the conjecture asserts a form of determinacy. On analytic sets of reals, those that are empty or continuous images of $\omega^\omega$, the Menger game is determined, since an analytic set with the Menger property is $\sigma$-compact. On wild sets the game need not be determined, and Fremlin and Miller proved that a wild enough set exists without any extra axioms.

**Exercises.**

(1) Let $X$ be a countable space. Show that Player II has a winning strategy in the Rothberger game on $X$, and that Player I has a winning strategy in the point-open game on $X$.
(2) Show that a space with the Menger property is Lindelöf. Deduce that an uncountable discrete space does not have the Menger property, and describe a winning strategy for Player I in the Menger game on such a space.
(3) Show that the Menger property is inherited by closed subspaces and preserved by continuous surjections.
(4) Show that Player II has a winning strategy in the Menger game on $[0, 1]$, but that Player I has a winning strategy in the Rothberger game on $[0, 1]$. (Hint: in round $n$, let Player I cover $[0, 1]$ by relatively open intervals of length at most $2^{-n-2}$; a finite family of intervals covering $[0, 1]$ has total length at least 1.)
(5) $(*)$ Show that the space $\omega^\omega$ of infinite sequences of natural numbers, with the product topology, does not have the Menger property, and hence is not $\sigma$-compact. (Hint: let $\mathcal{U}_n$ consist of the sets $\{x \in \omega^\omega : x(n) = k\}$ for $k \in \omega$.) Deduce from Theorem 14.1(1) that Player I has a winning strategy in the Menger game on $\omega^\omega$, and find one directly.

## 15. Applications to descriptive set theory

The games of this section are played with natural numbers, and their payoff sets are sets of infinite sequences of natural numbers, which descriptive set theory sorts by their complexity. We begin by recalling the space in which these sets live and the classes into which they are sorted.

**Definitions.** The space $\omega^\omega$ of infinite sequences of natural numbers carries the product topology. For a finite sequence $s$ of natural numbers, the basic open set $N_s = \{x \in \omega^\omega : s \subseteq x\}$ consists of the sequences that begin with $s$. A set is open if the set is a union of basic open sets. (By tradition the space $\omega^\omega$ is called "the Baire space". The name is not to be confused with the notion of a Baire space in Definition 1.1, although the space $\omega^\omega$, being completely metrizable, is a Baire space in that sense as well.) The *Borel sets* of a topological space form the smallest family of subsets of the space that contains the open sets and is closed under complements and countable unions. They are arranged in levels: the open sets and

the closed sets form the first level, the $F_\sigma$ sets and the $G_\delta$ sets the second, and so on through the countable ordinals. Here an $F_\sigma$ set is a countable union of closed sets, and a $G_\delta$ set is a countable intersection of open sets. The level at which a Borel set first appears is the *Borel rank* of the set. A subset of a separable completely metrizable space is *analytic* if the subset is empty or is a continuous image of $\omega^\omega$. Among such spaces are $\omega^\omega$, the Cantor space $2^\omega$ of infinite binary sequences, and $\mathbb{R}$.

**The problem.** Here is the simplest infinite game imaginable: fix a set $E \subseteq \omega^\omega$ of infinite sequences of natural numbers. In the *Gale–Stewart game* $G(E)$, Player I and Player II take turns calling out natural numbers, Player I playing $x(0), x(2), \dots$ and Player II playing $x(1), x(3), \dots$. Player II wins if the resulting sequence $x$ belongs to $E$. In the language of Definition 2.2, the game $G(E)$ is the pair $(\omega^{<\omega}, E)$, with moves in $A = \omega$. Recall that a game is determined if one of the players has a winning strategy (see Remarks 2.3(4)). Finite games are always determined, by Zermelo's theorem. In chess, where draws are possible, the same argument shows that one of the players can force a win or both can force a draw, although nobody knows which. For infinite games determinacy can fail: using the axiom of choice one can build sets $E$ for which neither player has a winning strategy in $G(E)$. A nonprincipal ultrafilter $\mathcal{U}$ on $\omega$ gives such a set. In Ulam's game $\mathrm{Ul}(\mathcal{U})$ of Section 13, which is not determined (see Theorem 13.1), the players alternately choose non-empty finite binary strings. Player II wins if the concatenation of the strings is the characteristic function of a set in $\mathcal{U}$. Once the countably many non-empty finite binary strings are numbered, Ulam's game becomes a Gale–Stewart game, since calling out a number then means playing a string. The problem is to determine which sets, and in particular which Borel sets, give determined games, and what the answer reveals about the structure of the Borel sets.

15.1. **History of the problem and of the solution using infinite games.** The story of this problem is unusual: the game came first, and the applications came later. Gale and Stewart introduced these games in 1953 [GS53]. They proved that $G(E)$ is determined when $E$ is open or closed, constructed an undetermined game from a well-ordering of the reals, and asked how far determinacy extends. Wolfe proved in 1955 that games with $F_\sigma$ or $G_\delta$ payoff sets are determined [Wol55], and Davis extended the result in 1964 to the next level of the Borel hierarchy [Dav64]. In 1970 Martin proved that analytic games are determined if there is a measurable cardinal [Mar70]. In 1971 Friedman showed that any proof of Borel determinacy must use uncountably many iterations of the power set operation [Fri71]. The problem therefore lay beyond the methods of ordinary analysis. That a statement about sets of real numbers should need, for its proof, sets of sets of sets of ... of reals, iterated uncountably many times, was something new.

In 1975 Donald A. Martin solved the problem [Mar75]: every Borel game is determined. His proof uses iterated power sets of the reals of every countable length, and by Friedman's theorem no proof can do with less. Even before Martin's proof, Wadge had compared sets of reals through a game of his own. In work of 1972, written up in his thesis of 1983 [Wad83], he showed that determinacy makes the Borel sets almost linearly ordered under continuous reducibility. Martin and Monk showed that this ordering is well-founded. Determinacy was later extended far beyond the Borel sets: Martin and Steel proved in 1989 that projective games are determined if there are infinitely many Woodin cardinals [MS89]. Projective games are those whose payoff sets are obtained from the Borel sets by finitely many applications of the operations of taking continuous images and complements.

*Recognition.* Martin's proof appeared in the *Annals of Mathematics*. Martin, Steel and Woodin received the Karp Prize of the Association for Symbolic Logic in 1988 for their work on determinacy. The proof of projective determinacy by Martin and Steel appeared in the *Journal of the American Mathematical Society*.

15.2. **The games.** Besides the Gale–Stewart game $G(E)$ of the problem, we need a game for comparing two sets $E, F \subseteq \omega^\omega$. We use a synchronous version of the *Wadge game*, with no passing moves. In this game $G_{\mathrm{W}}(E, F)$, Player I plays $x(0), x(1), \dots$ and Player II plays $y(0), y(1), \dots$ alternately, with Player I first, and Player II wins if

$$x \in E \quad \text{if and only if} \quad y \in F.$$

Player II thus tries to imitate Player I: whenever his sequence lands in $E$, hers should land in $F$, and vice versa. The play is the single sequence $(x(0), y(0), x(1), y(1), \dots)$; the Wadge game is therefore a Gale–Stewart game, whose payoff set is Borel when $E$ and $F$ are. A strategy of Player II turns the sequence of Player I into her own, one term at a time (see Figure 21).

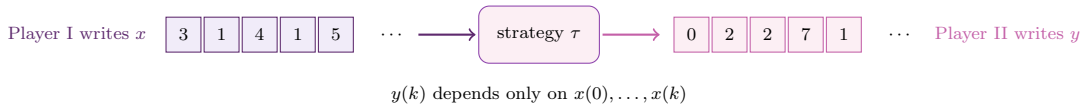



FIGURE 21. A strategy $\tau$ of Player II in the Wadge game $G_{\mathrm{W}}(E, F)$, for sets $E, F \subseteq \omega^\omega$, drawn as a machine that reads the sequence $x$ of Player I and writes her sequence $y$ one term at a time.

15.3. **The main result of this section.** The main result is Martin's theorem, together with Wadge's consequence for the comparison of Borel sets. For $E, F \subseteq \omega^\omega$ write $E \leq_W F$, and say that $E$ is *Wadge reducible* to $F$, if there is a continuous $f\colon \omega^\omega \to \omega^\omega$ with $E = f^{-1}(F)$.

**Theorem 15.1 (Martin, 1975; Wadge, 1972).**

(1) Borel determinacy. *For every Borel set $E \subseteq \omega^\omega$, the Gale–Stewart game $G(E)$ is determined.*
(2) Wadge's lemma. *For all Borel sets $E, F \subseteq \omega^\omega$, either $E \leq_W F$ or $F \leq_W \omega^\omega \setminus E$.*

By part (2), any two Borel sets are comparable under $\leq_W$, up to complements. Martin and Monk showed that $\leq_W$ is also well-founded on the Borel sets. Two sets $E, F \subseteq \omega^\omega$ are called *Wadge equivalent* when $E \leq_W F$ and $F \leq_W E$. Comparability and well-foundedness together show that the Borel sets, up to Wadge equivalence, form a well-founded hierarchy that refines the classical Borel hierarchy and is almost linear. More precisely, two Borel sets that are not comparable under $\leq_W$ are, up to Wadge equivalence, a set and the complement of the set. Every antichain of Borel sets under $\leq_W$ therefore has at most two elements. The position of a Borel set in this hierarchy measures the complexity of the set far more finely than the Borel rank does.

15.4. **Solving the problem with the games.**

**Open games.** The idea is the maxim "never step into a position from which you are already lost", and the openness of $E$ guarantees that this cautious play is enough. Suppose that $E$ is open and that Player II has no winning strategy in $G(E)$. We shall call a finite position *good for Player I* if Player II has no winning strategy from this position; by hypothesis, the empty position is good. If a position with Player I to move is good for him, some move leads to a good position. Otherwise Player II would have a winning strategy from every successor position and hence from the position itself. If a position with Player II to move is good for Player I, so is every position reached from this position in one move. Indeed, a move of Player II to a position that is not good would give her a winning strategy from the position she moved from. Player I can therefore play in such a way that every position of the play is good; let $x$ be a play in which he does so. If $x$ were in $E$, the openness of $E$ would provide an initial segment $s$ of $x$ with $N_s \subseteq E$. Player II would then win from the position $s$ whatever the players did, and this position would not be good for Player I. Consequently, the play $x$ lies outside $E$: the rule "move to a good position" is a winning strategy for Player I, and the game is determined.

**Borel games.** Martin's proof [Mar75] reduces every Borel game to a clopen game on a larger tree. In outline, for each Borel set $E$ he builds a game $G'$ whose moves are more complicated objects than natural numbers. Along with the game $G'$ he constructs a continuous map from plays of $G'$ to plays of $G(E)$ that sends strategies to strategies. The payoff set of $G'$, the preimage of $E$ under this map, is clopen. The argument for open games works on any tree, and the game $G'$ is therefore determined; the transfer of strategies makes $G(E)$ determined as well. The construction proceeds by induction on the Borel rank of $E$, and at each step the moves of the new game are sets of positions of the old one. For this reason Martin's proof climbs through the iterated power sets of $\mathbb{R}$, in agreement with Friedman's theorem.

**Wadge's lemma.** Let $E, F \subseteq \omega^\omega$ be Borel sets. The payoff set of the Wadge game $G_{\mathrm{W}}(E, F)$ is then Borel, and by Borel determinacy one of the players has a winning strategy in this game. If Player II has one, her strategy is a blueprint for a continuous function. Player I feeds in an input sequence $x$ one term at a time, and Player II's step-by-step answers are the output $y$. Each term of the output depends only on finitely many terms of the input, so the map $x \mapsto y$ is continuous. Since the strategy is winning, the input $x$ lies in $E$ if and only if the output $y$ lies in $F$, whence $E \leq_W F$. If Player I has one, his strategy turns each sequence $y$ of moves of Player II continuously into a sequence $x$, with $y \in F$ if and only if $x \notin E$; this map witnesses $F \leq_W \omega^\omega \setminus E$. The game thus produces the reduction itself, and the comparability of Borel sets is literally the determinacy of a game.

**Exercises.**

(1) For each of the following sets $E \subseteq \omega^\omega$, decide which player has a winning strategy in $G(E)$, and describe one: (a) the sequences with $x(1) > x(0)$; (b) the sequences with $x(1) < x(0)$; (c) the sequences for which $\{x(2n) : n \in \omega\}$ is finite; (d) the sequences with $x(2n+1) = x(0)$ for infinitely many $n$.

(2) Show that $\leq_W$ is reflexive and transitive, and that $E \leq_W F$ implies $\omega^\omega \setminus E \leq_W \omega^\omega \setminus F$.

(3) Let $E \subseteq \omega^\omega$ be closed. Show that $G(E)$ is determined, by adapting the argument for open games with the roles of the players exchanged. (Hint: call a position good for Player II if Player I has no winning strategy from the position.)

(4) Let $E$ be the set of $x \in \omega^\omega$ with $x(n) = 0$ for some $n$, and let $F = N_{(0)}$ be the set of sequences beginning with 0. Show that Player I has a winning strategy in $G_{\mathrm{W}}(E, F)$, and that $E$ is not Wadge reducible to $F$. Find a continuous $f$ with $F = f^{-1}(\omega^\omega \setminus E)$, as Theorem 15.1(2) predicts.

(5) (∗) Let $E \subseteq \omega^\omega$ be open but not closed. Show that $F \leq_W E$ for every open set $F \subseteq \omega^\omega$. (Hint: fix $z \in \overline{E} \setminus E$; in $G_{\mathrm{W}}(F, E)$, let Player II copy $z$ until an initial segment $p$ of Player I's sequence satisfies $N_p \subseteq F$.)

Part

# IV

# Applications of non-topological games

The games of Part III are infinite, and whether they are determined is a deep question: Martin's theorem settles the question for Borel payoff sets, while the axiom of choice provides games that are not determined. The games of this part are finite, played on boards and on families of finite sets. For these games, determinacy is Zermelo's elementary theorem, recalled in the problem at the beginning of Section 15. The spirit of the earlier parts is nevertheless at work here, far from topology: a global question becomes a local one, what to play next.

**Roadmap**

In this part we prove two classical theorems by games. In each of them, a statement about a large family of objects (all two-colorings of a set, all points of a square) is reduced to a statement about a finite game. In Section 16 a random coloring shows that a good coloring exists. There, the potential function of Erdős and Selfridge takes the place of the random choice and builds such a coloring one element at a time (see Theorem 16.1). In Section 17 we show that the board game Hex cannot end in a draw; from this fact about finite boards we derive the Brouwer fixed point theorem in dimension two (see Theorem 17.1). The two sections can be read in any order, and neither needs more than Section 2 from Part I.

## 16. Applications to the probabilistic method

**The problem.** The probabilistic method proves that combinatorial objects exist by showing that a random object has the required property with positive probability. Consider, for example, a positive integer $n$ and a family $\mathcal{F}$ of $n$-element sets with fewer than $2^{n-1}$ members. Erdős proved in 1963 that the elements of these sets can be colored red and blue so that no member of $\mathcal{F}$ is monochromatic. Indeed, a random coloring makes a given member of $\mathcal{F}$ monochromatic with probability $2^{1-n}$; the expected number of monochromatic members is therefore less than 1, and some coloring has none. Proofs of this kind give no construction; the problem is to turn them into explicit, deterministic, efficient constructions.

16.1. **History of the problem and of the solution using combinatorial games.** The probabilistic method goes back to a 1943 paper of Szele on tournaments. Erdős made the method famous in 1947 with his lower bound for diagonal Ramsey numbers [Erd47]. The coloring theorem above is from 1963 [Erd63]. In 1973 Erdős and Selfridge studied the games now called Maker–Breaker games [ES73]. In these games the players alternately claim elements of a board, and Maker tries to claim all the elements of some winning set. Erdős and Selfridge gave a strategy for Breaker based on a potential function. This potential is the expected number of winning sets that Maker would complete if each unclaimed element were given to Maker or to Breaker by the toss of a fair coin. Breaker, they showed, can play so that this expectation never grows over a move of Breaker followed by a move of Maker. With one player instead of two, the same potential argument solves the problem for the coloring theorem and related results. The single player colors the elements one at a time and never lets the expected number of monochromatic sets reach 1.

The argument was recognized as a general derandomization technique, the method of conditional probabilities. Spencer systematized the method in 1987 [Spe87]. Raghavan extended the method in 1988 [Rag88] with pessimistic estimators: efficiently computable upper bounds on the conditional probabilities of failure. These estimators take the place of the conditional probabilities when the latter are too costly to compute. Beck built a theory of positional games on the Erdős–Selfridge theorem and presented the theory in his book of 2008 [Bec08].

*Recognition.* Erdős received the Wolf Prize in 1983/84.

16.2. **The game.** Let $\mathcal{F}$ be a family of subsets of a finite set $V$. In the *Maker–Breaker game* on $(V, \mathcal{F})$, Maker is Player I and Breaker is Player II. The players alternately claim unclaimed elements of $V$, Player I first, until all elements are claimed. Player I wins if he has claimed all the elements of some $A \in \mathcal{F}$, and Player II wins otherwise. Tic-tac-toe is the example to keep in mind: the set $V$ is the board, and the family $\mathcal{F}$ is the set of lines. The Maker–Breaker game differs from tic-tac-toe in that Breaker wins merely by preventing Maker from completing a line. To define lines in general, let $n \geq 2$ and $d \geq 1$ be integers. On the board $\{1, \dots, n\}^d$, a *line* is a set $\{c_1, \dots, c_n\}$ of cells subject to two conditions. First, each of the $d$ coordinates is either the same in all the cells, or equal to $m$ in $c_m$ for $1 \leq m \leq n$, or equal to $n + 1 - m$ in $c_m$ for $1 \leq m \leq n$. Second, at least one coordinate is not the same in all the cells. Ordinary tic-tac-toe is the case $n = 3$, $d = 2$.

16.3. **The main result of this section.** The main result is the Erdős–Selfridge theorem, together with two classical consequences: a derandomized coloring and a result on multidimensional tic-tac-toe.

**Theorem 16.1 (Erdős–Selfridge, 1973).** *Let $\mathcal{F}$ be a family of subsets of a finite set $V$.*

(1) *If $\sum_{A \in \mathcal{F}} 2^{-|A|} < \frac{1}{2}$, then in the Maker–Breaker game on $(V, \mathcal{F})$ in which Maker moves first, Breaker has a winning strategy.*
(2) *If $\sum_{A \in \mathcal{F}} 2^{1-|A|} < 1$, then a coloring of $V$ in two colors with no monochromatic member of $\mathcal{F}$ can be computed in time polynomial in $|V|$ and $|\mathcal{F}|$.*
(3) *If $n \geq 2$ and $d \geq 1$ are integers with $\frac{1}{2}\big((n+2)^d - n^d\big) < 2^{n-1}$, then the first player cannot force a win in tic-tac-toe on the board $\{1, \dots, n\}^d$, the game in which the players alternately claim cells and the first to claim all the cells of a line wins.*

Part (3) follows from part (1) because the board (see Figure 22 for the case $n = 3$, $d = 2$) has $\frac{1}{2}\big((n+2)^d - n^d\big)$ lines of $n$ cells each. Indeed, in tic-tac-toe on this board the second

player uses Breaker's strategy and occupies a cell of every line. By the Hales–Jewett theorem (1963), on the other hand, the first player wins when $d$ is large compared with $n$ [HJ63].

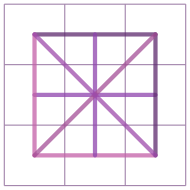


FIGURE 22. The eight winning lines of ordinary tic-tac-toe, each of three cells.

16.4. **Solving the problem with the game.** The proof of Theorem 16.1(1) turns the expectation of the problem into a potential function that Player II never lets grow. At any stage of a play of the Maker–Breaker game on $(V, \mathcal{F})$, we call $A \in \mathcal{F}$ *alive* if Player II has claimed none of the elements of $A$. We denote by $u(A)$ the number of unclaimed elements of a set $A \in \mathcal{F}$ at that stage, and for $v \in V$ we define

$$\Phi = \sum_{A \text{ alive}} 2^{-u(A)}, \qquad w(v) = \sum_{A \text{ alive}, v \in A} 2^{-u(A)}.$$

If the remaining elements were assigned to the two players by fair coin tosses, the quantity $\Phi$ would be the expected number of sets completed by Player I. A claim of $v$ by Player I increases $\Phi$ by $w(v)$, and a claim of $v$ by Player II decreases $\Phi$ by $w(v)$. By hypothesis the potential starts below $\frac{1}{2}$, and after the first move of Player I the potential is below 1 (see Exercise 1 at the end of this section). Player II's strategy is to claim, at each move, an unclaimed element $v$ of maximal weight $w(v)$. Her claim of an element kills the alive sets that contain the element and leaves the other terms unchanged; hence the claim increases no weight. Player I's next claim therefore raises $\Phi$ by at most the weight that Player II has just removed. Over each pair of moves the potential does not increase, and therefore stays below 1. If Player I completed a set $A$, the term of $A$ in $\Phi$ would be $2^0 = 1$, which is impossible since $\Phi < 1$ and all terms are nonnegative; Breaker thus wins. For part (2) one player colors the elements one at a time. At each step the player chooses a color that does not increase the expected number of monochromatic sets (such a color exists: see Exercise 4). The expectation starts below 1, never increases, and at the end is an integer, hence 0. Each conditional expectation is a sum of $|\mathcal{F}|$ explicit terms, and the whole computation takes time polynomial in $|V|$ and $|\mathcal{F}|$. The proof of part (2) is more than an existence proof: the proof is an explicit algorithm, short enough to write in a few lines of Python.

The game adds robustness to the probabilistic method. A probabilistic argument shows that a random play is good on average; a strategy forces a good outcome one move at a time, even against an adversary who chooses half of the moves. Forcing the outcome move by move is what makes the argument constructive, and the simplicity of each move is what makes the argument efficient.

**Exercises.**

(1) Check the bookkeeping in the proof of Theorem 16.1(1): a claim of $v$ by Player I increases $\Phi$ by exactly $w(v)$, and a claim of $v$ by Player II decreases $\Phi$ by exactly $w(v)$. Check further that, under the hypothesis of part (1), we have $\Phi < 1$ after the first move of Player I.

(2) Show that the board $\{1,\dots,n\}^2$ has $2n+2 = \frac{1}{2}\big((n+2)^2 - n^2\big)$ lines, and find the least $n$ for which Theorem 16.1(3) applies with $d=2$. Do the same for $d=3$, where the number of lines is $3n^2+6n+4$.

(3) Consider the Maker–Breaker game on the $3\times 3$ tic-tac-toe board, with $\mathcal{F}$ the set of the 8 lines of the board. Show that $\sum_{A\in\mathcal{F}} 2^{-|A|} = 1$, so Theorem 16.1(1) does not apply, and show that Maker has a winning strategy. (Hint: Maker takes the center and then creates two threats at once.) Explain why Maker's win does not contradict the fact that ordinary tic-tac-toe is a draw.

(4) Let $\mathcal{F}$ be a family of subsets of a finite set $V$ with $\sum_{A\in\mathcal{F}} 2^{1-|A|} < 1$, as in Theorem 16.1(2). Explain why, in the proof of Theorem 16.1(2), one of the two colors for the next element never increases the conditional expected number of monochromatic members of $\mathcal{F}$.

(5) (∗) Let $n \geq 1$ be an integer, and let $V$ be the set of vertices of a complete binary tree with $n$ levels. Let $\mathcal{F}$ be the family of the $2^{n-1}$ paths from the root to a leaf. Show that $\sum_{A\in\mathcal{F}} 2^{-|A|} = \frac{1}{2}$ and that Maker has a winning strategy. Thus the strict inequality in Theorem 16.1(1) cannot be weakened.

## 17. Applications to fixed point theory

**The problem.** The *Brouwer fixed point theorem* says that, for every positive integer $n$, every continuous map $f\colon [0,1]^n \to [0,1]^n$ has a fixed point. The original proofs used algebraic topology. The problem is to find an elementary, combinatorial proof that exposes the finite core of the theorem and, ideally, yields a method for locating approximate fixed points.

17.1. **History of the problem and of the solution using combinatorial games.** The theorem was proved in general by Hadamard in 1910 and Brouwer in 1912 [Bro12], with the methods of the new algebraic topology. A combinatorial route was opened by Sperner's lemma of 1928 on labeled triangulations [Spe28], from which Knaster, Kuratowski and Mazurkiewicz derived the fixed point theorem in 1929. Scarf turned this line of thought into an algorithm for computing approximate fixed points in 1967 [Sca67]. Meanwhile the game of Hex was invented by Piet Hein in 1942 and again, apparently independently, by John Nash at Princeton in 1948 or 1949. Nash observed that the game cannot end in a draw, and he proved with a strategy-stealing argument that the first player has a winning strategy. A winning strategy for the second player could be adopted by the first, since an extra stone never hurts its owner, and so the second player has none; in a finite game without draws, the first player must then have a winning strategy (see Exercise 4 at the end of this section).

David Gale found the game-theoretic solution in 1979 [Gal79]. He proved that the Brouwer fixed point theorem is an easy consequence of the fact that Hex cannot end in a draw, and he extended both statements to $n$ dimensions.

*Recognition.* Gale's paper, in the *American Mathematical Monthly*, received a Lester R. Ford Award of the Mathematical Association of America in 1980, the year in which Gale shared the John von Neumann Theory Prize. Nash received the Nobel Memorial Prize in Economic Sciences in 1994 and the Abel Prize in 2015.

17.2. **The game.** Hex is played on a rhombus-shaped board of hexagonal cells. Anyone who has played the game knows the feeling that somebody *must* win: a complete blockade of one player seems to be, at the same time, a connection for the other. To make this precise, we fix an integer $k \geq 1$ and describe the board by the centers of the cells, the points $(i,j)$ with

integers $1 \le i, j \le k$. The cell $(i, j)$ is adjacent to those of the cells $(i \pm 1, j)$, $(i, j \pm 1)$, $(i+1, j+1)$ and $(i-1, j-1)$ that lie on the board. Player I owns the bottom and top rows (where $j = 1$ and $j = k$), and Player II the left and right columns (where $i = 1$ and $i = k$). The players alternately place stones on empty cells, Player I first. A *chain* is a sequence of cells, each adjacent to the next. Player I wins if his stones contain a chain from the bottom row to the top row; Player II wins if hers contain one from the left column to the right column.

17.3. **The main result of this section.** The main result is the Hex theorem and, as a consequence, the fixed point theorem in dimension two.

**Theorem 17.1 (Gale, 1979).**

(1) The Hex theorem. *If $k \ge 1$ is an integer and every cell of the $k \times k$ Hex board is occupied by a stone of Player I or of Player II, then exactly one of the players has a chain of his or her stones joining his or her two sides of the board.*
(2) *The Hex theorem implies the Brouwer fixed point theorem for $[0,1]^2$.*

Part (1) has an $n$-dimensional counterpart, which Gale also proved [Gal79]. For integers $n \ge 2$ and $k \ge 1$, the counterpart concerns $n$-dimensional Hex, played by $n$ players on the grid $\{1, \dots, k\}^n$, and asserts that a full board has at least one winner. This statement implies the Brouwer fixed point theorem for $[0,1]^n$.

17.4. **Solving the problem with the game.**

**No draws.** Gale's idea is to follow the boundary between the stones of the two players, which we picture in two colors, until the boundary can go no further. We add an extra row of Player I's stones above and below the board, and an extra column of Player II's stones on either side. Starting at a corner of the enlarged board, between the two colors, we walk along the edges of the hexagons that separate a stone of Player I from one of Player II. At each vertex of the walk three cells meet, two of one color and one of the other; the walk therefore continues in exactly one way. Moreover, the walk never returns to a vertex, since a vertex already visited has no unused separating edge. The board is finite, so the walk ends, and the walk can end only at another corner of the enlarged board. The stones of one color along one side of the walk then form a winning chain. The two players cannot both win: a bottom-to-top chain and a left-to-right chain would have to cross, and on the hexagonal board crossing chains share a cell. The walk on a full board is shown in Figure 23.

**Brouwer's theorem.** The idea is that a winning chain must cross from a side where the map pushes points one way to a side where the map pushes them the other way, and continuity forbids such a reversal between neighboring cells. Suppose that a continuous map $f = (f_1, f_2)\colon [0,1]^2 \to [0,1]^2$ has no fixed point. For points $z = (z_1, z_2)$ and $z' = (z_1', z_2')$ of $[0,1]^2$, we write $|z - z'|_\infty = \max(|z_1 - z_1'|, |z_2 - z_2'|)$. By compactness there is $\varepsilon > 0$ with $\max(|f_1(z) - z_1|, |f_2(z) - z_2|) \ge \varepsilon$ for all $z \in [0,1]^2$, and by uniform continuity there is $\delta$ with $0 < \delta < \varepsilon$ such that $|f_1(z) - f_1(z')| < \varepsilon$ and $|f_2(z) - f_2(z')| < \varepsilon$ whenever $|z - z'|_\infty \le \delta$. We choose an integer $k \ge 2$ with $1/(k-1) \le \delta$ and place at each cell $(i, j)$ the point $z = ((i-1)/(k-1), (j-1)/(k-1))$. We give the cell $(i, j)$ to Player II if $|f_1(z) - z_1| \ge \varepsilon$, and to Player I otherwise, in which case $|f_2(z) - z_2| \ge \varepsilon$. According to the Hex theorem, one of the players has a winning chain. Suppose first that Player II has a chain from left to right. On the left side we have $z_1 = 0 \le f_1(z)$, whence $f_1(z) - z_1 \ge \varepsilon$; on the right side, the inequality $z_1 = 1 \ge f_1(z)$ forces $z_1 - f_1(z) \ge \varepsilon$. The sign of $f_1(z) - z_1$ must change along the chain, so two adjacent cells of the chain, carrying points $z$ and $z'$, satisfy $f_1(z) - z_1 \ge \varepsilon$

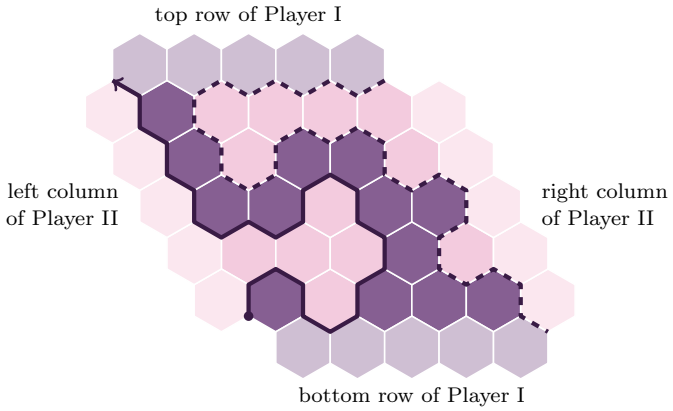



FIGURE 23. The no-draw argument of Theorem 17.1 on a full $5 \times 5$ board, with the stones of Player I in plum and those of Player II in pink, the extra rows and columns paler, and the walk that starts at the lower left corner, the one from the lower right corner being dashed.

and $z_1' - f_1(z') \geq \varepsilon$. Adding these inequalities, we obtain $(f_1(z) - f_1(z')) + (z_1' - z_1) \geq 2\varepsilon$, which is impossible because both terms are less than $\varepsilon$. A chain of Player I leads to the same contradiction, with $f_2$ in place of $f_1$. Hence every continuous map of $[0,1]^2$ into itself has a fixed point.

Seen through the game, the topological content of the theorem is the fact that Hex cannot end in a draw. The argument also has a finite form. Taken in the contrapositive, the argument shows that for every continuous map $f\colon [0,1]^2 \to [0,1]^2$ and every $\varepsilon > 0$, some point $z$ of a sufficiently fine grid satisfies $|f(z) - z|_\infty < \varepsilon$. By compactness these approximate fixed points have a limit point as $\varepsilon$ tends to 0, and by continuity this limit point is a fixed point. Combinatorial proofs of fixed point theorems, from Sperner's lemma to Scarf's algorithm, enter computation in exactly this finite form.

**Exercises.**

(1) List the adjacent pairs of cells of the $2 \times 2$ Hex board. Then check directly that in each of the 16 ways of filling the board exactly one player has a winning chain, as Theorem 17.1(1) asserts.

(2) Show that on the $3 \times 3$ board Player I has a winning strategy. (Hint: open with the center cell $(2,2)$, whose neighbors $(1,1)$ and $(2,1)$ lie in the bottom row and whose neighbors $(2,3)$ and $(3,3)$ lie in the top row. Player I then answers a move of Player II in one of these two pairs of neighbors with the other cell of the pair.)

(3) In the proof of Theorem 17.1(2), check that the points $z$ and $z'$ placed at adjacent cells satisfy $|z - z'|_\infty \leq \delta$, and carry out the argument for a winning chain of Player I, using $f_2$.

(4) $(*)$ Show that Player I has a winning strategy in Hex on the $k \times k$ board for every $k \geq 1$. (Hint: the reflection $(i,j) \mapsto (j,i)$ preserves adjacency and exchanges the sides of the two players. Combine Theorem 17.1(1), the determinacy of finite games, and the fact that an extra stone never hurts its owner, to show that Player II cannot have a winning strategy.)

Part
# V

# Ehrenfeucht–Fraïssé games and Fraïssé limits

This part begins with a question that needs no logic to state. The order of the rational numbers contains a copy of every finite linear order, and every order-preserving matching between two finite sets of rationals extends to a symmetry of the whole order. Which other countable structures are built in the same way, from a class of finite pieces, with every symmetry between pieces extending to a symmetry of the whole? Fraïssé answered the question in 1954: such a structure exists exactly when the pieces can be glued together along common parts, and the structure is then unique. The rational order, the random graph and the rational Urysohn space arise in this way, and so, once finite pieces are replaced by finitely generated ones, do the algebraic closure of a finite field and the field of algebraic numbers. The first section of the part is devoted to these structures, the Fraïssé limits. The section assumes no background in logic and can be read on its own.

The proof that a Fraïssé limit is unique matches the elements of two structures one at a time, alternately in each, and this back-and-forth matching is a game. Cut short after finitely many rounds, the game becomes the subject of the rest of the part. Like the games of Part IV, which were played on boards, this game is finite. The board, however, is a pair of mathematical structures (two linear orders, say, or two graphs), and either structure may be infinite. The purpose of the game is to compare the two structures through the eyes of first-order logic. First-order logic can say that a linear order has exactly seventeen elements, or that the order has no largest element. As we show in Corollary 21.5 below, first-order logic cannot say that a finite linear order has an even number of elements. A proof of such a negative statement must account for every sentence at once, and there are infinitely many sentences, some of them very long.

The game deals with the sentences by turning their quantifiers into moves. Suppose that a sentence $\exists x\, \varphi(x)$ holds in one structure and fails in the other. Player I then plays a witness for $x$ in the structure where the sentence holds. Whatever element Player II chooses in the other structure, the formula $\varphi$ fails there, and the contest moves on to the simpler formula $\varphi$, with one more pair of elements matched. Each round thus answers one quantifier, and a sentence whose quantifiers are nested $n$ deep is played out in $n$ rounds. The Ehrenfeucht–Fraïssé theorem says that this correspondence is exact: for two structures for a finite relational vocabulary, Player II can survive $n$ rounds precisely when the structures satisfy the same

sentences with at most $n$ nested quantifiers (see Theorem 20.3). For each fixed quantifier depth, agreement on infinitely many sentences thereby becomes a question about one winning strategy. In much the same way, the largeness of a set in the sense of Baire category became, in Part I, a fact about what Player II can enforce.

Like the Banach–Mazur game, our form of the game comes from Poland, but from Warsaw rather than Lwów. Between the wars the two cities shared Polish mathematics: Lwów, around Banach and Steinhaus and the journal *Studia Mathematica*, was the home of functional analysis. Warsaw, around Sierpiński and Mazurkiewicz and the journal *Fundamenta Mathematicae*, was the home of set theory, topology and logic. In Warsaw Ehrenfeucht found the game, by recasting as a game a characterization of elementary equivalence that Fraïssé had given in his thesis. Ehrenfeucht published the game in *Fundamenta Mathematicae* a year before Mycielski and Steinhaus proposed the axiom of determinacy (see Section 31).

**Roadmap**

In Section 18 we construct the Fraïssé limits, first from classes of finite structures and then from classes of finitely generated ones, with amalgamation supplying the extensions needed for back-and-forth. A catalog of examples from orders, graphs, algebra, topology and analysis comes before the construction, and the proofs for each example come after. The section uses the real-valued structures recalled in Appendix A, assumes no background in logic, and ends with the first steps from Fraïssé limits into logic. In Section 19 we define the game and watch the game separate a triangle from a path in two rounds. In Section 20 we prove the Ehrenfeucht–Fraïssé theorem (see Theorem 20.3). The main construction is a Hintikka formula: a finite description of everything a tuple satisfies up to a prescribed quantifier rank. In Section 21 we apply the theorem. Player II wins on any two dense linear orders without endpoints, and on any two long finite linear orders by halving gaps. The second strategy yields the result on parity announced above, and a counterpart for the connectivity of finite graphs. In Section 22 we first take the players away, which leaves Fraïssé's back-and-forth systems. We then let the game run for infinitely many rounds, which yields Cantor's theorem on countable dense orders once more. With ultrafilters, the same passage to infinitely many rounds leads to a theorem of Keisler and Shelah: elementarily equivalent structures have isomorphic ultrapowers. In Section 23 we survey the games that do for other logics what the game of this part does for first-order logic, from infinitary and finite-variable logics to current research on the expressive power of games. The history, which in the earlier parts accompanied each problem, is told in Section 24 at the end, because a single story runs through all the sections. From Part I the part needs only the language of games of Section 2, and Section 18 needs nothing from the earlier parts. The paragraph on ultrapowers also uses Łoś's theorem and saturated models, and in Section 23 we state results and give references rather than proofs.

## 18. Fraïssé limits

The guiding example of this section is the order of the rational numbers, which has two features. The finite pieces of the rationals are exactly the finite linear orders: every finite linear order appears inside the rationals, and every finite set of rationals, ordered as in the rationals, is a finite linear order. Moreover, every order-preserving matching between two finite sets of rationals extends to an order-preserving bijection of the rationals onto themselves. A countable structure whose finite pieces are, in this sense, exactly the members

of a class of finite structures is called *finitely universal* for that class, and a structure with the second feature is called *ultrahomogeneous*. Countable structures with both features occur throughout mathematics: among orders and graphs, in algebra, and in topology and analysis. We first make the two notions precise and describe a catalog of such structures, without proofs. Fraïssé discovered that a countable structure that is finitely universal for a class of finite structures and ultrahomogeneous is unique up to isomorphism, and that such a structure exists exactly when the class satisfies three simple conditions. We prove his theorem first for structures built from relations, such as orders and graphs, where the pieces are finite and the proofs are shortest. We then extend the theorem to the algebraic structures of the catalog, whose natural pieces are finitely generated rather than finite, and we show that every structure of the catalog has the two features. At the end of the section we show how the theorem leads into logic. The section assumes no background in logic, can be read independently of the rest of the part, and is meant as an entrance to the subject.

**Structures.** We use Keisler's general real-valued structures [Kei23], in the notation of Aubone, Dueñez and Iovino [ADI26]; Appendix A recalls the definitions, and their paper gives the details. In this section the vocabularies are countable and have a single sort descriptor $M$. A structure $\mathfrak{M}$ for such a vocabulary $V = (\mathbb{M}, F, P)$ consists of a set $M^{\mathfrak{M}}$, the *universe* of $\mathfrak{M}$, of a function $f^{\mathfrak{M}}$ on the universe for each function symbol $f$ in $F$, and of a function $p^{\mathfrak{M}}$ from a power of the universe to the interval $[-r_p, r_p]$ for each predicate symbol $p$ in $P$. A predicate is *two-valued* if the predicate takes only the values 1 and $-1$, read as true and false; a two-valued predicate is a *relation* in the traditional sense. A graph, for instance, is a structure for a vocabulary with one binary predicate symbol $E$ with a priori bound 1, interpreted by the two-valued predicate whose value at a pair of vertices is 1 when the vertices are adjacent and $-1$ otherwise; a linear order is a structure for a vocabulary with one binary predicate symbol $<$, read as "smaller than". A vocabulary is *relational* if the vocabulary has no function symbols, and therefore no constant symbols, and a structure is *classical* if every predicate of the structure is two-valued. Except where we say otherwise, the vocabulary is finite and relational and the structures are classical. We write structures in fraktur letters, $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{M}$, $\mathfrak{N}$ and so on, and their universes in the corresponding roman letters, $A$, $B$, $M$, $N$, so that $M = M^{\mathfrak{M}}$; when $\mathfrak{A}$ is a substructure of $\mathfrak{M}$, the universe $A$ is a subset of $M$. We also allow the *empty structure*, with no elements, so that an extension may start with no elements already matched; this is our one departure from the conventions of Appendix A, in which every universe is non-empty.

For a structure $\mathfrak{M}$ and a subset $S$ of $M$, the *induced substructure* on $S$ has universe $S$, with each relation of $\mathfrak{M}$ restricted to elements of $S$. For a graph, the induced subgraph on a set $S$ of vertices keeps every edge of the graph between two vertices of $S$ and no other edge. An *embedding* of a structure $\mathfrak{A}$ into a structure $\mathfrak{M}$ is an injective map $f\colon \mathfrak{A} \to \mathfrak{M}$ that preserves both the truth and the failure of each relation: for graphs, adjacent vertices go to adjacent vertices and non-adjacent vertices to non-adjacent vertices. Equivalently, an embedding is an isomorphism of $\mathfrak{A}$ onto an induced substructure of $\mathfrak{M}$. An *automorphism* of $\mathfrak{M}$ is an isomorphism of $\mathfrak{M}$ onto itself, a symmetry of the whole structure. An isomorphism between finite substructures of structures $\mathfrak{M}$ and $\mathfrak{N}$ is called a *finite partial isomorphism* from $\mathfrak{M}$ to $\mathfrak{N}$. The notions on which the section rests are the following.

- The *age* of a structure $\mathfrak{M}$ is the class of finite structures that embed into $\mathfrak{M}$, including all their isomorphic copies.
- A countable structure $\mathfrak{M}$ is *finitely universal* for a class $\mathcal{K}$ of finite structures if the age of $\mathfrak{M}$ is $\mathcal{K}$: every member of $\mathcal{K}$ embeds into $\mathfrak{M}$, and every finite substructure of

$\mathfrak{M}$ belongs to $\mathcal{K}$. The first condition says that $\mathfrak{M}$ is large enough to contain every member of $\mathcal{K}$, and the second that $\mathfrak{M}$ contains nothing outside $\mathcal{K}$. The rationals are not finitely universal for the class of finite partial orders, since a partial order with two incomparable elements does not embed into a linear order; and the random graph of the catalog below is not finitely universal for the class of finite triangle-free graphs, since the random graph contains triangles.

- A structure $\mathfrak{M}$ is *ultrahomogeneous* if every isomorphism between finite substructures of $\mathfrak{M}$ extends to an automorphism of $\mathfrak{M}$.

The age records the finite pictures that can be found inside $\mathfrak{M}$. Ultrahomogeneity says that $\mathfrak{M}$ cannot tell apart two finite pictures that look alike: whenever two finite pieces of $\mathfrak{M}$ are isomorphic, some symmetry of $\mathfrak{M}$ carries the first onto the second.

**Back and forth.** The tool that runs through this section is an argument that goes back to Cantor.

**Proposition 18.1 (Cantor).** *Let $\mathfrak{M}$ and $\mathfrak{N}$ be countable dense linear orders without endpoints. Then every finite partial isomorphism from $\mathfrak{M}$ to $\mathfrak{N}$ extends to an isomorphism of $\mathfrak{M}$ onto $\mathfrak{N}$.*

*Proof.* Fix enumerations of $M$ and of $N$, and let $p_0$ be a finite partial isomorphism from $\mathfrak{M}$ to $\mathfrak{N}$. For a forth step, let $p$ be a finite partial isomorphism, and let $a$ be the first element of $M$, in the enumeration, outside the domain of $p$. The elements of the domain of $p$ split into those below $a$ and those above $a$, and the images of the first lie below the images of the second. Since $\mathfrak{N}$ is dense and has no endpoints, some element $b$ of $N$ lies above the images of the elements below $a$ and below the images of the elements above $a$, and $p \cup \{(a, b)\}$ is again a finite partial isomorphism. A back step treats the first element of $N$ outside the range of $p$ in the same way, with the roles of $\mathfrak{M}$ and $\mathfrak{N}$ exchanged. Alternating forth and back steps from $p_0$ gives an increasing sequence $p_0 \subseteq p_1 \subseteq \cdots$ of finite partial isomorphisms. The union of the sequence is defined on all of $M$, has range all of $N$, and preserves the order, since any two elements lie in the domain of some $p_j$; the union is therefore an isomorphism of $\mathfrak{M}$ onto $\mathfrak{N}$ that extends $p_0$. ☺

With $p_0$ empty, the proposition says that any two countable dense linear orders without endpoints are isomorphic, which is Cantor's theorem of 1895. With $\mathfrak{N} = \mathfrak{M}$, the proposition says that every finite partial isomorphism of such an order extends to an automorphism. The matching of elements alternately in the two structures, which Cantor's own proof did not use, is the *back-and-forth* method.

**Example 18.2 (Ages).** We compute four ages.

(1) The orders $(\mathbb{N}, <)$, $(\mathbb{Z}, <)$ and $(\mathbb{Q}, <)$ have the same age, the class of all finite linear orders. Indeed, every finite subset of a linear order is a linear order, and a linear order with $n$ elements is isomorphic to the set $\{1, \ldots, n\}$, which lies inside each of the three orders.

(2) A countably infinite set with no relations has as age the class of all finite sets with no relations.

(3) The *infinite complete graph*, on a countably infinite set of vertices any two of which are adjacent, has as age the class of finite complete graphs.

(4) Let $\mathcal{P}$ be the graph with vertex set $\mathbb{N}$ whose edges are the pairs $\{2k, 2k+1\}$ for $k \in \mathbb{N}$, so that $\mathcal{P}$ consists of infinitely many disjoint edges. The age of $\mathcal{P}$ is the class of finite graphs in which every vertex has at most one neighbor. An induced subgraph of $\mathcal{P}$ consists of whole edges and of single vertices whose partner is missing; conversely, a

finite graph with $m$ edges and $l$ vertices without neighbors embeds into $\mathcal{P}$ by sending the edges to $m$ edges of $\mathcal{P}$ and the remaining vertices to one endpoint each of $l$ further edges.

**Example 18.3 (Ultrahomogeneous structures).** The structures of Example 18.2, except $(\mathbb{N}, <)$ and $(\mathbb{Z}, <)$, are ultrahomogeneous.

(1) The order $(\mathbb{Q}, <)$ is ultrahomogeneous. By Proposition 18.1 with $\mathfrak{N} = \mathfrak{M}$, every isomorphism between finite substructures extends to an automorphism.
(2) A countably infinite set with no relations is ultrahomogeneous. An isomorphism between finite subsets is now an arbitrary bijection, and the complements of the two finite subsets are both countably infinite, so the bijection extends to a permutation of the whole set.
(3) The infinite complete graph is ultrahomogeneous for the same reason, since every permutation of the vertices is an automorphism.
(4) The graph $\mathcal{P}$ of disjoint edges is ultrahomogeneous. Let $p\colon \mathfrak{A} \to \mathfrak{B}$ be an isomorphism between finite induced subgraphs of $\mathcal{P}$, and for a vertex $x$ write $x'$ for the vertex joined to $x$. For every vertex $a$ of $A$ whose partner $a'$ lies outside $A$, the vertex $p(a)$ has no neighbor in $B$, and so the partner $p(a)'$ lies outside $B$; send $a'$ to $p(a)'$. After these assignments, the domain and the range of the extended map are both unions of finitely many whole edges. Infinitely many edges lie outside the domain and infinitely many outside the range, and a bijection between these two sets of edges, matching endpoints with endpoints, completes an automorphism.

An automorphism preserves every property that can be described in terms of the relations: being the least element, being the immediate successor of a given element, having a given number of neighbors. To prove that a structure is not ultrahomogeneous, one therefore needs only to exhibit two isomorphic finite pieces that sit differently in the whole structure.

**Example 18.4 (Structures that are not ultrahomogeneous).** Each of the following structures has two isomorphic finite pieces that no automorphism matches.

(1) In $(\mathbb{N}, <)$ the map $1 \mapsto 0$ is an isomorphism between one-element substructures. Every automorphism of $(\mathbb{N}, <)$ fixes the least element 0, so no automorphism sends 1 to 0.
(2) In $(\mathbb{Z}, <)$, the map that sends 0 to 0 and 1 to 2 is an isomorphism from the substructure $\{0, 1\}$ onto the substructure $\{0, 2\}$. An automorphism $\sigma$ of $(\mathbb{Z}, <)$ with $\sigma(0) = 0$ must send 1, the least integer above 0, to the least integer above $\sigma(0)$, which is 1 and not 2.
(3) In the path with three vertices $x - y - z$, the map $x \mapsto y$ is an isomorphism between one-vertex subgraphs. No automorphism sends $x$, which has one neighbor, to $y$, which has two.

The first two items show that a single age may belong to several countable structures, of which at most one, here $(\mathbb{Q}, <)$, is ultrahomogeneous. By Theorem 18.9 below, the same uniqueness holds for every age.

**A catalog.** The structures of Example 18.3 are simple. We now describe richer countable structures that are finitely universal and ultrahomogeneous, going from orders and graphs to algebra, and then to topology and analysis. For each structure we name the class of pieces, and we indicate the way of gluing two pieces along a common part on which the construction of the structure rests. This gluing is called *amalgamation* below; algebraists know instances of amalgamation as pushouts, composita and tensor products. The proofs come after the construction, in the same order.

*Orders.* The order $(\mathbb{Q}, <)$ is finitely universal for the class of finite linear orders, and ultrahomogeneous. Two finite linear orders with a common suborder are glued along that suborder by interleaving their other elements in each gap. If each rational is colored red or blue so that both colors meet every open interval, the result is finitely universal for the class of finite linear orders with two colors, and ultrahomogeneous.

*Graphs.* There is a countable graph, the *random graph* or *Rado graph*, that is finitely universal for the class of all finite graphs, and ultrahomogeneous. The Rado graph has a striking description: for any two disjoint finite sets $U$ and $V$ of vertices, some vertex outside $U \cup V$ is adjacent to every vertex of $U$ and to no vertex of $V$. A countable graph whose edges are chosen by independent tosses of a fair coin is, with probability 1, isomorphic to the Rado graph. Two finite graphs with a common induced subgraph are glued along that subgraph by adding no edges between their other vertices. Forbidding triangles gives the *Henson graph*, which is finitely universal for the class of finite triangle-free graphs and ultrahomogeneous, and the same pattern yields universal ultrahomogeneous tournaments and partial orders.

*From relations to operations.* In algebra, structures carry operations, such as addition and multiplication, and constants, such as 0 and 1. A piece of an algebraic structure should be closed under the operations: inside a field, the natural pieces are subfields, not arbitrary finite sets of elements. Closing a finite set under the operations may, however, produce an infinite set. The smallest subfield of $\mathbb{C}$ that contains $\sqrt{2}$ is the field $\mathbb{Q}(\sqrt{2})$, which is infinite although one element describes the field; similarly, a line in a vector space over $\mathbb{Q}$ is infinite and is described by one vector. For algebraic structures, the finite pieces of the rational order are therefore replaced by *finitely generated* pieces, the smallest substructures that contain a given finite set. Over a finite field these pieces are again finite, but over $\mathbb{Q}$ they are not. We give the precise definitions after the construction (see Theorem 18.23 below). For now, an algebraic structure is finitely universal for a class of finitely generated structures if the finitely generated substructures of the structure are, up to isomorphism, exactly the members of the class, and ultrahomogeneous if every isomorphism between finitely generated substructures extends to an automorphism. In a structure built from relations alone every subset counts as a piece, so the finitely generated pieces are the finite ones, and nothing changes for orders and graphs.

*Vector spaces over a finite field.* Let $p$ be a prime, and let $V$ be a vector space of countably infinite dimension over the field $\mathbb{F}_p$ with $p$ elements. The space $V$ is finitely universal for the class of finite vector spaces over $\mathbb{F}_p$, and ultrahomogeneous: every linear isomorphism between two finite subspaces of $V$ extends to a linear automorphism of $V$. Two finite vector spaces with a common subspace are glued along that subspace by a pushout.

*The algebraic closure of a finite field.* The algebraic closure $\overline{\mathbb{F}}_p$ of $\mathbb{F}_p$ is finitely universal for the class of finite fields of characteristic $p$, and ultrahomogeneous. The automorphism group of $\overline{\mathbb{F}}_p$ is isomorphic to the profinite completion $\widehat{\mathbb{Z}}$ of the integers, and the powers of the Frobenius map $x \mapsto x^p$ form a dense subgroup. Two finite fields with a common subfield are glued inside a larger finite field.

*The field of algebraic numbers.* Here the pieces are infinite. The field $\overline{\mathbb{Q}}$ of algebraic numbers is finitely universal for the class of number fields, the finite extensions of $\mathbb{Q}$, and ultrahomogeneous. The automorphism group of $\overline{\mathbb{Q}}$ is the absolute Galois group of $\mathbb{Q}$, a central object of algebraic number theory, which thus appears as the symmetry group of the structure assembled from the number fields. Two number fields with a common subfield $K$ are glued over $K$ by a quotient of their tensor product over $K$.

*Finite groups.* In 1959 Philip Hall constructed [Hal59] a countable group $\mathbb{H}$ that is finitely universal for the class of finite groups, and ultrahomogeneous; in $\mathbb{H}$, every isomorphism between finite subgroups is even induced by conjugation by an element of $\mathbb{H}$. Two finite groups with a common subgroup are glued along that subgroup inside a finite group, by a theorem of B. H. Neumann [Neu60].

*Topology: Boolean algebras and the Cantor set.* The countable atomless Boolean algebra $\mathbb{B}$ is finitely universal for the class of finite Boolean algebras, and ultrahomogeneous. By Stone duality, $\mathbb{B}$ is the algebra of clopen subsets of the Cantor set $2^\omega$, and the automorphisms of $\mathbb{B}$ correspond to the homeomorphisms of $2^\omega$. Ultrahomogeneity of $\mathbb{B}$ thus says that every bijection between the pieces of two partitions of $2^\omega$ into finitely many non-empty clopen sets is induced by a homeomorphism of $2^\omega$. Gluing two finite Boolean algebras along a common subalgebra corresponds, dually, to a fiber product of finite sets.

*Analysis: metric and Banach spaces.* The rational Urysohn space is a countable metric space with rational distances at most 1 that is finitely universal for the class of finite metric spaces with such distances, and ultrahomogeneous: every isometry between finite subsets of the space extends to an isometry of the whole space onto itself. The completion of the rational Urysohn space is Urysohn's universal space, which contains an isometric copy of every separable metric space of diameter at most 1. Two finite metric spaces with a common subspace are glued along that subspace by taking, as the distance between two of their other points, the length of the shortest route through the common subspace, capped at 1. In functional analysis the corresponding object is the Gurarij space [Gur66], a separable Banach space built from the finite-dimensional normed spaces. The Gurarij space has the two features only approximately: for finite-dimensional normed spaces $E \subseteq F$ and a real number $\varepsilon > 0$, every isometric embedding of $E$ into the Gurarij space extends to a linear map $T$ from $F$ into the space with $(1-\varepsilon)\|x\| \leq \|Tx\| \leq (1+\varepsilon)\|x\|$ for every $x \in F$. Lusky proved in 1976 that the Gurarij space is unique up to linear isometry [Lus76], and Ben Yaacov placed the spaces of Urysohn and Gurarij in a common theory of Fraïssé limits of metric structures [BY15]. These approximate limits lie beyond the exact setting of these notes, but they follow the same pattern.

**Which classes are ages?** We now show how such structures are constructed, beginning with structures built from relations. Let $\mathfrak{M}$ be any structure with age $\mathcal{K}$. Two properties of $\mathcal{K}$ hold automatically. First, an induced substructure of a finite substructure of $\mathfrak{M}$ is again a finite substructure of $\mathfrak{M}$, so $\mathcal{K}$ is closed under induced substructures. Second, two members of $\mathcal{K}$ embed into $\mathfrak{M}$, and the induced substructure of $\mathfrak{M}$ on the union of the two images is a finite substructure into which both members embed. When $\mathfrak{M}$ is moreover ultrahomogeneous, a third property appears, and Fraïssé's insight was that this third property is the heart of the matter. Suppose that $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ belong to $\mathcal{K}$ and that $f\colon \mathfrak{A} \to \mathfrak{B}$ and $g\colon \mathfrak{A} \to \mathfrak{C}$ are embeddings. Choose embeddings $e_B\colon \mathfrak{B} \to \mathfrak{M}$ and $e_C\colon \mathfrak{C} \to \mathfrak{M}$. The two copies $e_B(f(A))$ and $e_C(g(A))$ of $\mathfrak{A}$ inside $\mathfrak{M}$ are isomorphic through the map $e_B(f(a)) \mapsto e_C(g(a))$, for $a \in A$. By ultrahomogeneity this map extends to an automorphism $\sigma$ of $\mathfrak{M}$. Now $\sigma \circ e_B$ and $e_C$ embed $\mathfrak{B}$ and $\mathfrak{C}$ into the induced substructure $\mathfrak{D}$ of $\mathfrak{M}$ on the union of their images, and they agree on $A$: we have $\sigma \circ e_B \circ f = e_C \circ g$. Thus any two extensions of a member of $\mathcal{K}$ can be fitted together over that member inside a third member of $\mathcal{K}$. These three properties define the classes from which Fraïssé's construction starts.

**Definition 18.5.** Let $\mathcal{K}$ be a non-empty class of finite classical structures for a fixed finite relational vocabulary, closed under isomorphism and containing structures of arbitrarily large finite size. We call $\mathcal{K}$ a *Fraïssé class* if $\mathcal{K}$ has the following properties.

(HP) *Hereditary property.* Every induced substructure of a member of $\mathcal{K}$ belongs to $\mathcal{K}$.
(JEP) *Joint embedding property.* Any two members of $\mathcal{K}$ embed into a third member.
(AP) *Amalgamation property.* If $\mathfrak{A}, \mathfrak{B}, \mathfrak{C} \in \mathcal{K}$ and $f\colon \mathfrak{A} \to \mathfrak{B}$, $g\colon \mathfrak{A} \to \mathfrak{C}$ are embeddings, there are $\mathfrak{D} \in \mathcal{K}$ and embeddings $u\colon \mathfrak{B} \to \mathfrak{D}$, $v\colon \mathfrak{C} \to \mathfrak{D}$ such that $u \circ f = v \circ g$.

A structure $\mathfrak{D}$ as in (AP), together with the embeddings $u$ and $v$, is an *amalgam* of $\mathfrak{B}$ and $\mathfrak{C}$ over $\mathfrak{A}$.

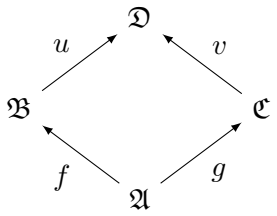



FIGURE 24. The amalgamation property: two extensions $\mathfrak{B}$ and $\mathfrak{C}$ of $\mathfrak{A}$ embed into a common $\mathfrak{D}$ so that the two routes from $\mathfrak{A}$ to $\mathfrak{D}$ agree.

In practice an amalgam is built in two steps (see Figure 25). First replace $\mathfrak{B}$ and $\mathfrak{C}$ by isomorphic copies whose universes meet exactly in a common copy of $A$, with $f$ and $g$ becoming inclusions; then decide which relations hold between the new elements of $B$ and the new elements of $C$, so that the resulting structure on $B \cup C$ belongs to $\mathcal{K}$. The only requirement is that the relations inside $\mathfrak{B}$ and inside $\mathfrak{C}$ stay as they were. Amalgamation does not require the images of $B$ and $C$ to meet only in the copy of $A$: an amalgam may glue a new element of $B$ to a new element of $C$, and in Example 18.21 below gluing is unavoidable.

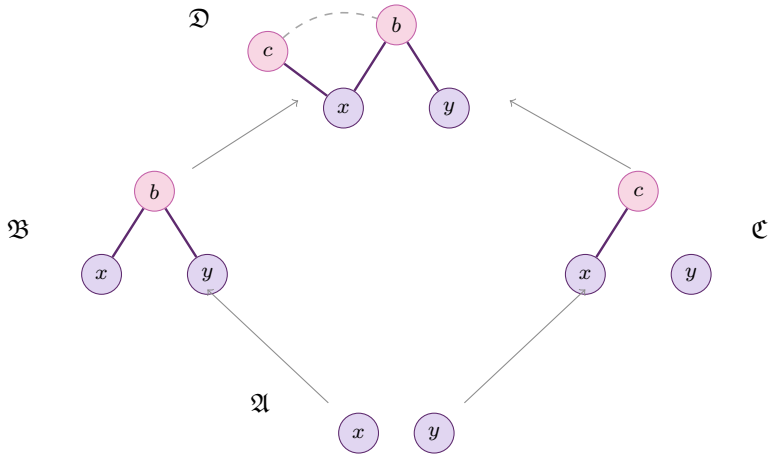



FIGURE 25. The amalgam of two graphs $\mathfrak{B}$ and $\mathfrak{C}$ over a common induced subgraph $\mathfrak{A}$ keeps every edge of $\mathfrak{B}$ and of $\mathfrak{C}$ and adds no edge between the new vertices $b$ and $c$.

Since the empty structure belongs to $\mathcal{K}$ by (HP), (JEP) is the case of amalgamation over the empty structure; we name the joint embedding property separately because, in the construction, joint embedding supplies universality. There are only countably many isomorphism types of finite classical structures for a finite relational vocabulary, so no further countability

hypothesis on $\mathcal{K}$ is needed. The argument above shows that the age of a countably infinite ultrahomogeneous structure is a Fraïssé class; the hypothesis on sizes holds because such a structure has finite substructures of every finite size.

**Example 18.6 (Classes that are not Fraïssé classes).** Each of the three properties can fail.

(1) *The hereditary property fails for finite connected graphs.* The path $x - y - z$ is connected, but the induced subgraph on $\{x, z\}$ has no edge and is not connected. The joint embedding and amalgamation properties hold: join the two given graphs, or the new vertices of the two extensions, by enough edges to make the result connected.

(2) *The joint embedding property fails for graphs that are complete or have no edges.* The class of finite graphs that are either complete or edgeless has the hereditary property. A graph into which both a single edge and a pair of non-adjacent vertices embed has an edge and a non-edge, so the graph is neither complete nor edgeless. Amalgamation fails as well, since amalgamation over the empty structure is the joint embedding property.

(3) *The amalgamation property fails for finite forests*, the finite graphs without cycles. The hereditary and joint embedding properties hold; for the second, take disjoint unions. Let $\mathfrak{A}$ consist of two non-adjacent vertices $x$ and $y$, let $\mathfrak{B}$ be the path $x-b-y$, and let $\mathfrak{C}$ be the path $x-c_1-c_2-y$; all three are forests. In an amalgam $\mathfrak{D}$, the image of $b$ is adjacent to $x$ and to $y$. The image of $b$ cannot be the image of $c_1$, which is not adjacent to $y$, nor the image of $c_2$, which is not adjacent to $x$. Hence $\mathfrak{D}$ contains the cycle formed by $x$, $b$, $y$, $c_2$, $c_1$, with five distinct vertices, and $\mathfrak{D}$ is not a forest (see Figure 26).

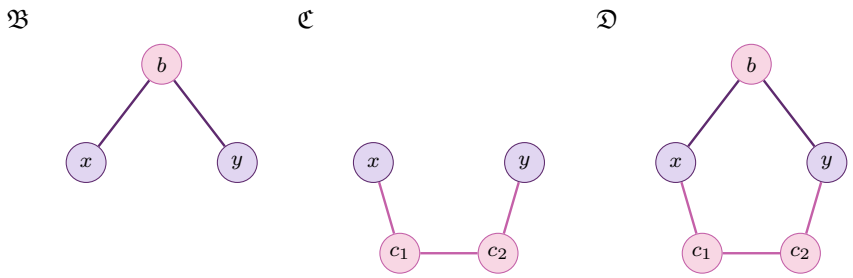



FIGURE 26. Two forests $\mathfrak{B}$ and $\mathfrak{C}$ that contain the same two non-adjacent vertices $x$ and $y$ can be amalgamated over $\{x, y\}$ only in a graph $\mathfrak{D}$ that contains a cycle with five vertices.

**The extension property.** The construction and every example below turn on one property of the limit.

**Definition 18.7.** Let $\mathcal{K}$ be a class of finite structures and let $\mathfrak{M}$ be a structure. We say that $\mathfrak{M}$ has the *extension property* for $\mathcal{K}$ if, whenever $\mathfrak{A} \subseteq \mathfrak{B}$ are members of $\mathcal{K}$ with $\mathfrak{A}$ an induced substructure of $\mathfrak{B}$, every embedding $f\colon \mathfrak{A} \to \mathfrak{M}$ extends to an embedding $g\colon \mathfrak{B} \to \mathfrak{M}$.

For empty $\mathfrak{A}$, the extension property asks that every member of $\mathcal{K}$ embed into $\mathfrak{M}$. The property can be read as a game of one move. A challenger points at a finite piece $f(A)$ of $\mathfrak{M}$ and describes a larger finite structure $\mathfrak{B}$ around a copy of $\mathfrak{A}$; the structure $\mathfrak{M}$ must answer by finding, around $f(A)$, a copy of $\mathfrak{B}$. For $(\mathbb{Q}, <)$ the challenge "find a rational between these two" is always met; for $(\mathbb{Z}, <)$ the challenge "find an integer between 0 and 1" is not. Challenges that add one element suffice.

**Lemma 18.8.** *Let $\mathcal{K}$ be a class of finite structures with the hereditary property, and let $\mathfrak{M}$ be a structure. Suppose that, whenever $\mathfrak{A} \subseteq \mathfrak{B}$ are members of $\mathcal{K}$ with $\mathfrak{A}$ an induced substructure of $\mathfrak{B}$ and $B \setminus A$ a single element, every embedding $\mathfrak{A} \to \mathfrak{M}$ extends to an embedding $\mathfrak{B} \to \mathfrak{M}$. Then $\mathfrak{M}$ has the extension property for $\mathcal{K}$.*

*Proof.* Let $\mathfrak{A} \subseteq \mathfrak{B}$ be members of $\mathcal{K}$, let $f\colon \mathfrak{A} \to \mathfrak{M}$ be an embedding, and list the elements of $B \setminus A$ as $b_1, \dots, b_m$. For $0 \le j \le m$, let $\mathfrak{B}_j$ be the induced substructure of $\mathfrak{B}$ on $A \cup \{b_1, \dots, b_j\}$; every $\mathfrak{B}_j$ belongs to $\mathcal{K}$ by (HP). Starting from $f$, extend the embedding from $\mathfrak{B}_{j-1}$ to $\mathfrak{B}_j$ for $j = 1, \dots, m$, one element at a time, by the hypothesis. The last extension is an embedding of $\mathfrak{B}_m = \mathfrak{B}$ that extends $f$. ☺

**Theorem 18.9 (Fraïssé).** *Let $\mathcal{K}$ be a Fraïssé class for a finite relational vocabulary. Then there is a countably infinite ultrahomogeneous structure whose age is $\mathcal{K}$, and any two such structures are isomorphic.*

This structure is called the *Fraïssé limit* of $\mathcal{K}$ [Fra54a, Fra54b]. The proof has three parts, and each part can be presented separately.

(1) *Construction.* We build the limit as the union of a growing chain of finite structures in $\mathcal{K}$. At each stage we first add a copy of the next structure on a list of all members of $\mathcal{K}$, by (JEP); then we answer, by (AP), every one-move challenge of the extension property that concerns the structure built so far. Challenges that arise later are answered at later stages.

(2) *Ultrahomogeneity gives the extension property.* A symmetry of the structure moves any copy of $\mathfrak{B}$ to a copy that lies around the prescribed copy of $\mathfrak{A}$.

(3) *The extension property gives back-and-forth.* Two countable structures with the extension property can be matched element by element, alternately in each, as in the proof of Proposition 18.1; matching every element of both structures produces an isomorphism.

*Proof. Construction.* List representatives $\mathfrak{B}_0, \mathfrak{B}_1, \dots$ of the isomorphism types in $\mathcal{K}$. We build finite structures

$$\mathfrak{C}_0 \subseteq \mathfrak{C}_1 \subseteq \mathfrak{C}_2 \subseteq \cdots$$

in $\mathcal{K}$, each an induced substructure of the next, starting with the empty structure $\mathfrak{C}_0$. At stage $n$, use (JEP) to obtain a member of $\mathcal{K}$ into which both $\mathfrak{C}_n$ and $\mathfrak{B}_n$ embed. Replace that member by an isomorphic copy in which the embedding of $\mathfrak{C}_n$ is an inclusion, and call the copy $\mathfrak{D}$. Next list the *requests* of stage $n$: one request for each index $i \le n$, each induced substructure $\mathfrak{A}$ of $\mathfrak{B}_i$, and each embedding $f\colon \mathfrak{A} \to \mathfrak{C}_n$. The request asks that $f$ extend to an embedding of $\mathfrak{B}_i$ into $\mathfrak{C}_{n+1}$. There are finitely many requests, since there are finitely many choices of $i$ and $\mathfrak{A}$ and finitely many maps from each $A$ into $C_n$. The list is fixed before processing begins, and every request refers to $\mathfrak{C}_n$, even as $\mathfrak{D}$ grows.

Each request can be met by amalgamation. Suppose that the structure built so far in stage $n$ is $\mathfrak{D}$ and that the next request is $f\colon \mathfrak{A} \to \mathfrak{C}_n$, with $\mathfrak{A} \subseteq \mathfrak{B}_i$. Amalgamate $\mathfrak{D}$ and $\mathfrak{B}_i$ over $\mathfrak{A}$, using $f$ followed by the inclusion of $\mathfrak{C}_n$ into $\mathfrak{D}$ on one side and the inclusion of $\mathfrak{A}$ into $\mathfrak{B}_i$ on the other. Replace the amalgam by an isomorphic copy in which the embedding of $\mathfrak{D}$ is an inclusion. The new structure contains $\mathfrak{D}$ as an induced substructure, so the requests already met remain met, and the embedding of $\mathfrak{B}_i$ into the new structure extends $f$. After the whole list has been processed, call the resulting structure $\mathfrak{C}_{n+1}$.

Let $\mathfrak{M}$ be the union of the chain, the structure whose universe is $M = \bigcup_{n\geq 0} C_n$ and whose relations are the unions of those of the structures $\mathfrak{C}_n$. Every finite substructure of $\mathfrak{M}$ is a substructure of some $\mathfrak{C}_n$, so (HP) shows that the age of $\mathfrak{M}$ is contained in $\mathcal{K}$. Every $\mathfrak{B}_i$ embeds into $\mathfrak{C}_{i+1}$, so the age of $\mathfrak{M}$ is exactly $\mathcal{K}$. The union is countable, and infinite because $\mathcal{K}$ contains arbitrarily large structures. To check the extension property, take members $\mathfrak{A} \subseteq \mathfrak{B}$ of $\mathcal{K}$ and an embedding $f\colon \mathfrak{A} \to \mathfrak{M}$. Choose an isomorphism $h\colon \mathfrak{B} \to \mathfrak{B}_i$. The image $f(A)$ is finite, so the map $f \circ (h|_A)^{-1}$ embeds the induced substructure of $\mathfrak{B}_i$ on $h(A)$ into some $\mathfrak{C}_n$ with $n \geq i$. The corresponding request of stage $n$ supplies an embedding $g\colon \mathfrak{B}_i \to \mathfrak{C}_{n+1}$ that extends $f \circ (h|_A)^{-1}$. Then $g \circ h$ is an embedding of $\mathfrak{B}$ into $\mathfrak{M}$ that extends $f$.

*Ultrahomogeneity gives the extension property.* Let $\mathfrak{N}$ be an ultrahomogeneous structure with age $\mathcal{K}$. Given members $\mathfrak{A} \subseteq \mathfrak{B}$ of $\mathcal{K}$ and an embedding $f\colon \mathfrak{A} \to \mathfrak{N}$, choose an embedding $e\colon \mathfrak{B} \to \mathfrak{N}$, which exists because $\mathfrak{B}$ belongs to the age of $\mathfrak{N}$. The map $e(a) \mapsto f(a)$, for $a \in A$, is an isomorphism between finite substructures of $\mathfrak{N}$. Extend the map to an automorphism $\sigma$ of $\mathfrak{N}$. Then $\sigma \circ e$ is an embedding of $\mathfrak{B}$ into $\mathfrak{N}$ that extends $f$.

*The extension property gives back-and-forth.* Let $\mathfrak{M}$ and $\mathfrak{N}$ be countably infinite structures with age $\mathcal{K}$ and the extension property for $\mathcal{K}$, and fix enumerations of $M$ and of $N$. We show that every finite partial isomorphism $p_0$ from $\mathfrak{M}$ to $\mathfrak{N}$ extends to an isomorphism. For a forth step, let $p$ be the current map and let $a$ be the first element of $M$, in the enumeration, outside the domain of $p$. Apply the extension property in $\mathfrak{N}$ to the induced substructure $\mathfrak{A}$ of $\mathfrak{M}$ on the domain of $p$, to the induced substructure $\mathfrak{B}$ of $\mathfrak{M}$ on the domain together with $a$, and to the embedding $p\colon \mathfrak{A} \to \mathfrak{N}$; the extension sends $a$ to an element of $N$ and extends $p$. For a back step, apply the same argument, with the roles of $\mathfrak{M}$ and $\mathfrak{N}$ exchanged, to $p^{-1}$ and to the first element of $N$ outside the range of $p$, and invert the extension. Alternating forth and back steps gives an increasing sequence $p_0 \subseteq p_1 \subseteq \cdots$ of finite partial isomorphisms. The union of the sequence is defined on all of $M$, has range all of $N$, and preserves and reflects every relation, since each finite tuple lies in the domain of some $p_j$; hence the union is an isomorphism of $\mathfrak{M}$ onto $\mathfrak{N}$.

Taking $\mathfrak{N} = \mathfrak{M}$ proves that the structure constructed in the first part is ultrahomogeneous. Taking $p_0$ to be the empty map proves uniqueness, since by the second part every countably infinite ultrahomogeneous structure of age $\mathcal{K}$ has the extension property. ☺

The proof yields a practical test, which we use in every example below.

**Corollary 18.10.** *Let $\mathcal{K}$ be a Fraïssé class, and let $\mathfrak{M}$ be a countably infinite structure whose age is contained in $\mathcal{K}$. Suppose that, whenever $\mathfrak{A} \subseteq \mathfrak{B}$ are members of $\mathcal{K}$ with $\mathfrak{A}$ an induced substructure of $\mathfrak{B}$ and $B \setminus A$ a single element, every embedding $\mathfrak{A} \to \mathfrak{M}$ extends to an embedding $\mathfrak{B} \to \mathfrak{M}$. Then $\mathfrak{M}$ is the Fraïssé limit of $\mathcal{K}$.*

*Proof.* By Lemma 18.8, the structure $\mathfrak{M}$ has the extension property for $\mathcal{K}$. With $\mathfrak{A}$ empty, the extension property shows that every member of $\mathcal{K}$ embeds into $\mathfrak{M}$, so the age of $\mathfrak{M}$ is exactly $\mathcal{K}$. The back-and-forth part of the proof of Theorem 18.9, with $\mathfrak{N} = \mathfrak{M}$, shows that $\mathfrak{M}$ is ultrahomogeneous. ☺

To present a Fraïssé limit, one therefore follows the same four steps: describe the class and check the hereditary property; check the joint embedding property; build an amalgam; and exhibit a countable structure with the right age that meets every one-element challenge. The examples below follow these steps.

**Example 18.11 (The construction for linear orders).** For the class of finite linear orders, the construction can be followed by hand. A one-element challenge for a finite order $\mathfrak{C}$ asks for a new element in a prescribed position: between two given consecutive elements of

$C$, below the least element, or above the greatest. Start with $C_1 = \{0\}$, and let $C_{n+1}$ consist of $C_n$, the midpoint of every two consecutive elements of $C_n$, one element 1 less than the least element and one element 1 greater than the greatest. Thus

$$C_1 = \{0\}, \qquad C_2 = \{-1, 0, 1\}, \qquad C_3 = \{-2, -1, -\tfrac{1}{2}, 0, \tfrac{1}{2}, 1, 2\},$$

and so on. The union of the sets $C_n$ is a countable set of dyadic rationals. The union is dense: if $x < y$ lie in $C_n$, the midpoint of two consecutive elements of $C_n$ between $x$ and $y$ lies in $C_{n+1}$, strictly between $x$ and $y$; the union has no endpoints, since every $C_{n+1}$ extends $C_n$ by one element at each end. By Proposition 18.1, the union is isomorphic to $(\mathbb{Q}, <)$. The general construction differs from this one only in bookkeeping: the requests of each stage cover every finite structure listed so far, not only one-element extensions. The first four stages are drawn in Figure 27.

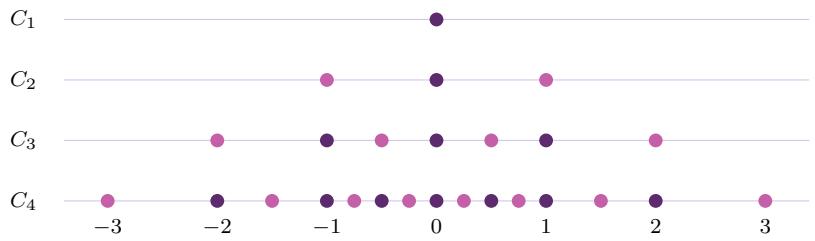



FIGURE 27. The first four stages of the construction of Example 18.11, with the elements added at each stage in pink: a midpoint in every gap and one new element beyond each end.

**Example 18.12 (Finite sets).** For a vocabulary whose only predicate symbol is the logical unity constant, a structure is just a set, and an embedding is an injective map. The class of all finite sets has the hereditary property, disjoint unions give joint embeddings, and the union of two sets meeting exactly in $A$ is an amalgam over $\mathfrak{A}$. The limit is a countably infinite set: a one-element challenge asks for a point outside a given finite set $f(A)$, and there are infinitely many.

**Example 18.13 (Colored sets).** Let the vocabulary have one unary predicate symbol $R$, interpreted by a relation $R$ read as "$x$ is red"; an element outside $R$ is blue. The class of all finite colored sets is a Fraïssé class: amalgamate by taking the union of copies meeting exactly in $A$, each element keeping the same color. The limit is $\mathbb{N}$ with the even numbers red and the odd numbers blue. A one-element challenge asks for a new point of a prescribed color outside a finite set $f(A)$, and each color has infinitely many points. With $k$ unary relations, each element has one of $2^k$ combinations of colors, and the limit is a countable set with infinitely many elements of each combination.

**Example 18.14 (Equivalence relations).** Let the vocabulary have one binary predicate symbol $E$, and let $\mathcal{K}$ be the class of finite sets with an equivalence relation $E$ that interprets $E$. The hereditary property holds because the restriction of an equivalence relation is an equivalence relation, and disjoint unions, with no new equivalences, give joint embeddings. For amalgamation, let $\mathfrak{B}$ and $\mathfrak{C}$ be members whose universes meet exactly in the universe $A$ of a common substructure $\mathfrak{A}$. On $B \cup C$ declare $x$ and $y$ equivalent if they are equivalent in $\mathfrak{B}$, or equivalent in $\mathfrak{C}$, or if one lies in $B$, the other in $C$, and both are equivalent to a common element of $A$. The new relation is transitive because two elements of $A$ that are equivalent in $\mathfrak{C}$ are equivalent in $\mathfrak{A}$, and hence in $\mathfrak{B}$. The new relation agrees with the old ones on $B$

and on $C$, so the resulting structure on $B \cup C$ is an amalgam. The limit is $\mathbb{N} \times \mathbb{N}$ with $(m, n)$ equivalent to $(m', n')$ when $m = m'$: infinitely many classes, each infinite. A one-element challenge asks for a new point outside $f(A)$ that is equivalent either to a prescribed element $f(a)$ or to no element of $f(A)$. In the first case, take an unused point of the class of $f(a)$, which is infinite; in the second, take a point of a class that does not meet $f(A)$, which exists because $f(A)$ meets only finitely many of the infinitely many classes. Restricting the class to equivalence relations with at most $k$ classes gives another Fraïssé class, whose limit has exactly $k$ classes, each infinite.

**Example 18.15 (Finite linear orders).** The class of all finite linear orders is a Fraïssé class. The hereditary property is immediate, and two finite orders embed into the order obtained by placing one after the other. For amalgamation, replace the orders by copies meeting exactly in the common suborder $\mathfrak{A}$, with the given embeddings becoming inclusions. If $\mathfrak{A}$ is empty, place one order after the other. Otherwise, in each gap of $A$, including the two outer gaps, place the new elements from the first order before those from the second, preserving their order within each group. The result is a finite linear order containing both orders as induced suborders. The Fraïssé limit is $(\mathbb{Q}, <)$. A one-element challenge asks for a rational in a prescribed gap of a finite set of rationals: between two consecutive elements, below all of them, or above all of them. Density answers the first kind of challenge, and the absence of endpoints answers the other two.

**Example 18.16 (Colored linear orders).** Let the vocabulary have a binary predicate symbol $<$ and a unary predicate symbol $R$, and let $\mathcal{K}$ be the class of finite linear orders in which every element is red, in $R$, or blue, outside $R$. The amalgam of Example 18.15, with every element keeping the same color, shows that $\mathcal{K}$ is a Fraïssé class. A one-element challenge now asks for an element of a prescribed color in a prescribed gap. The limit is therefore $(\mathbb{Q}, <)$ with a set $R$ of red rationals such that every non-empty open interval contains both red and blue rationals. For example, let a rational be red if the denominator in lowest terms is odd and blue if the denominator is even. Every open interval contains a fraction $m/3^k$ and a fraction $m'/2^{k'}$ with $m'$ odd and $k' \geq 1$, so both colors occur in every interval.

**Example 18.17 (The random graph).** All finite graphs form a Fraïssé class. To amalgamate two graphs, replace them by copies meeting exactly in the common induced subgraph, with the given embeddings becoming inclusions. Keep all edges of both copies, and add no edge between the remaining vertices. The Fraïssé limit is the *countable random graph*, also called the *Rado graph* [Cam97]. A one-element challenge describes a new vertex by the neighbors of the new vertex among the vertices of $f(A)$. The extension property therefore takes a particularly simple form: for disjoint finite sets $U$ and $V$ of vertices, there is a vertex outside $U \cup V$ adjacent to every vertex of $U$ and to no vertex of $V$ (see Figure 28). By Corollary 18.10, any countable graph with this property is the Rado graph. Two very different constructions produce such a graph.

(1) *Rado's construction* [Rad64]. Take $\mathbb{N}$ as the set of vertices, and for $i < j$ let $i$ and $j$ be adjacent if the $i$-th binary digit of $j$ is 1, counting digits from 0 at the right. For example, among the vertices smaller than $6 = 110_2$, the vertex 6 is adjacent exactly to 1 and 2. Given disjoint finite sets $U$ and $V$, choose an integer $m$ larger than every element of $U \cup V$ and let $j = 2^m + \sum_{u \in U} 2^u$. Then $j$ exceeds every element of $U \cup V$, the $u$-th digit of $j$ is 1 for $u \in U$, and the $v$-th digit is 0 for $v \in V$. Hence $j$ is adjacent to every vertex of $U$ and to no vertex of $V$.

(2) *Coin flips* [ER63]. Take $\mathbb{N}$ as the set of vertices and, for each pair of distinct vertices independently, flip a fair coin to decide whether the pair is an edge. Fix disjoint finite

sets $U$ and $V$ with $k$ elements in total. A vertex outside $U \cup V$ has the required adjacencies with probability $2^{-k}$, and different vertices involve different coin flips. The probability that none of the first $N$ vertices outside $U \cup V$ has the required adjacencies is $(1 - 2^{-k})^N$, which tends to 0 as $N$ tends to infinity. So the extension property fails for $U$ and $V$ with probability 0. There are only countably many pairs $U$, $V$, and a countable union of events of probability 0 has probability 0. Hence, with probability 1, the coin flips produce the Rado graph.

Thus the word "random" is justified twice: no probability is needed to construct the graph, but a random countable graph is almost surely the Rado graph.

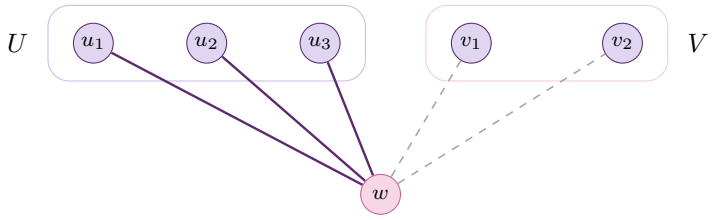



FIGURE 28. The extension property of the Rado graph: for disjoint finite sets $U$ and $V$ of vertices, some vertex $w$ outside $U \cup V$ is adjacent to every vertex of $U$ and to no vertex of $V$.

**Example 18.18 (Triangle-free graphs).** A graph is *triangle-free* if no three vertices are pairwise adjacent. The class of finite triangle-free graphs has the hereditary property, and disjoint unions give joint embeddings. The amalgam of Example 18.17, which adds no edges between new vertices, works here as well. Indeed, a triangle in the amalgam cannot contain both a new vertex of $\mathfrak{B}$ and a new vertex of $\mathfrak{C}$, since such vertices are not adjacent; so the triangle would lie inside $\mathfrak{B}$ or inside $\mathfrak{C}$, which are triangle-free. The limit is the *Henson graph* [Hen71], a countable triangle-free graph that contains a copy of every finite triangle-free graph. A one-element challenge adds a vertex adjacent exactly to a set $U$ of given vertices, and the enlarged graph is triangle-free precisely when no two vertices of $U$ are adjacent. The extension property therefore reads: for disjoint finite sets $U$ and $V$ of vertices such that no two vertices of $U$ are adjacent, there is a vertex outside $U \cup V$ adjacent to every vertex of $U$ and to no vertex of $V$. For each $n \geq 3$, the same argument applies to graphs that contain no $n$ pairwise adjacent vertices.

**Example 18.19 (Tournaments).** A *tournament* is a structure with one binary relation $\to$ such that, for any two distinct elements $x$ and $y$, exactly one of $x \to y$ and $y \to x$ holds, and $x \to x$ never holds. Think of a round-robin competition in which every two players meet once and $x \to y$ means that $x$ beat $y$. The class of finite tournaments has the hereditary property. For amalgamation, let $\mathfrak{B}$ and $\mathfrak{C}$ be members whose universes meet exactly in the universe $A$ of a common substructure $\mathfrak{A}$, and for every new element $b$ of $\mathfrak{B}$ and new element $c$ of $\mathfrak{C}$ declare $b \to c$. The result is a tournament containing $\mathfrak{B}$ and $\mathfrak{C}$ as induced substructures, and with $\mathfrak{A}$ empty the same recipe gives joint embeddings. The limit is the *random tournament*, characterized by the extension property: for disjoint finite sets $U$ and $V$ of players, there is a player outside $U \cup V$ who beats every player of $V$ and loses to every player of $U$. As for graphs, a countable tournament whose results are decided by independent fair coins is almost surely the random tournament.

**Example 18.20 (Partial orders).** A *strict partial order* is a structure with one binary relation $<$ that is irreflexive and transitive; two elements may be incomparable. The class

of finite strict partial orders has the hereditary property, and disjoint unions, with no new comparabilities, give joint embeddings. For amalgamation, let $\mathfrak{B}$ and $\mathfrak{C}$ be members whose universes meet exactly in the universe $A$ of a common substructure $\mathfrak{A}$. On $B \cup C$, declare $x < y$ if $x < y$ in $\mathfrak{B}$ or in $\mathfrak{C}$, or if one of $x$ and $y$ lies in $B$, the other in $C$, and $x < a < y$ for some element $a$ of $A$. The new relation is transitive and irreflexive; the key point, as for equivalence relations, is that two elements of $A$ comparable in $\mathfrak{C}$ are comparable in the same way in $\mathfrak{A}$, and hence in $\mathfrak{B}$. The relation agrees with the old ones on $B$ and on $C$, so the resulting structure on $B \cup C$ is an amalgam. The limit is the *generic partial order*, a countable partial order that contains a copy of every finite partial order and in which every isomorphism between finite suborders extends to an automorphism.

**Example 18.21 (Disjoint edges).** Let $\mathcal{K}$ be the class of finite graphs in which every vertex has at most one neighbor. By Example 18.2(4) and Example 18.3(4), $\mathcal{K}$ is the age of the ultrahomogeneous graph $\mathcal{P}$ of disjoint edges, so $\mathcal{K}$ is a Fraïssé class and $\mathcal{P}$ is the limit. The amalgams of this class show why amalgamation must allow gluing. Let $\mathfrak{A}$ be a single vertex $a$, let $\mathfrak{B}$ be an edge $a - b$, and let $\mathfrak{C}$ be an edge $a - c$. In any amalgam, the images of $b$ and $c$ are both neighbors of the image of $a$, which has at most one neighbor, so the two images coincide. The smallest amalgam is the edge $a - b$ itself, with $c$ glued to $b$. The graph obtained from $\mathcal{P}$ by adding infinitely many vertices without neighbors has the same age but is not ultrahomogeneous: no automorphism moves a vertex without neighbors to a vertex with one.

**From finite to finitely generated structures.** The algebraic structures of the catalog have operations and constants, and we now extend the theory to them. One might try to replace each operation by a relation, the graph of the operation, but in the resulting relational structure every subset is a substructure, and the pieces would no longer be closed under the operations. The vocabularies of Appendix A may contain function symbols, among them constant symbols, and from now on we allow them; the vocabulary is still countable and has a single sort descriptor. A *substructure* of $\mathfrak{M}$ is then given by a subset of $M$ that contains the constants and is closed under the operations, with the predicates and the operations of $\mathfrak{M}$ restricted to that subset. An *embedding* is an injective map that preserves the operations and the constants and preserves both the truth and the failure of each relation; equivalently, an embedding of $\mathfrak{A}$ into $\mathfrak{M}$ is an isomorphism of $\mathfrak{A}$ onto a substructure of $\mathfrak{M}$. For a subset $X$ of $M$, the intersection of all substructures of $\mathfrak{M}$ whose universes contain $X$ is again a substructure, the substructure *generated* by $X$. The elements of that substructure are obtained from the elements of $X$ and the constants by finitely many applications of the operations. A structure is *finitely generated* if the structure is generated by a finite subset. A finitely generated structure for a countable vocabulary is countable, since only countably many expressions combine finitely many generators by the operations. For a relational vocabulary every subset is a substructure, and the finitely generated structures are the finite ones. For example, for the vocabulary of rings, with function symbols for $+$, $-$ and $\cdot$ and constant symbols for 0 and 1, the substructure of $\mathbb{C}$ generated by $\sqrt{2}$ is the ring $\mathbb{Z}[\sqrt{2}]$; if a function symbol for $x \mapsto x^{-1}$, interpreted with $0^{-1} = 0$, is added to the vocabulary, the substructure generated by $\sqrt{2}$ is the field $\mathbb{Q}(\sqrt{2})$.

The *age* of a structure is now the class of finitely generated structures that embed into the structure, including all their isomorphic copies. Finite universality, ultrahomogeneity and the extension property are defined as before, with "finitely generated" in place of "finite" and "substructure" in place of "induced substructure".

**Definition 18.22.** Let $\mathcal{K}$ be a non-empty class of finitely generated structures for a fixed countable vocabulary, closed under isomorphism and with only countably many isomorphism types. We call $\mathcal{K}$ a *Fraïssé class* if $\mathcal{K}$ has the joint embedding and amalgamation properties of Definition 18.5 and the following hereditary property: every finitely generated substructure of a member of $\mathcal{K}$ belongs to $\mathcal{K}$.

For classical structures and a finite relational vocabulary the countability condition holds automatically, as we observed after Definition 18.5; for fields and groups the condition has to be checked. We no longer require structures of arbitrarily large size, so the limit in the next theorem may be finite.

**Theorem 18.23 (Fraïssé).** *Let $\mathcal{K}$ be a Fraïssé class of finitely generated structures for a countable vocabulary. Then there is a countable ultrahomogeneous structure whose age is $\mathcal{K}$, and any two such structures are isomorphic.*

The structure is again called the *Fraïssé limit* of $\mathcal{K}$. We follow the presentation in Section 7.1 of the book of Hodges [Hod93].

*Proof.* The proof of Theorem 18.9 applies with three changes.

*One generator at a time.* In Lemma 18.8, the challenges that add one element become challenges that add one generator: $\mathfrak{B}$ is generated by $A$ together with one more element. If $b_1, \ldots, b_m$ are elements of $B$ such that $A \cup \{b_1, \ldots, b_m\}$ generates $\mathfrak{B}$, the structures $\mathfrak{B}_j$ of the proof are the substructures generated by $A \cup \{b_1, \ldots, b_j\}$, which are finitely generated and belong to $\mathcal{K}$ by the hereditary property.

*Countably many requests.* A finitely generated structure may have infinitely many finitely generated substructures, and infinitely many embeddings into the structure built so far, so a single stage can no longer meet all the requests that arise at that stage. Start the chain with $\mathfrak{C}_0 = \mathfrak{B}_0$. A request that refers to $\mathfrak{C}_k$ is a triple $(i, \mathfrak{A}, f)$, where $\mathfrak{A}$ is a finitely generated substructure of $\mathfrak{B}_i$ and $f \colon \mathfrak{A} \to \mathfrak{C}_k$ is an embedding. There are only countably many such requests, since $B_i$ has only countably many finite subsets and an embedding of $\mathfrak{A}$ into the countable structure $\mathfrak{C}_k$ is determined by the images of finitely many generators of $\mathfrak{A}$. There is at least one, since $\mathfrak{C}_k$ is isomorphic to some $\mathfrak{B}_i$. Once $\mathfrak{C}_k$ is built, list the requests that refer to $\mathfrak{C}_k$ as $r_{k,0}, r_{k,1}, \ldots$, with repetitions if there are only finitely many. Fix a bijection $\pi$ of $\omega$ onto $\omega \times \omega$ such that the first coordinate of $\pi(n)$ is at most $n$ for every $n \in \omega$; the inverse of the Cantor pairing function has this property. At stage $n$, write $\pi(n) = (k, j)$ and $r_{k,j} = (i, \mathfrak{A}, f)$. As before, first embed $\mathfrak{C}_n$ and $\mathfrak{B}_n$ into a common member of $\mathcal{K}$, and then amalgamate the result with $\mathfrak{B}_i$ over $\mathfrak{A}$, using $f$ followed by the inclusion of $\mathfrak{C}_k$ on one side and the inclusion of $\mathfrak{A}$ into $\mathfrak{B}_i$ on the other; a copy of the amalgam that contains $\mathfrak{C}_n$ as a substructure is $\mathfrak{C}_{n+1}$. Every finitely generated substructure of the union $\mathfrak{M}$ of the chain, whose universe is $M = \bigcup_{n \in \omega} C_n$, is generated by finitely many elements, which lie in a single $C_n$, and therefore belongs to $\mathcal{K}$. For the extension property, an embedding $f \colon \mathfrak{A} \to \mathfrak{M}$ of a member $\mathfrak{A}$ of $\mathcal{K}$ has image in some $C_k$, since $f(A)$ is generated by the images of finitely many generators of $\mathfrak{A}$. The corresponding request refers to $\mathfrak{C}_k$, and the request is met at the stage $n$ with $\pi(n) = (k, j)$, for the appropriate $j$; that stage satisfies $n \geq k$, so $\mathfrak{C}_k$ has already been built.

*Back-and-forth over finitely generated substructures.* The partial isomorphisms are now isomorphisms between finitely generated substructures, and a forth step passes from the domain $A$ of the current isomorphism to the substructure generated by $A$ and the next element of $M$, which is again finitely generated. For uniqueness, the back-and-forth construction starts with the isomorphism between the substructures of $\mathfrak{M}$ and of $\mathfrak{N}$ generated by the empty set: the substructure of $\mathfrak{M}$ generated by the empty set belongs to $\mathcal{K}$ and therefore embeds into

$\mathfrak{N}$, and the image of that embedding is the substructure of $\mathfrak{N}$ generated by the empty set, since embeddings preserve constants. ☺

**Corollary 18.24.** *Let $\mathcal{K}$ be a Fraïssé class of finitely generated structures for a countable vocabulary, and let $\mathfrak{M}$ be a countable structure whose age is contained in $\mathcal{K}$. Suppose that, whenever $\mathfrak{A} \subseteq \mathfrak{B}$ are members of $\mathcal{K}$ with $\mathfrak{A}$ a substructure of $\mathfrak{B}$ and $\mathfrak{B}$ generated by $A$ and one more element, every embedding $\mathfrak{A} \to \mathfrak{M}$ extends to an embedding $\mathfrak{B} \to \mathfrak{M}$. Then $\mathfrak{M}$ is the Fraïssé limit of $\mathcal{K}$.*

*Proof.* As in the proof of Corollary 18.10, the structure $\mathfrak{M}$ has the extension property for $\mathcal{K}$, and back-and-forth shows that $\mathfrak{M}$ is ultrahomogeneous. Let $\mathfrak{B}$ be a member of $\mathcal{K}$, let $\mathfrak{A}$ be the substructure of $\mathfrak{B}$ generated by the empty set, and let $\mathfrak{M}_0$ be the substructure of $\mathfrak{M}$ generated by the empty set, which belongs to $\mathcal{K}$. By the joint embedding property, $\mathfrak{B}$ and $\mathfrak{M}_0$ embed into a common member $\mathfrak{D}$ of $\mathcal{K}$. Embeddings preserve constants, so $\mathfrak{A}$ and $\mathfrak{M}_0$ are both mapped onto the substructure of $\mathfrak{D}$ generated by the empty set, and composing the first map with the inverse of the second gives an embedding of $\mathfrak{A}$ into $\mathfrak{M}$. The extension property extends this embedding to an embedding of $\mathfrak{B}$ into $\mathfrak{M}$, so the age of $\mathfrak{M}$ is exactly $\mathcal{K}$. ☺

We now return to the algebraic structures of the catalog. In these examples we follow the usage of algebra and denote a vector space, a field, a group or a Boolean algebra and its universe by the same roman letter.

**Example 18.25 (Vector spaces over a finite field).** Let $p$ be a prime, and use the vocabulary of abelian groups, with function symbols for $+$ and $-$ and a constant symbol for 0. A vector space over $\mathbb{F}_p$ is an abelian group in which $px = 0$ for every element $x$; multiplication by a scalar is repeated addition, so the substructures of the space are the subspaces. Let $\mathcal{K}$ be the class of finite vector spaces over $\mathbb{F}_p$. The hereditary property holds, there is one isomorphism type for each finite dimension, and direct sums give joint embeddings. For amalgamation, let $f\colon A \to B$ and $g\colon A \to C$ be embeddings, and let $D$ be the quotient of $B \oplus C$ by the subspace $\{(f(a), -g(a)) : a \in A\}$, the pushout of $f$ and $g$. Send $b \in B$ to the class of $(b, 0)$ and $c \in C$ to the class of $(0, c)$. The first map is injective, since $(b, 0) = (f(a), -g(a))$ forces $g(a) = 0$, hence $a = 0$ and $b = 0$; the second is injective for the same reason, and the two maps agree on $A$, since $(f(a), 0)$ and $(0, g(a))$ differ by an element of the subspace. Let $V$ be the space of finitely supported sequences of elements of $\mathbb{F}_p$, a countable space of infinite dimension. Every finitely generated subspace of $V$ is finite. A one-generator challenge for $V$ consists of an embedding $f\colon A \to V$ and a space $B$ generated by $A$ and one element $b$ outside $A$, so that $B = A \oplus \mathbb{F}_p b$. Choose a vector $v \in V$ outside the finite subspace $f(A)$, and extend $f$ linearly by $b \mapsto v$. By Corollary 18.24, the space $V$ is the Fraïssé limit of $\mathcal{K}$.

**Example 18.26 (Finite fields).** Use the vocabulary of rings, and let $\mathcal{K}$ be the class of finite fields of characteristic $p$. A substructure of a field is a subring that contains 1. A finite subring of a field has no zero divisors, and a finite ring without zero divisors is a field, since multiplication by a non-zero element is injective and hence surjective. The hereditary property therefore holds, and there is one isomorphism type for each number $p^n$ of elements. We use two facts from the theory of finite fields: the field $\mathbb{F}_{p^n}$ with $p^n$ elements has exactly one subfield with $p^d$ elements for each divisor $d$ of $n$, and the automorphisms of $\mathbb{F}_{p^n}$ are the powers of the Frobenius map $\varphi(x) = x^p$. For amalgamation, let $f\colon A \to B$ and $g\colon A \to C$ be embeddings, where $B$ has $p^m$ elements and $C$ has $p^n$ elements, and let $D = \mathbb{F}_{p^l}$, where $l$ is the least common multiple of $m$ and $n$. Choose embeddings $u_0\colon B \to D$ and $v\colon C \to D$. The

embeddings $u_0 \circ f$ and $v \circ g$ map $A$ onto the unique subfield $E$ of $D$ with as many elements as $A$, so $(v \circ g) \circ (u_0 \circ f)^{-1}$ is an automorphism of $E$, the restriction of a power $\varphi^r$ of the Frobenius map of $D$. The embeddings $u = \varphi^r \circ u_0$ and $v$ satisfy $u \circ f = v \circ g$; with $A$ the prime field $\mathbb{F}_p$, the same argument gives joint embeddings. The algebraic closure $\overline{\mathbb{F}}_p$ is countable, and finitely many of the elements of $\overline{\mathbb{F}}_p$ lie in a common finite subfield, so the age of $\overline{\mathbb{F}}_p$ is contained in $\mathcal{K}$. For a one-generator challenge, let $A \subseteq B$ be finite fields and $f\colon A \to \overline{\mathbb{F}}_p$ an embedding. The field $\overline{\mathbb{F}}_p$ has exactly one subfield with $p^n$ elements for each $n$, so some embedding $e\colon B \to \overline{\mathbb{F}}_p$ exists, and the maps $e$ and $f$ send $A$ onto the same subfield. As in the amalgamation, a power of the Frobenius map of $\overline{\mathbb{F}}_p$ corrects $e$ on $A$ and yields an embedding of $B$ that extends $f$. By Corollary 18.24, the field $\overline{\mathbb{F}}_p$ is the Fraïssé limit of $\mathcal{K}$.

**Example 18.27 (Number fields).** Add to the vocabulary of rings a function symbol for $x \mapsto x^{-1}$, interpreted with $0^{-1} = 0$, so that the substructures of a field are the subfields, and let $\mathcal{K}$ be the class of number fields. A subfield of a number field has finite degree over $\mathbb{Q}$, so the hereditary property holds. By the primitive element theorem, every number field is generated by a root of an irreducible polynomial with rational coefficients, and there are only countably many isomorphism types. The field $\mathbb{Q}$ is the substructure generated by the empty set in every number field, so joint embedding is the case of amalgamation over $\mathbb{Q}$. For amalgamation, let $f\colon A \to B$ and $g\colon A \to C$ be embeddings, and regard $B$ and $C$ as vector spaces over $A$ through $f$ and $g$. The tensor product $B \otimes_A C$ is a non-zero commutative ring of finite dimension over $A$; for a maximal ideal $\mathfrak{m}$ of the ring, the quotient $D = (B \otimes_A C)/\mathfrak{m}$ is a field of finite degree over $\mathbb{Q}$. The maps that send $b$ to the class of $b \otimes 1$ and $c$ to the class of $1 \otimes c$ are ring homomorphisms between fields, hence embeddings, and the two maps agree on $A$, since $f(a) \otimes 1 = 1 \otimes g(a)$. The field $\overline{\mathbb{Q}}$ is countable, and finitely many algebraic numbers generate a number field, so the age of $\overline{\mathbb{Q}}$ is contained in $\mathcal{K}$. For a one-generator challenge, let $B = A(b)$ and let $f\colon A \to \overline{\mathbb{Q}}$ be an embedding. Apply $f$ to the coefficients of the minimal polynomial of $b$ over $A$; the resulting polynomial is irreducible over $f(A)$ and has a root $\beta$ in $\overline{\mathbb{Q}}$, which is algebraically closed. The isomorphism of $A$ onto $f(A)$ then extends to an isomorphism of $A(b)$ onto $f(A)(\beta)$ that sends $b$ to $\beta$. By Corollary 18.24, the field $\overline{\mathbb{Q}}$ is the Fraïssé limit of $\mathcal{K}$. Ultrahomogeneity is the classical theorem that every isomorphism between subfields of $\overline{\mathbb{Q}}$ of finite degree extends to an automorphism of $\overline{\mathbb{Q}}$.

**Example 18.28 (Finite groups).** Use the vocabulary of groups, with function symbols for multiplication and inversion and a constant symbol for the identity, and let $\mathcal{K}$ be the class of finite groups. The substructures of a group are the subgroups, so the hereditary property holds, and direct products give joint embeddings. Amalgamation is a theorem of B. H. Neumann, proved with his permutational products [Neu60], which we do not reproduce. The Fraïssé limit of $\mathcal{K}$ is a countable group into which every finite group embeds, whose finitely generated subgroups are finite, and in which every isomorphism between finite subgroups extends to an automorphism; the limit is Hall's group $\mathbb{H}$ of the catalog, constructed directly by Hall.

**Example 18.29 (Finite Boolean algebras and the Cantor set).** Use the vocabulary of Boolean algebras, with function symbols for meet, join and complement and constant symbols for 0 and 1, and let $\mathcal{K}$ be the class of finite Boolean algebras. Every finite Boolean algebra is isomorphic to the algebra $\mathcal{P}(X)$ of all subsets of a finite set $X$, the set of atoms of the algebra. A subalgebra of a finite Boolean algebra is a finite Boolean algebra, and there is one isomorphism type for each finite number of atoms. An embedding of $\mathcal{P}(Z)$ into $\mathcal{P}(X)$ has the form $S \mapsto s^{-1}(S)$ for a surjection $s\colon X \to Z$, namely the map that sends $x$ to the element $z$ of $Z$ such that the image of $\{z\}$ contains $x$. For amalgamation, let the

embeddings of $\mathcal{P}(Z)$ into $\mathcal{P}(X)$ and $\mathcal{P}(Y)$ come from surjections $s\colon X \to Z$ and $t\colon Y \to Z$. The fiber product $W = \{(x, y) \in X \times Y : s(x) = t(y)\}$ projects onto $X$ and onto $Y$, and the two projections agree after composition with $s$ and $t$; the corresponding embeddings of $\mathcal{P}(X)$ and $\mathcal{P}(Y)$ into $\mathcal{P}(W)$ therefore agree on $\mathcal{P}(Z)$. Joint embedding is amalgamation over the two-element algebra $\{0, 1\}$, the substructure generated by the empty set. Let $\mathbb{B}$ be the algebra of clopen subsets of the Cantor set $2^\omega$, which is countable, since every clopen set is a finite union of basic sets $N_u$. Every finitely generated Boolean algebra is finite, so the age of $\mathbb{B}$ is contained in $\mathcal{K}$. For a one-generator challenge, let $A \subseteq B$ be finite Boolean algebras with $B$ generated by $A$ and one element $b$, and let $f\colon A \to \mathbb{B}$ be an embedding. Each atom $a$ of $A$ splits in $B$ into $a \wedge b$ and $a \wedge \neg b$, and the atoms of $B$ are those of these elements that are not 0. The images $f(a)$ of the atoms of $A$ form a partition of $2^\omega$ into non-empty clopen sets. Every non-empty clopen set $U$ splits into two non-empty clopen sets, since $U$ contains a basic set $N_u$, and $U$ is the union of $N_{u0}$ and $U \setminus N_{u0}$, which contains $N_{u1}$. Choose a clopen set $c$ that, for each atom $a$ of $A$, contains $f(a)$ when $a \wedge \neg b = 0$, misses $f(a)$ when $a \wedge b = 0$, and otherwise meets $f(a)$ in a non-empty proper clopen subset of $f(a)$. The extension of $f$ that sends $b$ to $c$ is an embedding of $B$ into $\mathbb{B}$, since the atoms of $B$ go to non-empty sets. By Corollary 18.24, the algebra $\mathbb{B}$ is the Fraïssé limit of $\mathcal{K}$; the countable atomless Boolean algebra of the catalog is $\mathbb{B}$.

**Real-valued predicates.** So far the structures have been classical. The construction also applies to relational structures whose predicates take arbitrary values in their intervals, as the structures of Appendix A allow. A predicate symbol $p$ of arity $n$ is then interpreted in a structure $\mathfrak{M}$ by a function $p^{\mathfrak{M}}\colon M^n \to [-r_p, r_p]$, and no metric or continuity of the interpretations is assumed. An *exact embedding* $f\colon \mathfrak{M} \to \mathfrak{N}$ is an injective map satisfying

$$p^{\mathfrak{N}}(f(a_1), \dots, f(a_n)) = p^{\mathfrak{M}}(a_1, \dots, a_n)$$

for every predicate symbol $p$ of arity $n$ and every tuple $(a_1, \dots, a_n) \in M^n$; for classical structures, exact embeddings are the embeddings defined above. A finite substructure is obtained by restricting the predicates to a finite subset. The age is the class of finite structures that embed into the structure, including their isomorphic copies. Ultrahomogeneity again means that every isomorphism between finite substructures extends to an automorphism.

**Proposition 18.30.** *Let $\mathcal{K}$ be a non-empty class of finite real-valued structures in a fixed finite relational vocabulary. Suppose that $\mathcal{K}$ is closed under isomorphism, contains structures of arbitrarily large finite size, and has only countably many isomorphism types. If $\mathcal{K}$ has the hereditary, joint embedding and amalgamation properties for exact embeddings, then there is a countably infinite ultrahomogeneous structure whose age is $\mathcal{K}$, and any two such structures are isomorphic.*

*Proof.* In the proof of Theorem 18.9, the finiteness of the relational vocabulary and the classical values of the predicates were used only to obtain a countable list of representatives. That list is now supplied by the countability hypothesis. At each stage there are still only finitely many substructures of each listed finite structure and finitely many embeddings into the current stage. Thus the same amalgamation construction meets every extension request. The union of the chain preserves the predicate values already assigned, has age $\mathcal{K}$, and has the extension property. Back-and-forth proves ultrahomogeneity and uniqueness as before. The union of the increasing sequence of finite partial isomorphisms preserves every predicate value, since each finite tuple lies in the domain of some stage. Alternately covering enumerations of both structures makes that union an isomorphism. ☺

The countability hypothesis is essential here, even with a single unary predicate. For each $r \in [0,1]$, let $\mathfrak{A}_r$ be the one-element structure whose unique element has predicate value $r$. The structures $\mathfrak{A}_r$ are pairwise non-isomorphic. A structure containing an exact copy of every $\mathfrak{A}_r$ must realize every value in $[0,1]$, so the universe of that structure cannot be countable. In fact, the class of all finite structures with one $[0,1]$-valued unary predicate has the hereditary, joint embedding and amalgamation properties: amalgamation identifies the common part and otherwise takes disjoint copies. The three properties therefore do not by themselves give a countable limit in the real-valued setting.

Restricting all predicate values to a fixed countable subset of $[0,1]$ guarantees the countability hypothesis for a finite relational vocabulary. On a finite universe, each predicate has only finitely many entries to specify; there are therefore only countably many structures of each finite size, up to isomorphism. The other hypotheses of Proposition 18.30 must still be checked.

**Example 18.31 (Rational labels).** Let $Q = \mathbb{Q} \cap [0,1]$ and consider all finite structures for a vocabulary with one unary predicate symbol $p$ with a priori bound 1 whose values lie in $Q$; think of each element as carrying a rational label. The class is closed under isomorphism and induced substructures and contains arbitrarily large structures. Disjoint unions give joint embeddings; identifying the common part in otherwise disjoint copies gives amalgamation. The limit is the structure $\mathfrak{L}$ with universe $Q \times \mathbb{N}$ and $p^{\mathfrak{L}}(q,n) = q$. Every finite structure in the class embeds into $\mathfrak{L}$, since each value has infinitely many realizations. For each $q \in Q$, the set $\{q\} \times \mathbb{N}$ consists exactly of the elements with predicate value $q$. A finite partial isomorphism restricts to a bijection between two finite subsets of each such set. The complements of these subsets in $\{q\} \times \mathbb{N}$ are both countably infinite, so the bijection extends to a permutation of $\{q\} \times \mathbb{N}$. Combining these permutations gives an automorphism, which proves ultrahomogeneity, and Proposition 18.30 gives uniqueness. The construction concerns countable structures; the limit involves neither a completion nor approximate embeddings.

**Example 18.32 (The rational Urysohn space).** A finite metric space with distances at most 1 is a finite structure for a vocabulary with one binary predicate symbol $d$ with a priori bound 1, interpreted by a function $d$ satisfying $d(x,y) = 0$ if and only if $x = y$, symmetry, and the triangle inequality $d(x,z) \le d(x,y) + d(y,z)$. An exact embedding is then a distance-preserving map, an *isometric embedding*. Let $\mathcal{K}$ be the class of finite metric spaces with distances in $Q = \mathbb{Q} \cap [0,1]$. The class is closed under isomorphism and subspaces, contains arbitrarily large spaces, and has only countably many isomorphism types. For amalgamation, let $\mathfrak{B}$ and $\mathfrak{C}$ be metric spaces in $\mathcal{K}$ whose universes meet exactly in the universe $A$ of a common subspace $\mathfrak{A}$, and for new points $b \in B \setminus A$ and $c \in C \setminus A$ set

$$d(b,c) = \min\Big(1,\ \min_{a \in A}\big(d(b,a) + d(a,c)\big)\Big),$$

where the inner minimum is read as 1 when $\mathfrak{A}$ is empty. The new distance is rational, positive, and at most 1. The triangle inequality survives. For instance, for $b, b' \in B$ and a new point $c$ of $C$, either $d(b',c) = 1 \ge d(b,c)$, or $d(b',c) = d(b',a) + d(a,c)$ for some $a \in A$ and then

$$d(b,b') + d(b',c) \ge d(b,a) + d(a,c) \ge d(b,c);$$

for two points $b, b'$ of $B$ and a new point $c$ of $C$, the bound $d(b,b') \le d(b,c) + d(c,b')$ follows in the same way from $d(a,c) + d(c,a') \ge d(a,a')$ in $\mathfrak{C}$, and the remaining cases are symmetric. With $\mathfrak{A}$ empty, the formula places $\mathfrak{B}$ and $\mathfrak{C}$ at distance 1 from each other and gives joint embeddings. The limit, by Proposition 18.30, is the *rational Urysohn space* of diameter 1, a countable metric space that contains an isometric copy of every finite metric space with distances in $Q$. A one-element challenge asks for a new point $x$ at prescribed

positive distances $r_a \in Q$ from the points $a$ of a finite set, where the numbers $r_a$ satisfy $|r_a - r_{a'}| \le d(a, a') \le r_a + r_{a'}$, the conditions that make the enlarged space metric. The completion of the rational Urysohn space is Urysohn's universal space of diameter 1, which contains an isometric copy of every separable metric space of diameter at most 1.

| **Class of finite structures** | **Amalgam over $A$** | **Fraïssé limit** |
|---|---|---|
| sets | union | countably infinite set |
| colored sets | union | infinitely many points of each color |
| equivalence relations | join classes through $A$ | infinitely many infinite classes |
| linear orders | interleave in each gap of $A$ | $(\mathbb{Q}, <)$ |
| colored linear orders | interleave in each gap of $A$ | $(\mathbb{Q}, <)$ with both colors dense |
| graphs | no new edges | Rado graph |
| triangle-free graphs | no new edges | Henson graph |
| tournaments | new elements of $B$ beat those of $C$ | random tournament |
| partial orders | comparabilities through $A$ | generic partial order |
| graphs of degree at most one | glue the partners | infinitely many disjoint edges |
| vector spaces over $\mathbb{F}_p$ | pushout | space of countable dimension |
| finite fields of characteristic $p$ | a larger finite field | $\overline{\mathbb{F}}_p$ |
| number fields | quotient of the tensor product | $\overline{\mathbb{Q}}$ |
| finite groups | Neumann's theorem | Hall's universal group |
| finite Boolean algebras | fiber product of the sets of atoms | clopen algebra of the Cantor set |
| metric spaces with distances in $\mathbb{Q} \cap [0, 1]$ | shortest route through $A$, capped at 1 | rational Urysohn space of diameter 1 |

TABLE 2. Fraïssé classes of this section, with an amalgam and the limit of each.

The examples are collected in Table 2. The passage from Cantor's theorem to Fraïssé's theorem is thus a passage from one order to a general construction from finite and finitely generated structures, which reaches orders, graphs, fields, groups, Boolean algebras and metric spaces alike.

**From Fraïssé limits to logic.** The two features of a Fraïssé limit can be written down in the language of Appendix A, and the passage from a structure to the sentences that describe the structure is the beginning of model theory. Consider the Rado graph, and add to the vocabulary of graphs a binary predicate symbol $=$ with a priori bound 1, interpreted in every graph by equality: the value of $=$ at a pair of vertices is 1 when the two vertices coincide and $-1$ otherwise. For non-negative integers $m$ and $n$, let $\bar{x} = (x_1, \dots, x_m)$ and $\bar{y} = (y_1, \dots, y_n)$ be tuples of distinct variables, and let $\theta_{m,n}(\bar{x}, \bar{y})$ be the formula

$$\begin{aligned}(x_1 = y_1) \vee \cdots \vee (x_m = y_n) \\ \vee \sup_z \big( E(z, x_1) \wedge \cdots \wedge E(z, x_m) \wedge -E(z, y_1) \wedge \cdots \wedge -E(z, y_n) \\ \wedge -(z = y_1) \wedge \cdots \wedge -(z = y_n) \big),\end{aligned}$$

in which an empty disjunction is omitted and an empty conjunction stands for the logical unity constant 1. In a graph, the formula $\theta_{m,n}$ is two-valued, and the axiom $\theta_{m,n} \geq 1$, required of all values of $\bar{x}$ and $\bar{y}$, says the following: whenever no vertex of $\bar{x}$ is a vertex of $\bar{y}$, some vertex outside $\bar{y}$ is adjacent to every vertex of $\bar{x}$ and to no vertex of $\bar{y}$. Since no vertex is adjacent to itself, such a vertex also lies outside $\bar{x}$. By Corollary 18.10, every countable graph that satisfies all these axioms is the Rado graph; in the language of model theory, the axioms, together with those of graphs, form a theory with exactly one countable model up to isomorphism.

A second theorem connects the Rado graph with finite graphs. Glebskiĭ, Kogan, Liogon'kiĭ and Talanov in 1969 [GKLT69], and independently Fagin in 1976 [Fag76], proved a *zero-one law*. Let $\theta$ be a sentence for the vocabulary of graphs with $=$, built from the atomic formulas by the connectives $-$, $\vee$, $\wedge$, sup and inf, so that $\theta$ is two-valued in every graph. Then the proportion of the graphs on the vertex set $\{1, \ldots, n\}$ in which $\theta$ has value 1 tends, as $n$ tends to infinity, either to 1 or to 0, and the limit is 1 exactly when $\theta$ has value 1 in the Rado graph. Fagin's proof rests on the axioms $\theta_{m,n} \geq 1$ and on the uniqueness of their countable model, that is, on the back-and-forth argument of this section. Run for finitely many rounds, the back-and-forth argument becomes the Ehrenfeucht–Fraïssé game, the subject of the next sections (see Section 19), which measures how far two structures agree on sentences.

**Exercises.**

(1) Show that the cycle with five vertices is ultrahomogeneous and that the cycle with six vertices is not.
(2) Show that $(\mathbb{Q}, <)$ with the integers removed is isomorphic to $(\mathbb{Q}, <)$, and that $(\mathbb{Z}, <)$ and $(\mathbb{Q} \times \mathbb{Z}, <_{\text{lex}})$ are not ultrahomogeneous.
(3) Show that the class of finite equivalence relations with at most two classes is a Fraïssé class, and identify the limit. Show that the structure with infinitely many classes, all of size 2 except one class of size 3, is not ultrahomogeneous.
(4) Show that the Rado graph remains the Rado graph after the removal of one vertex, and also after the removal of finitely many edges. Show that the complement of the Rado graph, in which two distinct vertices are adjacent exactly when they are not adjacent in the Rado graph, is again the Rado graph.
(5) Show that every vertex of the Rado graph has infinitely many neighbors, and that the Rado graph and the Henson graph are not isomorphic.
(6) In the construction of Example 18.11, show that the union of the sets $C_n$ is exactly the set of dyadic rationals.
(7) Show that the class of finite graphs in which every vertex has at most two neighbors does not have the amalgamation property. (Amalgamate, over a single vertex $x$, a path with three vertices whose middle vertex is $x$ and a triangle containing $x$.)
(8) For a vocabulary with a binary predicate symbol $<$ and a unary predicate symbol $R$, show that the class of finite linear orders in which $R$ holds of at most one element has the hereditary and joint embedding properties but not the amalgamation property.
(9) Verify transitivity of the amalgam of partial orders in Example 18.20.
(10) Show that the class of finite abelian groups, for the vocabulary of abelian groups, is a Fraïssé class.
(11) Show that the class of all finite fields does not have the joint embedding property.
(12) Show that a countable structure whose age is the class of finite structures with one unary predicate taking values in $\mathbb{Q} \cap [0, 1]$, and which is ultrahomogeneous, is isomorphic to the structure $\mathfrak{L}$ of Example 18.31.

## 19. Elementary equivalence and the game

**The problem.** Two structures $\mathfrak{M}$ and $\mathfrak{N}$ in the same vocabulary are *elementarily equivalent*, written $\mathfrak{M} \equiv \mathfrak{N}$, if they satisfy the same first-order sentences. Isomorphic structures are elementarily equivalent, but the converse fails: the orders $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ are not isomorphic, since only the first is countable. Yet, as we show in Proposition 21.1 below, no first-order sentence tells the two orders apart. The problem is to find a tool that decides whether two given structures are elementarily equivalent without examining the sentences one by one. The same tool should also serve to prove that a given property of structures cannot be expressed by a first-order sentence.

The answer is the game sketched at the opening of this part, and in this section we make the game precise. Two notions come first: the structures on which the game is played, and the partial isomorphisms that Player II must preserve. We work with relational vocabularies, which keep the bookkeeping of the game to a minimum; function and constant symbols can be traded for relations, as we explain in Remark 20.9 below.

**Definition 19.1.** A *relational vocabulary* is a set $V$ of relation symbols, each with an *arity*, which is an integer $r \geq 1$. A *$V$-structure* $\mathfrak{M}$ consists of a non-empty set $M$, the *universe* of $\mathfrak{M}$, together with a relation $R^{\mathfrak{M}} \subseteq M^r$ for each symbol $R \in V$ of arity $r$.

In the language of Section 18, a $V$-structure is a classical structure for a relational vocabulary, written in the traditional notation of first-order logic: the relation $R^{\mathfrak{M}}$ is the set of tuples at which the corresponding two-valued predicate takes the value 1. From now on, $\mathfrak{M}$ and $\mathfrak{N}$ denote structures with universes $M$ and $N$, and we use the letters $a, a', a_1, \ldots$ for elements of $M$ and $b, b', b_1, \ldots$ for elements of $N$. Our examples are linear orders and graphs: a linear order is a structure in the vocabulary $\{<\}$ whose binary relation is a strict linear order. A graph is a structure in the vocabulary $\{E\}$ whose relation is symmetric and irreflexive. All tuples are finite; for a tuple $\bar{a} = (a_1, \ldots, a_k) \in M^k$ and an element $a \in M$ we write $\bar{a}a$ for the tuple $(a_1, \ldots, a_k, a)$, and similarly $\bar{a}\bar{a}^*$ for the concatenation of two tuples.

**Definition 19.2.** Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for a relational vocabulary $V$, let $k$ be a non-negative integer, and let $\bar{a} = (a_1, \ldots, a_k) \in M^k$ and $\bar{b} = (b_1, \ldots, b_k) \in N^k$ be tuples. The map $\bar{a} \mapsto \bar{b}$ is a *partial isomorphism* if

(1) for all $i, l \in \{1, \ldots, k\}$, we have $a_i = a_l$ if and only if $b_i = b_l$, and
(2) for every symbol $R \in V$ of arity $r$ and all indices $i_1, \ldots, i_r \in \{1, \ldots, k\}$, the following equivalence holds:

$$(a_{i_1}, \ldots, a_{i_r}) \in R^{\mathfrak{M}} \quad \text{if and only if} \quad (b_{i_1}, \ldots, b_{i_r}) \in R^{\mathfrak{N}}.$$

Condition (1) says that $a_i \mapsto b_i$ is a well-defined bijection between the sets $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_k\}$, and condition (2) says that this bijection is an isomorphism between the substructures with these domains. For $k = 0$ both conditions are vacuous, and the empty map is a partial isomorphism. Deleting, repeating or reordering matched pairs preserves a partial isomorphism. Each equality or relation tested after such a change was already tested before the change.

**Definition 19.3.** Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for a relational vocabulary, let $k$ and $n$ be non-negative integers, and let $\bar{a} = (a_1, \ldots, a_k) \in M^k$ and $\bar{b} = (b_1, \ldots, b_k) \in N^k$ be tuples. The *Ehrenfeucht–Fraïssé game* $\mathrm{EF}_n(\mathfrak{M}, \bar{a}, \mathfrak{N}, \bar{b})$ has $n$ rounds. In round $t$, where $1 \leq t \leq n$, Player I chooses an element of $M$ or an element of $N$, and Player II answers with an element of the other structure. We write $a_{k+t} \in M$ and $b_{k+t} \in N$ for the two elements chosen in round $t$, whichever player chose them. Player II wins the play if the map $(a_1, \ldots, a_{k+n}) \mapsto$

$(b_1, \dots, b_{k+n})$ is a partial isomorphism, and Player I wins otherwise. For $k = 0$ we write $\mathrm{EF}_n(\mathfrak{M}, \mathfrak{N})$.

Player I is often called *Spoiler*, since he tries to expose a difference, and Player II *Duplicator*, since she tries to copy his moves. Player I may switch structures from one round to the next, and this freedom is essential. The tuples $\bar{a}$ and $\bar{b}$ record elements matched before the game starts. These tuples cost nothing, and they make induction on the number of rounds possible, because after one round of $\mathrm{EF}_n(\mathfrak{M}, \bar{a}, \mathfrak{N}, \bar{b})$ what remains is a game of the same kind with $n-1$ rounds. For $n = 0$ there are no rounds, and Player II wins $\mathrm{EF}_0(\mathfrak{M}, \bar{a}, \mathfrak{N}, \bar{b})$ if and only if $\bar{a} \mapsto \bar{b}$ is a partial isomorphism.

**Example 19.4.** Let $\mathfrak{M}$ be the triangle, the graph with three pairwise adjacent vertices, and let $\mathfrak{N}$ be the path $b_1 - b_2 - b_3$, in which $b_1$ and $b_3$ are not adjacent. Player II wins $\mathrm{EF}_1(\mathfrak{M}, \mathfrak{N})$, since a single vertex matched with a single vertex is always a partial isomorphism (the relation $E$ is irreflexive in both graphs). Player I wins $\mathrm{EF}_2(\mathfrak{M}, \mathfrak{N})$: he chooses $b_1$ and then $b_3$ in $\mathfrak{N}$. If Player II answers with two different vertices of the triangle, they are adjacent, while $b_1$ and $b_3$ are not. If she answers twice with the same vertex, condition (1) of Definition 19.2 fails. The winning strategy of Player I mirrors the sentence $\exists x \, \exists y \, (\neg x = y \wedge \neg E(x, y))$, which holds in $\mathfrak{N}$ and fails in $\mathfrak{M}$: each quantifier becomes a move, played in the structure where the sentence is true. In the next section we turn this correspondence into a theorem (see Theorem 20.3). The play is drawn in Figure 29.

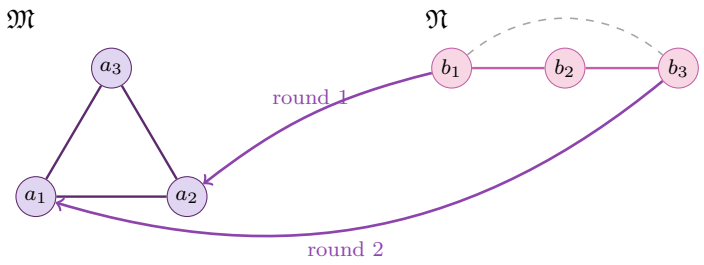



FIGURE 29. Player I chooses $b_1$ and then $b_3$ in the path, and whatever two vertices of the triangle Player II chooses in answer, they are adjacent while $b_1$ and $b_3$ are not.

**Remark 19.5.** The game is an instance of the games $(T, W)$ of Definition 2.2. A move of Player I is a pair $(1, a)$ with $a \in M$ or a pair $(2, b)$ with $b \in N$, the first coordinate naming the structure. Player II answers a move $(1, a)$ with some $(2, b)$ where $b \in N$, and a move $(2, b)$ with some $(1, a)$ where $a \in M$. The coordinates keep the structures apart even when $M$ and $N$ meet. After the $2n$-th move both players play a dummy move $*$ forever, and the payoff set $W$ consists of the plays whose first $2n$ moves satisfy the winning condition. Two consequences of this description are used below. First, the tree of positions depends only on $\mathfrak{M}$, $\mathfrak{N}$ and $n$; the tuples $\bar{a}$ and $\bar{b}$ enter only through the payoff set. Second, a play is decided by the first $2n$ moves. The game is therefore determined (see Exercise 4 at the end of Section 2), and exactly one player has a winning strategy (see Remarks 2.3(3)).

**Exercises.**

(1) Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for a relational vocabulary, and let $f \colon M \to N$ be an isomorphism of $\mathfrak{M}$ onto $\mathfrak{N}$. Show that Player II has a winning strategy in $\mathrm{EF}_n(\mathfrak{M}, \mathfrak{N})$ for every $n \geq 0$, by answering each move with the image of the move under $f$ or under $f^{-1}$.

(2) Let $V$ be the empty vocabulary, so a $V$-structure is just a non-empty set, and let $\mathfrak{M}$ and $\mathfrak{N}$ be $V$-structures with $m$ and $m'$ elements, where $m$ and $m'$ are positive integers. Show that, for every $n \geq 0$, Player II has a winning strategy in $\mathrm{EF}_n(\mathfrak{M}, \mathfrak{N})$ if and only if $m = m'$ or $m, m' \geq n$.
(3) In the setting of Example 19.4, suppose that Player I must choose all his elements in the triangle $\mathfrak{M}$. Show that Player II then wins the game with two rounds. Thus, in the example, the choice of the structure in which Player I plays is essential.
(4) Show that Player I wins $\mathrm{EF}_2((\omega, <), (\mathbb{Z}, <))$, but that Player II wins if Player I must choose both of his elements in the same structure. Thus the freedom to switch structures from one round to the next is essential. (Hint: Player I starts with the least element of $\omega$.)
(5) Let $C_4$ be the cycle with four vertices and $P_4$ the path with four vertices. Show that Player II wins $\mathrm{EF}_2(C_4, P_4)$ and that Player I wins $\mathrm{EF}_3(C_4, P_4)$. (Hint: an end of the path has only one neighbor.)

## 20. The Ehrenfeucht–Fraïssé theorem

In Example 19.4 the winning strategy of Player I followed a sentence quantifier by quantifier. The Ehrenfeucht–Fraïssé theorem, which we prove in this section, says that the correspondence is no accident. By this theorem (see Theorem 20.3), for structures for a finite relational vocabulary, Player I wins the game with $n$ rounds exactly when some sentence of quantifier rank at most $n$ holds in one of the structures and fails in the other. The number of rounds corresponds to the depth of nested quantifiers. We first define this depth, called quantifier rank, and fix the notation used in the proof.

**Definition 20.1.** Let $V$ be a relational vocabulary. The variables are $v_1, v_2, v_3, \ldots$, and the letters $x$, $y$, $z$, $u$, $w$, with or without indices, stand for variables.

(1) The *atomic formulas* are the constant $\top$ (true), the equations $x = y$, and the formulas $R(x_1, \ldots, x_r)$ with $R \in V$ of arity $r$. The *formulas* are built from the atomic formulas by negation $\neg$, conjunction $\wedge$ and existential quantification $\exists x$; the formulas $\varphi \vee \psi$, $\varphi \to \psi$ and $\forall x\, \varphi$ abbreviate $\neg(\neg\varphi \wedge \neg\psi)$, $\neg(\varphi \wedge \neg\psi)$ and $\neg\exists x\, \neg\varphi$.
(2) The *quantifier rank* $\mathrm{qr}(\varphi)$ is defined by recursion: $\mathrm{qr}(\alpha) = 0$ for atomic $\alpha$, $\mathrm{qr}(\neg\varphi) = \mathrm{qr}(\varphi)$, $\mathrm{qr}(\varphi \wedge \psi) = \max\{\mathrm{qr}(\varphi), \mathrm{qr}(\psi)\}$, and $\mathrm{qr}(\exists x\, \varphi) = \mathrm{qr}(\varphi) + 1$.

The quantifier rank measures the depth of nesting of quantifiers, not their number. For example, the sentence $\exists x\, \exists y\, (x < y)$ has rank 2, while the sentence $\exists x\, (x < x) \wedge \exists y\, (y = y)$, which also has two quantifiers, has rank 1. For the abbreviations, the recursion yields

$$\mathrm{qr}(\forall x\, \varphi) = \mathrm{qr}(\varphi) + 1 \qquad \text{and} \qquad \mathrm{qr}(\varphi \vee \psi) = \max\{\mathrm{qr}(\varphi), \mathrm{qr}(\psi)\}.$$

If $\bar{x} = (x_1, \ldots, x_k)$ is a list of distinct variables containing the free variables of $\varphi$, we write $\varphi(\bar{x})$. In that case, for $\bar{a} \in M^k$ we write $\mathfrak{M} \models \varphi(\bar{a})$ if $\varphi$ holds in $\mathfrak{M}$ when each $x_i$ is interpreted as $a_i$. In this notation Tarski's clause for $\exists$ reads as follows. Suppose that $\varphi = \exists z\, \psi$, and let $\bar{x}'$ and $\bar{a}'$ be obtained from $\bar{x}$ and $\bar{a}$ by deleting the entry of $z$ when $z$ occurs in $\bar{x}$. Then the free variables of $\psi$ lie in $\bar{x}'z$, and we have $\mathfrak{M} \models \varphi(\bar{a})$ if and only if $\mathfrak{M} \models \psi(\bar{a}', a)$ for some $a \in M$. Two formulas $\varphi(\bar{x})$ and $\psi(\bar{x})$ are *logically equivalent* if they hold of the same tuples in every $V$-structure.

**Definition 20.2.** Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for a relational vocabulary, let $k$ and $n$ be non-negative integers, and let $\bar{a} \in M^k$ and $\bar{b} \in N^k$ be tuples. We write $(\mathfrak{M}, \bar{a}) \equiv_n (\mathfrak{N}, \bar{b})$ if every formula $\varphi(\bar{x})$ of quantifier rank at most $n$, with $\bar{x}$ a list of $k$ distinct variables, satisfies

$$\mathfrak{M} \models \varphi(\bar{a}) \quad \text{if and only if} \quad \mathfrak{N} \models \varphi(\bar{b}).$$

For $k = 0$ we write $\mathfrak{M} \equiv_n \mathfrak{N}$: the two structures satisfy the same sentences of quantifier rank at most $n$.

Since every sentence has some quantifier rank, we have $\mathfrak{M} \equiv \mathfrak{N}$ if and only if $\mathfrak{M} \equiv_n \mathfrak{N}$ for every $n \geq 0$.

**Theorem 20.3 (Fraïssé, 1954; Ehrenfeucht, 1961).** *Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for a finite relational vocabulary.*

(1) *For all $k, n \geq 0$, $\bar{a} \in M^k$ and $\bar{b} \in N^k$, we have $(\mathfrak{M}, \bar{a}) \equiv_n (\mathfrak{N}, \bar{b})$ if and only if Player II has a winning strategy in $\mathrm{EF}_n(\mathfrak{M}, \bar{a}, \mathfrak{N}, \bar{b})$.*
(2) *For every $n \geq 0$, we have $\mathfrak{M} \equiv_n \mathfrak{N}$ if and only if Player II has a winning strategy in $\mathrm{EF}_n(\mathfrak{M}, \mathfrak{N})$.*
(3) *We have $\mathfrak{M} \equiv \mathfrak{N}$ if and only if Player II has a winning strategy in $\mathrm{EF}_n(\mathfrak{M}, \mathfrak{N})$ for every $n \geq 0$.*

The finite relational vocabulary has a specific role in the proof. On a fixed finite list of variables, there are only finitely many atomic formulas: equalities, applications of the finitely many relation symbols, and $\top$. If a unary function symbol $f$ were allowed, the same variable could occur in the infinitely many atomic formulas $f(x) = x$, $f(f(x)) = x$, and so on. All have quantifier rank 0, so bounding quantifier rank would not bound this complexity. The relational assumption avoids this problem and lets us build finite descriptions of game positions. Functions can instead be represented by their graphs, but this translation need not preserve quantifier rank; we return to this point in Remark 20.9.

In Example 19.4, the theorem says that the triangle and the path of three vertices satisfy the same sentences of quantifier rank 1, and that some sentence of rank 2 separates them. The proof has two directions. To obtain a strategy from $\equiv_n$, we construct a single formula describing a tuple through the game with $n$ rounds remaining (see Lemma 20.7). To pass from a strategy to $\equiv_n$, we use induction on formulas, matching each quantifier with a round. Two elementary observations about the game prepare the way.

**Lemma 20.4.** *Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for a relational vocabulary, let $k$ and $n$ be non-negative integers, and let $\bar{a} \in M^k$ and $\bar{b} \in N^k$ be tuples. Suppose that Player II has a winning strategy in $\mathrm{EF}_n(\mathfrak{M}, \bar{a}, \mathfrak{N}, \bar{b})$.*

(1) *The map $\bar{a} \mapsto \bar{b}$ is a partial isomorphism.*
(2) *Player II still has a winning strategy if any of the initially matched pairs are deleted, repeated or reordered, including if all of them are deleted.*

*Proof.* Fix a winning strategy for Player II. In any play that follows the strategy, the initial matching together with the $n$ new pairs is a partial isomorphism. Restricting this matching to the initial pairs proves (1); when $n = 0$, assertion (1) is the winning condition itself.

For (2), use the same strategy after changing the list of initial pairs. The legal moves are unchanged, so every play following this strategy also occurs in the original game. The final matching of such a play differs from the original one only by the deletion, repetition or reordering of initial pairs. Every equality and relation tested in the new matching was therefore already tested in the original one, with the same truth value. Thus the new matching is a partial isomorphism, and the strategy still wins. ☺

The next lemma separates the first round from the rest of the game.

**Lemma 20.5.** *Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for a relational vocabulary, let $n$ be a positive integer and $k$ a non-negative integer, and let $\bar{a} \in M^k$ and $\bar{b} \in N^k$ be tuples. Player II has a winning strategy in $\mathrm{EF}_n(\mathfrak{M}, \bar{a}, \mathfrak{N}, \bar{b})$ if and only if she can answer every opening move so*

*that she has a winning strategy for the remaining $n-1$ rounds, with the first pair added to the initial matching.*

*Proof.* Suppose that the first round matches $a \in M$ with $b \in N$, in either order. What remains is precisely $\mathrm{EF}_{n-1}(\mathfrak{M}, \bar{a}a, \mathfrak{N}, \bar{b}b)$: the remaining moves and the final matching are the same in both games.

If Player II has a winning strategy for $n$ rounds, let $b$ be the response of the strategy to an opening move $a \in M$. Continue to use that strategy, with the first round kept as part of the history. The continued strategy wins the shorter game. The same argument applies to opening moves in $N$.

Conversely, for each opening move choose a response and a winning strategy for the resulting shorter game. Player II gives the chosen response and then follows that strategy. Every resulting play ends with a partial isomorphism, so she wins the game with $n$ rounds. ☺

When Player I opens in $M$, the condition says that every $a \in M$ has a response $b \in N$ from which Player II can win the shorter game: the *forth* condition. Opening moves in $N$ give the *back* condition, with the roles of the structures reversed.

**Hintikka formulas.** We now express the winning condition by a formula. With no rounds left, the formula records the equalities and relations among the matched elements. For $n+1$ rounds, the formula lists the rank-$n$ descriptions of the one-element extensions. Existential quantifiers require every listed description to occur; a universal quantifier rules out all others. The two requirements are the forth and back conditions of Lemma 20.5. The resulting formulas are called *Hintikka formulas*.

To make this construction a first-order formula, we must show that each list is finite, even when the structures are infinite. The families in the following definition will establish this finiteness.

Fix a finite relational vocabulary $V$ and an enumeration of all $V$-formulas. For a finite set $\Phi$ of formulas, we write $\bigwedge \Phi$ for the conjunction of the members of $\Phi$ in the order of the enumeration ($\top$ if $\Phi = \varnothing$), and $\bigvee \Phi$ for $\neg \bigwedge\{\neg\varphi : \varphi \in \Phi\}$. Both are definite formulas, whose quantifier rank is the largest rank of a member of $\Phi$ (0 if $\Phi = \varnothing$). The first holds of a tuple exactly when every member of $\Phi$ does, and the second exactly when some member does. For $k \geq 0$, let $\Lambda_k$ be the set of atomic formulas with variables among $v_1, \ldots, v_k$, together with their negations.

**Definition 20.6.** Let $V$ be a finite relational vocabulary. For $n, k \geq 0$, the sets $H_{n,k}$ of formulas are defined by recursion on $n$:

$$H_{0,k} = \Big\{ \bigwedge \Delta : \Delta \subseteq \Lambda_k \Big\},$$
$$H_{n+1,k} = \Big\{ \bigwedge \big\{ \exists v_{k+1}\, \chi : \chi \in S \big\} \wedge \forall v_{k+1} \bigvee S \ : \ S \subseteq H_{n,k+1} \Big\}.$$

For a $V$-structure $\mathfrak{M}$ and $\bar{a} \in M^k$, the *Hintikka formula of rank $n$* of $\bar{a}$ in $\mathfrak{M}$ is defined by recursion on $n$:

$$\varphi^0_{\bar{a}} = \bigwedge \big\{ \alpha \in \Lambda_k : \mathfrak{M} \models \alpha(\bar{a}) \big\},$$
$$\varphi^{n+1}_{\bar{a}} = \bigwedge \big\{ \exists v_{k+1}\, \varphi^n_{\bar{a}a} : a \in M \big\} \wedge \forall v_{k+1} \bigvee \big\{ \varphi^n_{\bar{a}a} : a \in M \big\}.$$

The notation $\varphi^n_{\bar{a}}$ suppresses the structure $\mathfrak{M}$. Although the definition ranges over all elements of $M$, only finitely many distinct extension formulas occur. Part (1) of the next lemma proves the finiteness by induction, justifying the conjunctions and disjunctions at each stage.

**Lemma 20.7.** *Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for a finite relational vocabulary $V$, let $n$ and $k$ be non-negative integers, and let $\bar{a} \in M^k$ and $\bar{b} \in N^k$ be tuples.*

(1) *The set $H_{n,k}$ consists of finitely many formulas of quantifier rank at most $n$ with free variables among $v_1, \dots, v_k$, and contains $\varphi^n_{\bar{a}}$.*
(2) *The formula $\varphi^n_{\bar{a}}$ holds of $\bar{a}$ in $\mathfrak{M}$.*
(3) *We have $\mathfrak{N} \models \varphi^n_{\bar{a}}(\bar{b})$ if and only if Player II has a winning strategy in $\mathrm{EF}_n(\mathfrak{M}, \bar{a}, \mathfrak{N}, \bar{b})$.*

*Proof.* We prove the three assertions by induction on $n$, for all tuple lengths at once.

(1) The set $\Lambda_k$ is finite because the relational vocabulary and the variable list are finite. Hence $H_{0,k}$ is finite and contains $\varphi^0_{\bar{a}}$; the formulas of $H_{0,k}$ have rank 0 and free variables among $v_1, \dots, v_k$. For the induction step, assume that $H_{n,k+1}$ is finite and contains every formula $\varphi^n_{\bar{a}a}$. The set

$$S = \{\varphi^n_{\bar{a}a} : a \in M\}$$

is therefore finite, even if $M$ is infinite. The definition of $H_{n+1,k}$ includes the formula built from this $S$, namely $\varphi^{n+1}_{\bar{a}}$. There are only finitely many subsets of $H_{n,k+1}$, so $H_{n+1,k}$ is finite. Each resulting formula has rank at most $n+1$ and free variables among $v_1, \dots, v_k$.

(2) The rank-0 formula holds by definition. For the induction step, each extended tuple $\bar{a}a$, with $a \in M$, satisfies the formula $\varphi^n_{\bar{a}a}$ in $\mathfrak{M}$ by induction. Thus each existential requirement has a witness, and every element satisfies one of the listed extension formulas. Both conjuncts of $\varphi^{n+1}_{\bar{a}}$ therefore hold.

(3) At rank 0, the formula records every atomic formula true of $\bar{a}$ and the negation of every one false of $\bar{a}$. The formula therefore holds of $\bar{b}$ exactly when $\bar{a} \mapsto \bar{b}$ is a partial isomorphism, which is the winning condition for zero rounds. At the next rank, the definition says that $\mathfrak{N} \models \varphi^{n+1}_{\bar{a}}(\bar{b})$ exactly when both conditions hold:

(i) for every $a \in M$ there is $b \in N$ such that $\mathfrak{N} \models \varphi^n_{\bar{a}a}(\bar{b}b)$;
(ii) for every $b \in N$ there is $a \in M$ such that $\mathfrak{N} \models \varphi^n_{\bar{a}a}(\bar{b}b)$.

By induction, the displayed satisfaction condition means that Player II can win $n$ rounds from the extended tuples $\bar{a}a$ and $\bar{b}b$. Thus (i) and (ii) say that every opening move has a response leaving her a winning strategy for $n$ more rounds. By Lemma 20.5, conditions (i) and (ii) together are equivalent to winning $n+1$ rounds. ☺

*Proof of Theorem 20.3.* (1) Suppose first that $(\mathfrak{M}, \bar{a}) \equiv_n (\mathfrak{N}, \bar{b})$. By Lemma 20.7(1),(2), the formula $\varphi^n_{\bar{a}}(v_1, \dots, v_k)$ has rank at most $n$ and holds of $\bar{a}$ in $\mathfrak{M}$; hence the formula holds of $\bar{b}$ in $\mathfrak{N}$. Part (3) of the same lemma gives a winning strategy for Player II in $\mathrm{EF}_n(\mathfrak{M}, \bar{a}, \mathfrak{N}, \bar{b})$.

For the converse, we show that a winning strategy preserves the truth of every formula whose quantifier rank is no larger than the number of rounds. We argue by induction on the formula $\varphi$. Let $\bar{x}$ list distinct variables containing the free variables of $\varphi$, and let $\bar{a}$ and $\bar{b}$ be tuples of the same length as this list. If Player II can win $j \geq \mathrm{qr}(\varphi)$ rounds from these tuples, we must show that

$$\mathfrak{M} \models \varphi(\bar{a}) \quad \text{if and only if} \quad \mathfrak{N} \models \varphi(\bar{b}).$$

Atomic formulas are preserved because the initial matching is a partial isomorphism (see Lemma 20.4(1)). The cases $\neg\varphi$ and $\varphi \wedge \psi$ follow from the induction hypothesis with the same $\bar{x}$, $j$ and tuples.

Suppose now that $\varphi = \exists z\, \psi$. The idea is to play a witness in one structure and use Player II's response as a witness in the other. If $z$ occurs in the variable list, first remove $z$ from the list, together with the corresponding matched pair of elements, since the quantifier supplies a new value for $z$. Write $\bar{x}'$, $\bar{a}'$ and $\bar{b}'$ for the resulting lists, whether or not an entry

was removed. Any such deletion leaves Player II's winning strategy intact by Lemma 20.4(2). There is at least one round available, since $j \geq \mathrm{qr}(\psi) + 1$. Suppose that $\mathfrak{M} \models \varphi(\bar{a})$, and choose $a \in M$ with $\mathfrak{M} \models \psi(\bar{a}', a)$. Play this witness as the opening move. Her winning strategy supplies a response $b \in N$ and wins the remaining $j - 1$ rounds from the extended tuples (see Lemma 20.5). The induction hypothesis for $\psi$, with the list $\bar{x}'z$ and $j - 1 \geq \mathrm{qr}(\psi)$, gives $\mathfrak{N} \models \psi(\bar{b}', b)$, hence $\mathfrak{N} \models \varphi(\bar{b})$. The implication from $\mathfrak{N}$ to $\mathfrak{M}$ is the same, using the back condition instead. The claim follows, and the claim for formulas of rank at most $n$ is (1). No step in this direction used the finiteness of the vocabulary.

Part (2) is the case $k = 0$ of (1), and part (3) follows from (2), since $\mathfrak{M} \equiv \mathfrak{N}$ if and only if $\mathfrak{M} \equiv_n \mathfrak{N}$ for all $n$. ☺

The Hintikka formula thus determines the *rank-n type* of a tuple: the set of formulas of rank at most $n$ that the tuple satisfies. Since only finitely many Hintikka formulas are needed for tuples of a fixed length, we also obtain a finiteness result for arbitrary formulas of bounded rank.

**Corollary 20.8.** *Let $V$ be a finite relational vocabulary, and let $n$ and $k$ be non-negative integers.*

(1) *For $V$-structures $\mathfrak{M}$, $\mathfrak{N}$ and tuples $\bar{a} \in M^k$, $\bar{b} \in N^k$, we have $(\mathfrak{M}, \bar{a}) \equiv_n (\mathfrak{N}, \bar{b})$ if and only if $\mathfrak{N} \models \varphi^n_{\bar{a}}(\bar{b})$.*

(2) *Every formula $\varphi(v_1, \ldots, v_k)$ of quantifier rank at most $n$ is logically equivalent to $\bigvee S$ for some $S \subseteq H_{n,k}$. Hence, up to logical equivalence, there are at most $2^{|H_{n,k}|}$ formulas of rank at most $n$ with free variables among $v_1, \ldots, v_k$.*

*Proof.* Part (1) combines Theorem 20.3(1) with Lemma 20.7(3). For (2), let $S$ be the set of Hintikka formulas $\varphi^n_{\bar{a}}$, where $\mathfrak{M}$ ranges over all $V$-structures and $\bar{a}$ over the tuples in $M^k$ with $\mathfrak{M} \models \varphi(\bar{a})$; the set $S$ is a subset of $H_{n,k}$. If $\mathfrak{M} \models \varphi(\bar{a})$, then the member $\varphi^n_{\bar{a}}$ of $S$ holds of $\bar{a}$ (see Lemma 20.7(2)), and therefore $\bar{a}$ satisfies $\bigvee S$. Conversely, if $\bar{b} \in N^k$ satisfies $\bigvee S$, there are $\mathfrak{M}$ and $\bar{a}$ with $\mathfrak{M} \models \varphi(\bar{a})$ and $\mathfrak{N} \models \varphi^n_{\bar{a}}(\bar{b})$. Part (1) then gives $(\mathfrak{M}, \bar{a}) \equiv_n (\mathfrak{N}, \bar{b})$, whence $\mathfrak{N} \models \varphi(\bar{b})$. ☺

**Remark 20.9.** Finiteness of the vocabulary is used only in the implication from $\equiv_n$ to a winning strategy, through the finiteness of $\Lambda_k$ in Lemma 20.7(1); the converse holds in every relational vocabulary. The implication from $\equiv_n$ to a winning strategy fails for infinite vocabularies. Indeed, two structures may be elementarily equivalent while Player I wins in a single round (see Exercise 8 at the end of Section 21).

Function and constant symbols can be absorbed into relational vocabularies by replacing each function symbol $f$ of arity $r$ with a relation symbol $R_f$ of arity $r + 1$, interpreted as the graph $\{(\bar{a}, a) : f(\bar{a}) = a\}$. Each constant symbol is replaced by a unary relation symbol interpreted as the singleton containing the value of that constant. Formulas translate back and forth, so two structures are elementarily equivalent exactly when their relational versions are, and Theorem 20.3(3) characterizes elementary equivalence in every finite vocabulary. The quantifier rank, however, is not preserved, since nested terms must be unnested by quantifiers: the formula $f(f(x)) = y$ becomes $\exists z\,(R_f(x, z) \wedge R_f(z, y))$.

## 21. Linear orders, parity and connectivity

The Ehrenfeucht–Fraïssé theorem (see Theorem 20.3) turns questions about sentences into questions about strategies. Showing that two structures satisfy the same sentences of quantifier rank at most $n$ becomes the task of describing a winning strategy for Player II in the game with $n$ rounds. In this section we describe two such strategies, both on linear orders.

On dense linear orders without endpoints she survives any number of rounds; on finite linear orders she survives $n$ rounds as soon as both orders have at least $2^n - 1$ elements. The second strategy has two consequences for finite structures: no first-order sentence says that a finite linear order has an even number of elements, and none says that a finite graph is connected.

**Dense linear orders.** Let $\bar{a}$ and $\bar{b}$ be tuples of the same length in two linear orders. The map $\bar{a} \mapsto \bar{b}$ is a partial isomorphism exactly when $a_i < a_l \iff b_i < b_l$ and $a_i = a_l \iff b_i = b_l$ for all $i, l$, that is, when the two tuples have the same order type. Player II's strategy is the one of the back-and-forth proof of Cantor's theorem: answer every move with an element in the same position relative to the elements already chosen.

**Proposition 21.1.** *Let $\mathfrak{M}$ and $\mathfrak{N}$ be dense linear orders without endpoints, let $k \geq 0$, and let $\bar{a} \in M^k$ and $\bar{b} \in N^k$ be such that $\bar{a} \mapsto \bar{b}$ is a partial isomorphism. Then Player II has a winning strategy in $\mathrm{EF}_n(\mathfrak{M}, \bar{a}, \mathfrak{N}, \bar{b})$ for every $n \geq 0$. In particular, we have $\mathfrak{M} \equiv \mathfrak{N}$; an example is $(\mathbb{Q}, <) \equiv (\mathbb{R}, <)$.*

*Proof.* We argue by induction on $n$, for all tuples at once; for $n = 0$ the hypothesis is the winning condition. When $n \geq 1$, by Lemma 20.5 and the induction hypothesis we need only show that for every $a \in M$ there is $b \in N$ with $\bar{a}a \mapsto \bar{b}b$ a partial isomorphism, and symmetrically. By symmetry we treat only the first condition. If $a = a_i$ for some $i$, take $b = b_i$; otherwise let $P^- = \{i : a_i < a\}$ and $P^+ = \{i : a_i > a\}$, which partition $\{1, \dots, k\}$. For $i \in P^-$ and $l \in P^+$ we have $a_i < a_l$, hence $b_i < b_l$. Consequently, there is $b \in N$ with $b_i < b < b_l$ for all $i \in P^-$ and $l \in P^+$. Density provides such an element if both sets are non-empty, and the absence of endpoints if exactly one is. If both are empty, which happens only when $k = 0$, the existence of $b$ is trivial. Then $a_i < a \iff b_i < b$ and $a_i > a \iff b_i > b$ for every $i$, and neither $a_i = a$ nor $b_i = b$, so $\bar{a}a \mapsto \bar{b}b$ is a partial isomorphism.

Applied to the empty tuples, the argument shows that Player II wins $\mathrm{EF}_n(\mathfrak{M}, \mathfrak{N})$ for every $n$, and Theorem 20.3(3) gives $\mathfrak{M} \equiv \mathfrak{N}$. ☺

The example given in the problem at the beginning of Section 19 is thereby settled.

**Finite linear orders.** For an integer $m \geq 1$, we write $L_m$ for the linear order $(\{1, \dots, m\}, <)$. In $\mathrm{EF}_n(L_m, L_{m'})$ Player II must keep track of distances. She adds an imaginary element 0 below each order, and imaginary elements $m + 1$ and $m' + 1$ above $L_m$ and $L_{m'}$, and she watches the gaps between consecutive chosen elements. With $j$ rounds remaining, Player II keeps corresponding gaps equal in length or both at least $2^j$ long. The splitting lemma below shows how she can preserve this condition after each move.

**Definition 21.2.** Fix an integer $j \geq 0$; for integers $d$ and $d'$, we write $d \approx_j d'$ if $d = d'$ or $d, d' \geq 2^j$. Let $m, m' \geq 1$ and $k \geq 0$ be integers, and let $\bar{a} \in \{1, \dots, m\}^k$ and $\bar{b} \in \{1, \dots, m'\}^k$ be given. Write $a_- = 0$, $a_+ = m + 1$, $b_- = 0$, $b_+ = m' + 1$, and $K = \{-, 1, \dots, k, +\}$. Indices $p, q \in K$ are *consecutive* for $\bar{a}$ if $a_p < a_q$ and no $s \in K$ satisfies $a_p < a_s < a_q$. The tuples $\bar{a}$ and $\bar{b}$ are *$j$-similar* if

(1) for all $p, q \in K$ the equivalences $a_p < a_q \iff b_p < b_q$ and $a_p = a_q \iff b_p = b_q$ hold, and
(2) for all $p, q \in K$ consecutive for $\bar{a}$, the gaps satisfy $a_q - a_p \approx_j b_q - b_p$.

By condition (1), two indices $p$ and $q$ are consecutive for $\bar{a}$ exactly when they are consecutive for $\bar{b}$; hence the definition is symmetric. For $j \geq 1$, the relation $d \approx_j d'$ implies $d \approx_{j-1} d'$. The heart of the strategy is arithmetic: when Player I cuts a gap into two pieces, Player II can cut the corresponding gap so that corresponding pieces are similar one level down.

**Lemma 21.3.** *Let $j \geq 1$, and let $d$, $d'$, $e$ be positive integers with $d \approx_j d'$ and $e < d$. Then there is an integer $e'$ with $0 < e' < d'$ such that*

$$e \approx_{j-1} e' \qquad \text{and} \qquad d - e \approx_{j-1} d' - e'.$$

*Proof.* Copy a short piece, measured from the corresponding end, and cut at distance $2^{j-1}$ from the left end when both pieces are long. If $d = d'$, take $e' = e$; otherwise $d, d' \geq 2^j$, and there are three cases.

If $e < 2^{j-1}$, take $e' = e$. Then $0 < e' < 2^{j-1} < d'$, and $d - e > 2^{j-1}$ and $d' - e' > 2^{j-1}$.

If $e \geq 2^{j-1}$ and $d - e < 2^{j-1}$, take $e' = d' - (d - e)$. Then $d' - e' = d - e > 0$, so $e' < d'$, and $e' > d' - 2^{j-1} \geq 2^{j-1}$; thus $e, e' \geq 2^{j-1}$.

If $e \geq 2^{j-1}$ and $d - e \geq 2^{j-1}$, take $e' = 2^{j-1}$. Then $0 < e' < 2^j \leq d'$, $e, e' \geq 2^{j-1}$, and $d - e \geq 2^{j-1}$, $d' - e' \geq 2^j - 2^{j-1} = 2^{j-1}$. ☺

**Proposition 21.4.** *Let $m$ and $m'$ be positive integers.*

(1) *If $j, k \geq 0$ and $\bar{a} \in \{1, \ldots, m\}^k$ and $\bar{b} \in \{1, \ldots, m'\}^k$ are $j$-similar, then Player II has a winning strategy in $\mathrm{EF}_j(L_m, \bar{a}, L_{m'}, \bar{b})$.*

(2) *If $n \geq 0$ and $m, m' \geq 2^n - 1$, then $L_m \equiv_n L_{m'}$.*

*Proof.* (1) The proof is by induction on $j$, for all tuples at once. For $j = 0$, condition (1) of Definition 21.2, restricted to the indices $1, \ldots, k$, says that $\bar{a} \mapsto \bar{b}$ is a partial isomorphism.

For $j \geq 1$, by Lemma 20.5 and the induction hypothesis Player II need only answer every move of Player I so that the extended tuples are $(j-1)$-similar. By symmetry, assume that Player I plays $a \in \{1, \ldots, m\}$.

If $a = a_i$ for some $i \in \{1, \ldots, k\}$, Player II answers $b = b_i$. The new index then behaves exactly like $i$: the new index has the same comparisons, and the consecutive pairs containing the new index have the same gaps as those containing $i$. Hence the extended tuples are $(j-1)$-similar.

Otherwise $a$ differs from every $a_p$, $p \in K$. Let $p, q \in K$ be indices such that $a_p$ is the largest value below $a$ and $a_q$ the smallest value above $a$; then $p$ and $q$ are consecutive for $\bar{a}$. Write $d = a_q - a_p$, $d' = b_q - b_p$ and $e = a - a_p$; these numbers satisfy $0 < e < d$ and $d \approx_j d'$. Let $e'$ be given by Lemma 21.3, and let Player II answer $b = b_p + e'$; since $0 \leq b_p < b < b_q \leq m' + 1$, the answer $b$ is an element of $\{1, \ldots, m'\}$.

Every $s \in K$ satisfies $a_s \leq a_p$ or $a_s \geq a_q$, whence condition (1) gives $b_s \leq b_p$ or $b_s \geq b_q$. Therefore

$$a_s < a \quad \text{if and only if} \quad a_s \leq a_p \quad \text{if and only if} \quad b_s \leq b_p \quad \text{if and only if} \quad b_s < b,$$

likewise $a_s > a \iff b_s > b$, and neither $a_s = a$ nor $b_s = b$; condition (1) therefore holds for the extended tuples. Except for the pairs with values $a_p$ and $a_q$, the consecutive pairs of the old tuples remain consecutive in the extended tuples. The only other consecutive pairs of the extended tuples are the new pairs joining the new index to the indices with values $a_p$ and $a_q$. The old pairs satisfy condition (2) for $j - 1$ because they satisfy the condition for $j$, and the new pairs have gaps $e, d - e$ and $e', d' - e'$, which satisfy the condition by the choice of $e'$.

(2) The empty tuples have the single consecutive pair $(-, +)$, with gaps $m+1, m'+1 \geq 2^n$, so they are $n$-similar. By (1) Player II wins $\mathrm{EF}_n(L_m, L_{m'})$, and Theorem 20.3(2) gives $L_m \equiv_n L_{m'}$. ☺

In the play of $\mathrm{EF}_2(L_6, L_5)$ shown in Figure 30, Player II follows this strategy. The bound $2^n - 1$ cannot be lowered: if $m \neq m'$ and $\min\{m, m'\} < 2^n - 1$, then Player I wins $\mathrm{EF}_n(L_m, L_{m'})$ (see Exercises 2 and 3).

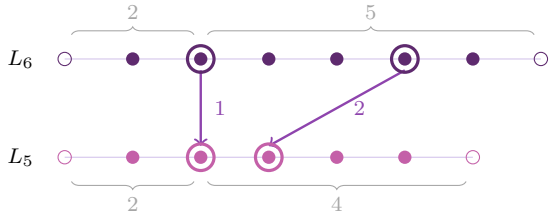



FIGURE 30. A play of $\mathrm{EF}_2(L_6, L_5)$ in which Player II follows the strategy of Proposition 21.4, with the imaginary elements drawn as open circles.

**Parity and connectivity.** Every first-order sentence has some quantifier rank $n$, and by Proposition 21.4(2) the sentence takes the same truth value in all finite linear orders with at least $2^n - 1$ elements. A property that changes from each length to the next, as parity does, therefore cannot be expressed by a sentence.

**Corollary 21.5.** *No first-order sentence $\theta$ in the vocabulary $\{<\}$ holds in exactly the finite linear orders with an even number of elements.*

*Proof.* Suppose $\theta$ were such a sentence, of quantifier rank $n$. Since $2^n, 2^n + 1 \geq 2^n - 1$, we obtain $L_{2^n} \equiv_n L_{2^n+1}$ from Proposition 21.4(2), so the two orders agree on $\theta$, although exactly one of them has an even number of elements. ☺

For connectivity we define, inside a linear order, a graph that is connected exactly when the order has odd size, and we transfer sentences about the graph to sentences about the order. Let $\delta(x, y)$ be a formula in the vocabulary $\{<\}$ with free variables among the distinct variables $x$ and $y$. For a $\{<\}$-structure $\mathfrak{M}$, let $\mathfrak{M}^\delta$ be the $\{E\}$-structure with universe $M$ and

$$E^{\mathfrak{M}^\delta} = \big\{(a, a') \in M^2 : \mathfrak{M} \models \delta(a, a')\big\}.$$

**Lemma 21.6.** *Let $\delta(x, y)$ be a formula in the vocabulary $\{<\}$. For every sentence $\theta$ in the vocabulary $\{E\}$ there is a sentence $\theta^*$ in the vocabulary $\{<\}$ such that, for every $\{<\}$-structure $\mathfrak{M}$, we have $\mathfrak{M} \models \theta^*$ if and only if $\mathfrak{M}^\delta \models \theta$.*

*Proof.* First rename the bound variables of $\delta$ so that they do not occur in $\theta$. For variables $u, w$ of $\theta$, let $\delta(u, w)$ be the result of substituting $u$ for $x$ and $w$ for $y$ simultaneously. By the substitution lemma, the formula $\delta(u, w)$ holds in $\mathfrak{M}$ under the assignment $u \mapsto a$, $w \mapsto a'$ if and only if $\mathfrak{M} \models \delta(a, a')$. For each subformula $\varphi$ of $\theta$, let $\varphi^*$ be obtained by replacing every atomic subformula $E(u, w)$ with $\delta(u, w)$; the free variables of $\varphi^*$ lie among those of $\varphi$. We show by induction on $\varphi$ that, for every list $\bar{x}$ of distinct variables containing the free variables of $\varphi$ and every tuple $\bar{a}$ in $M$ of the same length, we have

$$\mathfrak{M} \models \varphi^*(\bar{a}) \quad \text{if and only if} \quad \mathfrak{M}^\delta \models \varphi(\bar{a}).$$

For $\varphi = E(u, w)$ the claim is the substitution lemma together with the definition of $E^{\mathfrak{M}^\delta}$; the constant $\top$ and the equations are unchanged, and their truth does not depend on the relations. The steps for $\neg$ and $\wedge$ are immediate. For $\varphi = \exists z\, \psi$ we have $\varphi^* = \exists z\, \psi^*$, and since $\mathfrak{M}$ and $\mathfrak{M}^\delta$ have the same universe, Tarski's clause reduces the claim to the induction hypothesis for $\psi$ at the tuples $\bar{a}'a$, $a \in M$. The case $\varphi = \theta$, with the empty list, is the lemma. ☺

The translation may raise the quantifier rank, but only the rank of $\theta^*$ matters below. For $m \geq 3$, let $G_m$ be the graph with vertex set $\{1, \ldots, m\}$ in which $i$ is joined to $i+2$ for $1 \leq i \leq m-2$. In addition, the vertex $m$ is joined to 2, and the vertex $m-1$ to 1 (see Figure 31).

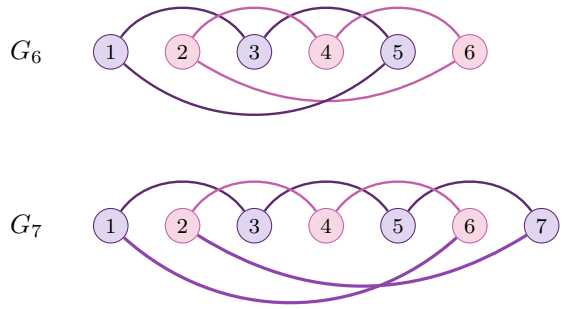



FIGURE 31. The graphs $G_6$ and $G_7$ of Lemma 21.7, with the edges from $i$ to $i+2$ above the vertices and the extra edges below.

**Lemma 21.7.** (1) *For $m \geq 3$, the graph $G_m$ is connected if and only if $m$ is odd.*
(2) *There is a formula $\delta(x, y)$ in the vocabulary $\{<\}$ such that $L_m^\delta = G_m$ for every $m \geq 3$.*

*Proof.* (1) The edges joining $i$ to $i+2$ form a path through the odd vertices $1, 3, 5, \ldots$ and a path through the even vertices $2, 4, 6, \ldots$, both non-empty since $m \geq 3$. If $m$ is odd, the edge joining the odd vertex $m$ to the even vertex 2 connects the two paths. If $m$ is even, every edge joins vertices of the same parity (the extra edges join $m$ to 2 and $m-1$ to 1), so no path leads from 1 to 2.

(2) Successor, the first two elements and the last two elements of a finite linear order are first-order definable:

$$\begin{aligned} \mathrm{succ}(x,y) &= x < y \wedge \neg\exists z\,(x < z \wedge z < y), & \mathrm{plus}_2(x,y) &= \exists z\,\big(\mathrm{succ}(x,z) \wedge \mathrm{succ}(z,y)\big), \\ \mathrm{first}(x) &= \neg\exists z\,(z < x), & \mathrm{second}(x) &= \exists z\,\big(\mathrm{first}(z) \wedge \mathrm{succ}(z,x)\big), \\ \mathrm{last}(x) &= \neg\exists z\,(x < z), & \mathrm{penult}(x) &= \exists z\,\big(\mathrm{last}(z) \wedge \mathrm{succ}(x,z)\big), \end{aligned}$$

with nested bound variables renamed apart from each other and from $x$ and $y$. In $L_m$ the formulas $\mathrm{succ}(i, i')$ and $\mathrm{plus}_2(i, i')$ hold exactly when $i' = i+1$ and $i' = i+2$, respectively. The formulas first and second define the elements 1 and 2; the last two elements, $m-1$ and $m$, are defined by penult and last. Consequently, the formula

$$\varepsilon(x,y) = \mathrm{plus}_2(x,y) \vee \big(\mathrm{last}(x) \wedge \mathrm{second}(y)\big) \vee \big(\mathrm{penult}(x) \wedge \mathrm{first}(y)\big)$$

holds of $(i, i')$ exactly when $i' = i+2$ or $(i, i')$ is one of the pairs $(m, 2)$ and $(m-1, 1)$, and the formula $\delta(x,y) = \varepsilon(x,y) \vee \varepsilon(y,x)$ defines the edges of $G_m$. ☺

**Corollary 21.8.** *No first-order sentence $\theta$ in the vocabulary $\{E\}$ holds in exactly the connected finite graphs.*

*Proof.* Suppose $\theta$ were such a sentence. Let $\delta$ be as in Lemma 21.7(2) and $\theta^*$ as in Lemma 21.6. For every $m \geq 3$ we have

$$L_m \models \theta^* \iff G_m \models \theta \iff G_m \text{ is connected} \iff m \text{ is odd}.$$

Let $n = \mathrm{qr}(\theta^*)$ and $m = 2^n + 2$; then $m \geq 3$ and $m, m+1 \geq 2^n - 1$. By Proposition 21.4(2) we obtain $L_m \equiv_n L_{m+1}$, and the two orders agree on $\theta^*$, although exactly one of $m$ and $m+1$ is odd. ☺

Neither corollary mentions a game; in both, the game reduced a statement about all sentences to a single strategy, the halving of gaps in Lemma 21.3.

**Exercises.**

(1) Show that $L_m \equiv_1 L_{m'}$ for all $m, m' \geq 1$, by describing a winning strategy for Player II in $\mathrm{EF}_1(L_m, L_{m'})$. Show that Player I wins $\mathrm{EF}_2(L_1, L_2)$, and find a sentence of quantifier rank 2 true in $L_2$ and false in $L_1$.

(2) Show that Player I wins $\mathrm{EF}_2(L_2, L_3)$ and that Player II wins $\mathrm{EF}_2(L_3, L_4)$, describing the strategies directly. Thus the orders $L_3$ and $L_4$ satisfy the same sentences of quantifier rank 2, while $L_2$ and $L_3$ do not, and for $n = 2$ the bound $2^n - 1$ of Proposition 21.4(2) cannot be lowered.

(3) (∗) Let $n \geq 1$ be an integer, and let $m \neq m'$ be positive integers with $\min\{m, m'\} < 2^n - 1$. Show that Player I wins $\mathrm{EF}_n(L_m, L_{m'})$, and conclude that $L_m \equiv_n L_{m'}$ holds if and only if $m = m'$ or $m, m' \geq 2^n - 1$. (Hint: show by induction on $j$ that Player I wins in $j$ rounds from any pair of tuples that violates condition (1) of Definition 21.2. Show, in the same induction, that he also wins in $j$ rounds from any pair of tuples having a consecutive pair of indices whose gaps have lengths $d \neq d'$ with $\min\{d, d'\} < 2^j$. If $d < d'$, he plays in the longer gap, at distance $\min\{2^{j-1}, \lfloor d'/2 \rfloor\}$ from the left end of the gap.)

(4) Show that $(\mathbb{Z}, <) \equiv_2 (\mathbb{Q}, <)$, but that Player I wins $\mathrm{EF}_3((\mathbb{Z}, <), (\mathbb{Q}, <))$. Find a sentence of quantifier rank 3 that distinguishes the two orders.

(5) Draw the graphs $G_5$ and $G_8$ of Lemma 21.7. Show that for odd $m \geq 3$ the graph $G_m$ is a single cycle through all the vertices. Show also that for even $m \geq 6$ the graph $G_m$ consists of two disjoint cycles, one through the odd and one through the even vertices. Compute the quantifier rank of the formula $\delta$ in the proof of Lemma 21.7(2).

(6) Show that there is no first-order sentence in the vocabulary $\{E\}$ that holds in a finite graph if and only if the graph is bipartite. (Hint: for $m \geq 3$, the cycle obtained from $L_m$ by joining each $i < m$ to $i + 1$ and $m$ to 1 is bipartite exactly when $m$ is even. Argue with these cycles as in the proof of Corollary 21.8.)

(7) (∗) Let $(\mathbb{Z} + \mathbb{Z}, <)$ denote the order obtained by placing one copy of $(\mathbb{Z}, <)$ after another. Show that $(\mathbb{Z}, <) \equiv (\mathbb{Z} + \mathbb{Z}, <)$. (Hint: adapt the invariant used for finite linear orders, regarding two elements in different copies as being infinitely far apart.)

(8) (∗) Let $V = \{P_0, P_1, P_2, \dots\}$ be an infinite vocabulary of unary relation symbols. Let $\mathfrak{M}$ be the $V$-structure with universe $\mathbb{N}$ in which $P_i$ is interpreted as $\{a \in \mathbb{N} : a > i\}$. Let $\mathfrak{N}$ be obtained from $\mathfrak{M}$ by adding a new element $\infty$ that lies in every $P_i$. Show that $\mathfrak{M} \equiv \mathfrak{N}$, but that Player I wins $\mathrm{EF}_1(\mathfrak{M}, \mathfrak{N})$. (Hint: for every integer $r \geq 0$, the reducts of $\mathfrak{M}$ and $\mathfrak{N}$ to the vocabulary $\{P_0, \dots, P_r\}$ are isomorphic.) Thus parts (2) and (3) of Theorem 20.3 fail for infinite vocabularies, as stated in Remark 20.9.

## 22. Back-and-forth systems and ultrapowers

Thus far the game has lasted a fixed finite number of rounds. In this section we look at the game from two further sides. Without the players, the game becomes Fraïssé's original formulation, a family of partial isomorphisms that can be extended back and forth. For countable dense linear orders the extension strategy works indefinitely, and arranging that

every element is eventually chosen produces an isomorphism. For suitable ultrapowers, saturation supplies the extensions through a transfinite sequence of moves. In Section 18 the same extensions became a construction, the Fraïssé limits.

**Back-and-forth systems.** We begin with the static form. Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for the same finite relational vocabulary. A *finite partial isomorphism* from $\mathfrak{M}$ to $\mathfrak{N}$ is a map $p$ from a finite subset of $M$ to $N$ such that $\bar{a} \mapsto p(\bar{a})$ is a partial isomorphism for an enumeration $\bar{a}$ of the domain of $p$. Consider a sequence $I_0, \dots, I_n$ of sets of finite partial isomorphisms, with $I_n \neq \varnothing$. Such a sequence is a *back-and-forth system* of length $n$ between $\mathfrak{M}$ and $\mathfrak{N}$ if, for $1 \leq j \leq n$ and $p \in I_j$, every $a \in M$ has some $b \in N$ with $p \cup \{(a,b)\} \in I_{j-1}$, and every $b \in N$ has some $a \in M$ with the same property. Strategies of Player II and back-and-forth systems carry the same information, one dynamically and the other statically. Suppose that Player II wins $\mathrm{EF}_n(\mathfrak{M}, \mathfrak{N})$, and let $I_j$ be the set of those $p$ for which she wins $\mathrm{EF}_j(\mathfrak{M}, \bar{a}, \mathfrak{N}, p(\bar{a}))$, with $\bar{a}$ enumerating the domain of $p$. These sets form a back-and-forth system: the enumeration does not matter by Lemma 20.4(2), and the conditions are those of Lemma 20.5. Conversely, given a back-and-forth system $I_0, \dots, I_n$, induction on $j$ shows that Player II wins $\mathrm{EF}_j(\mathfrak{M}, \bar{a}, \mathfrak{N}, p(\bar{a}))$ for every $p \in I_j$. For $j = 0$ the claim holds because $p$ is a partial isomorphism, and for $j \geq 1$ the claim follows from Lemma 20.5, since $\bar{a}a \mapsto p(\bar{a})b$ is a reindexing of an enumeration of $p \cup \{(a,b)\}$ and Lemma 20.4(2) applies. For any $p \in I_n$, Player II wins $\mathrm{EF}_n(\mathfrak{M}, \bar{a}, \mathfrak{N}, p(\bar{a}))$; after the deletion of all the initially matched pairs, she still has a winning strategy, now in $\mathrm{EF}_n(\mathfrak{M}, \mathfrak{N})$, by Lemma 20.4(2). Thus Theorem 20.3(2) is Fraïssé's theorem, which asserts that $\mathfrak{M} \equiv_n \mathfrak{N}$ if and only if there is a back-and-forth system of length $n$ between them. For dense linear orders without endpoints, the set of all finite partial isomorphisms serves as every $I_j$, as the proof of Proposition 21.1 shows. If both orders are countable, Player II can keep the matched pairs a partial isomorphism for $\omega$ rounds. In that case, if Player I enumerates both universes, alternating between them, the union of the matched pairs is an isomorphism. Cantor's theorem (see Proposition 18.1) appears here as the infinite form of the game.

**Ultrapowers.** Ultrafilters give a second, algebraic answer to the key problem: two structures are elementarily equivalent if and only if they have isomorphic ultrapowers. Keisler [Kei61] proved the characterization in 1961 under the generalized continuum hypothesis (GCH), and Shelah [She71] removed the hypothesis in 1971. One direction is Łoś's theorem: every structure embeds elementarily, by the diagonal map, in every ultrapower of the structure. For the other direction, elementarily equivalent finite structures are isomorphic. For infinite structures we give the proof under GCH in the form usually presented today, with Keisler's good ultrafilters. Let $\lambda$ be an infinite cardinal at least the cardinalities of the structures and of the vocabulary, and let $\mathcal{D}$ be a countably incomplete $\lambda^+$-good ultrafilter on $\lambda$. Such ultrafilters exist under GCH by Keisler's theorem of 1964 [Kei64], and in ZFC by Kunen's theorem of 1972 [Kun72]. Kunen builds them from an independent family of $2^\lambda$ functions from $\lambda$ to $\lambda$. For elementarily equivalent infinite $\mathfrak{M}$ and $\mathfrak{N}$, the ultrapowers $\mathfrak{M}^\lambda/\mathcal{D}$ and $\mathfrak{N}^\lambda/\mathcal{D}$ are $\lambda^+$-saturated and have cardinality $2^\lambda$, which is $\lambda^+$ under GCH. Now play the game for $\lambda^+$ rounds between the two ultrapowers; Player II keeps the map between the chosen elements elementary. She answers each move by realizing, in the other structure, the image under this map of the type of the new element over the elements already chosen. Since fewer than $\lambda^+$ elements have been chosen, saturation provides the required realization, and at limit stages the union of elementary maps is elementary. If Player I enumerates both structures, alternating between them, the completed play is an isomorphism. Shelah removed GCH by building the ultrafilter and the isomorphism together, in a single transfinite

construction [She71]. For structures of cardinality at most $\lambda$ his ultrafilter lies on $\mu = 2^\lambda$ rather than on $\lambda$, and the construction has length $2^\mu$. At each stage the filter is enlarged just enough to carry out one more back-and-forth step or to decide one more subset of $\mu$. A family of $2^\mu$ functions on $\mu$, independent modulo the filter, guarantees that the construction can continue. The method generalizes Kunen's proof that good ultrafilters exist in ZFC [Kun72].

## 23. Extensions to other logics

The proof of Theorem 20.3 matches one feature of the logic, the existential quantifier, with one feature of the game, a round. Change the logic, and the game changes with the logic: for most logics studied today there is a game of this kind, and the game characterizes equivalence in that logic as the game with $n$ rounds characterizes agreement on the sentences of quantifier rank at most $n$. In this section we describe the principal examples, without proofs; the details are in the surveys of Hodges and Väänänen [HV19] and of Abramsky and Reggio [AR24]. The books of Ebbinghaus and Flum [EF99], of Libkin [Lib04], of Immerman [Imm99], of Grohe [Gro17] and of Väänänen [Vää11] contain these details as well.

**Infinitary logic.** The infinitary logic $L_{\infty\omega}$ allows conjunctions and disjunctions of arbitrary sets of formulas. Karp [Kar65] showed that two structures are equivalent in this logic exactly when there is a non-empty set of finite partial isomorphisms with the back and forth properties of Lemma 20.5. Equivalently, such a set of partial isomorphisms exists exactly when Player II can survive the game for $\omega$ rounds. Games of transfinite length, and *dynamic* games in which Player I fixes the length as play proceeds, measure equivalence in stronger infinitary logics and are a central tool for uncountable structures [Vää11].

**Pebble games and counting.** Suppose that the players have only $k$ pairs of pebbles, which may be lifted and reused. In each round, Player I places one pebble of a pair on an element of either structure, and Player II places the other pebble of the pair on an element of the other structure. Player II must keep the map between the currently pebbled elements a partial isomorphism throughout the play. The resulting *pebble game* with $n$ rounds characterizes equivalence for formulas of quantifier rank at most $n$ with at most $k$ variables, each of which may be quantified many times. The game of unbounded length characterizes the logic $L^k$ of all formulas with at most $k$ variables on finite structures [Bar77, Imm82, Poi82]. In the *bijective pebble game* of Hella [Hel96], Player I first selects a pair of pebbles to lift or place. Player II then chooses a bijection between the two universes, and Player I chooses an element of the first structure; the selected pebbles are placed on that element and on the image of the element under the bijection. If no bijection exists, Player II loses. On finite structures, Player II can survive indefinitely in this variant exactly when the structures agree on the sentences of $C^k$, which adds the counting quantifiers $\exists^{\geq i} x$ to $L^k$. Cai, Fürer and Immerman proved that equivalence in $C^{k+1}$ coincides with indistinguishability by the $k$-dimensional Weisfeiler–Leman algorithm, a standard heuristic for graph isomorphism. With a counting version of the pebble game, the same authors constructed, for each $k$, non-isomorphic graphs that $C^k$ does not distinguish [CFI92]. The same hierarchy now bounds the power of the graph neural networks of machine learning: message-passing networks distinguish no more graphs than the 1-dimensional algorithm [MRF+19, XHLJ19]. The logical expressiveness of these networks is measured by fragments of $C^2$ [BKM+20, Gro21]. In the *existential* pebble game, Player I moves only in $\mathfrak{M}$. Player II must preserve equality and each relation that holds among the pebbled elements, but need not preserve the failure of a relation. The resulting matching is called a *partial homomorphism*. Kolaitis and Vardi introduced this game to study the expressive power of Datalog. They later showed that the winning strategies in the game

correspond to the consistency algorithms for constraint satisfaction problems [KV00]. Pebble games for logics with generalized quantifiers were developed by Kolaitis and Väänänen [KV95] and by Hella [Hel96].

**Monadic second-order logic and locality.** For monadic second-order logic, Player I may also choose a subset of one structure, and Player II answers with a subset of the other. The matched elements must preserve both the original relations and membership in corresponding chosen subsets. A variant in which the players first choose colorings and then play the first-order game is due to Fagin [Fag75]. With this variant he proved that connectivity of finite graphs is not definable in monadic existential second-order logic; Ajtai and Fagin refined the method to separate directed from undirected reachability [AF90]. Many non-definability proofs avoid explicit strategies by appeal to *locality*: the locality theorems of Hanf [Han65] and Gaifman [Gai82] describe first-order formulas using neighborhoods of bounded radius, together with information about how these neighborhoods occur throughout the structure. On finite structures these two theorems can be proved with games [FSV95, Lib04]. In modal logic the game starts from two pointed structures, each with a distinguished current world. Player I moves from either current world along the accessibility relation, and Player II must make a corresponding move in the other structure. Player II loses if no response exists or if the current worlds disagree on a propositional letter, initially or after any round. With finitely many propositional letters, the $n$-round game characterizes equivalence for modal formulas of modal depth at most $n$, and Player II survives forever exactly when the two pointed structures are bisimilar. Bisimulations appear, as p-relations, in van Benthem's thesis, in which he proves that modal logic is the bisimulation-invariant fragment of first-order logic [vB76].

**Current directions.** A categorical approach packages the games into constructions that can be studied together. The following paragraph assumes the language of category theory and may be skipped. Abramsky, Dawar and Wang showed that the $k$-pebble game is encoded by a comonad on the category of relational structures [ADW17]. Abramsky and Shah found comonads for the Ehrenfeucht–Fraïssé and modal games as well [AS18, AS21]. For finite structures, this *game comonad* approach identifies equivalence in the extension of each logic by counting quantifiers with isomorphism in the coKleisli category of the comonad. Equivalence in the logic itself becomes a form of bisimulation between coalgebras. The coalgebras also capture combinatorial parameters: a structure admits a coalgebra for the $k$-pebble comonad exactly when the treewidth of the structure is less than $k$. For the Ehrenfeucht–Fraïssé comonad of depth $k$, a structure admits a coalgebra exactly when the tree-depth of the structure is at most $k$. Theorems are thus proved for all the games at once; examples are the Lovász-type theorems on homomorphism counts of Dawar, Jakl and Reggio [DJR21]. The axiomatic theory of arboreal categories of Abramsky and Reggio [AR23] is another example (see the survey [AR24]). A second line of research counts quantifiers instead of nesting them. Immerman's separability game [Imm81], rediscovered as the multi-structural game, is played on two sets of structures and gives lower bounds on the number of quantifiers needed to express a property [FLRV21, CFI+24]. The Ehrenfeucht–Fraïssé game itself is computationally hard. For every fixed vocabulary containing at least one binary and one ternary relation symbol, Pezzoli proved that deciding which player wins $\mathrm{EF}_n(\mathfrak{M}, \mathfrak{N})$, for finite structures and $n$ given as input, is PSPACE-complete [Pez99]. For a fixed $n$ and a fixed finite vocabulary, by contrast, one need only check which sentences of the finite set $H_{n,0}$ of Definition 20.6 hold in $\mathfrak{M}$ and in $\mathfrak{N}$ (see Corollary 20.8). This check can be carried out in logarithmic space.

**Real-valued structures.** For the real-valued structures of Appendix A, which are Keisler's general structures [Kei23], a direct analogue of the game can be described for a finite relational vocabulary $V$. Fix a number of rounds $n \geq 0$ and a tolerance $\varepsilon > 0$. The players choose elements of two $V$-structures $\mathfrak{M}$ and $\mathfrak{N}$ as before, producing paired tuples $\bar{a}$ and $\bar{b}$. Player II wins if

$$|\alpha^{\mathfrak{M}}(\bar{a}) - \alpha^{\mathfrak{N}}(\bar{b})| \leq \varepsilon$$

for every atomic formula $\alpha$ in the selected variables. The vocabulary has no equality symbol, so the pairs need not define a bijection between the selected elements.

If Player II wins these games for every finite number of rounds and every positive tolerance, the two structures assign the same value to every sentence of $L^V$. The proof follows the induction used for Theorem 20.3. The connectives $s\cdot$, $+$ and $\vee$ are Lipschitz functions of the values of their arguments, so small errors in the arguments give small errors in the values; at a supremum $\sup_x$, Player I chooses an approximate witness and Player II supplies a response. For each fixed formula, this induction makes the error tend to zero with the atomic tolerance. The argument uses no metric and no continuity of the predicates on either universe. Function symbols require additional care: even a finite vocabulary can produce infinitely many atomic tests through nested terms, just as in the discussion after Theorem 20.3.

Metric versions have also been studied. Hart describes approximate Ehrenfeucht–Fraïssé games for continuous model theory [Har23]. Hirvonen and Puljujärvi establish the correspondence between finite game length and quantifier rank for finite relational metric vocabularies [HP24]. These metric results should be distinguished from the general real-valued setting just described.

## 24. Notes and historical remarks

Elementary equivalence was introduced by Tarski in Warsaw around 1930 (see [HV19]) and studied systematically by Tarski and Vaught [TV57]. The program of Tarski's school, to classify the complete first-order theories of familiar structures, needed methods for proving that two structures satisfy the same sentences. The method that Fraïssé would find is rooted in an argument of Cantor.

Cantor proved in 1895 that any two countable dense linear orders without endpoints are isomorphic [Can95]. His proof enumerates both orders but builds the isomorphism in one direction only. Each element of the first order is sent, in turn, to the first element of the second enumeration that fits. A separate argument shows that every element of the second order is reached; the symmetric back-and-forth argument appears first in Huntington [Hun04] and Hausdorff [Hau14] (see [Sil94]).

Fraïssé, who was then teaching in Algiers, characterized elementary equivalence by systems of partial isomorphisms in his thesis. The thesis, written under René de Possel, was defended in Paris in 1953 and published in Algiers in 1954 [Fra54b]. Ehrenfeucht, a student of Mostowski at the University of Warsaw, recast the characterization as a game, with an announcement in 1957 [Ehr57] and a full publication in 1961 [Ehr61]. Among Ehrenfeucht's applications was a game proof of the results of Mostowski and Tarski on the elementary equivalence of well-orders. The formulas $\varphi^n_{\bar{a}}$ of Definition 20.6 go back to Hintikka's distributive normal forms of 1953 [Hin53]. Karp's theorem of 1965 [Kar65] extended Fraïssé's characterization to infinitary logic. The names Spoiler and Duplicator are due to Joel Spencer, in the early 1990s (see [HV19]).

In the same year as his thesis, Fraïssé published the construction of limits, in a paper whose title announces the extension to relations of properties of orders [Fra54a]: Cantor's

theorem was the model, and amalgamation the new ingredient. The random graph has several discoverers. Erdős and Rényi showed in 1963 that a countable random graph is almost surely isomorphic to one fixed graph [ER63], and Rado gave the explicit construction by binary digits in 1964 [Rad64]. Henson's triangle-free graphs appeared in 1971 [Hen71], and Cameron's survey tells the story of the random graph [Cam97]. Several structures of the catalog are older than Fraïssé's theorem or were found independently of the theorem. Steinitz proved the uniqueness of algebraic closures in 1910 [Ste10], Urysohn constructed his universal metric space in a paper published in 1927 [Ury27], and Philip Hall constructed his universal locally finite group in 1959 [Hal59]. Jónsson extended Fraïssé's construction to uncountable structures in 1956 and 1960 [Jón56, Jón60], and the form for finitely generated structures that we use is presented by Hodges [Hod93].

In finite model theory, which grew from the 1970s onward in connection with databases and computational complexity, the games became one of the principal methods for proving that properties are not definable. An early example is Fagin's proof of 1975 that connectivity is not expressible in monadic existential second-order logic [Fag75].

The ultrapower characterization of elementary equivalence has a separate story. Keisler proved the characterization in 1961 under GCH [Kei61], and the hypothesis stood for ten years, until Shelah removed the hypothesis in a paper submitted in February 1971 [She71]. Kunen's construction of good ultrafilters in ZFC [Kun72] was then still to appear, and the two arguments were found at nearly the same time (see [She71, pp. 224–225]).

Part

# VI

# Banach–Mazur games in computational complexity

In Part II the rules of the Banach–Mazur game were changed in the sizes of the moves, in the winning condition and in the topology. In this part we change a rule of another kind: we limit how much Player II may compute. The result, due to Mehlhorn and Lutz, is a notion of Baire category inside complexity classes, a notion that separates classes ordinary category cannot tell apart. The story belongs with those of Part II. Since we assume no background in computer science, however, the definitions that the story requires take more room than theirs. For this reason we give the story a part of its own.

**Roadmap**

In Section 25 we state the problem and recall the history of the subject. In Section 26 we make computation a mathematical object: we define binary strings and Turing machines, the classes P, E and ESPACE, and Boolean circuits. In Section 27 we define the game, in which Player II must compute her moves quickly, and the notion of a meager class inside a complexity class. The main result, a theorem of Lutz, is stated in Section 28 and proved in Section 29. By this theorem, the classes E and ESPACE are not meager in themselves, and the class P is meager in E. The same theorem shows that the languages with small circuits form a class that is meager in ESPACE. The part follows the plan of the sections of Part II (the problem, the history, the game, the main result, and how the game solves the problem). From Part I the part needs only the Banach–Mazur game of Section 3 and Theorem 3.5. We prove every step, except a few facts about programming Turing machines, which we state precisely and for which we refer to the textbook of Arora and Barak [AB09].

## 25. Category inside complexity classes

**The problem.** A *language* is a set of finite binary strings. A *complexity class* is a class of languages singled out by a bound on the resources that a computer needs to decide membership in these languages. Three such classes concern us (see Section 26 below for precise definitions). The class P consists of the languages that a computer can decide in time polynomial in the length of the input. A language belongs to the class E if a computer can decide the language in time $2^{cn}$ on inputs of length $n$, for some constant $c$. The class ESPACE, finally, consists of the languages that a computer can decide with memory of size $2^{cn}$ on such inputs. We would like to say that "most" languages in E lie outside P, or that "most" languages in ESPACE cannot be computed by small Boolean circuits. Ordinary Baire category and the coin-tossing measure on the space of all languages are useless for this purpose. Since there are only countably many programs, every complexity class is countable, hence meager and of measure zero, whence the classes P and E look equally small. The problem is to find a notion of "small" that works *inside* a complexity class.

25.1. **History of the problem and of the solution using topological games.** The first step was taken by Kurt Mehlhorn in 1973 [Meh73]. He introduced resource-bounded Baire category for classes of computable functions, a category in which the descriptions of meager sets must themselves be computable with limited resources. Jack Lutz developed the theory inside complexity classes such as E and ESPACE, in his thesis of 1987 and in a paper published in 1990 [Lut90]. In the same two works he reformulated the theory in terms of the Banach–Mazur game played with resource-bounded strategies. Among other results, he proved that languages with high circuit complexity are typical in ESPACE in this sense. He introduced the companion notion of resource-bounded measure in the same paper, and developed that notion in a paper of 1992 [Lut92]. Both notions became standard tools for quantitative questions about complexity classes.

*Recognition.* Mehlhorn received the Leibniz Prize of the Deutsche Forschungsgemeinschaft in 1986.

## 26. Strings, machines and complexity classes

Before the cost of a computation can be measured, the computation itself has to become a mathematical object. Turing achieved this transformation in 1936 with a device whose every step is a finite and explicit operation. The device consists of a finite table of instructions acting on a strip of paper that is divided into squares and has no end. The machines below have several such strips. The price of this simplicity is that real programs become long; the reward is that the time and the memory of a computation are honest natural numbers. In this section we collect the definitions that the rest of the part needs, assuming nothing from computer science. They are of four kinds: binary strings and the standard order, Turing machines together with time and space, the three complexity classes of the problem, and Boolean circuits.

**Binary strings.** We use the notation of Section 2 for finite sequences, with $A = \{0, 1\}$. A *binary string* is an element of $\{0, 1\}^{<\omega}$, and we write $\lambda$ for the empty string, the empty sequence $\langle\rangle$ of Section 2. For binary strings $u$ and $v$, the *concatenation* $uv$ is the string $u$ followed by the string $v$. When $v$ is the string of length one whose only term is a bit $a$, the concatenation $uv$ is the extension $u^\frown a$ of Section 2. We write $u \subsetneq v$ when $v$ extends $u$ and differs from $u$. The *standard order* lists the binary strings by length and, among strings of

the same length, in dictionary order:

$$\lambda,\ 0,\ 1,\ 00,\ 01,\ 10,\ 11,\ 000,\ \ldots.$$

The *position* of a string is the number of strings that precede the string in the standard order. Thus the string $\lambda$ has position 0, the string 0 has position 1, and the string 00 has position 3. For $n \in \omega$, every string of length $n$ is the binary expansion, written with exactly $n$ digits (leading zeros included), of a unique integer between 0 and $2^n - 1$. We call this integer the *value* of the string; for instance, the string 011 has value 3. Writing a 1 in front of a string of length $n$ adds $2^n$ to the value of the string.

**Lemma 26.1.** *Let $n \in \omega$ be given.*

(1) *The strings of length $n$ occupy the positions $2^n - 1$ through $2^{n+1} - 2$, and the string of length $n$ with value $j$, where $0 \leq j < 2^n$, has position $2^n - 1 + j$. Equivalently, every binary string $x$ has position $m - 1$, where $m$ is the value of the string $1x$.*

(2) *For every binary string $w$, the positions $0, 1, \ldots, |w| - 1$ include those of all strings of length less than $n$ if and only if $|w| \geq 2^n - 1$.*

*Proof.* (1) For each $l < n$ there are $2^l$ strings of length $l$, all of which precede every string of length $n$. Exactly $2^0 + 2^1 + \cdots + 2^{n-1} = 2^n - 1$ strings therefore precede the strings of length $n$, which begin at position $2^n - 1$. Among strings of length $n$, dictionary order is the order of values, so the string with value $j$ is preceded by exactly $j$ strings of length $n$ and has position $2^n - 1 + j$. If $x$ has length $n$ and value $j$, the string $1x$ has value $2^n + j$, which exceeds the position of $x$ by one.

(2) By part (1), the strings of length less than $n$ are those with positions $0, 1, \ldots, 2^n - 2$, and all these positions are less than $|w|$ exactly when $|w| \geq 2^n - 1$. ☺

**Languages as points of the Cantor space.** A *language* is a set of binary strings. The *characteristic sequence* of a language $L$ is the infinite binary sequence $\chi_L \in 2^\omega$ defined as follows. For $m \in \omega$, the term of $\chi_L$ with index $m$ is 1 if the string at position $m$ belongs to $L$, and 0 otherwise. The map $L \mapsto \chi_L$ is a bijection from the set of all languages onto the Cantor space $2^\omega$ of infinite binary sequences, with the product topology. Through this bijection we regard every class of languages as a subset of $2^\omega$. For a binary string $u$, the set $N_u = \{y \in 2^\omega : u \subseteq y\}$ of the sequences that begin with $u$ is clopen, and the sets $N_u$ form a base for the topology of $2^\omega$. Each binary string $w$ describes the languages $L$ with $\chi_L \in N_w$, all of which agree about the strings at positions $0, 1, \ldots, |w| - 1$; we say that $w$ *settles* these strings. The strings that $w$ settles form an initial segment of the standard order. By Lemma 26.1(2), a string that settles all strings of length less than $n$, for $n \in \omega$, has length at least $2^n - 1$ (see Figure 32). This relation between the length of an input and the length of a string that settles the input is exponential. In consequence, time polynomial in the length of such a string will become time $2^{cn}$ on inputs of length $n$, for some constant $c$.

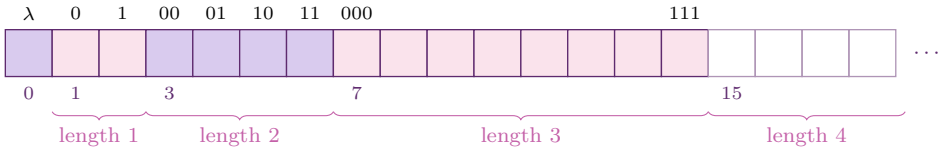



FIGURE 32. The characteristic sequence of a language, cut into blocks that begin at the indices 0, 1, 3, 7 and 15 written below the cells.

**Turing machines.** A Turing machine has finitely many internal states and a fixed number of tapes, each a row of cells indexed by $\omega$ and scanned by a head. At each step the machine reads the symbols under the heads. The current state, together with these symbols, determines what the machine writes, where the heads move and which state the machine enters next.

**Definition 26.2.** Let $k \geq 2$ be an integer. A *Turing machine with $k$ tapes* is a quintuple $M = (\Gamma, Q, q_{\text{start}}, q_{\text{halt}}, \delta)$ with the following properties.

(1) The *alphabet* $\Gamma$ is a finite set that contains four distinct symbols: 0, 1, the *start symbol* $\triangleright$ and the *blank* $\square$.
(2) The set $Q$ of *states* is finite, and $q_{\text{start}}$ and $q_{\text{halt}}$ are distinct elements of $Q$.
(3) The *transition function* $\delta$ is a function from $(Q \setminus \{q_{\text{halt}}\}) \times \Gamma^k$ to $Q \times \Gamma^{k-1} \times \{-1, 0, 1\}^k$.

The tapes are numbered $1, \ldots, k$; tape 1 is the *input tape*, and tape $k$ is the *output tape*. A *configuration* of $M$ consists of a state, a function from $\omega$ to $\Gamma$ for each tape (the contents of the tape), and a position in $\omega$ for each head. The *initial configuration on input* $x \in \{0,1\}^{<\omega}$ has the state $q_{\text{start}}$, all heads at cell 0, the symbol $\triangleright$ in cell 0 of every tape, the bits of $x$ in cells $1, \ldots, |x|$ of the input tape, and blanks everywhere else. Let a configuration with a state $q \neq q_{\text{halt}}$ be given, in which head $j$ is at cell $h_j$ and scans the symbol $\gamma_j$, for $1 \leq j \leq k$, and write $\delta(q, \gamma_1, \ldots, \gamma_k) = (q', \gamma_2', \ldots, \gamma_k', d_1, \ldots, d_k)$. The *next* configuration is obtained by writing $\gamma_j'$ in cell $h_j$ of tape $j$ for $2 \leq j \leq k$, moving head $j$ to cell $\max(h_j + d_j, 0)$ for $1 \leq j \leq k$, and entering the state $q'$. The *run* of $M$ on $x$ is the sequence of configurations that begins with the initial configuration on input $x$ and in which each configuration is followed by the next one, until a configuration with the state $q_{\text{halt}}$ is reached. If that happens, the machine $M$ *halts* on $x$, and the number of steps taken is the *time* of $M$ on $x$; the *output* of $M$ on $x$ is the longest binary string $y$ whose bits occupy cells $1, \ldots, |y|$ of the output tape in the final configuration, and the *space* of $M$ on $x$ is the number of cells of tapes $2, \ldots, k$ at which some head stands during the run.

The input tape is never written on, since $\delta$ prescribes symbols only for tapes 2 through $k$. Each head moves at most one cell per step, and a machine whose time on an input is $t$ therefore has space at most $(k-1)(t+1)$ on that input. For the same reason, the output of a machine that halts within $t$ steps has length at most $t$. Indeed, the output tape is blank beyond cell 0 at the start, and the head of the output tape needs $|y|$ steps to reach cell $|y|$, which holds the last bit of an output $y$.

**Definition 26.3.** Let $M$ be a Turing machine, let $L$ be a language, let $f\colon \{0,1\}^{<\omega} \to \{0,1\}^{<\omega}$ be a function, and let $t\colon \omega \to \omega$ be given.

(1) The machine $M$ *decides* $L$ if $M$ halts on every binary string $x$, with output 1 if $x \in L$ and output 0 if $x \notin L$.
(2) The machine $M$ *computes* $f$ if $M$ halts on every binary string $x$ with output $f(x)$.
(3) The machine $M$ runs *within time* $t$ if $M$ halts on every binary string $x$ with time at most $t(|x|)$, and *within space* $t$ if $M$ halts on every binary string $x$ with space at most $t(|x|)$.

A function $t\colon \omega \to \omega$ is *polynomially bounded* if there are integers $c, d \geq 1$ with $t(m) \leq cm^d + c$ for every $m \in \omega$.

**Definition 26.4.** Let $L$ be a language.

(1) The language $L$ belongs to the class P if, for some integer $d \geq 1$, some Turing machine decides $L$ within the time $m \mapsto m^d + d$.
(2) The language $L$ belongs to the class E if, for some integer $c \geq 1$, some Turing machine decides $L$ within the time $m \mapsto c\,2^{cm}$.

(3) The language $L$ belongs to the class ESPACE if, for some integer $c \geq 1$, some Turing machine decides $L$ within the space $m \mapsto c\,2^{cm}$.

Every polynomially bounded function is bounded by $m \mapsto m^d + d$ for a suitable integer $d \geq 1$ (if $t(m) \leq cm^k + c$ for integers $c, k \geq 1$ and every $m \in \omega$, the integer $d = k + 2c$ serves). Consequently, the class P consists of the languages that some machine decides within some polynomially bounded time. The inclusions $\mathrm{P} \subseteq \mathrm{E} \subseteq \mathrm{ESPACE}$ are the subject of Exercise 1 at the end of this part.

**Algorithms described in words.** Writing down transition functions is lengthy and unilluminating; in the proofs below we describe machines instead as algorithms in words, as is customary in the subject. Since such algorithms often receive several binary strings at once, we fix one way of writing several binary strings as one. For a positive integer $r$ and binary strings $u_1, \ldots, u_r$, we form the string $\langle u_1, \ldots, u_r \rangle$ by writing, for $j = 1, \ldots, r-1$ in turn, every bit of $u_j$ twice followed by the two bits 01. We then write the bits of $u_r$, each once; the result is a single binary string, not the sequence $(u_1, \ldots, u_r)$ of Section 2. Given $r$, the strings $u_1, \ldots, u_r$ can be read off from $\langle u_1, \ldots, u_r \rangle$: reading the bits of this string two at a time, one meets the pair 01 exactly at the end of each of $u_1, \ldots, u_{r-1}$. The bits that remain after the last of these pairs form $u_r$. A binary string can also describe a Turing machine, by listing the alphabet, the states and the table of the transition function of the machine. A single machine can read such a description and imitate the machine described. We use the following standard facts, for which we refer to the book of Arora and Barak (see [AB09, Chapters 1 and 4]); the statements about cells follow from the same constructions as those about steps. In part (3) of Fact 26.5, two machines are *isomorphic* if one is obtained from the other by renaming the states and the symbols of the alphabet other than 0, 1, ▷ and □. Isomorphic machines have the same time, space and output on every input.

**Fact 26.5.**

(1) Elementary operations. *There are a polynomially bounded function $\eta_0$ and an integer $c_0 \geq 1$ such that the following operations on strings of length at most $m$, for $m \in \omega$, take at most $\eta_0(m)$ steps and at most $c_0(m+1)$ cells: copying a string, comparing two strings, appending a bit, reading or writing the bit at a position given in binary, and adding, subtracting, multiplying or comparing natural numbers written in binary.*

(2) Subroutines. *Let $f\colon \{0,1\}^{<\omega} \to \{0,1\}^{<\omega}$ and $t, r\colon \omega \to \omega$ be functions, and let $M$ be a machine that computes $f$ within the time $t$ and within the space $r$. There are an integer $c_1 \geq 1$ and a polynomially bounded function $\eta_1$, both depending on $M$, such that an algorithm may compute $f(y)$ for every string $y$ that the algorithm has written, with at most $c_1 t(|y|) + \eta_1(|y| + |f(y)|)$ steps and at most $c_1 r(|y|) + \eta_1(|y| + |f(y)|)$ cells; the cells may be used again afterwards.*

(3) Descriptions. *There is an assignment $\alpha \mapsto M_\alpha$ of a Turing machine to every binary string $\alpha$ such that every Turing machine is isomorphic to $M_\alpha$ for some $\alpha$.*

(4) A universal machine. *There is a Turing machine $\mathfrak{U}$ with the following property. For every binary string $\alpha$ there is an integer $C_\alpha \geq 1$ such that, for every binary string $x$ and every $t \in \omega$, if $M_\alpha$ halts on $x$ within $t$ steps, then the machine $\mathfrak{U}$ halts on $\langle \alpha, x \rangle$ within $C_\alpha(t+1)^2$ steps, with the same output as $M_\alpha$ on $x$.*

(5) Clocks. *For every machine $M$ there is a polynomially bounded function $\eta_M$ such that an algorithm, given a binary string $y$ and a natural number $t$ written in binary, may run $M$ on $y$ for at most $t$ steps, and learn whether $M$ halted within these steps and with which output, with at most $\eta_M(|y| + t)$ steps and at most $\eta_M(|y| + t)$ cells.*

The machine of part (4) is the relaxed universal machine, with the bound $C_\alpha(t+1)^2$, that Arora and Barak construct on the way to their efficient universal machine. The bound of the efficient machine is $C_\alpha t \log t$ for $t \geq 2$ (see [AB09, Theorem 1.9]); the square is all we need. Arora and Barak write the input of the universal machine in the order $(x, \alpha)$. Since $\alpha$ comes first in $\langle \alpha, x \rangle$, the machine $\mathcal{U}$ can read $x$ in place, and the order does not matter.

**Boolean circuits.** A computer chip evaluates a Boolean function by wiring together gates, each of which combines one or two incoming bits. Unlike a Turing machine, a circuit has a fixed number of inputs, and a language therefore needs a separate circuit for each length.

**Definition 26.6.** Let $n$ and $s$ be positive integers, and let $u_1, \dots, u_n$ be distinct symbols, the *input variables*. A *circuit with $n$ inputs and $s$ gates* consists of gates $g_1, \dots, g_s$, one of which is designated as the *output gate*. Each gate $g_j$, for $1 \leq j \leq s$, has a *type*, which is AND, OR or NOT, and an ordered pair $(a_j, b_j)$ of *sources*, each of which is an input variable or a gate $g_l$ with $l < j$; the two sources of a gate of type NOT coincide. For $z = (z_1, \dots, z_n) \in \{0,1\}^n$, the input variable $u_m$ takes the value $z_m$ for $1 \leq m \leq n$, and the values of the gates are defined in the order $g_1, \dots, g_s$: a gate of type AND, OR or NOT takes, respectively, the minimum of the values of the sources of the gate, their maximum, or 1 minus the value of the source. The circuit *computes* the function $f\colon \{0,1\}^n \to \{0,1\}$ whose value at $z$ is the value of the output gate. The *truth table* of a function $f\colon \{0,1\}^n \to \{0,1\}$ is the binary string of length $2^n$ whose bit with index $j$, for $0 \leq j < 2^n$, is the value of $f$ at the string of length $n$ with value $j$.

For a language $L$ and a positive integer $n$, let $f_{L,n}\colon \{0,1\}^n \to \{0,1\}$ be the function that takes the value 1 exactly at the strings of length $n$ that belong to $L$. By Lemma 26.1(1), the truth table of $f_{L,n}$ is the block of the characteristic sequence $\chi_L$ formed by the terms with indices $2^n - 1$ through $2^{n+1} - 2$. We write $\mathrm{SIZE}(2^n/n)$ for the class of languages $L$ for which there is $n_0 \in \omega$ such that, for every integer $n \geq \max(n_0, 1)$, the function $f_{L,n}$ is computed by a circuit with $n$ inputs and at most $2^n/n$ gates.

## 27. The game

The game is the Banach–Mazur game of Section 3 on the Cantor space, played with the basic clopen sets $N_w$ and written in terms of the strings $w$ that name them. Player II aims at the complement of a given class of languages.

**Definition 27.1.** Let $\mathcal{C}$ be a class of languages. In the game $\mathrm{G}(\mathcal{C})$, the players alternately choose binary strings, each a proper extension of the previous one: Player I chooses $w_0$, Player II chooses $w_1$ with $w_0 \subsetneq w_1$, Player I chooses $w_2$ with $w_1 \subsetneq w_2$, and so on, forever. The language *produced* by the play is the language $L$ such that every string $w_k$, for $k \in \omega$, is an initial segment of $\chi_L$. Player II wins if $L \notin \mathcal{C}$.

A *positional strategy* for Player II is a function $\tau\colon \{0,1\}^{<\omega} \to \{0,1\}^{<\omega}$ such that $w \subsetneq \tau(w)$ for every binary string $w$. A play is *consistent* with $\tau$ if $w_{2j+1} = \tau(w_{2j})$ for every $j \in \omega$, and the strategy $\tau$ is *winning* in $\mathrm{G}(\mathcal{C})$ if every play consistent with $\tau$ produces a language outside $\mathcal{C}$.

Since $|w_k| \geq k$ for every $k \in \omega$, the strings of a play are initial segments of exactly one sequence in $2^\omega$, and the language produced by the play is therefore well defined. The strategies of Section 2 may consult all the moves made so far, whereas a positional strategy sees only the current string. We restrict Player II in this way because the resource bounds below measure what she computes from the current string. Without resource bounds the restriction costs her nothing: by the next proposition, she has a winning positional strategy

exactly when $\mathcal{C}$ is meager. By Theorem 3.5(1), the class $\mathcal{C}$ is meager exactly when Player II has a winning strategy in the Banach–Mazur game $\mathrm{BM}(2^\omega, 2^\omega \setminus \mathcal{C})$.

**Proposition 27.2.** *Let $\mathcal{C}$ be a class of languages, regarded as a subset of $2^\omega$. Player II has a winning positional strategy in $\mathrm{G}(\mathcal{C})$ if and only if $\mathcal{C}$ is meager in $2^\omega$.*

*Proof.* The idea is that of the proof of Theorem 3.5(1): a nowhere dense set can be avoided in a single move. As in that proof, a winning strategy yields countably many dense open sets whose common points are all produced by plays consistent with the strategy.

Suppose first that $\mathcal{C} = \bigcup_{m\in\omega} E_m$, where each set $E_m$ is nowhere dense, and let $F_m = \overline{E_0 \cup \cdots \cup E_m}$ for $m \in \omega$. A finite union of nowhere dense sets is nowhere dense, and each $F_m$ is therefore a closed set with empty interior. Let $w$ be a binary string. The open set $N_w \setminus F_{|w|}$ is non-empty and therefore contains a basic set $N_v$. Since $N_v \subseteq N_w$, the string $v$ extends $w$, and so does the string $v0$. We let $\tau(w)$ be the first string $u$ in the standard order such that $w \subsetneq u$ and $N_u \cap F_{|w|} = \emptyset$, which exists because $u = v0$ qualifies. Consider a play consistent with $\tau$, and let $L$ be the language that the play produces. For every $j \in \omega$, the sequence $\chi_L$ lies in $N_{w_{2j+1}}$, which misses $F_{|w_{2j}|}$, and therefore $\chi_L \notin E_m$ for every $m \leq |w_{2j}|$. Since $|w_{2j}| \geq 2j$, every $m \in \omega$ satisfies $m \leq |w_{2j}|$ for some $j$, whence $L \notin \mathcal{C}$. The strategy $\tau$ is thus winning.

Conversely, let $\tau$ be a winning positional strategy for Player II, and for $k \in \omega$ let $G_k$ be the union of the sets $N_{\tau(w)}$ over all binary strings $w$ with $|w| \geq k$. Each set $G_k$ is open, and each is also dense. Indeed, every non-empty open set contains a basic set $N_v$, and if $w$ is the string $v$ followed by $k$ zeros, then $|w| \geq k$ and $N_{\tau(w)} \subseteq N_w \subseteq N_v$. Let $L$ be a language with $\chi_L \in \bigcap_{k\in\omega} G_k$; we build a play consistent with $\tau$ that produces $L$. Since $\chi_L \in G_0$, there is a string $w_0$ with $\tau(w_0) \subseteq \chi_L$. Suppose that $w_0, \ldots, w_{2j}$ have been chosen, for some $j \in \omega$, with $\tau(w_{2j}) \subseteq \chi_L$, and let $w_{2j+1} = \tau(w_{2j})$. Since $\chi_L \in G_{|w_{2j+1}|+1}$, there is a string $w$ with $|w| > |w_{2j+1}|$ and $\tau(w) \subseteq \chi_L$. Both $w_{2j+1}$ and $w$ are initial segments of $\chi_L$ (the string $w$ because $w \subsetneq \tau(w) \subseteq \chi_L$), and the string $w$ is the longer of the two; hence $w_{2j+1} \subsetneq w$, and Player I may play $w_{2j+2} = w$. Every string of the resulting play is an initial segment of $\chi_L$, whence the play produces $L$; since the play is consistent with the winning strategy $\tau$, we have $L \notin \mathcal{C}$. Consequently, the class $\mathcal{C}$ is contained in $\bigcup_{k\in\omega}(2^\omega \setminus G_k)$. The complement of a dense open set is closed with empty interior, hence nowhere dense, and the class $\mathcal{C}$ is therefore meager. ☺

Every countable class of languages is meager in $2^\omega$, since a single point of $2^\omega$ forms a closed set with empty interior. By part (3) of Fact 26.5, there are only countably many Turing machines up to isomorphism. Since isomorphic machines decide the same languages, and every language in P, E or ESPACE is decided by some machine, these three classes are countable. By Proposition 27.2, Player II has winning positional strategies in G(P), in G(E) and in G(ESPACE) alike, and ordinary category cannot tell these classes apart.

The resource-bounded version of the game adds one rule: *Player II is not allowed to think for too long.* Each of her moves must be computed from the current string in time polynomial in the length of that string (or, when we work in ESPACE, in space polynomial in that length). Player I, by contrast, is under no restriction at all.

**Definition 27.3.** Let $\mathcal{C}$ be a class of languages, and let $\tau$ be a positional strategy for Player II. The strategy $\tau$ is *computable in polynomial time* if some Turing machine computes $\tau$ within some polynomially bounded time, and *computable in polynomial space* if some Turing machine computes $\tau$ within some polynomially bounded space. The class $\mathcal{C}$ is *meager in* E if Player II

has a winning positional strategy in $\mathrm{G}(\mathcal{C})$ that is computable in polynomial time, and *meager in* ESPACE if she has one that is computable in polynomial space.

Lutz calls a class $\mathcal{C}$ meager in E when Player II has such a strategy in the game for the class $\mathcal{C} \cap \mathrm{E}$; our condition is stronger, and Theorem 28.1 below holds for both. By Proposition 27.2, a class that is meager in E or in ESPACE is meager in $2^\omega$. The new notions are much finer: the class E is meager in $2^\omega$, but, as we show in Theorem 28.1 below, the same class is not meager in E.

## 28. The main result

A notion of smallness that declares everything small is useless. The theory rests on a resource-bounded version of the Baire category theorem, part (1) of Theorem 28.1 below, which we prove in the next section. Since the class E is not meager in E, no class meager in E contains E (a winning strategy in the game for a class of languages is also winning in the game for each subclass of that class). By part (2) of the same theorem, the class P is meager in E, and the notion therefore separates P from E, which ordinary category could not do.

**Theorem 28.1 (Lutz, 1987 and 1990).** *Let* P*,* E *and* ESPACE *be the classes of Definition 26.4, let* $\mathrm{SIZE}(2^n/n)$ *be the class of languages with small circuits defined after Definition 26.6, and let meagerness in* E *and in* ESPACE *be understood as in Definition 27.3.*

(1) Baire category theorem. *The class* E *is not meager in* E*, and the class* ESPACE *is not meager in* ESPACE*.*
(2) *The class* P *is meager in* E*.*
(3) *The class* $\mathrm{SIZE}(2^n/n)$ *is meager in* ESPACE*.*

By parts (1) and (3), high circuit complexity is typical in ESPACE. Indeed, the class ESPACE is not meager in itself, whereas the languages computed by circuits with at most $2^n/n$ gates at every large length $n$ form a class meager in ESPACE. Lutz [Lut92] also proved the measure counterpart in a stronger form: in resource-bounded measure, almost every language in ESPACE has, at every large length $n$, no circuit with at most $2^n/n$ gates. Category cannot give this strengthening, as the following strategy shows. Suppose that, at each move, Player II appends zeros until she has settled every string of the least length of which the current string settles no string. Every language $L$ produced then has, at infinitely many lengths $n$, the constant function 0 as $f_{L,n}$, which a circuit with two gates computes. This strategy is computable in polynomial time. Consequently, the languages with no circuit of at most $2^n/n$ gates at every large length $n$ form a class that is meager in E, hence in ESPACE. As on the real line (see Remark 1.2(2)), a class can thus be small in the sense of category and large in the sense of measure.

## 29. Solving the problem with the game

The game makes the definition of category effective in a natural way: it suffices to limit the resources of Player II. Each proof below is then the construction of an efficient strategy, or, for the Baire category theorem, of a play that an efficient algorithm can follow.

**The Baire category theorem: the classes** E **and** ESPACE **are not meager in themselves.** The idea is that Player I can play so predictably that the whole play becomes computable from Player II's strategy, with modest resources.

*Proof of Theorem 28.1(1).* Let $\tau$ be a positional strategy for Player II that is computable in polynomial time. To show that $\tau$ is not winning in G(E), we exhibit a play consistent with $\tau$ that produces a language in E. Fix a machine $M$ and integers $c, d \geq 1$ such that $M$ computes $\tau$ within the time $p$, where $p(m) = cm^d + c$ for $m \in \omega$; the function $p$ is increasing. Player I always appends a single 0: he plays $w_0 = 0$ and $w_{2j+2} = w_{2j+1}0$ for every $j \in \omega$, while Player II plays $w_{2j+1} = \tau(w_{2j})$. Let $L$ be the language produced by this play.

The following algorithm decides $L$: given a binary string $x$, compute the position $m_x$ of $x$ (see Lemma 26.1(1)) and set $w = 0$. While $|w| \leq m_x$, repeat the following round: replace $w$ by $\tau(w)$, and if $|w| \leq m_x$, append 0 to $w$. Stop as soon as $|w| > m_x$, and output the bit of $w$ with index $m_x$. After each round, the string $w$ is one of the strings $w_{2j+1}$ or $w_{2j+2}$ of the play, hence an initial segment of $\chi_L$. The output is therefore the term of $\chi_L$ with index $m_x$, which is 1 exactly when $x \in L$.

We bound the time of the algorithm. Each round lengthens $w$ by at least two, except possibly the last, and there are therefore at most $m_x/2 + 1$ rounds. Every string to which $\tau$ is applied has length at most $m_x$. On such a string, the computation of $\tau$ takes at most $p(m_x)$ steps and produces a string of length at most $p(m_x)$. By parts (1) and (2) of Fact 26.5, each round therefore takes at most $c_1 p(m_x) + \eta(p(m_x) + m_x)$ steps, for an integer $c_1 \geq 1$ and an increasing polynomially bounded function $\eta$ that do not depend on $x$. By the same two parts, computing $m_x$ from $x$ takes at most $\eta_0(|x| + 1)$ steps. Since $|x| \leq m_x \leq 2^{|x|+1}$ by Lemma 26.1(1), the total time on $x$ is thus at most $r(2^{|x|+1})$ for an increasing polynomially bounded function $r$ that does not depend on $x$. If $c', d' \geq 1$ are integers with $r(m) \leq c'm^{d'} + c'$ for every $m \in \omega$, then

$$r(2^{n+1}) \;\leq\; c'\, 2^{d'(n+1)} + c' \;\leq\; 2c'2^{d'} \cdot 2^{d'n} \qquad \text{for every } n \in \omega,$$

and the time on $x$ is therefore at most $c''2^{c''|x|}$ for the integer $c'' = 2c'2^{d'}$, independent of $x$ and at least $d'$. Hence $L \in$ E; as the play is consistent with $\tau$, the strategy $\tau$ is not winning in G(E). Since $\tau$ was arbitrary, the class E is not meager in E.

For ESPACE, let $\tau$ be computable in polynomial space, and let Player I play as before. Fix a machine $M$ and integers $c, d \geq 1$ such that $M$ computes $\tau$ within the space $p$, where $p(m) = cm^d + c$ for $m \in \omega$. The same algorithm decides the language $L$ produced and halts on every input, because $M$ halts on every input. Every string to which $\tau$ is applied has length at most $m_x$, and the image of the string under $\tau$ has length at most $p(m_x)$, since the output of $M$ occupies cells that a head of $M$ visits. The cells used in one round can be used again in the next. Hence, by parts (1) and (2) of Fact 26.5, the algorithm uses at most $c_1 p(m_x) + \eta(p(m_x) + m_x)$ cells, for an integer $c_1 \geq 1$ and an increasing polynomially bounded function $\eta$ that do not depend on $x$. As in the case of time, this number is at most $c''2^{c''|x|}$ for an integer $c'' \geq 1$ that does not depend on $x$, whence $L \in$ ESPACE. The class ESPACE is therefore not meager in ESPACE. ☺

**The class P is meager in E: diagonalization, one move at a time.** Player II must ensure that the language produced disagrees with every language in P on at least one string. Each of her moves disposes of a few more of the languages in P. To make the phrase "every language in P" concrete, we first list these languages, with a machine of known cost behind each entry of the list. We use the Cantor pairing function $\pi\colon \omega \times \omega \to \omega$, given by $\pi(a, b) = (a + b)(a + b + 1)/2 + b$; the function $\pi$ is a bijection and satisfies $a \leq \pi(a, b)$ and $b \leq \pi(a, b)$ for all $a, b \in \omega$. For $a \in \omega$, we write $\alpha_a$ for the binary string at position $a$, whose length is at most $a$ by Lemma 26.1(1). For $i \in \omega$, let $(a, b)$ be the pair with $\pi(a, b) = i$, and let $L_i$ be the language decided by the algorithm $D_i$ that works as follows. Given a binary

string $x$, the algorithm $D_i$ runs the universal machine $\mathcal{U}$ of Fact 26.5(4) on $\langle \alpha_a, x\rangle$ for at most $|x|^b + b$ steps. The output of $D_i$ on $x$ is 1 if $\mathcal{U}$ halted within these steps with output 1, and 0 otherwise.

**Lemma 29.1.**

(1) *Every language in* P *is equal to* $L_i$ *for some* $i \in \omega$.
(2) *There are an algorithm* $D$ *and an increasing polynomially bounded function* $\eta$ *with the following property: for all* $a, b \in \omega$ *and every binary string* $x$*, the algorithm* $D$*, given* $i = \pi(a, b)$ *written in binary together with* $x$*, produces the output of* $D_i$ *on* $x$ *within* $\eta(|x|^b + |x| + i + 1)$ *steps.*

*Proof.* (1) Let $L \in \mathrm{P}$ be given, and let $M$ be a machine and $d \geq 1$ an integer such that $M$ decides $L$ within the time $m \mapsto m^d + d$. By part (3) of Fact 26.5, and since every binary string is $\alpha_a$ for some $a \in \omega$, we may assume that $M = M_{\alpha_a}$ for some $a \in \omega$. By part (4), the machine $\mathcal{U}$ halts on $\langle \alpha_a, x\rangle$ with the output of $M$ within $C(|x|^d + d + 1)^2$ steps for every binary string $x$, where $C = C_{\alpha_a}$. For every $n \in \omega$ we have

$$C(n^d + d + 1)^2 \;\leq\; C(d+2)^2(n+1)^{2d} \;\leq\; K \max(n, 1)^{2d}, \qquad \text{where } K = 4^d C(d+2)^2.$$

Let $b = K^{2d+1}$, an integer at least $2d+1$. If $n \leq K$, the right-hand side is at most $K^{2d+1} = b$, and if $n > K$, that side is less than $n^{2d+1} \leq n^b$. In either case, the right-hand side is at most $n^b + b$. Within the budget of $|x|^b + b$ steps, the machine $\mathcal{U}$ therefore halts with the output of $M$ on $x$, and the algorithm $D_i$ with $i = \pi(a, b)$ decides $L$, whence $L = L_i$.

(2) Given $i$ in binary and $x$, the algorithm $D$ finds $a$ and $b$ by trying the pairs with $a+b \leq i$, and then writes $\alpha_a$, which has length at most $a \leq i$. The algorithm then forms the string $\langle \alpha_a, x\rangle$, of length at most $2i + |x| + 2$, and computes the number $|x|^b + b$ in binary by at most $b$ multiplications of numbers with at most $b(|x| + 1) + 1$ bits. By part (1) of Fact 26.5, these preparations take a number of steps bounded by a polynomially bounded function of $|x|+i+1$. By part (5), running $\mathcal{U}$ for at most $|x|^b + b$ steps takes at most $\eta_{\mathcal{U}}(2i + |x| + 2 + |x|^b + b)$ steps, where $b \leq i$. The sum of the two bounds is at most $\eta(|x|^b + |x| + i + 1)$ for a suitable increasing polynomially bounded function $\eta$. ☺

We now describe the strategy of Player II. From now on, we write log for the logarithm to base 2. Let $w$ be the current string, and let $n$ be the least length such that $w$ settles no string of length $n$. By Lemma 26.1(1), for every $m \in \omega$, the string $w$ settles no string of length $m$ exactly when $|w| \leq 2^m - 1$. The length $n$ is therefore the least integer with $2^n > |w|$. If $n \geq 1$, then the string $w$ settles some string of length $n-1$, and we have $2^{n-1} \leq |w| \leq 2^n - 1$. For $n \geq 1$ and each integer $i$ with $0 \leq i \leq \log n$, let $x_{n,i}$ be the string of length $n$ with value $i$, which exists because $i \leq \log n < 2^n$. For each such $n$, the strings $x_{n,i}$ are distinct, because their values are. Player II answers with the string $\tau(w)$ of length $2^{n+1} - 1$ that extends $w$, and thus settles all strings of length at most $n$. The new bits of $\tau(w)$ are determined as follows. For every integer $i$ with $0 \leq i \leq \log n$, the bit at the position of $x_{n,i}$ is 1 if $D_i$ outputs 0 on $x_{n,i}$, and 0 otherwise. Every other bit settled by $\tau(w)$ but not by $w$ is 0. For $n = 0$, the strings $x_{n,i}$ do not occur, and her answer $\tau(w)$ is the one-bit string 0. Player I may then extend the string as much as he likes, but he cannot undo what she wrote.

*Proof of Theorem 28.1(2).* First, the strategy $\tau$ is winning in G(P). Consider a play consistent with $\tau$, and let $L$ be the language that the play produces. Each move of Player II settles all strings of length $n$, where $n$ is the least length of which the string before the move settles no string. At her next move, the least such length is therefore greater than $n$. The lengths $n$ that she treats thus increase strictly from move to move and tend to infinity. Let $i \in \omega$ be

given, and consider a move of Player II at which $n \geq 2^i$. Since $i \leq \log n$, at that move she places $x_{n,i}$ in $L$ exactly when $D_i$ outputs 0 on $x_{n,i}$, that is, exactly when $x_{n,i} \notin L_i$. Hence $L \neq L_i$, and since $i$ was arbitrary, part (1) of Lemma 29.1 gives $L \notin \mathrm{P}$.

Next, the strategy $\tau$ is computable in polynomial time. If $n < 16$, the string $w$ has length less than $2^{16}$, and there are only finitely many such strings; the time spent on them is bounded by a constant. Suppose now that $n \geq 16$, and let $i \leq \log n$ be given, with $i = \pi(a, b)$. We have $b \leq i \leq \log n$, and part (2) of Lemma 29.1 bounds by $\eta(n^{\log n} + n + \log n + 1)$ the time that the algorithm $D$ of that lemma takes to produce the output of $D_i$ on $x_{n,i}$. The function $y \mapsto \sqrt{y} - \log y$ on $[16, \infty)$ vanishes at $y = 16$ and is increasing, since the derivative $1/(2\sqrt{y}) - \log e/y$ is positive when $\sqrt{y} > 2 \log e$; hence $(\log n)^2 \leq n$, which gives $n^{\log n} = 2^{(\log n)^2} \leq 2^n$. Each of the at most $\log n + 1 \leq n$ computations of $D_i$ therefore takes at most $\eta(3 \cdot 2^n)$ steps. By Fact 26.5(1), computing the positions of the strings $x_{n,i}$ by Lemma 26.1(1) and writing the $2^{n+1} - 1 - |w|$ new bits take a number of steps bounded by a polynomially bounded function of $2^n$. Since $2^n \leq 2|w|$, the total time is bounded by a polynomially bounded function of $|w|$, and the class P is therefore meager in E. ☺

The one clever point is the restriction $i \leq \log n$. Player II cannot afford to run $D_i$ on short strings when $i$ is large: the $n^b$ steps of the simulation could exceed her time budget. That budget is a polynomial in $2^n$ fixed once and for all. A machine that runs for $n^{1000}$ steps on strings of length $n$, for instance, needs $10^{1000}$ steps at length $n = 10$, where Player II writes only about $2^{10}$ bits. A strategy that simulated every $D_i$ at lengths that grow too slowly with $i$ would therefore obey no polynomial bound at all. She waits instead until $n \geq 2^i$, when the strings are long enough that the simulation costs little relative to $|w|$. This "diagonalization with delay" is a standard device of complexity theory, which the game makes appear natural.

**Circuit size: a counting argument, one length at a time.** Here the resource bound is space instead of time. The key is Shannon's counting argument (1949): for every positive integer $n$, there are $2^{2^n}$ Boolean functions of $n$ variables, one for each truth table, but far fewer small circuits. For a positive integer $s$, a circuit with $s$ gates can be written down by listing the type and the two sources of each gate. A crude count of such lists already shows that, for every sufficiently large $n$, some function needs more than $2^n/(10n)$ gates (see Exercise 5 at the end of this part). To reach the bound $2^n/n$, one must observe that renaming the gates of a circuit does not change the function that the circuit computes.

**Lemma 29.2 (Shannon).** *For every integer $n \geq 64$, some function $f\colon \{0,1\}^n \to \{0,1\}$ is computed by no circuit with $n$ inputs and at most $2^n/n$ gates.*

*Proof.* Write $s = \lfloor 2^n/n \rfloor$, and for an integer $t$ with $1 \leq t \leq s$, let $\Phi_t$ be the set of functions $f\colon \{0,1\}^n \to \{0,1\}$ for which the least number of gates of a circuit computing $f$ is $t$. We show that $|\Phi_1 \cup \cdots \cup \Phi_s| < 2^{2^n}$.

Let $f \in \Phi_t$ be given, and let $\mathcal{K}$ be a circuit with $t$ gates that computes $f$. No two gates of $\mathcal{K}$ have the same type and the same ordered pair of sources. Indeed, if $g_j$ and $g_l$ did, with $j < l$, they would take the same value at every input. Replacing $g_l$ by $g_j$ wherever $g_l$ is a source or the output gate, and deleting $g_l$, would give a circuit with $t-1$ gates that computes $f$, contrary to the choice of $t$.

A *labeled description* with $t$ gates is a pair consisting of a function and an element of $\{1, \ldots, t\}$, the output label. The function assigns to each $j \in \{1, \ldots, t\}$ a type and an ordered pair of elements of $\{u_1, \ldots, u_n\} \cup \{1, \ldots, t\}$. There are at most $t\,(3(n+t)^2)^t$ labeled descriptions with $t$ gates. For every permutation $\sigma$ of $\{1, \ldots, t\}$, the circuit $\mathcal{K}$ yields a labeled description $\Delta_\sigma$, defined as follows. The function of $\Delta_\sigma$ assigns to $\sigma(j)$ the type of $g_j$ and

the pair obtained from the sources of $g_j$ by replacing each gate $g_l$ with the label $\sigma(l)$. The output label of $\Delta_\sigma$ is $\sigma(o)$, where $g_o$ is the output gate of $\mathcal{K}$. Fix an input $z \in \{0,1\}^n$, and give the input variable $u_m$ the value $z_m$ for $1 \le m \le n$. Call an assignment of values to the labels of $\Delta_\sigma$ *admissible* if each label receives the value prescribed by the type of the label. Depending on that type, the value is the minimum of the values of the sources of the label, their maximum, or 1 minus the value of the source. Giving the label $\sigma(j)$ the value of $g_j$ at $z$, for $1 \le j \le t$, is admissible. Any two admissible assignments agree at $\sigma(1), \sigma(2), \ldots, \sigma(t)$ in turn, by induction, because the sources of $\sigma(j)$ lie among the input variables and the labels $\sigma(1), \ldots, \sigma(j-1)$. The description $\Delta_\sigma$ alone thus determines a unique admissible assignment at $z$, whose value at the output label is $f(z)$. We claim that the $t!$ descriptions $\Delta_\sigma$ are distinct. Suppose that $\Delta_\sigma = \Delta_\rho$, and let $\theta = \rho^{-1}\sigma$. Then, for every $j$, the gate $g_{\theta(j)}$ has the type of $g_j$, and the sources of $g_{\theta(j)}$ are those of $g_j$ with each gate $g_l$ replaced by $g_{\theta(l)}$. By induction on $j$ we show that $\theta(j) = j$. If $\theta(l) = l$ for all $l < j$, the sources of $g_j$, which are input variables or gates $g_l$ with $l < j$, are unchanged. In that case the gate $g_{\theta(j)}$ has the same type and the same sources as $g_j$, and since no two gates of $\mathcal{K}$ are alike in this way, we have $\theta(j) = j$. Hence $\sigma = \rho$, and the claim is proved.

Different functions cannot share a labeled description, since a description determines, at every input, the value at the output label. Each $f \in \Phi_t$ therefore has $t!$ labeled descriptions with $t$ gates, shared with no other function, and

$$|\Phi_t| \;\le\; \frac{t\,(3(n+t)^2)^t}{t!} \;\le\; t\Big(\frac{A}{t}\Big)^t, \qquad \text{where } A = 3e(n+s)^2,$$

because $t! \ge (t/e)^t$ (one of the terms of the series $e^t = \sum_{m\in\omega} t^m/m!$ is $t^t/t!$) and $n+t \le n+s$. The function $t \mapsto t\log(A/t)$ has derivative $\log(A/t) - \log e$, which is positive for $t < A/e$. The number $s$ is less than $A/e = 3(n+s)^2$; hence $(A/t)^t \le (A/s)^s$ for $1 \le t \le s$, and

$$|\Phi_1 \cup \cdots \cup \Phi_s| \;\le\; s^2\Big(\frac{A}{s}\Big)^s.$$

Since $n \ge 64$, we have $s \ge n$, hence $n+s \le 2s$ and $A/s \le 12e\,s$; we also have $\log(12e) < 5.03$. Taking logarithms, and using $\log s \le n - \log n$ and $\log n \ge 6$, we obtain

$$\log|\Phi_1\cup\cdots\cup\Phi_s| \;\le\; 2n+s(\log s+5.03) \;\le\; 2n+\frac{2^n}{n}(n-\log n+5.03) \;\le\; 2^n+2n-\frac{0.97\cdot 2^n}{n} \;<\; 2^n,$$

where the last inequality holds because $2^n > 3n^2$. There are $2^{2^n}$ functions from $\{0,1\}^n$ to $\{0,1\}$, and some of them lie in no $\Phi_t$ with $t \le s$. Such a function is computed by no circuit with at most $s$ gates, that is, with at most $2^n/n$ gates. ☺

Player II turns this count into a strategy. Let $w$ be the current string, and let $n$ be the least length such that $w$ settles no string of length $n$. If $n < 64$, she extends $w$ by zeros until the new string settles every string of length $n$. Otherwise she considers the functions computed by circuits with $n$ inputs and at most $2^n/n$ gates. By brute force, she finds the first binary string $T$ of length $2^n$ in dictionary order that is the truth table of none of these functions. Her search runs through the strings of length $2^n$ one by one. For each of them, she checks every circuit with at most $2^n/n$ gates, evaluated at all $2^n$ inputs, to see whether the circuit computes the function with that truth table. She then fills the remaining positions before the strings of length $n$ with 0 and appends $T$. The resulting string settles the strings of length $n$, and makes $T$ the truth table of $f_{L,n}$ for the language $L$ produced. The search takes an astronomical amount of *time*, but very little *space*: at any moment she needs to remember only one candidate truth table, one circuit and the values of the gates of that circuit. Space, unlike time, can be used again, and Player II is charged only for space.

*Proof of Theorem 28.1(3).* By Lemma 29.2, the string $T$ exists whenever $n \geq 64$, and the strategy is well defined. First, the strategy is winning in G(SIZE($2^n/n$)). Consider a play consistent with the strategy, and let $L$ be the language that the play produces. As in the proof of part (2), each move of Player II settles all strings of length $n$, where $n$ is the least length of which the string before the move settles no string. The lengths $n$ so treated increase strictly from move to move. There are therefore infinitely many $n \geq 64$ for which the truth table of $f_{L,n}$ is the string $T$ of one of her moves. For these $n$, no circuit with at most $2^n/n$ gates computes $f_{L,n}$, and hence $L \notin \mathrm{SIZE}(2^n/n)$.

It remains to be shown that the strategy is computable in polynomial space. If $n < 64$, the string $w$ has length less than $2^{64}$, and the finitely many such strings require a bounded number of cells. Suppose now that $n \geq 64$; the string $w$ settles some string of length $n-1$, whence $|w| \geq 2^{n-1}$ by Lemma 26.1(1). A circuit with at most $2^n/n$ gates is described by at most $2^n/n$ triples, followed by the $n+1$ bits that name the output gate. Each triple consists of a type, written with two bits, and two sources, written in binary with $n+1$ bits each (there are at most $n + 2^n/n \leq 2^{n+1}$ possible sources). The description of such a circuit has at most $(2n+4)2^n/n + n + 1 \leq 3 \cdot 2^n$ bits, because $n \geq 64$; the values of the gates of the circuit at one input take at most $2^n/n$ bits. Player II stores the candidate $T$, the description of the current circuit, the current input, the values of the gates and a few counters. By part (1) of Fact 26.5, these data therefore fit into a number of cells bounded by a constant multiple of $2^n$. The cells are used again for each new candidate and each new circuit. The answer itself has length $2^{n+1} - 1$. Since $2^n \leq 2|w|$, the space used is bounded by a polynomially bounded function of $|w|$, and the class SIZE($2^n/n$) is therefore meager in ESPACE. ☺

## Exercises

(1) Show that P $\subseteq$ E $\subseteq$ ESPACE, and that a class of languages that is meager in E is meager in ESPACE.

(2) Consider the game G($\mathcal{C}$) without resource bounds, where $\mathcal{C}$ is the class of languages that contain the empty string $\lambda$. Show that Player I has a winning strategy, and explain why this agrees with Proposition 27.2.

(3) Show that $(\log n)^2 > n$ for every integer $n$ with $5 \leq n \leq 15$, whereas $(\log n)^2 \leq n$ for $1 \leq n \leq 4$ and for $n \geq 16$. Explain why the proof of Theorem 28.1(2) needs the inequality $(\log n)^2 \leq n$ only for large $n$.

(4) Show that the class of finite languages is meager in E, by exhibiting a positional strategy for Player II that is computed in time linear in the length of the current string.

(5) (∗) Show that, for all positive integers $n$ and $s$, there are at most $s^2\,(3(n+s)^2)^s$ circuits with $n$ inputs and at most $s$ gates. Deduce, without the renaming argument in the proof of Lemma 29.2, that for every sufficiently large $n$ some function from $\{0,1\}^n$ to $\{0,1\}$ is computed by no circuit with at most $2^n/(10n)$ gates.

Part
# VII

# The Scottish Book, von Neumann, Ulam, Texas, Toronto, and the triumph of games, logic, topology and combinatorics

We now return to Lwów, and from the mathematics of these notes to the people behind that mathematics: where they worked, what happened to them, and how their problems traveled. Many of them were very young. When the Scottish Book was started, in the summer of 1935, Stanisław Ulam was twenty-six, Stanisław Mazur was thirty, Mark Kac was twenty, and Stefan Banach, who led the discussions at the café table, was forty-three. The journey of the Scottish Book and of the people who carried the notebook is shown in Figure 33, and the timeline in Figure 34 is worth keeping at hand throughout.

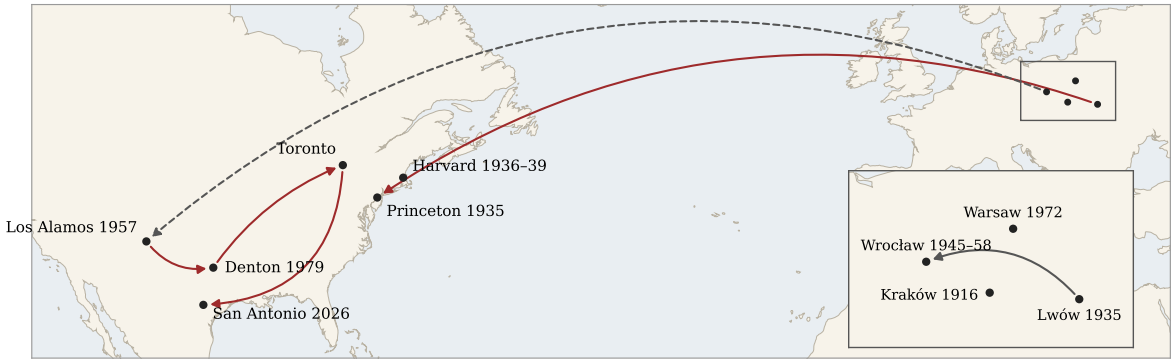



Figure 33. The journey of the Scottish Book, with Ulam's crossing in 1935 and the road from Los Alamos to Denton, Toronto and San Antonio in red, the typewritten copy that Hugo Steinhaus sent from Wrocław to Ulam in 1956 dashed, and an inset of Poland.

## 30. The marble tables of the Scottish Café

Between the two world wars, the city of Lwów (then in southeastern Poland, now Lviv, in Ukraine) was home to one of the liveliest mathematical communities of the century. Led

| | |
|---|---|
| 1916 | Steinhaus meets Banach and Nikodým in the Planty park, Kraków. |
| 1920 | Banach's doctorate in Lwów, without a university degree. |
| 1930 | Ulam on measures defined on all subsets of a set (measurable cardinals). |
| 1932 | Banach's *Théorie des opérations linéaires*. |
| 1935 | 17 July: Problem 1, by Banach. Mazur's Problem 43; 4 August: Banach's answer. End of the year: Ulam goes to Princeton. |
| 1936 | 6 November: Mazur's Problem 153, the goose problem. |
| 1939 | 20 August: Ulam leaves Poland for the last time. September: war. |
| 1941 | 31 May: last problem, number 193, by Steinhaus. July: the Lwów professors are murdered. |
| 1945 | 31 August: Banach dies in Lwów. Steinhaus moves to Wrocław. |
| 1946 | 2 July: first entry in the New Scottish Book, by Steinhaus. |
| 1946 | Ulam's idea of the Monte Carlo method. |
| 1957 | Ulam's English translation, mimeographed at Los Alamos; Oxtoby publishes the first proof of Banach's solution of Problem 43. |
| 1962 | Mycielski and Steinhaus propose the Axiom of Determinacy. |
| 1972 | Mazur gives Enflo a live goose in Warsaw. |
| 1975 | Martin proves Borel determinacy. |
| 1979 | The Scottish Book Conference in Denton, organized by Mauldin. |
| 1981 | *The Scottish Book*, edited by Mauldin (second edition 2015). |
| 1994–98 | Distortion problem (Odell and Schlumprecht); Gowers's dichotomy; Fields Medal for Gowers. |
| 2026 | The joint UT San Antonio–Toronto seminar. |

FIGURE 34. A timeline of the Scottish Book.

by Stefan Banach and Hugo Steinhaus, with Stanisław Mazur, Juliusz Schauder, Stanisław Ulam, Mark Kac and many others, the Lwów School did much to create modern functional analysis and, through Mazur's Problem 43, gave topology its first infinite game.

The school began with a chance meeting. One evening in 1916, walking in the Planty, the park around the old town of Kraków, Steinhaus overheard the words "Lebesgue integral" and stopped to talk to the two young men who had said them. They were Stefan Banach and Otto Nikodým. Banach, then twenty-four, had studied engineering in Lwów without finishing his degree, and was teaching himself mathematics. Steinhaus mentioned a problem on which he had been working; a few days later Banach came back with the solution, and their first joint paper followed. In 1920, with Steinhaus's support, Banach was allowed to submit a doctoral thesis in Lwów without a university degree; the thesis, published in 1922, contains what we now call Banach spaces. Steinhaus, whose own contributions to mathematics were many, later said that Banach was his greatest scientific discovery. In 1935, fifteen years after the thesis, Banach opened the Scottish Book with Problem 1, which asks when a metric space can be remetrized so as to become complete and compact while every sequence that converged in the old metric still converges in the new one; within weeks he had also answered Mazur's Problem 43.

The school's real seminar room was the *Kawiarnia Szkocka*, the Scottish Café, on Plac Akademicki (see Figure 35). The mathematicians spent hours at its tables over coffee and, as the stories go, wrote their formulas directly on the marble tabletops, which were wiped clean when the café closed. To save at least the problems, Banach's wife, Łucja Banachowa, is said to have provided a large notebook, which was kept by the headwaiter and brought to the table on request. Thus was born the *Scottish Book* (*Księga Szkocka*); the first problem, by Banach, is dated 17 July 1935. The sessions could be long: Ulam recalled one with Mazur

and Banach that lasted seventeen hours, interrupted only for meals [Ula76]. He also recorded Banach's view of what a mathematician does: "Good mathematicians see analogies between theorems or theories, the very best ones see analogies between analogies." Problems came with prizes: a cup of coffee, a small beer, five beers, a bottle of wine, 100 grams of caviar, a dinner at the Hotel George, a fondue in Geneva, a bottle of whisky "of measure $> 0$" (John von Neumann, 1937), a kilo of bacon (Stanisław Saks, 1940), and a live goose (see Section 36).




FIGURE 35. Plac Akademicki in Lwów, the square of the Scottish Café, on a hand-colored postcard of 1909 (Kraków: Wydawnictwo Salonu Malarzy Polskich; National Library of Poland, Polona, public domain).

## 31. Problem 43

Mazur's Problem 43, the subject of these notes, was written in the summer of 1935, in the first weeks of the book. In the Introduction we saw what the problem asks and how Banach answered on 4 August (see Figure 1), and in Section 7 we recalled Oxtoby's paper of 1957, whose proof, the first to be published, turned the answer into a theorem. Right after Problem 43 in the book, Ulam proposed a variant in which the players choose binary digits, one digit at a time. This variant later inspired Steinhaus and Jan Mycielski to propose, in 1962, a new axiom for set theory, the *Axiom of Determinacy* [MS62]; the game of Section 13 that bears Ulam's name is the variant in which the players choose non-empty finite binary strings. A question about the size of sets of real numbers, asked over coffee, had become a question in the foundations of mathematics.

## 32. Von Neumann's invitation and the bridge to America

John von Neumann met Ulam in Warsaw in 1935 and invited him to spend a few months at the Institute for Advanced Study in Princeton. Ulam sailed for America in December 1935.

Ulam was the most prolific author of the Scottish Book: forty problems are his alone, and about twenty more are joint with others, among them Banach, Mazur and Józef Schreier. His problems have had a long life. Problem 19 asks whether a solid of uniform density that floats in water in equilibrium in every position must be a ball. In 2022 Dmitry Ryabogin showed that the answer is no: there is a strictly convex body of revolution in $\mathbb{R}^3$, of density $1/2$ and not a ball, that floats in equilibrium in every position [Rya22]. Ulam's earlier work, of 1930, on measures defined on all subsets of a set led to the measurable cardinals (see Section 39 below).

At a tea in von Neumann's house, Ulam met G. D. Birkhoff, who asked whether he would like to join the Society of Fellows at Harvard. From 1936 to 1939 Ulam divided his time between Harvard and Lwów, returning each summer to Poland and to the tables of the Scottish Café. On 20 August 1939 he left Poland for the last time, sailing from Gdynia with his seventeen-year-old brother Adam. Twelve days later Germany invaded Poland; on 17 September the Soviet Union followed. The brothers' father, Józef, who had seen them off at the quay, was killed in the Holocaust, together with their sister Stefania and much of the family.

## 33. War, and the survival of the book

The war destroyed the Lwów School. From September 1939 the city was under Soviet occupation, and the café meetings went on: the problem for which Stanisław Saks offered a kilo of bacon dates from February 1940. German troops entered Lwów on 30 June 1941. On the night of 3–4 July, in a planned action against the Polish intelligentsia, they arrested professors of the city's universities together with members of their families, and shot them on the Wuleckie Hills. Among them were the mathematicians Antoni Łomnicki and Włodzimierz Stożek, together with two of Stożek's sons. Kazimierz Bartel, geometer and former prime minister of Poland, was arrested on 2 July and killed later that month.

The killing did not stop there. Saks was shot by the Gestapo in Warsaw in 1942. Herman Auerbach died the same year, deported from the Lwów ghetto to the Bełżec death camp. Juliusz Schauder, of the Schauder fixed point theorem, was murdered in 1943. Mark Kac had left for the United States in 1938; his parents and his brother were killed in 1942. Steinhaus survived by hiding in the countryside under a false name.

Banach survived because of Rudolf Weigl, a biologist who had developed a vaccine against typhus. The vaccine was grown in lice, and the lice had to be fed on human blood: "louse feeders" let cages of lice feed on their legs. The German army needed the vaccine, and the papers of Weigl's institute protected those who worked there; Weigl used that protection to employ mathematicians, scientists, members of the underground and Jews. Banach fed lice there, and so did several other members of the school. Weigl also smuggled vaccine into the Warsaw and Lwów ghettos. In 2003 Yad Vashem recognized him as Righteous Among the Nations. Banach lived to see the end of the war; he died of lung cancer in Lwów on 31 August 1945.

The book survived too. The last entry of the book, Problem 193, had been written by Steinhaus on 31 May 1941, a month before the Germans arrived. Ulam recalled that Mazur had once proposed, in case the city was bombed, to place the manuscripts and the Scottish

Book in a case and to bury the case near a goal post of a football field outside the city; whether he did so is not known. What is known is that after the war Banach's son, Stefan Banach Jr., brought the book to Wrocław, where Steinhaus had moved.

## 34. Ulam after Lwów

The mathematics of the café went on in Ulam's life (see Figure 36) in unexpected ways. In 1943, after asking von Neumann for war work, he received an invitation, signed by Hans Bethe, to an unnamed project near Santa Fe; in February 1944 he arrived at Los Alamos and joined the Manhattan Project. In 1946, while recovering from encephalitis and brain surgery, Ulam played a great deal of solitaire. He wondered what the chance of winning a game was, and saw that it was easier to estimate that chance by playing many random games than by counting. With von Neumann he turned this idea into the *Monte Carlo method*, named by their colleague Nicholas Metropolis after Monte Carlo, where an uncle of Ulam's liked to gamble [MU49]; the method is now among the most widely used in science. With Edward Teller he conceived the Teller–Ulam design of the hydrogen bomb. In the early 1950s, with Enrico Fermi and John Pasta, he simulated on the MANIAC computer a chain of oscillators that, unexpectedly, did not reach thermal equilibrium; their report appeared in 1955 [FPU55]. The programs were written by Mary Tsingou, whose contribution received wider recognition decades later, when the problem became known as the Fermi–Pasta–Ulam–Tsingou problem [Dau08].

## 35. The 1957 Los Alamos mimeograph

Steinhaus had a typewritten copy of the book made and in 1956 sent the copy to Ulam, who was by then at the Los Alamos Scientific Laboratory in New Mexico. Ulam translated the book into English and, in his words, "had some three hundred mimeographed copies of the translation made. I had to pay for this myself—Los Alamos is a government laboratory, and one cannot use taxpayers' money for such frivolous purposes." The copies circulated among mathematicians around the world, and for a generation they were how many mathematicians came to know the problems.

## 36. Problem 153: a live goose

The best-known prize of the Scottish Book was paid thirty-six years after it was offered. Mazur's Problem 153 of 6 November 1936 asked a question about approximating continuous functions of two variables and offered a live goose. The problem turned out to be equivalent to the approximation problem formulated by Grothendieck in 1955, the question whether every Banach space has the *approximation property*. A Banach space $X$ has this property if every compact operator from a Banach space into $X$ is a norm limit of operators of finite rank. A negative answer would also settle Banach's basis problem, the question whether every separable Banach space has a *Schauder basis*, a sequence $(e_k)_{k\in\mathbb{N}}$ of vectors such that every vector of the space is the sum of a norm-convergent series $\sum_{k\in\mathbb{N}} a_k e_k$ for unique scalars $a_k$. Indeed, a space with a Schauder basis has the approximation property. In 1972 the young Swedish mathematician Per Enflo constructed a separable Banach space without the approximation property, and therefore without a Schauder basis [Enf73]. Mazur, then in Warsaw, kept his word: that same year, in a ceremony at the Stefan Banach Center, he presented the twenty-eight-year-old Enflo with a live goose. According to several accounts, the ceremony was broadcast on Polish television.

Enflo is also a concert pianist. In 1956, at the age of eleven, he won the Swedish competition for young pianists, and he won it again in 1961; among his teachers was Géza Anda. He has gone on performing ever since.

## 37. The road to Denton: Dan Mauldin and North Texas

In the 1970s the story moved to Texas. R. Daniel Mauldin, a measure theorist and descriptive set theorist at North Texas State University (now the University of North Texas) in Denton, took on the task of finding out what had become of the 193 problems. In May 1979 he organized the *Scottish Book Conference* in Denton.

The list of speakers reads like a family reunion. Ulam, Mark Kac, Antoni Zygmund, Paul Erdős and Andrzej Granas all lectured, Ulam on the anecdotal history of the book, Kac on his personal history of the book, and Zygmund on Steinhaus and the development of Polish mathematics. R. D. Anderson and D. A. Martin spoke about work that went back to the Scottish Book. Martin had proved Borel determinacy in 1975: every infinite game of perfect information in which the players choose natural numbers, and whose payoff set is a Borel subset of the space $\omega^\omega$ of infinite sequences of natural numbers, is determined (see Section 15). John Oxtoby, who had published the first proof of Banach's solution of Problem 43, toasted a "condensation of Poles" at the conference; R. H. Bing, a pillar of Texas topology, also offered a toast.

From the conference came the book edited by Mauldin, *The Scottish Book: Mathematics from the Scottish Café* (Birkhäuser, 1981) [Mau81]. The book contains lectures from the conference, Ulam's "anecdotal history" of the book among them, and the problems, with commentaries on the later history of each. A second edition appeared in 2015 [Mau15].

## 38. The New Scottish Book

The tradition did not die in Lwów. In Wrocław, where much of the Lwów university was reestablished after the war, Steinhaus started the *New Scottish Book* (*Nowa Księga Szkocka*). The first entry, by Steinhaus, is dated 2 July 1946; problems were entered regularly until 1979, and occasionally until 1987. The second edition of Mauldin's book includes a short history of the New Scottish Book and a selection of the problems entered there.

## 39. Set-theoretic topology and Ramsey spaces at Toronto

The questions of the Scottish Book on measure, category and infinite games also run through set-theoretic topology, a field in which the University of Toronto has long been a leading center.

In 1930 Ulam proved that a countably additive probability measure that is defined on all subsets of a set and vanishes on points can exist only if the set is very large: at least as large as a weakly inaccessible cardinal, and, if the measure takes only the values 0 and 1, at least as large as a strongly inaccessible one [Ula30]. In particular, such a measure on all subsets of the real line can exist only if the continuum is at least as large as a weakly inaccessible cardinal; in fact, the least cardinal that carries such a measure is then weakly inaccessible and no larger than the continuum. Along the way Ulam introduced what are now called Ulam matrices. A set carries a two-valued measure of this kind exactly when the set is at least as large as a measurable cardinal, and measurable cardinals, with the larger cardinals that followed, became a standard tool for consistency proofs in topology. At Toronto, Franklin D. Tall and his collaborators applied such set-theoretic methods to classical problems of point-set topology, most notably the normal Moore space problem. Infinite games of the

Banach–Mazur type, and their relatives such as the Menger game and the point-open game of Part III, are now everyday tools in the field.

Stevo Todorčević, at Toronto, developed the theory of Ramsey spaces, which unifies the Galvin–Prikry theorem, Ellentuck's theorem, Gowers's $\mathrm{FIN}_k$ theorem and much else in a single axiomatic framework [Tod10]. Here $\mathrm{FIN}_k$, for a positive integer $k$, is the set of finitely supported functions from $\mathbb{N}$ to $\{0, 1, \dots, k\}$ that take the value $k$. The circle closes in the present: Tall is one of the authors, with Eduardo Dueñez and José Iovino of UT San Antonio and Tonatiuh Matos-Wiederhold and Luciano Salvetti of Toronto, of recent work on the complexity of deep computations through the topology of function spaces [DIMW$^+$26]; that work is the subject of the seminar for which these notes were written.

## 40. The triumph of games: combinatorics, logic and the real-valued setting

Texas also became a center of the Banach space theory that Banach's 1932 monograph, *Théorie des opérations linéaires*, and the Scottish Book had launched. Edward Odell at the University of Texas at Austin and Thomas Schlumprecht at Texas A&M University were at the heart of the Texas school.

In 1991 Schlumprecht constructed the first arbitrarily distortable Banach space [Sch91], a space that admits, for every $\lambda > 1$, an equivalent norm whose supremum on the original unit sphere of every infinite-dimensional subspace exceeds $\lambda$ times the infimum of the new norm there. In 1994 Odell and Schlumprecht used his ideas to solve the distortion problem, the question whether $\ell_2$ admits, for some $\lambda > 1$, an equivalent norm of this kind: the Hilbert space $\ell_2$ is itself arbitrarily distortable [OS94]. Schlumprecht's space was also the starting point of W. T. Gowers and B. Maurey, who in 1991 (published in 1993) answered in the negative the old unconditional basic sequence problem, the question whether every infinite-dimensional Banach space contains an unconditional basic sequence, a sequence that is a Schauder basis of the closed subspace that the sequence spans and whose expansions converge whatever the order of their terms (see Section 12). They constructed a "hereditarily indecomposable" space, one in which no infinite-dimensional closed subspace is the direct sum of two infinite-dimensional closed subspaces; in particular, no infinite-dimensional subspace has an unconditional basis [GM93].

The last step needed a different kind of mathematics. Banach's homogeneous space problem, from his 1932 monograph, asks whether every *homogeneous* Banach space, an infinite-dimensional space $X$ isomorphic to every infinite-dimensional closed subspace of $X$, is isomorphic to $\ell_2$. To answer the problem, Gowers turned to infinite combinatorics and logic.

The bridge was infinite-dimensional Ramsey theory, the subject of Section 11: the Galvin–Prikry theorem, Silver's theorem and Ellentuck's theorem, all proved by logicians, say that definable colorings of the infinite subsets of $\omega$ have infinite homogeneous sets. Gowers saw that this machinery could be imported into Banach space theory once translated into a continuous, real-valued setting, with block sequences of vectors in place of infinite sets of integers, and with a winning strategy in a two-player game in place of a homogeneous set. His Ramsey theorem for this game yields his dichotomy theorem: every infinite-dimensional Banach space contains an infinite-dimensional closed subspace that either has an unconditional basis or is hereditarily indecomposable. The dichotomy, combined with the theorem of Gowers and Maurey that a hereditarily indecomposable space $Y$ is isomorphic to no proper subspace of $Y$ and with a theorem of Komorowski and Tomczak-Jaegermann, implies that every infinite-dimensional homogeneous Banach space is isomorphic to $\ell_2$ [Gow96] (see Section 12).

When Gowers received the Fields Medal in 1998, the cycle was complete. Infinite games, born at a café table in Lwów in 1935, had passed through the foundations of mathematics (determinacy) and through the hands of logicians and combinatorialists (infinite Ramsey theory); by way of Texas, Paris and Cambridge, the games had come back to analysis to settle a problem posed by Banach himself. The notebook that survived a war, and the people who survived it, or did not, and the colleagues who carried their problems across an ocean, are part of that triumph.

## 41. Epilogue: A Scottish Book of our own

The Scottish Book was not a journal but a notebook on a café table, open to anyone who sat down, where a problem could be written before being understood and a prize could be a beer or a goose. That spirit is easy to revive. We propose a Scottish Book of our own for the participants of this seminar, in San Antonio and in Toronto: a shared notebook of the problems that come up in our discussions, each signed and dated, with a prize chosen by the author of the problem, and with room for the answers, partial results and counterexamples that follow. Students are especially welcome: some of the first authors of the original book were in their twenties, and some of the best problems of that book are short questions that anyone can state.

## Further reading

Ulam's autobiography [Ula76] and Kac's [Kac85] are the best first-hand accounts of the Lwów School and of what came after. Kałuża's biography of Banach [Kał96], written with a reporter's eye, tells the story of his life and of his city. Mauldin's book [Mau81, Mau15] contains the problems themselves, with commentaries on the fate of each.

## Appendix A. Real-valued structures

In this appendix we recall the structures and the first-order formulas used in Section 18, in the notation of Aubone, Dueñez and Iovino [ADI26], whose structures are Keisler's general structures [Kei23]. We use only the first-order part of their language, with finitary disjunctions; their paper treats the infinitary language $L_{\omega_1\omega}$, which also allows countable joins, and contains the details that we omit. As in that paper, the symbols of a vocabulary are printed in this color, and the variables, the connectives and the formal expressions built from them in this color.

**Vocabularies.** A *sort descriptor* is a purely syntactic symbol $M$. Let $\mathbb{M}$ be a non-empty set of sort descriptors. A *function symbol* for $\mathbb{M}$ is a purely syntactic symbol $f$ given together with a natural number $n$, called the *arity* of $f$, and an $(n+1)$-tuple $(M_i)_{i\leq n}$ of sort descriptors in $\mathbb{M}$, called the *extended arity* of $f$; we indicate the arity and the extended arity by writing $f\colon \prod_{i<n} M_i \to M_n$. A function symbol of arity 0 is called a *constant symbol* of sort $M_0$. A *predicate symbol* for $\mathbb{M}$ is a purely syntactic symbol $p$ given together with a natural number $n$, called the arity of $p$, an $n$-tuple $(M_i)_{i<n}$ of sort descriptors in $\mathbb{M}$, called the extended arity of $p$, and a real number $r_p \geq 0$, called the *a priori bound* of $p$; we write $p\colon \prod_{i<n} M_i \to [-r_p, r_p]$. A *vocabulary* is a triple $V = (\mathbb{M}, F, P)$ consisting of a non-empty set $\mathbb{M}$ of sort descriptors, a set $F$ of function symbols for $\mathbb{M}$ and a set $P$ of predicate symbols for $\mathbb{M}$; the set $P$ contains a nullary predicate symbol 1, the *logical unity constant*, with a priori bound 1. The vocabulary is *countable* if $\mathbb{M}$, $F$ and $P$ are countable. There is no equality symbol; when equality is needed, a binary predicate symbol $=$ with a priori bound 1 is added to the vocabulary and interpreted as described below.




FIGURE 36. Stanisław Ulam in 1962, photographed by Alfred Eisenstaedt for *LIFE*.

**Structures.** A *$V$-structure* is a triple $\mathfrak{M} = \langle (M^{\mathfrak{M}})_{M\in\mathbb{M}}, (f^{\mathfrak{M}})_{f\in F}, (p^{\mathfrak{M}})_{p\in P}\rangle$ in which the interpretation $M^{\mathfrak{M}}$ of each sort descriptor is a non-empty set, the interpretation of each function symbol $f\colon \prod_{i<n} M_i \to M_n$ is a function $f^{\mathfrak{M}}\colon \prod_{i<n} M_i^{\mathfrak{M}} \to M_n^{\mathfrak{M}}$, the interpretation of each predicate symbol $p\colon \prod_{i<n} M_i \to [-r_p, r_p]$ is a function $p^{\mathfrak{M}}\colon \prod_{i<n} M_i^{\mathfrak{M}} \to [-r_p, r_p]$, and $1^{\mathfrak{M}} = 1$. A nullary symbol is interpreted on the one-point set $\prod_{i<0} M_i^{\mathfrak{M}}$, and the interpretation is identified with the value at that point. No topology is prescribed with respect to which the functions or the predicates must be continuous. The value 1 stands for true and the value $-1$ for false, and a predicate is *two-valued* if the predicate takes only these two values. Classical first-order structures correspond to the case in which every predicate is two-valued. A predicate symbol $=$ is *interpreted by equality* if $=^{\mathfrak{M}}(a, b)$ is 1 when $a = b$ and $-1$ otherwise.

**Terms and formulas.** For each sort descriptor $M$ we fix a sequence $(x_i^M)_{i<\omega}$ of distinct *variables of sort $M$*, the sequences for distinct sort descriptors having no symbol in common. The *terms of sort $M$* are the variables of sort $M$ and the expressions $f(\bar\tau)$, where $f\colon \prod_{i<n} M_i \to M$ is a function symbol and $\bar\tau = (\tau_i)_{i<n}$ is a tuple of terms with $\tau_i$ of sort $M_i$; in particular, every constant symbol of sort $M$ is a term of sort $M$. For a predicate symbol $p\colon \prod_{i<n} M_i \to [-r_p, r_p]$ and such a tuple $\bar\tau$, the expression $p(\bar\tau)$ is an *atomic $V$-formula*, with a priori bound $r_p$. The *$V$-formulas* are built from the atomic formulas by the logical connectives $s\cdot$, one for each rational number $s$, the connectives $+$ and $\vee$, and the connectives $\sup_x$, one for each variable $x$. The a priori bounds of $s\cdot\phi$, $\phi+\psi$, $\phi\vee\psi$ and $\sup_x\phi$ are $|s|r_\phi$, $r_\phi + r_\psi$,

$\max\{r_\phi, r_\psi\}$ and $r_\phi$, respectively. The set of all $V$-formulas is the *first-order language* $L^V$. The free variables of a formula are defined as usual, with $\sup_x$ binding $x$, and a *$V$-sentence* is a formula without free variables. We write $\phi(\bar{x})$ to indicate that every variable free in $\phi$ is among the variables of the tuple $\bar{x}$.

**Interpretation.** Let $\mathfrak{M}$ be a $V$-structure. A term $\tau(\bar{x})$ is interpreted in $\mathfrak{M}$ as a function $\tau^{\mathfrak{M}}$ from the product of the sorts of the variables of $\bar{x}$ to the sort of $\tau$, by composing the interpretations of the function symbols. A formula $\xi(\bar{x})$ is interpreted as a function $\xi^{\mathfrak{M}}$ on the same product, by recursion: the value of $p(\bar{\tau})$ at $\bar{a}$ is $p^{\mathfrak{M}}$ applied to the values of the terms at $\bar{a}$; the connectives $s\cdot$, $+$ and $\vee$ are interpreted as multiplication by $s$, as the sum and as the maximum; and $\sup_x$ is interpreted as the supremum over the sort of $x$. By induction on formulas, every formula $\xi$ takes its values in $[-r_\xi, r_\xi]$, and the value of a sentence is a real number in that interval. We write $-\phi$ for the formula $(-1)\cdot\phi$, the minimum $\phi \wedge \psi$ for $-(-\phi \vee -\psi)$ and the infimum $\inf_x \phi$ for $-\sup_x -\phi$. For formulas that are two-valued in a structure, the connectives $-$, $\vee$ and $\wedge$ are negation, disjunction and conjunction there, and $\sup_x$ and $\inf_x$ are the existential and the universal quantifier over the sort of $x$. For a formula $\phi(\bar{x})$ and a real number $r$, the *axiom* $\phi \geq r$ asserts that $\phi^{\mathfrak{M}}(\bar{a}) \geq r$ for all tuples $\bar{a}$; a two-valued formula $\theta$ is asserted by the axiom $\theta \geq 1$.

**Substructures and embeddings.** Let $\mathfrak{M}$ and $\mathfrak{N}$ be $V$-structures. An *embedding* of $\mathfrak{M}$ into $\mathfrak{N}$ is a family $h = (h_M)_{M \in \mathbb{M}}$ of injective maps $h_M \colon M^{\mathfrak{M}} \to M^{\mathfrak{N}}$ that commutes with the interpretations of the function symbols and preserves the values of the predicate symbols:

$$h_{M_n}\big(f^{\mathfrak{M}}(a_0, \dots, a_{n-1})\big) = f^{\mathfrak{N}}\big(h_{M_0}(a_0), \dots, h_{M_{n-1}}(a_{n-1})\big),$$
$$p^{\mathfrak{N}}\big(h_{M_0}(a_0), \dots, h_{M_{n-1}}(a_{n-1})\big) = p^{\mathfrak{M}}(a_0, \dots, a_{n-1}).$$

An *isomorphism* is an embedding whose maps are bijections, and an *automorphism* of $\mathfrak{M}$ is an isomorphism of $\mathfrak{M}$ onto itself. A *substructure* of $\mathfrak{N}$ is given by non-empty subsets of the sorts that contain the interpretations of the constant symbols and are closed under the interpretations of the function symbols, with the interpretations of the symbols restricted to these subsets; the inclusion of a substructure is an embedding, and the image of an embedding is a substructure. The intersection of the substructures of $\mathfrak{N}$ that contain a given family of subsets of the sorts is the substructure *generated* by the family, and a structure is *finitely generated* if the structure is generated by finitely many elements; in Section 18 we also allow the empty structure, which arises in this way when the vocabulary has no constant symbols. By induction on formulas, an embedding preserves the values of the formulas built without the connectives $\sup_x$, and an isomorphism preserves the values of all formulas.

## References


[AB09] Sanjeev Arora and Boaz Barak, *Computational complexity: A modern approach*, Cambridge University Press, Cambridge, 2009.

[AB15] Leandro F. Aurichi and Angelo Bella, *When is a space Menger at infinity?*, Appl. Gen. Topol. **16** (2015), no. 1, 75–80.

[ABV18] Jinpeng An, Victor Beresnevich, and Sanju Velani, *Badly approximable points on planar curves and winning*, Adv. Math. **324** (2018), 148–202.

[ADI26] Olivia Aubone, Eduardo Dueñez, and José Iovino, *Model existence and Scott's theorem for infinitary real-valued logic: a topological perspective*, arXiv:2611.02247, 2026.

[ADW17] Samson Abramsky, Anuj Dawar, and Pengming Wang, *The pebbling comonad in finite model theory*, Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), IEEE, 2017, pp. 1–12.

[AF90] Miklós Ajtai and Ronald Fagin, *Reachability is harder for directed than for undirected finite graphs*, J. Symbolic Logic **55** (1990), no. 1, 113–150.

[Ale24] P. Alexandroff, *Sur les ensembles de la première classe et les ensembles abstraits*, C. R. Acad. Sci. Paris **178** (1924), 185–187.

[An16] Jinpeng An, *Two-dimensional badly approximable vectors and Schmidt's game*, Duke Math. J. **165** (2016), no. 2, 267–284.

[AR23] Samson Abramsky and Luca Reggio, *Arboreal categories: an axiomatic theory of resources*, Log. Methods Comput. Sci. **19** (2023), no. 3, Paper No. 14.

[AR24] ———, *An invitation to game comonads*, 2024, arXiv:2407.00606.

[AS18] Samson Abramsky and Nihil Shah, *Relating structure and power: comonadic semantics for computational resources*, 27th EACSL Annual Conference on Computer Science Logic (CSL 2018), LIPIcs, vol. 119, Schloss Dagstuhl, 2018, pp. 2:1–2:17.

[AS21] ———, *Relating structure and power: comonadic semantics for computational resources*, J. Logic Comput. **31** (2021), no. 6, 1390–1428.

[Asp68] Edgar Asplund, *Fréchet differentiability of convex functions*, Acta Math. **121** (1968), 31–47.

[Ban30] S. Banach, *Théorème sur les ensembles de première catégorie*, Fund. Math. **16** (1930), 395–398.

[Bar77] Jon Barwise, *On Moschovakis closure ordinals*, J. Symbolic Logic **42** (1977), no. 2, 292–296.

[Bec08] József Beck, *Combinatorial games: Tic-tac-toe theory*, Cambridge University Press, Cambridge, 2008.

[BJ95] Tomek Bartoszyński and Haim Judah, *Set theory: On the structure of the real line*, A K Peters, Wellesley, MA, 1995.

[BKM+20] Pablo Barceló, Egor V. Kostylev, Mikaël Monet, Jorge Pérez, Juan Reutter, and Juan Pablo Silva, *The logical expressiveness of graph neural networks*, International Conference on Learning Representations (ICLR), 2020.

[Boo70] David Booth, *Ultrafilters on a countable set*, Ann. Math. Logic **2** (1970), 1–24.

[Bou96] Ahmed Bouziad, *Every čech-analytic Baire semitopological group is a topological group*, Proc. Amer. Math. Soc. **124** (1996), 953–959.

[BPV11] Dzmitry Badziahin, Andrew Pollington, and Sanju Velani, *On a problem in simultaneous Diophantine approximation: Schmidt's conjecture*, Ann. of Math. (2) **174** (2011), no. 3, 1837–1883.

[Bro12] L. E. J. Brouwer, *Über Abbildung von Mannigfaltigkeiten*, Math. Ann. **71** (1912), 97–115.

[BY15] I. Ben Yaacov, *Fraïssé limits of metric structures*, J. Symbolic Logic **80** (2015), 100–115.

[Cam97] P. J. Cameron, *The random graph*, The Mathematics of Paul Erdős II, Algorithms Combin., vol. 14, Springer, Berlin, 1997, pp. 333–351.

[Can95] Georg Cantor, *Beiträge zur Begründung der transfiniten Mengenlehre*, Math. Ann. **46** (1895), 481–512.

[Car37] Henri Cartan, *Théorie des filtres*, C. R. Acad. Sci. Paris **205** (1937), 595–598.

[Čec37] Eduard Čech, *On bicompact spaces*, Ann. of Math. (2) **38** (1937), no. 4, 823–844.

[CFI92] Jin-Yi Cai, Martin Fürer, and Neil Immerman, *An optimal lower bound on the number of variables for graph identification*, Combinatorica **12** (1992), no. 4, 389–410.

[CFI+24] Marco Carmosino, Ronald Fagin, Neil Immerman, Phokion Kolaitis, Jonathan Lenchner, and Rik Sengupta, *Multi-structural games and beyond*, Log. Methods Comput. Sci. **20** (2024), no. 4, Paper No. 27.

[Cho58] G. Choquet, *Une classe régulière d'espaces de Baire*, C. R. Acad. Sci. Paris **246** (1958), 218–220.

[Cho69] ———, *Lectures on analysis. Vol. I: Integration and topological vector spaces*, Mathematics Lecture Note Series, W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[Chr81] Jens Peter Reus Christensen, *Joint continuity of separately continuous functions*, Proc. Amer. Math. Soc. **82** (1981), 455–461.

[Dan85] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*, J. Reine Angew. Math. **359** (1985), 55–89.

[Dan86] ———, *Bounded orbits of flows on homogeneous spaces*, Comment. Math. Helv. **61** (1986), 636–660.

[Dau08] Thierry Dauxois, *Fermi, Pasta, Ulam, and a mysterious lady*, Physics Today **61** (2008), no. 1, 55–57.

[Dav64] Morton Davis, *Infinite games of perfect information*, Advances in Game Theory, Annals of Mathematics Studies, no. 52, Princeton University Press, 1964, pp. 85–101.

[DIMW+26] E. Dueñez, J. Iovino, T. Matos-Wiederhold, L. Salvetti, and F. D. Tall, *Complexity of deep computations via topology of function spaces*, 2026, arXiv:2601.00528.

[DJR21] Anuj Dawar, Tomáš Jakl, and Luca Reggio, *Lovász-type theorems and game comonads*, Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), IEEE, 2021, pp. 1–13.

[EF99] Heinz-Dieter Ebbinghaus and Jörg Flum, *Finite model theory*, second ed., Springer Monographs in Mathematics, Springer, Berlin, 1999.

[EHMR84] P. Erdős, A. Hajnal, A. Máté, and R. Rado, *Combinatorial set theory: Partition relations for cardinals*, Studies in Logic and the Foundations of Mathematics, vol. 106, North-Holland Publishing Co., Amsterdam, 1984.

[Ehr57] Andrzej Ehrenfeucht, *Application of games to some problems of mathematical logic*, Bull. Acad. Polon. Sci. Cl. III **5** (1957), 35–37.

[Ehr61] ———, *An application of games to the completeness problem for formalized theories*, Fund. Math. **49** (1961), 129–141.

[Ell57] Robert Ellis, *Locally compact transformation groups*, Duke Math. J. **24** (1957), 119–125.

[Ell74] Erik Ellentuck, *A new proof that analytic sets are Ramsey*, J. Symbolic Logic **39** (1974), 163–165. MR 0349393 (50 #1887)

[Enf73] Per Enflo, *A counterexample to the approximation problem in Banach spaces*, Acta Math. **130** (1973), 309–317.

[Eng89] Ryszard Engelking, *General topology*, second ed., Sigma Series in Pure Mathematics, vol. 6, Heldermann Verlag, Berlin, 1989, Translated from the Polish by the author. MR MR1039321 (91c:54001)

[ER52] Paul Erdős and Richard Rado, *Combinatorial theorems on classifications of subsets of a given set*, Proc. London Math. Soc. (3) **2** (1952), 417–439.

[ER56] P. Erdős and R. Rado, *A partition calculus in set theory*, Bull. Amer. Math. Soc. **62** (1956), no. 5, 427–489.

[ER63] P. Erdős and A. Rényi, *Asymmetric graphs*, Acta Math. Acad. Sci. Hungar. **14** (1963), 295–315.

[Erd47] Paul Erdős, *Some remarks on the theory of graphs*, Bull. Amer. Math. Soc. **53** (1947), 292–294.

[Erd63] ———, *On a combinatorial problem*, Nordisk Mat. Tidskr. **11** (1963), 5–10.

[ES73] Paul Erdős and John L. Selfridge, *On a combinatorial game*, J. Combin. Theory Ser. A **14** (1973), 298–301.

[Fag75] Ronald Fagin, *Monadic generalized spectra*, Z. Math. Logik Grundlagen Math. **21** (1975), 89–96.

[Fag76] R. Fagin, *Probabilities on finite models*, J. Symbolic Logic **41** (1976), 50–58.

[Far98] Ilijas Farah, *Semiselective coideals*, Mathematika **45** (1998), 79–103.

[Fef65] Solomon Feferman, *Some applications of the notions of forcing and generic sets*, Fund. Math. **56** (1965), 325–345.

[FLRV21] Ronald Fagin, Jonathan Lenchner, Kenneth W. Regan, and Nikhil Vyas, *Multi-structural games and number of quantifiers*, Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), IEEE, 2021, pp. 1–13.

[FM88] David H. Fremlin and Arnold W. Miller, *On some properties of Hurewicz, Menger, and Rothberger*, Fund. Math. **129** (1988), 17–33.

[FPU55] E. Fermi, J. Pasta, and S. Ulam, *Studies of nonlinear problems. I*, Tech. Report LA-1940, Los Alamos Scientific Laboratory, 1955.

[Fra54a] R. Fraïssé, *Sur l'extension aux relations de quelques propriétés des ordres*, Ann. Sci. École Norm. Sup. **71** (1954), 363–388.

[Fra54b] Roland Fraïssé, *Sur quelques classifications des systèmes de relations*, Publ. Sci. Univ. Alger. Sér. A **1** (1954), 35–182.

[Fri71] Harvey M. Friedman, *Higher set theory and mathematical practice*, Ann. Math. Logic **2** (1971), 325–357.

[Fro67] Zdeněk Frolík, *Sums of ultrafilters*, Bull. Amer. Math. Soc. **73** (1967), 87–91.

[FSV95] Ronald Fagin, Larry J. Stockmeyer, and Moshe Y. Vardi, *On monadic NP vs. monadic co-NP*, Inform. and Comput. **120** (1995), no. 1, 78–92.

[Gai82] Haim Gaifman, *On local and non-local properties*, Proceedings of the Herbrand Symposium (Logic Colloquium '81), Stud. Logic Found. Math., vol. 107, North-Holland, Amsterdam, 1982, pp. 105–135.

[Gal78] Fred Galvin, *Indeterminacy of point-open games*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **26** (1978), 445–449.

[Gal79] David Gale, *The game of Hex and the Brouwer fixed-point theorem*, Amer. Math. Monthly **86** (1979), 818–827.

[GKLT69] Y. V. Glebskiĭ, D. I. Kogan, M. I. Liogon'kiĭ and V. A. Talanov, *Range and degree of realizability of formulas in the restricted predicate calculus*, Kibernetika (Kiev) (1969), no. 2, 17–28.

[GM93] W. Timothy Gowers and Bernard Maurey, *The unconditional basic sequence problem*, J. Amer. Math. Soc. **6** (1993), no. 4, 851–874.

[Gow96] W. Timothy Gowers, *A new dichotomy for Banach spaces*, Geom. Funct. Anal. **6** (1996), no. 6, 1083–1093.

[Gow02] ———, *An infinite Ramsey theorem and some Banach-space dichotomies*, Ann. of Math. (2) **156** (2002), no. 3, 797–833. MR 1954235 (2005a:46032)

[GP73] F. Galvin and K. Prikry, *Borel sets and Ramsey's theorem*, J. Symbolic Logic **38** (1973), 193–198. MR 337630

[Gro17] Martin Grohe, *Descriptive complexity, canonisation, and definable graph structure theory*, Lecture Notes in Logic, vol. 47, Cambridge University Press, Cambridge, 2017.

[Gro21] ———, *The logic of graph neural networks*, Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), IEEE, 2021, pp. 1–17.

[GS53] David Gale and F. M. Stewart, *Infinite games with perfect information*, Contributions to the Theory of Games, Vol. II, Annals of Mathematics Studies, no. 28, Princeton University Press, 1953, pp. 245–266.

[Gur66] V. I. Gurariĭ, *Spaces of universal placement, isotropic spaces and a problem of Mazur on rotations of Banach spaces*, Sibirsk. Mat. Ž. **7** (1966), 1002–1013. MR 34 #585

[Hal59] P. Hall, *Some constructions for locally finite groups*, J. London Math. Soc. **34** (1959), 305–319.

[Han65] William Hanf, *Model-theoretic methods in the study of elementary logic*, The Theory of Models (J. W. Addison, L. Henkin, and A. Tarski, eds.), North-Holland, Amsterdam, 1965, pp. 132–145.

[Har23] Bradd Hart, *An introduction to continuous model theory*, Model Theory of Operator Algebras (Isaac Goldbring, ed.), De Gruyter, Berlin, 2023, pp. 83–132.

[Hau14] Felix Hausdorff, *Grundzüge der Mengenlehre*, Veit, Leipzig, 1914.

[Hel96] Lauri Hella, *Logical hierarchies in PTIME*, Inform. and Comput. **129** (1996), no. 1, 1–19.

[Hen71] C. W. Henson, *A family of countable homogeneous graphs*, Pacific J. Math. **38** (1971), 69–83.

[Hin53] K. Jaakko J. Hintikka, *Distributive normal forms in the calculus of predicates*, Acta Philosophica Fennica, vol. 6, Societas Philosophica, Helsinki, 1953.

[HJ63] A. W. Hales and R. I. Jewett, *Regularity and positional games*, Trans. Amer. Math. Soc. **106** (1963), 222–229.

[Hod93] Wilfrid Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, Cambridge, 1993. MR 94e:03002

[HP24] Åsa Hirvonen and Joni Puljujärvi, *Ehrenfeucht–Fraïssé games for continuous first-order logic*, 2024, arXiv:2402.16662.

[Hun04] Edward V. Huntington, *The continuum and other types of serial order, with an introduction to Cantor's transfinite numbers*, Harvard University Press, Cambridge, MA, 1904.

[Hur26] Witold Hurewicz, *Über eine Verallgemeinerung des Borelschen Theorems*, Math. Z. **24** (1926), 401–421.

[HV19] Wilfrid Hodges and Jouko Väänänen, *Logic and games*, The Stanford Encyclopedia of Philosophy, 2019.

[Imm81] Neil Immerman, *Number of quantifiers is better than number of tape cells*, J. Comput. System Sci. **22** (1981), no. 3, 384–406.

[Imm82] ———, *Upper and lower bounds for first order expressibility*, J. Comput. System Sci. **25** (1982), 76–98.

[Imm99] ———, *Descriptive complexity*, Graduate Texts in Computer Science, Springer, New York, 1999.

[JN76] S. A. Jalali-Naini, *The monotone subsets of Cantor space, filters and descriptive set theory*, Ph.D. thesis, University of Oxford, 1976.

[Jón56] B. Jónsson, *Universal relational systems*, Math. Scand. **4** (1956), 193–208.

[Jón60] ———, *Homogeneous universal relational systems*, Math. Scand. **8** (1960), 137–142.

[Kac85] Mark Kac, *Enigmas of chance: An autobiography*, Harper & Row, New York, 1985.

[Kał96] Roman Kałuża, *Through a reporter's eyes: The life of Stefan Banach*, Birkhäuser, Boston, 1996.

[Kar65] Carol R. Karp, *Finite-quantifier equivalence*, The Theory of Models (J. W. Addison, L. Henkin, and A. Tarski, eds.), North-Holland, Amsterdam, 1965, pp. 407–412.

[Kas83] Ilias G. Kastanas, *On the Ramsey property for sets of reals*, J. Symbolic Logic **48** (1983), 1035–1045.

[Kec95] Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995. MR 1321597

[Kei61] H. J. Keisler, *Ultraproducts and elementary classes*, Nederl. Akad. Wetensch. Proc. Ser. A 64 = Indag. Math. **23** (1961), 477–495. MR 140396

[Kei64] H. Jerome Keisler, *Good ideals in fields of sets*, Ann. of Math. (2) **79** (1964), no. 2, 338–359.

[Kei23] H. Jerome Keisler, *Model theory for real-valued structures*, Beyond First-Order Model Theory, Vol. 2 (José Iovino, ed.), CRC Press, 2023, arXiv:2005.11851, pp. 43–101.

[KKM01] Petar S. Kenderov, Ivaylo S. Kortezov, and Warren B. Moors, *Topological games and topological groups*, Topology Appl. **109** (2001), 157–165.

[KM96] D. Y. Kleinbock and G. A. Margulis, *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, Sinaĭ's Moscow Seminar on Dynamical Systems, Amer. Math. Soc. Transl. Ser. 2, vol. 171, Amer. Math. Soc., Providence, RI, 1996, pp. 141–172.

[KM98] ———, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. of Math. (2) **148** (1998), no. 1, 339–360.

[KM99] P. S. Kenderov and W. B. Moors, *Fragmentability and sigma-fragmentability of Banach spaces*, J. London Math. Soc. (2) **60** (1999), 203–223.

[Kőn27] D. Kőnig, *Über eine Schlußweise aus dem Endlichen ins Unendliche*, Acta Sci. Math. (Szeged) **3** (1927), 121–130.

[Kun72] Kenneth Kunen, *Ultrafilters and independent sets*, Trans. Amer. Math. Soc. **172** (1972), 299–306. MR 0314619 (47 #3170)

[Kun76] ———, *Some points in $\beta N$*, Math. Proc. Cambridge Philos. Soc. **80** (1976), 385–398.

[Kun80] ———, *Weak P-points in $N^*$*, Topology, Vol. II (Proc. Fourth Colloq., Budapest, 1978), Colloq. Math. Soc. János Bolyai, vol. 23, North-Holland, 1980, pp. 741–749.

[KV95] Phokion G. Kolaitis and Jouko Väänänen, *Generalized quantifiers and pebble games on finite structures*, Ann. Pure Appl. Logic **74** (1995), 23–75.

[KV00] Phokion G. Kolaitis and Moshe Y. Vardi, *A game-theoretic approach to constraint satisfaction*, Proceedings of the 17th National Conference on Artificial Intelligence (AAAI), 2000, pp. 175–181.

[KW13] Dmitry Kleinbock and Barak Weiss, *Modified Schmidt games and a conjecture of Margulis*, J. Mod. Dyn. **7** (2013), 429–460.

[Lav76] Richard Laver, *On the consistency of Borel's conjecture*, Acta Math. **137** (1976), 151–169.

[Lib04] Leonid Libkin, *Elements of finite model theory*, Texts in Theoretical Computer Science. An EATCS Series, Springer, Berlin, 2004.

[Lus76] W. Lusky, *The Gurarij spaces are unique*, Arch. Math. (Basel) **27** (1976), no. 6, 627–635. MR 55 #6177

[Lut90] Jack H. Lutz, *Category and measure in complexity classes*, SIAM J. Comput. **19** (1990), no. 6, 1100–1131.

[Lut92] ———, *Almost everywhere high nonuniform complexity*, J. Comput. System Sci. **44** (1992), 220–258.

[Mar70] Donald A. Martin, *Measurable cardinals and analytic games*, Fund. Math. **66** (1970), 287–291.

[Mar75] ———, *Borel determinacy*, Ann. of Math. (2) **102** (1975), 363–371.

[Mar91] G. A. Margulis, *Dynamical and ergodic properties of subgroup actions on homogeneous spaces with applications to number theory*, Proceedings of the International Congress of Mathematicians (Kyoto, 1990), Vol. I, Math. Soc. Japan, Tokyo, 1991, pp. 193–215.

[Mat77] A. R. D. Mathias, *Happy families*, Ann. Math. Logic **12** (1977), 59–111.

[Mau81] R. Daniel Mauldin (ed.), *The Scottish Book: Mathematics from the Scottish Café*, Birkhäuser, Boston, 1981.

[Mau15] R. Daniel Mauldin (ed.), *The Scottish Book: Mathematics from the Scottish Café, with selected problems from the New Scottish Book*, second ed., Birkhäuser/Springer, Cham, 2015.

[Meh73] Kurt Mehlhorn, *On the size of sets of computable functions*, 14th Annual Symposium on Switching and Automata Theory, IEEE, 1973, pp. 190–196.

[Men24] Karl Menger, *Einige überdeckungssätze der Punktmengenlehre*, Sitzungsberichte Wiener Akad. **133** (1924), 421–444.

[Mic63] E. Michael, *The product of a normal space and a metric space need not be normal*, Bull. Amer. Math. Soc. **69** (1963), no. 3, 375–376.

[MRF+19] Christopher Morris, Martin Ritzert, Matthias Fey, William L. Hamilton, Jan Eric Lenssen, Gaurav Rattan, and Martin Grohe, *Weisfeiler and Leman go neural: higher-order graph neural networks*, Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 4602–4609.

[MS62] Jan Mycielski and Hugo Steinhaus, *A mathematical axiom contradicting the axiom of choice*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **10** (1962), 1–3.

[MS89] Donald A. Martin and John R. Steel, *A proof of projective determinacy*, J. Amer. Math. Soc. **2** (1989), 71–125.

[MU49] Nicholas Metropolis and S. Ulam, *The Monte Carlo method*, J. Amer. Statist. Assoc. **44** (1949), 335–341.

[Nam74] Isaac Namioka, *Separate continuity and joint continuity*, Pacific J. Math. **51** (1974), 515–531.

[Neu60] B. H. Neumann, *Permutational products of groups*, J. Austral. Math. Soc. **1** (1960), 299–310.

[NP75] I. Namioka and R. R. Phelps, *Banach spaces which are Asplund spaces*, Duke Math. J. **42** (1975), 735–750.

[NW65] C. St. J. A. Nash-Williams, *On well-quasi-ordering transfinite sequences*, Proc. Cambridge Philos. Soc. **61** (1965), 33–39.

[OS94] Edward Odell and Thomas Schlumprecht, *The distortion problem*, Acta Math. **173** (1994), no. 2, 259–281.

[Oxt57] John C. Oxtoby, *The Banach–Mazur game and Banach category theorem*, Contributions to the Theory of Games, Vol. III, Annals of Mathematics Studies, vol. 39, Princeton University Press, 1957, pp. 159–163.

[Oxt80] J. C. Oxtoby, *Measure and category*, second ed., Graduate Texts in Mathematics, vol. 2, Springer-Verlag, New York, 1980.

[Paw94] Janusz Pawlikowski, *Undetermined sets of point-open games*, Fund. Math. **144** (1994), 279–285.

[Pez99] Elena Pezzoli, *Computational complexity of Ehrenfeucht–Fraïssé games on finite structures*, Computer Science Logic (CSL '98) (Berlin), Lecture Notes in Comput. Sci., vol. 1584, Springer, 1999, pp. 159–170.

[Poi82] Bruno Poizat, *Deux ou trois choses que je sais de $L_n$*, J. Symbolic Logic **47** (1982), no. 3, 641–658.

[Rad64] R. Rado, *Universal graphs and universal functions*, Acta Arith. **9** (1964), 331–340.

[Rag88] Prabhakar Raghavan, *Probabilistic construction of deterministic algorithms: approximating packing integer programs*, J. Comput. System Sci. **37** (1988), 130–143.

[Ram30] F. P. Ramsey, *On a Problem of Formal Logic*, Proc. London Math. Soc. (2) **30** (1930), no. 4, 264–286.

[Rec94] Ireneusz Recław, *Every Lusin set is undetermined in the point-open game*, Fund. Math. **144** (1994), 43–54.

[Rud56] Walter Rudin, *Homogeneity problems in the theory of čech compactifications*, Duke Math. J. **23** (1956), 409–419.

[Rud69] M. E. Rudin, *A new proof that metric spaces are paracompact*, Proc. Amer. Math. Soc. **20** (1969), 603.

[Rya22] Dmitry Ryabogin, *A negative answer to Ulam's Problem 19 from the Scottish Book*, Ann. of Math. (2) **195** (2022), no. 3, 1111–1150.

[Sca67] Herbert Scarf, *The approximation of fixed points of a continuous mapping*, SIAM J. Appl. Math. **15** (1967), 1328–1343.

[Sch66] Wolfgang M. Schmidt, *On badly approximable numbers and certain games*, Transactions of the American Mathematical Society **123** (1966), no. 1, 178–199.

[Sch69] ———, *Badly approximable systems of linear forms*, J. Number Theory **1** (1969), 139–154.

[Sch91] Thomas Schlumprecht, *An arbitrarily distortable Banach space*, Israel J. Math. **76** (1991), no. 1-2, 81–95.

[She71] S. Shelah, *Every two elementarily equivalent models have isomorphic ultrapowers*, Israel J. Math. **10** (1971), 224–233. MR 297554

[She84] Saharon Shelah, *Can you take Solovay's inaccessible away?*, Israel J. Math. **48** (1984), 1–47.

[Sie38] Wacław Sierpiński, *Fonctions additives non complètement additives et fonctions non mesurables*, Fund. Math. **30** (1938), 96–99.

[Sil70] J. Silver, *Every analytic set is Ramsey*, J. Symbolic Logic **35** (1970), 60–64. MR 332480

[Sil94] Charles L. Silver, *Who invented Cantor's back-and-forth argument?*, Modern Logic **4** (1994), no. 1, 74–78.

[Spe28] Emanuel Sperner, *Neuer Beweis für die Invarianz der Dimensionszahl und des Gebietes*, Abh. Math. Sem. Univ. Hamburg **6** (1928), 265–272.

[Spe87] Joel Spencer, *Ten lectures on the probabilistic method*, CBMS-NSF Regional Conference Series in Applied Mathematics, no. 52, SIAM, Philadelphia, 1987.

[SR83] Jean Saint-Raymond, *Jeux topologiques et espaces de Namioka*, Proc. Amer. Math. Soc. **87** (1983), 499–504.

[Ste10] E. Steinitz, *Algebraische Theorie der Körper*, J. Reine Angew. Math. **137** (1910), 167–309.

[Ste75] Charles Stegall, *The Radon–Nikodým property in conjugate Banach spaces*, Trans. Amer. Math. Soc. **206** (1975), 213–223.

[Sto36] M. H. Stone, *The theory of representations for Boolean algebras*, Trans. Amer. Math. Soc. **40** (1936), 37–111.

[Sto48] A. H. Stone, *Paracompactness and product spaces*, Bull. Amer. Math. Soc. **54** (1948), 977–982.

[Tal80] Michel Talagrand, *Compacts de fonctions mesurables et filtres non mesurables*, Studia Math. **67** (1980), 13–43.

[Tar30] Alfred Tarski, *Une contribution à la théorie de la mesure*, Fund. Math. **15** (1930), 42–50.

[Tel84] Rastislav Telgársky, *On games of Topsøe*, Math. Scand. **54** (1984), 170–176.

[Tel87] R. Telgársky, *Topological games: On the 50th anniversary of the Banach–Mazur game*, Rocky Mountain J. Math. **17** (1987), no. 2, 227–276.

[Tod10] Stevo Todorcevic, *Introduction to Ramsey spaces*, Annals of Mathematics Studies, vol. 174, Princeton University Press, Princeton, NJ, 2010. MR 2603812 (2011i:05263)

[TV57] Alfred Tarski and Robert L. Vaught, *Arithmetical extensions of relational systems*, Compositio Math. **13** (1957), 81–102.

[Ula30] Stanisław Ulam, *Zur Masstheorie in der allgemeinen Mengenlehre*, Fund. Math. **16** (1930), 140–150.

[Ula76] S. M. Ulam, *Adventures of a mathematician*, Charles Scribner's Sons, New York, 1976.

[Ury27] P. Urysohn, *Sur un espace métrique universel*, Bull. Sci. Math. **51** (1927), 43–64, 74–90.

[Vää11] Jouko Väänänen, *Models and games*, Cambridge Studies in Advanced Mathematics, vol. 132, Cambridge University Press, Cambridge, 2011.

[vB76] Johan van Benthem, *Modal correspondence theory*, Ph.D. thesis, University of Amsterdam, 1976.

[Wad83] William W. Wadge, *Reducibility and determinateness on the Baire space*, Ph.D. thesis, University of California, Berkeley, 1983.

[Wim82] Edward L. Wimmers, *The Shelah P-point independence theorem*, Israel J. Math. **43** (1982), 28–48.

[Wol55] Philip Wolfe, *The strict determinateness of certain infinite games*, Pacific J. Math. **5** (1955), 841–847.

[XHLJ19] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka, *How powerful are graph neural networks?*, International Conference on Learning Representations (ICLR), 2019.